

Horizontal transfer of large plasmid with type IV secretion system and mosquitocidal genomic island with excision and integration capabilities in *Lysinibacillus sphaericus*

Peiling Geng,¹ Jiao Cheng,² Zhiming Yuan,²
Hairong Xiong,¹ Haiying Wang¹ and Xiaomin Hu^{1*}

¹College of Life Science, South-Central University for Nationalities, Wuhan, 430074, China.

²Wuhan Institute of Virology, Chinese Academy of Sciences, Wuhan, 430071, China.

Summary

We identified a ~30-kb genomic island (named GI8) carrying the binary toxin gene operon *binA/binB* on both the chromosome and large pBsph plasmid in the mosquitocidal *Lysinibacillus sphaericus* C3-41 strain. We found that GI8 is related to the occurrence of *binA/binB* within *L. sphaericus* and displays excision and integration capability by recognizing the *attB* region, which consists of a 2-nt target site (AT) flanked by an 11-nt imperfect inverted repeat. pBsph and two pBsph-like plasmids (p2362 and p1593) were found to carry a type IV secretion system (T4SS) and displayed transmissibility within a narrow host range specific to *L. sphaericus*. GI8 can be co-transferred with pBsph as a composite element by integration into its *attB* site, then excised from pBsph and re-integrated into the chromosomal *attB* site in the new host. The potential hosts of GI8, regardless of whether they are toxic or non-toxic to mosquito larvae, share good collinearity at the chromosomal level. Data indicated that the appearance of the mosquitocidal *L. sphaericus* lineage was driven by horizontal transfer of the T4SS-type conjugative plasmid and GI8 with excision and specific integration capability.

Introduction

Conjugative transfer is a common horizontal genetic transfer (HGT) strategy for large DNA-fragment translocation across cell envelopes. Two types of mobile genetic elements (MGEs), i.e., conjugative plasmids and integrative conjugative elements (ICEs) (including conjugative transposons), exhibit self-transmission capability by conjugation (Norman *et al.*, 2009; Johnson and Grossman, 2015). The latter are large DNA segments typically containing recombination and conjugation modules. Integrative mobilizable elements (IMEs) are another recently identified MGE, which differ from ICEs in that they cannot self-transfer but can be mobilized by ‘co-resident’ ICEs or conjugative plasmids (Guédon *et al.*, 2017; Parmeciano *et al.*, 2019). Both ICEs and IMEs carry genes that can help hosts adapt to specific ecological niches (e.g., antibiotic resistance, pathogenesis, symbiosis, and/or metabolic pathway) (Kohler *et al.*, 2019; Durrant *et al.*, 2020; Botelho and Schulenburg, 2021). The genetic basis of conjugational transfer in Gram-negative (G^-) bacteria has been well explored and is reported to be based on the type IV secretion system (T4SS). This system typically includes a relaxase (Mob) nicking *oriT* along with other accessory proteins, as well as a coupling protein (T4CP) acting as a substrate receptor and mating-pair formation (MPF) proteins (VirBx or homologues) constituting the mating channel (Alvarez-Martinez and Christie, 2009; de la Cruz *et al.*, 2010; Waksman and Fronzes, 2010). The single-strand conjugative transfer system (T4SS) in Gram-positive (G^+) bacteria is suggested to encode homologous components to T4SS in G^- bacteria. Nevertheless, unlike the latter, which utilizes an extracellular pilus, G^+ conjugation relies on surface adhesins to build contacts between donor and recipient cells (Grohmann *et al.*, 2018). Some conjugative plasmids and ICEs in G^+ bacteria also use a double-stranded transfer system, which involves a replication initiator protein (Rep) and a FtsK/SpoIIIE domain-containing protein (Tra) for DNA translocation (Grohmann *et al.*, 2018).

Received 17 November, 2020; revised 5 March, 2021; accepted 14 March, 2021. *For correspondence. E-mail xiaomin.hu@mail.scuec.edu.cn; Tel. (+86)13986247699; Fax (+86)2767842689.

Several *Lysinibacillus sphaericus* isolates, which are ubiquitous in natural environments, display larvicidal features against mosquitoes (Ge *et al.*, 2011). The crystals consisting of BinA/BinB produced at the sporulation stage are the main contributors to mosquitocidal activity (Berry *et al.*, 2012). The commercial formulations of mosquitocidal *L. sphaericus*, such as 2362 (VectoLex® and Spherimos®), C3-41 (Jianbao®), 1593 (Spicbiomoss®), and 2297 (VectoLex G®), have greatly contributed to the integrative control of mosquitoes and mosquito-borne diseases (e.g., malaria, dengue fever, filariasis, and yellow fever) (Xiong *et al.*, 2010). However, their long-term use can induce mosquito resistance, which is likely due to the single effector binary toxin (Su *et al.*, 2018). Therefore, it is important to explore the diversity and evolution of mosquitocidal *L. sphaericus*.

The genetic structure and sequence of chromosomes of mosquitocidal isolates are highly conserved, in contrast to non-toxic isolates, which display large heterogeneity (Ge *et al.*, 2011). The binary operon *binA/binB* is located within a 30-kb genomic island (GI8) in *L. sphaericus* and shows high mosquitocidal activity (Xu *et al.*, 2015). In *L. sphaericus* C3-41, a copy of GI8 is found in the chromosome and in the large pBsph plasmid (178 kb) (Fu *et al.*, 2017a). GI8 is predicted to encode integrase and transposase (Xu *et al.*, 2015), indicating it is probably an ICE or IME. Two plasmids displaying almost identical backbones to pBsph were recently identified by genome sequencing in two *L. sphaericus* isolates (OT4b.25 and III(3)7) and were predicted to encode the VirB4, VirB6, and VirD4 homologues of the conjugative T4SS complex (Xu *et al.*, 2015; Rey *et al.*, 2016a, b) related to conjugative transfer characteristics as well as pBsph. Therefore, it is hypothesized that the pBsph-like plasmids and GI8 are potential MGEs, and their HGT has contributed to the transmission of the binary operon *binA/binB* and the emergence of the toxic lineage within the *L. sphaericus* population. Therefore, in the current study, we explored the occurrence and genetic background of the large plasmid and 30-kb IME-type mosquitocidal island within the *L. sphaericus* population and further analysed the HGT capabilities of the pBsph-like plasmids and GI8.

Results

Revision of genetic organization of pBsph

Based on predicted functions, we identified four genetic regions related to conjugation in the pBsph plasmid (Table 1, Fig. 1). (i) The pBsph_001-017 region was predicted to carry two essential components of the ssDNA conjugation machinery, i.e., the type IV coupling protein gene *virD4* (pBsph_002) and relaxosome-encoding gene

complex consisting of relaxase (MOB) (probably encoded by pBsph_001 and pBsph_005 as both carry MobC domain I) and auxiliary proteins. This region likely plays a role in the nicking of dsDNA and binding of ssDNA at the origin of transfer (Garcillan-Barcia *et al.*, 2009). (ii) The pBsph_039-058 region was predicted to encode the core MPF region, containing the T4SS marker *virB4* (pBsph_052) and several other *vir* homologues (e.g., pBsph_053 and pBsph_054 were predicted to carry six and nine transmembrane helices, respectively, and both displayed homology to *virB6*). (iii) The pBsph_063-138 region was predicted to encode the regulative region (e.g., pBsph_063, 098, 101, 113, and 138 were predicted as helix-turn-helix or DNA-binding domains containing transcriptional regulators). (iv) The pBsph_173-186 region was predicted to encode proteins with putative surface-associated and pilus assembly functions, which are the genetic basis for G+ conjugation to establish and maintain mating pairs (Grohmann *et al.*, 2018; Kohler *et al.*, 2019). Thus, pBsph was thought to carry a T4SS-like system.

GI8, which carries mosquitocidal binary toxin genes, as identified in previous research (Xu *et al.*, 2015), was re-analysed in this study and found to contain 23 open reading frames (ORFs), corresponding to pBsph_144–166 in pBsph (and Bsph_3182 and Bsph_3204 in the C3-41 strain chromosome). Both pBsph_144 and pBsph_145 were predicted to encode the XerC and XerD homologues, respectively, which belong to the tyrosine recombinase family. pBsph_147 to pBsph_149 were found to be homologous to the *gerX* operon located in the 44.8-kb class II transposon TnXO1 recognized in pXO1 (Van der Auwera and Mahillon, 2005). pBsph_158 was the largest ORF (7170 bp) with a predicted non-ribosomal peptide synthase (gramicidin S synthetase 2). Moreover, GI8 was found inserted in an *umuDC*-like operon encoding DNA polymerase V. This operon included predicted interrupted *yolD* genes (pBsph_143 and pBsph_167) and UV-damage repair protein gene *uvrX* (Table 2, Fig. 1). Previous studies have suggested that the *umuDC* operon belongs to the SOS-regulon and is often associated with MGEs (Murli *et al.*, 2000; Permina *et al.*, 2002). These results suggest that GI8 is likely to be an ICE or IME.

Occurrence of pBsph-like plasmid and GI8 in *L. sphaericus*

Five of the 22 tested *L. sphaericus* isolates (Table 3) (C3-41, 2362, 1593, 2317-2, and SSII-1) were found to carry the plasmid replicon genes pBsph-*tubRZ* by polymerase chain reaction (PCR) (data not shown). Plasmids from these isolates were visualized on agarose gel (Fig. 2). Results demonstrated that smaller-sized pBsph-like plasmids were distributed in 2362, 1593, 2317-2, and

Table 1. Conjugation related genes and predicted function in pBsph.

