Evolutionary Genomics of *Bacillus thuringiensis*

Dissertation

for the award of the degree
"Doctor rerum naturalium"
of the Georg-August-Universität Göttingen

within the doctoral program Biology
of the Georg-August University School of Science (GAUSS)

submitted by

Jacqueline Hollensteiner

from Detmold

Göttingen, 2017

Thesis committee

Prof. Dr. Rolf Daniel, Department of Genomic and Applied Microbiology, Institute of Microbiology and Genetics, Georg-August University Göttingen

PD Dr. Michael Hoppert, Department of General Microbiology, Institute of Microbiology and Genetics, Georg-August University Göttingen,

Dr. Heiko Liesegang, Department of Genomic and Applied Microbiology, Institute of Microbiology and Genetics, Georg-August University Göttingen

Member of the Examination Board

Referent: **Prof. Dr. Rolf Daniel**, Department of Genomic and Applied Microbiology, Institute of Microbiology and Genetics, Georg-August University Göttingen

Coreferent: **PD Dr. Michael Hoppert**, Department of General Microbiology, Institute of Microbiology and Genetics, Georg-August University Göttingen

Additional Members

Prof. Dr. Kai Heimel Department of Molecular Microbiology and Genetics, Institute of Microbiology and Genetics, Georg-August University Göttingen

Prof. Dr. Gerhard Braus, Department of Molecular Microbiology and Genetics, Institute of Microbiology and Genetics, Georg-August University Göttingen

PD. Dr. Kramer, Department of Molecular Genetics, Institute of Microbiology and Genetics, Georg-August University Göttingen

Prof Dr. Stefan Vidal, Department of Crop Science, Agricultural Entomology, Georg-August University Göttingen

Date of oral examination: 21.02.2017

List of publications

1. **Complete Genome Sequence of the nematicidal *Bacillus thuringiensis* MYBT18246 strain**

   Jacqueline Hollensteiner, Anja Poehlein, Cathrin Spröer, Boyke Bunk, Anna E. Sheppard, Philip Rosenstiel, Hinrich Schulenburg and Heiko Liesegang; *Standard in Genomic Science* (2017), 12:48, DOI 10.1186/s40793-017-0259-x

2. **Complete Genome sequence of the nematicidal *Bacillus thuringiensis* MYBT18247**

   Jacqueline Hollensteiner, Anja Poehlein, Cathrin Spröer, Boyke Bunk, Anna E. Sheppard, Philip Rosenstiel, Hinrich Schulenburg, Heiko Liesegang; *Journal of Biotechnology* (2017), 260; 48–52

3. **Genome-Based Identification of Active Prophage Regions by Next Generation Sequencing in *Bacillus licheniformis* DSM13**

   Robert Hertel, David Pintor Rodriguez, **Jacqueline Hollensteiner**, Sascha Dietrich, Andreas Leimbach, Michael Hoppert, Heiko Liesegang and Sonja Volland; *PLoS One* (2015), 10(3):e0120759

4. **Host-pathogen coevolution: the selective advantage of *Bacillus thuringiensis* virulence and its Cry toxin genes**

   Leila Masri, Antoine Branca, Anna E. Sheppard, Andrei Papkou, David Laehnemann, Patrick S. Guenther, Swantje Prahl, Manja Saebelfeld, **Jacqueline Hollensteiner,** Heiko Liesegang, Elzbieta Brzuszkiewicz, Rolf Daniel, Nicolaas K. Michiels, Rebecca D. Schulte, Joachim Kurtz, Philip Rosenstiel, Arndt Telschow, Erich Bornberg-Bauer and Hinrich Schulenburg**,** *PLoS Biology* (2015), 13(6): e1002169

5. ***Bacillus thuringiensis* and *Bacillus weihenstephanensis* inhibit the growth of phytopathogenic *Verticillium* species**

   Jacqueline Hollensteiner, Stefani Maria Díaz Valerio, Anna Kolarzyk, Rebekka Harting, Anja, Poehlein, Kai Nesemann, Elzbieta Brzuskiewicz, Gerhard Braus, Susanna Braus-Stromeyer, Rolf Daniel and Heiko Liesegang; *Front. Microbiol.* (2016), 7:2171. doi:10.3389/fmicb.2016.02171

Poster presentations

1. **Virulence plasmids from nematocidal *Bacillus thuringiensis* genomes: a rich source for compatible new shuttle vectors for *B. subtilis* clade strains.**
   Jacqueline Hollensteiner, Elzbieta Brzuszkiewicz, David Gutiérrez, Robert Hertel, Hinrich Schulenburg, Rolf Daniel, Heiko Liesegang
   „Genomes 2014“International conference in Paris, France (2014)
2. **Competetive Infection: Comparative genomics of three nematocidal *Bacillus thuringiensis* strains.**
   Jacqueline Hollensteiner, Andrea Thürmer, Boyke Bunk, Phillip Rosenstiel, Hinrich Schulenburg, Heiko Liesegang
   ProkaGENOMICS 2015 - 6th European Conference on Prokaryotic and Fungal Genomics in Göttingen, Germany (2015)
3. **The green impact: bacterioplankton response towards a phytoplankton spring bloom in the southern North Sea.**
   Bernd Wemheuer, Franziska Wemheuer, **Jacqueline Hollensteiner**, Frauke-Dorothee Meyer, Sonja Voget, Rolf Daniel
   ProkaGENOMICS 2015 - 6th European Conference on Prokaryotic and Fungal Genomics in Göttingen, Germany (2015)
4. **Beneficial infections: *Bacillus cereus* sensu lato isolates from root microbiome fight phytophathogneic *Verticillium* species**.
   Jacqueline Hollensteiner, Anna Kolarzyk; Rebekka Harting, Franziska Wemheuer, Kai Nesemann, Gerhard Braus, Susanna Braus-Stromeyer, Rolf Daniel, Heiko Liesegang
   VAAM 2016- Annual Conference 2016 of the Association for General and Applied Microbiology, Jena, Germany (2016)
5. **Identification of adaptative genomic determinants in host-parasite co-evolution.** Ludovic Mallet, Barbara Milutinovic, **Jacqueline Hollensteiner**, Heiko Liesegang, Hinrich Schulenburg, Erich Bornberg-bauer, Joachim Kurtz
   DFG SPP 1399 Host-Parasite Coevolution Symposium, Münster, Germany **(**2016)

TABLE OF CONTENTS

CHAPTER I INTRODUCTION

I.1 General Introduction

After World War II, the green revolution started, describing the progress in the field of agricultural research by the development of new technologies which resulted in an enormous increase in crop production and enabled the world food production to sustain currently ~7.44 billion people (Stiftung Weltbevölkerung: „Weltbevölkerungstag 2016: 7,44 Milliarden Menschen leben auf der Erde“). However, insect pests and microbial plant diseases challenged the agricultural food production. As consequence of the indiscriminate usage of agrochemicals and chemical fertilizers, in the 1950s, a natural imbalance of enemies and pests occurred associated with resistance building of pest populations against insecticides. Moreover, their negative environmental effects on non-target beneficial insects and vertebrates were noticed (Sansinenea, 2012). Attention was brought to biological control agents including insect pathogens such as *Bacillus thuringiensis*. The story of success of the entomopathogenic bacterium *B. thuringiensis* began because of its broad host spectrum including insects, nematodes, mites, ticks and some protozoa (Schnepf et al., 1998). Furthermore, the abilities of an easy formulation, suitable shelf life, stability, and that the species is a classified by the Food and Drug Administration (FDA) as *general recognized as safe* (GRAS) organism lead to the acceptance as biological control agent (Bravo et al., 2011, 2013; Bulla, 1975; Drobniewski, 1993; Mendelsohn et al., 2003; Walter et al., 2010). Today, *B. thuringiensis* is used in agriculture, forestry, and pest control worldwide because of their specific toxicity against target insects which is mainly based on their production of insecticidal crystal proteins (ICPs) (Sansinenea, 2012). Recently, the industrial production has extended the production of transgenic crops and other biocontrol formulations (Plumer, 2016).

I.1.1 History of *Bacillus thuringiensis* – an environmental friendly alternative

B. thuringiensis was and still is one of the most famous organisms with huge impact on the development of biocontrol research in the last century (Figure 1). In the year 1901, *B. thuringiensis* was first isolated from dead silkworm-larvae by Shigetane Ishiwata (Shigetane, 1901). The bacterium was named *Bacillus sotto* because it caused death to a huge number of silkworms in Japan (Shigetane, 1901). Later, in 1915 the German scientist Ernst Berliner isolated a similar bacterium from dead Mediterranean flour moth larvae in the German state of Thuringia and did the first valid morphological description of *B. thuringiensis* (Ernst, 1915). The application of *B. thuringiensis* as biopesticide started early, in the late 1920s in Hungary and at the beginning of the 1930s in Yugoslavia (Husz, 1928; Vouk and Klas, 1931). In 1928 the first commercial product "Sporein" based on a spore formulation was available (Aronson et al., 1986). Due to the "Green Revolution" a great development took place in agriculture. New methods and substances were available to increase the crop productivity such as chemical fertilizers but also by mechanization which led to a decrease of used biocontrol agents including *B. thuringiensis*. *Bacillus thuringiensis* started his comeback in the field of biocontrol in 1951 (Figure 1) (Sansinenea, 2012). Researchers did great progress in the following years. In 1954 parasporal crystals have been identified as causative agent of the host specific toxicity of *B. thuringiensis* (Angus, 1954; Hannay, 1953), new subspecies were isolated and new formulations and commercial products were created (Aizawa and Iida, 1963; Dulmage, 1970; Krieg et al., 2009; Kurstak, 1970). Zakharyan *et al.* found in 1976, that the mechanism of toxicity of *B. thuringiensis* and spore formation is based on genomic level and localized on a plasmid (Sansinenea, 2012). Notably, in 1990 *B. thuringiensis* was used active as pest control in targeting a nematode which causes the human "river eye disease" in West Africa (Guillet et al., 1990). In 1992 Beegle and Yamato *et al*. (Beegle et al., 1992) found out, that the world market for only *B. thuringiensis kurstaki*-based products was estimated at $ 20-25 million in the United States. The biggest milestone was the creation of the first transgenic plant expressing *B. thuringiensis* toxins in 1993 (van Frankenhyzen, 1993).

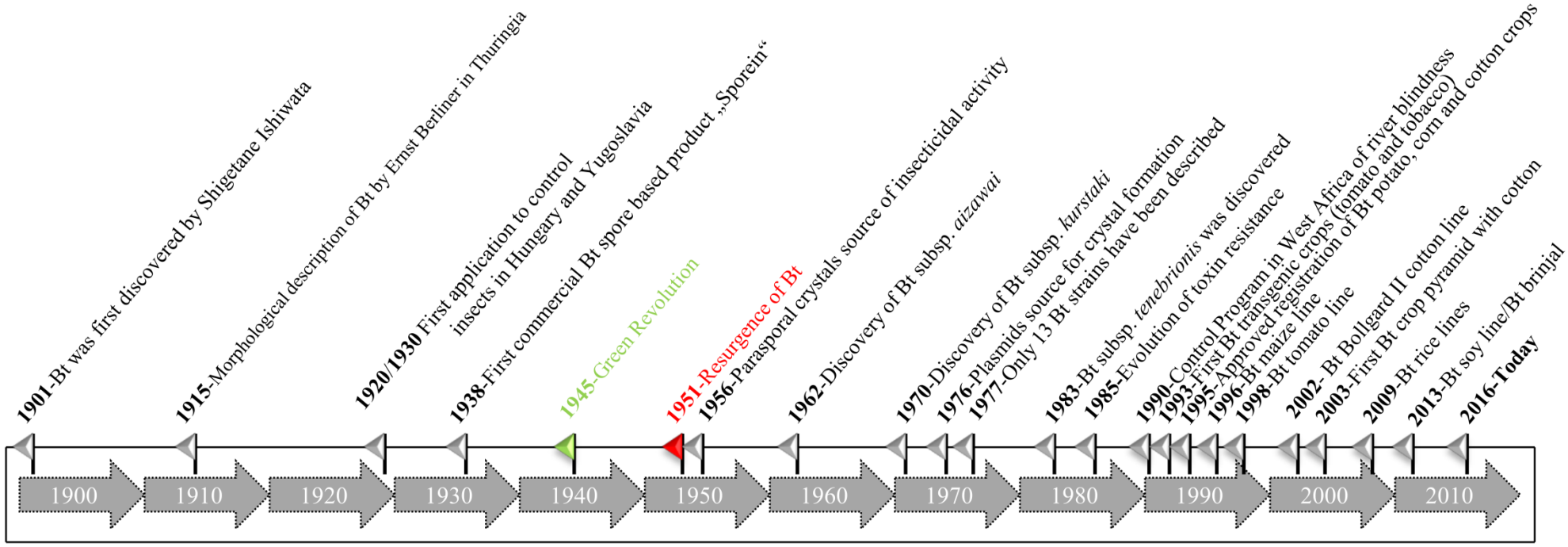

Figure 1. Timeline displaying the main milestones for *Bacillus thuringiensis* research (grey flags). Milestones are colored.

Since then, new transgenic plant lines for major crop plants have been introduced including *B. thuringiensis* field corn, potato, cotton, tomato, rice, soy, brinjal to circumvent the building of resistances (Carrière et al., 2015; Choudhary et al., 2014; James and Krattiger, 1996; Li et al., 2014; Sanahuja et al., 2011).

This made the agricultural industry highly profitable. But since the competition for being the market leader is high, expensive, and always new innovations are needed for the upcoming resistance building they get more and more consolidated (McGaughey, 1984; Plumer, 2016).

Due to the latest fusion of Bayer and Monsanto in 2016 there are only three big companies left: Bayer-Monsanto, ChemChina-Syngenta and Dow-Dupont (Plumer, 2016). However, there is a great amount of scientific work ongoing in *B. thuringiensis* research, involving the discovery and evolution of new Cry- (Crystal protein), Cyt- (Cytolytic protein), Vip- (Vegetative insecticidal protein) and Sip- (Secreted insecticidal protein) –toxins (Ye et al., 2012). The mode of action of such toxins and the mechanisms of resistance building in insects are all fundamental and promising studies for improving and broaden the application of *B. thuringiensis* as biocontrol agent (Carrière et al., 2015; Jurat-Fuentes and Crickmore, 2016).

I.1.2 *Bacillus thuringiensis* biology, ecology and lifestyle

B. thuringiensis (Bt) is a gram-positive, facultative anaerobic, endospore-forming and wide spread bacterium (Aronson et al., 1986; Bravo et al., 1998; Schnepf et al., 1998). The defining feature of *B. thuringiensis* is the formation of insecticidal crystal proteinaceous (ICP's) inclusion bodies during sporulation and stationary phase of the growth cycle (Schnepf et al., 1998). Those insecticidal toxins are encoded by *cry* toxin genes which are primarily encoded on plasmids but could also reside in the chromosome (Kronstad et al., 1983; Schnepf et al., 1998; Wang et al., 2013). The crystal toxins (Cry) are active against a broad range of insects such as Lepidoptera, Diptera, and Coleoptera but also against other orders like Hymenoptera, Homoptera, Orthoptera, Mallophaga and nematodes, mites, ticks and protozoa (Aronson et al., 1986; Fernández-Ruvalcaba et al., 2010; Ibrahim et al., 2010; Manasherob et al., 1998; Schnepf et al., 1998). In 2016, 37 closed *B. thuringiensis* genomes are publicly available, which vary in their genome size (5.3-6.8 Mbp) as well as in the number of plasmids (1-15 replicons). The GC content ranges from 34 to 35% and classifies thus

the species as low GC-content organism (https://www.ncbi.nlm.nih.gov/genome/genomes/486?_2016_20_09). The genome plasticity of *B. thuringiensis* is flexible. Genomic flexibility is a measurable value of bacteria describing the ability to adapt or evolve under selective pressures. The degree of genomic flexibility of bacteria is complex and includes mobile elements and site-specific recombination systems (Argolo-Filho and Loguercio, 2014; Bennett, 2004). Other mobile elements, including bacteriophages, integrated conjugative elements (ICE), transposases, insertion sequence (IS) elements, and transposons, are thus supposed to be the main driving forces for either adaptation and evolution (Brüssow et al., 2004; Frost et al., 2005). Moreover, the degree of the genetic flux differs and is dynamic between replicons, such as chromosomes and plasmids. Plasmids not only have a higher degree of self DNA re-assortment events compared to chromosomes. They are important mobile elements, taking part as horizontal gene transfer agents conferring new DNA to other bacteria (Bennett, 2004). Besides, mutation rates are non-neutral and genes that are highly expressed have a higher mutation rates resulting in a non-random genomic positional distribution (Lind and Andersson, 2008; Sharp and Li, 1987).

As it could be seen so far, *B. thuringiensis* is distributed worldwide and can be isolated from nearly everywhere: soil, aquatic habitats, phylloplane, dust, insects, and feces of arid birds (Burges and Hurst, 1977; Donovan et al., 1988; Iriarte et al., 2000; Martin et al., 1989; Poopathi et al., 2014; Schnepf et al., 1998; Smith and Couche, 1991) (Figure 2).

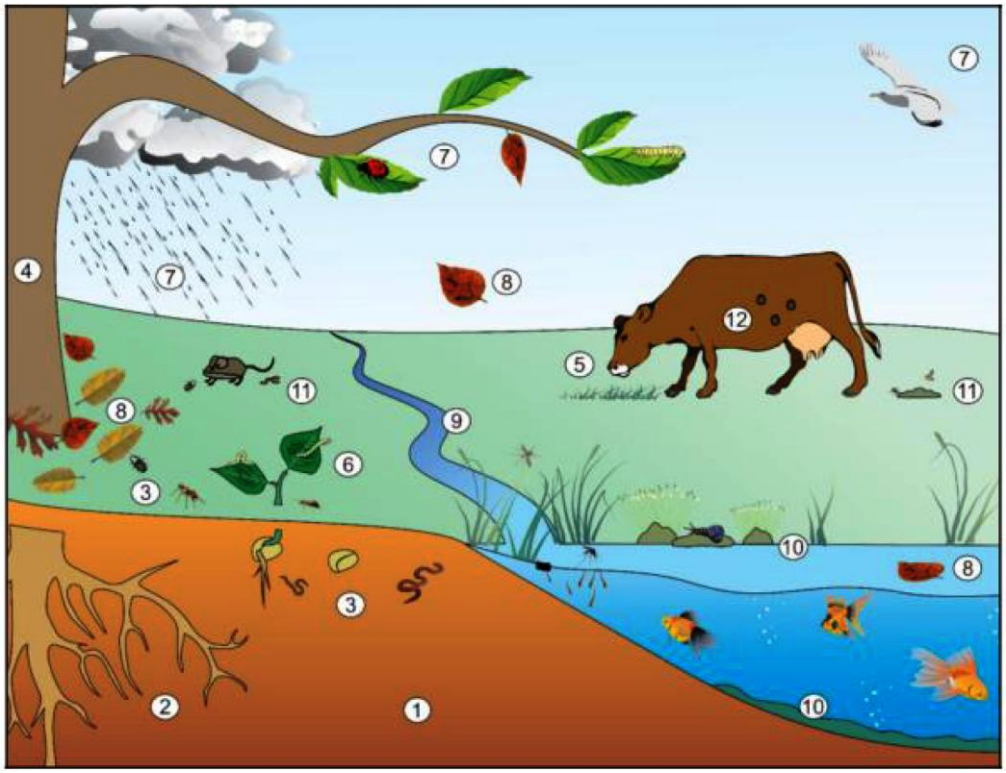

Figure 2. Ecological niches and potential dissemination pathways of *Bacillus thuringiensis* are indicated by numbers. (1) soil, (2) rhizosphere, (3) insects, nematodes and worms, (4) endophytic as rhizosphere colonizer, (5) paratenic, (6) pathogenic in susceptible hosts, (7) feces and rain, (8) composted leaves, (9) rain, (10) in water living invertebrates or vertebrates and in aquatic plants and sediments ,(11) coprophagy, (12) mites and ticks. The figure was taken from Filho *et al.* (Argolo-Filho and Loguercio, 2014).

The availability of *B. thuringiensis* depends basically on the production of spores which are passively dispersed into a variety of different ecological niches (Argolo-Filho and Loguercio, 2014). The lifecycle of *B. thuringiensis* (Figure 3A) can be subdivided into the vegetative phase, where *B. thuringiensis* proliferates by medial division (Errington, 2003) and into the sporulation phase, where *B. thuringiensis* produces spores and toxins to be prepared for unfavorable conditions (Jensen et al., 2003). If spores and the toxic crystals are dispersed and ingested by a potential host, *B. thuringiensis* has the opportunity to pass two different lifecycles, namely the paratentic and the infectious lifecycle (Dubois et al., 2012; Jensen et al., 2003) (Figure 3B).

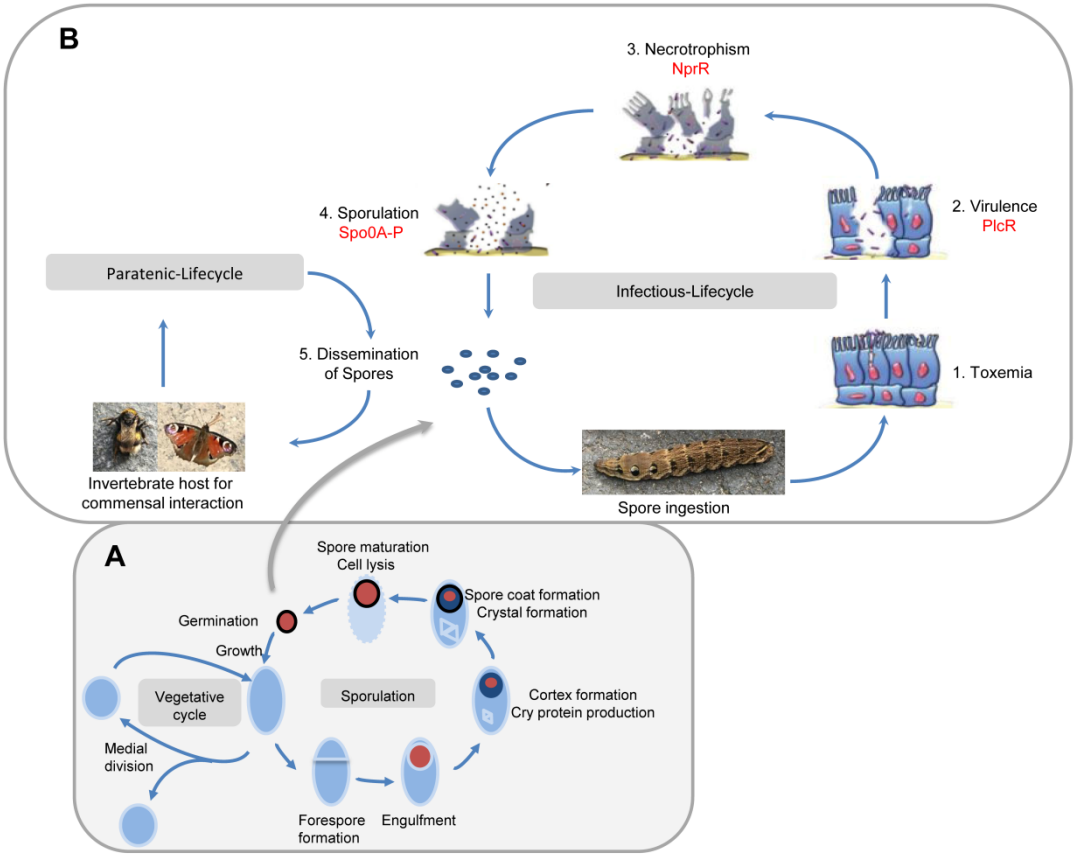

Figure 3. Different lifecycles of *B. thuringiensis*. (A) General lifecycle, (B) Two sub-lifecycles: (Right) Infectious-lifecycle, (Left) Paratenic-lifecycle, (1-5) Different lifestyles at different lifecycle stages are displayed: (Red) expression of master regulators at different stages. The components of the figure were modified from Dubois *et al.* (Dubois et al., 2012), Jensen *et al.* (Jensen et al., 2003), Bechtel *et al.* (Bechtel et al., 1976) and Errington (Errington, 2003)**.**

If the potential host is not susceptible against Cry toxins, *B. thuringiensis* infect in a paratenic way and is re-introduced into the environment through feces (Argolo-Filho and Loguercio, 2014). In a susceptible host, Cry toxins induce cell lysis and *B. thuringiensis* spores germinate (Dubois et al., 2012). Different lifestyle switches, controlled by cell-cell communication systems including specific transcriptional regulators, are passed through such as toxemia, virulence, necrotrophism and sporulation. During virulence, the pleiotropic regulator PlcR activates the production of various virulence factors to weaken the intestinal barrier till it is causing death to the host (Dubois et al., 2012; Perchat et al., 2016). *Bacillus thuringiensis* then switches to the nectrotrophic lifestyle which is regulated by the pleiotropic regulator NprR (Dubois et al., 2012). Genes involved in food supply, general bacterial fitness factors, and genes involved in survival of the bacteria are activated (Perchat et al., 2016; Raymond et al., 2010). The last lifestyle shift is towards sporulation, regulated by Spo0A-P (Dubois et

al., 2012). Sporulation is prevented by NprR only, based on Spo0F dephosphorylation. In contrast, bound NprR-NprX prevent dephosphorylation of Spo0A resulting in sporulation initiation (Perchat et al., 2016). These various complex quorum sensing signal transduction systems such as PlcR-PapR, NprR-NprX, and Rap-Phr are present within genomes of *B. thuringiensis* and enables it to adapt to different ecological niches and switch between a variety of different lifestyles such as free-living, endophytic, symbiotic, pathogenic, paratenic or saprophytic (Dubois et al., 2012; Perchat et al., 2016; Raymond et al., 2010; Rose Gomes, 2012; Swiecicka et al., 2002; Wilcks et al., 2008) (Schmidt et al., 2011; Verplaetse et al., 2015, 2016). Therefore, the question of the primary ecological niche is hard to answer. Within the scientific community, two major concepts are discussed: *B. thuringiensis* as insect pathogen and *B. thuringiensis* as saprophytic bacterium. Argolo-Filho *et al*. (Argolo-Filho and Loguercio, 2014) introduced the terminus of *B. thuringiensis* as "environmental pathogen" due to its ability to spend a part of its lifecycle out of its host and by getting in contact with a potential host it is causing disease without the need of specific conditions. This is also in consensus with previous definitions that *B. thuringiensis* is an entomopathogenic bacterium with additional activity against mites, ticks and protozoa (Schnepf et al., 1998).

I.1.3 Virulence Factors of *Bacillus thuringiensis*

As environmental entomopathogenic bacterium, *B. thuringiensis* needs several virulence factors to invade, persist and infect a potential host. The host defense mechanism and barriers such as the cuticle, the intestinal microbiota, the gut peristalsis as well as chemical defenses as the pH, proteases, antimicrobial compounds and the immune system must be overcome (Argolo-Filho and Loguercio, 2014; Raymond et al., 2010; Vallet-Gely et al., 2008). *Bacillus thuringiensis* encodes an armory of different virulence factors to guarantee the successful infection including a variety of proteinacous toxins, α-exotoxins, endochitinases, hemolysins, metalloproteases, β-exotoxins, bacteriocines and antibiotics including Zwittermicin A. (Celandroni et al., 2014; Fedhila et al., 2002; Liu et al., 2014; Nair et al., 2004; Salazar-Marroquín et al., 2016; Sampson and Gooday, 1998; Xiaoxia et al., 2012).

The best investigated virulence factors are the δ–endotoxins which are subdivided into two families, namely Cry and Cyt. Additionally, two other toxin groups, Vip (vegetative

insecticidal proteins) and Sip (secreted insecticidal proteins), have similar properties as Cry and Cyt. A common feature of *B. thuringiensis* is the broad host range based on the high number of various toxins where each toxin is highly specific against a specific host (Crickmore et al., 2016) (Figure 4).

In 2016, 965 toxins (including Cry, Cyt, Vip and Sips) are listed showing toxicity against Lepidoptera, Diptera, Coleoptera, Gastropoda, Hymenoptera, Hemiptera, Rhabditida, and human cancer cell lines (Ali et al., 2010; De Maagd et al., 2001; Ohba et al., 2009; Palma et al., 2014; Wei et al., 2003; Whiteley and Höfte, 1989). Activity against mites, protozoa, and ticks were also documented (Erban et al., 2009; Fernández-Ruvalcaba et al., 2010) (http://www.btnomenclature.info/ accessed 12.10.2016) (Figure 4).

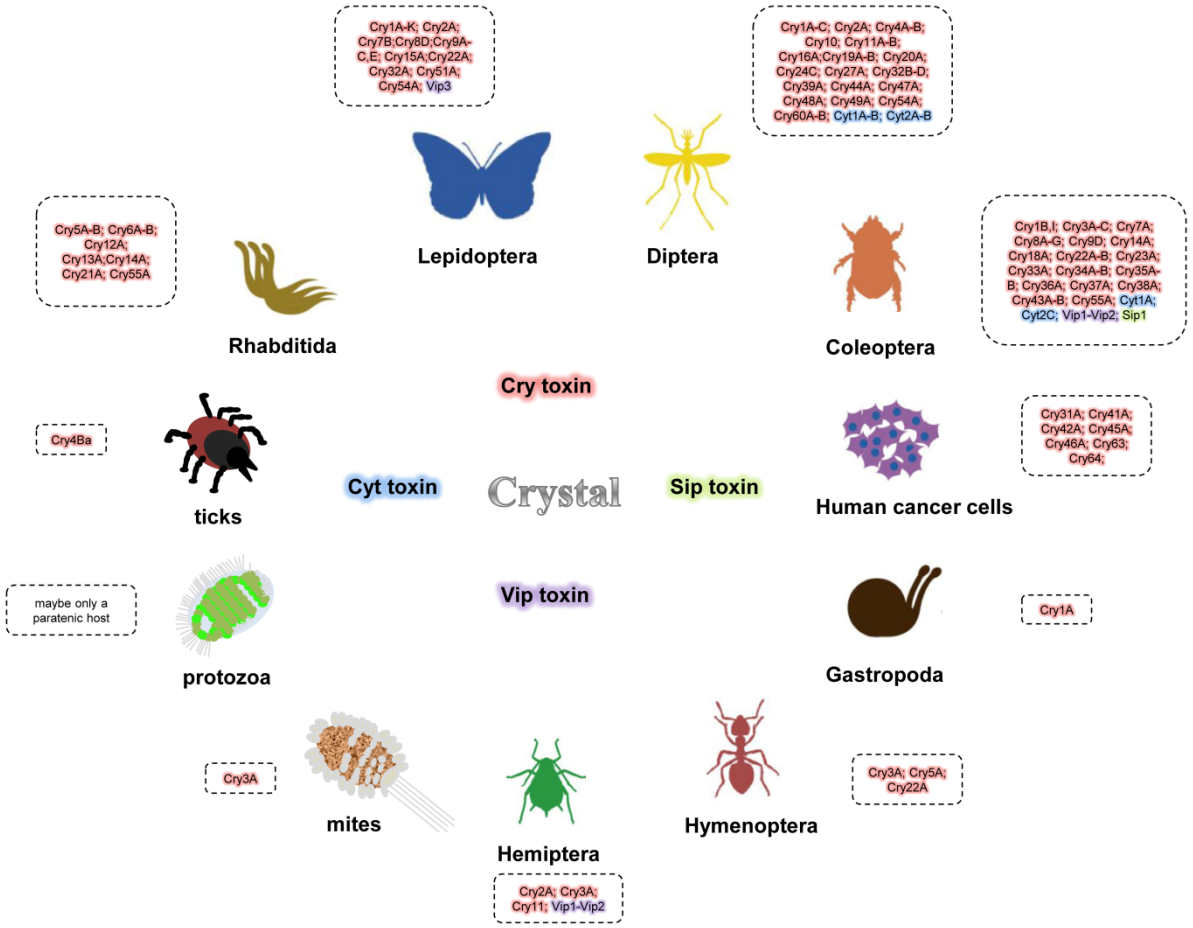

Figure 4. Overview of hosts susceptible to toxins of *Bacillus thuringiensis*. The four different protein toxin classes, such as crystal (Cry), cytolytic (Cyt), vegetative insecticidal protein (Vip) and secreted insecticidal proteins (Sip), and their host range are depicted by the colors red, blue, purple, and green, respectively. Cry toxins, where no host specificity was determined, were excluded. Additionally, no host was reported for Vip4 or Cyt3. The components of the figure were modified and updated from Palma *et al.* (Palma et al., 2014).

Due to the rapid development in next-generation sequencing (NGS) technologies and increased availability of new bioinformatic genome analysis tools for the detection of toxins, Crickmore *et al.* (Crickmore et al., 2016) established a nomenclature for all Bt-toxins to avoid redundant investigations and misleading namings (Figure 5).

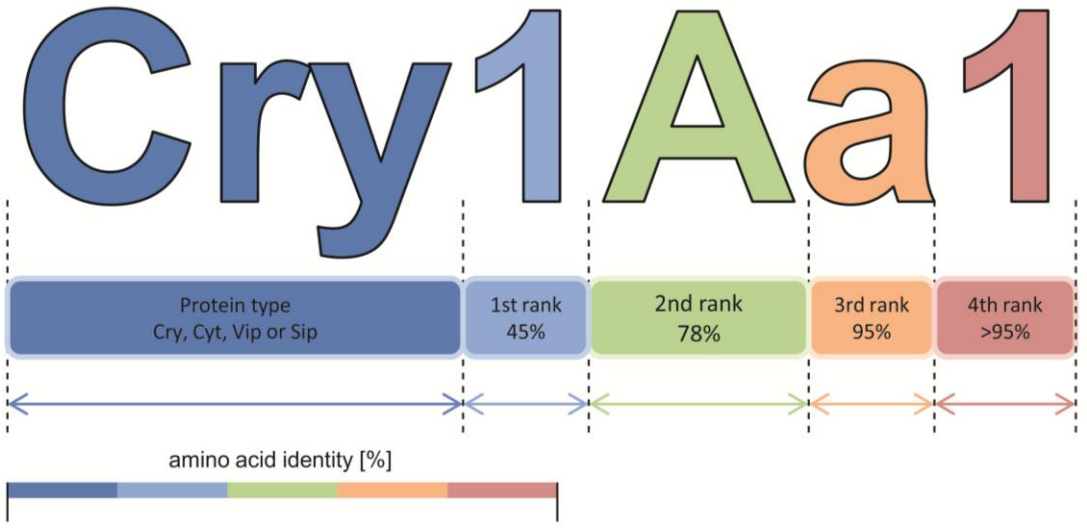

Figure 5. Nomenclature for crystal (Cry) proteins, cytolytic (Cyt) proteins, vegetative insecticidal proteins (Vip) and secreted insecticidal proteins (Sip) based on amino acid similarity (Crickmore et al., 2016). The components of the figure were modified from Chakroun *et al.* (Chakroun et al., 2016) and Palma *et al*. (Palma et al., 2014).

The toxin classification scheme assigns four ranks based on the degree of amino acid sequence identity compared to known accepted toxins including Cry, Cyt, Vip and Sips. A protein with a sequence identity lower than 45% is assigned as Cry-like toxin. Proteins with more than 45, 78 and 95% sequence identity to known toxins are grouped to rank 1, rank 2, or rank 3, respectively. Proteins with over 95% identity belong to rank 4, and the protein has no or only minor amino acid differences. For example, a potential protein candidate shares 88% sequence similarity to Cry1Aa1 is classified based on the sequence identity as Rank 2 toxin, and named as "Cry1Oa1", respectively. The "O" as second Rank must be used, because other Cry names, such as Cry1B*xx*-N*xx* have been already assigned in the Cry-toxin database (Crickmore et al., 2016).

I.1.3.1 Cry toxins

The Cry proteins are pore-forming toxins and can be subdivided into three different homology groups: (i) three-domain toxin group, (ii) Mtx group of toxins, and the (iii) Bin group toxins (Krishnan, 2013). Additionally, there are a few Cry toxins known which have no sequence similarity to any of these groups such as Cry6, Cry22, Cry37, Cry46 and Cry55. Furthermore, a subgroup of split-toxins was documented. The basic features of each toxin group are summarized in Figure 6.

	Cry				Cyt	Vip				Sip
Number of toxin sequences	787				38	139				1
Subgroup of toxins	3-Domain Toxins	Mtx	Bin	Unique Toxins	Cyt	Vip1	Vip2	Vip3	Vip4	Sip1
Number of toxin sequences	725	17	27	18	38	15	20	103	1	1
Group sepecific feature	Conserved three-domain structure	Show features of the ETX-MTX2 family	Binary toxins, need a partner	Unique compared to the other toxins	Cytolytic	Vegetative secreted	Vegetative secreted	Vegetative secreted, atypical signal sequences	Vegetative secreted	Vegetative secreted
Conserved domains	N-terminal domain, Middle-domain, C-terminal domain	ETX/MTX2 domain	Ricin B-domain Toxin_10 domain	Variable	Delta-endotoxin CytB family	Receptor-binding domain	Cytotoxic A domain	Vip3A_N domain	anthrax protective antigen PA14 domain, Binary_ToxB exotoxin domain	ETX/MTX2 domain
Mode of action	Pore formation	Pore formation?	Pore formation	Unknown	Pore forming or acting as detergent	Vip1/2 binary toxins, pore formation		Pore formation?	Unknown, maybe part of a binary toxin	Pore formation?

3-Domain Toxins	Mtx Toxins	Bin Toxins	Unique Toxins
Cry1-Cry5	Cry15	Cry35-Cry34*	Cry6
Cry7- Cry14	Cry23*	Cry36	Cry22
Cry16- Cry21	Cry33	Cry49-Cry48*	Cry37*
Cry24-Cry32	Cry38		Cry46
Cry39-Cry44	Cry45		Cry55
Cry47-Cry48	Cry51		Cry34
Cry50	Cry60		
Cry52-Cry54	Cry64		
Cry56-Cry59			
Cry61-Cry63			
Cry65-Cry72			

Figure 6. List of *Bacillus thuringiensis* toxins. Sub-homology groups, number of deposited sequences, specific group features, conserved domains and the mode of action are listed. The Cry toxins were in detail listed regarding to their homology group. Unique toxins are not related to other Bt toxins.

Three-domain toxins

In total, 787 Crystal toxin sequences including 725 three-domain Cry toxins which are active against various insects, nematodes and ticks have been assigned to different host organisms (Figure 4 and Figure 6) (http://www.btnomenclature.info/ accessed 12.10.2016). Three-domain toxins can vary remarkably in their amino acid composition but all share a highly conserved domain structure (Bravo et al., 2008; De Maagd et al., 2001, 2003; Pardo-López et al., 2013). However, amino acid sequence alignments of three-domain toxins revealed five conserved blocks in matured Cry toxins and three additional blocks in the protoxin (Höfte and Whiteley, 1989; Schnepf et al., 1998). Xu *et al*. (Xu et al., 2014) summarized that each domain is important for toxicity and is responsible for a distinct function. Domain I (N-domain) is responsible for the insertion of the toxin into the apical membrane, the pore formation and the maintenance of receptor binding (Ben-Dov, 2014; Schnepf et al., 1998). Domain II (middle-domain) is the most variable domain and suspected to be important for toxin-receptor specificity. Domain III (C-domain) is involved in the maintenance of structural integrity of the toxin, binding assistance, penetration and formation of pores (Schnepf et al., 1998). The general toxic mechanism comprises ten steps: (1) ingestion of spores and crystals (2) solubilization of crystals by gut pH, (3) proteolytic processing of the protoxin by gut proteases, (4) toxin binding to receptors, (5) toxin insertion into the membrane, (6) formation of lytic pores, (7) paralysis of the host, (8) intestinal rupture, (9) septicemia and (10) host death (Höfte and Whiteley, 1989; Palma et al., 2014; Schnepf et al., 1998; Vachon et al., 2012). Today two modes of action are hypothized the "Sequential Binding Model" where the Cry protoxin is proteolytically cleaved to bind to receptors, followed by oligomerization to form a pore in the cell membrane and the "Signaling Pathway Model" where the necrotic death is induced by a signaling pathway activated by Cry toxin binding instead of lytic pore formation (Vachon et al., 2012).

Mtx toxins

The second group of Cry toxins is the ETX_MTX2 family and comprises recently 17 sequences (http://www.btnomenclature.info/ accessed 12.10.2016, Figure 6). The defining feature of this toxins is the conserved ETX domain belonging to the ETX/MTX2 family (Pfam 03318) (Palma et al., 2014). Mtx2 and Mtx3 are also known from mosquitocidal *Lysinibacillus sphaericus* and show similarities to other toxins such

as the aerolysins from *Aeromonas hydrophila*, the alpha toxins of *Clostidium septicum* and cytotoxins of *Pseudomonas aeruginosa* (Knapp et al., 2010). Member of this Mtx-toxin group are Cry15, Cry23, Cry 33, Cry38, Cry45, Cry51, Cry60, Cry60 and Cry64. Most of them are not fully investigated but their host specificity ranges from Lepidoptera (Cry15, Cry51) to Coleoptera (Cry23, Cry33, Cry38) to Diptera (Cry60) and human cancer cell lines (Cry45, Cry64) (Palma et al., 2014) (Figure 4). Because of the related structure of Mtx toxins to aerolysins which are pore-forming toxins, the above mentioned toxins are supposed to act in a similar mode of action (Palma et al., 2014).

Bin-like toxins

The last subgroup of Cry toxins are the Bin-like toxins, named as such because they show sequence homology to mosquitocidal binary toxin components (toxic BinA/binding moiety BinB) from *Lysinibacillus sphaericus* (Charles et al., 1996). Today the group comprises three basic members (Cry35, Cry36 and Cry49) and 27 different toxin sequences are deposited (http://www.btnomenclature.info/ accessed 12.10.2016, Figure 6). Notably, Cry36 seems to be toxic on its own. Following pairs are known to act as binary toxins: Cry23/Cry37; Cry34/Cry35, Cry48/Cry49 (Jones et al., 2007; Palma et al., 2014; Schnepf et al., 2005).

Split-toxins

Split-toxins are the last group of special toxins which are poorly understood. The toxic gene product is split on two Open Reading Frames (ORF) separated by an intergenic region where ORF1 encode the N-terminus and ORF2 the C-terminus (Krishnan, 2013). This gene splitting was observed for some *cry* gene pairs (Lenane et al., 2008; Ohgushi et al., 2005). Why and how the splitting of cry toxin genes have evolved is unknown.

I.1.3.2 Cyt toxins

Cytolytic (Cyt) toxins have cytolytic and hemolytic activity, are encoded by *cyt* genes and are divided into three families: Cyt1-Cyt3 (Bravo et al., 2008; Butko, 2003). Today 38 different sequences are available comprising 13 Cyt1, 24 Cyt2 and one Cyt3 (http://www.btnomenclature.info/ accessed 12.10.2016, Figure 6). The toxicity of Cyt toxins is highly specific to target hosts and has been reported for Diptera, Lepidoptera,

Coleoptera, nematodes and cancer cells (Guerchicoff et al., 2001; Guillem and Porcar, 2012; Soberón et al., 2013). Cyt toxins are hydrophobic and exhibit in contrast to Cry toxins a strong hemolytic activity *in vitro* and especially against Dipteran *in vivo* (Butko, 2003; Schnepf et al., 1998). The mode of action is not fully understood but Butko (Butko, 2003) presented two ways of action depending on the toxin concentration. He could show that at low Cyt-toxin concentrations oligomeric pores are formed in the cell membranes of the insect gut whereas at high toxin concentration Cyt have detergent-like function by rupturing the cells. There is also evidence of synergism between several Cyt and Cry or Cyt and Vip3 toxin proteins. Sayyed *et al.* (Sayyed et al., 2001) presented the synergism of Cy1Aa and Cry11Aa toxins, which is also promising in the reduction of resistance building (Pé Rez et al., 2005). Nevertheless, antagonistic effects were observed as well (Del Rincón-Castro et al., 1999).

I.1.3.3 Vip and Sip toxins

In contrast to Cry toxins, all Vip (vegetative insecticidal proteins) and Sip (secreted insecticidal proteins) are mainly produced during the vegetative growth phase and secreted into the medium (Donovan et al., 2006; Estruch et al., 1996). However, the secretion can be extended into the sporulation phase (Estruch et al., 1996). Vip are promising toxins for the agricultural usage as they show no sequence homology to δ-endotoxins so that cross-resistance in insects are unlikely (Rice, 1999). Vip proteins are grouped into four different families, Vip1-Vip4. Today 139 vegetative insecticidal proteins are reported (15 Vip1 proteins, 20 Vip2 proteins, 103 Vip3 proteins and only one Vip4 as well as one Sip1 protein) (http://www.btnomenclature.info/ accessed 12.10.2016, Figure 6). Those toxin families share nearly no sequence homology with each other and differ in their protein structures (Chakroun et al., 2016). Vip1 and Vip2 belong to the binary toxins, carried on and expressed from the same operon, are related to mammalian A+B toxins from *Bacillus anthracis* or *Clostridium perfringens* (Chakroun et al., 2016) where Vip1 represent the B component and Vip2 the A component (Barth et al., 2004). They are active against some members of Coleoptera and Hemiptera (Sattar and Maiti, 2011). In general, Vip1 toxins show a highly conserved N-terminus which is important for pore formation, and a weak conserved C-terminus. In contrast, Vip2 penetrates the cell and causes the final toxic effect (Jucovic et al., 2008). Additionally, Vip2 contains two domains, an ADP ribosyltransferase

domain and a NAD-binding domain and both termini of Vip2 are conserved (Han et al., 1999). The mode of action of Vip1/Vip2 toxins is similar to Cry-toxins, the protoxin Vip1 is cleaved by a midgut protease and thereby activated, it binds to receptors and is inserted into the membrane (Leuber et al., 2006). Afterwards, Vip2 binds to oligomeric Vip1 and enters the host cell through pores formed by Vip1 or by endocytosis (Barth et al., 2004; Leuber et al., 2006). In the cell Vip2 prevents polymerization of monomeric actin by transferring the ADP-ribose group to NAD (Han et al., 1999). Vip3 toxins are common in *B. thuringiensis* strains and are active against Lepidoptera (Estruch et al., 1996). The protein structure is similar to Vip1, the N-terminus is highly conserved comprising a signal peptide sequence, whereas the C-terminus is variable (Rang et al., 2005; Wu et al., 2007). However, the signal peptide sequences which are important for protein translocation into the membrane are atypical because they are not processed during secretion (Li et al., 2007; De Maagd et al., 2003).The exact mode of action of Vip3 and their secretion mechanisms are unknown but seem to be similar to the pore formation mechanism of Cry toxins (Chakroun et al., 2016; Yu et al., 1997). Synergistic and antagonistic effects for Vip3 toxins in combination with other toxins, such as Cry1 were observed (Lemes et al., 2014). Notably for Vip4, no target hosts have been found till now (Chakroun et al., 2016) and remains cryptic due to the lack of available information. An anthrax protective antigen PA14 domain and the Binary_ToxB exotoxin domain were identified in Vip4 which is known from binary Vip1 and other toxins (Palma et al., 2014).

Sip is also a secreted toxin, active against Coleopteran larvae but their amino acid sequence homology is too low to be grouped to the other Vip toxins (Donovan et al., 2006). The toxin shows homology to the Mtx3 mosquitocidal toxin of *Lysinibacillus sphaericus* (Palma et al., 2014). The mode of action is unknown but pore formation was strongly suggested by Donovan *et al*. (Donovan et al., 2006).

I.1.3.4 Additional Virulence factors

Besides the well known Cry, Cyt, Vip and Sip toxins, *B. thuringiensis* comprises a plethora of other virulence factors which efficiently act as supporting factors. For example other groups of toxins are described for *B. thuringiensis*, such as hemolysins including hemolytic enterotoxin (Hbl), cereolysin (CerAB) and cytotoxin (CytK), and the non-hemolytic enterotoxin (Nhe), which all are important during the infection and

controlled by the quorum sensing transcriptional regulator PlcR indicating that virulence is a coordinated activity of an infecting population (Granum et al., 1999). Moreover, they form pores in insect haemocytes and thus increase the acquisition of the nutrients from host cells (Argolo-Filho and Loguercio, 2014; Kim et al., 2015). Thuringiensin is a non proteinaceous β-exotoxin and is highly toxic to diverse taxa. It is thermostable and inhibits the eukaryotic DNA-dependant RNA polymerase, which leads to an altered development (Beebee and Bond, 1973). Due to the high toxicity, *B. thuringiensis* strain producing thuringiensin are forbidden in biological control (World Health Organization-WHO, 1999). Besides toxins, *B. thuringiensis* contains several enzymes acting as virulence factors such as α–exotoxins like phospholipase C, thermolabile proteins which hydrolyze phospholipids of host cell membranes and degenerate haemocytes (Argolo-Filho and Loguercio, 2014). Endochitinases degrade chitin which is part of the midgut peritrophic membrane of many insects as well as in the exoskeleton of arthropods. Sampson *et al.* could show that both exochitinases and endochitinases have a supporting effect on Cry/Vip toxins (Sampson and Gooday, 1998). Furthermore, endogenous metalloproteases such as InhA and Bmp1 revealed toxicity and synergistic activity in combination with Cry toxins against Lepidopera and nematodes, respectively (Fedhila et al., 2002; Luo et al., 2013).

One additional host defense mechanism is the specific host microbiome, which helps to protect them against invaders. To overcome this specific barrier *B. thuringiensis* produces secondary metabolites such as bacteriocins and microcins which all are excellent substances to control gram-positive or gram-negative host bacteria by affecting their growth and their viability (de la Fuente-Salcido et al., 2013).

I.2 The *Bacillus cereus sensu lato* group - Phylogeny of *Bacillus thuringiensis*

Bacillus thuringiensis is a member of the *Bacillus cereus sensu lato* (Bcsl) species group (Priest et al., 1994). This group comprises seven different species defined as either pathogens, opportunists or as well as environmental pathogens (Argolo-Filho and Loguercio, 2014; Priest et al., 2004). The characteristic of these species group is the close genetic relation, which is causing problems in phylogenetic- and taxonomic-classification's of species group members. Each species member contains

one conserved chromosome but their extra-chromosomal equipment such as plasmids and phages is highly diverse. Notably, the virulence factors of pathogenic species are mostly located on transmissible plasmids and crucial for discrimination. The most famous member is the human pathogen *B. anthracis*, which is the causative agent of anthrax (Mock and Fouet, 2001). The virulence is based on the protective antigen (PA), the lethal factor (LF) and the endemia factor (EF). The three factors are encoded on the virulence plasmid pXO1 (Bhatnagar and Batra, 2001). Additionally, the poly-c-D-glutamic acid capsule is important to withstand phagocytosis and is encoded on the plasmid pXO2. *Bacillus cereus* is an opportunistic human and mammal pathogen and well known for food poisoning. Two different types of syndromes are observed, the diarrheal syndrome, characterized by abdominal pain and diarrhea, and the emetic syndrome, leading to nausea and vomiting (Agata et al., 1996; Jensen et al., 2003). Similar to *B. anthracis*, the virulence factors of emetic *B. cereus* strains, such as cereulide, are encoded on a big virulence pXO1-like plasmid (Ehling-Schulz et al., 2006). The main feature of *B. thuringiensis* is the production of crystalline proteinaceous inclusions during the stationary and sporulation phase. There are four less understood and investigated species, such as *B. weihenstephanensis, B. mycoides, B. pseudomycoides* and *B. cytotoxicus.* Those species were primarily classified based on biological and ecological properties. *B. weihenstephanensis* is mainly distinguished on the basis of their psychrotolerance but may also have endophytic or enteropathogenic potential (Lechner et al., 1998; Thorsen et al., 2006). The saprophytic organism *B. mycoides* produces rhizoid colonies (Nakamura and Jackson, 1995). *B. pseudomycoides* was considered as separate group from *B. mycoides* in 1998 by Nakamura based on fatty acid composition (Nakamura, 1998). On agar they form cream, opaque and usually rhizoid colonies and can't be distinguished from *B. mycoides*. The last member is *B. cytotoxicus,* which is a thermotolerant organism and occasionally associated with food poisoning (Guinebretière et al., 2013). With a genome size of 4 Mb it is the smallest of all Bcsl species group members (https://www.ncbi.nlm.nih.gov/genome/?term=Bacillus%20cytotoxicus.)

I.2.1 The molecular taxonomy of the *Bacillus cereus sensu lato* group

Due to the high level of sequence homology among the Bcsl group members several approaches have been used for classification, such as 16 rRNA gene sequence (Chen

and Tsen, 2002), amplified fragment length polymorphism (AFLP) (Hill et al., 2004), restriction fragment length polymorphism (RFLP) in small subunit rRNA sequences (Joung and Côté, 2001), multi-locus enzyme electrophoresis (MEE) (Carlson et al., 1994), multi-locus sequence analysis (MLSA) (Priest et al., 2004), whole genome analysis, estimation of average nucleotide identities (ANI) (Arahal, 2014), Genome-Blast Distance Phylogeny method (GBDP) (Auch et al., 2006; Patil and McHardy, 2013) and the Feature Frequency Profile (FFP) method (Wang and Ash, 2015). The oversupply of available methods is the basic problem because researches could simply choose between methods based on social reasons (quickest or cheapest) or on scientific reasons (provides the best classification for their own purposes). Since the beginning of next-generation sequencing (NGS), new species members were announced and the classification of many strains based on insufficient methods such as 16S rRNA, pathogenicity factor determination or phenotypic appearance, are questionable (Jeong et al., 2016; Li et al., 2015; Liu et al., 2015). All this lead to the quandary that the intra and inter phylogenetic relationships of Bcsl group members are highly unclear. The debate to classify all strains of the Bcsl group into a single species, to broaden the spectrum of species members or to re-classify single strains is recently discussed (Liu et al., 2015; Okinaka et al., 2016; Wang and Ash, 2015).

I.3 *Bacillus thuringiensis*-Genetics' and Phages

Bacteriophages, or short phages, are the most abundant and diverse biological entities on earth and have been found in nearly every ecological niches where bacteria are available (Brüssow and Hendrix, 2002). This includes soil, water, food, sewage, the North Sea but also extreme environments such as deserts, hot springs, and polar inland waters (Breitbart et al., 2004; Davis et al., 1985; Kumari et al., 2010; Lin et al., 2010; Lucena et al., 2006; Prigent et al., 2005; Säwström et al., 2008; Wichels et al., 1998). Besides, phages were also identified in human and animal body fluids, feces and rumen (Bachrach et al., 2003; Gantzer et al., 2002; Keller and Traub, 1974; Nigutová et al., 2008). Bacteriophages are small viruses which infect specifically bacteria and were first described in 1910 and in 1915 independently by Twort and d'Herelle (D'Herelle, 1917; Twort, 1915). Two major lifestyles of phages are described, the lytic and the lysogenic lifestyle (Figure 7).

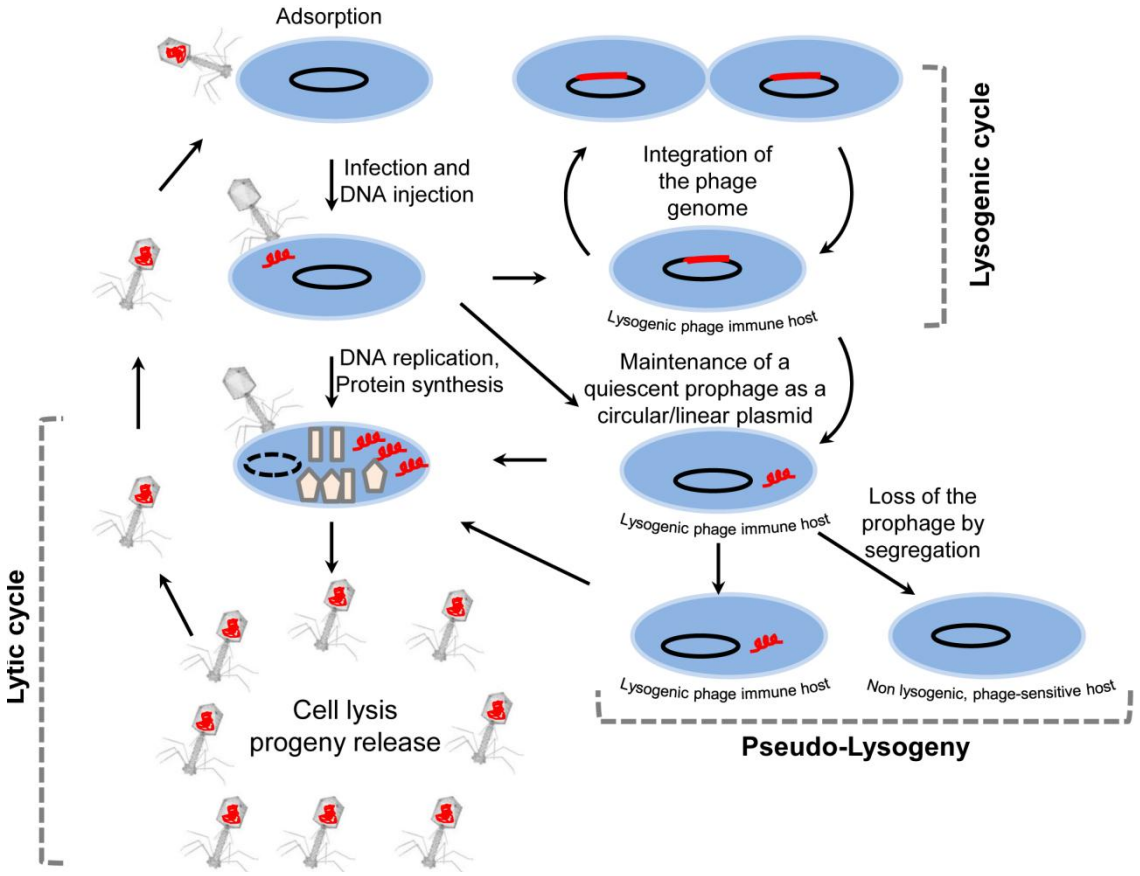

Figure 7. General lifestyles of phages. Virulent phages are only able to conduct the lytic lifecycle where the host is lysed at the end to release the phage progeny. Temperate phages have the opportunity to enter both, the lytic or the lysogenic lifecycle where the phage DNA is integrated into the bacterial chromosome and is replicated together with the host. They can also enter a "Pseudo-Lysogenic" lifecycle where the phage DNA remains as linear or circular extra-chromosomal element. The components of the figure were modified from Fortier *et al* (Fortier and Sekulovic, 2013).

By definition virulent phages only perform the lytic lifecycle, where the phage DNA is injected into a host cell, the host cell replication is reversed to phage DNA replication where phage proteins are produced and in the end the host cell is lysed for phage progeny release (Madigan and Martinko, 2013). In contrast, temperate phages can enter both lifecycles, the lytic and the lysogenic. In the lysogenic lifecycle the phage DNA is injected and directly incorporated into the host chromosome. Besides, the phage DNA can be also inserted in plasmids or self-circularize or stay as linear extra-chromosomal element in the host cell. However, considering the genetic foundation of the lytic and the lysogenic lifestyle it needs only a specific phage sensor/repressor which inhibits the required gene transcription of the lytic cycle resulting in a lysogenic lifestyle (Fortier and Sekulovic, 2013). Studies of the gene activities of integrated prophages genome confirm that a viable prophage needs only very few active genes to be replicated with the host chromosome (R. Hertel, pers. communication). In case of *B. licheniformis*

prophages, the repressor genes are linked to the universal stress response and thus can be inactivated by general stressors like Mitomycin C (Hertel et al., 2015).

There is still a lack of information regarding the high diversity of phages, due to the limited availability of prophage sequence data and because most phages are experimentally not investigated. Since next generation sequencing methods developed quickly, not only the available genome information of bacteria increased at the National Center for Biotechnology Information (NCBI) which hosts the International Nucleotide Sequence Database Collaboration comprising GenBank, DDBJ and EMBL. Also, the availability of prophage data, which were sequenced alongside with their hosts have dramatically raised (Krupovic et al., 2016). It became obvious that in pathogens the abundance of prophages and their contribution to virulence were highly underestimated (Hayashi et al., 2001; Matos et al., 2013; Wang et al., 2010; Winstanley et al., 2009). However, Brüssow *et al.* (Brüssow et al., 2004) already discussed the ability of temperate phages to modify the lifestyles, fitness, virulence and evolution of bacteria in many different ways, which is also named "lysogenic conversion". Notably, prophages dominantly confer advantageous genes, also named morons, to the host bacterium. Morons are not essential for prophages but often confer virulence to the bacterium resulting in the evolution of new pathogens (Brüssow et al., 2004). Morons are additionally defined as autonomous genetic elements and it is not fully investigated if they are part of prophages or if they use prophages to mobilize themselves (Brüssow et al., 2004). The evolution of new pathogens was observed several times for a broad spectrum of bacteria, including gram-positive and gram-negative bacteria, human and animal pathogens such as *Corynebacterium diphteriae*, *Clostridium botulinum*, *Pseudomonas aeruginosa*, *Stapphylococcus aureus*, *Streptococcus pyogenes*, *Vibrio cholera*, *Neisseria meningitides*, *Salmonella enterica*, *Shigella flexneri*, *E. coli* (EAHEC) O104:H4 GOS1 and many more (Barksdale and Arden, 1974; Brüssow et al., 2004; Brzuszkiewicz et al., 2011; Freeman, 1951). Morons can include not only powerful toxins genes. Toxins are only a subset of diverse virulence factors that could strongly affect the lifestyle, fitness, the potential to adapt and pathogenicity of bacteria. Other factors that boost virulence include proteases, antigens, resistance factors, adhesins, regulatory proteins that can increase the virulence or the fitness and thereby promote bacterial evolution (Sharma et al., 2016). Additionally, other evolutionary effects can be observed as well like gene disruption, which can lead to loss of important

functions such as virulence (Moumen et al., 2012). Prophages are also able to rearrange with taxonomically close related prophage regions in the chromosome, prophage regulators can activate or silence bacterial genes, and prophages can confer a protection against other close related phage invaders (Canchaya et al., 2004; Guan et al., 1999; Mavris et al., 1997; Spanier and Cleary, 1980).

Gillis *et al.* (Gillis and Mahillon, 2014) reviewed the specific group of phages preying on members of the *Bacillus cereus sensu lato* group and their contribution on genetic levels. Generally, phage classification is based on their morphology, host specificity and nucleic acid type. Additionally, the mode of infection, morphogenesis, phylogeny, serology and the sensitivity to physical and chemical agents are considered (Sharma et al., 2016). Recently, the classification of phages was revised by the International Committee on Taxonomy of Virus (ICTV) Release 2015 (http://www.ictvonline.org/virustaxonomy.asp) (Krupovic et al., 2016). Phages infecting members of the *Bacillus cereus senus lato* group belong to the order of *Caudovirales* including three families, *Myoviridae* (isometic head, contractile tail, small base plate), *Siphoviridae* (Isometric head, long non-contractile tail) or *Podoviridae* (isometic head, short non-contractile tail) and to an additional family the *Tectiviridae* (Isometric virion with apical spikes) (Gillis and Mahillon, 2014). Those phages comprise double-stranded DNA as genome (Ackermann, 2006; Gillis and Mahillon, 2014; Krupovic et al., 2016). Especially, these phages have diverse lifestyles (Figure 7), such as the virulent lifestyle, temperate lifestyle, the co-integrated lifestyle (also into plasmids), and the self-replicating lifestyle (either in a circular state or linear state) (Fortier and Sekulovic, 2013). Several subcategories can be made based on the most dominant feature of phages. Transducing phages, chromosomal-plasmid state phages, gamma phages, jumbo phages as well as defective phages are known which all have specific effects on the host bacterium (Bobay et al., 2014; Canchaya et al., 2003; Casjens, 2003; Van Etten et al., 2010; Fouts et al., 2006; Gillis and Mahillon, 2014; Griffiths et al., 2005; Hendrix, 2009).

In contrast, only little is known about phages of other Bcsl group members because the interest relies always on the primary pathogens (*B. anthracis*, *B. cereus* and *B. thuringiensis*). Additionally, only a fraction of bacterial genome sequences are available for the other strains in the public data repositories compared. However, the knowledge of the contribution of phages to the lifestyle and evolution of

B. thuringiensis is still fragmentary considering phage diversity, phage mechanisms, high number of unknown phages, high number of unknown putative new virulence factors in combination with the lack of basic phage research.

I.4 Evolution of *Bacillus thuringiensis*

The evolutionary biology was founded by Charles Darwin in 1859 in his book "On the Origin of Species by Means of Natural Selection, or the Preservation of Favored Races in the Struggle for Life"(Darwin, 1859). He first described the natural selection, which is a major mechanism in evolution, including two postulates "survival of the fittest" and "struggle for life". Besides, Darwin showed that variation, selection and isolation are important in population changes of organisms. Phenotypes in a population are thereby a result from the genotype, environmental influences, interaction partners and the selective pressure between them (Mitchell-Olds et al., 2007). In general, bacteria are always under selective pressure. Especially, micro-parasites such as *B. thuringiensis* have to antagonize their host organisms (Masri et al., 2015). This host-parasite interaction can remain over several generations and lead to rounds of adaptation and counter-adaptation resulting in fast evolutionary changes for both organisms (Schulte et al., 2013). This effect is also described as co-evolution or "Red Queen" effect (Rabajante et al., 2015; van Valen, 1973). Both organisms try to develop new functions and try to increase their fitness to adapt to their hosts, competitors or to enter other different ecological niches (Brüssow et al., 2004).

In general, spontaneously phenotypic variants are produced based on genetic variation which includes vertical gene transfer as well as horizontal gene transfer (Arber, 2014; Gogarten et al., 2002; Wolf et al., 2002). Changes occur in the bacterial genome depending on long-term adaptations and include nucleotide mutations, insertions or deletions with estimated evolutionary rate of 10^{-10} to 10^{-9} substitutions per nucleotide (Ochman et al., 1999; Wilson et al., 1987). Besides, fast-term evolutionary events occur with a rate of 10^{-7} to 10^{-5} substitutions per nucleotide and are represented by various mobile elements through horizontal-gene transfer (HGT) (Brüssow et al., 2004; Didelot et al., 2016). The rate of evolution in bacteria is highly diverse and depends on available horizontal-gene transfer mechanisms. *Bacillus thuringiensis* comprises many different genetic mobile elements to evolve and adapt rapidly to altered circumstances, which fits

to the overall "picture" of *B. thuringiensis* as ecological-all-rounder. There are three HGT mechanisms that contribute to evolution in *B. thuringiensis*: (i) transformation - uptake of foreign DNA or RNA (ii) transduction - bacteriophages provide new DNA, which can be beneficial or disadvantageous, and (iii) conjugation - DNA is transferred between a donor and a recipient (Thomas and Nielsen, 2005). Mobile elements confer additional genetic plasticity and include transposons, insertion elements, plasmids, bacteriophages, genomic islands, group I intron and group II intron elements. Some of these are supposed to be selfish DNA elements and are able to accumulate in the genome of bacteria over time (Werren, 2011). This includes transposons class II and insertion sequence (IS) elements which are mobile and are able to mobilize genes (Kidwell and Lisch, 2001; Kronstad and Whiteley, 1984; Lereclus et al., 1986).

Group I and group II introns are other selfish DNA elements and are able to self-splice themselves (Lambowitz and Zimmerly, 2011; Saldanha et al., 1993). Group I introns were found inserted in bacteriophages (Edgell et al., 2000). Werren suggested in 2011 (Werren, 2011) that those selfish DNA elements confer maintenance to plasmids or bacteriophages in the bacterial genome. However, in the *B. thuringiensis* species group both bacteriophages and plasmids represent unique genetic material and are spread via HGT by transduction and conjugation, respectively (Thomas and Nielsen, 2005). Genomic islands or pathogenic islands are also distributed via HGT and confer virulence factors but were not reported for *B. thuringiensis* so far (Kolstø et al., 2002).

One mechanism to counteract evolutionary processes and limit HGT is the clustered, regularly interspaced, short palindromic repeats (CRISPR) mechanism. CRISPR is considered as a bacterial immune system to defend against invading DNA of bacteriophages and plasmids (Boyaval et al., 2007; Horvath and Barrangou, 2010; Marrafini and Sontheimer, 2008). If a prophage or a plasmid invades into a cell, a short piece of sequence is incorporated as spacer into the CRISPR system for memory. If this virus or plasmid invades again, they are recognized and cleaved by specific enzymes (Marraffini, 2010). A second mechanism that limits HGT is the bacterial restriction-modification system, leading to an epigenetic alteration due to SNP based evolution and methylation guided DNA repairs systems. Those changes can affect transformation and conjugation between two bacterial strains carrying different restriction-modification systems (Arber, 2014).

Those mechanisms can limit the HGT and also reduce the evolutionary potential but these mechanisms were not reported for the whole *B. thuringiensis* species group.

I.5 General study aims

Nowadays, many evolutionary processes and involved key factors are known but how they influence natural genetic variation and phenotypic traits in bacteria is under debate. However, the mechanism of evolution as well as the adaptive potential of *B. thuringiensis* as a species is poorly understood. Especially, strains highly specific to various ecological niches also with regard to their specific hosts, the complex lifestyle switches and the selection of different virulence factors provided makes *B. thuringiensis* to a good model for the study of evolution. A glimpse is known about single evolutionary mechanisms which partially contribute to "rapid evolution", and there is a lack of information about how a given selection regime determines the various opportunities of *B. thuringiensis* as a species. The aim of this study was to investigate the different genomic mechanisms to gain insights in the emerging of successful new strains from the complex and divers evolutionary puzzle of the species *B. thuringiensis.* Therefore I wanted to investigate:

- whole genomes of nematicidal *B. thuringiensis* strains with focus on virulence factors, fitness factors, methylation pattern, metabolic properties and mobile elements such as bacteriophages, IS elements, transposases, which could contribute to a successful infection and show new insights into how *B. thuringiensis* is able to evolve, adapt and survive in various ecological niches and where the broad host range has its origin (Chapter II.1 & Chapter II.2).
- how NGS methods can be successfully implemented not only in *in depth* bacteriophage research, furthermore how phages can impact biotechnology and can be used in biotechnological relevant strains (*B. licheniformis* DSM13) where strain optimization is crucial for a sufficient production (Chapter II.3).
- the fast evolution of nematicidal *B. thuringiensis* strains in a continuous arms race with their host organism *Caenorhabditis elegans* by looking at the general genetic trait mechanism under selection, including adaptive changes in real time using large-scale phenol-typing, population next generation sequencing, and genetic analysis of the identified candidate genes (Chapter II.4).

Furthermore, *B. thuringiensis* is well known as biocontrol agent against a broad spectrum of insects, nematodes, mites and protozoa depending on their armory of toxins (Cry, Cyt, Vip and Sip toxins). However, the interaction structures such as competition, amensalism, expoitation, neutralism, commensalism, mutualism, or symbiosis in complex environments between animals, plants, fungi and *B. thuringiensis* has been rarely studied in detail. This includes also the question of the antifungal potential of *B. thuringiensis* and their potential as control agent against other pests such as phytopathogenic fungi. Therefore the second aim of this thesis was to investigate the anti-fungal potential of host-plant associated Bcsl group species, especially *B. thuringiensis,* against wilt causing phytopathogenic *Verticillia.*

Therefore I investigated:

i. the natural occurrence of wildtype *B. thuringiensis* and other Bcsl group species members sampled from *Solanum lycopersicum* (tomato) with the plant as primary ecological niche (Chapter II.5).
ii. the phylogeny of new isolates especially in the whole context of the non-trivial classification of Bcsl group members (Chapter II.5).
iii. the potential of wildtype *B. thuringiensis* and other Bcsl group species members against two phytopathogenic fungi *Verticillium dahliae* JR2 and *Vertillium longisporum* 43 which are either natural pathogens of *Solanum lycopersicum* or *Brassicacea*, respectively (oil seed rape) (Chapter II.5).

I.6 References

Ackermann, H. W. (2006). "Classification of Bacteriophages," in *The bacteriophages*, ed. R. Calendar (New York: Oxford University Press), 746.

Agata, N., Ohta, M., and Mori, M. (1996). Production of an emetic toxin, cereulide, is associated with a specific class of *Bacillus cereus*. *Curr. Microbiol.* 33, 67–69. doi:10.1007/s002849900076.

Aizawa, K., and Iida, S. (1963). Nucleic acids extracted from the virus polyhedra of the silkworm, *Bombyx mori* (Linnaeus). *J. Insect Path* 5, 344–348.

Ali, B. A., Salem, H. H., Wang, X. M., Huang, T. H., Xie, Q. D., and Zhang, X. Y. (2010). Effect of *Bacillus thuringiensis* var. *israelensis* Endotoxin on the Intermediate Snail Host of Schistosoma japonicum. doi:10.3923/crb.2010.37.41.

Angus, T. A. (1954). A Bacterial Toxin paralysing Silkworm Larvæ. *Nature* 173, 545–546. doi:10.1038/173545a0.

Arahal, D. R. (2014). "Whole-Genome Analyses: Average Nucleotide Identity," in *Methods in Microbiology: New Approaches to Prokaryotic Systematics*, eds. M. Goodfellow, I. Sutcliffe, and J. Chun (Oxford: Elsevier Ltd; Academic Press), 103–122. doi:http://dx.doi.org/10.1016/S0580-9517(14)00027-0.

Arber, W. (2014). Horizontal Gene Transfer among Bacteria and Its Role in Biological Evolution. *Life* 4, 217–24. doi:10.3390/life4020217.

Argolo-Filho, R. C., and Loguercio, L. L. (2014). *Bacillus thuringiensis* is an environmental pathogen and host-specificity has developed as an adaptation to human-generated ecological niches. *Insects* 5, 62–91. doi:10.3390/insects5010062.

Aronson, A. I., Beckman, W., and Dunn, P. (1986). *Bacillus thuringiensis* and related insect pathogens. *Microbiol. Rev.* 50, 1–24.

Auch, A. F., Henz, S. R., Holland, B. R., and Göker, M. (2006). Genome BLAST distance phylogenies inferred from whole plastid and whole mitochondrion genome sequences. *BMC Bioinformatics* 7, 350. doi:10.1186/1471-2105-7-350.

Bachrach, G., Leizerovici-Zigmond, M., Zlotkin, A., Naor, R., and Steinberg, D. (2003). Bacteriophage isolation from human saliva. *Lett. Appl. Microbiol.* 36, 50–53. doi:10.1046/j.1472-765X.2003.01262.x.

Barksdale, L., and Arden, S. B. (1974). Persisting bacteriophage infections, lysogeny, and phage conversions. *Annu. Rev. Microbiol.* 28, 265–299. doi:10.1146/annurev.mi.28.100174.001405.

Barrangou, R., Fremaux, C., Deveau, H., Boyaval, P., Richards, M., Moineau, S., et al. (2007). CRISPR Provides Acquired Resistance Against Viruses in Prokaryotes. *Science.* 315, 1709–1712. doi:10.1126/science.1138140.

Barth, H., Aktories, K., Popoff, M. R., and Stiles, B. G. (2004). Binary bacterial toxins: Biochemistry, Biology, and Applications of common *Clostridium* and *Bacillus* proteins. *Microbiol Mol Biol Rev* 68, 373–402. doi:10.1128/MMBR.68.3.373-402.2004.

Bechtel, D. B., Bulla, L. A., Kramer, K. J., Bechtel, D. B., and David-, L. I. (1976). Electron Microscope Study of Sporulation and Parasporal Crystal Formation in *Bacillus thuringiensis*. *J. Bacteriol.* 127, 1472–1481.

Beebee, T. J., and Bond, R. P. (1973). Effect of the exotoxin of *Bacillus thuringiensis* on normal and ecdysone-stimulated ribonucleic acid polymerase activity in intact nuclei from the fat-body of Sarcophaga bullata larvae. *Biochem J* 136, 1–7.

Beegle, C. C., and Yamamoto, T. (1992). History of *Bacillus thuringiensis* Berliner research and development. *Can. Entomol.* 124, 587–616. doi:10.4039/Ent124587-

4.
Ben-Dov, E. (2014). *Bacillus thuringiensis* subsp. *israelensis* and Its Dipteran-Specific Toxins. *Toxins (Basel).* 6, 1222–1243. doi:10.3390/toxins6041222.
Bennett, P. M. (2004). Genome plasticity: insertion sequence elements, transposons and integrons, and DNA rearrangement. *Methods Mol. Biol.* 266, 71–113. doi:10.1385/1-59259-763-7:071.
Bhatnagar, R., and Batra, S. (2001). Anthrax toxin. *Crit Rev Microbiol* 27, 167–200. doi:10.1080/20014091096738.
Bobay, L.-M., Touchon, M., and Rocha, E. P. C. (2014). Pervasive domestication of defective prophages by bacteria. *Proc. Natl. Acad. Sci.* 111, 12127–12132. doi:10.1073/pnas.1405336111.
Bravo, A., Gill, S. S., and Soberón, M. (2008). Mode of action of *Bacillus thuringiensis* Cry and Cyt toxins and their potential for insect control. *Natl. Institutes Acess* 49, 423–435. doi:10.1016/j.toxicon.2006.11.022.
Bravo, A., Gómez, I., Porta, H., García-Gómez, B. I., Rodriguez-Almazan, C., Pardo, L., et al. (2013). Evolution of *Bacillus thuringiensis* Cry toxins insecticidal activity. *Microb. Biotechnol.* doi:10.1111/j.1751-7915.2012.00342.x.
Bravo, A., Likitvivatanavong, S., Gill, S. S., and Soberón, M. (2011). *Bacillus thuringiensis*: A story of a successful bioinsecticide. *Insect Biochem. Mol. Biol.* 41, 423–431. doi:10.1016/j.ibmb.2011.02.006.
Bravo, A., Sarabia, S., Lopez, L., Ontiveros, H., Abarca, C., Ortiz, A., et al. (1998). Characterization of cry Genes in a Mexican *Bacillus thuringiensis* Strain Collection. *Appl. Environ. Microbiol.* 64, 4965–4972. Available at: http://aem.asm.org/content/64/12/4965.
Breitbart, M., Wegley, L., Leeds, S., Rohwer, F., and Schoenfeld, T. (2004). Phage Community Dynamics in Hot Springs These include : Phage Community Dynamics in Hot Springs. *Appl. Environ. Microbiol.* 70, 1633–1640. doi:10.1128/AEM.70.3.1633.
Brüssow, H., Canchaya, C., and Hardt, W.-D. (2004). Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. *Microbiol. Mol. Biol. Rev.* 68, 560–602. doi:10.1128/MMBR.68.3.560–602.2004.
Brüssow, H., and Hendrix, R. W. (2002). Phage Genomics: Small is beautiful. *Cell* 108, 13–16. doi:10.1016/S0092-8674(01)00637-7.
Brzuszkiewicz, E., Thürmer, A., Jörg, S., Leimbach, A., Liesegang, H., Meyer, F.-D., et al. (2011). Genome sequence analyses of two isolates from the recent *Escherichia coli* outbreak in Germany reveal the emergence of a new pathotype: Entero-Aggregative-Haemorrhagic *Escherichia coli* (EAHEC). *Arch. Microbiol.* 193, 883–891. doi:10.1007/s00203-011-0725-6.
Bulla, L. A. (1975). Bacteria as insect pathogens. *Annu. Rev. Microbiol.* 29, 163–190. doi:10.1146/annurev.mi.29.100175.001115.
Burges, D., and Hurst, J. A. (1977). Ecology of *Bacillus thuringiensis* in storage moths. *J. Invertebr. Pathol.* 30, 131–139. doi:10.1016/0022-2011(77)90210-5.
Butko, P. (2003). Cytolytic toxin Cyt1A and its mechanism of membrane damage: Data and hypotheses. *Appl. Environ. Microbiol.* 69, 2415–2422. doi:10.1128/AEM.69.5.2415-2422.2003.
Canchaya, C., Fournous, G., and Brüssow, H. (2004). The impact of prophages on bacterial chromosomes. *Mol. Microbiol.* 53, 9–18. doi:10.1111/j.1365-2958.2004.04113.x.
Canchaya, C., Proux, C., Fournous, G., Bruttin, A., and Brüssow, H. (2003). Prophage Genomics. *Microbiol. Mol. Biol. Rev.* 67, 238–276. doi:10.1128/MMBR.67.2.238-

276.2003.

Carlson, C. R., Caugant, D. A., and Kolsto, A. B. (1994). Genotypic diversity among *Bacillus cereus* and *Bacillus thuringiensis* strains. *Appl. Environ. Microbiol.* 60, 1719–1725.

Carrière, Y., Crickmore, N., and Tabashnik, B. E. (2015). Optimizing pyramided transgenic Bt crops for sustainable pest management. *Nat. Biotechnol.* 33, 161–168. doi:10.1038/nbt.3099.

Casjens, S. (2003). Prophages and bacterial genomics: What have we learned so far? *Mol. Microbiol.* 49, 277–300. doi:10.1046/j.1365-2958.2003.03580.x.

Celandroni, F., Salvetti, S., Senesi, S., and Ghelardi, E. (2014). *Bacillus thuringiensis* membrane-damaging toxins acting on mammalian cells. *FEMS Microbiol. Lett.* 361, 95–103. doi:10.1111/1574-6968.12615.

Chakroun, M., Banyuls, N., Bel, Y., Escriche, B., and Ferré, J. (2016). Bacterial Vegetative Insecticidal Proteins (Vip) from Entomopathogenic Bacteria. *Microbiol. Mol. Biol. Rev.* 80, 329–350. doi:10.1128/MMBR.00060-15.

Charles, J.-F., Nielson-LeRoux, C., and Delécluse, A. (1996). Bacillus Sphaericus Toxins: Molecular Biology and Mode of Action. *Annu. Rev. Entomol.* 41, 451–472. doi:10.1146/annurev.en.41.010196.002315.

Chen, M. L., and Tsen, H. Y. (2002). Discrimination of *Bacillus cereus* and *Bacillus thuringiensis* with 16S rRNA and gyrB gene based PCR primers and sequencing of their annealing sites. *J. Appl. Microbiol.* 92, 912–919. doi:10.1046/j.1365-2672.2002.01606.x.

Choudhary, B., Nasiruddin, K. M., and Gaur, K. (2014). The Status of Commercialized Bt Brinjal in Bangladesh. Ithaca, NY.

Crickmore, N., Baum, J., Bravo, A., Lereclus, D., Narva, K., Sampson, K., et al. (2016). *Bacillus thuringiensis* toxin nomenclature. Available at: http://www.btnomenclature.info/.

D'Herelle, F. (1917). On an invisible microbe antagonistic to dysentery bacilli. *Comptes Rendus Acad. des Sci.* 165, 373–375.

Darwin, C. M. A. (1859). *The origin of species by means of natural selection, or the preservation of favoured races in the struggle for life*. 1st ed. , ed. J. Murray London.

Davis, C., Silveira, N. F., and Fleet, G. H. (1985). Occurrence and properties of bacteriophages of *Leuconostoc oenos* in Australian wines. *Appl. Environ. Microbiol.* 50, 872–876.

Didelot, X., Walker, S. A., Peto, T. E., Crook, D. W., Wilson, D. J., and Walker, A. S. (2016). Within-host evolution of bacterial pathogens. *Nat. Rev. Microbiol.* 14, 150–62. doi:10.1038/nrmicro.2015.13.

Donovan, W. P., Engleman, J. T., Donovan, J. C., Baum, J. A., Bunkers, G. J., Chi, D. J., et al. (2006). Discovery and characterization of Sip1A: A novel secreted protein from *Bacillus thuringiensis* with activity against coleopteran larvae. *Appl. Microbiol. Biotechnol.* doi:10.1007/s00253-006-0332-7.

Donovan, W. P., Gonzalez, J. M., Gilbert, M. P., and Dankocsik, C. (1988). Isolation and characterization of EG2158, a new strain of *Bacillus thuringiensis* toxic to coleopteran larvae, and nucleotide sequence of the toxin gene. *Mol. Gen. Genet.* 214, 365–72.

Drobniewski, F. A. (1993). *Bacillus cereus* and related species. *Clin. Microbiol. Rev.* 6, 324–338. doi:10.1128/CMR.6.4.324.

Dubois, T., Faegri, K., Perchat, S., Lemy, C., Buisson, C., Nielsen-LeRoux, C., et al. (2012). Necrotrophism Is a Quorum-Sensing-Regulated Lifestyle in *Bacillus*

thuringiensis. *PLoS Pathog.* 8, e1002629. doi:10.1371/journal.ppat.1002629.

Dulmage, H. (1970). Insecticidal activity of HD-1, a new isolate of *Bacillus thuringiensis* var. *alesti*. *J. Invertebr. Pathol.* 15, 232–239.

Edgell, D. R., Belfort, M., and Shub, D. A. (2000). Barriers to intron promiscuity in bacteria. *J. Bacteriol.* 182, 5281–5289. doi:10.1128/JB.182.19.5281-5289.2000.

Ehling-Schulz, M., Fricker, M., Grallert, H., Rieck, P., Wagner, M., and Scherer, S. (2006). Cereulide synthetase gene cluster from emetic *Bacillus cereus*: structure and location on a mega virulence plasmid related to *Bacillus anthracis* toxin plasmid pXO1. *BMC Microbiol.* 6, 20. doi:10.1186/1471-2180-6-20.

Erban, T., Nesvorna, M., Erbanova, M., and Hubert, J. (2009). *Bacillus thuringiensis* var. *tenebrionis* control of synanthropic mites (Acari: Acaridida) under laboratory conditions. *Exp. Appl. Acarol.* 49, 339–346. doi:10.1007/s10493-009-9265-z.

Ernst, B. (1915). Ueber die Schlaffsucht der Mehlmottenraupe *Ephestia kuhniella* und ihren Erreger *Bacillus thuringiensis* n. sp. *Angew. Entomol.* 2, 21–56.

Errington, J. (2003). Regulation of endospore formation in *Bacillus subtilis*. *Nat Rev Microbiol* 1, 117–126. doi:10.1038/nrmicro750.

Estruch, J. J., Warren, G. W., Mullins, M. a, Nye, G. J., Craig, J. a, and Koziel, M. G. (1996). Vip3A, a novel *Bacillus thuringiensis* vegetative insecticidal protein with a wide spectrum of activities against lepidopteran insects. *Proc. Natl. Acad. Sci. U. S. A.* 93, 5389–5394. doi:10.1073/pnas.93.11.5389.

Van Etten, J. L., Lane, L. C., and Dunigan, D. D. (2010). DNA Viruses: The Really Big Ones (Giruses). *Annu. Rev. Microbiol* 64, 83–99. doi:10.1146/annurev.micro.112408.134338.

Fedhila, S., Nel, P., and Lereclus, D. (2002). The InhA2 metalloprotease of *Bacillus thuringiensis* strain 407 is required for pathogenicity in insects infected via the oral route. *J. Bacteriol.* 184, 3296–3304. doi:10.1128/JB.184.12.3296-3304.2002.

Fernández-Ruvalcaba, M., Peña-Chora, G., Romo-Martínez, A., Hernández-Velázquez, V., Bravo De La Parra, A., and De La Rosa, D. P. (2010). Evaluation of *Bacillus thuringiensis* pathogenicity for a strain of the tick, *Rhipicephalus microplus*, resistant to chemical pesticides. *J. Insect Sci.* 10, 186. doi:10.1673/031.010.14146.

Fortier, L.-C., and Sekulovic, O. (2013). Importance of prophages to evolution and virulence of bacterial pathogens. *Virulence* 4, 354–65. doi:10.4161/viru.24498.

Fouts, D. E., Rasko, D. A., Cer, R. Z., Jiang, L., Fedorova, N. D. N. B., Shvartsbeyn, A., et al. (2006). Sequencing *Bacillus anthracis* Typing Phages Gamma and Cherry Reveals a Common Ancestry. *J. Bacteriol.* 188, 3402–3408. doi:10.1128/JB.188.9.3402.

van Frankenhyzen, K. (1993). "The challenge of *Bacillus thuringiensis*.," in *Bacillus thuringiensis, an environmental biopesticide: theory and practice.*, eds. P. E. Entwistle, J. S. Cory, M. J. Bailey, and S. Higgs (Chichester: Wiley), 1–35.

Freeman, V. J. (1951). Studies on the virulence of bacteriophage-infected strains of *Corynebacterium diphtheriae*. *J. Bacteriol.* 61, 675–688.

Frost, L. S., Leplae, R., Summers, A. O., and Toussaint, A. (2005). Mobile genetic elements: the agents of open source evolution. *Nat. Rev. Microbiol.* 3, 722–732. doi:10.1038/nrmicro1235.

Gantzer, C., Henny, J., and Schwartzbrod, L. (2002). Bacteroides fragilis and *Escherichia coli* bacteriophages in human faeces. *Int. J. Hyg. Environ. Health* 205, 325–8. doi:10.1078/1438-4639-00152.

Gillis, A., and Mahillon, J. (2014). Phages preying on *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*: Past, present and future. *Viruses* 6, 2623–2672. doi:10.3390/v6072623.

Gogarten, J. P., Doolittle, W. F., and Lawrence, J. G. (2002). Prokaryotic evolution in light of gene transfer. *Mol. Biol. Evol.* 19, 2226–2238. doi:10.1093/oxfordjournals.molbev.a004046.

Granum, P. E., O'Sullivan, K., and Lund, T. (1999). The sequence of the non-haemolytic enterotoxin operon from *Bacillus cereus*. *FEMS Microbiol. Lett.* 177, 225–229. doi:10.1016/S0378-1097(99)00312-2.

Griffiths, J. A., Miller, H. J., Suzuki, T. D., Lewontin, C. R., and Gelbart, M. W. (2005). "The genetics of bacteria and their viruses," in *An introduction to genetic analysis* (New York: H.W., Freeman and Company), 1-172.

Guan, S., Bastin, A. D., and Verma, N. K. (1999). Functional analysis of the O antigen glucosylation gene cluster of *Shigella flexneri* bacteriophage SfX. *Microbiology* 145, 1263–1273.

Guerchicoff, A., Delécluse, A., and Rubinstein, C. P. (2001). The *Bacillus thuringiensis cyt* Genes for Hemolytic Endotoxins Constitute a Gene Family. *Appl. Environ. Microbiol.* 67, 1090–1096. doi:10.1128/AEM.67.3.1090-1096.2001.

Guillem, M., and Porcar, M. (2012). Ecological Mysteries: is *Bacillus thuringiensis* a Real Insect Pathogen? *Bt Res.* 3, 1–2. doi:10.5376/bt.2012.03.0001.

Guillet, P., Kurtak, D. C., Philippon, B., and Meyer, R. (1990). "Use of *Bacillus thuringiensis israelensis* for Onchocerciasis Control in West Africa," in *Bacterial Control of Mosquitoes & Black Flies* (Dordrecht: Springer Netherlands), 187–201.

Guinebretière, M. H., Auger, S., Galleron, N., Contzen, M., de Sarrau, B., de Buyser, M. L., et al. (2013). *Bacillus cytotoxicus* sp. *nov*. is a novel thermotolerant species of the *Bacillus cereus* group occasionally associated with food poisoning. *Int. J. Syst. Evol. Microbiol.* 63, 31–40. doi:10.1099/ijs.0.030627-0.

Han, S., Craig, J. A., Putnam, C. D., Carozzi, N. B., and Tainer, J. A. (1999). Evolution and mechanism from structures of an ADP-ribosylating toxin and NAD complex. *Nat. Struct. Biol.* 6, 932–936. doi:10.1038/13300.

Hannay, C. L. (1953). Crystalline inclusions in aerobic spore-forming bacteria. *Nature* 172, 1004.

Hayashi, T., Makino, K., Ohnishi, M., Kurokawa, K., Ishii, K., Yokoyama, K., et al. (2001). Complete genome sequence of enterohemorrhagic *Escherichia coli* O157 : H7 and genomic comparison with a laboratory strain K-12. *DNA Res.* 8, 11–22. doi:10.1093/dnares/8.1.11.

Hendrix, R. W. (2009). Jumbo Bacteriophages. *Curr. Top. Microbiol. Immunol.* 328, 229–240.

Hertel, R., Rodríguez, D. P., Hollensteiner, J., Dietrich, S., Leimbach, A., Hoppert, M., et al. (2015). Genome-Based Identification of Active Prophage Regions by Next Generation Sequencing in *Bacillus licheniformis* DSM13. *PLoS One* 10, e0120759. doi:10.1371/journal.pone.0120759.

Hill, K. K., Ticknor, L. O., Okinaka, R. T., Asay, M., Blair, H., Bliss, K. A., et al. (2004). Fluorescent Amplified Fragment Length Polymorphism Analysis of *Bacillus anthracis* , *Bacillus cereus* , and *Bacillus thuringiensis* Isolates. 70, 1068–1080. doi:10.1128/AEM.70.2.1068.

Höfte, H., and Whiteley, H. R. (1989). Insecticidal crystal proteins of *Bacillus thuringiensis*. *Microbiol. Rev.* 53, 242–255.

Horvath, P., and Barrangou, R. (2010). CRISPR/Cas, the Immune System of Bacteria and Archaea. *Science (80-.).* 327, 167–170. doi:10.1126/science.1179555.

Husz, B. (1928). *Bacilllus thuringiensis* Berl. A bacterium pathogenic to corn borer larvae. *Int Corn Vor. Invest Sci Rep* 1, 191–193.

Ibrahim, M. A., Griko, N., Junker, M., and Bulla, L. A. (2010). *Bacillus thuringiensis*:

A genomics and proteomics perspective. *Bioeng. Bugs* 1, 31–50. doi:10.4161/bbug.1.1.10519.

Iriarte, J., Porcar, M., Lecadet, M.-M., and Caballero, P. (2000). Isolation and Characterization of *Bacillus thuringiensis* Strains from Aquatic Environments in Spain. *Curr. Microbiol.* 40, 402–408. doi:10.1007/s002840010078.

James, C., and Krattiger, A. F. (1996). Global Review of the Field Testing and Commercialization of Transgenic Plants: 1986 to 1995 The First Decade of Crop Biotechnology.

Jensen, G. B., Hansen, B. M., Eilenberg, J., and Mahillon, J. (2003). The hidden lifestyles of *Bacillus cereus* and relatives. *Environ. Microbiol.* 5, 631–640. doi:10.1046/j.1462-2920.2003.00461.x.

Jeong, H., Jo, S. H., Hong, C. E., and Park, J. M. (2016). Genome Sequence of the Endophytic Bacterium *Bacillus thuringiensis* Strain KB1 , a Potential Biocontrol Agent against Phytopathogens. *Genome Announc.* 4, 9–10. doi:10.1128/genomeA.00279-16.Copyright.

Jones, G. W., Nielsen-Leroux, C., Yang, Y., Yuan, Z., Dumas, V. F., Gomes Monnerat, R., et al. (2007). A new Cry toxin with a unique two-component dependency from *Bacillus sphaericus*. *FASEB J.* 21, 4112–4120. doi:10.1096/fj.07-8913com.

Joung, K. B., and Côté, J. C. (2001). Phylogenetic analysis of *Bacillus thuringiensis* serovars based on 16S rRNA gene restriction fragment length polymorphisms. *J. Appl. Microbiol.* 90, 115–122. doi:10.1096/fj.07-8913com.

Jucovic, M., Walters, F. S., Warren, G. W., Palekar, N. V., and Chen, J. S. (2008). From enzyme to zymogen: Engineering Vip2, an ADP-ribosyltransferase from *Bacillus cereus*, for conditional toxicity. *Protein Eng. Des. Sel.* 21, 631–638. doi:10.1093/protein/gzn038.

Jurat-Fuentes, J. L., and Crickmore, N. (2016). Specificity determinants for Cry insecticidal proteins: Insights from their mode of action. *J. Invertebr. Pathol.* doi:10.1016/j.jip.2016.07.018.

Keller, R., and Traub, N. (1974). The Characterization of Bacteroides fragilis Bacteriophage Recovered from Animal Sera : observations on the Nature of Bacteroides Phage Carrier Cultures. *J. gen. Virol.* 24, 179–189.

Kidwell, M. G., and Lisch, D. R. (2001). Perspective: transposable elements, parasitic DNA, and genome evolution. *Evolution* 55, 1–24.

Kim, M.-J., Han, J.-K., Park, J.-S., Lee, J.-S., Lee, S.-H., Cho, J.-I. Il, et al. (2015). Various enterotoxin and other virulence factor genes widespread among *Bacillus cereus* and *Bacillus thuringiensis* strains. *J. Microbiol. Biotechnol.* 25, 872–879. doi:10.4014/jmb.1502.02003.

Knapp, O., Stiles, B., and Popoff, M. R. (2010). The Aerolysin-Like Toxin Family of Cytolytic, Pore-Forming Toxins. *Open Toxinology J.* 3, 53–68. doi:10.2174/1875414701003020053.

Kolstø, A. B., Lereclus, D., and Mock, M. (2002). "Genome Structure and Evolution of the *Bacillus cereus* Group," in *Pathogenicity Islands and the Evolution of Pathogenic Microbes*, eds. J. Hacker and Kaper B.J. (Springer Berlin Heidelberg), 95–105.

Krieg, A., Huger, A. M., Langenbruch, G. A., and Schnetter, W. (2009). *Bacillus thuringiensis* var. *tenebrionis*: ein neuer, gegenüber Larven von Coleopteren wirksamer Pathotyp. *Zeitschrift für Angew. Entomol.* 96, 500–508. doi:10.1111/j.1439-0418.1983.tb03704.x.

Krishnan, V. (2013). Investigation of parasporins , the cytotoxic proteins from the bacterium *Bacillus thuringiensis*, *Department of Biochemistry, School of Life*

Sciences, Univeristy of Sussex.

Kronstad, J. W., Schnepf, H. E., and Whiteley, H. R. (1983). Diversity of locations for *Bacillus thuringiensis* crystal protein genes. *J. Bacteriol.* 154, 419–428.

Kronstad, J. W., and Whiteley, H. R. (1984). Inverted Repeat Sequences Flank a *Bacillus thuringiensis* Crystal Protein Gene. *J. Bacteriol.* 160, 95–102.

Krupovic, M., Dutilh, B. E., Adriaenssens, E. M., Wittmann, J., Vogensen, F. K., Sullivan, M. B., et al. (2016). Taxonomy of prokaryotic viruses: update from the ICTV bacterial and archaeal viruses subcommittee. *Arch. Virol.* 161, 1095–1099. doi:10.1007/s00705-015-2728-0.

Kumari, S., Harjai, K., and Chhibber, S. (2010). Isolation and characterization of *Klebsiella pneumoniae* specific bacteriophages from sewage samples. *Folia Microbiol. (Praha).* 55, 221–227. doi:10.1007/s12223-010-0032-7.

Kurstak, E. (1970). Donnees sur l'epizootie bacterienne naturelle provoguee par un *Bacillus* du type *Bacillus thuringiensis* sur *Ephestia kuhniella* Zeller. *Enomophage Mem Hors Ser* 2, 245–247.

de la Fuente-Salcido, N. M., Casados-Vázquez, L. E., and Barboza-Corona, J. E. (2013). Bacteriocins of *Bacillus thuringiensis* can expand the potential of this bacterium to other areas rather than limit its use only as microbial insecticide. *Can. J. Microbiol.* 59, 515–22. doi:10.1139/cjm-2013-0284.

Lambowitz, A. M., and Zimmerly, S. (2011). Group II introns: Mobile ribozymes that invade DNA. *Cold Spring Harb. Perspect. Biol.* 3, 1–19. doi:10.1101/cshperspect.a003616.

Lechner, S., Mayr, R., Francis, K. P., Prüss, B. M., Kaplan, T., Wiessner-Gunkel, E., et al. (1998). *Bacillus weihenstephanensis* sp. *nov*. is a new psychrotolerant species of the *Bacillus cereus* group. *Int. J. Syst. Bacteriol.* 48 Pt, 1373–1382. doi:10.1099/00207713-48-4-1373.

Lemes, A. R. ita N., Davolos, C. C. hiaradia, Legori, P. C. ristina B. C. rialesi, Fernandes, O. A. parecido, Ferré, J., Lemos, M. V. ictor F., et al. (2014). Synergism and antagonism between *Bacillus thuringiensis* Vip3A and Cry1 proteins in *Heliothis virescens*, *Diatraea saccharalis* and *Spodoptera frugiperda*. *PLoS One* 9, e107196. doi:10.1371/journal.pone.0107196.

Lenane, I. J., Bagnall, N. H., Josh, P. F., Pearson, R. D., Akhurst, R. J., and Kotze, A. C. (2008). A pair of adjacent genes, cry5Ad and orf2-5Ad, encode the typical N- and C-terminal regions of a Cry5Adelta-endotoxin as two separate proteins in *Bacillus thuringiensis* strain L366. *FEMS Microbiol. Lett.* 278, 115–120. doi:10.1111/j.1574-6968.2007.00987.x.

Lereclus, D., Mahillon, J., Menou, G., and Lecadet, M. M. (1986). Identification of Tn4430, a transposon of *Bacillus thuringiensis* functional in *Escherichia coli*. *Mol. Gen. Genet.* 204, 52–7.

Leuber, M., Orlik, F., Schiffler, B., Sickmann, A., and Benz, R. (2006). Vegetative insecticidal protein (Vip1Ac) of *Bacillus thuringiensis* HD201: evidence for oligomer and channel formation. *Biochemistry* 45, 283–8. doi:10.1021/bi051351z.

Li, C., Xu, N., Huang, X., Wang, W., Cheng, J., Wu, K., et al. (2007). *Bacillus thuringiensis* Vip3 mutant proteins: Insecticidal activity and trypsin sensitivity. *Biocontrol Sci. Technol.* 17, 699–708. doi:10.1080/09583150701527177.

Li, G., Wang, Y., Liu, B., and Zhang, G. (2014). Transgenic *Bacillus thuringiensis* (Bt) Rice Is Safer to Aquatic Ecosystems than Its Non-Transgenic Counterpart. *PLoS One* 9, e104270. doi:10.1371/journal.pone.0104270.

Li, Q., Xu, L. Z., Zou, T., Ai, P., Huang, G. H., Li, P., et al. (2015). Complete genome sequence of *Bacillus thuringiensis* strain HD521. *Stand. Genomic Sci.* 10, 62.

doi:10.1186/s40793-015-0058-1.

Lin, L., Hong, W., Ji, X., Han, J., Huang, L., and Wei, Y. (2010). Isolation and characterization of an extremely long tail Thermus bacteriophage from Tengchong hot springs in China. *J. Basic Microbiol.* 50, 452–456. doi:10.1002/jobm.201000116.

Lind, P. A., and Andersson, D. I. (2008). Whole-genome mutational biases in bacteria. *Proc Natl Acad Sci U S A* 105, 17878–17883. doi:10.1073/pnas.0804445105.

Liu, X., Ruan, L., Peng, D., Li, L., Sun, M., and Yu, Z. (2014). Thuringiensin: A Thermostable Secondary Metabolite from *Bacillus thuringiensis* with Insecticidal Activity against a Wide Range of Insects. *Toxins.* 6, 2229–2238. doi:10.3390/toxins6082229.

Liu, Y., Lai, Q., Göker, M., Meier-Kolthoff, J. P., Wang, M., Sun, Y., et al. (2015). Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Sci. Rep.* 5, 14082. doi:10.1038/srep14082.

Lucena, F., Ribas, F., Duran, A. E., Skraber, S., Gantzer, C., Campos, C., et al. (2006). Occurrence of bacterial indicators and bacteriophages infecting enteric bacteria in groundwater in different geographical areas. *J. Appl. Microbiol.* 101, 96–102. doi:10.1111/j.1365-2672.2006.02907.x.

Luo, X., Chen, L., Huang, Q., Zheng, J., Zhou, W., Peng, D., et al. (2013). *Bacillus thuringiensis* metalloproteinase Bmp1 functions as a nematicidal virulence factor. *Appl. Environ. Microbiol.* 79, 460–468. doi:10.1128/AEM.02551-12.

De Maagd, R. A., Bravo, A., Berry, C., Crickmore, N., and Schnepf, H. E. (2003). Structure, Diversity, and Evolution of Protein Toxins from Spore-Forming Entomopathogenic Bacteria. *Annu. Rev. Genet* 37, 409–33. doi:10.1146/annurev.genet.37.110801.143042.

De Maagd, R. A., Bravo, A., and Crickmore, N. (2001). How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. *Trends Genet.* 17, 193–199. doi:10.1016/S0168-9525(01)02237-5.

Madigan, M. T., and Martinko, J. M. (2013). *Brock Mikrobiologie.* Pearson Studium ein Imprint von Pearson Deutschland.