CDS	Strand	Start	End	Gene size (bp)	Putative protein size (aa)	Annotation	Additional interesting homologues
pBsph_001	+	439	1674	1236	411	Hypothetical protein	MobC domain 1 by CONJscan (e-value: 5.6e-06)
pBsph_002	+	1818	4811	2994	997	VirD4	DNA-binding domain, 5 TMHs
pBsph_004	+	5739	6875	1137	378	Hypothetical protein	S1 RNA-binding domain
pBsph_005	+	6879	8015	1137	378	Hypothetical protein	MobC domain 1 by CONJscan (e-value: 2.4e-05)
pBsph_007	+	8291	8839	549	182	ATPase/ATP-binding protein	
pBsph_008	+	8879	9367	489	162	Hypothetical protein	27% identity to efflux RND transporter permease subunit (Pseudoalteromonas sp. H105)
pBsph_015	+	15 146	16 525	1380	459	DNA primase	DUF3854 domain
pBsph_016	+	16 537	16 827	291	96	Single-stranded DNA-binding protein	
pBsph_017	+	16 856	17 455	600	199	Single-strand binding protein	
pBsph_039	+	34 878	35 363	486	161	Outer membrane beta-barrel protein	
pBsph_041	+	36 856	41 313	4458	1485	S-layer homology domain-containing protein	Low similarity to VirB6 by SecreT4 (e-value: 0.006)
pBsph_043	+	42 730	43 629	900	299	Dinucleotide-utilizing protein	
pBsph_044	+	43 598	45 508	1911	636	DNA polymerase III subunit gamma	MobF domain 1 by CONJscan (e-value: 9.4e-05)
pBsph_045	+	45 539	46 411	873	290	DNA polymerase III delta subunit domain protein	
pBsph_046	+	46 579	47 760	1182	393	DNA polymerase III subunit beta	
pBsph_049	+	49 359	49 724	366	121	Hypothetical protein	25%–35% identity to TrbC/VirB2 family protein and DNA translocase FtsK, 2 TMHs
pBsph_050	+	49 784	50 095	312	103	Hypothetical protein	25%–35% identity to PrgI family protein, 2 TMHs
pBsph_052	+	51 252	53 147	1896	631	VirB4	
pBsph_053	+	53 170	55 998	2829	942	VirB6	6 TMHs
pBsph_054	+	55,961	57,901	1941	646	VirB6	9 TMHs
pBsph_055	+	57 914	60 097	2184	727	Peptidoglycan DD-metalloendopeptidase family protein	1 TMH
pBsph_056	+	60 110	60 748	639	212	Etratricopeptide repeat-containing protein	1 TMH
pBsph_058	+	61 416	61 700	285	94	Amino acid ABC transporter substrate-binding protein	
pBsph_063	+	64 591	64 959	369	122	XRE family transcriptional regulator	Helix-turn-helix domain
pBsph_064	+	64 962	65 513	552	183	Rad52/22 double-strand break repair protein	
pBsph_071	+	70 137	73 133	2997	998	Helicase/N-6 DNA methylase	
pBsph_076	+	75 466	75 768	303	100	Hypothetical protein	Helicase conserved C-terminal domain
pBsph_087	+	81 095	82 201	1107	368	Tyrosine recombinase xerS	
pBsph_088	+	83 113	83 700	588	195	ABC transporter substrate-binding protein	
pBsph_090	–	84 175	84 600	426	141	Hypothetical protein	4 TMHs
pBsph_094	–	87 836	88 447	612	203	DNA-binding protein/thermonuclease	
pBsph_098	–	90 419	90 646	228	75	DNA-binding XRE family transcriptional regulator	Helix-turn-helix domain, low similarity to conjugative transfer signal peptidase TraF (Nitrosomonas sp. AL212) e-value: 0.097
pBsph_101	–	92 148	93 134	987	328	Hypothetical protein	Helix-turn-helix domain, excisionase family DNA binding domain
pBsph_113	–	100 641	100 967	327	108	Hypothetical protein	AbrB/MazE/SpoVT family DNA-binding domain

(Continues)

Table 1. Continued

CDS	Strand	Start	End	Gene size (bp)	Putative protein size (aa)	Annotation	Additional interesting homologues
pBsph_114	-	101 068	101 361	294	97	Hypothetical protein	ATP-grasp domain
pBsph_123	-	110 703	111 593	891	296	DNA modification methylase	
pBsph_124	-	111 776	112 036	261	86	Hypothetical protein	Low identity and coverage to ATP-dependent helicase
pBsph_126	-	112 513	113 445	933	310	DNA polymerase III subunit delta	
pBsph_138	-	120 983	121 666	684	227	Transcriptional regulator	Helix-turn-helix domain
pBsph_173	-	157 704	160 817	3114	1037	Bacterial surface protein	1 TMH
pBsph_174	-	160 833	162 041	1209	402	Hypothetical protein	S-layer homology domain
pBsph_177	-	164 081	164 611	531	176	Hypothetical protein	Low identity (evalue: 0.094) to virB9-2 gene product (Rickettsia bellii RML369-C)
pBsph_179	-	165 051	165 518	468	155	ABC transporter permease	1 TMH
pBsph_180	-	165 617	167 656	2040	679	Hypothetical protein	8 TMHs
pBsph_181	-	167 650	169 509	1860	619	ATPase	Homologous to TrbB by SecReT4 software (7e-08) and VirB11 (Helicobacter pylori P12) (1.3e-5), 1 TMH
pBsph_182	-	169 512	170 291	780	259	Hypothetical protein	26% identity to ParA family protein
pBsph_184	-	171 039	171 989	951	316	Flagellar basal body P-ring biosynthesis protein	1 TMH
pBsph_185	-	172 006	173 508	1503	500	Serine/threonine protein kinase/Flp pilus assembly protein	1 TMH
pBsph_186	-	173 585	174 520	936	311	Hypothetical protein	34% identity to metallophosphoesterase

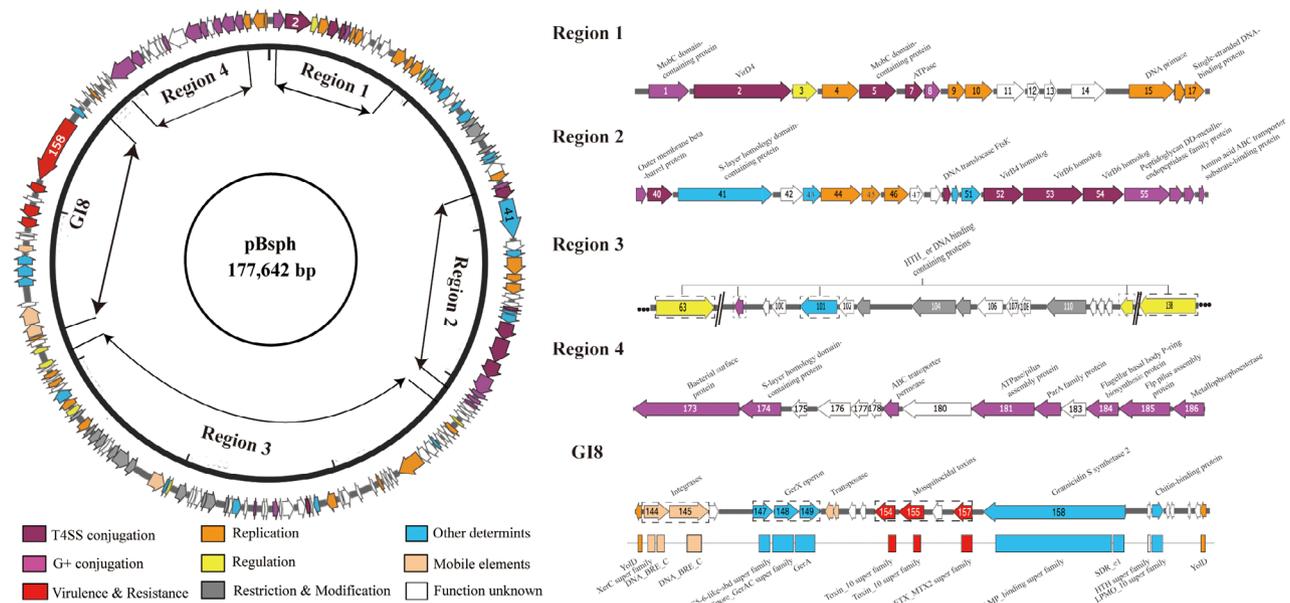


Fig 1. Genetic structure of pBsph. Left: circular genetic map displaying predicted CDSs and functional region assignment of pBsph. Right: four genetic regions carrying main putative CDSs related to conjugation in pBsph, and homologous analysis of main putative CDSs in GI8. [Color figure can be viewed at wileyonlinelibrary.com]

Table 2. Genetic structure of GI8.

ORFs		Gene size (bp)	Putative protein size (aa)	Product annotation	Conserved domain
In chromosome	In pBsph				
Bsph_3182	pBsph_144	1365	454	Phage integrase family protein	XerC super family; DNA_BRE_C
Bsph_3183	pBsph_145	2148	715	Phage integrase family protein	DNA_BRE_C
Bsph_3184	pBsph_146	501	166	Hypothetical protein	
Bsph_3185	pBsph_147	1077	358	Spore germination protein	SLC5-6-like_sbd super family
Bsph_3186	pBsph_148	1401	466	Spore germination protein	GerA
Bsph_3187	pBsph_149	1122	373	Spore germination protein XC precursor	Spore_GerAC Super family
Bsph_3188	pBsph_150	834	277	IS3 family transposase with a point mutation	
Bsph_3190	pBsph_152	549	182	Hypothetical protein	
Bsph_3191	pBsph_153	318	105	Hypothetical protein	
Bsph_3192	pBsph_154	1113	370	41.9 kDa insecticidal toxin	Toxin_10 super family
Bsph_3193	pBsph_155	1347	448	Larvicidal toxin 51 kDa protein	Toxin_10 super family
Bsph_3194	pBsph_156	531	176	Hypothetical protein	
Bsph_3195	pBsph_157	1014	337	Hypothetical Mtx2/3 toxin-like protein	ETX_MTX2 Super family
Bsph_3196	pBsph_158	7572	2523	Amino acid adenylation	AMP-binding super family; SDR_e1
Bsph_3197	pBsph_159	123	40	Hypothetical protein	
Bsph_3198	pBsph_160	240	79	Hypothetical protein	HTH Super family
Bsph_3199	pBsph_161	663	220	Chitin-binding protein	LPMO_10 Super family
Bsph_3200	pBsph_162	216	71	Hypothetical protein	
Bsph_3201	pBsph_163	249	82	Hypothetical protein	
Bsph_3202	pBsph_164	138	45	Hypothetical protein	
Bsph_3203	pBsph_165	186	61	Hypothetical protein	YolD
Bsph_3204	pBsph_166	342	113	Hypothetical protein	YolD

SSII-1, named p2362, p1593, p2317-2, and pSSII-1, respectively.

As shown in Fig. 3A, the occurrence of GI8 was established by PCR, with the predicted *attL* and *attR* sequences detected by *chro-a/b* and *chro-c/d*, respectively, and *xerC* and *xerD* sequences detected by *xerC-F/R* and *xerD-F/R*, respectively, which are the characteristic elements of GI8. In total, 12 isolates, including C3-41, G725 (pBsph-cured derivative of C3-41), IAB881, 2362, IAB872, LP1-G, 47-6b, 1593, 2317-2, IAB59, 2297, and Bs-197 were positive for all four primer pairs (Fig. 3Ba-d), suggesting the presence of GI8. The occurrence of binary toxin genes appeared to be associated with GI8 as they displayed a co-existing relationship (Table 3). Six isolates (i.e., 2173, Cok31, SSII-1, KellenQ, 2314-2, and Dak614) were *chro-a/d*-positive for the predicted *attB* target (Fig. 3Be), and the resulting product sequences were matched to the joined flanking sequence of GI8 (Fig. 3C), indicating the absence of GI8 but the presence of *attB* as a potential integrative host.

Identification of integrative sites of GI8 in chromosomes and plasmids of *L. sphaericus*

The chromosome sequences of 1593 (CP064068), SSII-1 (CP064070), and KellenQ (CP064067) obtained in this study were compared with all other *L. sphaericus* genomes published (up to December 2020), including

C3-41 (NC_010382.1), 2362 (CP015224.1), OT4b.25 (CP014643.1), III(3)7 (CP014856.1), DH-B01 (CP045583), and DSM 28 (NZ_CP019980) (Hu *et al.*, 2008; Hernandez-Santana *et al.*, 2016; Rey *et al.*, 2016a, b; Wang *et al.*, 2020). Those isolates with high larvicidal activity (e.g., C3-41, 2362, OT4b.25, III(3)7, and 1593) were conserved, whereas avirulent isolates (e.g., DH-B01 and DSM 28) were diverse. The mobile elements (genomic islands and prophages) were mainly located on hypervariable regions, typically carrying tandem repeat sequences with abnormal GC content compared to the average. Furthermore, the toxin-encoding genes were highly correlated with the genomic island (Fig. 4A). Dot-plot analysis showed good collinearity among the isolates carrying GI8 (C3-41, 2362, OT4B.25, III(3)7, and 1593). Interestingly, the two isolates lacking GI8 but carrying *attB* (i.e., SSII-1 and KellenQ) exhibited good collinearity with the five toxic isolates (Fig. 4B), although the former only showed low toxicity (due to the mosquitocidal toxins produced during the vegetative growth phase, which are unstable and easily degraded by host protease; Yang *et al.*, 2007) and the latter was avirulent.

The complete sequences of p1593 (CP064069, 157 721 bp) and pSSII-1 (CP064071, 153 702 bp) obtained in this study, together with the published sequences of pOT4b.25 (NZ_CP014644, 176 083 bp) and pIII(3)7 (NZ_CP014857, 173 793 bp), were compared with pBsph (CP000818, 177 642 bp) (Fig. 5). Both

Table 3. Characterization of the tested *L. sphaericus* isolates.

Strain	Resource	Serotype ^a	Toxicity ^b	G18 ^c	attB ^d	pBsph(-like) plasmid	Frequency of acceptance of pBsph(-like) plasmid ^e		
							pBsphΔBin	p2362Δ9-10	P1593Δ9-10
C3-41	China	H ₅	H	+	-	pBsph	ND	ND	ND
G725Δ0498 ^f	China	H ₅	H	+	-	-	$3.9 \times 10^{-9} \pm 1.1 \times 10^{-3}$	$4.6 \times 10^{-3} \pm 3.7 \times 10^{-3}$	$3.2 \times 10^{-3} \pm 3.2 \times 10^{-3}$
2362	USA	H ₅	H	+	-	p2362	$<10^{-9}$	ND	ND
1593	Indonesia	H ₅	H	+	-	p1593	$<10^{-9}$	ND	ND
2317-2	USA	H ₅	H	+	-	p2317-2	$<10^{-9}$	ND	ND
IAB881	Ghana	H ₃	H	+	-	-	$3.5 \times 10^{-9} \pm 3.1 \times 10^{-9}$	$<10^{-9}$	$<10^{-9}$
LP1-G	Singapore	H ₃	H	+	-	-	$3.8 \times 10^{-8} \pm 6.2 \times 10^{-8}$	$<10^{-9}$	$<10^{-9}$
47-6b	China	H ₆	H	+	-	-	ND	ND	ND
IAB59	Ghana	H ₆	H	+	-	-	$<10^{-9}$	ND	ND
IAB872	Ghana	H ₄₈	H	+	-	-	$3.5 \times 10^{-7} \pm 2.2 \times 10^{-7}$	$<10^{-9}$	$<10^{-9}$
2297	USA	H ₂₅	H	+	-	-	$1.5 \times 10^{-7} \pm 1.5 \times 10^{-7}$	$<10^{-9}$	$<10^{-9}$
Bs-197	India	H ₁	M	+	-	-	ND	ND	ND
2173	USA	H ₂₆	M	-	+	-	ND	ND	ND
Kellen Q	USA	H ₁	L	-	+	-	$2.3 \times 10^{-6} \pm 7.8 \times 10^{-7}$	$5.6 \times 10^{-6} \pm 4.6 \times 10^{-6}$	$4.1 \times 10^{-6} \pm 2.9 \times 10^{-6}$
SSII-1	India	H ₂	L	-	+	pSSII-1	$<10^{-9}$	ND	ND
Cok31	Turkey	H ₉	L	-	+	-	$<10^{-9}$	$<10^{-9}$	$<10^{-9}$
2314-2	Thailand	H ₉	N	-	+	-	ND	ND	ND
Dak614	France	H ₄	N	-	+	-	ND	ND	ND
NRS1693	USA	H ₂	N	-	+	-	$<10^{-9}$	ND	ND
Bs208-6	China	ND	ND	-	-	-	ND	ND	ND
Bs227-3	China	ND	ND	-	-	-	ND	ND	ND
SO7001	France	ND	ND	-	-	-	ND	ND	ND

^aThe serotype results were referred from the study by de Barijac et al. (1985).

^bThe activity against *Culex quinquefasciatus* larvae which was determined by Ge et al. (2011). H, high toxicity; M, moderate toxicity; N, nontoxic; L, low toxicity.

^cThe isolates with PCR positive to chro-a/b, chro-c/d, xerC-F/R and xerD-F/R and carrying the binary operon were deemed as carrying G18.

^dThe isolates with PCR positive to chro-a/d and the sequences of the products matching to the joined flanking sequences of G18 were deemed as carrying attB.

^eThe strains used as recipients in mating experiments were resistant derivatives as described in material and methods. Three independent mating experiments were performed. The frequency was determined as transconjugants/donor \pm SE. ND: not determined. $<10^{-9}$. No transconjugants were detected.

^fpBsph-cured mutant of C3-41 in which a spectinomycin gene inserted into Bsph_0498 which encodes a restriction-modification enzyme in the chromosome of G725.

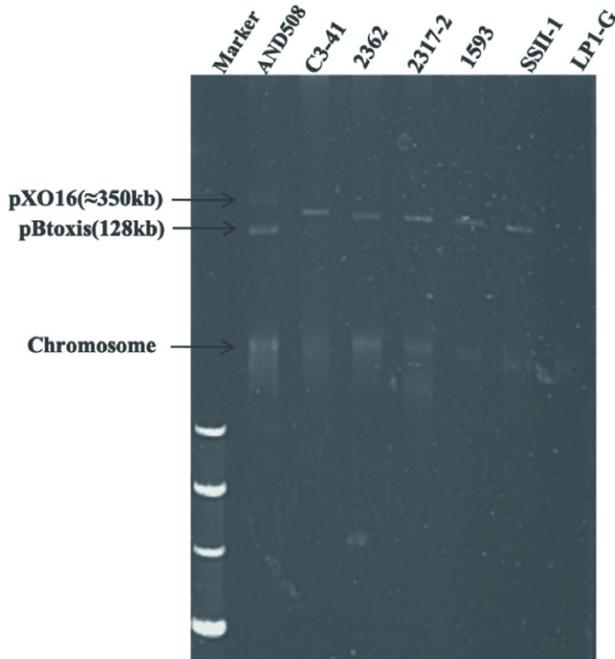


Fig 2. Large plasmid profile of *L. sphaericus* isolates containing *Bs-tubRZ*. Here, AND508 was used as a reference, which is a wild *B. thuringiensis israelensis* isolate containing four large plasmids, with 350-kb pXO16 and 128-kb pBtoxis indicated by arrows as DNA markers. *L. sphaericus* LP1-G negative to *Bs-tubR* was used as a negative control. Marker: Trans15k DNA Marker. [Color figure can be viewed at wileyonlinelibrary.com]

pOT4b.25 and pIII(3)7 were found to carry large fragment duplications (18 kb and 15 kb, respectively) matching pBsph_076-pBsph_102 and pBsph_001-pBsph_017 respectively. These duplications are likely due to the translocation of MGEs, which often results in changes in neighbouring DNA fragments by excision, insertion, or duplication, or raw data processing errors during sequencing and assembly. The main difference between pBsph and the four pBsph-like plasmids (i.e., p1593, pOT4b.25, pIII(3)7, and pSSII-1) was that GI8 was present in the former but absent in the latter. The flanking *umuDC*-like operon of GI8 in pBsph was also carried by the four pBsph-like plasmids, which lacked the joint target site 'AT'. In addition, p1593, pOT4b.25, and pIII(3)7 carried a ~13-kb fragment of unknown function flanked by a duplicated intron-encoded protein (IEP) gene (named Bsph.I1), which has homologues in many plasmids originating from the *B. cereus* group (Van der Auwera *et al.*, 2008; Hu *et al.*, 2009). Only one Bsph.I1 was predicted in pBsph and none in pSSII-1. Interestingly, a deficient type I-C CRISPR-Cas system, which contained three CRISPR spacers and *cas5*, *cas8*, and *cas7* but lacked *cas1*, *cas2*, *cas3*, and *cas4*, was found neighbouring the 'intron flanking region' in all five plasmids. In addition, fragment indels and rearrangements between

pSSII-1 and the other four plasmids were observed downstream of the predicted CRISPR-Cas system within the 'intron flanking region'.

A comparison of the *attB* flanking sequences with or without GI8 among the chromosomes and plasmids of the sequenced *L. sphaericus* was conducted. Results showed that GI8 was specifically inserted into a 2-nt (AT) core site within *umuDC*-like operons on the chromosomes of C3-41, 2362, OT4b.25, III(3)7, and 1593, and the pBsph plasmid of C3-41, which disrupted intact *attB* and resulted in an additional A and T or C compared to the chromosomes of DH-B01, SSII-1, KellenQ, and DSM 28 (Fig. 6A and B).

Intraspecies HGT capability of pBsph and pBsph-like plasmids

The plasmid pBsph Δ Bin (which contains the kanamycin resistance marker inserted into the binary operon of pBsph) was transferred into G725 Δ 0498 (a plasmid-cured derivative of C3-41 with a spectinomycin resistance marker inserted in the restriction-modification REase in the chromosome) at a frequency of $3.9 \times 10^{-3} \pm 1.1 \times 10^{-3}$ transconjugants/donor. It was also transferred into the *L. sphaericus* isolates IAB881, LP1-G, IAB872, 2297, and KellenQ, but at much lower frequencies (10^{-9} to 10^{-6} transconjugants/donor) (Table 3). PCR verified that the resulting transconjugants had a recipient background but carried pBsph Δ Bin (data not shown). The obtained transconjugants were used as donors in a new round of mating experiments, as the pBsph Δ Bin plasmids can transfer (defined as retroconjugational capability) at similar frequencies (Table 4). The transconjugants were passage cultured 40–50 times without antibiotics, and 99%–100% of colonies retained resistance to kanamycin and were PCR-positive to the plasmid replicon gene *tubR*, indicating that pBsph Δ Bin in the transconjugants was very stable. No mobilization of pBC16 by pBsph Δ Bin was observed. Other *Bacillus* spp. isolates, e.g., plasmid-cured derivatives of *B. thuringiensis israelensis* GBJ002, *B. thuringiensis kurstaki* AW05-R⁺, *B. cereus* Bt661-R⁺, AND1309-R⁺, AND1315-R⁺, and AND2022-R⁺ (Hu *et al.*, 2004), were also used as recipients in the mating experiments but did not show acceptance of pBsph Δ Bin. These data suggest that pBsph has a very narrow host range, possibly specific to *L. sphaericus*, and transmissibility between homologous hosts sharing identical chromosomal backgrounds was much higher than that between heterologous hosts.