Manasherob, R., Ben-dov, E., and Zaritsky, A. (1998). Germination , Growth , and Sporulation of *Bacillus thuringiensi*s subsp . *israelensis* in Excreted Food Vacuoles of the Protozoan *Tetrahymena pyriformis*. *Appl. Environ. Microbiol.* 64, 1750–1758.

Marraffini, L. a (2010). Impact of CRISPR immunity on the emergence of bacterial pathogens. *Future Microbiol.* 5, 693–5. doi:10.2217/fmb.10.38.

Marrafini, L. A., and Sontheimer, E. J. (2008). CRISPR Interference Limits Horizontal Gene Transfer in Staphylococci by Targeting DNA. *Science.* 322, 1843–1845. doi:10.1126/science.1165771.CRISPR.

Martin, P. A. W., Traverst, R. S., Travers, R. S., and Traverst, R. S. (1989). Worldwide Abundance and Distribution of *Bacillus thuringiensis* Isolates Worldwide Abundance and Distribution of *Bacillus thuringiensis* Isolates. *Appl. Environ. Microbiol.* 55, 2437–2442.

Masri, L., Branca, A., Sheppard, A. E., Papkou, A., Laehnemann, D., Guenther, P. S., et al. (2015). Host–Pathogen Coevolution: The Selective Advantage of *Bacillus thuringiensis* Virulence and Its Cry Toxin Genes. *PLOS Biol.* 13, e1002169. doi:10.1371/journal.pbio.1002169.

Matos, R. C., Lapaque, N., Rigottier-Gois, L., Debarbieux, L., Meylheuc, T., Gonzalez-Zorn, B., et al. (2013). *Enterococcus faecalis* Prophage Dynamics and Contributions to Pathogenic Traits. *PLoS Genet.* 9, e1003539.

doi:10.1371/journal.pgen.1003539.

Mavris, M., Manning, P. A., and Morona, R. (1997). Mechanism of bacteriophage SfII-mediated serotype conversion in *Shigella flexneri*. *Mol. Microbiol.* 26, 939–50. doi:10.1046/j.1365-2958.1997.6301997.x.

McGaughey, W. H. (1984). Insect Resistance to the Biological Insecticide *Bacillus thuringiensis*. *Science.* 229, 193–195.

Mendelsohn, M., Kough, J., Vaituzis, Z., and Matthews, K. (2003). Are Bt crops safe? *Nat. Biotechnol.* 21, 1003–1009. doi:10.1038/nbt0903-1003.

Mitchell-Olds, T., Willis, J. H., and Goldstein, D. B. (2007). Which evolutionary processes influence natural genetic variation for phenotypic traits? *Nat. Rev. Genet.* 8, 845–856. doi:10.1038/nrg2207.

Mock, M., and Fouet, A. (2001). Anthrax. *Annu. Rev. Microbiol.* 55, 647–671.

Moumen, B., Nguen-The, C., and Sorokin, A. (2012). Sequence Analysis of Inducible Prophage phIS3501 Integrated into the Haemolysin II Gene of *Bacillus thuringiensis* var *israelensis* ATCC35646. *Genet. Res. Int.* 2012, e543286. doi:10.1155/2012/543286.

Nair, J. R., Narasimman, G., and Sekar, V. (2004). Cloning and partial characterizaton of zwittermicin a resistance gene cluster from *Bacillus thuringiensis* subsp. *kurstaki* strain HD1. *J. Appl. Microbiol.* 97, 495–503. doi:10.1111/j.1365-2672.2004.02312.x.

Nakamura, L. K. (1998). *Bacillus pseudomycoides* sp. *nov*. *Int. J. Syst. Bacteriol.* 48 Pt 3, 1031–5. doi:10.1099/00207713-48-3-1031.

Nakamura, L. K., and Jackson, M. A. (1995). Clarification of the taxonomy of *Bacillus mycoides*. *Int. J. Syst. Bacteriol.* 45, 46–49.

Nigutová, K., Štyriak, I., Javorský, P., and Pristaš, P. (2008). Partial characterization of *Enterococcus faecalis* bacteriophage F4. *Folia Microbiol.* 53, 234–236. doi:10.1007/s12223-008-0033-y.

Ochman, H., Elwyn, S., and Moran, N. A. (1999). Calibrating bacterial evolution. *Proc. Natl. Acad. Sci. U. S. A.* 96, 12638–43. doi:10.1073/pnas.96.22.12638.

Ohba, M., Mizuki, E., and Uemori, A. (2009). Parasporin, a new anticancer protein group from *Bacillus thuringiensis*. *Anticancer Res.* 29, 427–433.

Ohgushi, A., Saitoh, H., Wasano, N., Uemori, A., and Ohba, M. (2005). Cloning and Characterization of Two Novel Genes, *cry*24B and *s1orf2*, from a Mosquitocidal Strain of *Bacillus thuringiensis* serovar *sotto*. *Curr. Microbiol.* 51, 131–136. doi:10.1007/s00284-005-7529-3.

Okinaka, T. R., Keim, P., Okinaka, R. T., and Keim, P. (2016). The Phylogeny of *Bacillus cereus sensu lato*. *Microbiol. Spectr.* 4, TBS-0012-2012. doi:10.1128/microbiolspec.TBS-0012-2012.

Palma, L., Muñoz, D., Berry, C., Murillo, J., and Caballero, P. (2014). *Bacillus thuringiensis* Toxins: An Overview of Their Biocidal Activity. *Toxins.* 6, 3296–3325. doi:10.3390/toxins6123296.

Pardo-López, L., Soberón, M., and Bravo, A. (2013). *Bacillus thuringiensis* insecticidal three-domain Cry toxins: Mode of action, insect resistance and consequences for crop protection. *FEMS Microbiol. Rev.* 37, 3-22. doi:10.1111/j.1574-6976.2012.00341.x.

Patil, K. R., and McHardy, A. C. (2013). Alignment-free genome tree inference by learning group-specific distance metrics. *Genome Biol. Evol.* 5, 1470–1484. doi:10.1093/gbe/evt105.

Pé Rez, C., Fernandez, L. E., Sun, J., Folch, J. L., Gill, S. S., Soberón, M., et al. (2005). *Bacillus thuringiensis* subsp. *israelensis* Cyt1Aa Synergizes Cry11Aa Toxin by

Functioning as a Membrane-bound Receptor. *Proc. Natl. Acad. Sci. U. S. A.* 102, 18303–18308. doi:10.1073/pnas.0505494102.

Perchat, S., Talagas, A., Poncet, S., Lazar, N., Li de la Sierra-Gallay, I., Gohar, M., et al. (2016). How Quorum Sensing Connects Sporulation to Necrotrophism in *Bacillus thuringiensis*. *PLOS Pathog.* 12, e1005779. doi:10.1371/journal.ppat.1005779.

Plumer, B. (2016). Why Bayer's massive deal to buy Monsanto is so worrisome - Vox. Available at: http://www.vox.com/2016/9/14/12916344/monsanto-bayer-merger [Accessed September 15, 2016].

Poopathi, S., Thirugnanasambantham, K., Mani, C., Ragul, K., and Sundarapandian, S. M. (2014). Isolation of mosquitocidal bacteria (*Bacillus thuringiensis*, *B.sphaericus* and *B. cereus*) from excreta of arid birds. *Indian J. Exp. Biol.* 52, 739–47.

Priest, F. G., Barker, M., Baillie, L. W. J., Holmes, E. C., and Maiden, M. C. J. (2004). Population Structure and Evolution of the *Bacillus cereus* Group. *Society* 186, 7959–7970. doi:10.1128/JB.186.23.7959.

Priest, F. G., Kaji, D. A., Rosato, Y. B., and Canhos, V. P. (1994). Characterization of *Bacillus thuringiensis* and related bacteria by ribosomal RNA gene restriction fragment length polymorphisms. *Microbiology* 140, 1015–1022. doi:10.1099/13500872-140-5-1015.

Prigent, M., Leroy, M., Confalonieri, F., Dutertre, M., and DuBow, M. S. (2005). A diversity of bacteriophage forms and genomes can be isolated from the surface sands of the Sahara Desert. *Extremophiles* 9, 289–296. doi:10.1007/s00792-005-0444-5.

Rabajante, J. F., Tubay, J. M., Uehara, T., Morita, S., Ebert, D., and Yoshimura, J. (2015). Red Queen dynamics in multi-host and multi-parasite interaction system. *Sci. Rep.* 5, 10004. doi:10.1038/srep10004.

Rang, C., Gil, P., Neisner, N., Rie, J. Van, and Frutos, R. (2005). Novel Vip3-Related Protein from *Bacillus thuringiensis*. *Appl. Environ. Microbiol.* 71, 6276–6281. doi:10.1128/AEM.71.10.6276.

Raymond, B., Johnston, P. R., Nielsen-LeRoux, C., Lereclus, D., and Crickmore, N. (2010). *Bacillus thuringiensis*: An impotent pathogen? *Trends Microbiol.* 18, 189–194. doi:10.1016/j.tim.2010.02.006.

Rice, W. C. (1999). Specific primers for the detection of vip3A insecticidal gene within a *Bacillus thuringiensis* collection. *Lett. Appl. Microbiol.* 28, 378–382. doi:10.1046/j.1365-2672.1999.00536.x.

Del Rincón-Castro, M. C., Barajas-Huerta, J., and Ibarra, J. E. (1999). Antagonism between Cry1Ac1 and Cyt1A1 toxins of *Bacillus thuringiensis*. *Appl. Environ. Microbiol.* 65, 2049–2053.

Rose Gomes, M. (2012). Endophytic Colonization by Brazilian Strains of *Bacillus thuringiensis* on Cabbage Seedlings Grown *in Vitro*. *Bt Res.* 3, 11–19. doi:10.5376/bt.2012.03.0003.

Salazar-Marroquín, E. L., Galán-Wong, L. J., Moreno-Medina, V. R., Reyes-López, M. Á., and Pereyra-Alférez, B. (2016). Bacteriocins synthesized by *Bacillus thuringiensis*. *Rev. Med. Microbiol.* 27, 95–101. doi:10.1097/MRM.0000000000000076.

Saldanha, R., Mohr, G., Belfort, M., and Lambowith, M. A. (1993). Group I and group II introns. *FASEB J.* 7, 15–24.

Sampson, M. N., and Gooday, G. W. (1998). Involvement of chitinases of *Bacillus thuringiensis* during pathogenesis in insects. *Microbiology* 144, 2189–2194.

doi:10.1099/00221287-144-8-2189.

Sanahuja, G., Banakar, R., Twyman, R. M., Capell, T., and Christou, P. (2011). *Bacillus thuringiensis*: a century of research, development and commercial applications. *Plant Biotechnol. J.* 9, 283–300. doi:10.1111/j.1467-7652.2011.00595.x.

Sansinenea, E. (2013). *Bacillus thuringiensis* Biotechnology. 1st ed. , ed. Estibaliz Sansinenea Dordrecht: Springer Netherlands doi:10.1017/CBO9781107415324.004.

Sattar, S., and Maiti, M. K. (2011). Molecular characterization of a novel vegetative insecticidal protein from *Bacillus thuringiensis* effective against sap-sucking insect pest. *J. Microbiol. Biotechnol.* 21, 937–946. doi:10.4014/jmb.1105.05030.

Säwström, C., Lisle, J., Anesio, A. M., Priscu, J. C., and Laybourn-Parry, J. (2008). Bacteriophage in polar inland waters. *Extremophiles* 12, 167–175. doi:10.1007/s00792-007-0134-6.

Sayyed, A. H., Crickmore, N., and Wright, D. J. (2001). Cyt1Aa from *Bacillus thuringiensis* subsp. *israelensis* Is Toxic to the Diamondback Moth, *Plutella xylostella*, and Synergizes the Activity of Cry1Ac towards a Resistant Strain. *Appl. Environ. Microbiol.* 67, 5859–5861. doi:10.1128/AEM.67.12.5859-5861.2001.

Schmidt, T. R., Scott II, E. J., and Dyer, D. W. (2011). Whole-genome phylogenies of the family *Bacillaceae* and expansion of the sigma factor gene family in the *Bacillus cereus* species-group. *BMC Genomics* 12, 430. doi:10.1186/1471-2164-12-430.

Schnepf, E., Crickmore, N., Rie, J. Van, Lereclus, D., Baum, J., Feitelson, J., et al. (1998). *Bacillus thuringiensis* and Its Pesticidal Crystal Proteins. *Microbiol. Mol. Biol. Rev.* 62, 775–806.

Schnepf, H. E., Lee, S., Dojillo, J., Burmeister, P., Fencil, K., Morera, L., et al. (2005). Characterization of Cry34 / Cry35 Binary Insecticidal Proteins from Diverse *Bacillus thuringiensis* Strain Collections. *Appl. Environ. Microbiol.* 71, 1765–1774. doi:10.1128/AEM.71.4.1765.

Schulte, R. D., Makus, C., and Schulenburg, H. (2013). Host-parasite coevolution favours parasite genetic diversity and horizontal gene transfer. *J. Evol. Biol.* 26, 1836–1840. doi:10.1111/jeb.12174.

Sharma, S., Chatterjee, S., Datta, S., Prasad, R., Dubey, D., Rajesh, K. R., et al. (2016). Bacteriophages and its applications: an overview. *Folia Microbiol.* 1-39. doi:10.1007/s12223-016-0471-x.

Sharp, P. M., and Li, W. H. (1987). The rate of synonymous substitution in enterobacterial genes is inversely related to codon usage bias. *Mol. Biol. Evol.* 4, 222–230.

Shigetane, I. (1901). On a new type of severe flacherie (sotto disease). 114, 1–5.

Smith, R. a., and Couche, G. A. (1991). The phylloplane as a source of *Bacillus thuringiensis* variants. *Appl. Environ. Microbiol.* 57, 311–315.

Soberón, M., López-Díaz, J. A., and Bravo, A. (2013). Cyt toxins produced by *Bacillus thuringiensis*: A protein fold conserved in several pathogenic microorganisms. *Peptides* 41, 87–93. doi:10.1016/j.peptides.2012.05.023.

Spanier, G. J., and Cleary, P. P. (1980). Bacteriophage control of Antiphagocytic determinats in group A *Streptococci*. *J. exp. Med.* 152, 1393–1406.

Stiftung Weltbevölkerung: „Weltbevölkerungstag 2016: 7,44 Milliarden Menschen leben auf der Erde“ Available at: http://www.weltbevoelkerung.de/aktuelles/details/show/detail/News/weltbevoelkerungstag-2016-744-milliarden-menschen-leben-auf-der-erde.html [Accessed November 29, 2016].

Swiecicka, I., Fiedoruk, K., and Bednarz, G. (2002). The occurrence and properties of *Bacillus thuringiensis* isolated from free-living animals. *Lett. Appl. Microbiol.* 34, 194–198. doi:10.1046/j.1472-765x.2002.01070.x.

Thomas, C. M., and Nielsen, K. M. (2005). Mechanisms of, and barriers to, horizontal gene transfer between bacteria. *Nat.Rev.Microbiol.* 3, 711–721. doi:10.1038/nrmicro1234.

Thorsen, L., Hansen, B. M., Nielsen, K. F., Hendriksen, N. B., Phipps, R. K., and Budde, B. B. (2006). Characterization of emetic *Bacillus weihenstephanensis*, a new cereulide-producing bacterium. *Appl. Environ. Microbiol.* 72, 5118–5121. doi:10.1128/AEM.00170-06.

Twort, F. W. (1915). An Investigation on the Nature of Ultra-Microscopic Viruses. *Lancet* 186, 1241–1243. doi:http://dx.doi.org/10.1016/S0140-6736(01)20383-3.

Vachon, V., Laprade, R., and Schwartz, J.-L. (2012). Current models of the mode of action of *Bacillus thuringiensis* insecticidal crystal proteins: A critical review. *J. Invertebr. Pathol.* 111, 1–12. doi:10.1016/j.jip.2012.05.001.

van Valen, L. (1973). A new evolutionary law. *Evol. theory* 1, 1–30.

Vallet-Gely, I., Lemaitre, B., and Boccard, F. (2008). Bacterial strategies to overcome insect defences. *Nat Rev Microbiol* 6, 302–313. doi:10.1038/nrmicro1870.

Verplaetse, E., Slamti, L., Gohar, M., and Lereclus, D. (2015). Cell Differentiation in a *Bacillus thuringiensis* Population during Planktonic Growth, Biofilm Formation, and Host Infection. *MBio* 6, 1–10. doi:10.1128/mBio.00138-15.

Verplaetse, E., Slamti, L., Gohar, M., and Lereclus, D. (2016). Two distinct pathways lead *Bacillus thuringiensis* to commit to sporulation in biofilm. *Res. Microbiol.* S0923-2508, 30003–1. doi:10.1016/j.resmic.2016.03.006.

Vouk, V., and Klas, Z. (1931). Conditions influencing the growth of the insecticidal fungus *Metarrhizium anisopliae*. *Int Corn Borer Invest Sci Rept* 4, 24–45.

Walter, C., Fladung, M., and Boerjan, W. (2010). The 20-year environmental safety record of GM trees. *Nat. Biotechnol.* 28, 228–229. doi:10.1038/nbt0710-656.

Wang, A., and Ash, G. J. (2015). Whole Genome Phylogeny of *Bacillus* by Feature Frequency Profiles (FFP). *Sci. Rep.* 5, 13644. doi:10.1038/srep13644.

Wang, A., Pattemore, J., Ash, G., Williams, A., and Hane, J. (2013). Draft Genome Sequence of *Bacillus thuringiensis* Strain DAR 81934, Which Exhibits Molluscicidal Activity. *Genome Announc.* 1, e0017512. doi:10.1128/genomeA.00175-12.

Wang, X., Kim, Y., Ma, Q., Hong, S. H., Pokusaeva, K., Sturino, J. M., et al. (2010). Cryptic prophages help bacteria cope with adverse environments. *Nat. Commun.* 1, 147. doi:10.1038/ncomms1146.

Wei, J.-Z., Hale, K., Carta, L., Platzer, E., Wong, C., Fang, S.-C., et al. (2003). *Bacillus thuringiensis* crystal proteins that target nematodes. *Proc. Natl. Acad. Sci. U. S. A.* 100, 2760–2765. doi:10.1073/pnas.0538072100.

Werren, J. H. (2011). Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proc. Natl. Acad. Sci. U. S. A.* 108, 10863–10870. doi:10.1073/pnas.1102343108.

Whiteley, H. R., and Höfte, H. (1989). Insecticidal Crystal Proteins of *Bacilllus thuringiensis*. *Microbiol. Rev.* 53, 242–255.

Wichels, A., Biel, S. S., Gelderblom, H. R., Brinkhoff, T., Muyzer, G., and Schütt, C. (1998). Bacteriophage diversity in the North Sea. *Appl. Envir. Microbiol.* 64, 4128–4133.

Wilcks, A., Smidt, L., Bahl, M. I., Hansen, B. M., Andrup, L., Hendriksen, N. B., et al. (2008). Germination and conjugation of *Bacillus thuringiensis* subsp. *israelensis* in

the intestine of gnotobiotic rats. *J. Appl. Microbiol.* 104, 1252–1259. doi:10.1111/j.1365-2672.2007.03657.x.

Wilson, A. C., Ochman, H., and Prager, E. M. (1987). Molecular time scale for evolution. *Perspectives.* 3, 241–247.

Winstanley, C., Langille, M. G. I., Fothergill, J. L., Kukavica-Ibrulj, I., Paradis-Bleau, C., Sanschagrin, F., et al. (2009). Newly introduced genomic prophage islands are critical determinants of in vivo competitiveness in the liverpool epidemic strain of *Pseudomonas aeruginosa*. *Genome Res.* 19, 12–23. doi:10.1101/gr.086082.108.

Wolf, Y. I., Rogozin, I. B., Grishin, N. V., and Koonin, E. V. (2002). Genome trees and the tree of life. *Trends Genet.* 18, 472–479. doi:10.1016/S0168-9525(02)02744-0.

World Health Organization-WHO (1999). Microbial Pest Control Agent: *Bacillus thuringiensis*. Geneva.

Wu, J., Zhao, F., Bai, J., Deng, G., Qin, S., and Bao, Q. (2007). Evidence for Positive Darwinian Selection of Vip Gene in *Bacillus thuringiensis*. *J. Genet. Genomics* 34, 649–660. doi:10.1016/S1673-8527(07)60074-5.

Xiaoxia, L., Chen, L., Huang, Q., Zheng, J., Zhou, W., Peng, D., et al. (2012). Factor Virulence Bmp1 Functions as a Nematicidal *Bacillus thuringiensis* Metalloproteinase. *Appl. Environ. Microbiol.* 79, 460-468. doi:10.1128/AEM.02551-12.

Xu, C., Wang, B.-C., Yu, Z., and Sun, M. (2014). Structural Insights into *Bacillus thuringiensis* Cry, Cyt and Parasporin Toxins. *Toxins.* 6, 2732–2770. doi:10.3390/toxins6092732.

Ye, W., Zhu, L., Liu, Y., Crickmore, N., Peng, D., Ruan, L., et al. (2012). Mining new crystal protein genes from *Bacillus thuringiensis* on the basis of mixed plasmid-enriched genome sequencing and a computational pipeline. *Appl. Environ. Microbiol.* 78, 4795-4801. doi:10.1128/AEM.00340-12.

Yu, C. G., Mullins, M. A., Warren, G. W., Koziel, M. G., and Estruch, J. J. (1997). The *Bacillus thuringiensis* vegetative insecticidal protein Vip3A lyses midgut epithelium cells of susceptible insects. *Appl. Environ. Microbiol.* 63, 532–536.

CHAPTER II PUBLICATIONS

II.1 Complete genome sequence of the nematicidal *Bacillus thuringiensis* MYBT18246

Standard in Genomic Science

Complete Genome Sequence of the nematicidal *Bacillus thuringiensis* MYBT18246

Jacqueline Hollensteiner, Anja Poehlein, Cathrin Spröer, Boyke Bunk, Anna E. Sheppard, Philip Rosentstiel, Hinrich Schulenburg, Heiko Liesegang

Standard in Genomic Science (2017), 12:48 DOI 10.1186/s40793-017-0259-x

Authors' contributions

Performed experiments: CS, AS, **JH**

Genome Sequencing: CS, BB, PR

Genome Assembly: HL

Genome Annotation, analysis of data, comparative genomics: **JH**

Data submission **JH** and AP

Wrote the manuscript: **JH**, AP and HL

Supervision: AP

Conceived and designed the experiments: HS and HL

Hollensteiner *et al. Standards in Genomic Sciences* (2017) 12:48
DOI 10.1186/s40793-017-0259-x

Standards in Genomic Sciences

EXTENDED GENOME REPORT Open Access

Complete Genome sequence of the nematicidal *Bacillus thuringiensis* MYBT18246

Jacqueline Hollensteiner[1], Anja Poehlein[1], Cathrin Spröer[2], Boyke Bunk[2], Anna E. Sheppard[3,4], Philip Rosenstiel[5], Hinrich Schulenburg[4] and Heiko Liesegang[1*]

Abstract

Bacillus thuringiensis is a rod-shaped facultative anaerobic spore forming bacterium of the genus *Bacillus*. The defining feature of the species is the ability to produce parasporal crystal inclusion bodies, consisting of δ-endotoxins, encoded by *cry*-genes. Here we present the complete annotated genome sequence of the nematicidal *B. thuringiensis* strain MYBT18246. The genome comprises one 5,867,749 bp chromosome and 11 plasmids which vary in size from 6330 bp to 150,790 bp. The chromosome contains 6092 protein-coding and 150 RNA genes, including 36 rRNA genes. The plasmids encode 997 proteins and 4 t-RNA's. Analysis of the genome revealed a large number of mobile elements involved in genome plasticity including 11 plasmids and 16 chromosomal prophages. Three different nematicidal toxin genes were identified and classified according to the Cry toxin naming committee as *cry*13Aa2, *cry*13Ba1, and *cry*13Ab1. Strikingly, these genes are located on the chromosome in close proximity to three separate prophages. Moreover, four putative toxin genes of different toxin classes were identified on the plasmids p120510 (Vip-like toxin), p120416 (Cry-like toxin) and p109822 (two Bin-like toxins). A comparative genome analysis of *B. thuringiensis* MYBT18246 with three closely related *B. thuringiensis* strains enabled determination of the pan-genome of *B. thuringiensis* MYBT18246, revealing a large number of singletons, mostly represented by phage genes, morons and cryptic genes.

Keywords: *Bacillus thuringiensis*, *Bacillus cereus* sensu *lato*, Prophages, Parasporal crystal protein, Pan-Core-genome

Introduction

Bacillus thuringiensis is an ubiquitously distributed, rod-shaped, Gram-positive, spore forming, facultative anaerobic bacterium [1, 2]. *Bacillus thuringiensis* has been isolated from various ecological niches, including soil, aquatic habitats, phylloplane and insects [3–7]. The defining property of the species is the ability to produce parasporal protein crystals consisting of δ-endotoxins, which are predominantly encoded on plasmids [1, 8, 9]. These proteins are toxic towards a wide spectrum of invertebrates of the orders *Lepidoptera, Diptera, Coleoptera, Hymenoptera, Homoptera, Orthoptera, Mallophaga* and other species like *Gastropoda*, mites, protozoa and especially nematodes [7, 10–12]. In addition, *B. thuringiensis* produce additional toxins such as Cyt, Vip, and Sip toxins [13]. Cry toxins represent the largest group and can be subdivided into three different homology groups. In total, over 787 different Cry toxins have been identified, each exhibiting toxicity against a specific host organism [14]. It has been shown that *B. thuringiensis* strains can produce more than one Cry toxin resulting in a broad host range. As such, *B. thuringiensis* has been used widely as a biopesticide in agriculture for several decades [1, 2, 8, 13, 15, 16]. *Bacillus thuringiensis* is a member of the genus *Bacillus*, which are low GC-content, Gram-positive bacteria with a respiratory metabolism and the ability to form heat- and desiccation-resistant endospores [11, 17, 18]. Within this genus, *B. thuringiensis* is a member of the *Bacillus cereus* sensu *lato* species group which originally contained seven different species (*B. cereus*, *B. anthracis, B. thuringiensis, B. mycoides, B. pseudomycoides, B. weihenstephanensis, B. cytotoxicus* [17–25]). Historically, most pathogenic and phenotypic properties were used for strain classification. However, recent publications utilizing genomic criteria suggest that the species group should be extended by species *B. toyonensis* [26, 27]. Moreover, the three proposed species

* Correspondence: hlieseg@gwdg.de
[1]Department of Genomic and Applied Microbiology & Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, University of Göttingen, Göttingen, Germany
Full list of author information is available at the end of the article

"Bacillus gaemokensis"[28], *"Bacillus manliponensis"*[29] and *"Bacillus bingmayongensis"* [30] have been isolated and effectively published. However, these names had not yet appeared on a Validation List at the time of pulbication [31]. Due to the very close phylogenetic relationships, it has also been proposed to assign the eleven species to a single extended Bcsl species [32, 33]. The genome of Bcsl-members contains a highly conserved chromosome with regard to gene content, sequence similarity and genome synteny, while variation can be observed within mobile genomic elements such as prophages, insertion elements, transposons, and plasmids [34]. Due to the significance of Bcsl group members in human health, the food industry and agriculture, resolving the phylogeny is of great importance. Because of the highly conserved 16S rRNA-genes, the classical 16S phylogeny of Bcsl strains is inconclusive. Thus, a combination of 16S and a seven gene multi-locus sequence typing scheme have been used to establish taxonomic relationships within species of the Bcsl-group [35, 36]. Comparative genomics of the *cry*-gene loci has revealed remarkable proximity to elements of genome plasticity such as plasmids, transposons, insertion elements and prophages [2, 37–39]. The activity of these mobile elements has resulted in a magnitude of highly diverse plasmid sizes through rearrangements such as deletions and insertions, as well as migration of *cry*-genes into the bacterial chromosome [40]. The worldwide distribution of *B. thuringiensis* and its capacity to adapt to a diverse spectrum of invertebrate hosts is explained by the formation of spores and a remarkable variability in crystal protein families [13]. This toxin arsenal, especially the copy number of individual toxin genes, can be shaped by reciprocal co-adaptation with a nematode host, as previously demonstrated using controlled evolution experiments in the laboratory [41, 42]. The *B. thuringiensis* strain MYBT18246 described herein and its host *Caenorhabditis elegans* have been selected as a model system for such co-evolution experiments [41]. One aim of this sequencing project was to provide a high-quality reference genome sequence for the original *B. thuringiensis* MYBT18246 in order to obtain a detailed phylogeny and shed light on the evolution of this microparasite, with a particular focus on the presence of virulence factors, elements of genome plasticity and host adaptation factors. Here we present the genome of the nematicidal *B. thuringiensis* MYBT18246 and its comparative analysis to the three closest relatives identified by MLST phylogeny.

Organism information

Classification and features

Bacillus thuringiensis belongs to the genus *Bacillus* and has been isolated in the end of the nineteenth century [17, 20] and used as a biocontrol agent for several decades [7, 18, 21]. The strain *B. thuringiensis* MYBT18246 is a Gram-positive, rod-shaped and spore forming bacterium (Fig.1a), as most *B. thuringiensis* [7]. *Bacillus thuringiensis* MYBT18246 was isolated in the Schulenburg lab by AS from a mixture of genotypes present in the strain NRRL B-18246, originally provided by the Agricultural Research Service Patent Culture Collection (United States Department of Agriculture, Peoria, IL, USA) [43–45]. As a member of the species *B. thuringiensis*, *B. thuringiensis* MYBT18246 is facultative anaerobe, motile and is able to produce parasporal crystal toxins, which is the characteristic feature of this species [2]. Growth occurred at temperatures ranging from 10 to 48 °C and optimal growth was monitored at mesophil temperatures ranging from 28 to 37 °C [46]. The pH range of *B. thuringiensis* strains varies from pH 4.9 to 8.0, with the optimum documented as pH 7 [47, 48]. Strain *B. thuringiensis* MYBT18246 exhibits flat, opaque colonies with undulate, curled margins and produced crystals during the stationary phase (Fig. 1a-b). Characteristic features of *B. thuringiensis* MYBT18246 are listed in Table 1.

Extended feature descriptions

The cell size of *Bacillus thuringiensis* can vary from 0.5 × 1.2 μm - 2.5 × 10 μm [11]. Categorization into the

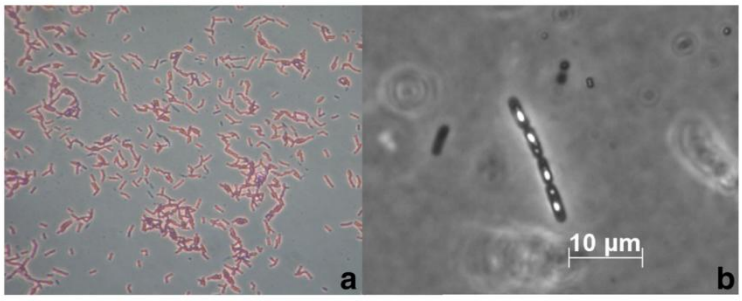

Fig. 1 Microscopic characteristics of *Bacillus thuringiensis* MYBT18246. **a** Light microscope analysis of Gram stained *B. thuringiensis* MYBT18246 cells (40×). **b** Phase contrast microscope analysis of sporulated and Cry-toxin producing cells of *B. thuringiensis* MYBT18246 (40×)

Table 1 Classification and general features of *B. thuringiensis* MYBT18246 [54]

MIGS ID	Property	Term	Evidence code[a]
	Classification	Domain *Bacteria*	TAS [86]
		Phylum *Firmicutes*	TAS [47]
		Class *Bacilli*	TAS [87, 88]
		Order *Bacillales*	TAS [18, 89]
		Family *Bacillaceae*	TAS [18, 90]
		Genus *Bacillus*	TAS [17, 18]
		Species *Bacillus thuringiensis*	TAS [46]
		Strain MYBT18246	IDA
	Gram stain	positive	IDA
	Cell shape	rod-shaped	IDA
	Motility	Motile	TAS [46]
	Sporulation	Spore-forming	IDA
	Temperature range	10–48 °C	TAS [46]
	Optimum temperature	28–37 °C	TAS [46]
	pH range; Optimum	4.9–8.0; 7.0	TAS [47, 48]
	Carbon source	Organic carbon source	NAS
MIGS-6	Habitat	Worldwide	TAS [7]
MIGS-6.3	Salinity	Salt tolerant	TAS [7]
MIGS-22	Oxygen requirement	Aerobic, facultative anaerobic	TAS [11]
MIGS-15	Biotic relationship	Free-living, microparasite of *C. elegans*	TAS [41]
MIGS-14	Pathogenicity	Nematode pathogen	TAS [41]
MIGS-4	Geographic location	not reported	
MIGS-5	Sample collection	not reported	
MIGS-4.1	Latitude	unreported	
MIGS-4.2	Longitude	unreported	
MIGS-4.4	Altitude	unreported	

[a]Evidence codes - IDA: Inferred from Direct Assay; TAS: Traceable Author Statement (i.e., a direct report exists in the literature); NAS: Non-traceable Author Statement (i.e., not directly observed for the living, isolated sample, but based on a generally accepted property for the species, or anecdotal evidence). These evidence codes are from the Gene Ontology project

group of Gram-positive organisms was confirmed by Gram staining, as shown in Fig. 1a. In Fig. 1b the production of Cry toxins can be observed. These toxins accumulate during the sporulation phase next to the endospore and build phase-bright inclusions [7]. *Bacillus thuringiensis* MYBT18246 exhibited 99% 16S rRNA sequence identity to other published Bcsl-members [49]. As a result of the high sequence similarity, a phylogenetic differentiation of *B. thuringiensis* MYBT18246 based on 16S phylogenetic differentiation of Bcsl group members is impossible (Fig. 2a). As an alternative, 23 *B. thuringiensis* strains, and a representative of each of the Bcsl group species were chosen for phylogenetic analysis using multilocus sequence typing as previously developed by Priest [36] (Fig. 2b). *Bacillus subtilis subsp. subtilis* str. 168 was selected as an outgroup to root the tree [17, 18]. The phylogenies were generated using the Neighbor-Joining method [50] and evolutionary distances were computed by the Maximum Composite Likelihood method [51]. In total, 217 MLST gene sequences were compared with 1000 bootstrap replicates. Phylogenetic analysis was conducted in MEGA7 [52]. All used reference sequences were retrieved from GenBank hosted at NCBI.

Genome sequencing information

Genome project history

Bacillus thuringiensis MYBT18246 was used in a co-evolution study with a *Caenorhabditis elegans* host. The original strain MYBT18246 was selected for sequencing in order to generate a reliable reference sequence for subsequent experiments [41, 42]. The genome sequence was analyzed to identify virulence factors and fitness factors contributing to the efficient infection of *C. elegans*. Additionally, the phylogenetic position of *B. thuringiensis* MYBT18246 in the Bcsl group was determined [41]. The complete genome sequence has been deposited in GenBank with the accession numbers (CP015350-CP015361) and in the integrated Microbial Genomes database with the Taxon ID 2671180122 [53]. A summary of the project information and its association with MIGS version 2.0 compliance [54] is shown in Table 2.

Growth conditions and genomic DNA preparation

Genomic DNA was isolated from *B. thuringiensis* MYBT18246 using the DNeasy blood and tissue kit (Qiagen, Hilden, Germany) for 454 pyrosequencing [55] and the Genomic-Tip 100/G Kit (Qiagen, Hilden, Germany) for Single Molecule real-time sequencing [56] according to the manufacturer's instructions. For SMRT-sequencing the procedure and Checklist: Greater than 10 kb Template Preparation Using AmPure PB Beads was used and blunt end ligation was applied overnight. Whole-genome sequencing was performed using a 454 GS-FLX system (Titanium GS70 chemistry; Roche Life Science, Mannheim, Germany) and on one SMRT Cell on the PacBio *RSII* system using P6-chemistry (Pacific Biosciences, Menlo Park, CA, USA).

Genome sequencing and assembly

A summary of the project information can be found in Table 2. 454-pyrosequencing was carried out at the Institute of Clinical Molecular Biology in Kiel, Germany and SMRT-sequencing at the DSMZ Braunschweig. First, approximately 331,000,454-reads with an average length of 600 bp were assembled using the Newbler 2.8 de novo assembler (Roche Diagnostics), resulting in 729

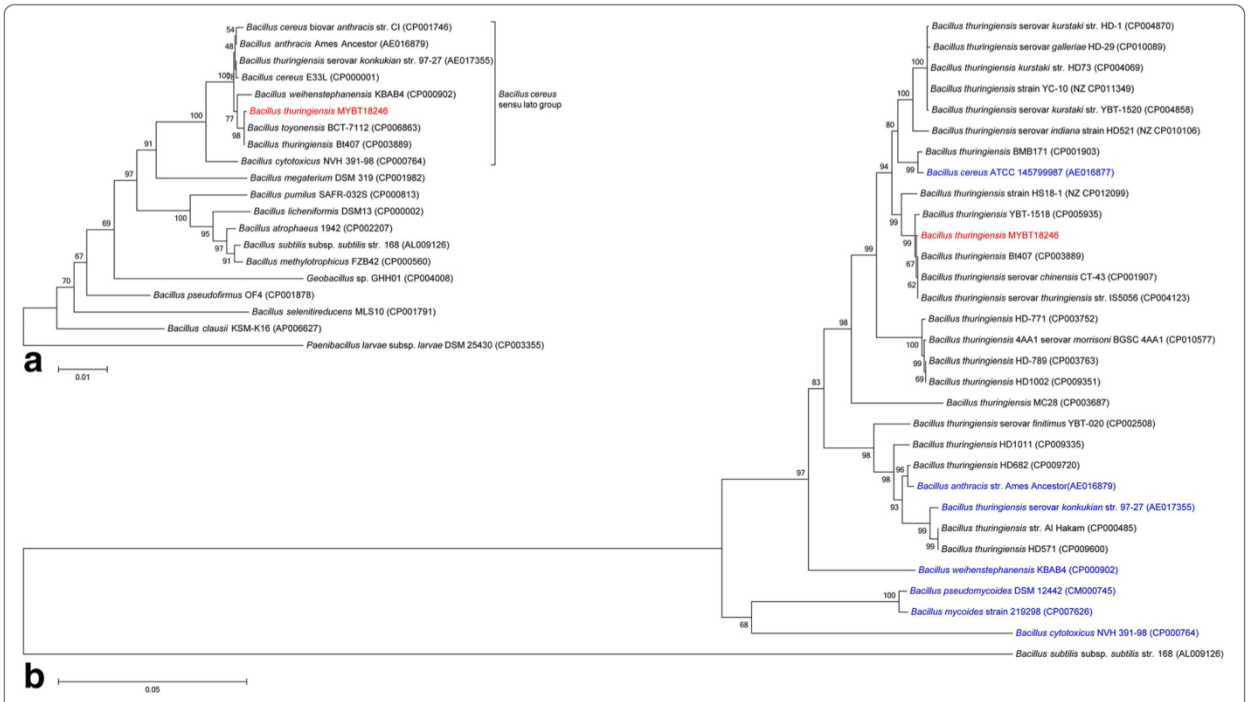

Fig. 2 Phylogenetic tree highlighting the taxonomic relation of *B. thuringiensis* MYBT18246 (red) based on **a**) 16rDNA amplicon within the *Bacillus* clade **b**) Multi-locus sequence typing within the *Bacillus cereus* sensu *lato* species group. GenBank accession numbers are given in parentheses. Comparison includes strains of the *Bacilli* clade or Bcsl group members (blue). *Paenibacillus larvae* subsp. *larvae* DSM 25430 or *Bacillus subtilis* subsp. *subtilis* str. 168 has been used as outlier to root the tree. Sequences were aligned using ClustalW 1.6 [91, 92]. The phylogenetic tree was constructed by using the Neighbor-Joining method [50] and evolutionary distances were computed by the Maximum Composite Likelihood method [51] within MEGA7.0 [52]. Numbers at the nodes are bootstrap values calculated from 1000 replicates

Table 2 Project information

MIGS ID	Property	Term
MIGS 31	Finishing quality	Complete
MIGS-28	Libraries used	Two genomic libraries: 454 pyrosequencing shotgun library, PacBio library
MIGS 29	Sequencing platforms	454 GS FLX system, PacBioRSII
MIGS 31.2	Fold coverage	18 × 454; 50 × PacBio
MIGS 30	Assemblers	Newbler 2.8; HGAP v2.3.0
MIGS 32	Gene calling method	Prodigal 2.6
	Locus Tag	BT246
	Genbank ID	CP015350-CP015361
	GenBank Date of Release	2016–07-15
	GOLD ID	Gp0020852
	BIOPROJECT	PRJNA290307
MIGS 13	Source Material Identifier	Department of Evolutionary Ecology and Genetics, CAU, Kiel
	Project relevance	Evolution

contigs with a coverage of 18 x. Repeats were resolved and gaps between contigs were closed using PCR with Sanger sequencing of the products with BigDye 3.0 chemistry and an ABI3730XL capillary sequencer (Applied Biosystems, Life Technology GmbH, Darmstadt, Germany). Manually editing in Gap4 (version 4.11) software of the Staden package [57] was performed to improve the sequence quality. For final gap closure PacBio sequencing was used. A total of 27,870 PacBio reads with a mean length of 14,053 bp were assembled using HGAP 2.0 [58], resulting in a coverage of 50 x, with further analysis using SMRT Portal (v2.3.0) [59]. Finally, both assemblies were combined, resulting in 12 contigs including a closed circular chromosome sequence of 5,867,749 bp. Eight additional contigs exhibited overlapping ends and were circularized to plasmid sequences ranging from 6.3 kb to 150 kb (Table 3). The assembly was checked for coverage drop downs and extremes of disparities including GC, AT, RY, and MK. Moreover, we determined the origin of replication of *B. thuringiensis* MYBT18246 by comparative analysis with OriC of eight other *B. thuringiensis* strains available in DoriC [60, 61]. These strains varied in chromosome size from 5.2 Mb to 5.8 Mb but all shared a similar GC-

Table 3 Summary of genome: one chromosome and 11 plasmids

Label	Size (Mb)	Topology	INSDC identifier	RefSeq ID
Chromosome	5.8	Circular	CP015350	NZ_CP015350.1
Plasmid 1	0.151	Circular	CP015351	NZ_CP015351.1
Plasmid 2	0.142	Circular	CP015352	NZ_CP015352.1
Plasmid 3	0.121	Circular	CP015353	NZ_CP015353.1
Plasmid 4	0.120	Circular	CP015354	NZ_CP015354.1
Plasmid 5	0.110	Circular	CP015355	NZ_CP015355.1
Plasmid 6	0.101	Circular	CP015356	NZ_CP015356.1
Plasmid 7	0.055	Linear	CP015357	NZ_CP015357.1
Plasmid 8	0.047	Circular	CP015358	NZ_CP015358.1
Plasmid 9	0.017	Linear	CP015359	NZ_CP015359.1
Plasmid 10	0.014	Linear	CP015360	NZ_CP015360.1
Plasmid 11	0.006	Circular	CP015361	NZ_CP015361.1

content of 35%. In total, including *B. thuringiensis* MYBT18246, two OriC regions were identified using the ORF-Finder [62]. One region was highly conserved with regard to OriC length (178/179 nt), OriC AT content (~0.69) and number of DnaA boxes (4). The second region varied in OriC length (564–767 nt) and OriC AT content (~0.67–0.7), but all had the same number of DnaA boxes (9). *B. thuringiensis* MYBT18246 showed the highest OriC similarities with both OriC regions of *B. thuringiensis* Bt407.

Genome annotation

Annotation was performed with Prokka v1.9 [63] using the manually curated *Bacillus thuringiensis* strain Bt407 [64] as a species reference and a comprehensive toxin protein database (including Cry, Cyt, Vip, and Sip toxins) as feature references. The Prokka pipeline was applied using prodigal for gene calling [65]. RNAmmer 1.2 [66] and Aragorn [67] were used for rRNA gene and t-RNA identification, respectively. Additionally, signal leader peptides were identified with SignalP 4.0 [68] and non-coding RNAs with an Infernal 1.1 search against the Rfam database [69]. Annotation of *cry* toxin genes were manually corrected and named according to the standards of the Cry toxin nomenclature by Crickmore [70]. Identified toxins were deposited at the *Bacillus thuringiensis* Toxin nomenclature database [14].

Genome properties

The genome of *B. thuringiensis* MYBT18246 consists of 12 replicons with a circular chromosome of 5,867,749 bp (Table 3). The GC content of the chromosome is 35% and the GC content of the plasmids ranges from 32 to 37%. The total number of protein coding genes is 7089 with 6092 genes on the chromosome and 997 genes on the plasmids. The genome harbors 12 rRNA clusters, 111 t-RNA genes, 5274 predicted protein-coding genes with assigned function and 1815 genes encoding proteins with unknown function (Table 4). All gene products have been assigned to COGs (Table 5). The genome sequence of *B. thuringiensis* MYBT18246 is available in GenBank (CP015350 for the chromosome and CP015351 - CP015361 for the plasmids).

Insights from the genome sequence

To investigate the phylogeny of *B. thuringiensis* MYBT18246 two approaches were used. First, nineteen *Bacillus* strains were chosen for 16S rRNA analysis within the *Bacillus* clade (Fig. 2a). The 16S rRNA phylogeny shows that *B. thuringiensis* MYBT18246 clusters with other Bcsl group members within the *Bacillus* clade. However, the low bootstrap values confirm the limitations of 16S rRNA as a discriminatory marker within the Bcsl species group. Second, we applied an MLST approach based on the scheme by Priest et al. [36]. This revealed that MYBT18246 clusters with the toxin cured *B. thuringiensis* Bt407, insecticidal *B. thuringiensis* serovar *chinensis* CT-43, and with the nematicidal *B. thuringiensis* YBT-1518 within the Bcsl phylogeny (Fig. 2b). Based on this phylogeny and the phenotypic defining feature of the *B. thuringiensis* species group (the ability to produce crystal toxins against invertebrates and nematodes), the strain *B. thuringiensis* MYBT18246 can be safely classified as nematicidal *B. thuringiensis*.

For a detailed analysis of encoded toxins in *B. thuringiensis* MYBT18246, we generated a local database consisting of all available Cry, Cyt, Vip and Sip protein sequences from UniProtKB [71] and GenBank [72]. The database was curated to generate a set of non-redundant reference toxins. In total, we identified three different

Table 4 Genome statistics

Attribute	Value	% of Total
Genome size (bp)	6,752,488	100
DNA coding (bp)	5,623,665	83.28
DNA G + C (bp)	2,389,665	35.39
DNA scaffolds	12	100
Total genes	7239	100
Protein coding genes	7089	97.9
RNA genes	151	2.09
Genes in internal clusters	2694	37.22
Genes with function prediction	5274	72.86
Genes assigned to COGs	4662	64.40
Genes with Pfam domains	5503	76.02
Genes with signal peptides	500	6.91
Genes with transmembrane helices	1863	25.74
CRISPR repeats	0	0

Table 5 Number of protein encoding genes associated with general COG functional categories

Code	Value	%	Description
J	226	3.19	Translation, ribosomal structure and biogenesis
A	0	0	RNA processing and modification
K	487	6.87	Transcription
L	625	8.82	Replication, recombination and repair
B	1	0.01	Chromatin structure and dynamics
D	59	0.83	Cell cycle control, Cell division, chromosome partitioning
V	141	1.99	Defense mechanisms
T	218	3.08	Signal transduction mechanisms
M	270	3.81	Cell wall/membrane biogenesis
N	64	0.90	Cell motility
U	79	1.12	Intracellular trafficking and secretion
O	121	1.71	Posttranslational modification, protein turnover, chaperones
C	214	3.02	Energy production and conversion
G	263	3.71	Carbohydrate transport and metabolism
E	420	5.93	Amino acid transport and metabolism
F	128	1.81	Nucleotide transport and metabolism
H	177	2.50	Coenzyme transport and metabolism
I	129	1.82	Lipid transport and metabolism
P	243	3.43	Inorganic ion transport and metabolism
Q	83	1.17	Secondary metabolites biosynthesis, transport and catabolism
R	653	9.22	General function prediction only
S	505	7.13	Function unknown
-	1978	27.9	Not in COGs

[a]The total number is based on the total number of protein coding genes in the genome

cry toxin genes in the *B. thuringiensis* MYBT18246 genome and classified them as *cry*13Aa2 (>95%), *cry*13Ba1 (<78%) and *cry*13Ab1 (<95%), based on the similarity scheme from the Cry-toxin naming committee by Crickmore [13, 70]. Notably, these *cry* toxin genes are encoded on the chromosome and not on extra-chromosomal elements as has been previously reported for the vast majority of *cry* toxin genes [7, 73, 74]. The toxin gene analysis revealed four additional putative toxin-like genes on plasmids with sequence similarity to *cry* genes and *vip* genes. A Pfam domain analysis using InterPro [75] revealed a p120510 encoded putative Vip--like toxin, a p120416 encoded putative Cry-like toxin and two p109822 encoded putative Bin-like toxins with potential for future studies.

Additionally, the *B. thuringiensis* MYBT18246 chromosome was screened for prophage regions by using the Phage Search Tool with default parameters. PHAST identifies prophage regions based on key genes from a reference database and defines the boundaries using a genomic composition-based algorithm. For a more detailed description see [76]. A total of 16 putative prophage loci were identified in the chromosome, including three that were associated with the previously identified chromosomally encoded *cry* toxin genes. As shown in Fig. 3, the *cry* toxins (displayed in red, track 4) are located in close proximity to identified prophage regions (displayed in blue, track 3). Furthermore, all *B. thuringiensis* MYBT18246 extra-chromosomal elements were also screened for prophages to check whether we could identify phages that reside in a linear or circular state in the host, as has been reported in 2013 by Fortier et al. [77]. Apparently, intact phage regions were identified according to the PHAST score system on p150790, p120416, p109822, p101287 and p46701.

The finding of prophage associated *cry* genes in strain MYBT18246 indicates that phages may serve as vectors for the transmission of virulence factors within the species *B. thuringiensis.* This resembles the previously described lysogenic conversion of pathogens by phages [78], supporting the idea that phages may represent a driving force for the distribution of fitness factors as well as virulence factors [78–80]. The finding that toxins, which are generally specific for a certain type of host organism, are located within a mobile genomic element in the chromosome of this bacterium, suggests that phages of strain MYBT18246 may contribute to adaptation to different hosts [81–83].