In addition, the pBsph-like plasmids p2362 (p2362 Δ 9–10) and p1593 (p1593 Δ 9–10) displayed similar HGT capability and host range as pBsph (Table 3). However, the transmissibility of pSSII-1 could not be determined as foreign DNA could not be transformed into the SSII-1

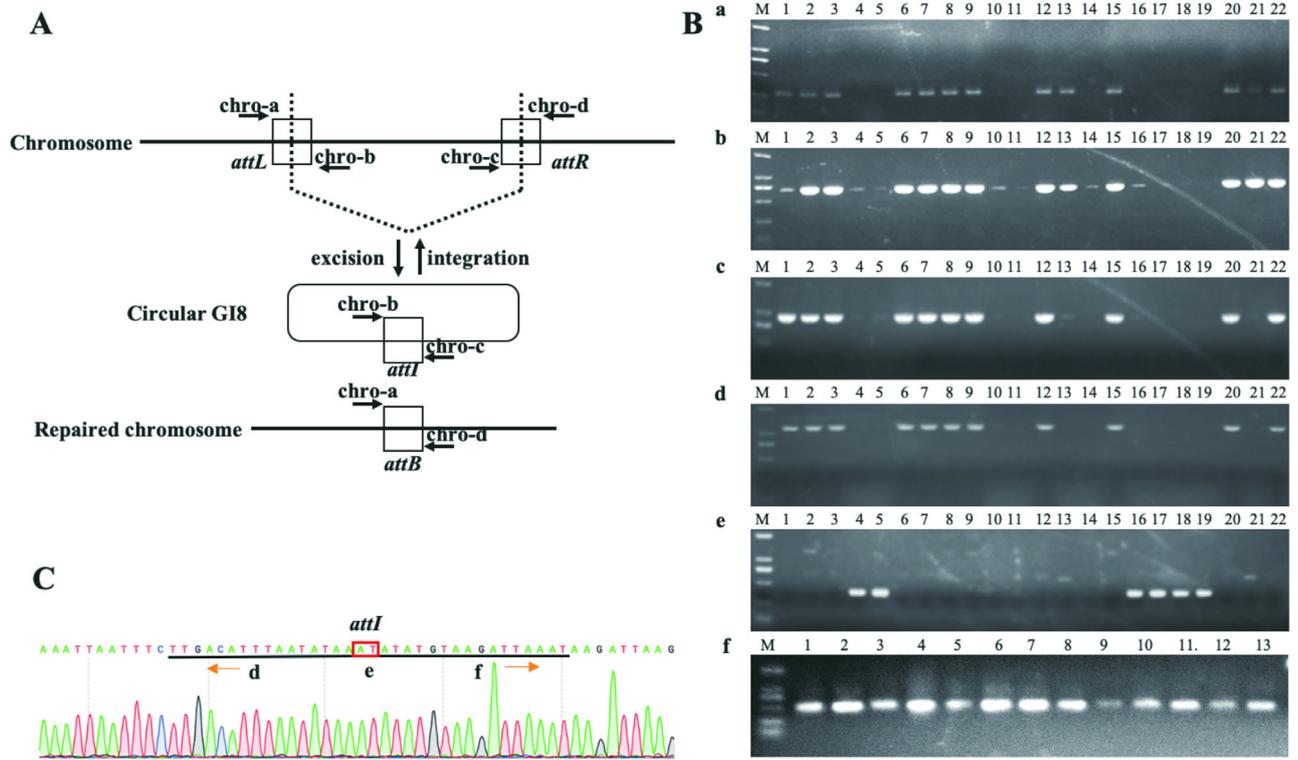


Fig 3. PCR detection of GI8-related genes in *L. sphaericus* isolates.

A. Schematic for detecting excision/integration of GI8 in *L. sphaericus*. Primer pairs chro-a/d and chro-b/c were used for detecting repaired chromosome (*attB*) and circular GI8 (*attI*), while chro-a/b and chro-c/d were used for left and right flanking sequences of GI8 (*attL* and *attR*, respectively).

B. PCR results using primer pairs chro-a/b, chro-c/d, xerC-F/R, xerD-F/R, and chro-a/d are shown as (a), (b), (c), (d), and (e), respectively. M: Trans2K Plus II. Lane 1–22: IAB881, 2362, Bs-197, 2173, 2314-2, IAB872, LP1-G, 47-6b, 1593, NRS1693, Bs208-6, 2317-2, Bs227-3, SO7001, IAB59, Cok31, SSII-1, KellenQ, Dak614, 2297, C3-41 and G725. (f) PCR results amplified by chro-b/c. M: Trans2K. Lane 1–13: IAB881, 2362, Bs-197, IAB872, LP1-G, 47-6b, 1593, 2317-2, IAB59, 2297, C3-41, G725, and KellenQ (pBC16, pBsphΔBin).

C. Sequencing of *attI* produced by G725. Regions d, f, and red-boxed region e indicate 11-nt incomplete reverse repeat on GI8 left and right flanking sequences and 2-nt forward repeat core recognition site. Directions are indicated by arrows. [Color figure can be viewed at wileyonlinelibrary.com]

host, which made it difficult to genetically tag selective markers in pSSII-1 for screening in the mating experiments.

Excision/cyclization and integration capability of GI8

As the C3-41 isolate contained two copies of GI8, we determined the excision/cyclization capability of GI8 in G725 (a pBsph-cured derivative strain of C3-41) and in KellenQ (pBC16, pBsphΔBin), the latter of which is a transconjugant-carrying pBsphΔBin originating from C3-41, representing chromosomal-born and plasmid-born locations respectively. As illustrated in Fig. 3A, the chro-b/c primer pair was used to capture the DNA band containing the cyclized *attI* sequences (Fig. 3Bf). The PCR product sequences matched the joint and cyclized sequences of the GI8 terminal, verifying that GI8 at both locations could be cyclized (Fig. 3C). Similarly, all 11 wild *L. sphaericus* isolates containing GI8 in their chromosomes were PCR-positive for chro-b/c (Fig. 3Bf).

The sequences of the obtained PCR products containing *attI* (by chro-b/c), *attL* (by chro-a/b), *attR* (by chro-c/d), and *attB* (by chro-a/d) were aligned, and their core sequences were recognized. The excision and cyclization of GI8 are illustrated in Fig. 6C. As shown in the figure, *attI* was 24 nt in total, consisting of a 2-nt target site (AT) flanked by an 11-nt imperfect inverted repeat (atttaatataaaatatgtgaagat). The integrated form in the chromosome indicated that a direct 'AT' repeat was generated to flank GI8, whereas, in the plasmid, only one intact 'AT' was observed on one side. In addition, *attB* was 18 nt in total (agacgcccaatttatatatt), in which the 2-nt target site was flanked by 8-nt DNA on the left and right respectively. The *attL* and *attR* sequences were agacgcccaatatgtgaagat and atttaatataaaatttatatatt respectively.

Remarkably, based on PCR screening, two (represented by B36) of the 186 colonies of the transconjugant KellenQ (pBC16, pBsphΔBin) were negative for *attB* in contrast to the positive signals in the other 184 colonies

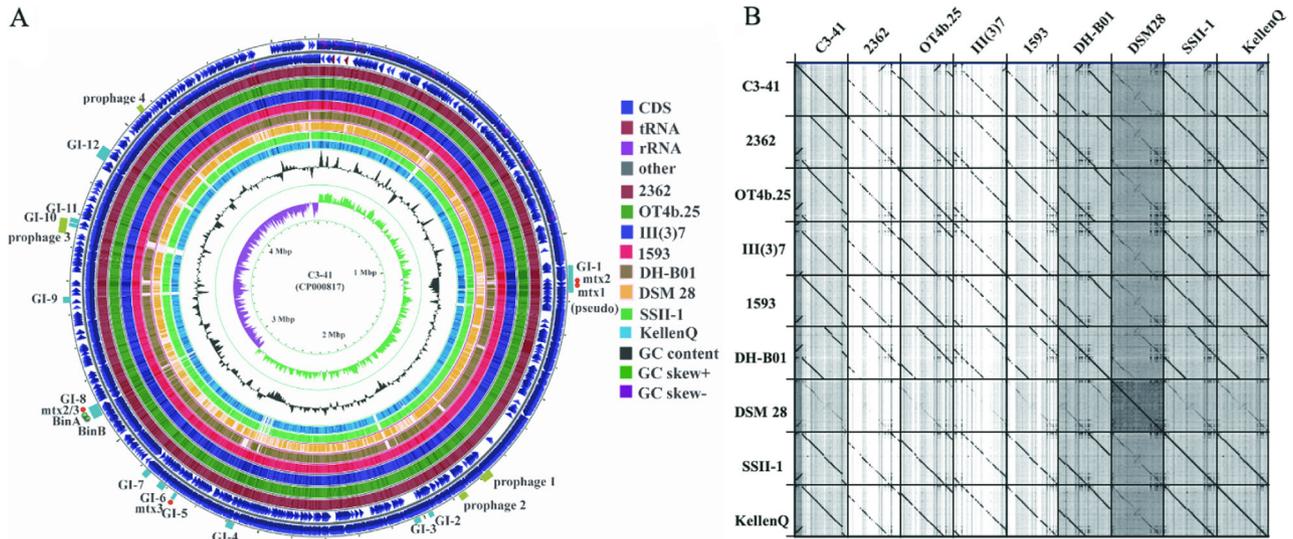


Fig 4. Chromosomal comparison of nine *L. sphaericus* strains.

A. Genomic island prediction and comparative analysis of nine *L. sphaericus* genomes by GCview. C3-41 was used as a reference at outermost circle. Genomic islands and prophages are shown as blue and olive rectangles; green and red solid circles represent binary toxins and toxins produced at vegetative period, respectively.

B. Pairwise dot-plot analysis comparison using Gepard (v1.40). [Color figure can be viewed at wileyonlinelibrary.com]

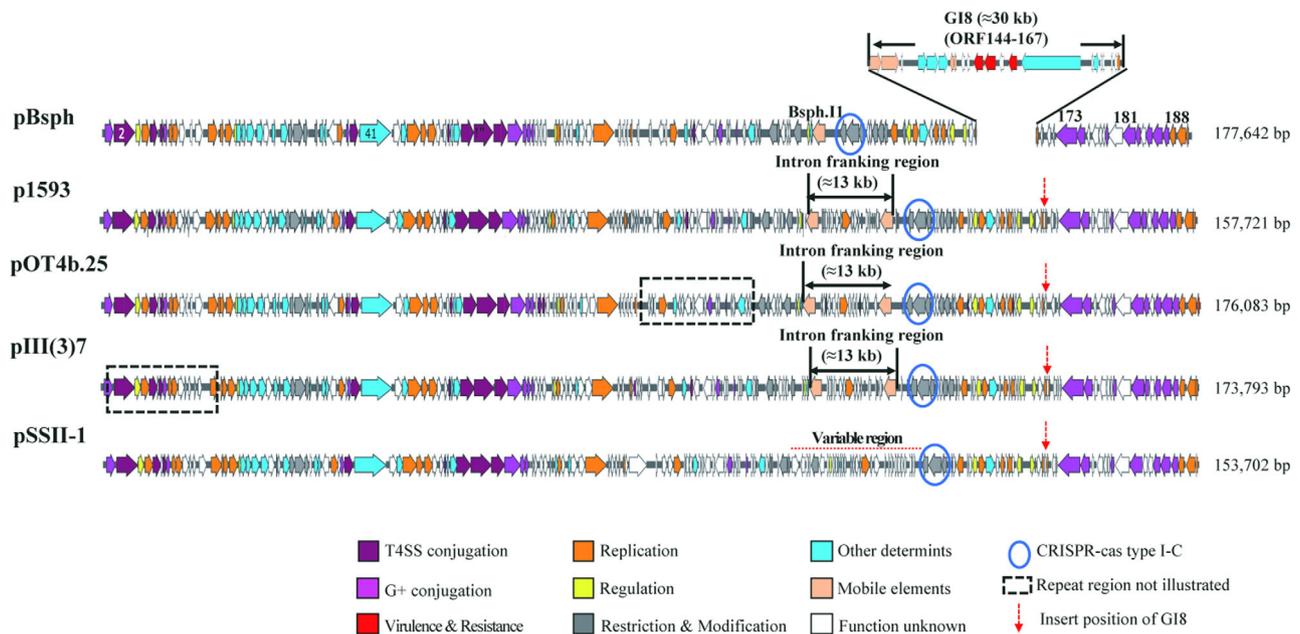


Fig 5. Sequencing comparison of G18 insertions in pBsph and pBsph-like plasmids. Predicted ORFs with different functions are indicated by different coloured arrows, with the region containing a CRISPR-Cas remnant highlighted with a blue ellipse. The 18-kb and 15-kb duplicated fragments presented in pOT4b.25 and pIII(3)7 are marked in black rectangles. Insertion position is indicated by a dotted red arrow. [Color figure can be viewed at wileyonlinelibrary.com]

(represented by A34); in contrast, all were positive for *attI* and the plasmid replicon gene. Data indicated that *attB* in B36 may be disrupted by the integration of the excised G18, which was verified by sequencing (Supporting Information Fig. S1A and B). This suggests that G18 may be co-transferred with the integrated pBsph plasmid from a

toxic donor into a recipient that is non-toxic to mosquitoes, followed by excision and reintegration into *attB* in the chromosome of the new host.

Production of the cyclized *attI*-containing band was on a similar level during all tested phases from 2 to 36 h no matter whether in Luria-Bertani (LB) medium or

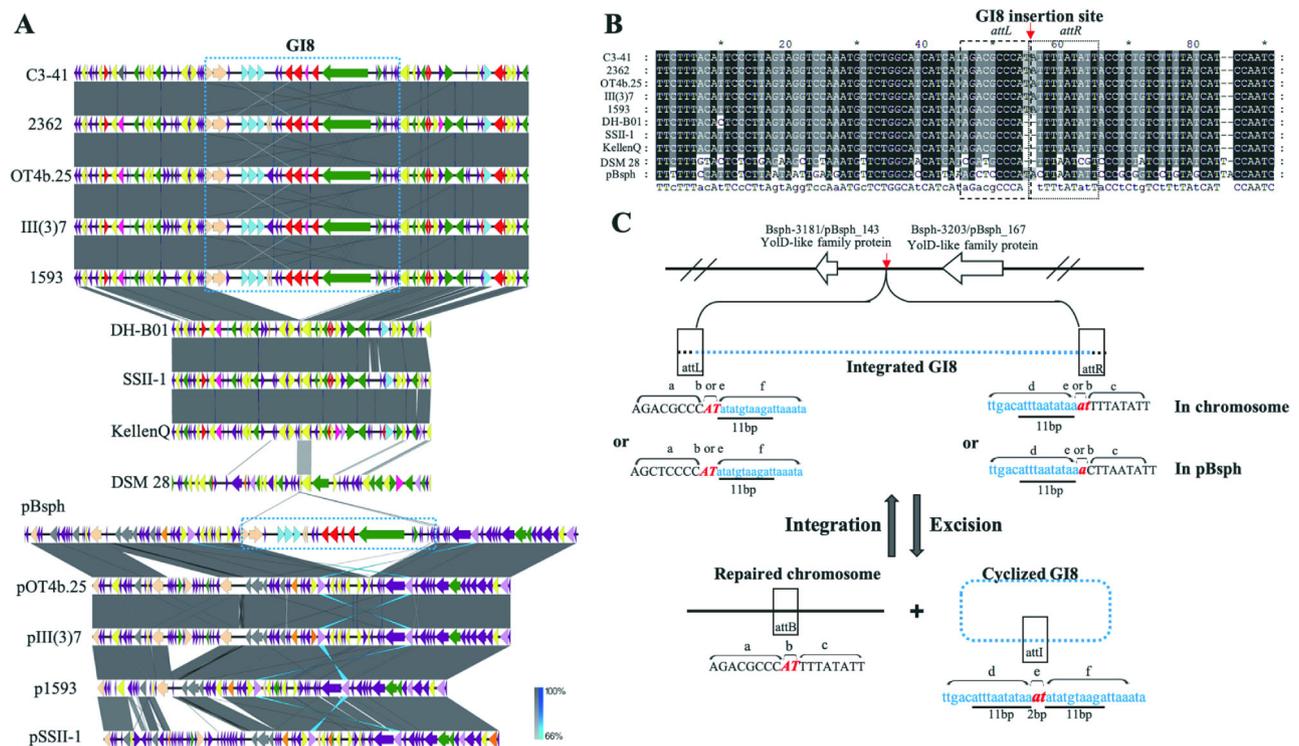


Fig 6. Comparison of flanking sequences of *attB* in chromosomes and plasmids to identify site-specific excision/integration of G18.

A. Comparison of flanking sequences of *attB* using Easyfig (v2.2.2). Different coloured arrows indicate genes with various functions. Plum: cell structure related; orange: DNA replication, transcription, and translation; sky blue: spore associated; pink: transport related; yellow: regulation, modification, and repair; Tan: mobile elements; red: virulence and resistance; grey: restriction and modification system; green: other function; and white: unknown function genes. Blue dotted boxes indicate G18.

B. Multiple alignments of *attB* sequences in avirulent isolates and G18 flanking sequences in high toxic strains. Both *attL* and *attR* containing 8 nt are indicated by dashed and dotted boxes, respectively, and G18 insertion site is indicated by red arrows.

C. Schematic of site-specific excision/integration of G18 in *L. spaerificus* C3-41. Chromosome of C3-41 contains a 30 612-bp insertion between BspH_3181 (YolD family protein) and BspH_3205 (Y family DNA polymerase) loci, corresponding to region between pBspH_143–167. G18 insertion site is indicated by red arrows. Double-stranded DNA (rectangular box) and circular G18 (dotted rectangular box) on chromosomal target containing 2-nt forward repeat ‘AT’ at target site. Regions a, b, and c on chromosomal targets are enclosed and indicate (a) left side; (b) 2-nt forward repeat, and (c) right side. Regions d, e, and f in circular G18 are enclosed and indicate (d) 11-nt incomplete reverse repeat on left; (e) 2-nt forward repeat, and (f) 11-nt incomplete reverse repeat on right. *attL*, *attR*, *attB*, and *attI* are represented by rectangular boxes and were used to detect positions of integrated chromosomes, repaired chromosomes, and extra-chromosomal circular using primers chro-a, chro-b, chro-c, and chro-d respectively. Directions are indicated by arrows. [Color figure can be viewed at wileyonlinelibrary.com]

Table 4. Retro-conjugative capability of pBspH(-like) plasmids within *L. spaerificus*.

Donor	Recipient	Frequency ^a
G725Δ0498 (pBspHΔBin) (Spc ⁺ Km ⁺)	G725 (pBC16) (Tet ⁺)	1.4×10^{-3} $\pm 1.4 \times 10^{-3}$
IAB872 (pBspHΔBin) (Rif ⁺ Km ⁺)	G725Δ0498 (Spc ⁺)	5.0×10^{-8} $\pm 6.2 \times 10^{-8}$
KellenQ (pBC16, pBspHΔBin) (Tet ⁺ Km ⁺)	G725Δ0498 (Spc ⁺)	2.2×10^{-7} $\pm 3.5 \times 10^{-7}$
G725Δ0498 (p2362Δ9-10) (Spc ⁺ Km ⁺)	G725 (pBC16) (Tet ⁺)	1.5×10^{-3} $\pm 1.1 \times 10^{-3}$
G725Δ0498 (p1593Δ9-10) (Spc ⁺ Km ⁺)	G725 (pBC16) (Tet ⁺)	2.1×10^{-4} $\pm 1.1 \times 10^{-4}$

^aThree independent mating experiments were performed. The frequency was determined as transconjugants/donor \pm SE.

sporulation-promoting minimal basal salts (MBS) medium. After treatment with different concentrations of mitomycin C (MMC) (0, 0.25, 0.5, 1, and 2 $\mu\text{g ml}^{-1}$) during growth, no

significant differences in the amounts of *attI* were observed either. These data suggest that the excision/cyclization of G18 can occur during whole growth phases independent of the nutrition and MMC-induction conditions. In contrast, when using a series of gradient amounts of template DNA (0, 12.5, 25, 50, 100, and 200 ng of genomic DNA), *attI* production increased gradually (Fig. 7), indicating that higher cell density can produce higher cyclized G18 production. Under the same conditions, primers chro-a/d were used to investigate the production of *attB*-containing products. However, compared to the strong bands containing intact *attB* observed in the positive control, in most cases, only weak and non-specific bands of predicted size were observed. This may be due to the mixed templates, which consisted of interrupted *attB* (complete chromosome integrated with G18) and intact *attB* (incomplete chromosome with G18 excised).

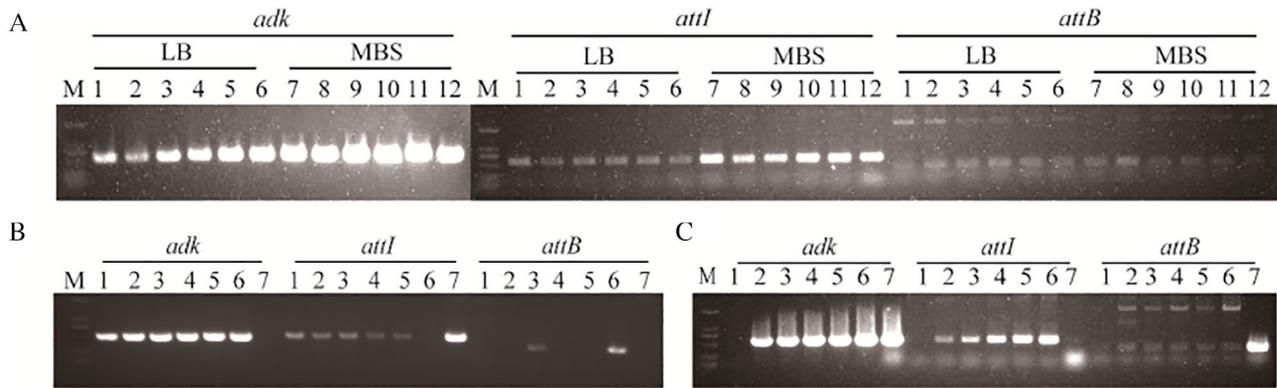


Fig 7. Detection of cyclization/excision under different growth conditions. Excision and circulation of GI8 in *L. sphaericus* IAB881 was measured by production of *attB* and *attI* using PCR. Detection of house-keeping gene *adk* was used as a reference.