Extended insights

Based on the proximity within the tree (Fig. 2b), the genomes of *B. thuringiensis* Bt407, *B. thuringiensis* serovar *chinensis* CT-43 and *B. thuringiensis* YBT-1518 were identified as closest relatives and selected for an in depth comparative analysis. Shared gene contents were determined, visualized and compared, with a focus on known virulence factors such as *cry* toxins and pathogenic driving forces such as phages. The analysis revealed unique as well as shared gene contents for each strain (Fig. 3). In Fig. 3 the outer rings represent the genes on the leading and lagging strand with COG classification. The inner rings (track 5–7) illustrate the orthologous genes of *B. thuringiensis* YBT-1518, *B. thuringiensis* CT-43, *B. thuringiensis* Bt407 in red (high similarity) to light yellow (low similarity), and white (no similarity). The

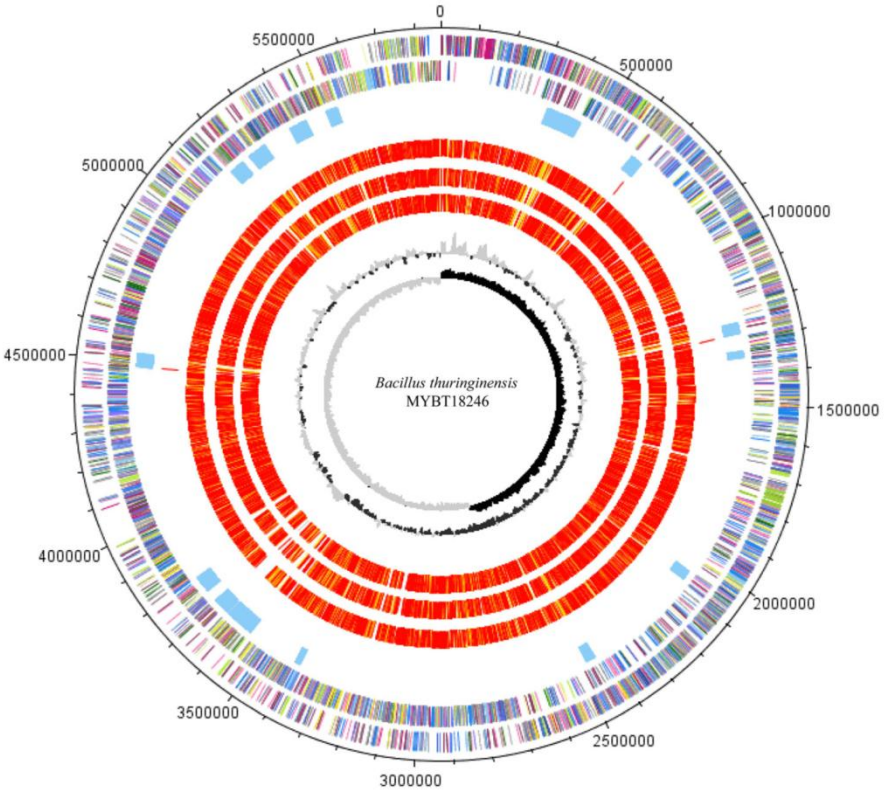

Fig. 3 Circular visualization of the genome comparison of *B. thuringiensis* MYBT18246 with 3 other sequenced *B. thuringiensis* strains. The tracks from the outside represent: (track 1–2) Genes encoded by the leading and lagging strand of *B. thuringiensis* MYBT18246 marked in COG colors [93]; (track 3) putative prophage regions, identified with PHAST in blue [76], (track 4) identified *cry* toxin genes in red; (track 5–7) orthologs for the genomes of *B. thuringiensis* YBT-1518 (CP005935.1), *B. thuringiensis* CT-43 (CP001907.1), *B. thuringiensis* Bt407 (CP003889.1) illustrated in red to light yellow, singletons in grey (grey: $<1e^{-20}$; light yellow: $1e^{-21}$–$1e^{-50}$; gold: $1e^{-51}$–$1e^{-90}$; light orange: $1e^{-91}$–$1e^{-100}$; orange: $1e^{-101}$–$1e^{-120}$; red: $>1e^{-121}$ (track 7) % GC plot (track 8), GC skew [(GC)/(G + C)]. Visualization was done with DNAPlotter [94]

circular representation of the chromosome comparison revealed that prophages are a major source of regional differences between the strains (Fig. 3). Additionally, the pan-genome of *B. thuringiensis* MYBT18246 compared to the three closest relatives was determined (Fig. 4). Orthologous genes between all four organisms were identified by comparing the whole genomes using Proteinortho [84] with a similarity cutoff of 50% and an E-value of $1e^{-10}$. Gbk-files were downloaded from NCBI and the protein sequences were extracted using cds_extractor v0.7.1 [85]. Detected paralogous genes are displayed in the Venn diagram in Fig. 4. All four strains share a core genome of 4298 genes. This is equivalent to 67% of each genome. *Bacillus thuringiensis* MYBT18246 shares 4 additional genes exclusively with *B. thuringiensis* Bt407, 17 genes with *B. thuringiensis* serovar *chinensis* CT-43 and 327 genes with *B. thuringiensis* YBT-1518. *Bacillus thuringiensis* serovar *chinensis* CT-43 and *B. thuringiensis* Bt407 share 398 orthologous genes. Notably, the genome of *B. thuringiensis* MYBT18246 contains 1242 orphan genes and thus two to threefold more singletons than the compared genomes. This result confirms the high degree of conservation of the four *Bacillus thuringiensis* strains (Fig. 2a and b) and it also refines the phylogenetic relationship of the strains to each other based on non-orthologous regions. Singletons are located on the chromosome as well as on extrachromosomal elements. The density of singletons is higher (2.5 fold) on the plasmids. Notably, all major chromosomal differences can be attributed to prophage regions. All gene products were assigned to COG categories and investigated for PFAM domains and Signal peptides (Table 6). In detail, those genes code for: (i) phage proteins, (ii) morons (virulence factors), (iii) a vast majority of proteins with cryptic function. This is supported by Fig. 3 which clearly shows that the regions of differences (track 5–7) directly correspond to the regions of identified phages (track 3). Moreover, the identified *cry* toxins (track 4) are adjacent to identified prophage regions and could be suggested as morons. Additionally, the singletons were screened for further virulence factors and genes encoding type-IV secretion system, C5-

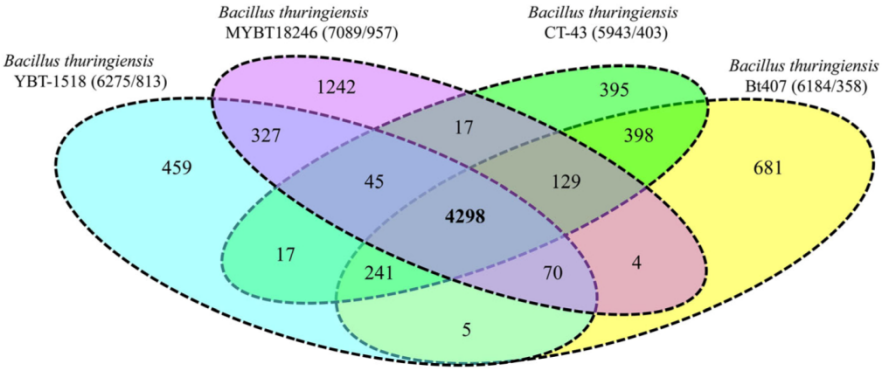

Fig. 4 Venn diagram of the genome comparison of *B. thuringiensis* MYBT18246 with other *B. thuringiensis* strains. Venn diagram displays the orthologous genes between *B. thuringiensis* MYBT18246 (CP015350-CP015361), *B. thuringiensis* YBT-1518 (CP005935-CP002486), *B. thuringiensis* serovar *chinensis* CT-43 (CP001907-CP001917) and *B. thuringiensis* Bt407 (CP003889-CP003898). Ortholog detection was performed with Proteinortho [84] including protein blast with a similarity cut-off of (50%) and an E-value of $1e^{-10}$. The total number of genes and paralogs are depicted under the corresponding species name. Open reading frames that were classified as pseudogenes were not included in this analysis

methyltransferase, type-restriction enzymes, sporulation, resistance and genes involved in genetic competence were identified. In particular, the finding of restriction-modification systems indicates a protection mechanism against other phages and plasmids and thus forms a putative barrier against further genomic modification.

Table 6 General genome features of *B. thuringiensis* MYBT18246 and close relatives

Genome features	Genome name			
	B. thuringiensis MYBT18246[a]	*B. thuringiensis* 407[b]	*B. thuringiensis* YBT-1518[c]	*B. thuringiensis* CT-43[d]
Sequencing status	Finished	Finished	Finished	Finished
Genome size (Mbp)	6.75	6.13	6.67	6.15
DNA coding (bp)	5,623,665	5,133,026	5,421,574	5,079,667
GC (%)	35.4	35.02	35.29	35.12
DNA scaffolds	12	10	7	11
Total gene count	7239	6442	6738	6252
Protein coding genes (%)	97.9	95.9	98.0	95.1
RNA genes	151	180	139	124
Genes in internal clusters	2694	489	370	334
Genes with function prediction	5274	4615	5193	4211
Genes assigned to COGs	4662	3634	3746	3505
Genes with Pfam domains	5503	4991	5333	4809
Genes with signal peptides	500	447	471	418
Genes with transmembrane helices	1863	1750	1854	1698
CRISPR repeats	0	2	0	2

Accession numbers: [a]CP015350, [b]CP003889, [c]CP005935, [d]CP001907

Conclusion

In this work we present the whole-genome sequence of *B. thuringiensis* MYBT18246 and its specific genome features. The genome includes three nematicidal *cry*13 gene variants located on the chromosome, which were named according to sequence similarity as stated by the Cry Toxin Nomenclature Committee, as *cry*13Aa2, *cry*13Ba1, and *cry*13Ab1. Four additional putative toxin genes were identified with low sequence similarity to other known toxins on plasmids: p120510 (Vip-like toxin), p120416 (Cry-like toxin) and p109822 (two Bin-like toxins). These toxins contained complete toxin domains, yet the activity against potential hosts should be elucidated in future studies. The genome comprises a large number of mobile elements involved in genome plasticity including eleven plasmids and sixteen chromosomal prophages. Both plasmids and prophages are important HGT elements indicating that they are an important driving force for the evolution of pathogens. The most striking finding is the close proximity of the chromosomal nematicidal *cry* toxin genes to three distinct prophages indicating a contribution of phages in defining the host range of this strain. *B. thuringiensis* MYBT18246 may show potential as a biocontrol agent against nematodes which should be addressed in future experiments.

Abbreviation

B: *Bacillus*; *B. thuringiensis*: *Bacillus thuringiensis*; Bcsl: *Bacillus cereus* sensu *lato*; Cry: Crystal; Cyt: Cytolytic; IMG: Integrated Microbial Genomes; MLST: Multi-locus sequence typing; PHAST: Phage Search Tool; Sip: Secreted insecticidal protein; SMRT: Single molecule real-time; Vip: Vegetative insectidal protein

Acknowledgements

We thank Sascha Dietrich and Andreas Leimbach for bioinformatic support. We as well thank Nicole Heyer and Simone Severitt for excellent technical assistance.

Funding
We thank the German Research Foundation (DFG-SPP1399, Grant LI 1690/2–1 to HL; SCHU 1415/9 to HS, and RO 2994/3 to PR) for financial support. We acknowledge support by the German Research Foundation and the Open Access Publication Funds of the Göttingen University.

Authors' contributions
HL and HS designed the study. AP supervised the genome analysis. JH corrected the annotation, analyzed the genome with focus on comparative genomics. JH, AP and HL wrote the manuscript. BB, CS and PR performed the genome sequencing. HL performed the assembly and the annotation of the sequence data. JH performed the microscopy of strain MYBT18246. AS and CS isolated the genomic DNA. All authors read and approved the final manuscript.

Competing interest
The authors declare that they have no competing interests.

Author details
[1]Department of Genomic and Applied Microbiology & Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, University of Göttingen, Göttingen, Germany. [2]Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany. [3]Present address: Nuffield Department of Medicine, University of Oxford, Oxford, UK. [4]Department of Evolutionary Ecology and Genetics, Zoological Institute, Christian-Albrechts University of Kiel, Kiel, Germany. [5]Institute of Clinical Molecular Biology, Christian-Albrechts University of Kiel, Kiel, Germany.

Received: 2 August 2016 Accepted: 8 August 2017
Published online: 24 August 2017

References

1. Bechtel DB, Bulla LA, Kramer KJ, Bechtel DB, David- LI. Electron microscope study of Sporulation and Parasporal crystal formation in *Bacillus thuringiensis*. J Bacteriol. 1976;127:1472–81.
2. Ibrahim MA, Griko N, Junker M, Bulla LA. *Bacillus thuringiensis*: a genomics and proteomics perspective. Bioeng Bugs. 2010;1:31–50.
3. Carozzi NB, Kramer VC, Warren G, Evola S, Koziel MG. Prediction of insecticidal activity of *Bacillus thuringiensis* strains by polymerase chain reaction product profiles. Appl Environ Microbiol. 1991;57:3057–61.
4. Iriarte J, Porcar M, Lecadet M-M, Caballero P. Isolation and characterization of *Bacillus thuringiensis* strains from aquatic environments in Spain. Curr Microbiol. 2000;40:402–8.
5. Smith RA, Couche GA. The phylloplane as a source of *Bacillus thuringiensis* variants. Appl Environ Microbiol. 1991;57:311–5.
6. Burges D, Hurst JA. Ecology of *Bacillus thuringiensis* in storage moths. J Invertebr Pathol. 1977;30:131–9.
7. Schnepf E, Crickmore N, Van RJ, Lereclus D, Baum J, Feitelson J, et al. *Bacillus thuringiensis* and its Pesticidal crystal proteins. Microbiol Mol Biol Rev. 1998;62:775–806.
8. Sanahuja G, Banakar R, Twyman RM, Capell T, Christou P. *Bacillus thuringiensis*: a century of research, development and commercial applications. Plant Biotechnol J. 2011;9:283–300.
9. Vidal-Quist JC, Castañera P, González-Cabrera J. Diversity of *Bacillus thuringiensis* strains isolated from citrus orchards in Spain and evaluation of their insecticidal activity against Ceratitis Capitata. J Microbiol Biotechnol. 2009;19:749–59.
10. Feitelson J. The *Bacillus thuringiensis* family tree. In: Kim L, editor. Advanced engineered pesticides. 1st ed. New York: CRC Press; 1993. p. 63–71.
11. De Vos P, Garrity GM, Jones D, Krieg NR, Ludwig W, Rainey FA, et al. Bergey's manual of systematic bacteriology. In: Vos P, Garrity G, Jones D, Krieg NR, Ludwig W, Rainey FA, Schleifer K-H, Whitman W, editors. Volume 3: the Firmicutes. 2nd ed. New York: Springer; 2009. p. 1–1317.
12. Sabha ME-S, Adayel SA, El-Masry SA, Alazazy H. *Bacillus thuringiensis* (Bt) toxin for the control of citrus trees snails. Researcher. 2013;5:26–32.
13. Palma L, Muñoz D, Berry C, Murillo J, Caballero P. *Bacillus thuringiensis* toxins: an overview of their Biocidal activity. Toxins (Basel). 2014;6:3296–325. doi:10.3390/toxins6123296.
14. Crickmore N. *Bacillus thuringiensis* Toxin nomenclature. http://www.btnomenclature.info/. Accessed 12 Oct 2016.
15. Aronson AI. The two faces of *Bacillus thuringiensis*: insecticidal proteins and post-exponential survival. Mol Microbiol. 1993;7:489–96.
16. Sanchis V, Bourguet D. *Bacillus thuringiensis*: applications in agriculture and insect resistance management. A review Agron Sustain Dev. 2008;28:11–20.
17. Cohn F. Untersuchungen über Bakterien. Beiträge zur Biologie der Pflanzen. 1872;1:127–224.
18. Skerman V, McGowan V, Sneath P. Approved lists of bacterial names. Int J Syst Bacteriol. 1980;30:225–420.
19. Scott E, Dyer DW. Divergence of the SigB regulon and pathogenesis of the *Bacillus cereus* Sensu Lato group. BMC Genomics. 2012;13:564. doi:10.1186/1471-2164-13-564.
20. Frankland GC, Frankland PF. Studies on some New Micro-Organisms obtained from Air. Philosophical Transactions of the Royal Society of London. 1887. p. 257–87.
21. Ernst B. Ueber die Schlaffsucht der Mehlmottenraupe (Ephestia kuhniella) und ihren Erreger *Bacillus thuringiensis* n. sp. Zeitschrift für Angew Entomol. 1915;2:21–56.
22. Flügge C. Die Mikroorganismen. 2.Auflage. Leipzig, Germany: F.C.W. Vogel; 1886.
23. Nakamura LK. *Bacillus pseudomycoides* sp. nov. Int J Syst Bacteriol. 1998;48:1031–5. doi:10.1099/00207713-48-3-1031.
24. Lechner S, Mayr R, Francis KP, Prüss BM, Kaplan T, Wießner-Gunkel E, et al. *Bacillus weihenstephanensis* sp. nov. is a new psychrotolerant species of the *Bacillus cereus* group. Int J Syst Bacteriol. 1998;48:1373–82. doi:10.1099/00207713-48-4-1373.
25. Guinebretière MH, Auger S, Galleron N, Contzen M, de Sarrau B, de Buyser ML, et al. *Bacillus cytotoxicus* sp. nov. is a novel thermotolerant species of the *Bacillus cereus* group occasionally associated with food poisoning. Int J Syst Evol Microbiol. 2013;63:31–40.
26. Jiménez G, Urdiain M, Cifuentes A, López-López A, Blanch AR, Tamames J, et al. Description of *Bacillus toyonensis* sp. nov., a novel species of the *Bacillus cereus* group, and pairwise genome comparisons of the species of the group by means of ANI calculations. Syst Appl Microbiol. 2013;36:383–91. doi:10.1016/j.syapm.2013.04.008.
27. Oren A, Garrity GM. List of new names and new combinations previously effectively, but not validly, published. Int J Syst Evol Microbiol. 2014;64:1–5. doi:10.1099/ijsem.0.000737.
28. Jung MY, Paek WK, Park IS, Han JR, Sin Y, Paek J, et al. *Bacillus gaemokensis* sp. nov., isolated from foreshore tidal flat sediment from the Yellow Sea. J Microbiol. 2010;48:867–71.
29. Jung MY, Kim JS, Paek WK, Lim J, Lee H, Kim PI, et al. *Bacillus manliponensis* sp. nov., a new member of the *Bacillus cereus* group isolated from foreshore tidal flat sediment. J Microbiol. 2011;49:1027–32.
30. Liu B, Liu GH, Hu GP, Cetin S, Lin NQ, Tang JY, et al. *Bacillus bingmayongensis* sp. nov., isolated from the pit soil of emperor Qin's Terra-cotta warriors in China. Antonie Van Leeuwenhoek. 2014;105:501–10.
31. NL; Validation of the publication of new names and new combinations previously effectively published outside the IJSB. List No. 65. Int J Syst Bacteriol. 1998;48(Pt 2):627.
32. Helgason E, Okstad OA, Caugant DA, Johansen HA, Fouet A, Mock M, et al. *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*–one species on the basis of genetic evidence. Appl Environ Microbiol. 2000;66:2627–30.
33. Okinaka TR, Keim P, Okinaka RT, Keim P. The Phylogeny of *Bacillus cereus* sensu lato. Microbiol Spectr. 2016;4 doi:10.1128/microbiolspec.
34. Rasko DA, Altherr MR, Han CS, Ravel J. Genomics of the *Bacillus cereus* group of organisms. FEMS Microbiol Rev. 2005;29:303–29.
35. Schmidt TR, Scott EJ II, Dyer DW. Whole-genome phylogenies of the family *Bacillaceae* and expansion of the sigma factor gene family in the *Bacillus cereus* species-group. BMC Genomics. 2011;12:430. doi:10.1186/1471-2164-12-430.
36. Priest FG, Barker M, Baillie LWJ, Holmes EC, Maiden MCJ. Population structure and evolution of the *Bacillus cereus* group. Society. 2004;186:7959–70.
37. Yu H, Zhang J, Huang D, Gao J, Song F. Characterization of *Bacillus thuringiensis* strain Bt185 toxic to the Asian cockchafer: *Holotrichia parallela*. Curr Microbiol. 2006;53:13–7.
38. Ackermann H-WW, Azizbekyan RR, Bernier RL, Barjac H, Saindoux S, Valero JR, et al. Phage typing of *Bacillus subtilis* and *B. thuringiensis*. Res Microbiol. 1995;146:643–57.
39. Mahillon J, Rezsöhazy R, Hallet B, Delcour J. IS231 and other *Bacillus thuringiensis* transposable elements: a review. Genetica. 1994;93:13–26.
40. Kronstad JW, Whiteley HR. Inverted repeat sequences flank a *Bacillus thuringiensis* crystal protein gene. J Bacteriol. 1984;160:95–102.
41. Masri L, Branca A, Sheppard AE, Papkou A, Laehnemann D, Guenther PS, et al. Host–pathogen Coevolution: the selective advantage of *Bacillus*

thuringiensis virulence and its cry toxin genes. PLoS Biol. 2015;13:e1002169. doi:10.1371/journal.pbio.1002169.
42. Schulte RD, Makus C, Hasert B, Michiels NK, Schulenburg H. Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. Proc Natl Acad Sci U S A. 2010;107:7359–64. doi:10.1073/pnas.1003113107.
43. Payne J. Isolates of *Bacillus thuringiensis* that are active against nematodes. 1992.
44. Payne J, Cannon R, Bagley A. *Bacillus thuringiensis* isolates for controlling acarides. 1993.
45. Schnepf H, Schwab G, Payne J, Narva K, Foncerrada L. Nematicidal proteins. 2001.
46. Barjac H, Frachon E. Classification of *Bacillus thuringiensis* strains. Entomophaga. 1990;35:233–40. doi:10.1007/BF02374798.
47. Gibbons NE, Murray RGE. Proposals concerning the higher Taxa of bacteria. Int J Syst Bacteriol. 1978;28:1–6. doi:10.1099/00207713-28-1-1.
48. West AW, Burges HD, Dixon TJ, Wyborn CH. Survival of *Bacillus thuringiensis* and *Bacillus cereus* spore inocula in soil: effects of pH, moisture, nutrient availability and indigenous microorganisms. Soil Biol Biochem. 1985;17:657–65.
49. Ash C, Farrow JA, Dorsch M, Stackebrandt E, Collins MD. Comparative analysis of *Bacillus anthracis*, *Bacillus cereus*, and related species on the basis of reverse transcriptase sequencing of 16S rRNA. Int J Syst Bacteriol. 1991; 41(3):343–6.
50. Saitou N, Nei M. The neighbor-joining method: a new method for reconstructing Phylogenetic trees. Mol Biol Evol. 1987;4:406–25.
51. Tamura K, Nei M, Kumar S. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc Natl Acad Sci U S A. 2004;101:11030–5.
52. Kumar S, Stecher G, Tamura K. MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. Mol Biol Evol. 2016;7:1870–4. doi:10.1093/molbev/msw054.
53. Markowitz VM, Chen IMA, Palaniappan K, Chu K, Szeto E, Grechkin Y, et al. IMG: the integrated microbial genomes database and comparative analysis system. Nucleic Acids Res. 2012;40:115–22.
54. Field D, Garrity G, Gray T, Morrison N, Selengut J, Tatusova T, et al. The minimum information about a genome sequences (MIGS) specification. Nat Biotechnol. 2008;26:541–7. doi:10.1038/1360.
55. Gilles A, Meglécz E, Pech N, Ferreira S, Malausa T, Martin J-F. Accuracy and quality assessment of 454 GS-FLX titanium pyrosequencing. BMC Genomics. 2011;12:245. doi:10.1186/1471-2164-12-245.
56. Rhoads A, Au KF. PacBio sequencing and its applications. Genomics, Proteomics Bioinforma. 2015;13:278–89. doi:10.1016/j.gpb.2015.08.002.
57. Staden R, Beal KF, Bonfield JK. The Staden package, 1998. In: Misener S, Krawtz AS, editors. Methods in molecular biology. New York: Humana Press; 1999. p. 115–30.
58. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, et al. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. Nat Methods. 2013;10:563–9. doi:10.1038/nmeth.2474.
59. PacBio Software Downloads. 2016. http://www.pacb.com/support/software-downloads/. Accessed 12 Apr 2016.
60. Gao F, Luo H, Zhang C-T. DoriC 5.0: an updated database of oriC regions in both bacterial and archaeal genomes. Nucleic Acids Res. 2013;41 Database issue:D90–3. doi:10.1093/nar/gks990.
61. DoriC: an updated database of bacterial and archaeal replication origins. http://tubic.tju.edu.cn/doric/. Accessed 15 Dec 2016.
62. Gao F, Zhang C-T. Ori-finder: a web-based system for finding oriCs in unannotated bacterial genomes. BMC Bioinformatics. 2008;9:79.
63. Seemann T. Prokka: rapid prokaryotic genome annotation. Bioinformatics. 2014;30:2068–9.
64. Sheppard AE, Poehlein A, Rosenstiel P, Liesegang H, Schulenburg H. Complete genome sequence of *Bacillus thuringiensis* strain 407 cry. Genome Announc. 2013;1:158–12.
65. Hyatt D, Chen G-L, Locascio PF, Land ML, Larimer FW, Hauser LJ. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinformatics. 2010;11:119. doi:10.1186/1471-2105-11-119.
66. Lagesen K, Hallin P, Rødland EA, Stærfeldt HH, Rognes T, Ussery DW. RNAmmer: consistent and rapid annotation of ribosomal RNA genes. Nucleic Acids Res. 2007;35:3100–8.
67. Laslett D, Canback B. ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. Nucleic Acids Res. 2004;32:11–6.
68. Petersen TN, Brunak S, von Heijne G, Nielsen H. SignalP 4.0: discriminating signal peptides from transmembrane regions. Nat Methods. 2011;8:785–6. doi:10.1038/nmeth.1701.
69. Griffiths-Jones S, Bateman A, Marshall M, Khanna A, Eddy SR. Rfam: an RNA family database. Nucleic Acids Res. 2003;31:439–41.
70. Crickmore N, Zeigler DR, Feitelson J, Schnepf E, Van Rie J, Lereclus D, et al. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. Microbiol Mol Biol Rev. 1998;62:807–13.
71. Bateman A, Martin MJ, O'Donovan C, Magrane M, Apweiler R, Alpi E, et al. UniProt: a hub for protein information. Nucleic Acids Res. 2015;43:D204–12.
72. Benson DA, Cavanaugh M, Clark K, Karsch-Mizrachi I, Lipman DJ, Ostell J, et al. GenBank. Nucleic Acids Res. 2013;41:36–42.
73. Carlson CR, Kolstø A-BB. A complete physical map of a *Bacillus thuringiensis* chromosome. J Bacteriol. 1993;175:1053–60.
74. Wang A, Pattemore J, Ash G, Williams A, Hane J. Draft genome sequence of Bacillus thuringiensis strain DAR 81934, which Exhibits Molluscicidal Activity. Genome Announc. 2013;1:e0017512. doi:10.1128/genomeA.00175-12.
75. Mitchell A, Chang HY, Daugherty L, Fraser M, Hunter S, Lopez R, et al. The InterPro protein families database: the classification resource after 15 years. Nucleic Acids Res. 2015;43:D213–21.
76. Zhou Y, Liang Y, Lynch KH, Dennis JJ, Wishart DS. PHAST: a fast phage search tool. Nucleic Acids Res. 2011;39(SUPPL. 2):347–52.
77. Fortier L-C, Sekulovic O. Importance of prophages to evolution and virulence of bacterial pathogens. Virulence. 2013;4:354–65. doi:10.4161/viru.24498.
78. Brüssow H, Canchaya C, Hardt W-D, Bru H. Phages and the evolution of bacterial pathogens: from genomic rearrangements to Lysogenic conversion. Microbiol Mol Biol Rev. 2004;68:560–602.
79. Barksdale L, Arden SB. Persisting bacteriophage infections, lysogeny, and phage conversions. Annu Rev Microbiol. 1974;28:265–99.
80. Freeman VJ. Studies on the virulence of bacteriophage-infected strains of *Corynebacterium diphtheriae*. J Bacteriol. 1951;61:675–88.
81. De Maagd RA, Bravo A, Crickmore N. How *Bacillus thuringiensis* has evolved specific toxins to colonize the insect world. Trends Genet. 2001;17:193–9.
82. Agaisse H, Lereclus D. How does *Bacillus thuringiensis* produce so much insecticidal crystal protein? J Bacteriol. 1995;177:6027–32.
83. Ben-Dov E. *Bacillus thuringiensis* subsp. *israelensis* and its Dipteran-specific toxins. Toxins (Basel). 2014;6:1222–43.
84. Lechner M, Findeiss S, Steiner L, Marz M, Stadler PF, Prohaska SJ. Proteinortho: detection of (co-)orthologs in large-scale analysis. BMC Bioinformatics. 2011;12:124. doi:10.1186/1471-2105-12-124.
85. Leimbach A. bac-genomics-scripts: Bovine *E. coli* mastitis comparative genomics editio. 2016. https://github.com/aleimba/bac-genomics-scripts/tree/master/cds_extractor ,Accessed 14 Mar 2016.
86. Woese CR, Kandler O, Wheelis ML. Towards a natural system of organisms: proposal for the domains Archaea, bacteria, and Eucarya. Proc Natl Acad Sci U S A. 1990;87:4576–9.
87. Euzéby J. List of new names and new combinations previously effectively, but not validly, published. Int J Syst Evol Microbiol. 2010;60:469–72.
88. Ludwig W, Schleifer K-H, Whitman WB. Class I. *Bacillis* class nov. Bergey's Man Syst Bacteriol. 2009;3:19–20.
89. Prévot AR. Dictionnaire des Bactéries Pathogènes In: Hauderoy P, Ehringer G, Guillot G, Magrou. J., Prévot AR, Rosset D, Urbain A (eds), Dictionnaire des Bactéries Pathogènes, Second Edition, Masson et Cie, Paris, 1953, p. 1–692.
90. Fischer A. Untersuchungen über bakterien. Jahrbücher für Wissenschaftliche Botanik. 1895;27:1–163.
91. Thompson JD, Higgins DG, Gibson TJ. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 1994;22:4673–80.
92. Ash C, Priest FG, Collins MD. Molecular identification of rRNA group 3 bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test. Proposal for the creation of a new genus *Paenibacillus*. Antonie Van Leeuwenhoek. 64:253–60. http://www.ncbi.nlm.nih.gov/pubmed/8085788. Accessed 6 Jun 2017.
93. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, et al. The COG database: an updated version includes eukaryotes. BMC Bioinformatics. 2003;4:41. doi:10.1186/1471-2105-4-41.
94. Carver T, Thomson N, Bleasby A, Berriman M, Parkhill J. DNAPlotter: circular and linear interactive genome visualization. Bioinformatics. 2009;25:119–20.

II.2 Complete Genome sequence of the nematicidal *Bacillus thuringiensis* MYBT18247

Journal of Biotechnology

Complete genome sequence of the nematicidal *Bacillus thuringiensis* MYBT18247

Jacqueline Hollensteiner, Anja Poehlein, Cathrin Spröer, Boyke Bunk, Anna E. Sheppard, Philip Rosenstiel, Hinrich Schulenburg, Heiko Liesegang

Journal of Biotechnology 260 (2017) 48–52

Author contributions

Performed the experiments: AS and CS

Genome sequencing: CS, BB and PR

Genome assemblies: HL

Genome annotation, correction and data submission **JH**

Data analysis: **JH**

Supervision: AP

Writing the manuscript: **JH**, AP and HL

Study design: HS and HL

Journal of Biotechnology 260 (2017) 48–52

Contents lists available at ScienceDirect

Journal of Biotechnology

journal homepage: www.elsevier.com/locate/jbiotec

Short Genome Communications

Complete genome sequence of the nematicidal *Bacillus thuringiensis* MYBT18247

Jacqueline Hollensteiner[a], Anja Poehlein[a], Cathrin Spröer[b], Boyke Bunk[b], Anna E. Sheppard[c,1], Philip Rosenstiel[d], Hinrich Schulenburg[c], Heiko Liesegang[a,*]

[a] Department of Genomic and Applied Microbiology & Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, University of Goettingen, Germany
[b] Leibniz Institute DSMZ-German Collection of Microorganisms and Cell Cultures, Braunschweig, Germany
[c] Department of Evolutionary Ecology and Genetics, Zoological Institute, Christian-Albrechts University of Kiel, Kiel, Germany
[d] Institute of Clinical Molecular Biology, Christian-Albrechts University of Kiel, Kiel, Germany

ARTICLE INFO

Keywords:
Bacillus thuringiensis
Genome sequence
Nematicidal
Crystal toxins
Evolution

ABSTRACT

The Gram-positive spore forming bacterium *Bacillus thuringiensis* MYBT18247 encodes three *cry* toxin genes, (*cry*6Ba2, *cry*6Ba3 and *cry*21-like) which are active against nematodes. For a better understanding of the evolution of virulence and *cry* toxins, we present here the complete genome sequence of *Bacillus thuringiensis* MYBT18247. Various additional virulence factors such as bacteriocins, proteases and hemolysins were identified. In addition, the methylome and the metabolic potential of the strain were analyzed and the strain phylogenetically classified.

Bacillus thuringiensis is a ubiquitous, Gram-positive, spore-forming, bacterium (Schnepf et al., 1998). Strains of the species are used as a biopesticide because of their ability to produce parasporal protein crystals (Bechtel and Bulla, 1976; Ibrahim et al., 2010). These protein crystals consist of δ-endotoxins which are active against a broad spectrum of invertebrates including species of the orders Lepidoptera, Diptera, Coleoptera, Hymenoptera, Homoptera, Orthoptera, Mallophaga as well as mites, protozoa and nematodes (Feitelson, 1993; Schnepf et al., 1998). Here we present the annotated genome sequence of *B. thuringiensis* MYBT18247 that was isolated and used for single as well as multiple infection co-evolution experiments within the nematode *Caenorhabditis elegans* (Masri et al., 2015; Schulte et al., 2010).

Genomic DNA was isolated using the DNeasy blood and tissue kit (Qiagen, Hilden, Germany) and the Genomic-Tip 100/G Kit (Qiagen, Hilden, Germany). For 454 pyrosequencing the genomic DNA was sheared (~700 bp), end repaired and universal barcoded sequencing adaptors were ligated (Rapid-GSFLX Titanium, Roche 454, Branford CT). The library preparation was done with the GS Titanium Sequencing Kit XLR70t (Roche 454, Branford CT). Illumina sequencing DNA libraries were generated with the Nextera XT DNA Library Prep Kit (Illumina, San Diego, USA). For SMRT-sequencing (Pacific Biosciences, Menlo Park, USA) the C2/P4 chemistry was applied for three SMRT-cells and C4/P6 chemistry were used for two additional SMRT-cells, respectively. Whole-genome sequencing was performed using the 454 GS-FLX instrument, the Genome analyzer IIx (Illumina, San Diego, USA) and the PacBio *RSII* system (Pacific Biosciences, Menlo Park, USA). The 454 shotgun sequencing produced 335,141 single-end reads with an average read length of 420 bp. The Newbler 2.8 *de novo* assembler (Roche Diagnostics) assembled the reads into 411 contigs with a coverage of 18 x. A hybrid assembly was performed with Mira 4.0.3 (http://mira-assembler.sourceforge.net/docs/DefinitiveGuideToMIRA.html) by using 4,000,000 (112 bp) paired-end Illumina reads and 30,952 PacBio reads (C2-chemistry) with an average read length of 5053 bp. An HGAP 2.3.0 assembly using 67,045 PacBio-reads (P6-chemistry) with a mean length of 13,802 bp resulted in an average coverage of 124.46 x. The assemblies were manually combined and contradictions were resolved by Sanger sequencing using BigDye 3.0 chemistry and an ABI3730XL capillary sequencer (Applied Biosystems, Life Technology GmbH, Darmstadt, Germany). All sequence positions were manually checked using Gap4 (v4.11) of the Staden package (Staden et al., 1999) to ensure the sequence quality.

Annotation was conducted with Prokka v.1.9 (Seemann, 2014). The initial annotation was performed using *Bacillus thuringiensis* Bt407 as a high quality species reference (Sheppard et al., 2013) and a comprehensive toxin protein database (including all Cry, Cyt, Vip, Sip toxins) as a feature reference set. The annotations of detected *cry* toxin genes were manually corrected and confirmed by Crickmore and deposited at

* Corresponding author at: Department of Genomic and Applied Microbiology & Göttingen Genomics Laboratory, Institute of Microbiology and Genetics, University of Goettingen, Grisebachstr. 8, 37077 Göttingen, Germany.
E-mail address: hlieseg@gwdg.de (H. Liesegang).
[1] Present address: Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom.

http://dx.doi.org/10.1016/j.jbiotec.2017.09.003
Received 12 January 2017; Received in revised form 8 September 2017; Accepted 8 September 2017
Available online 09 September 2017
0168-1656/

Table 1
Genome statistics of Bacillus thuringiensis MYBT18247.

Genome Feature	Value
Genome size (bp)	6,138,199
DNA coding (bp)	5,143,010
DNA G + C (bp)	2,159,739
DNA scaffolds	7
Total genes	6366
Protein coding genes	6210
RNA genes	156
rRNA genes	42
Genes in internal clusters	536
Genes with function prediction	4767
Genes assigned to COG	1443
Genes assigned to Mobilome	268
Genes with Pfam domains	5054
Genes with signal peptides	424
Genes with transmembrane helices	1692
CRISPR repeats	–

http://www.btnomenclature.info/(Crickmore et al., 1998).

The genome size of *B. thuringiensis* MYBT18247 is 6,138,199 bp with an average GC-content of 35% (Table 1). The data set consists of a chromosome of 5.6 Mbp and six plasmids which vary in size from 12.5 kb to 175 kb. One plasmid contig of 130 kb could not be circularized due to large repetitive regions at both contig ends. The chromosome encodes 6210 protein-coding and 156 RNA genes, including 14 rRNA clusters (5S, 16S, 23S rRNA). The plasmids encode 556 protein coding genes. There were 3727 genes assigned to COG database (Table 2) and 1114 genes were identified in MetaCyc metabolic pathways (Caspi et al., 2016). Three nematicidal *cry* toxin genes have been identified on plasmid p174778 (*cry*6Ba2; BTI247_60340), p113275 (*cry*6Ba3, BTI247_62380) and p15092 (*cry*21-like, BTI247_64000) (Table 3).

B. thuringiensis MYBT18247 comprises a plethora of virulence factors which might all contribute to pathogenicity. The chromosome encodes two chitinases, which enable the degradation of chitin in the midgut peritrophic membrane of many insects and have been identified as insect virulence factors (Sampson and Gooday, 1998). Additionally,

Table 2
COG categories of B. thuringiensis MYBT18247.

Letter	Name	Count	Percent [%]
E	Amino acid transport and metabolism	365	8.83
G	Carbohydrate transport and metabolism	227	5.49
D	Cell cycle control, cell division, chromosome partitioning	57	1.38
N	Cell motility	52	1.26
M	Cell wall/membrane/envelope biogenesis	216	5.22
B	Chromatin structure and dynamics	1	0.02
H	Coenzyme transport and metabolism	213	5.15
Z	Cytoskeleton	2	0.05
V	Defense mechanisms	111	2.69
C	Energy production and conversion	194	4.69
W	Extracellular structures	4	0.1
S	Function unknown	295	7.14
R	General function prediction only	368	8.9
P	Inorganic ion transport and metabolism	226	5.47
U	Intracellular trafficking, secretion, and vesicular transport	32	0.77
F	Lipid transport and metabolism	138	3.34
F	Nucleotide transport and metabolism	126	3.05
O	Posttranslational modification, protein turnover, chaperones	156	3.77
L	Replication, recombination and repair	190	4.6
Q	Secondary metabolites biosynthesis, transport and catabolism	89	2.15
T	Signal transduction mechanisms	185	4.48
K	Transcription	364	8.81
J	Translation, ribosomal structure and biogenesis	255	6.17
	Not in COG	2907	42.92

Table 3
Predicted virulence factors of B. thuringiensis MYBT18247.

Virulence factor	No. of proteins	Locus_Taq	Location
Cry toxins	3		
pesticidal crystal protein Cry6Ba2[a]		BTI247_60340	p174778
pesticidal crystal protein Cry6Ba3[a]		BTI247_62380	p113275
putative pesticidal crystal protein Cry21-like[a]		BTI247_64000	p15092
Bacteriocins	4		
Bacteriocin > 10 kD (bacteriocinIII)[b]			chromosome
LAP (linear azol(ine) containing peptides) [b]			chromosome
Head to tail cyclized bacteriocin IId [b]			p174778
Head to tail cyclized bacteriocin IId [b]			p81952
Proteases			
Bacillolysin	4		
*npr*1 [c]		BTI247_07440	chromosome
*npr*2 [c]		BTI247_24310	chromosome
nprM [c]		BTI247_29880	chromosome
*npr*3 [c]		BTI247_36120	chromosome
Collagenases	6		
*colA*1 [c]		BTI247_06950	chromosome
*colA*2_1 [c]		BTI247_32780	chromosome
*colA*4 [c]		BTI247_36420	chromosome
*colA*3 [c]		BTI247_37900	chromosome
*colA*5 [c]		BTI247_46580	chromosome
*colA*2_2 [c]		BTI247_59250	chromosome
Immune inhibitor A	5		
*ina*1_1 [c]		BTI247_08150	chromosome
*ina*1_2 [c]		BTI247_08170	chromosome
*ina*2_1 [c]		BTI247_15830	chromosome
*ina*3 [c]		BTI247_32580	chromosome
*ina*2_2 [c]		BTI247_33630	chromosome
Phopholipases	4		
cerA [c]		BTI247_08230	chromosome
phospholipase, patatin family		BTI247_22620	chromosome
plcA [c]		BTI247_37570	chromosome
ytpA [c]		BTI247_50630	chromosome
Chitinases	2		
chiA [c]		BTI247_05680	chromosome
chiD [c]		BTI247_39810	chromosome
Camelysins	4		
*calY*1_1 [c]		BTI247_07820	chromosome
*calY*1_2 [c]		BTI247_15800	chromosome
*calY*1_3 [c]		BTI247_59500	p174778
*calY*1_4 [c]		BTI247_63500	p81952
N-acyl homoserine lactonase	1		
AiiA [c]		BTI247_36860	chromosome
Enterotoxins			
Hemolysins			
Hemolysin BL			
*hblA*1_1 [c]		BTI247_21380	chromosome
*hblA*2 [c]		BTI247_21390	chromosome
hblD [c]		BTI247_26830	chromosome
*hblA*3 [c]		BTI247_26840	chromosome
*hblA*4_1 [c]		BTI247_26850	chromosome
*hblA*4_2 [c]		BTI247_26880	chromosome
*hblA*1_2 [c]		BTI247_60570	p130548
*hblA*1_3 [c]		BTI247_61620	p130548
Hemolysin II	3		
*hly*_1 [c]		BTI247_37800	chromosome
*hly*_2 [c]		BTI247_59850	p174778
*hly*_3 [c]		BTI247_63990	p15092
Hemolysin III	2		
*hlyIII*_1 [c]		BTI247_24750	chromosome

(continued on next page)

Table 3 *(continued)*

Virulence factor	No. of proteins	Locus_Taq	Location
hlyIII_2 [c]		BTI247_58090	chromosome
Hemolysin A	1		
tlyA [c]		BTI247_44490	chromosome
Gamma hemolysin	1		
hlgB [c]		BTI247_13860	Chromosome
Non-hemolysins	4		
nhe_1 [c]		BTI247_21370	chromosome
nhe_2 [c]		BTI247_60560p	p130548
nhe_3 [c]		BTI247_61630	p130548
Alveolysin [c]		BTI247_54170	chromosome

[a] all protein sequences of the genome were scanned with generic HMM models constructed from representative sequences of the known Cry/Cyt/VIP-toxins extracted from SwissProt and the Bt toxin database (http://www.btnomenclature.info/). Identified sequences were characterized by a procedure described by the international committee for cry-Toxins (Crickmore et al., 1998), verified by Crickmore and deposited at the BT toxin database.

[b] AntiSMASH3.0.5 (Weber et al., 2015) and BAGEL3 (van Heel et al., 2013) were used for identification.

[c] Predicted by Prodigal (Hyatt et al., 2010) and annotated by BLAST+ (Camacho et al., 2009) comparisons to reference-proteins from the Prokka genome annotation pipeline (Seemann, 2014). Annotation was primarily derived from *Bacillus thuringiensis* Bt407 as a high quality species reference (Sheppard et al., 2013).

proteases including four camelysins, six collagenases, and four phospholipases were identified and may play an important role in activating protoxins (Nisnevitch et al., 2010), destruction of the intestine of *C. elegans* (Peng et al., 2016) and in hydrolyzing phospholipids of host cell membranes (Hergenrother and Martin, 1997). Furthermore, five immune inhibitor A metalloproteases, four bacillolysins and an N-acyl homoserine lactonase were detected, suspicious for boosting the nematicidal or insecticidal activity, as well as overcoming the host immune system by cleaving host antibacterial peptides (Fedhila et al., 2002; Park et al., 2008; Raymond et al., 2010). The host gut microbiome is part of the immune response and bacterial secondary metabolites such as bacteriocins and microcins are able to suppress other pathogenic microbes. In the *B. thuringiensis* MYBT18247 genome eleven biosynthetic clusters including bacteriocins, siderophores, NRPS (non-ribosomal peptide synthetases) and terpene clusters were detected with AntiSMASH3.0.5 (Weber et al., 2015) (Table 4). Four potential bacteriocins were further validated with BAGEL3 (van Heel et al., 2013) (Table 3). In the chromosome a putative bacteriocin class III and a LAP (linear azol(ine) containing peptide) were classified. In the plasmids p174778 and p81952 one head to tail cyclized bacteriocin class IId (PF09221) were identified. The cyclic bacteriocins are active against a broad spectrum of Gram-positive and Gram-negative bacteria (Finn et al., 2016). *B. thuringiensis* MYBT18247 encodes as well genes for hemolysis and non-haemolytic enterotoxins (Table 3), which have an important effect during the infection of hosts (Argolo-Filho and Loguercio, 2014; Kim et al., 2015). However, a test on Columbia agar revealed a hemolytic negative phenotype. This might correlate to the observation that the corresponding genome locus misses the hemolysin BL lytic component L2, one essential part of the tripartite toxin. The plasmids encode a number of genes dedicated to genome fluidity such as transposases, insertion sequences, transcriptional regulators and recombinases.

The metabolic versatility of *B. thuringiensis* MYBT18247 has been evaluated using BIOLOG Phenotypic Microarrays (PM1-PM2). The strain utilized substrates assigned feeding into glycolysis, citrate cycle and pentose phosphate pathway. It metabolizes various carbon sources including sugars, sugar alcohols and sugar acids which have been found in fruits, fungi, insect compartments or plant compartments. Apparently, the strain is a generalist which is able to adapt to various ecological niches.

The methylome analysis of *B. thuringiensis* MYBT18247 was performed using the SMRT Portal v2.3.0 analysis platform. Four methylated N^6-methyladenine (m6A) motifs were observed. The non-palindromic motif (CRTANNNNNNNRTTNC/GNAAYNNNNNNNTAYG) was found to be methylated in more than 65% of all instances. Additionally, three N4-methylcytosin (m4C) motifs were detected with a methylation grade of 6–21% and in four putative motifs the identification of the methylation was not possible due to insufficient coverage. The SMRT data DNA methyltransferase recognition motifs are deposited in REBASE at http://rebase.neb.com/rebase/private/pacbio_Liesegang23.html (Roberts et al., 2015).

Phylogenetic classification of *B. thuringiensis* MYBT18247 within the *Bacillus cereus sensu lato* group, was performed by multi-locus-sequence typing (MLST) according to Priest et al. (Priest et al., 2004) (Fig. 1). The strain clusters within a subgroup, comprising *B. thuringiensis* MYBT18246, *B. thuringiensis* YBT-1518, *B. thuringiensis* Bt407, and *B. thuringiensis* serovar *chinensis* CT-43. Strikingly, all members of the cluster except *B. thuringiensis* 407, which is an artificially cured laboratory strain, encode at least one nematicidal or insecticidal *cry* toxin gene.

Summarized, *B. thuringiensis* MYBT18247 comprises a variety of promising genes, including the rare *cry*6Ba which show high potential for the development of new biotechnological relevant nematicidal and insecticidal control agents.

Nucleotide sequence accession numbers

The whole genome sequence has been deposited at the DDBJ/EMBL/GenBank with the accession numbers CP015250.1-CP015256.1. The strain is available from DSMZ (Braunschweig, Germany) under accession 104068.

Acknowledgements

We thank Dr. Andrea Thürmer, Nicole Heyer and Simone Severitt for sequencing support. This project was funded by the German

Table 4
Secondary metabolite clusters identified in B. thuringiensis MYBT18247 with antiSMASH3.0.

Cluster	Type	Location	Start	Stop	Similarity to known cluster	MIBiG BGC-ID
1	Other	chromosome	496,393	539,974	–	–
2	Bacteriocin	chromosome	1,444,119	1,457,984	–	–
3	Siderophore	chromosome	2,120,707	2,134,414	Petrobactin biosynthetic gene cluster (83% of genes show similarity)	BGC0000942_c1
4	NRPS	chromosome	2,433,678	2,483,403	Bacillibactin biosynthetic gene cluster (46% of genes show similarity)	BGC0000309_c1
5	Bacteriocin	chromosome	2,666,725	2,677,045	–	–
6	NRPS	chromosome	2,736,144	2,783,151	–	–
7	Bacteriocin	chromosome	2,800,158	2,810,424	–	–
8	NRPS	chromosome	3,472,892	3,532,849	–	–
9	Terpene	chromosome	3,659,931	3,681,784	Molybdenum cofactor biosynthetic gene cluster (11% of genes show similarity)	BGC0000916_c1
10	Bacteriocin	p174778	132,176	142,397	–	–
11	Bacteriocin	p81952	56,165	66,488	–	–

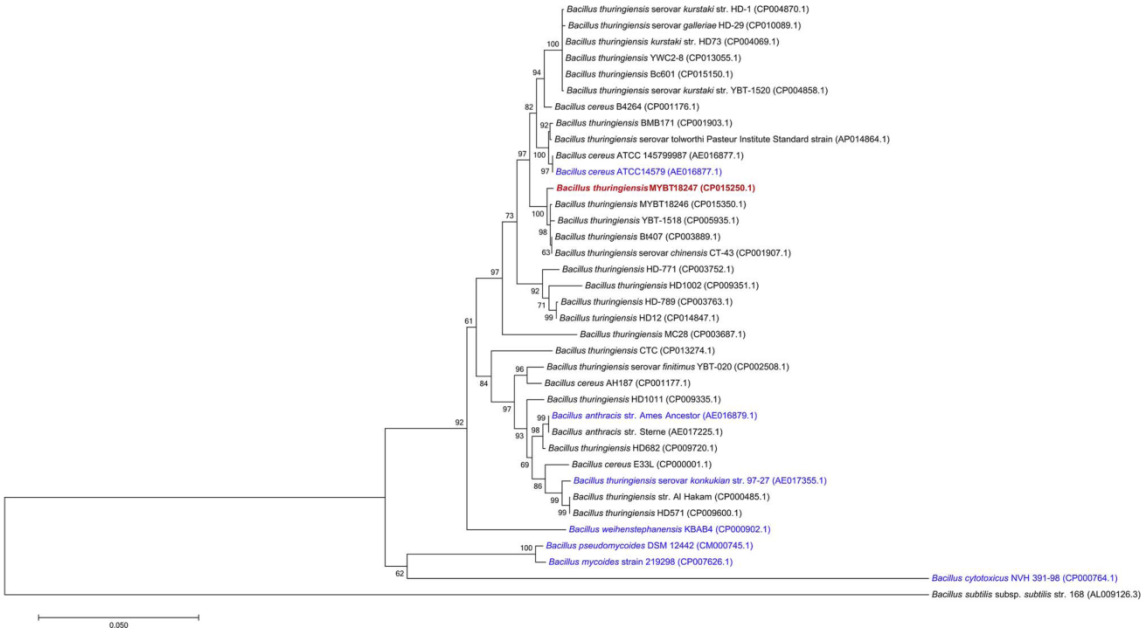

Fig. 1. Phylogenetic tree highlighting the taxonomic relation of *B. thuringiensis* MYBT18247 (red) based on Multi-Locus-Sequence-Typing within the *Bacillus cereus* sensu lato species group. GenBank accession numbers are given in parentheses. Comparison includes representative strains of Bcsl group members (blue). *Bacillus subtilis* subsp. *subtilis* str. 168 has been used as outlier to root the tree. Sequences were aligned using ClustalW 1.6 (Thompson et al., 1994). The phylogenetic tree was constructed by using the Neighbor-Joining method (Saitou and Nei, 1987) and evolutionary distances were computed by the Maximum Composite Likelihood method (Tamura et al., 2004) within MEGA7.0 (Kumar et al., 2016). Numbers at the nodes are bootstrap values calculated from 1000 replicates. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article).

Research Foundation (DFG-SPP1399, Grant LI 1690/2-1 to HL; SCHU 1415/9 to HS, and RO 2994/3 to PR). We acknowledge support by the German Research Foundation and the Open Access Publication Funds of the University of Goettingen.

References

Argolo-Filho, R.C., Loguercio, L.L., 2014. *Bacillus thuringiensis* is an environmental pathogen and host-specificity has developed as an adaptation to human-generated ecological niches. Insects 5, 62–91. http://dx.doi.org/10.3390/insects5010062.

Bechtel, D.B., Bulla, L.A., 1976. Electron Microscope Study of Sporulation and Parasporal Crystal Formation in *Bacillus thuringiensis*. J. Bacteriol. 127, 1472–1481.

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., Madden, T.L., 2009. BLAST+: architecture and applications. BMC Bioinf. 10, 1–9. http://dx.doi.org/10.1186/1471-2105-10-421.