A. Excision and circulation of GI8 in different media at various phases. Lanes 1–6: in LB medium at 2, 5, 8, 12, 24, and 36 h, respectively; Lanes 7–12: in MBS medium at 2, 5, 8, 12, 24, and 36 h respectively.

B. Excision and circulation of GI8 at different MMC concentrations. Lanes 1–5: MMC at final concentrations of 0, 0.25, 0.5, 1, and 2 $\mu\text{g ml}^{-1}$, respectively; Lanes 6–7: PCR products using genome of KellenQ or plasmid pMD-18T-*attI* as templates for control.

C. Excision and circulation of GI8 with different DNA template contents. Lanes 1–6: 0, 12.5, 25, 50, 100, and 200 ng of IAB881 genomic DNA respectively. Lane 7: KellenQ, used as a positive control of *attB* and negative control of *attI*. M: Trans 2K DNA Marker. [Color figure can be viewed at wileyonlinelibrary.com]

Intraspecies HGT capability of GI8

The conjugative transfer potential of GI8 within the *L. sphaericus* population was also tested. Two mutants with a kanamycin gene inserted into GI8, i.e., 1593-GI8-K and G725-GI8-K, were used as donors, with the former carrying the pBsph-like plasmid p1593 and the latter being plasmid cured. We found that both 1593-GI8-K and G725-GI8-K displayed an *attI* signal band, indicating that the insertion of kanamycin did not impair the excision/cyclization capability of GI8 (Supporting Information Fig. S2). However, neither donor produced transconjugants in the mating experiments when the resistant mutants of *L. sphaericus* containing the *attB* site (2314-2, 2173, Cok31, SSII-1, and KellenQ) or not containing the *attB* site (G725 Δ 0498) were used as recipients. This indicated that GI8 was likely not self-conjugative nor mobilized by the conjugative plasmid p1593 in 1593-GI8-K, or the transfer frequency was too low to be detected ($<10^{-9}$ transconjugants/donor).

Discussion

The four genetic regions in pBsph related to T4SS and the classic ICE/IME features of GI8 (containing *xerC/D*, *attL*, *attR*, toxin, and antibiotic-related genes) identified in the current study suggest that pBsph may be a novel T4SS-type conjugative plasmid carrying a co-resident ICE/IME-type mosquitocidal genomic island. The main difference between pBsph and the pBsph-like plasmids is the presence of GI8 in pBsph, but its absence in the other plasmids, indicating that these plasmids share

identical backbones and may act as vehicles for the transmission of GI8.

Indeed, the intraspecies conjugative transfer capabilities of pBsph, p2362, and p1593 were detected and the transmissibility between homologous hosts with identical chromosomal backgrounds was higher than that between heterologous hosts. When using 2362, 1593, and 2317-2 as recipients, no transconjugants were observed, although their chromosomal backgrounds were more similar than those of the other *L. sphaericus* strains tested. This may be because these isolates already harbour a pBsph-like plasmid that can repel invasion of a similar foreign plasmid. Previous studies have suggested that, except for naturally R-M deficient 2297 and NRS1693, *L. sphaericus* is difficult to genetically manipulate (Berry, 2012). Indeed, we attempted but failed to tag a resistant marker in pSSII-1, and thus the conjugative capability of pSSII-1 was not characterized. C3-41 carries a LspC3-41 R-M system, which is the main barrier for genetic manipulation. Our previous study observed that heterologous plasmids (e.g., pBU4 and pBC16) are difficult to transform into C3-41 as they are restricted by the LspC3-41 R-M system. We found that this barrier can be overcome if the LspC3-41 RM system is knocked out or the heterologous plasmids are protected by methylation of isoschizomer HaeIII (Fu *et al.*, 2017b). The five candidate recipients of pBsph (i.e., IAB881, IAB872, Lp1-G, 2297, and KellenQ) lack the LspC3-41 R-M system, and the cell-free extracts of IAB881, IAB872, and Lp1-G can completely digest the heterologous plasmid DNA (Fu *et al.*, 2017b), suggesting there are unknown R-M system(s) hindering genetic manipulation, which still

remain to be clarified. However, we found that these isolates could accept pBsph and the transfer frequencies using wild C3-41 and G725Δ0498 (plasmid-cured mutant G725 with deletion of the *IspC3-41* R-M system in the chromosome) as donor backgrounds exhibited no obvious differences (data not shown). This indicated the occurrence of ssDNA transfer, which is resistant to active recipient R-M systems (Possoz *et al.*, 2001).

The Gl8-carrying binary operon and integrative elements exhibited excision and integration capability. No auto-replicating element was predicted in Gl8. We performed real-time PCR to detect the replication of Gl8 in different growth phases but this failed. Excision/cyclization was more frequent under poor nutrition conditions (in MBS medium) than under rich conditions (in LB medium) (Fig. 7A). These results may help to understand the Gl8 life cycle, which includes: (i) excision and reintegration within the same host (intracellular); or (ii) co-transfer with an integrated plasmid to another host, followed by excision and integration with a chromosome or plasmid with replication capability (intercellular), as it needs to reside on a replicon to avoid being lost during cell partitioning; it also needs to be excised and integrated into a vehicle to transmit itself to other cells if it cannot stably co-exist with the original host cell when suffering nutritional or other survival crises. However, the induction condition mediating the SOS-response (MMC supplement in this study) had no promotion effect on excision. Previous studies have shown that *umuDC* gene products can increase DNA damage tolerance by regulating growth after DNA damage in both exponentially growing and stationary phase cells (Murli *et al.*, 2000). Here, the integration of Gl8 into the *umuDC* gene operon in the chromosome and into the pBsph plasmid may help it counteract the effects of MMC induction. We further explored the location diversity of Gl8 in this study: e.g., (i) on chromosomes in isolates that lack pBsph-like plasmids (e.g., 2297 and LP1-G); (ii) on chromosomes in isolates that carry pBsph-like plasmids (e.g., 2362, 1593, OT4b.25, and pIII(3)7); and (iii) on both the chromosome and large plasmid (i.e., C3-41). All Gl8-carrying isolates shared highly conserved chromosomal backgrounds. In addition, although SSII-1 carried a pBsph-like plasmid, neither the chromosome nor the plasmid carried Gl8. Remarkably, SSII-1 and the plasmid-less isolate KellenQ carrying the potential integration site (*attB*) in the chromosome also showed good collinearity with the isolates carrying Gl8, and the latter displayed the ability to accept pBsph and the integration of Gl8 into the *attB* site. These data help to trace the evolutionary origin of the mosquitocidal *L. sphaericus* lineage. As the excision/integration of Gl8 is constitutive, the T4SS-type plasmid may be a vehicle for Gl8 if it carries the target (i.e., *attB*). Our study suggested a potential interaction between Gl8 and pBsph-like plasmids. This corresponds to several recent

studies, which have suggested that ICEs/IMEs may interact with the host as a possible strategy to avoid population loss when cyclization occurs during cell division (Carraro *et al.*, 2017; Burrus, 2017; Huguet *et al.*, 2020).

Except for the two large fragments, i.e., 30-kb Gl8 and 13-kb 'intron flanking region', pBsph and pBsph-like plasmids were strikingly similar. This implies that the acquisition of pBsph and pBsph-like plasmids may be emerging events among *L. sphaericus*, and the R-M system may have helped protect homogenization. Previous studies have proposed that IEPs mediate DNA translocation and are associated with intron-like mobility and may also hijack neighbouring DNA fragments for co-transfer (Guhan and Muniyappa, 2003). This suggests that, in addition to Gl8, the 'intron flanking region' may also originate from HGT and may be associated with Gl8 acquisition. Interestingly, the 'intron flanking region' is neighbored by an incomplete type I-C CRISPR-Cas system. Recent studies have found that most CRISPR-Cas systems carried by the *B. cereus* group are deficient; intact and activated CRISPR-Cas systems may act as barriers to HGT, and their inactivation may help in the acquisition of MGEs and in adaptation to diverse environments (Zheng *et al.*, 2020). IEPs and intron-like intervening sequences thought to be involved in genetic mobility at the DNA or RNA level have been identified in other conjugative plasmids (e.g., B.th.11 and B.th.12 in pAW63 and pHT9727) but have not yet been well characterized (Van der Auwera and Mahillon, 2005, 2008). Whether Bsph.11 has mobility, how pBsph and other MGEs transfer, and what genetic elements are involved remain to be discovered.

As the transmission of plasmids and Gl8 is dynamic in natural environments, we speculate that the commercial *L. sphaericus* pesticides (e.g., VectoLex[®], Spherimos[®], Spicbiomoss[®], VectoLex G[®], and Jianbao[®]) applied in nature may produce mixed configurations with identical backbone elements. Jianbao[®] was originally developed from *L. sphaericus* C3-41, a unique isolate with Gl8 duplicates obtained from a mosquito-breeding site (Yuan *et al.*, 1999), which was probably driven by horizontal transmission of Gl8 and pBsph within dynamic *L. sphaericus* communities in the natural environment. Understanding how Gl8 is activated and excised and whether crosstalk exists between Gl8, pBsph, and the host will help clarify the ecology and hidden lifecycle of Gl8.

Materials and methods

Bacterial strains and growth conditions

The *L. sphaericus* strains used in this study are listed in Table 3 (de Barjac, 1985; Ge *et al.*, 2011). *Escherichia coli* JM109 was used for cloning the constructed plasmids and vectors. The *L. sphaericus* and *E. coli* cells were grown in LB medium at 30°C and 37°C

respectively. In addition, MBS (0.68% KH_2PO_4 , 1% tryptone, 0.2% yeast extract, 0.03% $\text{MgSO}_4 \cdot 7\text{H}_2\text{O}$, 0.002% $\text{CaCl}_2 \cdot 2\text{H}_2\text{O}$, 0.002% $\text{MnSO}_4 \cdot \text{H}_2\text{O}$, 0.002% $\text{ZnSO}_4 \cdot 7\text{H}_2\text{O}$, 0.002% $\text{Fe}_2(\text{SO}_4)_3$; pH 7.2) for sporulation was used to examine the effects of different nutrient conditions and growth phases on the induction of GI8 excision/integration in *L. sphaericus*. The antibiotic concentrations used were as follows: 50 $\mu\text{g ml}^{-1}$ kanamycin and 100 $\mu\text{g ml}^{-1}$ spectinomycin for *E. coli*; 30 $\mu\text{g ml}^{-1}$ kanamycin, 25 $\mu\text{g ml}^{-1}$ erythromycin, 200 $\mu\text{g ml}^{-1}$ spectinomycin, 10 $\mu\text{g ml}^{-1}$ tetracycline, and 100 $\mu\text{g ml}^{-1}$ rifampicin for *L. sphaericus*.

Primers and PCR

All primers used in this study are listed in the Supporting Information Table S1. Based on bioinformatics analysis of the genetic structure of GI8, four primer pairs (i.e., chro-a/b, chro-c/d, xerC-F/R, and xerD-F/R) were designed to detect the presence of typical elements in GI8. Here, chro-a/b and chro-c/d were designed based on the left and right ends of GI8 and were used for identification of the *attL* and *attR* sequences, respectively; xerC-F/R and xerD-F/R were used to amplify the predicted *xerC* and *xerD* genes in GI8, respectively. The PCR results positive for all four primer pairs were suggested to carry GI8. Moreover, the outward-facing primer pair chro-b/c was used to capture the cyclized GI8 linker, which contained the *attI* sequence; the inward-facing primer pair chro-a/d was used to detect the potential *attB* sequence in *L. sphaericus*, as illustrated in Fig. 3A.

DNA manipulation and construction of *L. sphaericus* mutants

For screening transconjugants generated by mating experiments, an antibiotic resistant gene was tagged into an ORF not related to conjugation within the tested mobile elements (pBsph/-like plasmids or GI8) by homologous recombination. Briefly, a resistance gene was sandwich assembled into the two flanking fragments of the deletion region amplified from the *L. sphaericus* C3-41 genome by PCR and cloned into the temperature-sensitive suicide plasmid pRN5101. Each pRN5101-derivate was methylated before electroporation as the functional R-M system in *L. sphaericus* C3-41 displays restriction activity to foreign DNA (Fu *et al.*, 2017b). The mutation occurred by double-crossover events during passage and was screened under sub-lethal temperature (42°C) to cure the pRN5101-derivate, followed by confirmational PCR analysis. By this method, pRN5101-GI8-K was constructed to insert a kanamycin-resistant marker in GI8 between Bsph_3196–3197 in the chromosomes of *L. sphaericus* G725 and 1593 respectively. The resulting mutants G725-GI8-K and 1593-GI8-K were used to explore the

HGT capability of GI8 located on the chromosome. C3-41 (pBsph Δ Bin), a similarly built mutant from previous study, in which a kanamycin gene was used to replace the binary toxin gene in GI8 and pBsph in C3-41 (Fu *et al.*, 2017a), was used as a donor to evaluate the transmissibility of pBsph. In addition, the constructed plasmid pRN9-10 (Fu *et al.*, 2018) was electroporated into *L. sphaericus* 2362, 1593, and SSII-1, and the transformants in which a kanamycin resistance gene was used to replace the ORF9-10 operon on plasmids not involved in transmission were used as donors to evaluate the transmissibility of pBsph-like plasmids p2362, p1593, and pSSII-1. Furthermore, plasmid pRN-MS0498 was constructed and then transformed into G725, a plasmid-cured derivative of *L. sphaericus* C3-41 (Ge *et al.*, 2014). The G725 Δ 0498 mutant generated by the replacement of Bsph_0498 (which encodes a R-M-related REase with a spectinomycin resistance marker) was used as the recipient in the mating experiments. Furthermore, rifampicin-resistant mutants were developed via spontaneous selection by culturing *L. sphaericus* cells under a series of gradually increasing sub-lethal concentrations of antibiotics, as described previously (Hu *et al.*, 2004). These spontaneous mutants were also used as recipients to survey the host range of pBsph and pBsph-like plasmids. Furthermore, to investigate the mobility of pBsph, the mobilizable plasmid pBC16 (from *Bacillus cereus*) carrying tetracycline resistance genes (Hu *et al.*, 2009) was methylated and then electroporated into *L. sphaericus*, with the resulting transformants then used as recipients.

Mating experiments

Conjugational transfer experiments were performed by colony (drop-to-drop) matings, as described previously (Hinneken *et al.*, 2019). In brief, 5 μl of overnight precultures of donor and recipient cells was placed on the same spot of a LB plate without antibiotics and incubated at 30°C for 4 h. The resulting colony was collected and resuspended in LB medium and then sprayed on the plate with double-selected antibiotics. The colonies grown after overnight incubation at 30°C were selected. Two rounds of PCR were performed against the recipient background and donor-derived transfer elements to exclude spontaneous mutants. All mating experiments were repeated three times independently. The transfer frequencies were calculated as the ratio of transconjugants to donor cells (T/D).

Detection of excision/cyclization and integration of GI8

The excision and cyclization of GI8 were determined by detecting the PCR products containing the *attB* and *attI*

sequences. Overnight bacteria were, respectively, inoculated into different liquid cultures at a dilution of 1:100, including LB and MBS media. The bacteria collected at different growth phases (2, 5, 8, 12, 24, and 36 h) were used for extraction of total DNA, with an equal amount of template (50 ng/10 μ l reaction). In addition, to investigate whether MMC promotes excision/circulation, cells pre-cultured in MBS medium for 5 h were supplemented with different concentrations of MMC (0, 0.25, 0.5, 1, and 2 μ g ml⁻¹), and induced for 1 h before DNA extraction. The house-keeping gene *adk* was detected by PCR in parallel as a reference. Furthermore, gradient template concentrations (0, 12.5, 25, 50, 100, and 200 ng of IAB881 genomic DNA) were used to detect the products produced by the excision and circulation of GI8.

To detect the integration capability of GI8, the trans-conjugant KellenQ (pBC16, pBsph Δ Bin) was passage cultured, serially diluted, and spread onto LB plates. In total, 186 single colonies were randomly selected, and primer pairs KellenQ-GI-F1/R1, KellenQ-GI-F2/R2, KellenQ-GI-F1/R2, and chro-b/c (to tag the flanking sequences of *attB* before and after GI8 integration in KellenQ) were used to detect whether the intact *attB* site in KellenQ (pBC16, pBsph Δ Bin) was damaged and whether integration of GI8 occurred.

Plasmid stability of transconjugants

The stability of pBsph Δ Bin in the transconjugants was investigated, as described previously (Hu *et al.*, 2004). Overnight cultures of each transconjugant were diluted 100-fold in fresh LB medium without antibiotics at 30°C and grown for 6 h. The dilution and inoculation were repeated 40–50 times every 6 h, and then appropriate dilutions were plated on non-selective LB plates. After overnight incubation at 30°C, 100 colonies were randomly selected and inoculated on LB plates with antibiotics. Stability was estimated as the percentage of cells carrying corresponding resistance.

Plasmid profile analysis

The large plasmids were extracted according to previously described methods (Andrup *et al.*, 2008). DNA was analysed by horizontal agarose gel electrophoresis (8–10 V cm⁻¹) in 0.5% agarose (SeaKem GTG) with 1 \times TBE buffer for 5 h in a precooled electrophoresis chamber.

Sequencing and bioinformatics analysis

The ORFs of pBsph were blasted using public databases (Nr and SwissProt) and those with similarity to T4SS and G+ conjugation were manually annotated. The ORFs

within GI8 were blasted with ICEberg 2.0 (<https://db-mml.sjtu.edu.cn/ICEberg/>) (Liu *et al.*, 2019). The *attL*, *attR*, *attI*, and *attB* sequences were determined by comparing the flanking and/or joint sequences of GI8 before and after excision.

The genomes of 1593, SSII-1, and KellenQ were sequenced according to standard Oxford Nanopore Technologies (ONT) (Benagen, Wuhan, China). Unicycler (v0.4.8) and Prokka were used for read assembly and genome annotation. CRISPR, GI, and prophage predictions were performed with CRISPRfinder (<https://crispr.i2bc.paris-saclay.fr/>), IslandViewer 4 (<http://www.pathogenomics.sfu.ca/islandviewer/>), and PHASTER (<http://phaster.ca/>), respectively (Grissa *et al.*, 2007; Arndt *et al.*, 2016; Bertelli *et al.*, 2017). The chromosomal sequence comparisons of 1593 (CP064068), SSII-1 (CP064070), and KellenQ (CP064067) obtained in this study, and those downloaded from public databases, including 2362 (CP015224), OT4b.25 (NZ_CP014643), III(3)7 (NZ_CP014856), DH-B01 (CP045583), and DSM 28 (NZ_CP019980), were conducted using GCview (http://stothard.afns.ualberta.ca/cgview_server/), Gepard (v1.40), and Easyfig v2.2.3 (Krumstiek *et al.*, 2007; Grant and Stothard, 2008; Sullivan *et al.*, 2011). In addition, the sequences of the pBsph-like plasmids p1593 (CP064069) and pSSII-1 (CP064071) obtained in this study, and pOT4b.25 (CP014644) and pIII(3)7 (CP014857) downloaded from public databases were compared with pBsph (CP000818) using local blast.

Acknowledgements

This project was supported by the National Key R&D Program (2017YFD0200400 and 2018ZX10101004), National Natural Science Foundation of China (NFSC) (31570007), Hubei Province Innovation Major Program (2018ABA093), and ‘Fundamental Research Funds for the Central Universities’ and ‘Talent Introduction Project’ of South-Central University for Nationalities (CZZ20001 and YZZ19016).

References

- Alvarez-Martinez, C.E., and Christie, P.J. (2009) Biological diversity of prokaryotic type IV secretion systems. *Microbiol Mol Biol Rev* **73**: 775–808.
- Andrup, L., Barfod, K.K., Jensen, G.B., and Smidt, L. (2008) Detection of large plasmids from the *Bacillus cereus* group. *Plasmid* **59**: 139–143.
- Arndt, D., Grant, J.R., Marcu, A., Sajed, T., Pon, A., Liang, Y., and Wishart, D.S. (2016) PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res* **44**: W16–W21.
- Berry, C. (2012) The bacterium, *Lysinibacillus sphaericus*, as an insect pathogen. *J Invertebr Pathol* **109**: 1–10.
- Bertelli, C., Laird, M.R., Williams, K.P., Simon Fraser University Research Computing, G, Lau, B.Y., Hoad, G., *et al.*