Caspi, R., Billington, R., Ferrer, L., Foerster, H., Fulcher, C.A., Keseler, I.M., Kothari, A., Krummenacker, M., Latendresse, M., Mueller, L.A., Ong, Q., Paley, S., Subhraveti, P., Weaver, D.S., Karp, P.D., 2016. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res. 44, D471–D480. http://dx.doi.org/10.1093/nar/gkv1164.

Crickmore, N., Zeigler, D.R., Feitelson, J., Schnepf, E., Van Rie, J., Lereclus, D., Baum, J., Dean, D.H., 1998. Revision of the nomenclature for the *Bacillus thuringiensis* pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. 62, 807–813.

Fedhila, S., Nel, P., Lereclus, D., 2002. The InhA2 metalloprotease of *Bacillus thuringiensis* strain 407 is required for pathogenicity in insects infected via the oral route. J. Bacteriol. 184, 3296–3304. http://dx.doi.org/10.1128/JB.184.12.3296-3304.2002.

Feitelson, J., 1993. The *Bacillus thuringiensis* family tree. In: Kim, L. (Ed.), Advanced Engineered Pesticides. CRC Press, New York, pp. 63–71.

Finn, R.D., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Mistry, J., Mitchell, A.L., Potter, S.C., Punta, M., Qureshi, M., Sangrador-Vegas, A., Salazar, G.A., Tate, J., Bateman, A., 2016. The Pfam protein families database: towards a more sustainable future. Nucleic Acids Res. 44, D279–D285. http://dx.doi.org/10.1093/nar/gkv1344.

Hergenrother, P.J., Martin, S.F., 1997. Determination of the kinetic parameters for phospholipase C (*Bacillus cereus*) on different phospholipid substrates using a chromogenic assay based on the quantitation of inorganic phosphate. Anal. Biochem. 251, 45–49. http://dx.doi.org/10.1006/abio.1997.2251.

Hyatt, D., Chen, G.-L., Locascio, P.F., Land, M.L., Larimer, F.W., Hauser, L.J., 2010. Prodigal: prokaryotic gene recognition and translation initiation site identification. BMC Bioinf. 11, 119. http://dx.doi.org/10.1186/1471-2105-11-119.

Ibrahim, M.A., Griko, N., Junker, M., Bulla, L.A., 2010. *Bacillus thuringiensis*: A genomics and proteomics perspective. Bioeng. Bugs 1, 31–50. http://dx.doi.org/10.4161/bbug.1.1.10519.

Kim, M.-J., Han, J.-K., Park, J.-S., Lee, J.-S., Lee, S.-H., Cho, J.-I., Kim, K.-S., 2015. Various enterotoxin and other virulence factor genes widespread among *Bacillus cereus* and *Bacillus thuringiensis* strains. J. Microbiol. Biotechnol. 25, 872–879. http://dx.doi.org/10.4014/jmb.1502.02003.

Kumar, S., Stecher, G., Tamura, K., 2016. MEGA7: molecular evolutionary genetics analysis version 7. 0 for bigger datasets. Mol. Biol. Evol. 7, 1870–1874. http://dx.doi.org/10.1093/molbev/msw054.

Masri, L., Branca, A., Sheppard, A.E., Papkou, A., Laehnemann, D., Guenther, P.S., Prahl, S., Saebelfeld, M., Hollensteiner, J., Liesegang, H., Brzuszkiewicz, E., Daniel, R., Michiels, N.K., Schulte, R.D., Kurtz, J., Rosenstiel, P., Telschow, A., Bornberg-Bauer, E., Schulenburg, H., 2015. Host-Pathogen coevolution: the selective advantage of bacillus thuringiensis virulence and its cry toxin genes. PLoS Biol. 13, e1002169. http://dx.doi.org/10.1371/journal.pbio.1002169.

Nisnevitch, M., Sigawi, S., Cahan, R., Nitzan, Y., 2010. Isolation, characterization and biological role of camelysin from *Bacillus thuringiensis* subsp. israelensis. Curr. Microbiol. 61, 176–183. http://dx.doi.org/10.1007/s00284-010-9593-6.

Park, S.J., Park, S.Y., Ryu, C.M., Park, S.H., Lee, J.K., 2008. The role of AiiA quorum quenching enzyme from *Bacillus thuringensis*, on the rhizosphere competence. J. Microbiol. Biotechnol. 18, 1518–1521.

Peng, D., Lin, J., Huang, Q., Zheng, W., Liu, G., Zheng, J., Zhu, L., Sun, M., 2016. A novel metalloproteinase virulence factor is involved in *Bacillus thuringiensis* pathogenesis in nematodes and insects. Environ. Microbiol. 18, 846–862. http://dx.doi.org/10.1111/1462-2920.13069.

Priest, F.G., Barker, M., Baillie, L.W.J., Holmes, E.C., Maiden, M.C.J., 2004. Population structure and evolution of the *Bacillus cereus* group. Society 186, 7959–7970. http://dx.doi.org/10.1128/JB.186.23.7959.

Raymond, B., Johnston, P.R., Nielsen-LeRoux, C., Lereclus, D., Crickmore, N., 2010. *Bacillus thuringiensis*: An impotent pathogen? Trends Microbiol. 18, 189–194. http://dx.doi.org/10.1016/j.tim.2010.02.006.

Roberts, R.J., Vincze, T., Posfai, J., Macelis, D., 2015. REBASE-a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Res. 43, D298–D299. http://dx.doi.org/10.1093/nar/gku1046.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sampson, M.N., Gooday, G.W., 1998. Involvement of chitinases of *Bacillus thuringiensis* during pathogenesis in insects. Microbiology 144, 2189–2194. http://dx.doi.org/10.1099/00221287-144-8-2189.

Schnepf, E., Crickmore, N., Rie, J., Van Lereclus, D., Baum, J., Feitelson, J., Zeigler, D.R., Dean, D.H., 1998. *Bacillus thuringiensis* and its pesticidal crystal proteins. Microbiol. Mol. Biol. Rev. 62, 775–806.

Schulte, R.D., Makus, C., Hasert, B., Michiels, N.K., Schulenburg, H., 2010. Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. Proc. Natl. Acad. Sci. U. S. A. 107, 7359–7364. http://dx.doi.org/10.1073/pnas.1003113107.

Seemann, T., 2014. Prokka: rapid prokaryotic genome annotation. Bioinformatics 30, 2068–2069. http://dx.doi.org/10.1093/bioinformatics/btu153.

Sheppard, A.E., Poehlein, A., Rosenstiel, P., Liesegang, H., Schulenburg, H., 2013. Complete genome sequence of *Bacillus thuringiensis* strain 407 cry. Genome Announc. 1, 158–212. http://dx.doi.org/10.1128/genomeA.00158-12.

Staden, R., Beal, K.F., Bonfield, J.K., 1999. The staden package, 1998. In: Misener, S., Krawtz, A.S. (Eds.), Methods in Molecular Biology. Humana Press, New York, pp. 115–130. http://dx.doi.org/10.1385/1-59259-192-2:115.

Tamura, K., Nei, M., Kumar, S., 2004. Prospects for inferring very large phylogenies by using the neighbor-joining method. Proc. Natl. Acad. Sci. U. S. A. 101, 11030–11035. http://dx.doi.org/10.1073/pnas.0404206101.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H.U., Bruccoleri, R., Lee, S.Y., Fischbach, M.A., Müller, R., Wohlleben, W., Breitling, R., Takano, E., Medema, M.H., 2015. antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. Nucleic Acids Res. 43, W237–W243. http://dx.doi.org/10.1093/nar/gkv437.

van Heel, A.J., de Jong, A., Montalbán-López, M., Kok, J., Kuipers, O.P., 2013. BAGEL3: automated identification of genes encoding bacteriocins and (non-)bactericidal posttranslationally modified peptides. Nucleic Acid Res. 41, W448–W453. http://dx.doi.org/10.1093/nar/gkt391.

II.3 Genome-based identification of active prophage regions by next generation sequencing in *Bacillus licheniformis* DSM13

PLoS One

Genome-Based Identification of Active Prophage Regions by Next Generation Sequencing in *Bacillus licheniformis* DSM13

Robert Hertel, David Pintor Rodriguez, **Jacqueline Hollensteiner**, Sascha Dietrich, Andreas Leimbach, Michael Hoppert, Heiko Liesegang and Sonja Volland

PLoS One (2015), 10(3):e0120759

Author contributions

Performed the experiments: RH, DPR, **JH**, MH, SV

Analyzed the data: RH, SD, HL, SV

Contributed reagents/material/analysis tools: SD and AL

Wrote the paper: RH, HL, SV

Conceived and designed the experiments: RH and HL

PLOS | ONE

RESEARCH ARTICLE

Genome-Based Identification of Active Prophage Regions by Next Generation Sequencing in *Bacillus licheniformis* DSM13

Robert Hertel, David Pintor Rodríguez, Jacqueline Hollensteiner, Sascha Dietrich, Andreas Leimbach, Michael Hoppert, Heiko Liesegang*, Sonja Volland*

Georg-August University Göttingen, Institute of Microbiology and Genetics, Department of Genomic and Applied Microbiology, Göttingen, Germany

* hlieseg@gwdg.de (HL); svollan@gwdg.de (SV)

OPEN ACCESS

Citation: Hertel R, Rodríguez DP, Hollensteiner J, Dietrich S, Leimbach A, Hoppert M, et al. (2015) Genome-Based Identification of Active Prophage Regions by Next Generation Sequencing in *Bacillus licheniformis* DSM13. PLoS ONE 10(3): e0120759. doi:10.1371/journal.pone.0120759

Academic Editor: Raymond Schuch, ContraFect Corporation, UNITED STATES

Received: September 30, 2014

Accepted: January 26, 2015

Published: March 26, 2015

Data Availability Statement: The sequencing raw data are available at the NIH short read archive (SRA) under the study number SRP035551.

Funding: This work was funded by the German Ministry of Education and Research (BMBF, http://www.bmbf.de/) and publication was supported by the Open Access Publication Funds of the Göttingen University. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Abstract

Prophages are viruses, which have integrated their genomes into the genome of a bacterial host. The status of the prophage genome can vary from fully intact with the potential to form infective particles to a remnant state where only a few phage genes persist. Prophages have impact on the properties of their host and are therefore of great interest for genomic research and strain design. Here we present a genome- and next generation sequencing (NGS)-based approach for identification and activity evaluation of prophage regions. Seven prophage or prophage-like regions were identified in the genome of *Bacillus licheniformis* DSM13. Six of these regions show similarity to members of the *Siphoviridae* phage family. The remaining region encodes the *B. licheniformis* orthologue of the PBSX prophage from *Bacillus subtilis*. Analysis of isolated phage particles (induced by mitomycin C) from the wild-type strain and prophage deletion mutant strains revealed activity of the prophage regions BLi_Pp2 (PBSX-like), BLi_Pp3 and BLi_Pp6. In contrast to BLi_Pp2 and BLi_Pp3, neither phage DNA nor phage particles of BLi_Pp6 could be visualized. However, the ability of prophage BLi_Pp6 to generate particles could be confirmed by sequencing of particle-protected DNA mapping to prophage locus BLi_Pp6. The introduced NGS-based approach allows the investigation of prophage regions and their ability to form particles. Our results show that this approach increases the sensitivity of prophage activity analysis and can complement more conventional approaches such as transmission electron microscopy (TEM).

Introduction

Phages are bacterial viruses with two replication modes. Lytic phages initiate their reproduction directly after successful DNA injection into the host cell, which leads to prompt cell lysis and phage particle liberation, whereas temperate phages integrate their DNA into the host genome and replicate in conjunction with the host, lacking a direct lethal effect [1]. These so-called prophages can have positive effects for their hosts like protection against related phages, mediation of antibiotic resistance, expanded metabolic capacities, or transfer of virulence

Competing Interests: The authors have declared that no competing interests exist.

factors (reviewed in [2]). Since such cooperation can last for a long time, the host may reduce its prophage to a minimal size retaining only the beneficial genes [1,3]. Under certain conditions prophages can switch back to the lytic life style and kill their hosts [1]. This behavior is observed if the host faces physiological stress, for example in industrial fermentation processes, where it can heavily interfere with the production or lead to a total collapse of the culture with the consequence of economic damage [4,5].

Since the beginning of the genomic era thousands of bacterial genomes have been sequenced including a large number of prophages. An organism can harbor several prophages and their genomes can account for up to 20% of the host genome [1,6]. This genetic prophage pool consists of approximately 23% of all currently known phage genetic material [7]. At the time of this study, the NCBI viral genome page comprised nearly 1000 defined bacteriophage genomes, including 34 *Bacillus* phage genomes.

Our strain of interest, *Bacillus licheniformis* DSM13, has been completely sequenced [8,9] and a recently updated genome annotation is available [10]. Strains of the species *B. licheniformis* are extensively used for industrial applications, especially as production platform for enzymes and antibiotics [11]. The lysogeny of the type strain *B. licheniformis* DSM13 was investigated on the equivalent strains LMD 50.16 and LMD 75.15 [12]. The experiments identified one type of phage particle for LMD 50.16 and two for LMD 75.15 [13]. The phage identified in both strains was a PBSX ortholog, a prophage which appears in many strains of the *subtilis* group. The additional phage was strain specific [13,14]. However, it is unknown which genomic regions encode the observed phage particles.

To investigate the activity of prophages in the genome of *B. licheniformis* DSM13, a combined bioinformatics and next generation sequencing (NGS)-based approach was applied. Prophage regions were predicted by Prophage Finder [15] and manually evaluated and annotated. A phylogenetic classification of the annotated prophage genes was performed based on their homology to known phage genes. And finally, the NGS-based analysis of the phage particle-derived DNA enabled the identification of three active prophage-encoding regions. We assigned the activity of two phage particles to the encoding genomic loci and identified a third new active region. This new method enables fast and precise investigation of putative active prophage regions, and creates a new base for strain evaluation.

Material and Methods

Strains and growth conditions

All *Bacillus* strains used in this investigation are listed in Table 1. Cells were grown in LB [16] at 37°C with vigorous shaking if not otherwise stated.

Table 1. *Bacillus* strains used in this study.

Strain	Description	Source	Ref.
B. licheniformis DSM13	Type strain	DSMZ Braunschweig	[8,9]
B. licheniformis MW3	Derivate of *B. licheniformis* DSM13; Δ*hsdR1*, Δ*mcrA*, Δ*hsdR2*	Prof. Dr. Friedhelm Meinhardt, University Münster	[17]
B. licheniformis ΔPBSX	Derivate of *B. licheniformis* MW3; Δ*hsdR1*, Δ*mcrA*, Δ*hsdR2*, ΔBLi_Pp2	Lab strain collection	[18]
B. licheniformis ΔPBSX-ΔBLi_Pp3	Derivate of *B. licheniformis* ΔPBSX; Δ*hsdR1*, Δ*mcrA*, Δ*hsdR2*, ΔBLi_Pp2, ΔBLi_Pp3	This study	

doi:10.1371/journal.pone.0120759.t001

Prophage identification

For prediction of prophage regions in the genome sequence of *B. licheniformis* DSM13, the web-based tool Prophage Finder [15] was used with standard parameters (E-value 0.5, Hits per Prophage 5, Hit Spacing 5500). The results were manually evaluated and putative phage genes were annotated using InterProScan [19]. Variations in the GC-content were identified using Artemis [20]. The genomes of *Bacillus licheniformis* 9945A (CP005965) [21] and *Bacillus subtilis* 168 (NC_000964) [22] were searched for orthologues of the prophage regions using the script rod_finder (https://github.com/aleimba/bac-genomics-scripts). The range of the prophage regions was defined based on annotation and GC-content composition, and refined by a search for insertion repeats flanking the regions. For this purpose 500 bp up- and downstream of the prophage regions were aligned using Clone Manager (Sci-Ed Software, USA) and putative repeats were manually evaluated.

Deletion of prophage regions

For the construction of the prophage double mutant *B. licheniformis* ΔPBSX-ΔBLi_Pp3, the markerless deletion protocol of Rachinger et al. [18] was applied on the prophage single mutant strain *B. licheniformis* ΔPBSX [18]. *B. licheniformis* ΔPBSX is based on *B. licheniformis* MW3, a restriction-modification (RM) system-negative derivative of *B. licheniformis* DSM13 [17]. All PCR reactions (50 µl) consisted of 200 µM deoxynucleotides, 100 ng of template DNA, 5 pmol of each primer and 0.5 U Phusion High-Fidelity DNA Polymerase (Thermo Scientific, Darmstadt, Germany). PCR products were purified after gel electrophoresis using a Qiaquick Gel extraction Kit (Qiagen, Hilden, Germany). Sanger sequencing was performed on an ABI3730XL capillary sequencer (BigDye 3.1 chemistry; Applied Biosystems, Darmstadt). The flanking regions of the deletion targets were amplified by PCR. Flank A was constructed using the respective primer pair A1/A2 and flank B using the primer pair B1/B2 (Table 2). Both flanks were fused via SOE-PCR and cloned into the temperature-sensitive vector pKVM2 using *Bam*HI and *Nco*I restriction sites (*Bam*HI, *Nco*I, FastAP and T4-DNA-ligase (Thermo Scientific, Darmstadt, Germany); *Escherichia coli* TOP10 (Invitrogen, Carlsbad, USA)). *Escherichia coli* S17-1 [23] was used for conjugative transfer [23] of the deletion vector. The deletion of the prophage region BLi_Pp3 was confirmed by Southern blot analysis.

Phage particle isolation

B. licheniformis overnight cultures were used to inoculate 4 ml LB in a test tube with a starting OD_{600} of 0.05. The cultures were grown at 37°C with vigorous shaking. Prophages were induced 2 hours after inoculation using mitomycin C (Sigma-Aldrich, St. Louis, USA) as induction agent with a final concentration of 0.5 µg/ml. A non-induced culture was used as control

Table 2. Oligonucleotides used in this study.

Description	Sequence*
FlankA1	AAA**GGATCC**GCTGTCGTCATCAGAGGGCTGTC
FlankA2	GGTCAAAATGTGGTCAAAAATCAATTTTAAACAAAAAACAAGC
FlankB1	AAATTGATTTTTGACCACTTTTTGACCTTCTTTGTAGTATCTTTTC
FlankB2	TTTT**CCATGG**CAGCAGCGGAAACGCATCCTTGC
testA	TCGATGTGTGACCGAGACGCGTAC
testB	CGAGTGACGACGAAGTTTCC

*Restriction sites within the oligonucleotide sequences are printed in bold.

doi:10.1371/journal.pone.0120759.t002

and cell growth was followed by optical density measurement (Klett-Summerson Photoelectric Colorimeter with a green filter; KLETT Mfg. Co. Inc., NY). Eight hours after induction the supernatant was separated from the cells via centrifugation at 2,500 g for 5 min, followed by sterile filtration with a 0.45μm pore size filter (Sarstedt, Nümbrecht, Germany). Lysozyme from chicken egg white (10 μg/ml, SERVA Heidelberg, Germany) was added to the filtrated supernatant to disrupt the cell walls of potentially remaining host cells. RNase A (Qiagen, Hilden, Germany) and DNase I (Roche Diagnostics, Mannheim, Germany) were added to a final concentration of 10 μg/ml each to the filtrate. The mixture was incubated overnight (16 h) at 37°C. After the enzymatic removal of free nucleic acids, the phage particles were sedimented by ultracentrifugation using a Sorvall Ultracentrifuge OTD50B with a 60Ti rotor applying 200,000 g for 4 h. The supernatant was discarded and the pellet was solved in 200 μl TMK buffer [24], and stored at 4°C or directly used for DNA isolation.

Phage DNA isolation

The DNA isolation was performed using a MasterPure DNA Purification Kit from Epicentre (Madison, WI, USA). 200 μl 2x T&C-Lysis solution containing 1 μl Proteinase K was added to 200 μl of the phage suspensions and incubated for 15 min at 65°C. After addition of 200 μl MPC solution, proteins were precipitated by centrifugation for 10 min at 10,000 g. The supernatant was transferred to a new tube, mixed with 670 μl cold isopropanol and incubated for 10 min at −20°C. DNA precipitation was performed by centrifugation for 10 min at 17,000 g and 4°C. The DNA pellet was washed twice with 150 μl 75% Ethanol, air-dried and re-suspended in DNase free water.

DNase I treatment control

To test the efficiency of the enzymatic removal of free nucleic acids in the phage suspension, we performed a control experiment using conditions comparable to the DNase I treatment of the phage particle isolation protocol. Different chromosomal DNA dilutions in LB medium (16 μg/ml, 8 μg/ml, 4 μg/ml, 2 μg/ml and 1 μg/ml) were incubated for 16 h at 37°C with and without 10 μg/ml DNase I (Roche Diagnostics, Mannheim, Germany). To test the efficiency of the treatment, 1 μl of each dilution was used as template for a PCR reaction with the primer pair testA/testB and the Phusion High-Fidelity DNA Polymerase (Thermo Scientific, Darmstadt, Germany), according to the recommendations of the supplier. The control PCRs are constructed to generate products with a size of 159 bp from chromosomal DNA.

Table 3. NGS read coverage of phage DNA preparations from *B. licheniformis* strains.

Strain for phage DNA preparation	Mapped reads	Unmapped reads
B. licheniformis DSM13	1,278,534	24,221
B. licheniformis MW3	3,195,030	41,897
B. licheniformis ΔPBSX	3,515,890	40,356
B. licheniformis ΔPBSX-ΔBLi_Pp3 (1. exp.)	5,384,670	1,788,078
B. licheniformis ΔPBSX-ΔBLi_Pp3 (2. exp.)	194,173	94,872
B. licheniformis ΔPBSX-ΔBLi_Pp3 (3. exp.)	2,030,737	53,687

The sequencing raw data are available at the NCBI short read archive (SRP035551).

doi:10.1371/journal.pone.0120759.t003

Next generation sequencing (NGS) and data processing

NGS phage DNA libraries were generated with the NEBNext ultra DNA kit (NEB, Ipswich, MA, USA) and Nextera XT DNA Sample Preparation Kit (Illumina, San Diego, USA), and the sequencing was performed on an Illumina GAii sequencer (Illumina, San Diego, USA). The generated sequence reads (Table 3) were mapped on the genome of *B. licheniformis* DSM13 (accession number AE017333.1) [10] using bowtie2 [25], allowing 2% aberration. The output SAM files were converted to tds-files by the program samtotds and visualized by TraV [26]. The sequencing raw data are available at the NIH short read archive (SRP035551). To compare phage DNA mappings with different read numbers, a coverage analysis was performed using NPKM (**n**ucleotide activity **p**er **k**ilobase of exon model per **m**illion mapped reads) values [10,26]. NPKM values are normalized mapping values, calculated from the coverage of a defined region in relation to the overall genome coverage. In this paper, activity is defined as the abundance of phage DNA reads, generated from phage particle derived DNA, which correlates to the phage activity. For showing the hotspots of phage activity, NPKM values were calculated for continuous 1 kb-segments over the whole *B. licheniformis* DSM13 genome by TraV [26] and used to draw phage activity graphs in MS-Excel.

Signal to noise ratio determination

The signal to noise ratios for the phage DNA mappings were calculated from base coverage determined by TraV [26]. The base coverage is defined as the number of reads mapped over a specific base during the mapping process. The average base coverage in the prophage regions (signal) and the nonprophage regions (noise) were calculated by counting the number of reads and dividing them by the respective region size. The signal to noise ratios were calculated by dividing the signal values by the noise values. In addition, the relative noise (%) in relation to the signal was calculated by setting the signal values to 100%.

Negative staining and transmission electron microscopy

For negative staining, a carbon film, evaporated on a freshly cleaved mica surface, was partially floated off the mica by dipping into a drop of phage particle suspension. The carbon-mica sandwich with the adsorbed biological material was then transferred to a drop of washing solution (deionized water) and then completely floated off on a drop of negative staining solution, where it was transferred onto a 400 mesh specimen grid. The staining solution was completely removed, resulting in a shallowly stained specimen [27,28]. Electron micrographs of phage particles were taken at calibrated magnifications with a JEM 1011 transmission electron microscope (Jeol Ltd., Eching, Germany). Images were taken with an Orius SC 1000 A CCD camera and processed with the digital micrograph image processing software (Gatan, Munich, Germany).

Protein comparison of prophage regions BLi_Pp1 – BLi_Pp7

The proteins of each prophage region were compared to the remaining six regions with a focus on putative phage repressor proteins. Therefore, the BLi_Pp1 – BLi_Pp7 protein sequences were extracted from the *B. licheniformis* DSM13 genbank file (accession number NC_006322.1) by the script cds_extractor (v0.6) (https://github.com/aleimba/bac-genomics-scripts). A BLASTP comparison was performed for every prophage region against a protein database comprising the protein sequences of the remaining six prophage regions. Subsequently, the resulting protein pairs were aligned using the Needleman-Wunsch-Algorithm [29]. The annotations of the prophage regions were searched for known and putative repressors and the

generated Needleman-Wunsch scores were used for the comparison of BLi_Pp1 – BLi_Pp7 (S1 Table), putative repressor proteins and related regulon proteins were marked in blue).

Phylogenetic classification of prophage regions

Phylogenetic classification of the identified prophage regions was done by comparison to known phages. In total, 979 bacterial phage genomes and the associated taxonomic information (Table 4) from NCBI's viral genome page were used (http://www.ncbi.nlm.nih.gov/genomes, viruses; data taken on June 5th 2013). A BLAST phage protein sequence database was created comprising 91,518 phage protein sequences, extracted from the genbank format-files by the script cds_extractor (v0.6) (https://github.com/aleimba/bac-genomics-scripts) and supplemented by the protein sequences of *B. subtilis* 168 PBSX phage (NC_000964). BLASTP comparisons were performed for all *B. licheniformis* phage proteins to the phage protein database. Subsequently, all hits were aligned by the needle program of EMBOSS suite [29] with default values. By comparing the similarity scores of all pairs with full-length alignment the best hit for every phage protein was determined (S2 Table). Hits with similarity $< 30\%$ were discarded.

Results and Discussion

Prophage identification

The genome sequence of *B. licheniformis* DSM13 was initially scanned with the prediction tool Prophage Finder [15]. The nine predicted prophage regions were evaluated based on a set of genomic features and thus, seven regions were annotated as prophage regions BLi_Pp1 – BLi_Pp7 (access code AE017333.1). The two remaining predicted regions were evaluated as false positives and not considered further. The criteria used for evaluation of the prophage predictions were similarity to known phage genes, GC content deviation and the presence of insertion repeats (Table 5). The final seven genomic loci are covering 4.7% of the complete genome.

The regions BLi_Pp1, BLi_Pp5 and BLi_Pp7 are encoded by genome loci of a size between 11 kb and 22 kb. In all three regions open reading frames (ORFs) with similarity to phage-related genes could be identified. An evaluation of the GC content of the prophage regions indicates a clear GC content deviation compared to the host genome (Table 5). Furthermore, the regions are flanked by insertion repeats. The region BLi_Pp1 is located next to a tRNA cluster which is similar to known phage integration sites [30,31]. A genome comparison with *Bacillus*

Table 4. Distribution of bacterial phage reference genomes used for the phylogenetic classification.

Phage family	Genetic material	Number of genomes
Microviridae	*ssDNA phages*	16
Inoviridae		32
Corticoviridae	*dsDNA phages*	1
Myoviridae		226
Plasmaviridae		1
Podoviridae		158
Siphoviridae		425
Tectiviridae		4
unclassified		62
Cytoviridae	*dsRNA phages*	5
Leviviridae	*ssRNA phages*	11

doi:10.1371/journal.pone.0120759.t004

Table 5. Prophage regions of *B. licheniformis* DSM13.

Prophage	GC content*	Size (bp)	Insertion repeats#	Location
BLi_Pp1	38.10	11,177	61 –6 bp	927,299–938,595
BLi_Pp2	47.05	27,509	no	1,317,754–1,345,262
BLi_Pp3	42.50	41,566	27 –1 bp	1,422,556–1,464,174
BLi_Pp4	44.73	38,319	256 –13 bp	1,504,028–1,542,847
BLi_Pp5	39.89	10,524	50 –3 bp	2,855,587–2,866,209
BLi_Pp6	40.95	44,793	18 bp	3,424,376–3,469,186
BLi_Pp7	35.58	21,733	19 –1 bp	4,155,490–4,177,258

*The average host genome GC content is 46.2. The size of the prophage regions refers to the sequence between the insertion repeats. For BLi_Pp2 the range from the first to the last identified phage gene was used since no insertion repeats were found.

Size of the insertion repeats with the number of variations in bp.

doi:10.1371/journal.pone.0120759.t005

licheniformis 9945A [21] and *Bacillus subtilis* 168 [22,32] revealed the absence of orthologous prophage regions in these strains, indicating that BLi_Pp1, BLi_Pp5 and BLi_Pp7 are part of the DSM13 accessory genome. The region BLi_Pp7 contains the genes of a restriction modification system, which has been reported as a gene class also identified in other prophages [33,34].

BLi_Pp3, BLi_Pp4 and BLi_Pp6 are prophage regions with a size between 38 kb and 45 kb. The GC contents of BLi_Pp3 and BLi_Pp6 also deviate from the host genome and all three loci are flanked by insertion repeats (Table 5). Again, no orthologous prophage regions could be found in *B. licheniformis* 9945A and *B. subtilis* 168.

Prophage BLi_Pp2 was bioinformatically predicted due to ORFs with similarity to phage genes. The strong similarity to the PBSX prophage locus of *B. subtilis* 168 [35] indicates that BLi_Pp2 represents a *B. licheniformis* orthologue of the PBSX prophage. For BLi_Pp2 no flanking insertion repeats could be identified and according to the GC-content, no clear sequence deviation was found (Table 5). However, this is in accordance with the PBSX insertion site in *B. subtilis*, where the genomic locus also shows no GC content deviation and no insertion repeats are present.

Comparison of BLi_Pp1 – BLi_Pp7

The similarity of the seven prophage regions to each other was investigated with a focus on putative phage repressor proteins. These proteins repress the induction of phages to remain the lysogenic state [36,37]. Repressor proteins can also confer immunity to the host strain against superinfection [38]. The annotations of the prophage regions were searched for known and putative repressors. The proteins of each prophage region were compared by BLASTP to the remaining six regions, and a global alignment of the resulting protein pairs was performed using the Needleman-Wunsch-Algorithm [29]. The resulting Needleman-Wunsch scores were used for the comparison of BLi_Pp1 – BLi_Pp7 (S1 Table).

The overall similarity between the prophages is low. BLi_Pp7 does not possess any similarity at all to the remaining six prophage regions (S1 Table). For BLi_Pp1 and BLi_Pp5 significant similarities (scores >30) were only observed for a few proteins (e.g. transposases/integrases, some hypothetical proteins and a terminase large subunit). No repressor proteins could be identified for BLi_Pp1, BLi_Pp5 and BLi_Pp7.

The similarities observed for the prophage regions BLi_Pp2, BLi_Pp3, BLi_Pp4 and BLi_Pp6 are more distinct. The comparison of BLi_Pp2 to BLi_Pp3 shows homology for the proteins BLi01317 – Bli01325. This BLi_Pp2 region comprises the transcriptional repressor

Xre and its regulon (XkdB-XkdC-XkdD-XtrA; S1 Table, marked in blue), which has been shown to control the induction of the PBSX phage in *B. subtilis* [37,39]. The respective open reading frames (ORFs) in BLi_Pp3 are the repressor YqaE (BLi01433) and parts of its regulon (BLi01444, BLi01443 and BLi01445, regulon proteins are marked in blue in S2 Table). YqaE (syn. SknR) was described as the repressor of the skin element, a prophage-like element of *B. subtilis* and belongs also to the Xre family of transcription repressors [40]. BLi_Pp4 and BLi_Pp6 do not possess any homology to the repressors of BLi_Pp2 or BLi_Pp3 and their regulon proteins. Only BLi03634 of prophage BLi_Pp6 is similar to the toxin YqaH of BLi_Pp3 (BLi01438, similarity score 78.3), which is in *B. subtilis* encoded by the skin element and executes its lethal effect by inhibition of host DNA replication [40]. Vice versa, BLi_Pp6 encodes only a putative HTH-type transcriptional repressor protein, which shows no homology to the remaining prophage region proteins. For BLi_Pp4 only an HTH-type DNA-binding protein was found, possessing no homology to the remaining prophage regions.

Interestingly, the highest similarity scores between BLi_Pp2 and BLi_Pp3 as well as BLi_Pp6 were observed for the holin-like protein XhlB and the membrane-associated protein XhlA (83% to 90% similarity, marked in green in S1 Table), which are essential for host cell lysis in the *B. subtilis* PBSX phage [41]. The associated endolysin XlyA, which is only found in BLi_Pp6, is not essential for functionality [41] and might have been replaced by a different endolysin in BLi_Pp3. BLi_Pp4 seems to encode a different type of host lysis system (BLi01567-BLi01569), for which no or only low similarity proteins could be identified in BLi_Pp2, BLi_Pp3 and BLi_Pp6.

Prophage classification

To classify the *B. licheniformis* DSM13 prophage regions, a comparison to publicly available phage genomes was performed. Initially a database comprising 979 bacterial phage genomes available at the NCBI database was constructed (Table 4). The protein sets encoded by the seven annotated prophage-like regions were compared by BLASTP with this database, followed by a global alignment of the resulting protein pairs using the Needleman-Wunsch-Algorithm [42]. For evaluation of similarity of prophage proteins, the best Needleman-Wunsch alignment score was considered, setting a cut off $\geq$ 30% for the full-length alignments (S2 Table). 89% of all prophage proteins encoded by region BLi_Pp2 showed top similarities to the proteins encoded by the *B. subtilis* prophage PBSX. In total, only four proteins exhibited the top similarity to non-PBSX phage proteins. The proteins of the remaining prophage regions showed similarities to members of the *Caudovirales*. Phages of this order harbor dsDNA within their phage particles [43]. An in-depth evaluation of protein homologies indicated that the majority of the investigated protein sequences match proteins encoded by members of the *Siphoviridae* phage family. However, an assignment of a more in depth taxonomic level was not possible due to the partly incomplete classification of the database hits and the comparably small database size (considering the huge pool of existing phages). Moreover, the mosaic nature of phage genomes [44] makes an unambiguous classification at sequence level very difficult. For example around 54% of the BLi_Pp6 proteins are homologous to the lambda-like *Bacillus* phage phi105 including 26% with top homologies (S2 Table, average similarity score 77), but the remaining BLi_Pp6 proteins are not similar at all to this phage.

Taken together, the prophage regions may be divided into prophages (BLi_Pp2, BLi_Pp3, BLi_Pp4 and BLi_Pp6) and prophage remnants (BLi_Pp1, BLi_P5). BLi_Pp1 and BLi_Pp5 with about 11 kb are probably too small to contain a complete genome of a *Siphoviradae* or *Myoviradae* phage [45,46]. The region BLi_Pp7 shows like BLi_Pp1 and BLi_Pp5 limited homology to known phage genes (S2 Table). However, BLi_Pp7 may have a sufficient size for a

complete phage genome. The fact that no phage repressor or repressor-like proteins were identified in BLi_Pp1, BLi_P5 and BLi_Pp7 could be additional evidence that these regions are prophage remnants consisting of harmless or beneficial genes as proposed by Lawrence et al., [3].

Phage particles and their genetic sources

To check if the identified prophage regions are able to generate phage particles, induction experiments with mitomycin C were performed using *B. licheniformis* DSM13 and different deletion mutants (Table 1). After induction, phage particles were purified, and isolated phage DNA was sequenced using NGS. The lab strain *B. licheniformis* MW3, a restriction-modification system-negative derivate of *B. licheniformis* DSM13, was also used in all experiments since the prophage deletion mutants were constructed in the MW3 background. The behavior of *B. licheniformis* DSM13 and MW3 was comparable during the experiments (for details see S1, S2 and S3 Figs.). To exclude a specific contamination with chromosomal host DNA, which would generate mappable reads, the supernatant of the induced *B. licheniformis* cultures was treated with lysozyme, DNase and RNase before the phage DNA was liberated from the particles. The efficiency of this procedure was tested in control experiments with different concentrations of chromosomal DNA. The control-PCRs were negative for all enzyme-treated samples, but generated a PCR product for all samples without nuclease treatment (S4 Fig.).

Activity of the PBSX-like phage region BLi_Pp2. The wild-type strain *B. licheniformis* DSM13 showed a clear decrease in turbidity 3 hours after induction (Fig. 1, red graph with filled symbols), which indicates bacterial lysis. Gel electrophoresis of phage DNA purified from the phage suspension showed two DNA bands with a size of approximately 13 kb and 44 kb (Fig. 2). In addition, the TEM-based investigation revealed a type of phage particles, which corresponds in size and appearance to PBSX-like phages (Fig. 3A). The DNA was sequenced by next generation sequencing (NGS) technology and mapped on the genome of *B. licheniformis* DSM13. Normalized activity values (NPKM values, for details see method section and [10,26]) were calculated from the read coverage and used to compare the phage DNA mappings from different experiments with prophage deletion mutants and the DSM13 wild type (Fig. 4). The mapping of the *B. licheniformis* DSM13 derived phage DNA shows a distribution of the reads over the whole genome (Fig. 4A). The coverage of the prophage regions is with a signal to noise ratio of 2.6 (Table 6) only 2.6 times increased. The relative noise over the whole genome is with 38% (Table 6), strongly increased. The general distribution of reads over the whole genome indicates random packing of chromosomal DNA and supports the assumption of a PBSX

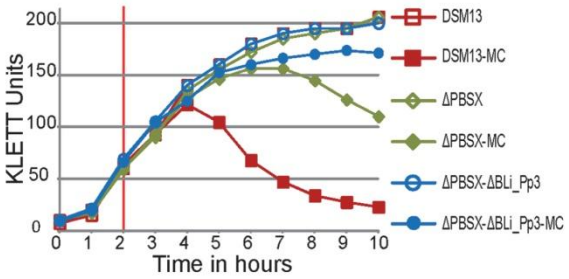

Fig 1. Growth behavior of *B. licheniformis* DSM13 and its derivatives after induction with mitomycin C. The KLETT Units represent the turbidity of the culture. The vertical red line at 2 hours marks the time point of induction with 0.5 μg/ml mitomycin C. Induced cultures (MC) are marked with filled symbols and non-induced cultures with open symbols. *B. licheniformis* DSM13 shows a loss in turbidity 3 hours after induction with mitomycin C and *B. licheniformis* ΔPBSX after 6 hours. The strain *B. licheniformis* ΔPBSX-ΔBLi_Pp3 exhibited no decrease of turbidity after mitomycin C treatment.

doi:10.1371/journal.pone.0120759.g001

ortholog prophage in *B. licheniformis* DSM13, since PBSX-like prophages are described to randomly pack 13 kb fragments of the host chromosome instead of their own genome with a size of about 30 kb [24,35,47,48]. The comparison of the results from strains *B. licheniformis* DSM13 and MW3 with *B. licheniformis* ΔPBSX, a BLi_Pp2 deletion mutant, strongly indicates that prophage BLi_Pp2 is responsible for the huge amount of prophage-unspecific sequence reads. The higher read coverage close to the origin of replication of the *B. licheniformis* DSM13 genome (Fig. 4A) can be explained through the higher available amount of DNA of this multicopy genomic region due to the replication start site [49]. The strongly increased read coverage observed at the BLi_Pp3 region (Fig. 4A) indicates this region as possible origin of the second phage DNA band of approximately 44 kb (Fig. 2). In addition, the prophage regions BLi_Pp2, BLi_Pp5, BLi_Pp6 and BLi_Pp7 exhibit a small read accumulation as well as a non-prophage region at 3.8 Mb. Considering the increased average coverage due to the genomic DNA from the PBSX orthologous phage, no further conclusions can be drawn on this observation at the moment.

The coverage gaps, which can be observed over the whole genome (Fig. 4A), correspond to repetitive regions. The applied mapping method considers only reads, which can be exclusively assigned to a single position within the genome. In addition, the *B. licheniformis* MW3 phage DNA mapping (S3 Fig.) exhibits an expected coverage gap at the BLi_Pp7 region, which corresponds to a deletion of the RM system described in [17].

Phage activity in the ΔPBSX mutant. A deletion mutant of the PBSX-orthologous region BLi_Pp2, designated as *B. licheniformis* ΔPBSX [18], was used in further induction experiments. Six hours after induction with mitomycin C the culture showed first lysis effects (Fig. 1). Gel electrophoresis of the isolated phage DNA revealed an exclusive DNA band of approximately 44 kb (Fig. 2), and TEM investigation of the phage suspension showed the presence of phage particles with a *Siphoviridae*-like morphology (Fig. 3B). In contrast, no phage

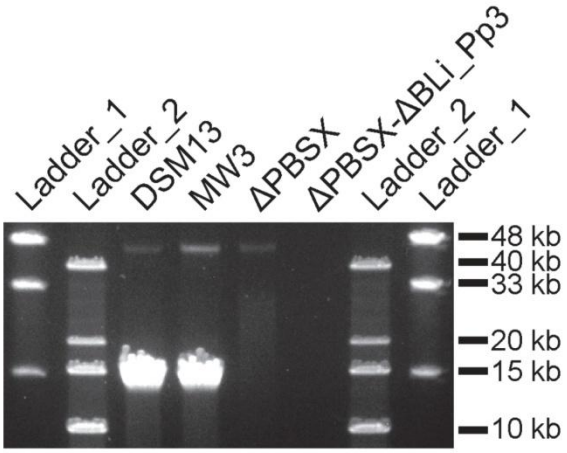

Fig 2. Pulsed field gel electrophoresis (PFGE) of phage DNA isolated from mitomycin C treated *B. licheniformis* DSM13 and its derivatives. The 1% gel was run for 18 h at 14°C using a voltage of 6V/cm and switch times ramped from 0.1–10 sec. The phage DNA preparations of *B. licheniformis* DSM13 and *B. licheniformis* MW3 show a strong band of approximately 13 kb and a weak band of approximately 44 kb. The phage DNA preparation of *B. licheniformis* ΔPBSX shows a weak band of approximately 44 kb, and for *B. licheniformis* ΔPBSX-ΔBLi_Pp3 no bands could be detected. NEB MidRange I PFG marker (ladder 1) and Invitrogen 1 kb DNA Extension Ladder (ladder 2) were used. The depicted figure has been cropped to improve clarity. S2 Fig. shows a full size gel picture. Uninduced cultures of *B. licheniformis* DSM13 and ΔPBSX do not show any DNA bands (data not shown).

doi:10.1371/journal.pone.0120759.g002

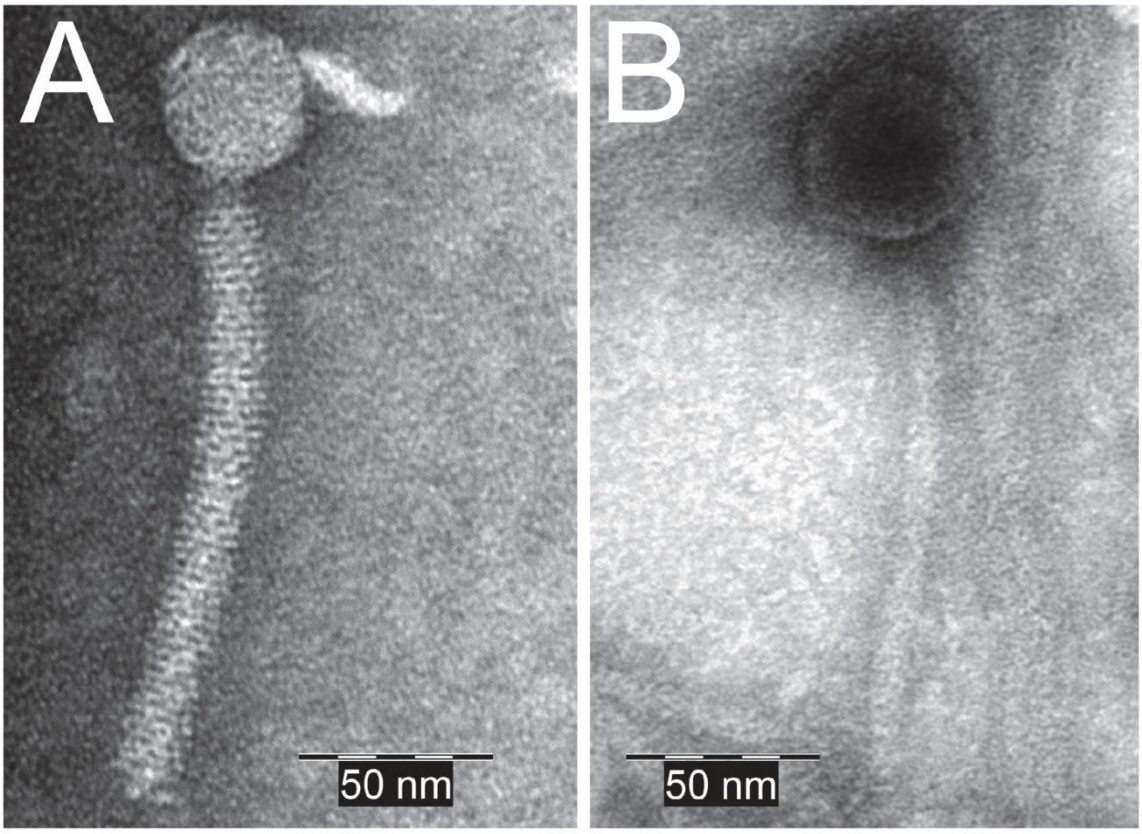

Fig 3. TEM micrographs of negatively stained phage particles detected in *B. licheniformis* strains after induction with mitomycin C. A: PBSX-like phage particle, detected in the phage suspension isolated from *B. licheniformis* DSM13. B: *Siphoviridae*-like phage particle, detected in the phage suspension of *B. licheniformis* ΔPBSX.

doi:10.1371/journal.pone.0120759.g003

particles from the PBSX-like particle type (Fig. 3A) could be detected. Thus, we conclude that the particles observed in the induction experiments with *B. licheniformis* DSM13 are encoded by the PBSX-like genome locus BLi_Pp2.

The mapping of NGS generated sequences from the second experiment revealed two genomic regions with increased read coverage, one at the locus annotated as BLi_Pp3 reaching into the adjacent downstream genomic region, and the second slight coverage increase at the locus BLi_Pp6 (Fig. 4B, green graph). Unlike the first experiment, no distinct read coverage over the whole genome was observed. The coverage of region BLi_Pp3 is strongly increased and declines for about 200 kb downstream, covering in addition the BLi_Pp4 prophage region. Due to the lack of coverage peaks at the boundaries of the BLi_Pp4 region, it remains unclear whether the locus BLi_Pp4 contributes to the phage particle-derived reads. Consequently, the NGS analysis confirms the activity of at least one more prophage region, seen as a DNA band of approximately 44 kb (Fig. 2) in agarose gel electrophoresis and as phage particle in TEM (Fig. 3B). Due to the high read coverage in the BLi_Pp3 region we expect this prophage to be active.

Activity of prophage regions BLi_Pp3 and BLi_Pp6. To confirm this assumption and to elucidate the role of BLi_Pp4, a *B. licheniformis* ΔPBSX-ΔBLi_Pp3 double mutant was

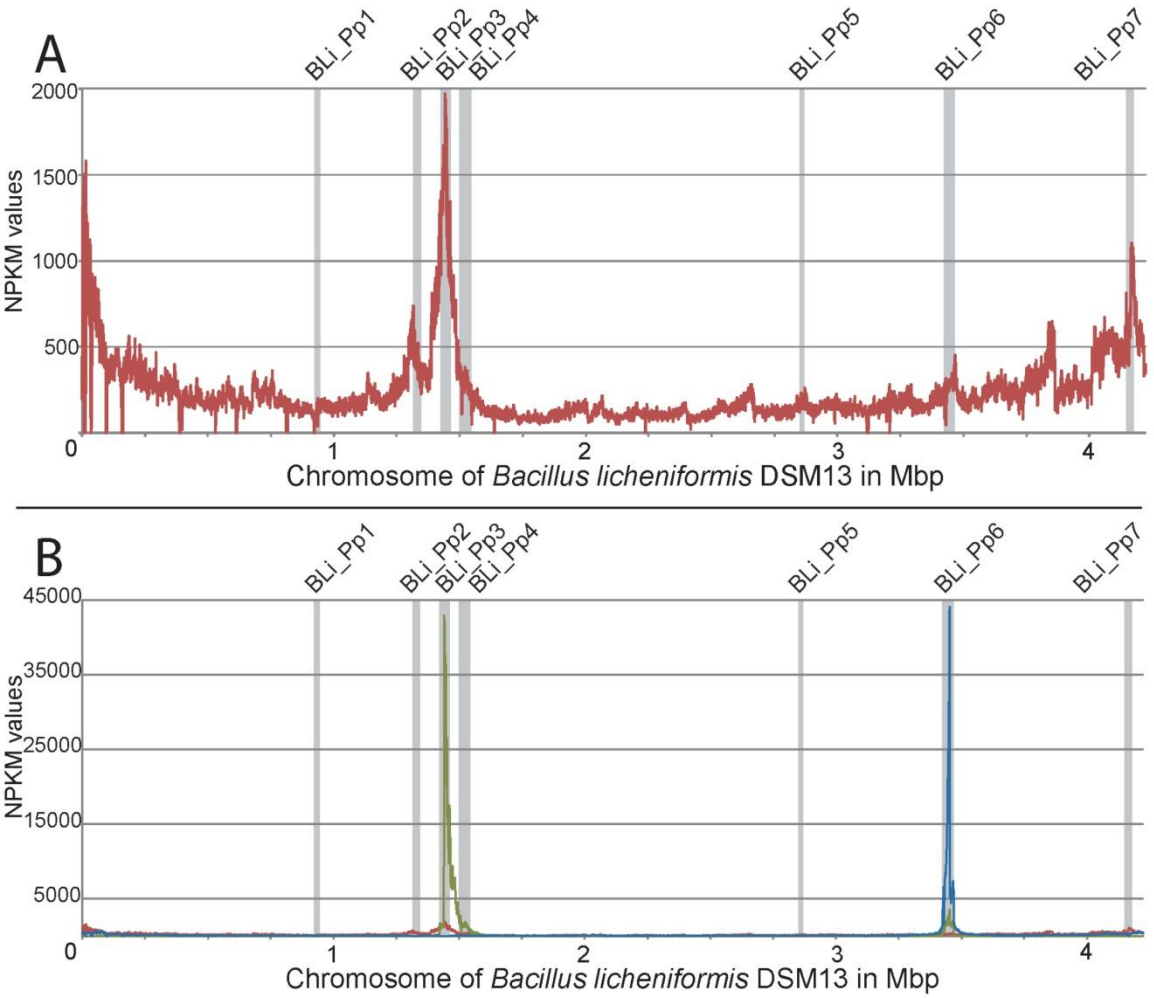

Fig 4. Phage activity graphs on the genome of *B. licheniformis* DSM13. *B. licheniformis* phage DNA was sequenced by next generation sequencing (NGS) and sequences were mapped on the genome of *B. licheniformis* DSM13. The results are displayed in NPKM values calculated by TraV [26]. Prophage regions BLi_Pp1 – BLi_Pp7 are marked with grey bars. A: The phage DNA read mapping of *B. licheniformis* DSM13 shows a even distribution of reads over the whole genome as well as read accumulations around the origin of replication and also in some further regions, especially in BLi_Pp3. B: The *B. licheniformis* ΔPBSX phage DNA read mapping (green graph) shows a strong read accumulation in the region BLi_Pp3 compared to the *B. licheniformis* DSM13 read mapping (red graph), and a slight read accumulation in region BLi_Pp6. The *B. licheniformis* ΔPBSX-ΔBLi_Pp3 phage DNA read mapping (blue graph, result of the first experiment is shown) exhibits a distinct read accumulation in BLi_Pp6.

doi:10.1371/journal.pone.0120759.g004

constructed on the background of *B. licheniformis* ΔPBSX [18]. In induction experiments with mitomycin C the double mutant culture exhibited only a slight reduction of turbidity, which might be rather due to the mitomycin C toxicity than to prophage-induced cell lysis (Fig. 1). The phage DNA preparation revealed neither a DNA band on an agarose gel (Fig. 2), nor was any DNA concentration measurable by NanoDrop (Peqlab, Erlangen, Germany). Nevertheless, this DNA preparation generated mappable reads after NGS sequencing (Fig. 4B, blue graph). Strikingly, the experiment showed a clear coverage peak within the BLi_Pp6 region. No phage particles could be observed in the phage suspension of *B. licheniformis* ΔPBSX-ΔBLi_Pp3 by TEM, although the BLi_Pp6 DNA might have been protected through a phage particle

Table 6. Results of the signal to noise ratio determination.