- (2017) IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res* **45**: W30–W35.
- Botelho, J., and Schulenburg, H. (2021) The role of integrative and conjugative elements in antibiotic resistance evolution. *Trends Microbiol* **29**: 8–18.
- Burrus, V. (2017) Mechanisms of stabilization of integrative and conjugative elements. *Curr Opin Microbiol* **38**: 44–50.
- Carraro, N., Durand, R., Rivard, N., Anquetil, C., Barrette, C., Humbert, M., and Burrus, V. (2017) *Salmonella* genomic Island 1 (SGI1) reshapes the mating apparatus of IncC conjugative plasmids to promote self-propagation. *PLoS Genet* **13**: e1006705.
- de la Cruz, F., Frost, L.S., Meyer, R.J., and Zechner, E.L. (2010) Conjugative DNA metabolism in Gram-negative bacteria. *FEMS Microbiol Rev* **34**: 18–40.
- Durrant, M.G., Li, M.M., Siranosian, B.A., Montgomery, S.B., and Bhatt, A.S. (2020) A bioinformatic analysis of integrative mobile genetic elements highlights their role in bacterial adaptation. *Cell Host Microbe* **27**: 140–153.
- Fu, P., Xiang, X., Ge, Y., Yuan, Z., and Hu, X. (2017a) Differential expression of duplicated binary toxin genes binA/binB in *Lysinibacillus sphaericus* C3-41. *Lett Appl Microbiol* **65**: 90–97.
- Fu, P., Ge, Y., Hu, Y., Yuan, Z., and Hu, X. (2018) A toxin-antitoxin system is essential for the stability of mosquitocidal plasmid pBsph of *Lysinibacillus sphaericus*. *Microbiol Res* **214**: 114–122.
- Fu, P., Ge, Y., Wu, Y., Zhao, N., Yuan, Z., and Hu, X. (2017b) The LspC3-41I restriction-modification system is the major determinant for genetic manipulations of *Lysinibacillus sphaericus* C3-41. *BMC Microbiol* **17**: 116.
- Garcillan-Barcia, M.P., Francia, M.V., and de la Cruz, F. (2009) The diversity of conjugative relaxases and its application in plasmid classification. *FEMS Microbiol Rev* **33**: 657–687.
- Ge, Y., Hu, X., Zheng, D., Wu, Y., and Yuan, Z. (2011) Allelic diversity and population structure of *Bacillus sphaericus* as revealed by multilocus sequence typing. *Appl Environ Microbiol* **77**: 5553–5556.
- Ge, Y., Hu, X., Zhao, N., Shi, T., Cai, Q., and Yuan, Z. (2014) A new tubRZ operon involved in the maintenance of the *Bacillus sphaericus* mosquitocidal plasmid pBsph. *Microbiology (Reading)* **160**: 1112–1124.
- Grant, J.R., and Stothard, P. (2008) The CGView server: a comparative genomics tool for circular genomes. *Nucleic Acids Res* **36**: W181–W184.
- Grissa, I., Vergnaud, G., and Pourcel, C. (2007) CRISPRFinder: a web tool to identify clustered regularly interspaced short palindromic repeats. *Nucleic Acids Res* **35**: W52–W57.
- Grohmann, E., Christie, P.J., Waksman, G., and Backert, S. (2018) Type IV secretion in Gram-negative and Gram-positive bacteria. *Mol Microbiol* **107**: 455–471.
- Guédon, G., Libante, V., Coluzzi, C., Payot, S., and Leblond-Bourget, N. (2017) The obscure world of integrative and mobilizable elements, highly widespread elements that pirate bacterial conjugative systems. *Genes (Basel)* **8**: 337.
- Guhan, N., and Muniyappa, K. (2003) Structural and functional characteristics of homing endonucleases. *Crit Rev Biochem Mol Biol* **38**: 199–248.
- de Barjac, H., Larget-Thiéry, I., Cosmao Dumanoir, V., and Ripouteau, H. (1985) Serological classification of *Bacillus sphaericus* strains on the basis of toxicity to mosquito larvae. *Appl Microbiol Biotechnol* **21**: 85–90.
- Hernandez-Santana, A., Gomez-Garzon, C., and Dussan, J. (2016) Complete genome sequence of *Lysinibacillus sphaericus* WHO reference strain 2362. *Genome Announc* **4**: 2169–8287.
- Hinneken, P., Kone, K.M., Fayad, N., Leprince, A., and Mahillon, J. (2019) pXO16, the large conjugative plasmid from *Bacillus thuringiensis* serovar *israelensis* displays an extended host spectrum. *Plasmid* **102**: 46–50.
- Hu, X., van der Auwera, G., Timmerly, S., Zhu, L., and Mahillon, J. (2009) Distribution, diversity, and potential mobility of extrachromosomal elements related to the *Bacillus anthracis* pXO1 and pXO2 virulence plasmids. *Appl Environ Microbiol* **75**: 3016–3028.
- Hu, X., Hansen, B.M., Eilenberg, J., Hendriksen, N.B., Smidt, L., Yuan, Z., and Jensen, G.B. (2004) Conjugative transfer, stability and expression of a plasmid encoding a cry1Ac gene in *Bacillus cereus* group strains. *FEMS Microbiol Lett* **231**: 45–52.
- Hu, X., Fan, W., Han, B., Liu, H., Zheng, D., Li, Q., et al. (2008) Complete genome sequence of the mosquitocidal bacterium *Bacillus sphaericus* C3-41 and comparison with those of closely related *Bacillus* species. *J Bacteriol* **190**: 2892–2902.
- Huguet, K.T., Rivard, N., Gameau, D., Palanee, J., and Burrus, V. (2020) Replication of the *Salmonella* genomic Island 1 (SGI1) triggered by helper IncC conjugative plasmids promotes incompatibility and plasmid loss. *PLoS Genet* **16**: e1008965.
- Johnson, C.M., and Grossman, A.D. (2015) Integrative and conjugative elements (ICEs): what they do and how they work. *Annu Rev Genet* **49**: 577–601.
- Kohler, V., Keller, W., and Grohmann, E. (2019) Regulation of gram-positive conjugation. *Front Microbiol* **10**: 1134.
- Krumsiek, J., Arnold, R., and Rattei, T. (2007) Gepard: a rapid and sensitive tool for creating dotplots on genome scale. *Bioinformatics* **23**: 1026–1028.
- Liu, M., Li, X., Xie, Y., Bi, D., Sun, J., Li, J., et al. (2019) ICEberg 2.0: an updated database of bacterial integrative and conjugative elements. *Nucleic Acids Res* **47**: D660–D665.
- Murli, S., Opperman, T., Smith, B.T., and Walker, G.C. (2000) A role for the umuDC gene products of *Escherichia coli* in increasing resistance to DNA damage in stationary phase by inhibiting the transition to exponential growth. *J Bacteriol* **182**: 1127–1135.
- Norman, A., Hansen, L.H., and Sorensen, S.J. (2009) Conjugative plasmids: vessels of the communal gene pool. *Philos Trans R Soc Lond B Biol Sci* **364**: 2275–2289.
- Parmeciano, D.I., Noto, G., Iriarte, A., Ramírez, M.S., Centrón, D., and Quiroga, C. (2019) ICE SXT vs. ICESh95: co-existence of integrative and conjugative elements and competition for a new host. *Sci Rep* **9**: 8045.
- Permina, E.A., Mironov, A.A., and Gelfand, M.S. (2002) Damage-repair error-prone polymerases of eubacteria: association with mobile genome elements. *Gene* **293**: 133–140.

- Possoz, C., Ribard, C., Gagnat, J., Pernodet, J.L., and Guerineau, M. (2001) The integrative element pSAM2 from *Streptomyces*: kinetics and mode of conjugal transfer. *Mol Microbiol* **42**: 159–166.
- Rey, A., Silva-Quintero, L., and Dussan, J. (2016a) Complete genome sequence of the larvicidal bacterium *Lysinibacillus sphaericus* strain OT4b.25. *Genome Announc* **4**: 2169–8287.
- Rey, A., Silva-Quintero, L., and Dussan, J. (2016b) Complete genome sequencing and comparative genomic analysis of functionally diverse *Lysinibacillus sphaericus* III(3) 7. *Genom Data* **9**: 78–86.
- Sullivan, M.J., Petty, N.K., and Beatson, S.A. (2011) Easyfig: a genome comparison visualizer. *Bioinformatics* **27**: 1009–1010.
- Su, T., Thieme, J., Ocegueda, C., Ball, M., and Cheng, M.L. (2018) Resistance to *Lysinibacillus sphaericus* and other commonly used pesticides in *Culex pipiens* (Diptera: Culicidae) from Chico, California. *J Med Entomol* **55**: 423–428.
- van der Auwera, G., and Mahillon, J. (2005) TnXO1, a germination-associated class II transposon from *Bacillus anthracis*. *Plasmid* **53**: 251–257.
- Van der Auwera, G., and Mahillon, J. (2008) Transcriptional analysis of the conjugative plasmid pAW63 from *Bacillus thuringiensis*. *Plasmid* **60**: 190–199.
- van der Auwera, G.A., Andrup, L., and Mahillon, J. (2005) Conjugative plasmid pAW63 brings new insights into the genesis of the *Bacillus anthracis* virulence plasmid pXO2 and of the *Bacillus thuringiensis* plasmid pBT9727. *BMC Genomics* **6**: 103.
- Waksman, G., and Fronzes, R. (2010) Molecular architecture of bacterial type IV secretion systems. *Trends Biochem Sci* **35**: 691–698.
- Wang, Y., Zhao, X., Tian, K., Meng, F., Zhou, D., Xu, X., et al. (2020) Identification and genome analysis of a novel 17beta-estradiol degradation bacterium, *Lysinibacillus sphaericus* DH-B01. *3 Biotech* **10**: 166.
- Xiong, W., Hu, X., and Yuan, Z. (2010) Application of *Bacillus sphaericus* in mosquito and vector control. *Chin J Vector Biol Control* **1**: 1–4.
- Xu, K., Yuan, Z., Rayner, S., and Hu, X. (2015) Genome comparison provides molecular insights into the phylogeny of the reassigned new genus *Lysinibacillus*. *BMC Genomics* **16**: 140.
- Yang, Y., Wang, L., Gaviria, A., Yuan, Z., and Berry, C. (2007) Proteolytic stability of insecticidal toxins expressed in recombinant bacilli. *Appl Environ Microbiol* **73**: 218–225.
- Yuan, Z., Neilsen-LeRoux, C., Pasteur, N., Delecluse, A., Charles, J.F., and Frutos, R. (1999) Cloning and expression of the binary toxin genes of *Bacillus sphaericus* C3-41 in a crystal minus *B. thuringiensis* subsp. *israelensis*. *Wei Sheng Wu Xue Bao* **39**: 29–35.
- Zheng, Z.Q., Zhang, Y.L., Liu, Z.Y., Dong, Z.X., Xie, C.S., Bravo, A., et al. (2020) The CRISPR-Cas systems were selectively inactivated during evolution of *Bacillus cereus* group for adaptation to diverse environments. *ISME J* **14**: 1479–1493.

Supporting Information

Additional Supporting Information may be found in the online version of this article at the publisher's web-site:

Figure S1 Integration of Gl8::kan into attB site in chromosome of *L. sphaericus* KellenQ. (A) PCR detection of *Bs-tubR*, *attI*, and *attB* is shown as Lanes 1–3, respectively. B36 is a colony screened from transconjugant KellenQ (pBC16, pBsphΔBin) with Gl8 insertion at *attB* in chromosome; and A34 is a colony of transconjugant KellenQ (pBC16, pBsphΔBin) without Gl8 insertion. Primer pair pTubR-F/R was used for detection of replicon of pBsph, *chro-b/c* for *attI*, and KellenQ-GI-F1/R2 for *attB* interruption by Gl8. (B) Sequencing verification of Gl8 cyclization and integration capability. Products *attB* (KellenQ-GI-F1/R2), *attI* (*chro-b/c*), *attL* (KellenQ-GI-F1/R1), and *attR* (KellenQ-GI-F2/R2) were sequenced to verify if intact *attB* in KellenQ is interrupted by Gl8::kan from pBsphΔBin. Blue and orange arrows indicate sequences from KellenQ chromosome and Gl8. Red box containing AT and regions a-f represent core sequences for excision/integration, as illustrated in Fig. 3C.

Figure S2 Detection of cyclization/excision and pBsph-like plasmids of Gl8-containing donors. Excision and circulation were measured by production of *attI* via PCR. Detection of house-keeping gene *adk* was used as a reference. Lanes 1–5 (*adk*): G725, G725-GI8-K, 1593, 1593-GI8-K, negative control; Lanes 6–10 (*tubR*): G725, G725-GI8-K, 1593, 1593-GI8-K, negative control; Lanes 11–15 (*attI*): G725, G725-GI8-K, 1593, 1593-GI8-K, negative control. M: Trans 2 K DNA Marker.

Table S1 Primers used in this study.