	DSM13	MW3	ΔPBSX	ΔPBSX ΔBLi_Pp3 (1. exp.)	ΔPBSX ΔBLi_Pp3 (2. exp.)	ΔPBSX ΔBLi_Pp3 (3. exp.)
Average base coverage in prophage regions (signal)	82.7	187	1447.2	1671.8	44.9	664.6
Average base coverage in non-prophage regions (noise)	31.5	79.7	27.2	68.3	1.4	5.4
Signal-to-noise ratio	2.6	2.3	53.2	24.5	32.1	123.1
Noise (%)	38.1%	42.6%	1.9%	4.1%	3.1%	0.8%

The table shows the average base coverage and signal to noise ratios for mappings of NGS experiments of phage DNA isolated from *B. licheniformis* DSM13 and MW3, and the mutant strains *B. licheniformis* ΔPBSX and ΔPBSX_ΔBLi_Pp3. The average base coverage in prophage regions (signal) and in non-prophage regions (noise) were calculated by counting the number of reads and dividing them by the respective region size. The signal to noise ratio (signal divided by noise) was calculated to compare experiments with different reads numbers. The relative noise (%) gives a percent value for the noise in relation to the signal (set 100%).

doi:10.1371/journal.pone.0120759.t006

considering the DNA isolation protocol. The identification of an active prophage locus in the absence of TEM-detectable phage particles and measurable or visible phage DNA highlights the superior sensitivity of NGS based phage identification.

In contrast to BLi_Pp6, the BLi_Pp4 prophage region showed no sequence read accumulation in the *B. licheniformis* ΔPBSX-ΔBLi_Pp3 mapping (Fig. 4B). Due to this observation we conclude that BLi_Pp4 does not form functional phage particles or it requires the presence of the genomic locus BLi_Pp3 for full activity. As neither phage particles nor phage DNA could be found in the phage suspension of *B. licheniformis* ΔPBSX-ΔBLi_Pp3, we further conclude that the phage particle observed in the single mutant strain *B. licheniformis* ΔPBSX (Fig. 3B) and the observed DNA band of approximately 44 kb (Fig. 2) correspond to the BLi_Pp3 prophage region.

Whether BLi_Pp3 has an impact on the expression of BLi_Pp6 is not known. However, BLi_Pp4 and BLi_Pp6 do not show any homology to the repressor regions of BLi_Pp2 and BLi_Pp3, and the overall similarity between the prophage regions is low. The DNA concentration in the phage DNA samples of *B. licheniformis* ΔPBSX-ΔBLi_Pp3 was not measurable. Thus, it is more likely that BLi_Pp6 is a low abundant phage and therefore its activity was not clearly visible in experiments with the single mutant strain *B. licheniformis* ΔPBSX.

Prophage region BLi_Pp7. The evaluation of the prophage region BLi_Pp7 was inconclusive. In *B. licheniformis* DSM13 a slight coverage increase for BLi_Pp7 was visible, but this might be due to the overall high read coverage generated by the random packing PBSX-like phage BLi_Pp2. In the *B. licheniformis* ΔPBSX mutant and the ΔPBSX-ΔBLi_Pp3 double mutant strains no increased read coverage was visible on the BLi_Pp7 loci after induction with mitomycin C. However, no conclusion on the BLi_Pp7 character can be drawn from the last observations. Since the mutant strains were constructed in the MW3 background, in which the RM system gene *hsdR*2 of the BLi_Pp7 region is deleted [17], the potential prophage locus is not in its native state. Due to the lack of mutants with DSM13 instead of MW3 genomic background, the activity of the BLi_Pp7 region has not been further investigated.

NGS-based phage identification

The use of NGS for the investigation of phage activity by sequencing of particle-protected phage DNA is a new, sensitive approach. It allows fast investigation of the capability of prophages to form particles and a specific assignment of the activity to a specific genome locus. A

recently published investigation of extra-chromosomal bacteriophages [50] also utilized an NGS approach. In this study, after the enrichment of extra-chromosomal DNA, NGS was performed to determine the appearance and state of these elements. Our approach targets particle-forming phages, thus showing their activity and particle content. Similar approaches utilize PCR analysis after induction with mitomycin C to show the presence of specific phage DNAs [50,51], but do not address the whole phage genome.

This study was conducted with the intention to investigate dsDNA phage genomes, and thus, dsDNA specific protocols were used for DNA isolation, library preparation and sequencing. To our knowledge there is no report of *Bacillus* phages that do not pack their genomes as dsDNA. However, there are phages, like members of the family *Inoviridae*, which pack their genomes as ssDNA and also have the ability to lysogenize their hosts [52]. Genome sequences of *Inoviridae*, deposited at the NCBI database, have a size of about 4.5 kb—10.5 kb, similar to the small prophage regions BLi_Pp1 (11 kb) and BLi_Pp5 (10.5 kb). The evaluation of the sequence comparisons does not indicate that the two regions belong to a known ssDNA generating phage class. However, we cannot exclude that BLi_Pp1 and BLi_Pp5 might belong to unknown classes of ssDNA or RNA packing phages and thus, would not have been detected with our methodology.

Sequencing of low abundant phages. The sequencing of phage DNA from the double mutant strain *B. licheniformis* ΔPBSX-ΔBLi_Pp3 revealed a distinct coverage peak for prophage region BLi_Pp6, although no DNA concentration was measurable or visible for this sample. However, a considerably high portion of unmapped reads was observed during the read mapping (Table 3). Consequently, the induction experiments for this strain where performed in triplicates (exp. 1–3).

For evaluation of the background noise of the read mappings, the relative percentage of noise was calculated (Table 6). The relative noise for *B. licheniformis* DSM13 and MW3 is 38.1% and 42.6% in non-prophage regions, respectively. For the *B. licheniformis* single mutant strains ΔPBSX, the majority of reads map within prophage regions and only 1.9% relative noise was detected. For the double mutant strain *B. licheniformis* ΔPBSX_ΔBLi_Pp3 the relative noises are comparably low for the first two experiments (4-3%) and for the third experiment even <1%. The high signal to noise ratios (24.5–123.1) and the comparable mappings for the *B. licheniformis* ΔPBSX_ΔBLi_Pp3 experiments (S5 Fig.) show that the results are unaffected by the number of unmapped reads.

Finally, we would like to point out that in our experiments with the *B. licheniformis* DSM13 wild-type strain the presence of the predominant PBSX-like prophage BLi_Pp2 hides the activities of the loci that generate low abundant phage particles. As PBSX-like prophages are distributed in many *Bacillus* strains [13], we expect a similar situation in the whole *B. subtilis* species complex. Thus, the application of our workflow, including the deletion of the predominant PBSX-like prophage, might result in the identification of currently unidentified phages.

Conclusion

Here we demonstrate how genomic methods can be used in the field of prophage analysis. Seven prophage-like regions could be identified in *B. licheniformis* DSM13 by comparative genomics, one PBSX-like prophage (BLi_Pp2) and six prophage regions with homology to known members of the *Siphoviridae*. The applied NGS approach confirmed the activity of two prophages from literature and revealed in addition a third active prophage. The read mapping and coverage analysis enabled the assignment of the active prophages to the genome regions BLi_Pp2, BLi_Pp3 and BLi_Pp6. The described workflow combines genomics and classical TEM- and gel-based methods for prophage analysis.

Supporting Information

S1 Fig. Prophage induction in *B. licheniformis* DSM13 and MW3. The KLETT Units represent the turbidity of the culture. The vertical red line at 2 hours marks the induction point with 0.5 μg/ml mitomycin C. Induced cultures (MC) are marked with filled symbols and non-induced cultures with open symbols. *B. licheniformis* DSM13 and MW3 show a loss in turbidity 3 hours after induction with mitomycin C.
(TIF)

S2 Fig. Pulsed field gel electrophoresis of phage DNA isolated from *B. licheniformis* DSM13 and its derivatives (full size). The 1% gel was run for 18 h at 14°C using a voltage of 6V/cm and switch times ramped from 0.1–10 sec. The phage DNA preparations of *B. licheniformis* DSM13 and *B. licheniformis* MW3 show a strong band of approximately 13 kb and a weak band of approximately 44 kb. The phage DNA preparation of *B. licheniformis* ΔPBSX shows a weak band of approximately 44 kb, and for *B. licheniformis* ΔPBSX-ΔBLi_Pp3 no bands could be detected. NEB MidRange I PFG marker and Invitrogen 1 kb DNA Extension Ladder were used.
(TIF)

S3 Fig. Phage activity graphs of *B. licheniformis* DSM13 and MW3 on the genome of *B. licheniformis* DSM13. After mitomycin C induction *B. licheniformis* phage DNA was isolated and sequenced by next generation sequencing (NGS). The sequences were mapped on the genome of *B. licheniformis* DSM13 and the NPKM value calculation was performed by TraV [26]. Prophage regions BLi_Pp1 – BLi_Pp7 are marked with grey bars. The read mappings of *B. licheniformis* DSM13 (red graph) and *B. licheniformis* MW3 (yellow graph) are comparable.
(TIF)

S4 Fig. Efficiency test for the DNase I treatment of the phage DNA preparation protocol. Different concentrations of chromosomal DNA were treated with DNase I and afterwards used for PCR. All PCRs with DNase I-treated chromosomal DNA samples did not result in a PCR product. All PCRs with non-treated chromosomal DNA generated an expected 159 bp control fragment. GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific) was used as length standard.
(TIF)

S5 Fig. Graphical comparison of three NGS experiments on *B. licheniformis* mutant ΔPBSX-ΔBLi_Pp3. Three independent phage DNA preparations were sequenced and mapped to the genome of *B. licheniformis* DSM13. The results are displayed in NPKM values calculated by TraV [26]. The three experiments (**A.—C.**, exp. 1. -3.) of *B. licheniformis* ΔPBSX-ΔBLi_Pp3 are comparable. All three mappings show a clear read accumulation at the BLi_Pp6 prophage region.
(TIF)

S1 Table. Annotation of the prophage regions BLi_Pp1 – BLi_Pp7 and results of the BLASTP comparison and Needleman-Wunsch alignment of the seven prophage regions to each other. The resulting Needleman-Wunsch similarity scores are colored according to similarity classes. Scores of 0–25 are not marked, low scores of 25–50 are marked in yellow, medium scores of 50–75 are marked in orange and high scores of 75–100 are marked in red. If more than one hit was found for one query protein the respective higher similarity score was used. Differences in the comparison of two prophage regions, dependent on the BLASTP direction, are due to paralogous proteins.
(XLSX)

S2 Table. Protein comparison of the *B. licheniformis* DSM13 prophage regions BLi_Pp1—BLi_Pp7 to known phage proteins.
(XLSX)

Acknowledgments

We would like to thank Stefania Neuber and Jessica Starke for performing induction experiments. We are grateful to Professor Friedhelm Meinhardt for kindly providing the strain *Bacillus licheniformis* MW3 and to Professor Rolf Daniel for valuable discussion.

Author Contributions

Conceived and designed the experiments: RH HL. Performed the experiments: RH DPR JH MH SV. Analyzed the data: RH SD HL SV. Contributed reagents/materials/analysis tools: SD AL. Wrote the paper: RH HL SV.

References

1. Casjens S. Prophages and bacterial genomics: what have we learned so far? Mol Microbiol. 2003; 49: 277–300. doi:10.1046/j.1365-2958.2003.03580.x PMID: 12886937
2. Bondy-Denomy J, Davidson AR. When a virus is not a parasite: the beneficial effects of prophages on bacterial fitness. J Microbiol. 2014; 52: 235–42. doi:10.1007/s12275-014-4083-3 PMID: 24585054
3. Lawrence JG, Hendrix RW, Casjens S. Where are the pseudogenes in bacterial genomes? Trends Microbiol. 2001; 9: 535–540. doi:10.1016/S0966-842X(01)02198-9 PMID: 11825713
4. Garneau JE, Moineau S. Bacteriophages of lactic acid bacteria and their impact on milk fermentations. Microb Cell Fact. BioMed Central Ltd; 2011; 10 Suppl 1: S20. doi:10.1186/1475-2859-10-S1-S20 PMID: 21995802
5. Wünsche L. Importance of bacteriophages in fermentation processes. Acta Biotechnol. 1989; 9: 395–419. doi:10.1002/abio.370090502
6. Canchaya C, Proux C, Fournous G, Bruttin A, Brussow H. Prophage Genomics. Microbiol Mol Biol Rev. 2003; 67: 238–276. doi:10.1128/MMBR.67.2.238-276.2003 PMID: 12794192
7. Casjens SR. Comparative genomics and evolution of the tailed-bacteriophages. Curr Opin Microbiol. 2005; 8: 451–8. doi:10.1016/j.mib.2005.06.014 PMID: 16019256
8. Veith B, Herzberg C, Steckel S, Feesche J, Maurer KH, Ehrenreich P, et al. The complete genome sequence of Bacillus licheniformis DSM13, an organism with great industrial potential. J Mol Microbiol Biotechnol. 2004; 7: 204–11. doi:10.1159/000079829 PMID: 15383718
9. Rey MW, Ramaiya P, Nelson BA, Brody-Karpin SD, Zaretsky EJ, Tang M, et al. Complete genome sequence of the industrial bacterium Bacillus licheniformis and comparisons with closely related Bacillus species. Genome Biol. 2004; 5: R77. doi:10.1186/gb-2004-5-10-r77 PMID: 15461803
10. Wiegand S, Dietrich S, Hertel R, Bongaerts J, Evers S, Volland S, et al. RNA-Seq of Bacillus licheniformis: active regulatory RNA features expressed within a productive fermentation. BMC Genomics. 2013; 14: 667. doi:10.1186/1471-2164-14-667 PMID: 24079885
11. Schallmey M, Singh A, Ward OP. Developments in the use of Bacillus species for industrial production. Can J Microbiol. 2004; 50: 1–17. doi:10.1139/w03-076 PMID: 15052317
12. Dawyndt P, Vancanneyt M, De Meyer H, Swings J. Knowledge accumulation and resolution of data inconsistencies during the integration of microbial information sources. IEEE Trans Knowl Data Eng. IEEE COMPUTER SOC; 2005; 17: 1111–1126. doi:10.1109/TKDE.2005.131
13. Steensma HY, Robertson LA, Elsas JD. The occurrence and taxonomic value of PBSX-like defective phages in the genus Bacillus. Antonie Van Leeuwenhoek. 1978; 44: 353–366. doi:10.1007/BF00394312 PMID: 582243
14. Steensma HY, Robertson LA. Lysogeny in Bacillus. FEMS Microbiol Lett. 1978; 3: 313–317. doi:10.1111/j.1574-6968.1978.tb01961.x
15. Bose M, Barber RD. Prophage Finder: a prophage loci prediction tool for prokaryotic genome sequences. In Silico Biol. 2006; 6: 223–7. Available: http://www.ncbi.nlm.nih.gov/pubmed/16922685 PMID: 16922685
16. Sambrook J, Russell DW. Molecular Cloning: A Laboratory Manual, Third Edition (3 volume set). Cold Spring Harbor Laboratory Press; 2001.

17. Waschkau B, Waldeck J, Wieland S, Eichstädt R, Meinhardt F. Generation of readily transformable Bacillus licheniformis mutants. Appl Microbiol Biotechnol. 2008; 78: 181–8. doi:10.1007/s00253-007-1278-0 PMID: 18046550

18. Rachinger M, Bauch M, Strittmatter A, Bongaerts J, Evers S, Maurer K-H, et al. Size unlimited markerless deletions by a transconjugative plasmid-system in Bacillus licheniformis. J Biotechnol. 2013; 167: 365–9. doi:10.1016/j.jbiotec.2013.07.026 PMID: 23916947

19. Quevillon E, Silventoinen V, Pillai S, Harte N, Mulder N, Apweiler R, et al. InterProScan: protein domains identifier. Nucleic Acids Res. 2005; 33: W116–20. doi:10.1093/nar/gki442 PMID: 15980438

20. Rutherford K, Parkhill J, Crook J, Horsnell T, Rice P, Rajandream M-A, et al. Artemis: sequence visualization and annotation. Bioinformatics. 2000; 16: 944–945. doi:10.1093/bioinformatics/16.10.944 PMID: 11120685

21. Rachinger M, Volland S, Meinhardt F, Daniel R, Liesegang H. First Insights into the Completely Annotated Genome Sequence of Bacillus licheniformis Strain 9945A. Genome Announc. 2013; 1. doi:10.1128/genomeA.00525-13

22. Barbe V, Cruveiller S, Kunst F, Lenoble P, Meurice G, Sekowska A, et al. From a consortium sequence to a unified sequence: the Bacillus subtilis 168 reference genome a decade later. Microbiology. 2009; 155: 1758–75. doi:10.1099/mic.0.027839-0 PMID: 19383706

23. Simon R, Priefer U, Pühler A. A Broad Host Range Mobilization System for In Vivo Genetic Engineering: Transposon Mutagenesis in Gram Negative Bacteria. Bio/Technology. Nature Publishing Group; 1983; 1: 784–791. doi:10.1038/nbt1183-784

24. Huang W, Marmur J. Characterization of inducible bacteriophages in Bacillus licheniformis. J Virol. 1970; 5: 237–46. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=375993&tool=pmcentrez&rendertype=abstract PMID: 4988269

25. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. Nature Publishing Group, a division of Macmillan Publishers Limited. All Rights Reserved.; 2012; 9: 357–9. doi:10.1038/nmeth.1923 PMID: 22388286

26. Dietrich S, Wiegand S, Liesegang H. TraV: A Genome Context Sensitive Transcriptome Browser. PLoS One. 2014; 9: e93677. doi:10.1371/journal.pone.0093677 PMID: 24709941

27. Holzenburg A, Hoppert M. Electron Microscopy in Microbiology (Royal Microscopical Society Microscopy Handbooks). Garland Science; 1998. p. 112.

28. Valentine RC, Shapiro BM, Stadtman ER. Regulation of glutamine synthetase. XII. Electron microscopy of the enzyme from Escherichia coli. Biochemistry. 1968; 7: 2143–2152. doi:10.1021/bi00846a017 PMID: 4873173

29. Rice P, Ian L, Bleasby A. The European Molecular Biology Open Software Suite EMBOSS: The European Molecular Biology Open Software Suite. Trends Genet. 2000; 16: 276–277. PMID: 10827456

30. Pope WH, Jacobs-Sera D, Best AA, Broussard GW, Connerly PL, Dedrick RM, et al. Cluster j mycobacteriophages: intron splicing in capsid and tail genes. PLoS One. 2013; 8: e69273. doi:10.1371/journal.pone.0069273 PMID: 23874930

31. Panis G, Franche N, Méjean V, Ansaldi M. Insights into the functions of a prophage recombination directionality factor. Viruses. 2012; 4: 2417–31. doi:10.3390/v4112417 PMID: 23202488

32. Kunst F, Ogasawara N, Moszer I, Albertini AM, Alloni G, Azevedo V, et al. The complete genome sequence of the gram-positive bacterium Bacillus subtilis. Nature. 1997; 390: 249–56. doi:10.1038/36786 PMID: 9384377

33. Schäfer A, Schwarzer A, Kalinowski J, Pühler A. Cloning and characterization of a DNA region encoding a stress-sensitive restriction system from Corynebacterium glutamicum ATCC 13032 and analysis of its role in intergeneric conjugation with Escherichia coli. J Bacteriol. 1994; 176: 7309–19. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=197120&tool=pmcentrez&rendertype=abstract PMID: 7961503

34. Baumgart M, Unthan S, Rückert C, Sivalingam J, Grünberger A, Kalinowski J, et al. Construction of a prophage-free variant of Corynebacterium glutamicum ATCC 13032 for use as a platform strain for basic research and industrial biotechnology. Appl Environ Microbiol. 2013; 79: 6006–15. doi:10.1128/AEM.01634-13 PMID: 23892752

35. Krogh S, O'Reilly M, Nolan N, Devine KM. The phage-like element PBSX and part of the skin element, which are resident at different locations on the Bacillus subtilis chromosome, are highly homologous. Microbiology. 1996; 142: 2031–2040. doi:10.1099/13500872-142-8-2031 PMID: 8760915

36. Meyer BJ, Kleid DG, Ptashne M. Lambda repressor turns off transcription of its own gene. Proc Natl Acad Sci U S A. 1975; 72: 4785–89. Available: http://www.ncbi.nlm.nih.gov/pubmed/1061069 PMID: 1061069

37. McDonnell GE, Wood H, Devine KM, McConnell DJ. Genetic control of bacterial suicide: regulation of the induction of PBSX in Bacillus subtilis. J Bacteriol. 1994; 176: 5820–30. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=196787&tool=pmcentrez&rendertype=abstract PMID: 8083174

38. Dhaese P, Dobbelaere M-R, Van Montagu M. The temperate B. subtilis phage Φ105 genome contains at least two distinct regions encoding superinfection immunity. Mol Genet Genomics. 1985; 200: 490–492.

39. Wood HE, Dawson MT, Devine KM, McConnell DJ. Characterization of PBSX, a defective prophage of Bacillus subtilis. J Bacteriol. 1990; 172: 2667–74. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=208911&tool=pmcentrez&rendertype=abstract PMID: 2110147

40. Kimura T, Amaya Y, Kobayashi K, Ogasawara N, Sato T. Repression of sigK intervening (skin) element gene expression by the CI-like protein SknR and effect of SknR depletion on growth of Bacillus subtilis cells. J Bacteriol. 2010; 192: 6209–16. doi:10.1128/JB.00625-10 PMID: 20889742

41. Krogh S, Jørgensen ST, Devine KM. Lysis genes of the Bacillus subtilis defective prophage PBSX. J Bacteriol. 1998; 180: 2110–7. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=107137&tool=pmcentrez&rendertype=abstract PMID: 9555893

42. Needleman SB, Wunsch CD. A general method applicable to the search for similarities in the amino acid sequence of two proteins. J Mol Biol. 1970; 48: 443–453. doi:10.1016/0022-2836(70)90057-4 PMID: 5420325

43. Ackermann H-W, Prangishvili D. Prokaryote viruses studied by electron microscopy. Arch Virol. 2012; 157: 1843–9. doi:10.1007/s00705-012-1383-y PMID: 22752841

44. Casjens SR, Thuman-Commike PA. Evolution of mosaically related tailed bacteriophage genomes seen through the lens of phage P22 virion assembly. Virology. 2011; 411: 393–415. doi:10.1016/j.virol.2010.12.046 PMID: 21310457

45. Ackermann H-W, Krisch HM, Comeau AM. Morphology and genome sequence of phage Φ1402: A dwarf myovirus of the predatory bacterium Bdellovibrio bacteriovorus. Bacteriophage. 2011; 1: 138–142. doi:10.4161/bact.1.3.15769 PMID: 22164347

46. Petrovski S, Dyson ZA, Seviour RJ, Tillett D. Small but sufficient: the Rhodococcus phage RRH1 has the smallest known Siphoviridae genome at 14.2 kilobases. J Virol. 2012; 86: 358–63. doi:10.1128/JVI.05460-11 PMID: 22013058

47. Haas M, Yoshikawa H. Defective bacteriophage PBSH in Bacillus subtilis. I. Induction, purification, and physical properties of the bacteriophage and its deoxyribonucleic acid. J Virol. 1969; 3: 233–47. Available: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=375757&tool=pmcentrez&rendertype=abstract PMID: 4975368

48. Jin T, Zhang X, Zhang Y, Hu Z, Fu Z, Fan J, et al. Biological and genomic analysis of a PBSX-like defective phage induced from Bacillus pumilus AB94180. Arch Virol. 2014; 159: 739–52. doi:10.1007/s00705-013-1898-x PMID: 24154951

49. Briggs GS, Smits WK, Soultanas P. Chromosomal replication initiation machinery of low-G+C-content Firmicutes. J Bacteriol. 2012; 194: 5162–70. doi:10.1128/JB.00865-12 PMID: 22797751

50. Utter B, Deutsch DR, Schuch R, Winer BY, Verratti K, Bishop-Lilly K, et al. Beyond the chromosome: the prevalence of unique extra-chromosomal bacteriophages with integrated virulence genes in pathogenic Staphylococcus aureus. PLoS One. 2014; 9: e100502. doi:10.1371/journal.pone.0100502 PMID: 24963913

51. Sozhamannan S, Chute MD, Mcafee FD, Fouts DE, Akmal A, Galloway DR, et al. The Bacillus anthracis chromosome contains four conserved, excision-proficient, putative prophages. BMC Genomics. 2006; 6. doi:10.1186/1471-2180-6-34

52. Kawasaki T, Nagata S, Fujiwara A, Satsuma H, Fujie M, Usami S, et al. Genomic characterization of the filamentous integrative bacteriophages φRSS1 and φRSM1, which infect Ralstonia solanacearum. J Bacteriol. 2007; 189: 5792–802. doi:10.1128/JB.00540-07 PMID: 17557818

Supplementary information

Supplementary information for this manuscript can be found at PLOS ONE website under the following address:
http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0120759#sec026

Additionally, supplementary figures and tables are provided along with the electronic version of this thesis (on DVD), under the following paths:

Additional Figures:

S1 Figure Supplementary Information/Chapter II.3/Figures/S1 Figure.tif
S2 Figure Supplementary Information/Chapter II.3/Figures/S2 Figure.tif
S3 Figure Supplementary Information/Chapter II.3/Figures/S3 Figure.tif
S4 Figure Supplementary Information/Chapter II.3/Figures/S4 Figure.tif
S5 Figure Supplementary Information/Chapter II.3/Figures/S5 Figure.tif

Additional Tables:

S1 Table Supplementary Information/Chapter II.3/Tables/S1 Table.xlsx
S2 Table Supplementary Information/Chapter II.3/Tables/S2 Table.xlsx

II.4 Host-Pathogen Coevolution: The Selective Advantage of *Bacillus thuringiensis* Virulence and It's Cry Toxin Genes

PLoS Biology

Host-Pathogen Coevolution: The Selective Advantage of *Bacillus thuringiensis* Virulence and Its Cry Toxin Genes

Leila Masri, Antoine Branca, Anna E. Sheppard, Andrei Papkou, David Laehnemann, Patrick S. Guenther, Swantje Prahl, Manja Saebelfeld, **Jacqueline Hollensteiner**, Heiko Liesegang, Elzbieta Brzuszkiewicz, Rolf Daniel, Nicolaas K. Michiels, Rebecca D. Schulte, Joachim Kurtz, Philip Rosenstiel, Arndt Telschow, Erich Bornberg-Bauer, Hinrich Schulenburg

PLoS Biology (2015), 13(6):e1002169

Author contributions

Performed the experiments: LM, AB, AES, AP, DL, PSG, SP, MS, **JH**

Analyzed the data: LM, AB, AES, AP, and HS.

Wrote the paper: LM, AB, AES, AP, **JH,** EB, HL, RD, NKM, RDS, JK, PR, AT, EBB and HS

Supported pathogen whole genome analysis: HL, EB and RD

Contributed to design and analysis of evolution experiment: NKM, RDS, JK, AT

Contributed to genome data analysis: PR and EBB

Conceived and designed the experiments: LM AB AES PR EBB HS.

PLOS | BIOLOGY

RESEARCH ARTICLE

Host–Pathogen Coevolution: The Selective Advantage of *Bacillus thuringiensis* Virulence and Its Cry Toxin Genes

Leila Masri[1,2☯¤a], Antoine Branca[3☯¤b], Anna E. Sheppard[1☯¤c], Andrei Papkou[1], David Laehnemann[1,2], Patrick S. Guenther[2], Swantje Prahl[1], Manja Saebelfeld[1], Jacqueline Hollensteiner[4], Heiko Liesegang[4], Elzbieta Brzuszkiewicz[4], Rolf Daniel[4], Nicolaas K. Michiels[2], Rebecca D. Schulte[5], Joachim Kurtz[3], Philip Rosenstiel[6], Arndt Telschow[3], Erich Bornberg-Bauer[3], Hinrich Schulenburg[1,2]*

1 Department of Evolutionary Ecology and Genetics, Zoological Institute, Christian-Albrechts-University of Kiel, Kiel, Germany, **2** Department of Animal Evolutionary Ecology, Institute of Evolution and Ecology, University of Tuebingen, Tuebingen, Germany, **3** Institute for Evolution and Biodiversity, University of Muenster, Muenster, Germany, **4** Goettingen Genomics Laboratory, Institute of Microbiology and Genetics, Georg-August-University of Goettingen, Goettingen, Germany, **5** Department of Behavioural Biology, University of Osnabrueck, Osnabrueck, Germany, **6** Institute for Clinical Molecular Biology, Christian-Albrechts-University, Kiel, Germany

☯ These authors contributed equally to this work.
¤a Current address: Institute of Science and Technology Austria, Klosterneuburg, Austria
¤b Current address: Laboratory Ecologie, Systématique et Evolution, CNRS-Univ Paris-Sud, UMR8079, Orsay, France
¤c Current address: Nuffield Department of Medicine, University of Oxford, Oxford, United Kingdom
* hschulenburg@zoologie.uni-kiel.de

OPEN ACCESS

Citation: Masri L, Branca A, Sheppard AE, Papkou A, Laehnemann D, Guenther PS, et al. (2015) Host–Pathogen Coevolution: The Selective Advantage of *Bacillus thuringiensis* Virulence and Its Cry Toxin Genes. PLoS Biol 13(6): e1002169. doi:10.1371/journal.pbio.1002169

Academic Editor: David S. Schneider, Stanford University, UNITED STATES

Received: August 3, 2014

Accepted: May 7, 2015

Published: June 4, 2015

Data Availability Statement: All of the relevant data can be found within the paper and its Supporting Information files, with the exception of Genome sequence data which are deposited at the European Nucleotide Archive http://www.ebi.ac.uk/ena/ (study accession number PRJEB5931).

Funding: We are very grateful for funding from the German Science Foundation (DFG) to HS (SCHU 1415/8, SCHU 1415/9), PR (RO 2994/3), EBB (BO 2544/7), HL (LI 1690/2), AT (TE 976/2), RDS (SCHU 2522/1), JK (KU 1929/4); from the Kiel Excellence Cluster Inflammation at Interfaces to HS and PR; and

Abstract

Reciprocal coevolution between host and pathogen is widely seen as a major driver of evolution and biological innovation. Yet, to date, the underlying genetic mechanisms and associated trait functions that are unique to rapid coevolutionary change are generally unknown. We here combined experimental evolution of the bacterial biocontrol agent *Bacillus thuringiensis* and its nematode host *Caenorhabditis elegans* with large-scale phenotyping, whole genome analysis, and functional genetics to demonstrate the selective benefit of pathogen virulence and the underlying toxin genes during the adaptation process. We show that: (i) high virulence was specifically favoured during pathogen–host coevolution rather than pathogen one-sided adaptation to a nonchanging host or to an environment without host; (ii) the pathogen genotype BT-679 with known nematocidal toxin genes and high virulence specifically swept to fixation in all of the independent replicate populations under coevolution but only some under one-sided adaptation; (iii) high virulence in the BT-679-dominated populations correlated with elevated copy numbers of the plasmid containing the nematocidal toxin genes; (iv) loss of virulence in a toxin-plasmid lacking BT-679 isolate was reconstituted by genetic reintroduction or external addition of the toxins. We conclude that sustained coevolution is distinct from unidirectional selection in shaping the pathogen's genome and life history characteristics. To our knowledge, this study is the first to characterize the pathogen genes involved in coevolutionary adaptation in an animal host–pathogen interaction system.

from the ISTFELLOW program (Co-fund Marie Curie Actions of the European Commission) to LM. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: AP is an associate member of the International Max-Planck Research School for Evolutionary Biology at the University of Kiel. The authors have declared that no other competing interests exist.

Abbreviations: AMOVA, analysis of molecular variance; FDR, false discovery rate.

Author Summary

Evolution can be extremely fast and dramatic, especially when infectious disease agents such as bacterial pathogens engage in a continuous arms race with their host organism. Rounds of novel pathogen attack strategies and associated host counterdefenses conspire to drive host–pathogen coevolution and biological innovation. To better understand the underlying genetic mechanisms and the exact trait characteristics under selection, we conducted experimental evolution using a simple host–pathogen model system (nematode versus bacterium) under controlled laboratory conditions. We analysed the associated adaptive changes in real time using large-scale phenotyping, population whole genome sequencing, and genetic analysis of the identified candidate genes. We show that coevolution (rather than one-sided adaptation) particularly favors and maintains pathogen virulence, and that two specific toxin genes significantly influence this virulence during coevolution.

Introduction

Antagonisms are often at the heart of rapid evolutionary change. One prime example for such antagonism is given by the interaction between host and pathogen. By definition, pathogens have a negative effect on host fitness, favouring selection of enhanced defence mechanisms in the affected hosts. If pathogen fitness depends on the host, then host defence can be detrimental for the pathogen, leading to selection for novel attack mechanisms. When the interaction persists over time, the ongoing cycles of adaptation and counteradaptation can produce one of the highest selective pressures known in nature [1–4]. There are numerous examples of the resulting rapid evolutionary responses during host–pathogen coevolution, including taxonomically diverse host systems such as bacteria [5,6], plants [7–9], invertebrates [10–13], and vertebrate animals [14]. In spite of its potential importance as a major driver of evolution, two core features of the coevolutionary dynamics are as yet only poorly understood [3,15]: (i) which trait functions are specifically under selection during coevolution when antagonists reciprocally coadapt to each other, rather than only one adapting while the other remains unchanged? (ii) Which genes and genetic mechanisms underlie adaptation during coevolution, particularly when rapid changes are required to keep up with the coevolving antagonist?

To date, very few studies have evaluated the selective consequences of coevolution relative to one-sided adaptation, and these have mainly used bacteria–phage interaction models [16–19]. For instance, the *Pseudomonas fluorescens*-infecting phage Φ2 was evolved in the presence of either its coevolving host or a nonevolving host, leading to the unique emergence and persistence of different phage infection varieties under the coevolution conditions [19]. Similarly, it is mainly bacteria–phage systems that have previously been used to dissect the genetics of host–pathogen coevolutionary change [16–18,20]. One of the recent examples was based on controlled coevolution of phage λ and its host *Escherichia coli*, resulting in a sequence of reciprocal adaptations in the phage to use different host receptors and in the host to alter the targeted receptors or prevent uptake of phage DNA [20]. Similar information for multicellular host interaction models is as yet scarce (e.g., [13]). Such information is essential for a more general understanding of the postulated impact of coevolution on the evolution of organisms [1,2,4] and, more precisely, it will help us understand which exact trait functions are under selection during sustained, reciprocally antagonistic interactions, how quickly changes can be achieved during evolution, and what the most likely resulting selection dynamics are [2,3,15].

Here, we addressed these questions by comparing alternative selection regimes in controlled evolution experiments, using an animal host interaction model, consisting of the nematode

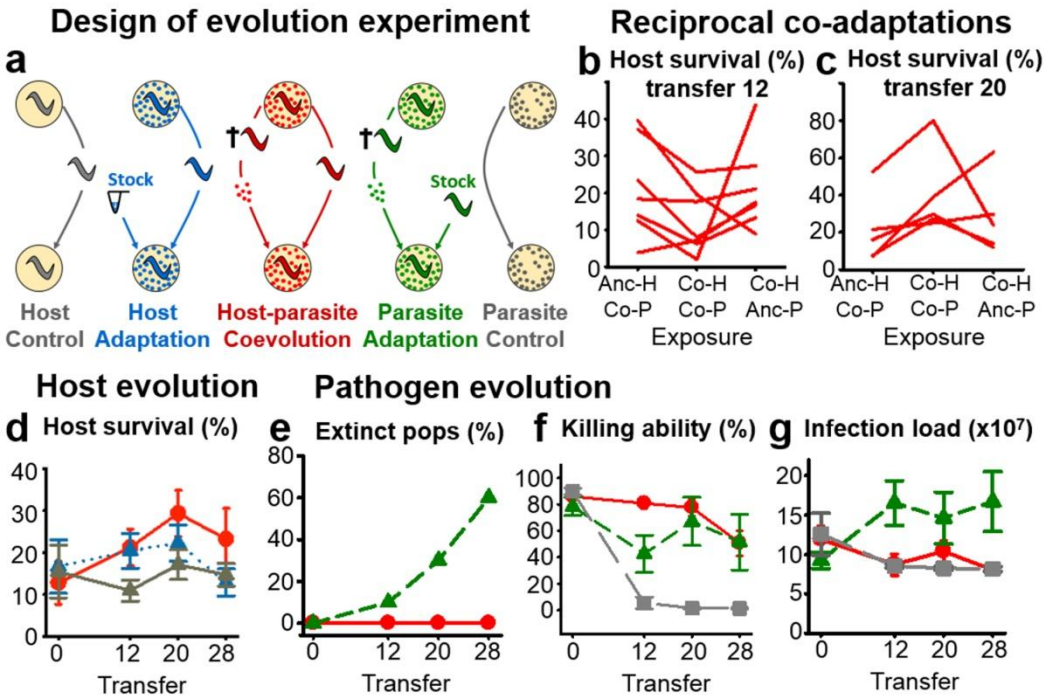

Fig 1. Experimental host–pathogen coevolution causes phenotypic changes in both antagonists. A, The five evolution treatments: (i) host control (grey) adapting to general laboratory conditions in the absence of the pathogen, (ii) host one-sided adaptation (blue) where the host adapted to the nonevolving, ancestral pathogen taken from a frozen stock culture at each transfer, (iii) host–pathogen coevolution (red) during which both antagonists were continuously forced to coevolve to each other, (iv) pathogen one-sided adaptation (green) where the pathogen adapted to the nonevolving, ancestral host population taken from a frozen stock culture at each transfer; and (v) pathogen control (grey) adapting to general laboratory conditions in the absence of the host. **B–C,** Analysis of reciprocal coadaptations in host survival and pathogen killing ability (*y*-axis) by comparing (along the *x*-axis) exposures of coevolved hosts with coevolved pathogens from the same replicate population and time point (indicated by Co-H Co-P in the middle of the panels) with either coevolved hosts from the same replicate exposed to ancestral pathogens (Co-H Anc-P, right side) or ancestral hosts exposed to coevolved pathogens from the same replicate (Anc-H Co-P, left side). Results are given for transfers 12 (**B**) and 20 (**C**) separately. The lines connect the results for particular replicate populations of the coevolution treatment. **D,** Survival of evolved host populations from different treatments (colors as in Fig 1A) upon exposure to the ancestral pathogen. **E,** Pathogen population extinctions under one-sided adaptation (green) and coevolution (red). **F–G,** Analysis of evolved pathogen populations from different treatments (colors as in Fig 1A) upon exposure to the ancestral host, including pathogen killing ability (measured as host death rate in %) (**F**) and pathogen infection load (**G**). Bars denote standard error. The original data is provided in S1 Data, and the results of the corresponding statistical analyses are given in S1 Table, S2 Table, S3 Table, S4 Table, and S5 Table.

doi:10.1371/journal.pbio.1002169.g001

Caenorhabditis elegans as host and its natural pathogen *Bacillus thuringiensis* [21]. The Gram-positive bacterium *B. thuringiensis* is of economic importance as a pest control agent [22,23] and infects insect or nematode hosts upon oral uptake via toxin-mediated destruction of intestinal cells and expression of additional virulence factors [22,23]. The interaction between *B. thuringiensis* and *C. elegans* was previously established as an experimental evolution model for studying the consequences of coevolution [21,24–27]. We have now used this interaction model for a new experimental design that consisted of five distinct evolution treatments (Fig 1A; see Materials and Methods), namely: (i) host control, where the host adapted to general laboratory conditions in the absence of pathogenic *B. thuringiensis*; (ii) host one-sided adaptation, where the host adapted to the ancestral pathogenic *B. thuringiensis*, taken from a frozen stock at each transfer step; (iii) host–pathogen coevolution, where both host and pathogen co-adapted to their continuously coevolving antagonist; (iv) pathogen one-sided adaptation, where the pathogen adapted to the ancestral *C. elegans* population, taken from a frozen stock at

each transfer step; and (v) pathogen control, where the pathogen adapted to general laboratory conditions in the absence of a host. The evolution experiment was specifically designed to provide identical conditions for the different treatments, except for the presence of a coevolving antagonist, a nonevolving antagonist (i.e., the ancestral antagonist), or no antagonist. Thus, the design allowed us to assess the unique consequences of coevolutionary adaptation, as opposed to one-sided adaptation and laboratory adaptation. The *C. elegans–B. thuringiensis* system was chosen because it enables a high level of control over the evolutionary interaction and subsequent analyses. This is due to the fact that *C. elegans* and *B. thuringiensis* can be efficiently purified from each other during the transfer steps, preventing any unintended coevolution from occurring in the evolution treatments, and both can be cryopreserved. This cryopreservation allows identical cultures of the stock ancestral antagonists to be used for the respective one-sided adaptation treatment throughout the entire experiment and enables evolved populations to be frozen for later parallel phenotypic characterizations.

Based on this experimental setup, we compared the consequences of reciprocal coevolution with those resulting from related selective pressures. We here present our findings on the evolved phenotypic changes across time, combined with results from population whole genome sequence analysis and a subsequent functional genetic assessment.

Results/Discussion

Differences in Experimental Selection Conditions Lead to Multiple Distinct Changes in Host and Pathogen

After completion of the evolution experiment (Fig 1A), phenotypic changes were assessed for both host and pathogen. Evolved *B. thuringiensis* and *C. elegans* material, which was frozen at regular intervals during the evolution experiment (see Materials and Methods), was thawed and examined in parallel to assess phenotypic variation across time and evolution treatments. We considered a total of 86 evolved pathogen and 91 evolved host populations (covering three evolution treatments for each antagonist, with three transfer time points per treatment and up to ten independent replicate populations per treatment per time point), as well as the ancestral populations. For these, traits of relevance for the interaction were studied in the presence of either the coevolving antagonist from the same replicate population at a specific time point or the respective ancestral antagonist (see below). We characterized several proxies for host resistance sensu lato (survival, body size, population size, and infection load), proxies for pathogen virulence sensu lato (killing ability and pathogen effects on host body size and host population size), and pathogen infection load (see Materials and Methods). We additionally assessed the ability of the pathogen to form biofilms and the resulting competitive advantage on different nutritional agar media (Materials and Methods), as biofilm formation was found to be common in some of the evolution treatments. Our statistical analysis of the obtained data always included adjustment of significance levels through the false discovery rate (FDR) [28] to account for increased type I errors.

Our analysis revealed specific coadaptations between host and pathogen during the coevolution treatment. In particular, we compared the survival of coevolved hosts exposed to coevolved pathogens from the same time point and replicate to that of coevolved hosts from the same replicate, which were exposed to ancestral pathogens, or to that of coevolved pathogens from the same replicate, which were exposed to ancestral hosts. For this analysis, we included all replicates, for which data was available for all three types of exposures (seven for transfers 12 and 28; five for transfer 20). In the absence of reciprocal coadaptations, we would expect the coevolved–coevolved combinations to produce phenotypic values that are either identical to or the average of those from the coevolved–ancestral combinations. This is clearly not the case for

transfers 12 and 20, for which the coevolved–coevolved combinations consistently produced either lower (for transfer 12) or higher (for transfer 20) survival than the corresponding exposures to the ancestral antagonists (Fig 1B and 1C). For these transfers, the comparisons are significantly different except for the comparison to the coevolved pathogens from transfer 20 exposed to the ancestral host, which still indicated a statistical trend ($p < 0.1$; S1 Table). At transfer 28, we could not identify any significant variation. We conclude that the experimental evolution conditions allowed the antagonists to reciprocally coadapt to each other within considerably short time periods of only 12 transfers.

We next assessed variation among evolution treatments for the host. Comparability of populations from different treatments was ensured by exposing all evolved hosts to the ancestral pathogens, followed by measurement of the various phenotypic traits. Our results showed that survival in the presence of the ancestral pathogen significantly increased during coevolution but not in other treatment conditions, suggesting evolution of increased host resistance sensu lato in the presence of a coevolving antagonist (Fig 1D, S2 Table, and S3 Table), generally consistent with above analysis of reciprocal coadaptations. None of the other considered traits showed significant variation (S2 Table, S3 Table).

We performed a similar comparison for the pathogen, based on the exposure of evolved pathogen populations from the various treatments to the ancestral host population. Intriguingly, pathogen one-sided adaptation caused extinction (i.e., no host killing, as pathogens were only transferred from dead hosts) in more than half of the pathogen populations. This was not observed under coevolution conditions (Fig 1E). The pathogen's ability to kill the host and reduce host body size and host population size (i.e., virulence sensu lato) was generally maintained during coevolution, while it was decreased transiently under pathogen one-sided adaptation and lost during pathogen control evolution (Fig 1F, S1 Fig, and S4 Table). The observed variation in killing ability was significant among all evolution treatments, while the pathogen's effect on host body size differed significantly between the control and each of the other treatments, and the pathogen's effect on host population size only showed a significant difference between the coevolution and control treatments (S5 Table). In contrast, pathogen infection load was significantly higher under one-sided adaptation than the other two treatments, which did not show significant variation from each other (Fig 1G, S4 Table, and S5 Table).

We subsequently assessed the bacteria's ability to form biofilms. Biofilm formation is a dynamic process, including (i) accumulation of planktonic bacterial cells; (ii) maturation of the bacterial community and first differentiation of cell types; (iii) production of a robust extracellular matrix by specifically differentiated cells (i.e., biofilm production); (iv) disintegration of an aged biofilm; and (v) dispersal of planktonic cells that emerge from the biofilm [29]. Biofilm formation was characterized using independent qualitative and quantitative measurements (Materials and Methods). The qualitative analysis revealed that biofilm formation was significantly more common during control evolution, being lost under coevolution and, to a lesser extent, one-sided adaptation (Fig 2A and 2B, S2 Fig, and S6 Table). Two separate quantitative analyses of individual clones or evolved populations consistently revealed that biofilm particle size was larger for control-evolved and avirulent bacteria (Fig 2C and 2D). To further investigate the causes of treatment variation in biofilm formation, we tested the bacteria's competitive ability under either low nutrient conditions (as used during experimental evolution; Materials and Methods) or high nutrient conditions. We directly competed two bacterial clones in a paired setup, either (i) a biofilm-forming versus a non-biofilm-forming clone (test comparison), (ii) two biofilm-forming clones (first control), and (iii) two non-biofilm-forming clones (second control). We found that the competiveness of biofilm-producing clones was significantly higher on the nutrient-poor medium but significantly lower under nutrient-rich conditions (Fig 2E, S7 Table). Our results thus suggest that the low nutrient medium used during the

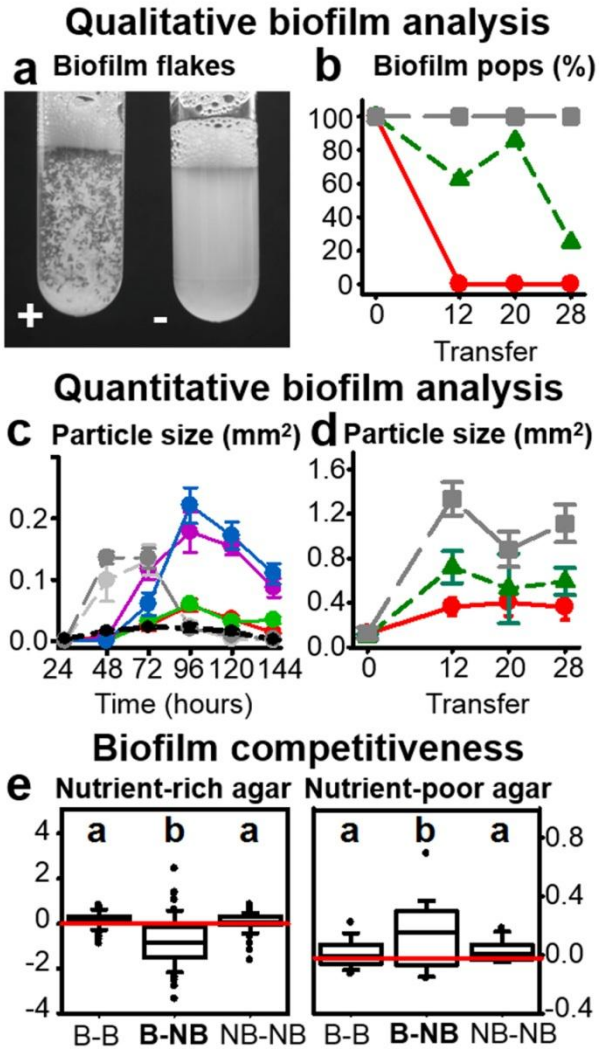

Fig 2. Variation of evolved *B. thuringiensis* in biofilm formation. A, Qualitative assessment of the presence (indicated by +) or absence (-) of biofilm flakes in evolved bacterial populations, washed off assay plates after 48 h of growth and inspected by eye in tubes. **B,** Ability of *B. thuringiensis* populations to form biofilms (% of populations) using the qualitative assay of Fig 2A. Red indicates coevolution, green one-sided adaptation, and grey control evolution. **C,** Quantification of the temporal dynamics of biofilm formation by measuring mean particle size during bacterial growth on plates across time; results for four evolved clones and three ancestral strains; grey shades indicate three of the ancestral strains (light grey: BT-247; dark grey: BT-246; black: BT-679), red a coevolved clone, green a one-sided adapted clone that is able to form biofilms, purple a one-sided adapted clone unable to form biofilms, and blue a non-biofilm-forming control-evolved clone. **D,** Similar quantification of mean biofilm particle size after growth on plates for 96 h for the evolved populations across transfers from the evolution experiment. Red indicates coevolution, green one-sided adaptation, and grey control evolution. **E**, Competitive ability of biofilm-forming (B) versus non-biofilm-forming (NB) clones on nutrient-rich nematode growth medium (left) or nutrient-poor peptone-free medium (right). Two clones always competed with each other in a paired setup. The value for the second listed phenotype was subtracted from the value for the first listed phenotype (e.g., B–NB, value for biofilm-producer minus value for non-biofilm-producer; see combinations on *x*-axis) to calculate a competitiveness index (*y*-axis). The red horizontal line indicates a value of zero (i.e., no difference). Different letters on top indicate significant variation between the different combinations. The original data is given in S2 Data and the corresponding statistical results in S6 Table and S7 Table.

doi:10.1371/journal.pbio.1002169.g002

main part of the evolution protocol (Materials and Methods) selectively favours the maintenance of biofilm-forming bacteria, which is of particular relevance in the absence of a host, as under the control conditions.

Taken together, our phenotypic analysis reveals that the different selection conditions of the evolution experiment result in distinct phenotypic changes in both antagonists, especially for the pathogen. Our results thus extend the few reports, all based on bacteria–phage interaction models, which previously contrasted coevolution with one-sided adaptation for the pathogen and identified unique changes under coevolutionary adaptation [16–19]. Importantly, our new findings now suggest that the different selection conditions favour pathogen characteristics that are likely of relevance during different phases of its life cycle and that may also increase extinction risk in the absence of a coevolving host (i.e., in the one-sided adaptation treatment). At first, prior to host infection, *B. thuringiensis* must ensure persistence in an unfavourable environment [23], for instance through biofilm formation [29,30]. This stage of the life cycle was under specific selection during control evolution without a host, where the pathogen was maintained in a nutrient-poor environment (Fig 2). The ability to form biofilms coincided with a loss of pathogenicity (Figs 1F and 2B), possibly suggesting a life history trade-off. Thereafter, following host entry, *B. thuringiensis* is known to pass through two phases during the infection process [23]. The first phase is characterized by toxin-mediated tissue damage, which weakens the host, easing access to nutritional resources. As toxins ultimately cause host death [22,23], selection on this step may lead to variation in host killing ability. Thus, the toxin effects appear of particular selective benefit during coevolution (Fig 1B, 1C, and 1F), possibly because of ongoing resistance evolution in the host (Fig 1B–1D). The second phase starts when the host is weakened or already dead, and when pathogens increase replication rate [23], leading to elevated infection load. During this latter phase, bacteria may have an advantage if they do not pay the cost of toxin production and/or replication of the toxin-containing plasmid and can thus grow faster than toxin-producing and/or plasmid-bearing cells [23,31]. This step appears to be under particular selection in the one-sided adaptation treatment, for which a significantly higher infection load was recorded (Fig 1G). Moreover, the duration of this second phase in the pathogen one-sided adaptation treatment (as determined by the characteristics of the ancestral host population used) seems to represent a "turning point" where chance can favour either toxin-bearing or toxin-lacking bacteria, as demonstrated below by the results of the genomic analysis (see below and Figs 3C and 4) and indicated also by the overall decreased virulence for this treatment (Fig 1F). If an avirulent genotype spreads to fixation in one of these populations, the bacteria would no longer be able to infect new hosts because of the absence of toxins, thus potentially explaining the high extinction rate under these conditions (Fig 1F; see below for further details).

Differences in Experimental Selection Conditions Favour Distinct Pathogen Genotypes

Genetic changes in the pathogen were explored through whole genome sequence analysis and a toxin gene screen of *B. thuringiensis* populations from three time points (transfers 0, 12, and 20). As a basis for our analysis, we first assembled reference genomes for five strains present in the ancestral population and established a novel analysis pipeline that ensured reliable variant detection in genetically variable populations (Fig 3A, S3 Fig, S4 Fig, S8 Table, S9 Table, Materials and Methods). This pipeline was used to analyse a total of 56 whole population genomes (three evolution treatments and two time points, with up to ten replicate populations each, plus the ancestral population). Based on the analysis, we identified a dramatic change in strain composition from the ancestral to the evolved populations. While several strains were

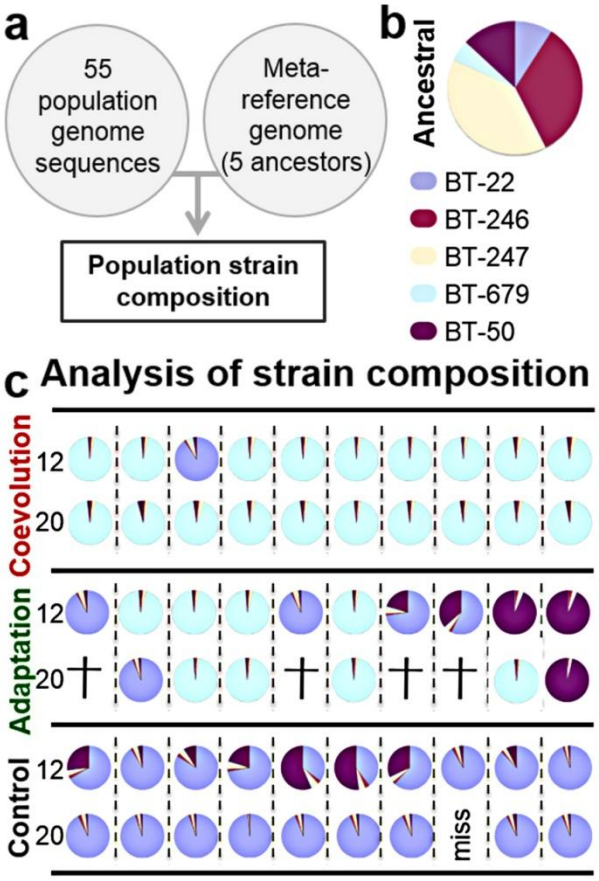

Fig 3. Broad-scale genomic analysis reveals clonal selection during experimental evolution. A, Genome analysis workflow: A metareference genome created from five genomes representative of the ancestral population was used for sequence read mapping and subsequent identification of strain composition for 55 evolved populations. **B–C,** Pie charts show pathogen strain composition of the ancestral and the evolved populations from ten replicates per treatment (horizontal axis) and two time points (transfer 12 and 20). Coloured slices indicate the relative abundance of the various *B. thuringiensis* strains. Crosses indicate extinction of replicates and "miss" that genetic material for the population was unavailable. The data is given in S3 Data.

doi:10.1371/journal.pbio.1002169.g003

abundant at the beginning (Fig 3B), most populations from transfer 12 and all from transfer 20 were dominated by single *B. thuringiensis* strains (Fig 3C, S10 Table). These results are consistent with the idea that bacterial adaptation is commonly determined by strong clonal interference or clonal competition, subsequently leading to rapid fixation of single genotypes (reviewed in [32]). Alternative selective dynamics such as balancing selection, which would have led to the coexistence of several genotypes, do not appear to be involved. The results are thus also in contrast with those from our previous evolution experiment with the same interaction model which, in contrast to the current work, included pathogen immigration, and found an increase in pathogen genotype diversity under coevolution conditions [21,26]. These opposing findings are nevertheless consistent with previous modelling results that immigration can enhance diversity during coevolution [33,34].

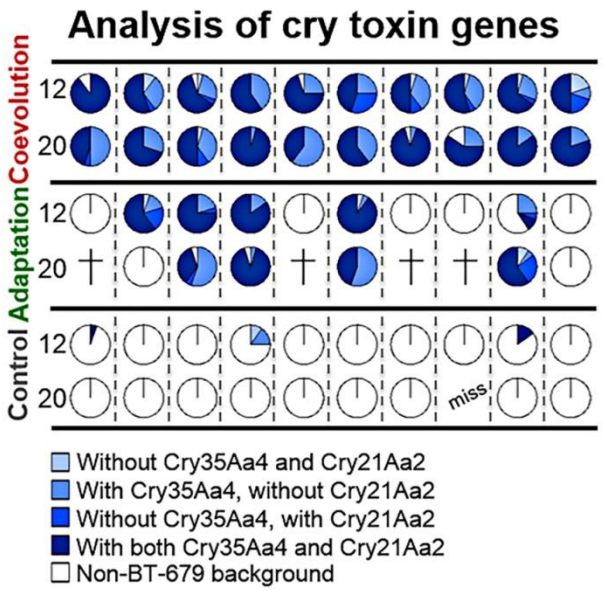

Fig 4. Frequency of BT-679 toxin genes *cry21Aa2* and *cry35Aa4* among the evolved replicate populations. The different shades of blue indicate alternative combinations of toxin genes present, as indicated. The toxin genes were all restricted to evolved clones of the BT-679 background (i.e., horizontal transfer was not detected). The top two rows refer to the coevolved, the middle two rows to one-sided adapted, and the bottom two rows to the control evolved replicate populations. Replicate populations are given along the horizontal axis. Data is shown for both transfer 12 and 20 and a total of 55 replicate populations. Crosses indicate extinction of replicates and "miss" that genetic material for the population was unavailable. The original data is shown in S4 Data.

doi:10.1371/journal.pbio.1002169.g004

At the strain level, we found that almost all coevolved and many one-sided adapted populations showed a high prevalence of BT-679 (Fig 3C), which is known to have stronger nematocidal effects than other pathogenic *B. thuringiensis* strains, such as BT-246 and BT-247, which are both present in the ancestral population [21,35]. Consistent with this observation, we found known nematocidal toxin genes to be almost exclusively restricted to the BT-679 genotype at transfers 12 and 20 and thus specifically enriched under coevolution and, to a lesser extent, one-sided adaptation conditions (Fig 4, S11–S13 Tables). This observation highlights the particular importance of toxin genes during the evolutionary interaction with a host, especially a coevolving host.

Interestingly, about half of the populations from the one-sided adaptation treatment are dominated at transfer 12 by the virulent, toxin-bearing BT-679, whereas the other half are dominated by the avirulent, toxin-lacking genotypes BT-22 and/or BT-50 (Figs 3C and 4). As indicated above, this result is most likely a consequence of the two phases that determine *B. thuringiensis* infection dynamics [22,23]. While the virulent genotypes are likely to have an advantage during the first phase (when the host is invaded and an infection is established), the avirulent genotypes that do not pay the cost of toxin production and/or plasmid replication are likely favoured during the second phase (when the host has already been weakened or killed) [22,23]. It seems that the nonevolving, ancestral host population used in the pathogen one-sided adaptation treatment produces a particular relative length of the two phases that favours both pathogen types to a similar extent over the entire infection cycle. Under these conditions, chance determines which of the two spreads to fixation, as observed across the one-sided

adapted pathogen populations at transfer 12. Moreover, the spread of an avirulent genotype in some populations may then also explain the high extinction rate found for this treatment (Fig 1E). Only populations dominated by avirulent and toxin-lacking genotypes at transfer 12 were extinct at transfer 20, whereas all populations dominated by BT-679 at transfer 12 persisted until transfer 20 (Figs 3C and 4). Similarly, the additional population that went extinct at transfer 28 was dominated by an avirulent genotype at transfer 20 (replicate population 10 on the far right in row 4 of Figs 3C and 4). Consequently, the relative length of the two-phase infection process seems to determine extinction or persistence of pathogenic *B. thuringiensis*, whereby persistence is apparently enhanced in the presence of coadapting host populations, as available in the coevolution treatment.

Specific Genomic Variants Are Selectively Favoured in the BT-679-Dominated Populations

We next assessed whether specific genetic changes were selectively favoured within the 27 BT-679-dominated populations under coevolution and one-sided adaptation conditions. We established two novel complementary analysis pipelines to identify candidate regions under selection based on either: (i) comparisons between coevolution and one-sided adaptation treatments or (ii) correlations between genetic and associated phenotypic variations across all BT-679-dominated populations (see Materials and Methods, Fig 5A, S14–S20 Tables). Genetic changes were surveyed for single nucleotide polymorphisms (SNPs; measured through their individual frequency or their effect on population genetic statistics like θ_W, π, and Tajima's D), structural variations, sequence region copy number, and presence of horizontally transferred fragments (Materials and Methods) [36].

Based on our analysis pipeline, we identified more than 100 significant regions from the treatment comparison and four regions from the correlational analysis (S20 Table, Fig 5, and S5 Fig). The relevance of these candidate regions is difficult to assess because many only contain genes with unknown function. However, three of these regions harbour genes previously implicated in bacterial interactions with a host (see more detailed descriptions in Materials and Methods). One of these refers to an approximately 65 kb region of a large plasmid, for which population genetic measures in a sliding window-based analysis consistently indicate significantly higher variation under coevolution conditions. This plasmid contains putative host-interacting genes encoding toxins, a membrane protein, germination proteins, and an acid phosphatase (Fig 5D, S20 Table).

Variation in the two remaining regions correlated significantly with virulence. One of these regions encompasses a gene with unknown function that contains an mviN domain previously linked to virulence in different pathogens [37–39], and for which the frequency of a deletion correlates negatively with virulence (Fig 5B and 5C, S20 Table). The second region refers to a plasmid with two known nematocidal toxin genes, *cry14Aa1* and *cry21Aa2* [40], for which copy number positively correlates with virulence (Fig 5B and 5C, S20 Table). Copy number variation for this plasmid yielded one of the highest significance levels if compared with the other significant candidate regions (S20 Table), possibly emphasizing its particular relevance for the observed variation in killing ability. Intriguingly, the mviN gene deletion frequency also correlated negatively with the copy number of the toxin-containing plasmid. To reassess this, we performed two types of pairwise analyses of the two genomic variations and killing ability. The nonparametric Spearman's rank correlation test confirmed a significant relationship between the mviN gene deletion frequency and both plasmid copy number and killing ability but not between plasmid copy number and killing ability (S21 Table). However, all three comparisons produced significant associations when assessed with weighted regression analysis, for

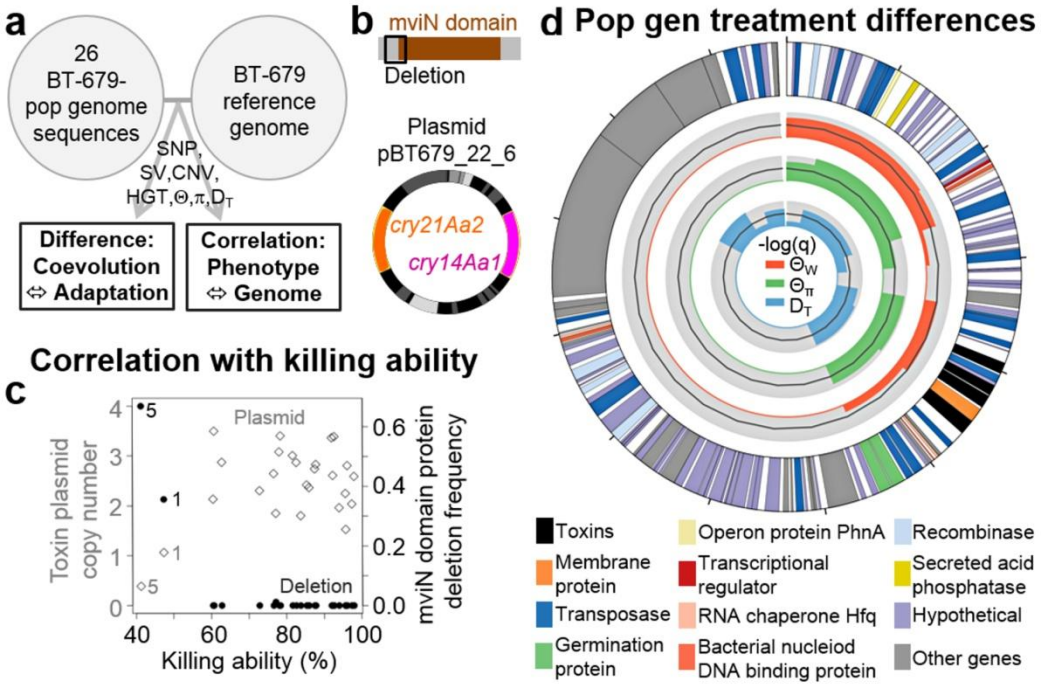

Fig 5. Fine-scale genomics and functional analysis demonstrate importance of nematocidal toxins and other genetic elements during adaptation. **A,** Workflow: Genomic variation of BT-679 populations was contrasted between treatments or correlated with phenotypic variation. **B,** mviN gene deletion and plasmid with cry toxins. **C,** Pathogen killing ability correlates negatively with mviN deletion frequency (left axis, filled circles) and positively with toxin plasmid copy number (right axis, open diamonds). The two most deviating values in all three considered traits were recorded for the same two populations (coevolved populations one and five, both from transfer 20, as indicated adjacent to the measured values), strongly indicating a link between reduced plasmid copy number, increased deletion frequency, and loss of virulence. **D,** Significant variation among the evolution treatments in population genomic statistics for the plasmid Bti_GWDALJX04I0LJH_51–405_fm319.5 (its structure is given in the outer circle). Using a sliding window-based analysis, approximately 65 kb of the plasmid yielded significant ANOVA FDR-corrected q-values, which are shown as-$\log_{10}$(q) on the light grey inner circles as coloured areas for the three inferred population genomic statistics (Θ_W, Θ_π, and D_T; see Materials and Methods). The 5% significance threshold is indicated by the dark grey line within each light grey circle. Thus, coloured areas above this line indicate significant variation among the evolution treatments. This region contains genes encoding for transposases, toxins with unknown effect, a membrane protein, a secreted acid phosphatase, and other proteins (outer circle and legend at the bottom). The total size of the plasmid is about 126 kb. The original data is given in S5 Data, S19 Table, and S20 Table. The results for the statistical analysis is provided in S14–S21 Tables.

doi:10.1371/journal.pbio.1002169.g005

which plasmid copy number was weighted by the inverse of its variance to take into account the accuracy variations in the copy number estimates, which are proportional to the inferred coverage (S21 Table). Even though the associations for plasmid copy numbers are influenced by outliers, our analysis strongly suggests that the three traits covary. Thus, a decrease in virulence in some populations seems to coincide with an increase in mviN deletion frequency and a reduction in plasmid copy number (especially for the coevolved pathogen replicate populations one and five from transfer 20, highlighted in Fig 5C). This tripartite association may then imply that cry toxin abundance (as likely influenced by plasmid copy number) functionally interacts with an intact mviN domain-containing protein to determine virulence. Further functional analysis of this tripartite association, and especially the role of the yet uncharacterized mviN domain-containing gene, represents a particular challenge for the future.

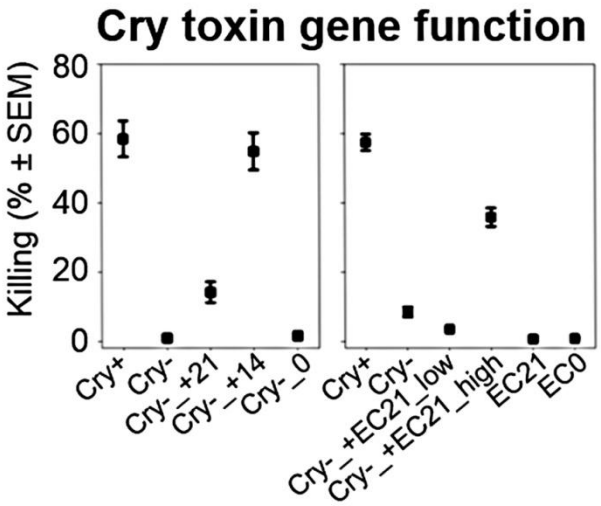

Fig 6. Virulence of BT-679 pathogens with or without nematocidal toxin genes. Mean virulence of plasmid-lacking BT-679 (Cry-) with reintroduced *Cry14Aa1* (+14) or *Cry21Aa2* (+21; left panel) or two concentrations of Cry21Aa2-expressing *E. coli* (+EC21_low, +EC21_high; right panel). Cry+, toxin gene plasmid-bearing BT-679; Cry-_0, empty vector control for BT-679; EC0, empty vector control for *E. coli*. The data is provided in S6 Data.

doi:10.1371/journal.pbio.1002169.g006

The Toxin Genes *cry14Aa1* and *cry21Aa2* Significantly Influence Pathogen Virulence

The nematocidal effects of the above implicated toxin genes *cry14Aa1* and *cry21Aa2* have previously been inferred from their heterologous expression in *E. coli* and subsequent exposure to *C. elegans* and other nematodes [40]. We thus asked whether their presence indeed explains the high killing ability of the strain BT-679, using a functional genetic approach based on a toxin-plasmid-lacking BT-679 variant (see Materials and Methods). Our analysis confirmed that loss of virulence in the toxin-plasmid-lacking variant (denoted Cry- in Fig 6) could be reconstituted to almost wildtype levels (denoted Cry+) by reintroduction of a plasmid with either of the two toxin genes (especially *cry14Aa1*, indicated by Cry-_+14 in Fig 6) or by addition of a high concentration of a Cry21Aa2-expressing *E. coli* (indicated by Cry-_+EC21_high in Fig 6; see also S22–S24 Tables). These results strongly suggest that the two toxin genes, and possibly their copy number, account for the nematocidal effects in BT-679 and may thus have been under positive selection under one-sided adaptation, and especially under during coevolution, where high virulence is particularly favoured by selection (Fig 1B, 1C, and 1F).

Conclusion

To our knowledge, this is the first experimental evolution study that dissects the phenotypic and genomic consequences of coevolution rather than one-sided adaption for a bacterial pathogen, thus extending previous experiments, which were all exclusively based on bacteria–phage interaction models [16–20]. Our results highlight that coevolution particularly favours pathogen virulence, while distinct pathogen life history traits were selected for under one-sided adaptation (i.e., high infection load) or adaptation in the absence of a host (i.e., environmental persistence through biofilm formation). To our knowledge, this study is also the first to dissect the genetic basis of coevolutionary adaptation in a bacterial pathogen. Among the identified

candidate genes, our analysis revealed a particular selective advantage of nematocidal toxin genes and their high copy number during the process of adaptation to a host organism, especially under coevolution conditions. Intriguingly, the selective advantage of high infection load under one-sided adaptation led to the loss of virulent and toxin-bearing *B. thuringiensis* genotypes in some populations, subsequently enhancing pathogen extinction. Our findings thus additionally suggest that the high levels of virulence often observed across *B. thuringiensis* natural isolates [41,42] may only be maintained if the target host is able to coadapt, indicating widespread coevolutionary interactions of *B. thuringiensis* and its various host taxa under natural conditions.

Materials and Methods

C. elegans and *B. thuringiensis* Material

The starting *C. elegans* host population was previously generated through consecutive crosses among 16 natural isolates (PB306, AB1, CB4858, CB4855, N2, JU400, MY16, JU319, PX174, MY1, PX179, JU345, CB4856, CB45507, RC301, and CB4852) [43]. These isolates cover the known worldwide genotype diversity for *C. elegans*. We adapted this genetically diverse population to our experimental conditions by maintaining it for ten generations at 19°C in 40 replicates in the presence of a nonpathogenic *B. thuringiensis* strain (DSM-350). This adaptation step served to minimize potential artifacts in the results of the main evolution experiment caused by predominance of environmental selection unrelated to the host–parasite interaction. These laboratory-adapted populations were mixed and cryopreserved in glycerol at −80°C in 200 aliquots (each containing approximately 5,000 worms) for later use in the main evolution experiment. Note that *C. elegans* larvae survive cryopreservation, thus allowing storage of worm populations for subsequent applications [44]. For all phenotypic experiments, hermaphroditic fourth instar larvae (L4) were used.

The starting pathogen population was similar to that used in our previous experiment [21] and consisted of a mixture of genotypes of the Gram-positive *B. thuringiensis*, including as dominant genotypes the strains MYBT18246 and MYBT18247 (both at an abundance of more than 10%; referred to hereafter and in the main text as BT-246 and BT-247, respectively) and also MYBT18679, MYBT22, and MYBT50 (less than 10% and more than 1%; referred to as BT-679, BT-22, and BT-50, respectively; see also Fig 3B of the main text). The host control treatment (see below) contained the non-nematocidal *B. thuringiensis* strain DSM-350. Prior to the evolution experiment, large quantities of *B. thuringiensis* cultures were prepared, aliquotted and conserved at −20°C for later use [21]. In all experiments, *B. thuringiensis* was used at a final concentration of 1.2×10^8 particles/ml, always mixed with the standard *C. elegans* food source *E. coli* OP50 (final *E. coli* concentration of 2×10^9 cells/ml).

Experimental Evolution

The evolution experiment consisted of five treatments (Fig 1A): (i) host control, during which the host adapted to general laboratory conditions in the absence of pathogenic *B. thuringiensis*; (ii) host one-sided adaptation, where the host was allowed to adapt to a nonevolving pathogenic *B. thuringiensis* taken from a frozen stock culture at each transfer step; (iii) host–pathogen coevolution, in which both antagonists were continuously forced to coevolve with each other; (iv) pathogen one-sided adaptation, where the parasite was allowed to adapt to a nonevolving *C. elegans* population taken from a frozen culture at each transfer step; and (v) pathogen control, during which the pathogen adapted to general laboratory conditions in the absence of the nematode host. The treatment protocols for the five evolution conditions were otherwise completely identical.

In particular, the evolution experiment was run at a temperature of 19°C and included transfers to fresh media twice per week. Each treatment was run in ten replicates for a total of 28 transfers (equivalent to 14 weeks). Host population size was set to 500 individuals at each transfer step. 5% of the ancestral host population was added at every second transfer to simulate immigration in order to reduce the likelihood of drift effects. All treatments were maintained in wormballs, which we established as environments for *C. elegans–B. thuringiensis* coevolution experiments [21]. These consist of two halves of a transparent plastic ball, which are filled with a thin layer of the respective medium, followed by addition of bacteria and worms and subsequent closure of the halves [21]. The evolving host populations (treatments (i)–(iii) above) were purified and synchronized at every second transfer step with alkaline hypochlorite:NaOH, which is only survived by nematode eggs [44], thus eliminating any bacteria present [21]. The resulting eggs were raised to L4 larvae on NGM plates with *E. coli*, and then a total of 500 worms (475 evolved worms and 25 from the ancestral stock culture as immigrants) were transferred to the next round of the evolution experiment. Nematodes for the pathogen one-sided adaptation treatment were thawed at each transfer step from frozen aliquots and then raised as above before addition to the wormballs. For *B. thuringiensis*, the host-adapting populations (treatments (iii)–(iv)) were always isolated from dead worms, which were specifically collected at each transfer step and maintained for two additional days in phosphate buffered saline (PBS), followed by pasteurization at 80°C for 10 min to eliminate bacterial contaminants [21], subsequent culturing on NGM plates for 3–5 d, mixing with *E. coli* food, and transfer to the next selection round. *B. thuringiensis* for the host one-sided adaptation treatment were always taken from frozen stock cultures, and those from the pathogen control treatment were directly washed off the wormballs, followed in both treatments by pasteurization and all subsequent steps listed above. Samples from all replicate populations of all treatments were cryopreserved at transfers 12, 20, and 28. The general experimental protocol is similar to that used for our previous evolution experiments [21], and the exact methods for most of the host side of the experiment were recently published in Masri et al. [27].

Phenotypic Analysis

Phenotypic changes across time and treatments were studied for the frozen host and parasite samples from transfer steps 0, 12, 20, and 28, using the same general environmental conditions as in the evolution experiment. These samples from the various transfer steps and treatments were characterized simultaneously and in random order to avoid artifacts because of observer bias and/or random environmental or temporal fluctuations. Both nematodes and bacteria were raised and purified prior to the experiments (alkaline hypochlorite:NaOH treatment for worms [44], pasteurization for bacteria). The hermaphroditic worms were used once they reached the L4 stage, and the final *B. thuringiensis* concentration was adjusted to 1.2×10^8 particles/ml. We first assessed the presence of reciprocal coadaptations during experimental coevolution by comparing survival rate (see assay description below) of coevolved hosts exposed to coevolved pathogens from the same time point and replicate with those of coevolved hosts from the same replicate population exposed to ancestral pathogens and those of ancestral hosts exposed to coevolved pathogens from the same replicate. Thereafter, we analysed variation among evolution treatments for the evolved hosts and pathogens separately. To ensure comparability of evolved populations from the various treatments, we always exposed all evolved hosts or all evolved pathogens to their respective ancestral antagonist. For this analysis, several phenotypic traits were characterized, as described in more detail below.

Changes in host resistance sensu lato (i.e., the ability of the host to survive pathogen exposure) and pathogen virulence sensu lato (i.e., the ability of the pathogen to kill the host) were

assessed by respectively measuring nematode survival and *B. thuringiensis* killing ability [21]. For this measure, 50 worms were exposed to *B. thuringiensis*, and the proportion of surviving hosts was counted after 48 h. Pathogen killing ability is simply represented by the inverse of this measure (i.e., the proportion of dead hosts). The results from these measurements are shown in Fig 1B, 1C, 1D, and 1F. As additional proxies of host resistance and pathogen virulence, we also examined the bacterium's effect on worm body size and population size [21]. For both traits, 35 L4 *C. elegans* were exposed to *B. thuringiensis*. After 48 h, body size was measured as whole worm area for four to six nematodes using differential interference contrast (DIC) microscopy (DM5000B microscope; Leica) and the program ImageJ 1.36b (http://rsb.info.nih.gov/ij/), followed by calculation of the average body size per replicate population for later statistical analysis. After five days of exposure, population size was determined by washing off all worms from the wormballs with 2 ml PBS, counting of animals in three 10 μl subsamples, and subsequent calculation of the total number of worms per replicate population. The results of these assays are shown in S1 Fig.

Infection load was quantified with a new protocol to characterize the ability of *B. thuringiensis* to ensure high abundance inside the host. For this assay, 35 worms were exposed to *B. thuringiensis–E. coli* mixtures (final concentrations respectively of 1.2×10^8 particles/ml and 2×10^9 cells/ml) on PFM plates. After 48 h, three to six live worms per replicate were transferred onto a 12 well microscopic slide, followed by body size measurements using ImageJ 1.36b. To remove bacteria adhering to the cuticle, the worms were carefully washed with approximately 20 μl sterile H_2O under a dissecting microscope, followed by their transfer into 1.5 ml tubes containing 100 μl H_2O. The number of externally associated bacteria, which could not be removed, was estimated by counting cells in the surrounding solution using standard Thoma counting chambers (0.1 mm depth). For each replicate, bacteria were subsequently extracted by sonicating the worms for 10 sec, 6 cycles at 60 Hz, followed by addition of four 1 mm Zirconia beads and vortexing for 3 sec. The number of extracted bacteria was quantified using Thoma chambers. The infection load was then calculated as the number of extracted bacteria minus the number of bacteria in the surrounding solution, adjusted for worm size and averaged per replicate population. The results are shown in Fig 1G.

The characteristics of biofilm formation were studied for all of the evolved replicate populations, and, for a separate set of assays, a selection of isolated clones from these populations. All evolved replicate populations from transfers 0, 12, 20, and 28 were characterized with two assays: (i) as a rough qualitative proxy of biofilm formation, we scored the proportion of replicate populations per treatment that produced clearly visible flakes (Fig 2A and 2B); (ii) biofilm formation was quantified by measuring average particle size produced by each replicate population. 20 μl per population were grown for 96 h at 19°C on NGM, washed off with 3 ml PBS, vortexed for 5 sec in 15 ml tubes, followed by measuring particle area (in mm^2) for the five largest particles within a random 20 μl sample of the culture using DIC microscopy and ImageJ. Four random 20 μl samples were assessed per replicate and averaged for subsequent statistical analysis. The results of this assay are shown in Fig 2D.

An additional analysis of the dynamics of biofilm formation was performed for four individual clones, isolated from transfer 20 from above selected populations. We confirmed that the isolated clones showed the same general characteristics as their source populations (i.e., one highly virulent, non-biofilm-forming clone from the coevolution treatment; one highly virulent, non-biofilm-forming clone from the one-sided adaptation treatment; one non-nematocidal, biofilm-forming clone from the one-sided adaptation treatment; and one non-nematocidal, biofilm-forming clone from the control). The dynamics of biofilm production were characterized by growing the selected four clones, as well as three of the ancestral strains (BT-246, BT-247, and BT-679), on 9 cm NGM plates in several replicates. Every 24 h, an entire

plate was washed off for particle size measurements as described above. The entire analysis was performed for a total of 144 h (Fig 2C).

Moreover, these four clones were further characterized by assessing their bacterial competitive ability under either low nutrient conditions on PFM or high nutrient conditions on NGM. Biofilm-forming and non-biofilm-forming bacteria (concentration of 1.2 x 10^9 particle/ml) from the selected evolved clones were streaked out along thin lines in parallel to each other at a distance of 5 mm and grown at 19°C for 96 hours on NGM and 21 days on PFM (due to the absence of nutrition, growth was substantially reduced under these conditions). Thereafter, the growth expansion of one clone in the direction of the other clone was measured as the distance from the original streak to the farthest area of the grown culture. An analogous measurement was taken for the competing bacterium. A competitiveness index was subsequently calculated for a particular clone by taking its growth expansion measurement and subtracting from it the respective measurement of the competitor. Thus, a competitiveness index of 0 indicates equality, whereas a positive index suggests higher competitiveness for the focal population (Fig 2E).

Statistical analysis of phenotypic data was based on JMP 9 (SAS). Variations between the treatments in almost all traits (exceptions are given below) were evaluated with a general linear model including transfer, treatment, and the interaction between the two as fixed factors and replicate population as a random factor nested within the treatment factor. For the analysis of reciprocal coadaptations, we used a general linear model based on ordinal logistic regression, using exposure type and replicate as factors. A factor effect test was used to assess the relative influence of the defined factors in the model, as implemented in JMP 9 (SAS). Variation in competitiveness was compared with the Mann-Whitney U test (MWU). Variation in the number of replicate populations per treatment and time point, which are able to produce biofilm flakes (qualitative assessment of biofilm formation), was assessed with a Fisher exact test. Graphs were generated with SigmaPlot version 11.0 (Systat Software Inc.). The results of the statistical analysis of the phenotypic data are shown in S1–S7 Tables.

Genome Sequencing

Draft genome sequences for five *B. thuringiensis* strains present in the starting population (BT-246, BT-247, BT-679, BT-22, and BT-50) were used as references for mapping of the population genomic data. For each of these strains, genomic DNA was isolated using a DNeasy Blood and Tissue Kit (Qiagen). Whole genome sequencing was performed using the Roche 454 Genome Sequencer FLX platform. The resulting reads were assembled using GS *De Novo* Assembler (Roche). For MYBT18679, a partially closed reference was generated through targeted PCR and Sanger sequencing, consisting of 31 scaffolds, including more than ten plasmids. A summary of the data and assemblies for each strain is shown in S8 Table.

For samples of the ancestral population and each of the evolved replicate populations from transfer 12 and 20, genomic DNA was isolated following the Qiagen DNeasy Blood and Tissue kit procedures for gram-positive bacteria. Prior to DNA extraction, 10 μl of the frozen bacterial populations were spread onto NGM plates and grown for 14–16 h at 25°C. Bacteria were washed off plates with 1 ml of autoclaved H_2O, followed by DNA extraction. For samples showing the biofilm phenotype, four replicates were extracted and pooled, while three replicates were extracted for the other samples. DNA quantity, measured with Qubit Fluorometric Quantitation, ranged between 9.13 ng/μL and 55.1 ng/μL. For Illumina sequencing, genomic paired-end libraries were prepared following standard methods [45]. Insert sizes (excluding adapters) ranged from 200–450 nucleotides. Libraries were sequenced using GAII or GAIIx Illumina sequencing instruments to yield paired 100mers. The Illumina image analysis pipeline with default parameters was used for image analysis, base-calling, and read filtering. Further

filtering served to remove adapter and PhiX contamination based on blast alignment (pairs with $\geq$ 14nt aligned at $\geq$ 98% were removed). The reads were subsequently processed with SeqPrep (https://github.com/jstjohn/SeqPrep) software to remove adapter sequences and merge overlapping read pairs. The raw read data are available from the ENA database (ENA; www.ebi.ac.uk/ena/) under study number PRJEB5931.

Comparison of Mapping Software

We assessed suitability of mapping software programs (Bowtie [46]; BWA [47]; MOSAIK [48]; SOAP [49]; and GSNAP [50]) to correctly align Illumina reads from population samples (thus including nucleotide and structural variation) to our concatenated metareference. We first simulated reads from three of the publicly available *B. thuringiensis* genome sequences (Genbank accession number NC_014171.1, NC_005957.1 and NC_008600.1) using the dwgsim tool from the dnaa 0.1.2 software suite (http://sourceforge.net/projects/dnaa/). A depth of coverage of 1x for each genome was generated corresponding to 55,000 reads of 100 bp with a fragment size of 350 bp per genome and maximum quality. A metareference was generated by aligning the five genomes using progressiveMAUVE [51] and polymorphic sites were recorded following ambiguity IUPAC codes to avoid counting them as mismatches in the alignment. For GSNAP, a SNP file was created to account for variation. Indel positions were kept without gaps. Usage of SOAP and Bowtie led to low mapping efficiency (S3 Fig), apparently because of imprecise alignment of polymorphic positions. MOSAIK and GSNAP performed equally well, while BWA aligned substantially fewer reads to the metareference (S3 Fig). Based on these results, SOAP and Bowtie were excluded from subsequent analyses.

A second set of simulated data was generated to test the influence of allele frequency biases, which are likely to be present in the evolved populations. Ten genomes of each of the three references were generated with SNP variation, resulting in a total of 30 different genome sequences that were simulated as 100 bp paired-end reads with 1,000x read depth. These produced the site frequency spectrum shown in S2 Fig. The reads were mapped with the three programs, followed by detection of SNPs using SNVer [52] and allele frequency calculations based on the number of SNP reads divided by the total read depth. Based on this data set, which is likely to be representative of the sequence data from our evolved populations, we found GSNAP to produce a site frequency spectrum most similar to the original distribution (S4 Fig). GSNAP was therefore used as mapping software.

Broad-Scale Genome Analysis: Strain Composition of Evolved Populations

Our strategy for estimating the frequency of the five ancestral strains in pooled population samples consisted of four steps (Fig 4A). We first generated a concatenated metareference based on the five ancestral strains. Secondly, the obtained reads of the considered population samples were mapped onto the metareference using GSNAP, resulting in 90%–97% mapping efficiency (S8 Table). Thirdly, we identified the polymorphic sites where only one of the five reference strains shows a substitution. For the population samples, we then determined the frequency of substitutions at each of these diagnostic polymorphic positions and took these as independent estimates of strain frequencies (S9 Table). Fourthly, as such frequency distributions are usually asymmetric (e.g., left-skewed), we calculated the mode of the distribution as the final frequency estimate, using the function mlv from package modeest on the R platform [53]. The results of the statistical comparison among the evolved populations are given in S10 Table, and the inferred values per strain are presented in Fig 3.

Fine-Scale Genome Analysis of BT-679-Dominated Populations: Variant Detection

For the fine-scale genome analysis, we focused on the evolved populations dominated by the BT-679 strain. These populations still showed substantial variation in both killing ability and infection load. The analysis was based on a four-step strategy (Fig 5A). Firstly, each read was mapped to each of the five reference genomes present in the starting population. Secondly, the edit distance between the reference and the mapped read (NM field in SAM format) was recorded and compared among the five alignments (referring to the five reference genomes). Only reads that produced the lowest edit distance to the BT-679 genome were considered for further analysis (i.e., they had the highest similarity to BT-679). Thus, reads with the same or lower edit distance to the non-BT-679 strains were excluded (see mapping statistics in S9 Table, where reads mapping uniquely to each reference and total reads mapped are reported). Thirdly, SNVer software version 0.4.1 [52] was used to identify SNPs and short indels using default parameters, except that the strand bias and the Fisher's exact test threshold was set to 20 instead of 30 (-u 20) to avoid frequency bias due to overfiltering [54]. Minimum mapping quality and base quality were set to 20, and the results are shown in S14 Table. Further filtering consisted of: (i) excluding positions for which an identified SNP was below the 2% or above the 98% quantile of the observed coverage distribution; (ii) excluding SNPs and short indels if a significant Fisher's exact test on strand bias was inferred (0.05 threshold); (iii) excluding SNPs if an indel is detected at the same position; and (iv) keeping SNPs with a minimum allele frequency (MAF) across the sample above 5%. Finally, we identified structural variations using Pindel version 0.2.4 [55] with default parameters except the following: –w 1 (1 million base bins) and –u 0.03 (maximum allowed mismatch rate). The results are summarized in S15 Table.

Fine-Scale Genome Analysis of BT-679-Dominated Populations: Copy Number Variation and Horizontal Gene Transfer

Several tools have been developed to detect copy number variations (CNVs) using depth of coverage (e.g., CNVnator [56] or Event Wise Testing [57]). However, these approaches have not been designed to account for pooled population samples where only some individuals may harbor a CNV, possibly leading to only a proportional coverage change below but not above the value of one. Therefore, we developed our own approach. Firstly, we used the average rank of each position instead of the raw or scale data in order to account for general coverage variations among samples. Secondly, we calculated the variance at each position for the rank of the depth of coverage across the samples. Thirdly, outliers were extracted using the getOutliers() function of the extremevalues package on the R platform with the method I and a normal fit. Adjacent outlier positions (i.e., with a distance of less than 100 bp) were considered to belong to the same coverage singularity. Fourthly, scale coverage relative to the median coverage of chromosomal contigs was calculated at candidate position to estimate the average copy number in each sample (S16 Table).

Following a similar approach, we also assessed copy number variation for each contig within the BT-679 reference by calculating the ratio of each contig over the average of all chromosomal contigs. The variance was estimated by random sampling of 10,000 positions. The results are presented in S17 Table.

Horizontal gene transfer (HGT) was evaluated by identifying non-BT-679 genome regions within the populations dominated by BT-679. For this, we extracted reads mapping uniquely and best to one of the non-BT-679 reference genomes. We only considered the thus identified putative HGT fragments, for which an indication of HGT from the same reference genome is continuously found across at least 1 kb. The frequency of each putative HGT was then

estimated through the ratio of the median coverage of the fragment over the median coverage of the chromosomal contigs of BT-679 (S18 Table).

Fine-Scale Genome Analysis of BT-679-Dominated Populations: Population Genetic Analysis

We calculated three different population genetics statistics in a sliding window approach with 5 kb steps and 10 kb window size, namely Watterson's θ, Tajima's π, and Tajima's D. To correct for coverage variation within a window and along the genome, coverage was taken as a proxy for the number of samples, and statistics were calculated using the adjusted Watterson's θ and Tajima's π estimates that specifically allow for sample size variation across the genome [36]. Only polymorphisms showing a frequency above 0.05 were considered. The results are summarized in S19 Table.

Fine-Scale Genome Analysis of BT-679-Dominated Populations: Statistical Analyses

The same statistical tests were performed on each dataset, which either contained the identified SNPs, short indels, pindel structural variants, CNVs detected through coverage variation, putative HGTs, or the population genetic characteristics. We excluded coevolution replicate 3 at transfer 12 from the analysis because it contained two genotypes at higher frequencies (BT-679 and BT-22), and it was thus not directly comparable to the other replicate populations dominated by BT-679. Two types of statistical analyses were performed. Firstly, a linear regression analysis was performed using genomic variation versus either killing ability or infection load. The linear regression (using R [53]) was weighted by the $\log_{10}$(coverage) on each dataset except of the population genetics statistics, because the read depth coverage is directly affecting the variance on frequency estimates. Secondly, an ANOVA was performed to compare the difference between treatments. The treatment effect was nested within transfer as follows:

$$Variable \sim Transfer + Treatment[Replicate] + Transfer * Treatment[Replicate]$$

Significance levels were adjusted using the FDR [28].

The statistical analysis identified a large number of genome regions (S20 Table). At least some of them, but possibly not all, may have influenced pathogen adaptation to either coevolving or nonchanging host. In order to identify the most relevant regions for such adaptive processes, we used the following statistical and functional criteria: (i) the relevant genome regions should have been identified through variation in at least four replicate populations (and thus not be the consequence of exceptional events in very few populations; note that under the latter conditions, homoscedasticity of the data—as required for ANOVA—may also be compromised); (ii) for the ANOVA approach, they should only show a treatment effect and not a transfer or an interaction effect, the latter of which may both indicate convergent evolution across treatments during the course of the experiment; (iii) for the analysis of horizontal gene transfer, treatment variance of the transferred region should exceed 0.04; and (iv) the identified variations should be of functional consequence; for example they should influence gene expression levels (i.e., changes in copy number) or directly influence gene function (nonsynonymous or frame-shift mutations, etc). The resulting list of candidate regions is presented below and highlighted in yellow and bold font type in S20 Table.

Two of the identified regions were found to covary with killing ability: the deletion frequency of the mviN virulence gene and the copy number of the cry toxin-containing plasmid. This relationship was reassessed through pairwise analysis of the three characteristics using two approaches: (i) Spearman's rank correlation analysis and (ii) weighted regression analysis,

estimated through the ratio of the median coverage of the fragment over the median coverage of the chromosomal contigs of BT-679 (S18 Table).

Fine-Scale Genome Analysis of BT-679-Dominated Populations: Population Genetic Analysis

We calculated three different population genetics statistics in a sliding window approach with 5 kb steps and 10 kb window size, namely Watterson's θ, Tajima's π, and Tajima's D. To correct for coverage variation within a window and along the genome, coverage was taken as a proxy for the number of samples, and statistics were calculated using the adjusted Watterson's θ and Tajima's π estimates that specifically allow for sample size variation across the genome [36]. Only polymorphisms showing a frequency above 0.05 were considered. The results are summarized in S19 Table.

Fine-Scale Genome Analysis of BT-679-Dominated Populations: Statistical Analyses

The same statistical tests were performed on each dataset, which either contained the identified SNPs, short indels, pindel structural variants, CNVs detected through coverage variation, putative HGTs, or the population genetic characteristics. We excluded coevolution replicate 3 at transfer 12 from the analysis because it contained two genotypes at higher frequencies (BT-679 and BT-22), and it was thus not directly comparable to the other replicate populations dominated by BT-679. Two types of statistical analyses were performed. Firstly, a linear regression analysis was performed using genomic variation versus either killing ability or infection load. The linear regression (using R [53]) was weighted by the $\log_{10}$(coverage) on each dataset except of the population genetics statistics, because the read depth coverage is directly affecting the variance on frequency estimates. Secondly, an ANOVA was performed to compare the difference between treatments. The treatment effect was nested within transfer as follows:

$$Variable \sim Transfer + Treatment[Replicate] + Transfer * Treatment[Replicate]$$

Significance levels were adjusted using the FDR [28].

The statistical analysis identified a large number of genome regions (S20 Table). At least some of them, but possibly not all, may have influenced pathogen adaptation to either coevolving or nonchanging host. In order to identify the most relevant regions for such adaptive processes, we used the following statistical and functional criteria: (i) the relevant genome regions should have been identified through variation in at least four replicate populations (and thus not be the consequence of exceptional events in very few populations; note that under the latter conditions, homoscedasticity of the data—as required for ANOVA—may also be compromised); (ii) for the ANOVA approach, they should only show a treatment effect and not a transfer or an interaction effect, the latter of which may both indicate convergent evolution across treatments during the course of the experiment; (iii) for the analysis of horizontal gene transfer, treatment variance of the transferred region should exceed 0.04; and (iv) the identified variations should be of functional consequence; for example they should influence gene expression levels (i.e., changes in copy number) or directly influence gene function (nonsynonymous or frame-shift mutations, etc). The resulting list of candidate regions is presented below and highlighted in yellow and bold font type in S20 Table.

Two of the identified regions were found to covary with killing ability: the deletion frequency of the mviN virulence gene and the copy number of the cry toxin-containing plasmid. This relationship was reassessed through pairwise analysis of the three characteristics using two approaches: (i) Spearman's rank correlation analysis and (ii) weighted regression analysis,

whereby plasmid copy number was weighted by its variance to reduce the influence of extreme values.

Fine-Scale Genome Analysis of BT-679-Dominated Populations: Overview of Identified Genome Regions

The linear regression analysis revealed four cases of a significant association between genome and killing ability but none with infection load. Three of the four significant cases may be of relevance for bacterial adaptation to either coevolving or nonchanging host as they could have functional consequences (i.e., they affect expression levels of genes or gene function itself). In particular, killing ability was found to correlate positively with the copy number of (i) the plasmid pBT679_22_6, which contains the nematocidal toxin genes *cry21Aa2* and *cry14Aa1* (q-value = 1.73E-07; Fig 5B and 5C); and (ii) the plasmid (or plasmid fragment) represented by the contig Bti_GWDALJX04IG4JR_1–226, containing a plasmid recombination enzyme, two hypothetical proteins and Parvulin-like peptidyl-prolyl isomerase (q-value = 0.016). Killing ability also correlated negatively with the frequency of a chromosomal deletion of 12 amino acids in a putative virulence factor containing an mviN domain (gene Bt_01995; q-value = 0.038; Fig 5B and 5C). Taken together, the region with the strongest effect (according to the q-value) refers to the plasmid that contains genes with known nematocidal effect and that is thus known to have a function in pathogen–host interactions. Interestingly, copy number of this plasmid and the mviN domain-containing protein deletion frequency not only correlate with killing ability but also with each other (Fig 5C, S21 Table). Here, the significant negative correlation indicates that a high protein deletion frequency coincides with a low plasmid copy number in the same replicate populations, suggesting a functional relationship between these.

The ANOVA approach yielded a comparatively large list of genome regions with a significant treatment effect, strongly suggesting that the imposed differences in selection conditions lead to changes in the favoured genomic variants and/or promoted horizontal gene transfer. In particular, a total of 53 genome regions were inferred from the SNP analysis, 3 from the Pindel-based structural analysis, 81 from the coverage-based copy number variation analysis, 66 from the assessment of horizontal gene transfer, and 35 from the population genetic analysis. Note that some of these regions overlap as a consequence of the different approaches used during the respective analyses. Based on the above outlined conservative criteria, only a few of the identified regions are likely of relevance for adaptation to either coevolving or nonchanging hosts (S20 Table), including (i) a recombinase or invertase (gene Bti_05865), for which the gain of a stop codon varies among treatments (inferred through SNP analysis); (ii) a predicted acetyltransferase or hydrolase (gene Bti_05100), which shows copy number variation between treatments; (iii) two horizontally transferred gene regions from the *B. thuringiensis* strain BT-246 containing a Cysteine protease, a recombinase or invertase, and several hypothetical proteins; (iv) one horizontally transferred gene region from BT-247, containing among others a transcriptional antiterminator; (v) one horizontally transferred 16S rRNA gene region from BT-50 (S5 Fig); and (vi) an approximately 65 kb region from the plasmid contig Bti_GWDALJX04I0LJH_51–405_fm319.5, consistently identified by the population genetic measures to vary among treatments and containing a variety of different genes such as those encoding toxins (with unknown effects), a membrane protein, a transposase, germination proteins, a secreted acid phosphatase, and others (Fig 5D). None of the above regions contains genes previously implicated in the bacterium's interaction with a host. The only exception may refer to some of the genes found in the 65 kb plasmid region, of which the toxin, the membrane protein, the acid phosphatase, or the germination protein genes could be speculated to contribute to host

interactions. The dissection of the above genes' exact role in shaping adaptation to either co-evolving or nonchanging hosts represents a particular challenge for future research.

Interestingly, 14 of the inferred cases of copy number variations refer to collagen triple helix repeats (S20 Table), possibly suggesting a role of these genes in general adaptation to a host environment, irrespective of whether the host is coadapting or not. It is similarly interesting to note that horizontal transfers mainly originated from two ancestral non-nematocidal *B. thuringiensis* strains that are mainly, yet not exclusively, favoured in the absence of the host. Of these, most transfers came from strain BT-22, encompassing 35 horizontally transferred fragments with a total length of 51 kb; whereas 19 fragments with a total length of 45 kb originated from BT-50. One of the transferred fragments refers to a phage that originated from BT-50 and spread through the BT-679-dominated populations across time, irrespective of the evolution treatment regime and possibly as a selfish element that does not contribute to bacterial adaptation to a host (S5 Fig).

Toxin Gene Screen

To identify genes for crystal toxin proteins that were present in the starting population of *B. thuringiensis*, we performed sequence similarity searches on the draft genome assemblies of the nematocidal strains BT-246, BT-247, and BT-679 using known cry toxin protein sequences as queries. The query sequences were derived from the cry toxin list available on the Bt toxin nomenclature webpage (http://www.lifesci.sussex.ac.uk/home/Neil_Crickmore/Bt/). Based on this analysis, we identified seven genes with high similarity to known cry toxin sequences: *cry13Aa1* in BT-246, *cry6Ba1* in BT-247, and *cry14Aa1*, *cry21Aa2*, *cry34Aa4*, *cry35Aa4*, and *cry38Aa1* in BT-679. *cry14Aa1* and *cry21Aa2* are located 8 kb apart on a 23 kb plasmid, while *cry34Aa4*, *cry35Aa4*, and *cry38Aa1* are all located within a 4 kb region on a separate plasmid. We also identified several additional putative cry toxin genes (<60% similarity to query sequences), but for practical reasons they were not considered in subsequent analyses.

To analyse the composition of crystal toxin genes in the evolved *B. thuringiensis* populations, we focused on *cry13Aa1*, *cry6Ba1*, *cry14Aa1*, *cry21Aa2*, and *cry35Aa4*. Twenty individual clones were isolated from each available evolved population from transfer 12 and 20 by plating the population on nematode growth medium (NGM) plates and picking single colonies, resulting in a total of 1,100 clones tested. The clones were grown overnight at 28°C in LB medium and then frozen at -20°C. This frozen material was used directly in PCRs with toxin-specific primers (S11 Table) and 15.6 μl reaction volumes containing 0.39 units GoTaq DNA Polymerase (Promega), 1x Green GoTaq reaction buffer, 0.2mM each dNTP, and 0.4 μM of each primer. Thermal cycling was performed with an initial denaturation step at 95°C for 2 min followed by 35 cycles of 30 sec 95°C, 30 sec 57°C, 90 sec 72°C, and then a final extension at 72°C for 10 min. *CodY* primers were included in each reaction to ensure integrity of the template. We additionally determined the chromosomal background of each clone by Sanger sequencing of part of the *codY* gene, amplified by PCR as above.

The chromosomal backgrounds were largely consistent with the whole genome data (Fig 3). The coevolution treatment was dominated by BT-679 and the control treatment by BT-22, while the adaptation treatment showed variation between replicates with virulent populations dominated by BT-679 and nonvirulent populations dominated by BT-22 or BT-50 (S12 Table). The toxin genes *cry14Aa1*, *cry21Aa2*, and *cry35Aa4* were only found in evolved BT-679 clones, thus remaining within the same chromosomal background. Their presence varied among these clones, whereby *cry14Aa1* and *cry21Aa2* showed the same pattern (i.e., both present or both absent) for all but seven clones and were both less abundant than *cry35Aa4* (Fig 4). *cry13Aa1* was found only once in a BT-246 background, while *cry6Ba1* was absent, consistent with the

very low abundance of BT-246 and BT-247 in the evolved material. The distribution of the BT-679 toxin genes differed significantly between coevolution and control conditions and between some of the coevolved and one-sided adapted populations (Fig 4, S13 Table).

Functional Analysis of Toxin Genes

We used two complementary approaches to assess the nematocidal effect of cry toxin genes from BT-679. On the one hand, we expressed one of the toxin genes, *cry21Aa2*, in *E. coli*, followed by *C. elegans* survival analysis. On the other hand, we introduced either *cry14Aa1* or *cry21Aa2* into a BT-679 variant that lost the 22.5 kb plasmid carrying these two toxin genes (denoted BT-679_Cry-), again followed by analysis of nematode survival.

For the former approach, the entire coding region of *cry21Aa2* was amplified by PCR (see below) and cloned into the expression vector pQE30 using standard procedures. Both the pQE30 with *cry21Aa2* and the empty vector were transferred into *E. coli* JM109 by electroporation and selection on ampicillin-containing medium (100 μg/ml). Prior to nematode survival experiments, *E. coli* was cultured at 37°C overnight in LB medium, containing ampicillin (100 μg/ml) and IPTG (200 μg/ml). The bacteria were washed twice and cell density was adjusted to OD_{600} = 5. Virulence of the resulting *E. coli* strains was assessed using exactly the same methods as described above for phenotypic analysis of the evolved material (chapter 1.3). The main exception was that we used an isogenic *C. elegans* strain (the standard laboratory strain N2) and standard Petri dishes instead of wormballs. Nematode survival was assessed after 48 h under six treatment conditions: (i) the ancestral BT-679 with cry toxins (BT-679_Cry+); (ii) the BT-679_Cry− strain lacking the two tested toxin genes; (iii) BT-679_Cry− combined with a low concentration of the *cry21Aa2*-expressing *E. coli* (a 1:10 dilution of the OD5-concentrated stock); (iv) BT-679_Cry− combined with a high concentration of *cry21Aa2*-expressing *E. coli* (the OD5-concentrated stock without any dilution); (v) only the *cry21Aa2*-expressing *E. coli* (at the OD5 stock concentration); and (vi) only the empty vector *E. coli* strain. In all cases, the empty-vector *E. coli* strain was added as food.

For the second approach, we first substituted *gfp* with a multiple cloning site (MCS) in the pHT315 p*AphA-gfp* plasmid that is used as an *E. coli–B. thuringiensis* shuttle vector (kindly provided by Christina Nielsen-LeRoux, Guyancourt, France). For this, the MCS of the pUC19 plasmid (Carl Roth, Germany) was amplified by PCR using Phusion High-Fidelity DNA polymerase (Thermo Scientific, Germany) and primers MCS_f and MCS_r (S21 Table). The PCR product was gel-purified (QIAquick Gel Extraction Kit and PCR purification kit, both Qiagen, Germany), digested with *Hin*dIII and *Xba*I, ligated with T4 DNA ligase (Thermo Scientific, Germany) into the respective sites of pHT315_p*AphA-gfp* to create pHT315_p*AphA*-MCS. This vector was introduced into *E. coli* Top10 (Invitrogen, US), grown in LB with ampicillin (100 μg/ml), and followed by plasmid isolation (QIAprep Spin Miniprep Kit, Qiagen, Germany). Thereafter, the entire coding regions of *cry21Aa2* and *cry14Aa1* were PCR-amplified using Phusion High-Fidelity DNA polymerase (Thermo Scientific, Germany) and the respective primers (S21 Table), digested with *Sal*I and *Pae*I, ligated into pHT315_p*AphA*-MCS, followed by transformation into *E. coli* Top10 and plasmid isolation as above. *B. thuringiensis* BT-679_Cry- was transformed with three different vectors (containing either *cry14Aa1*, *cry21Aa2*, or the red fluorescent protein (*rfp*) as a control), using electroporation with a Bio-Rad Gene Pulser (Bio-Rad, Germany), as described previously [58]. Transformants were grown in LB containing erythromycin (10μg/ml), and presence of the correct inserts was confirmed by Sanger sequencing. Prior to survival experiments, the *B. thuringiensis* strains were grown for four days at 19°C on NGM, washed in S buffer, and had the concentration adjusted to 1.2 x 10^8 particles/ml, generally following the procedures outlined above for the evolution

experiment. The survival experiment was performed using the same methods as above, including the N2 *C. elegans* strain and Petri dishes for exposure. The empty-vector *E. coli* strain was always added as food. Survival was tested for a total of five treatments: (i) the ancestral BT-679 containing all cry toxins (BT-679_Cry+); (ii) the BT-679_Cry- lacking the two toxin genes; (iii) the BT-679_Cry-, which contains the *cry14Aa1*-expressing plasmid (BT-679_Cry-_+14); (iv) the BT-679_Cry-, which contains the *cry21Aa2*-expressing plasmid (BT-679_Cry-_+21); and (v) and the BT-679_Cry- with the *rfp*-expressing plasmid as a control (BT-679_Cry-_0).

The results of the two assays are shown in Fig 6, and the statistical results are given in S23 and S24 Tables

Supporting Information

S1 Data. Results on phenotypic changes in both host and pathogen during experimental evolution, including data for reciprocal adaptations of coevolved hosts and pathogens, changes in evolved hosts exposed to ancestral pathogens (survival, population growth, body size, and infection load), and changes in evolved pathogens exposed to ancestral hosts (killing ability, infection load, pathogen effect on host population growth and host body size). The data are summarized in Fig 1, and S1 Fig, and the statistical results are shown in S1–S5 Tables.
(XLSX)

S2 Data. Results on the analysis of biofilm formation in evolved *B. thuringiensis*, including frequency of evolved populations able to form biofilms, quantitative analysis of particle size across time for specific bacterial clones, mean particle size for evolving *B. thuringiensis* populations, and competitive ability of either biofilm-forming or non-biofilm-forming clones on either nutrient-rich or nutrient-poor medium. The results are summarized in Fig 2, and the statistical results are given in S6 and S7 Tables.
(XLSX)

S3 Data. Results on the relative frequencies of *B. thuringiensis* strains in the evolving populations. A summary of the results is shown in Fig 3, and the statistical results are given in S10 Table.
(XLSX)

S4 Data. Results on the frequency of *B. thuringiensis* BT-679 cry toxin gene combinations in the evolving populations. The results are shown in Fig 4 and the statistical results given in S13 Table.
(XLSX)

S5 Data. Data on the relationship between killing ability, frequency of the mviN deletion, and copy number of the cry toxin plasmid, and also data on the annotation of the plasmid showing significant population genetic differences among evolution treatments. The results are summarized in Fig 5, and the statistical findings are shown in S14–S21 Tables.
(XLSX)

S6 Data. Original data for virulence of BT-679 pathogens with or without nematocidal cry toxin genes. The results are shown in Fig 6. The statistical analysis is given in S23 and S24 Tables.
(XLSX)

S7 Data. Results on the efficiency of read mapping by different software programs. The results are shown in S3 Fig.
(XLSX)

S8 Data. Results on site frequency spectrum of simulated genomes, as illustrated in S4 Fig.
(XLSX)

S9 Data. Results on two cases of horizontal gene transfer, as highlighted in S5 Fig.
(XLSX)

S1 Fig. Variation among evolved pathogen populations from different treatments (colors as in Fig 1A) upon exposure to the ancestral host. A, Variation in the pathogen's effect on host population size; and **B**, in the pathogen's effect on host body size. Bars show standard errors. S4 Table and S5 Table show the corresponding statistical results. The original data is provided in S1 Data.
(TIF)

S2 Fig. Electron micrograph of a solid and highly robust biofilm particle produced by control-evolved *B. thuringiensis*.
(TIF)

S3 Fig. Analysis of five mapping programs as to their ability to correctly align simulated reads from *B. thuringiensis* genomes. The five mapping programs are given along the *x*-axis. The *y*-axis presents the total number of reads mapped, classified in four categories following the samtools flagstat function: (i) reads unmapped (yellow bar area); (ii) not properly paired: both reads of a pair are mapped onto the reference genome but expected insert size and/or orientation is incorrect (red bar area); (iii) singletons: only one read of the pair is mapped (light blue area); and (iv) properly paired: both reads are mapped onto the reference genome with correct orientation and expected insert size (dark blue area). The data is shown in S7 Data.
(TIF)

S4 Fig. Site frequency spectrum of 30 simulated *B. thuringiensis* genome sequences derived from three reference genomes. Original spectrum relative to the reference genome NC_014171.1 (expected results on far left), and the results obtained with the mapping software BWA, MOSAIK, and GSNAP. The data is shown in S8 Data.
(TIF)

S5 Fig. Exemplary cases of horizontal transfers with significant variation between treatments or transfers. A, Significant variation among treatments for a 16S rRNA gene, horizontally transferred from the BT-50 ancestral strain to the BT-679 genotype. **B,** Horizontal transfer and spread of a phage from the non-nematocidal ancestral strain BT-50 in the BT-679 coevolved and one-sided adapted populations across time. The different replicate populations are given along the *x*-axis and DNA fragment frequency on the *y*-axis. Red indicates coevolution and green one-sided adaptation. The data is shown in S9 Data.
(TIF)

S1 Table. Analysis of reciprocal coevolutionary adaptation in comparison to adaptation to the ancestral antagonist. Pairwise comparison of the different exposure types. Here, the coevolved–coevolved exposures from a particular replicate population were compared with the corresponding exposures, in which the same coevolved host or pathogen replicate was confronted with the ancestral antagonist. Significant values after FDR adjustment are given in bold. The data is provided in S1 Data.
(DOCX)

S2 Table. Comparison between evolved and ancestral host phenotypes. Comparison between evolved (host coevolution, host one-sided adaptation, and host control) and ancestral hosts both exposed to ancestral pathogens using an analysis of variance. Degrees of freedom (df) are given for the comparison and the error (before and after comma, respectively). Significant values after FDR adjustment are given in bold. The data is provided in S1 Data.
(DOCX)

S3 Table. Analysis of changes in host phenotypes across time and treatments. Evolved host populations (host coevolution, host one-sided adaptation, and host control) were exposed to the ancestral pathogen; the defined models included evolution treatment, transfer, the interaction between the two as fixed factors, and replicate nested within treatment as a random factor. The specified models provide a better fit to the data than the corresponding minimal models ($p < 0.0001$). The table shows the results for the factor effect tests, none of which yielded a significant result. The data is provided in S1 Data.
(DOCX)

S4 Table. Comparison between evolved and ancestral pathogen phenotypes. Comparison between evolved (pathogen coevolution, pathogen one-sided adaptation, and pathogen control) and ancestral pathogens both exposed to ancestral hosts using an analysis of variance. Degrees of freedom (df) are given for the comparison and the error (before and after comma, respectively). Significant values after FDR adjustment are in bold. The data is provided in S1 Data.
(DOCX)

S5 Table. Analysis of the changes in pathogen phenotypes across time and treatments. Evolved pathogen populations (pathogen coevolution, pathogen one-sided adaptation, and pathogen control) were exposed to the ancestral host; the defined models included evolution treatment, transfer, the interaction between the two as fixed factors, and replicate nested within treatment as a random factor. The specified models provide a better fit to the data than the corresponding minimal models ($p < 0.0001$). The table shows the results for the factor effect tests. Significant probabilities are given in bold. The data is provided in S1 Data.
(DOCX)

S6 Table. Fisher exact test of differences in the number of bacterial populations able to form biofilm. The data is provided in S2 Data.
(DOCX)

S7 Table. Mann-Whitney U test of differences in bacterial competition. The data is provided in S2 Data.
(DOCX)

S8 Table. Statistics of the mapping of Illumina reads to the concatenated metareference.
(XLSX)

S9 Table. Statistics of the mapping of Illumina reads to the five different reference genomes. Table is presented as an Excel file.
(XLSX)

S10 Table. Statistical analysis of the variation in strain composition across evolution treatments and time. We used an extended analysis of molecular variance (AMOVA) adonis function in R package vegan. The defined model included evolution treatment, transfer, and the interactions between the two as fixed factors and replicate nested within treatment as random factor. The specified model provided a better fit to the data than the corresponding minimal

model ($p < 0.0001$). The table shows the effect tests for the fixed factors. Significant probabilities are given in bold. The data is provided in S3 Data.
(DOCX)

S11 Table. Genes and primers used for PCR-based toxin screen.
(DOCX)

S12 Table. Number of clones with particular chromosomal background in the evolved replicate populations.
(DOCX)

S13 Table. Statistical analysis of the variation in toxin gene composition across evolution treatments and time. The defined nominal logistic models included evolution treatment, transfer, and the interactions between the two as fixed factors and replicate nested within treatment as random factor. The specified models provided a better fit to the data than the corresponding minimal model ($p < 0.0001$). The table shows the effect tests for the fixed factors. Significant probabilities are given in bold. The data is provided in S4 Data.
(DOCX)

S14 Table. Results of the analysis of SNPs and short indels in the BT-679-dominated populations. Table is presented as an Excel file.
(XLSX)

S15 Table. Results of the Pindel analysis of structural variation in the BT-679-dominated populations. Table is presented as an Excel file.
(XLSX)

S16 Table. Results of the analysis of copy number variation in the BT-679-dominated populations. Table is presented as an Excel file.
(XLSX)

S17 Table. Results of the analysis of horizontal gene transfer to the BT-679 genotype. Table is presented as an Excel file.
(XLSX)

S18 Table. Results of the analysis of copy number variation for entire contigs in the BT-679-dominated populations. Table is presented as an Excel file.
(XLSX)

S19 Table. Results for the population genetic analysis of the BT-679-dominated populations. Table is presented as an Excel file.
(XLSX)

S20 Table. Summary of significant variations in the fine-scale genomic analysis of the BT-679-dominated populations. Table is presented as an Excel file.
(XLSX)

S21 Table. Pairwise analysis of the deletion frequency in the mviN virulence gene, the copy number of the cry toxin-containing plasmid and killing ability. Pairwise analysis was performed twice using: (i) the Spearman rank order correlation analysis (correlation parameter ρ_s); and (ii) weighted regression analysis (strength of relationship indicated by R^2), whereby plasmid copy number was weighted by its variance to reduce influence of outliers. Significant values after FDR adjustment are in bold. The data is provided in S5 Data.
(DOCX)

S22 Table. Primers used during functional analysis of cry toxins.
(DOCX)

S23 Table. Statistical analysis of nematode survival after exposure to cry-toxin-expressing *B. thuringiensis*. The data is provided in S6 Data.
(DOCX)

S24 Table. Statistical analysis of nematode survival after exposure to cry-toxin-expressing *E. coli*. The data is provided in S6 Data.
(DOCX)

Acknowledgments

We thank Philip Crain, Sylvia Cremer, Dieter Ebert, Michel Gohar, Francis Jiggins, Christina Nielsen-LeRoux, Thorsten Reusch, Vincent Sanchis, Jacqui Shykoff, Jakob Strauss, Arne Traulsen, Mathias Wegner, and the Michiels, Bornberg-Bauer and Schulenburg labs for advice; Martina Hohloch, Nadine Timmermeyer, Susanne Weller, Rania Nakad, Karoline Fritzsche, Antje Thomas, and the Kiel ICMB sequencing team (especially Markus Schilhabel, Melanie Friskovec, Melanie Schlapkohl, Daniela Esser) for technical support.

Author Contributions

Conceived and designed the experiments: LM AB AES PR EBB HS. Performed the experiments: LM AB AES AP DL PSG SP MS JH. Analyzed the data: LM AB AES AP HS. Wrote the paper: LM AB AES AP JH EB HL RD NKM RDS JK PR AT EBB HS. Supported pathogen whole genome analysis: HL EB RD. Contributed to design and analysis of evolution experiment: NKM RDS JK AT. Contributed to genome data analysis: PR EBB.

References

1. Woolhouse ME, Webster JP, Domingo E, Charlesworth B, Levin BR. Biological and biomedical implications of the co-evolution of pathogens and their hosts. Nat Genet. 2002; 32: 569–577. PMID: 12457190
2. Schmid-Hempel P. Evolutionary parasitology. Oxford, UK: Oxford University Press; 2011.
3. Dybdahl MF, Jenkins CE, Nuismer SL. Identifying the molecular basis of host-parasite coevolution: merging models and mechanisms. Am Nat. 2014; 184: 1–13. doi: 10.1086/676591 PMID: 24921596
4. Zaman L, Meyer JR, Devangam S, Bryson DM, Lenski RE, Ofria C. Coevolution Drives the Emergence of Complex Traits and Promotes Evolvability. Keller L, editor. PLoS Biol. 2014; 12: e1002023. doi: 10.1371/journal.pbio.1002023 PMID: 25514332
5. Weitz JS, Hartman H, Levin SA. Coevolutionary arms races between bacteria and bacteriophage. Proc Nat Acad Sci U S A. 2005; 102: 9535–9540. PMID: 15976021
6. Koskella B, Brockhurst MA. Bacteria-phage coevolution as a driver of ecological and evolutionary processes in microbial communities. FEMS Microbiol Rev. 2014;
7. Dodds PN, Rathjen JP. Plant immunity: towards an integrated view of plant-pathogen interactions. Nat Rev Genet. 2010; 11: 539–548. doi: 10.1038/nrg2812 PMID: 20585331
8. Karasov TL, Horton MW, Bergelson J. Genomic variability as a driver of plant-pathogen coevolution? Curr Opin Plant Biol. 2014; 18: 24–30. doi: 10.1016/j.pbi.2013.12.003 PMID: 24491596
9. Jousimo J, Tack AJM, Ovaskainen O, Mononen T, Susi H, Tollenaere C, et al. Disease ecology. Ecological and evolutionary effects of fragmentation on infectious disease dynamics. Science. 2014; 344: 1289–1293. doi: 10.1126/science.1253621 PMID: 24926021
10. Ebert D. Host-parasite coevolution: Insights from the Daphnia-parasite model system. Curr Opin Microbiol. 2008; 11: 290–301. doi: 10.1016/j.mib.2008.05.012 PMID: 18556238
11. King KC, Delph LF, Jokela J, Lively CM. The geographic mosaic of sex and the Red Queen. Curr Biol. 2009; 19: 1438–1441. doi: 10.1016/j.cub.2009.06.062 PMID: 19631541

12. Dupas S, Dubuffet A, Carton Y, Poirie M. Local, geographic and phylogenetic scales of coevolution in Drosophila-parasitoid interactions. Adv Parasitol. 2009; 70: 281–295. doi: 10.1016/S0065-308X(09)70011-9 PMID: 19773075

13. Morran LT, Schmidt OG, Gelarden IA, Parrish RC, Lively CM. Running with the Red Queen: Host-Parasite Coevolution Selects for Biparental Sex. Science. 2011; 333: 216–218. doi: 10.1126/science.1206360 PMID: 21737739

14. Kerr PJ. Myxomatosis in Australia and Europe: a model for emerging infectious diseases. Antiviral Res. 2012; 93: 387–415. doi: 10.1016/j.antiviral.2012.01.009 PMID: 22333483

15. Brockhurst MA, Koskella B. Experimental coevolution of species interactions. Trends Ecol Evol. 2013; 28: 367–375. doi: 10.1016/j.tree.2013.02.009 PMID: 23523051

16. Paterson S, Vogwill T, Buckling A, Benmayor R, Spiers AJ, Thomson NR, et al. Antagonistic coevolution accelerates molecular evolution. Nature. 2010; 464: 275–278. doi: 10.1038/nature08798 PMID: 20182425

17. Kashiwagi A, Yomo T. Ongoing phenotypic and genomic changes in experimental coevolution of RNA bacteriophage Qb and *Escherichia coli*. PLoS Genet. 2011; 7: e1002188. doi: 10.1371/journal.pgen.1002188 PMID: 21829387

18. Scanlan PD, Hall AR, Lopez-Pascua LD, Buckling A. Genetic basis of infectivity evolution in a bacteriophage. Mol Ecol. 2011; 20: 981–989. doi: 10.1111/j.1365-294X.2010.04903.x PMID: 21073584

19. Poullain V, Gandon S, Brockhurst MA, Buckling A, Hochberg ME. The evolution of specificity in evolving and coevolving antagonistic interactions between a bacteria and its phage. Evolution. 2008; 62: 1–11. PMID: 18005153

20. Meyer JR, Dobias DT, Weitz JS, Barrick JE, Quick RT, Lenski RE. Repeatability and contingency in the evolution of a key innovation in phage lambda. Science. 2012; 335: 428–432. doi: 10.1126/science.1214449 PMID: 22282803

21. Schulte RD, Makus C, Hasert B, Michiels NK, Schulenburg H. Multiple reciprocal adaptations and rapid genetic change upon experimental coevolution of an animal host and its microbial parasite. Proc Natl Acad Sci U S A. 2010; 107: 7359–7364. doi: 10.1073/pnas.1003113107 PMID: 20368449

22. Griffitts JS, Aroian RV. Many roads to resistance: how invertebrates adapt to Bt toxins. Bioessays. 2005; 27: 614–624. PMID: 15892110

23. Nielsen-LeRoux C, Gaudriault S, Ramarao N, Lereclus D, Givaudan A. How the insect pathogen bacteria *Bacillus thuringiensis* and *Xenorhabdus/Photorhabdus* occupy their hosts. Curr Opin Microbiol. 2012; 15: 220–231. doi: 10.1016/j.mib.2012.04.006 PMID: 22633889

24. Schulte RD, Makus C, Hasert B, Michiels NK, Schulenburg H. Host-parasite local adaptation after experimental coevolution of *Caenorhabditis elegans* and its microparasite *Bacillus thuringiensis*. Proc Biol Sci. 2011; 278: 2832–2839. doi: 10.1098/rspb.2011.0019 PMID: 21307053

25. Schulte RD, Hasert B, Makus C, Michiels NK, Schulenburg H. Increased responsiveness in feeding behaviour of *Caenorhabditis elegans* after experimental coevolution with its microparasite *Bacillus thuringiensis*. Biol Lett. 2012; 8: 234–236. doi: 10.1098/rsbl.2011.0684 PMID: 21880622

26. Schulte RD, Makus C, Schulenburg H. Host-parasite coevolution favours parasite genetic diversity and horizontal gene transfer. J Evol Biol. 2013; 26: 1836–1840. doi: 10.1111/jeb.12174 PMID: 23865952

27. Masri L, Schulte RD, Timmermeyer N, Thanisch S, Crummenerl LL, Jansen G, et al. Sex differences in host defence interfere with parasite-mediated selection for outcrossing during host-parasite coevolution. Ecol Lett. 2013; 16: 461–468. doi: 10.1111/ele.12068 PMID: 23301667

28. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc B. 1995; 57: 289–300.

29. Vlamakis H, Chai Y, Beauregard P, Losick R, Kolter R. Sticking together: building a biofilm the Bacillus subtilis way. Nature Reviews Microbiology. 2013; 11: 157–168. doi: 10.1038/nrmicro2960 PMID: 23353768

30. Lopez D, Vlamakis H, Kolter R. Biofilms. Cold Spring Harb Perspect Biol. 2010; 2: a000398. doi: 10.1101/cshperspect.a000398 PMID: 20519345

31. Agaisse H, Lereclus D. How does *Bacillus thuringiensis* produce so much insecticidal crystal protein? J Bacteriol. 1995; 177: 6027–6032. PMID: 7592363

32. Sniegowski PD, Gerrish PJ. Beneficial mutations and the dynamics of adaptation in asexual populations. Philos Trans R Soc Lond B Biol Sci. 2010; 365: 1255–1263. doi: 10.1098/rstb.2009.0290 PMID: 20308101

33. Morgan AD, Gandon S, Buckling A. The effect of migration on local adaptation in a coevolving host-parasite system. Nature. 2005; 437: 253–256. PMID: 16148933

34. Tellier A, Brown JK. Spatial heterogeneity, frequency-dependent selection and polymorphism in host-parasite interactions. BMC Evol Biol. 2011; 11: 319. doi: 10.1186/1471-2148-11-319 PMID: 22044632

35. Wang J, Nakad R, Schulenburg H. Activation of the *Caenorhabditis elegans* FOXO family transcription factor DAF-16 by pathogenic *Bacillus thuringiensis*. Dev Comp Immunol. 2012; 37: 193–201. doi: 10.1016/j.dci.2011.08.016 PMID: 21945834

36. Begun DJ, Holloway AK, Stevens K, Hillier LW, Poh YP, Hahn MW, et al. Population genomics: whole-genome analysis of polymorphism and divergence in *Drosophila simulans*. PLoS Biol. 2007; 5: e310. PMID: 17988176

37. Carsiotis M, Stocker BA, Weinstein DL, O'Brien AD. A *Salmonella typhimurium* virulence gene linked to *flg*. Infect Immun. 1989; 57: 3276–3280. PMID: 2680969

38. Ling JM, Moore RA, Surette MG, Woods DE. The *mviN* homolog in *Burkholderia pseudomallei* is essential for viability and virulence. Can J Microbiol. 2006; 52: 831–842. PMID: 17110975

39. Ulland TK, Buchan BW, Ketterer MR, Fernandes-Alnemri T, Meyerholz DK, Apicella MA, et al. Cutting edge: mutation of *Francisella tularensis mviN* leads to increased macrophage absent in melanoma 2 inflammasome activation and a loss of virulence. J Immunol. 2010; 185: 2670–2674. doi: 10.4049/jimmunol.1001610 PMID: 20679532

40. Wei JZ, Hale K, Carta L, Platzer E, Wong C, Fang SC, et al. *Bacillus thuringiensis* crystal proteins that target nematodes. Proc Natl Acad Sci U S A. 2003; 100: 2760–2765. PMID: 12598644

41. Bravo A, Gómez I, Porta H, García-Gómez BI, Rodriguez-Almazan C, Pardo L, et al. Evolution of *Bacillus thuringiensis* Cry toxins insecticidal activity: Evolution of Bt toxins. Microbial Biotechnology. 2013; 6: 17–26. doi: 10.1111/j.1751-7915.2012.00342.x PMID: 22463726

42. Crickmore N. The diversity of *Bacillus thuringiensis* d-endotoxins. In: Charles J, Delécluse A, Nielsen-LeRoux C, editors. Entomopathogenic bacteria: From laboratory to field application. Dordrecht: Kluwer Academic Publishers; 2000. pp. 65–79.

43. Teotonio H, Carvalho S, Manoel D, Roque M, Chelo IM. Evolution of outcrossing in experimental populations of *Caenorhabditis elegans*. PLoS ONE. 2012; 7: e35811. doi: 10.1371/journal.pone.0035811 PMID: 22540006

44. Stiernagle T. Maintenance of *C. elegans*. In: The *C. elegans* Research Community, editor. WormBook. 2006. p.

45. Bentley DR, Balasubramanian S, Swerdlow HP, Smith GP, Milton J, Brown CG, et al. Accurate whole human genome sequencing using reversible terminator chemistry. Nature. 2008; 456: 53–59. doi: 10.1038/nature07517 PMID: 18987734

46. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. Genome Biol. 2009; 10: R25. doi: 10.1186/gb-2009-10-3-r25 PMID: 19261174

47. Li H, Durbin R. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics. 2009; 25: 1754–1760. doi: 10.1093/bioinformatics/btp324 PMID: 19451168

48. Lee W-P, Stromberg MP, Ward A, Stewart C, Garrison EP, Marth GT. MOSAIK: A hash-based algorithm for accurate next-generation sequencing short-read mapping. PLoS ONE. 2014; 9: e90581. doi: 10.1371/journal.pone.0090581 PMID: 24599324

49. Li R, Li Y, Kristiansen K, Wang J. SOAP: short oligonucleotide alignment program. Bioinformatics. 2008; 24: 713–714. doi: 10.1093/bioinformatics/btn025 PMID: 18227114

50. Wu TD, Nacu S. Fast and SNP-tolerant detection of complex variants and splicing in short reads. Bioinformatics. 2010; 26: 873–881. doi: 10.1093/bioinformatics/btq057 PMID: 20147302

51. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. PLoS ONE. 2010; 5: e11147. doi: 10.1371/journal.pone.0011147 PMID: 20593022

52. Wei Z, Wang W, Hu P, Lyon GJ, Hakonarson H. SNVer: a statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data. Nucleic Acids Res. 2011; 39: e132. doi: 10.1093/nar/gkr599 PMID: 21813454

53. Team R. R: A language and environment for statistical computing. R Foundation for Statistical Computing Vienna Austria; 2010.

54. Kim SY, Lohmueller KE, Albrechtsen A, Li Y, Korneliussen T, Tian G, et al. Estimation of allele frequency and association mapping using next-generation sequencing data. BMC Bioinformatics. 2011; 12: 231. doi: 10.1186/1471-2105-12-231 PMID: 21663684

55. Ye K, Schulz MH, Long Q, Apweiler R, Ning Z. Pindel: a pattern growth approach to detect break points of large deletions and medium sized insertions from paired-end short reads. Bioinformatics. 2009; 25: 2865–28671. doi: 10.1093/bioinformatics/btp394 PMID: 19561018

56. Abyzov A, Urban AE, Snyder M, Gerstein M. CNVnator: an approach to discover, genotype, and characterize typical and atypical CNVs from family and population genome sequencing. Genome Res. 2011; 21: 974–984. doi: 10.1101/gr.114876.110 PMID: 21324876

57. Yoon S, Xuan Z, Makarov V, Ye K, Sebat J. Sensitive and accurate detection of copy number variants using read depth of coverage. Genome Res. 2009; 19: 1586–1592. doi: 10.1101/gr.092981.109 PMID: 19657104

58. Peng D, Luo Y, Guo S, Zeng H, Ju S, Yu Z, et al. Elaboration of an electroporation protocol for large plasmids and wild-type strains of *Bacillus thuringiensis*. J Appl Microbiol. 2009; 106: 1849–1858. doi: 10.1111/j.1365-2672.2009.04151.x PMID: 19291242

Supplementary information

Supplementary information for this manuscript can be found at the PLOS Biology website under the following address:
http://journals.plos.org/plosbiology/article?id=10.1371/journal.pbio.1002169#sec022

Additionally, supplementary figures and tables are provided along with the electronic version of the thesis (on DVD), under the following path:

Additional Data:

S1 Data	Supplementary Information/Chapter II.4/Data/S1 Data.xlsx
S2 Data	Supplementary Information/Chapter II.4/Data/S2 Data.xlsx
S3 Data	Supplementary Information/Chapter II.4/Data/S3 Data.xlsx
S4 Data	Supplementary Information/Chapter II.4/Data/S4 Data.xlsx
S5 Data	Supplementary Information/Chapter II.4/Data/S5 Data.xlsx
S6 Data	Supplementary Information/Chapter II.4/Data/S6 Data.xlsx
S7 Data	Supplementary Information/Chapter II.4/Data/S7 Data.xlsx
S8 Data	Supplementary Information/Chapter II.4/Data/S8 Data.xlsx
S9 Data	Supplementary Information/Chapter II.4/Data/S9 Data.xlsx

Additional Figures:

S1 Figure	Supplementary Information/Chapter II.4/Figures/S1 Fig.tif
S2 Figure	Supplementary Information/Chapter II.4/Figures/S2 Fig.tif
S3 Figure	Supplementary Information/Chapter II.4/Figures/S3 Fig.tif
S4 Figure	Supplementary Information/Chapter II.4/Figures/S4 Fig.tif
S5 Figure	Supplementary Information/Chapter II.4/Figures/S5 Fig.tif

Additional Tables:

S1 Table	Supplementary Information/Chapter II.4/Tables/S1 Table.docx
S2 Table	Supplementary Information/Chapter II.4/Tables/S2 Table.docx
S3 Table	Supplementary Information/Chapter II.4/Tables/S3 Table.docx
S4 Table	Supplementary Information/Chapter II.4/Tables/S4 Table.docx
S5 Table	Supplementary Information/Chapter II.4/Tables/S5 Table.docx
S6 Table	Supplementary Information/Chapter II.4/Tables/S6 Table.docx

S7 Table	Supplementary Information/Chapter II.4/Tables/S7 Table.docx
S8 Table	Supplementary Information/Chapter II.4/Tables/S8 Table.xlsx
S9 Table	Supplementary Information/Chapter II.4/Tables/S9 Table.xlsx
S10 Table	Supplementary Information/Chapter II.4/Tables/S10 Table.docx
S11 Table	Supplementary Information/Chapter II.4/Tables/S11 Table.docx
S12 Table	Supplementary Information/Chapter II.4/Tables/S12 Table.docx
S13 Table	Supplementary Information/Chapter II.4/Tables/S13 Table.docx
S14 Table	Supplementary Information/Chapter II.4/Tables/S14 Table.xlsx
S15 Table	Supplementary Information/Chapter II.4/Tables/S15 Table.xlsx
S16 Table	Supplementary Information/Chapter II.4/Tables/S16 Table.xlsx
S17 Table	Supplementary Information/Chapter II.4/Tables/S17 Table.xlsx
S18 Table	Supplementary Information/Chapter II.4/Tables/S18 Table.xlsx
S19 Table	Supplementary Information/Chapter II.4/Tables/S19 Table.xlsx
S20 Table	Supplementary Information/Chapter II.4/Tables/S20 Table.xlsx
S21 Table	Supplementary Information/Chapter II.4/Tables/S21 Table.docx
S22 Table	Supplementary Information/Chapter II.4/Tables/S22 Table.docx
S23 Table	Supplementary Information/Chapter II.4/Tables/S23 Table.docx
S24 Table	Supplementary Information/Chapter II.4/Tables/S24 Table.docx

II.5 *Bacillus thuringiensis* and *Bacillus weihenstephanensis* inhibit the growth of phytopathogenic *Verticillium* species

Frontiers in Microbiology

Bacillus thuringiensis and *Bacillus weihenstephanensis* inhibit the growth phytopathogenic *Verticillium* species

Authors

Hollensteiner J[1], Wemheuer F[1], Harting R[2], Kolarzyk AM[2,3], Diaz Valerio SM[1], Poehlein A[1], Brzuszkiewicz EB[1], Nesemann K[2], Braus-Stromeyer SA[2], Braus GH[2], Daniel R[1] and Liesegang H[1*]

Front. Microbiol. (2016), 7:2171. doi:10.3389/fmicb.2016.02171

Authors and Contributors

Performed Metagenome experiments: **JH** and FW

Bacterial isolation: AK, **JH**, EB and HL

Genome sequencing, data processing and Assembly: **JH** and AP

Genome annotation, data submission and data interpretation: **JH**, SDV and HL

Antagonistic assays: AK, RH and **JH**

Analyzed and Interpreted data: **JH**, FW, RH and HL

Supervision: RH, SB, GB, KN, RD and HL

Conceived and designed the experiments: HL, RH, RD, and **JH**

frontiers in Microbiology

ORIGINAL RESEARCH
published: 18 January 2017
doi: 10.3389/fmicb.2016.02171

Bacillus thuringiensis and *Bacillus weihenstephanensis* Inhibit the Growth of Phytopathogenic *Verticillium* Species

Jacqueline Hollensteiner[1], Franziska Wemheuer[1], Rebekka Harting[2], Anna M. Kolarzyk[2†], Stefani M. Diaz Valerio[1], Anja Poehlein[1], Elzbieta B. Brzuszkiewicz[1], Kai Nesemann[2], Susanna A. Braus-Stromeyer[2], Gerhard H. Braus[2], Rolf Daniel[1] and Heiko Liesegang[1*]

[1] Göttingen Genomics Laboratory, Department of Genomic and Applied Microbiology, Institute of Microbiology and Genetics, Georg-August-University, Gottingen, Germany, [2] Department of Molecular Microbiology and Genetics, Institute of Microbiology and Genetics and Göttingen Center for Molecular Biosciences, Georg-August-University, Gottingen, Germany

OPEN ACCESS

Edited by:
Yunrong Chai, Northeastern University, USA

Reviewed by:
Gao Xuewen, Nanjing Agricultural University, China
Dennis Halterman, Agricultural Research Service (USDA), USA

***Correspondence:**
Heiko Liesegang
hlieseg@gwdg.de

†Present Address:
Anna M. Kolarzyk, Department of Bacterial Genetics, Institute of Microbiology, University of Warsaw, Warsaw, Poland

Specialty section:
This article was submitted to Plant Biotic Interactions, a section of the journal Frontiers in Microbiology

Received: *06 October 2016*
Accepted: *23 December 2016*
Published: *18 January 2017*

Citation:
Hollensteiner J, Wemheuer F, Harting R, Kolarzyk AM, Diaz Valerio SM, Poehlein A, Brzuszkiewicz EB, Nesemann K, Braus-Stromeyer SA, Braus GH, Daniel R and Liesegang H (2017) Bacillus thuringiensis and Bacillus weihenstephanensis Inhibit the Growth of Phytopathogenic Verticillium Species. Front. Microbiol. 7:2171. doi: 10.3389/fmicb.2016.02171

Verticillium wilt causes severe yield losses in a broad range of economically important crops worldwide. As many soil fumigants have a severe environmental impact, new biocontrol strategies are needed. Members of the genus *Bacillus* are known as plant growth-promoting bacteria (PGPB) as well as biocontrol agents of pests and diseases. In this study, we isolated 267 *Bacillus* strains from root-associated soil of field-grown tomato plants. We evaluated the antifungal potential of 20 phenotypically diverse strains according to their antagonistic activity against the two phytopathogenic fungi *Verticillium dahliae* and *Verticillium longisporum*. In addition, the 20 strains were sequenced and phylogenetically characterized by multi-locus sequence typing (MLST) resulting in 7 different *Bacillus thuringiensis* and 13 *Bacillus weihenstephanensis* strains. All *B. thuringiensis* isolates inhibited *in vitro* the tomato pathogen *V. dahliae* JR2, but had only low efficacy against the tomato-foreign pathogen *V. longisporum* 43. All *B. weihenstephanensis* isolates exhibited no fungicidal activity whereas three *B. weihenstephanensis* isolates showed antagonistic effects on both phytopathogens. These strains had a rhizoid colony morphology, which has not been described for *B. weihenstephanensis* strains previously. Genome analysis of all isolates revealed putative genes encoding fungicidal substances and resulted in identification of 304 secondary metabolite gene clusters including 101 non-ribosomal polypeptide synthetases and 203 ribosomal-synthesized and post-translationally modified peptides. All genomes encoded genes for the synthesis of the antifungal siderophore bacillibactin. In the genome of one *B. thuringiensis* strain, a gene cluster for zwittermicin A was detected. Isolates which either exhibited an inhibitory or an interfering effect on the growth of the phytopathogens carried one or two genes encoding putative mycolitic chitinases, which might contribute to antifungal activities. This indicates that chitinases contribute to antifungal activities. The present study identified *B. thuringiensis* isolates from tomato roots which exhibited *in vitro* antifungal activity against *Verticillium* species.

Keywords: *Bacillus thuringiensis*, *Bacillus weihenstephanensis*, *Verticillium*, bacterial-fungal interaction, antifungal activity, biopesticide, plant pathogen, chitinase

INTRODUCTION

Verticillium wilt occurs in a wide range of plant species including economical important crops. Symptoms include wilting, stunting, vascular discoloration, and early senescence, which cause an annual yield loss of billion dollars worldwide (Pegg and Brady, 2002; Fradin and Thomma, 2006). Causative agents are different soil-borne *Verticillium* species such as *Verticillium dahliae* and *Verticillium longisporum*. *Verticillium longisporum* has a narrow host range mainly infecting *Brassicaceae* (Zeise and von Tiedemann, 2002). Wilting symptoms or crop losses are only observed in the greenhouse (Zeise and von Tiedemann, 2002). In contrast, *V. dahliae* is known as a vascular pathogen with a broad host range including trees, legumes, ornamental crops, and vegetables such as tomato (Pegg and Brady, 2002; Luo et al., 2014). This fungus forms extremely outlasting melanized resting structures (microsclerotia) that are able to survive in soil for many years (Schnathorst, 1981). As consequence, the control of this phytopathogenic fungus is difficult. As many soil fumigants have a severe environmental impact, no pesticide which meets the standards of a sustainable agriculture is currently available to control the expansion of *V. dahliae* and *V. longisporum* (Frank, 2003; Depotter et al., 2016). Other control strategies such as crop rotation, the usage of resistant plant species, and soil solarization have only limited efficiency (Haas and Défago, 2005; Depotter et al., 2016). They are either ineffective, time-consuming, costly, may have a severe environmental impact, or even could affect human health (Angelopoulou et al., 2014).

Many plant-associated bacteria have beneficial effects on their host plant by increasing nutrient availability (Scherling et al., 2009) or by modulating the plant immunity (Jones and Dangl, 2006). Moreover, some of these plant growth-promoting bacteria (PGPB) have been used as biological control agents against plant diseases and pests (Ahemad and Kibret, 2014). However, common fungal antagonistic bacteria such as *Pseudomonas* species have only limited inhibitory impact on *Verticillium* due to the long-term persistence of *Verticillium* microsclerotia in soil (Angelopoulou et al., 2014). Members of the endospore-forming genus *Bacillus* possess a high potential as new fungal antagonists. They provide several advantages compared to other biocontrol agents such as (i) a better life-shell, (ii) a dry-product formulation which contains a lower contamination rate, (iii) established large-scale treatment, and finally, (iv) a cheap and easy usage (Katan, 1981; Fravel, 2005; Haas and Défago, 2005). Some *Bacilli* synthesize antifungal compounds such as cyclic lipopeptides produced by non-ribosomal peptide synthetases (NRPS), polyketide synthases (PKS) or mycolytic enzymes like chitinases (Swiontek Brzezinska et al., 2014; Aleti et al., 2015). Three families of *Bacillus* lipopeptides are known to confer an antifungal effect: surfactins, iturins, and fengycins (Ongena and Jacques, 2008). Strains of the species *Bacillus thuringiensis* (Bt) have been used as biological control agent against insecticidal crop pests for decades (Schnepf et al., 1998).

Bacillus thuringiensis (Bt) is a member of the *Bacillus cereus sensu lato* (Bcsl) complex, which comprises seven species (Priest et al., 2004) including the well-investigated name-giving species *B. cereus* (Bc), an opportunistic human pathogen. Other Bcsl members are *B. anthracis* (Ba), the etiological agent of anthrax, and *B. weihenstephanensis* (Bw). The latter is the only described species, which is able to grow under psychrophilic conditions (Lechner et al., 1998; Helgason et al., 2000). All Bcsl members share a highly conserved chromosomal backbone and have diverse extra-chromosomal elements (Priest et al., 2004). Especially Ba, Bc, and Bt strains showed high identity in a previous study (Priest et al., 1994). Despite their obvious similarity, specific phenotypic characteristics including the presence or absence of virulence genes have been used for differentiation (Helgason et al., 2000). However, the grouping of those species is in a still ongoing debate and it was suggested to reclassify this group to a single species (Okinaka and Keim, 2016). Bt is able to produce parasporal protein crystals consisting of δ-endotoxins (Schnepf et al., 1998). The insecticidal effect of entomopathogenic Bt is attributed to the production of these crystal toxins. In addition to its insecticidal effects, Rocha et al. (2014) reported that Bt serovar *kurstaki* can prevent the growth of the maize pathogen *Fusarium verticilloides*. However, a systematic investigation on different Bt strains isolated from root-associated soil as antagonists of *Verticillium* species has not been performed to date.

The objective of this study was to isolate members of the genus *Bacillus* to investigate their antifungal potential against two important phytopathogenic fungi differing in their host ranges. For this purpose, we generated a *Bacillus* strain collection of 267 isolates from tomato root-associated soil by using an enrichment method for Bt. Twenty *Bacillus* strains from this collection were selected based on unique morphological traits and further investigated for their antifungal activity against *V. dahliae* and *V. longisporum*. The phenotypic and taxonomic classification of the bacterial isolates was determined within the Bcsl species complex. The genomes of the 20 selected strains were sequenced and mined for genes related to crystal toxin as well as antifungal compound production. Moreover, the genome sequences were screened for cyclic lipopeptides produced by NRPS/PKS clusters and mycolytic enzymes such as chitinases. The abundance of the genera within the root-associated bacterial community composition of tomato plants was determined based on the bacterial 16S rRNA amplicon sequencing.

Abbreviations: BGSC, *Bacillus* Genetic Stock Center; NRPS, non-ribosomal polypeptide synthetases; PKS, polyketide synthetases; RiPP, Ribosomally synthesized and post-translationally modified peptides; PGPB, plant growth promoting bacteria; Bcsl, *Bacillus cereus sensu lato;* Bt, *Bacillus thuringiensis;* Bw, *Bacillus weihenstephanensis;* Ba, *Bacillus anthracis;* Bm, *Bacillus mycoides;* Bp, *Bacillus pseudomycoides;* Bc, *Bacillus cereus;* Bcyt, *Bacillus cytotoxicus*.

MATERIALS AND METHODS

Soil Sampling and Extraction of Environmental DNA

Soil samples of *Solanum lycopersicum* were collected in a field plot near Göttingen (Germany, 51° 32′N, 9° 56′O) in June 2014. Samples of topsoil with 10 g each were taken next to roots (<3–5 mm) of three tomato plants and subsequently frozen at −80°C. Prior to DNA extraction, visible roots were removed.

Afterwards, samples were treated with mechanical disruption using a microdismembrator (B. Braun Biotech International, Melsungen, Germany) for 3 min with 2000 rpm. Disrupted samples were dissolved in 600 µl sterile water. The DNA was directly extracted using the PowerSoil® DNA isolation kit (MO BIO Laboratories, Inc., Carlsbad, USA) as described by the manufacturer. The quality and purity of extracted DNA was tested with gel electrophoresis and NanoDrop ND-1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, USA), respectively.

Amplification of 16S rRNA Gene

The V6–V8 region of the bacterial 16S rRNA gene was amplified with primers containing the Roche 454 pyrosequencing adaptors, keys as well as one unique MID per sample (underlined): F968 5′-CCATCTCATCCCTGCGTGTCTCCGAC-TCAG-$(dN)_{10}$-AA CGCGAAGAACCTTAC-3′ and R1401 5′-CCATCTCATCCC TGCGTGTCTCCGAC-TCAG-CGGTGTGTACAAGACCC-3′ (Nübel et al., 1996). The PCR (25 µl) contained following final concentrations: one-fold Phusion GC buffer (Thermo Scientific): 0.2 mM of each of the four deoxynucleoside triphosphates (Thermo Scientific), 0.2 µM of each primer, 0.04 U of Phusion high fidelity hot start DNA polymerase (Thermo Scientific), 5% DMSO and ∼4 ng of the isolated soil DNA as template. The following thermal cycling scheme was used: initial denaturation at 98°C for 30 s, 30 cycles of denaturation at 98°C for 15 s, annealing at 53°C for 30 s, extension at 72°C for 30 s, followed by an additional extension step at 72°C for 2 min. Negative controls were performed using the reaction mixture without template. Genomic DNA of *Bacillus* was used as positive control. Three independent PCRs were performed per sample. Obtained PCR products were controlled for appropriate size and subsequently purified using the peqGOLD Gel Extraction Kit (Peqlab, Erlangen, Germany, now VWR) as recommended by the manufacturer. PCR products were quantified using the Quant-iT dsDNA HS assay kit and a Qubit fluorometer as recommended by the manufacturer (Thermo Scientific). Purified PCR products from the three independent PCRs were subsequently pooled in equal amounts. The Göttingen Genomics Laboratory determined the 16S rRNA gene sequences employing the Roche GS-FLX+ pyrosequencer with Titanium chemistry as recommended by the manufacturer (Roche, Mannheim, Germany).

Processing and Analysis of 454 Pyrosequencing Derived Data

Pyrosequencing derived 16S rRNA gene data were preprocessed with QIIME version 1.8 (Caporaso et al., 2010). Preprocessing included the removal of short reads (<300 bp) as well as reads containing long homopolymer stretches (>8 bp) and too many primer mismatches (>3 bp) in the forward primer. Filtered data was subsequently denoised employing Acacia version 1.53b (Bragg et al., 2012). Remaining primer sequences were truncated employing cutadapt version 1.0 (Martin, 2011). Chimeric sequences were removed using USEARCH version 7.0.1090 (Edgar, 2010). For this purpose, sequences were first dereplicated in full-length mode and putative chimeras were initially removed using the UCHIME algorithm in *de novo* mode and subsequently in reference mode using the most recent SILVA database (SSURef 119 NR) as reference dataset (Edgar et al., 2011; Klindworth et al., 2013). Afterwards, processed sequences of all samples were joined and clustered in operational taxonomic units (OTUs) at 3% genetic dissimilarity according to Wemheuer et al. (2012) employing the UCLUST algorithm (Edgar, 2010). To determine taxonomy, a consensus sequence for each OTU was classified by BLAST alignment against the Silva SSURef 119 NR database (Camacho et al., 2009). Alpha diversity indices were calculated with QIIME as described by Wemheuer et al. (2014). Rarefaction curves were calculated in QIIME and subsequently interpolated in R (R Core Team, 2014) using the "drc" package [https://cran.r-project.org/web/packages/drc/]. The statistical analysis was performed in R.

Isolation and Enrichment of Bacteria

Bacterial strains were obtained from root-associated soil samples. Soil samples were enriched for *Bacillus* strains, especially for Bt subspecies as described by Patel et al. (2011) using a modified glucose yeast extract salt (GYS) sporulation medium, which contained the following ingredients per liter: 1 g glucose, 2 g yeast extract, 2 g $NH_4(SO_4)_2$, 0.06 g $MnSO_4$, 0.4 g $MgSO_4 \cdot 7H_2O$, 0.08 g $CaCl_2$ and 5 g KH_2PO_4. Resulting colonies were picked with sterile toothpicks and inoculated in 96-deep well plates containing 2 ml two-fold concentrated lysogeny broth (TLB; Bone and Ellar, 1989). After overnight shaking at 30°C with 50 rpm (Orbitron S-000119510, Infors HT, Bottmingen, CHE), all colonies were stamped on lysogeny broth (LB; Sambrook and Russell, 2001) agar plates using a sterile 96 deep well plate replicator. The plates were incubated at 30 and 21°C for 24 h.

Phenotypic Analysis of Isolates

Specific growth of cells was monitored after incubation at 30°C for 24 h using Bino Olympus SZX12, Olympus SC30 camera and the Olympus Cell Sense software. Different media were tested, nutrient rich liquid LB as well as modified solid nutrient limiting Simulated Xylem Medium (SXM). SXM was prepared according to (Dixon and Pegg, 1971) which contained following ingredients per liter: 2 g pectin, 4 g casein, 20 ml AspA (50x), 2 ml $MgSO_4$ (1M), 1 ml trace elements (1000x; de Serres and Hollaender, 1982), and 20 g agar. AspA (50x) contained the following ingredients per liter: 300 g $NaNO_3$; 26 g KCl and 76 g KH_2PO_4. Strains were tested for hemolytic activity on Columbia blood agar with Oxoid sheep-blood "plus" (Thermo Scientific). Hemolytic activity was monitored and measured after 48 h.

Classification of Isolated Bacterial Strains

Genomic DNA of each strain was extracted from overnight cultures grown in LB medium at 30°C with 150 rpm (New Brunswick Incubator Shaker Innova 2300, Neu-Isenburg, Germany) by using the MasterPure™ Complete DNA Purification Kit (Epicentre, Madison, USA) according to the manufacturer's instructions. The purified DNA served as template for multi-locus sequence typing (MLST) amplification by PCR according to Priest et al. (2004). Primers for seven house-keeping genes (*glpF, gmk*, *ilvD*, *pta*, *purH*, *pycA,* and *tpi*) are listed in Supplementary Table 1. The PCR mixture (50 µl)

contained following final concentrations: one-fold OptiBuffer, 3 mM $MgCl_2$, 0.04 U BIO-X-ACT™ short DNA polymerase (Bioline, London, UK), 1 mM deoxynucleoside triphosphates (Thermo Scientific), 0.4 μM of each primer, sterile water, and ~2 ng of bacterial DNA as template. Thirty cycles were conducted in a Mastercycler (Eppendorf, Hamburg, Germany). An initial denaturation at 98°C for 5 min was followed by 45 s denaturation at 98°C, 45 s annealing at 65°C, and 30 s primer extension at 72°C with a final step of 72°C for 5 min. PCR products were controlled for appropriate size and purified employing the QIAquick PCR purification kit (QIAGEN, Hilden, Germany) and eluted in 35 μl sterile water. Quantification of PCR products was performed via Nanodrop 1000 Spectrophometer (Thermo Scientific). The Göttingen Genomics Laboratory determined the sequences of the PCR products. For classification of the new isolates, all MLST genes from Bcsl representative genomes at time of analysis were considered. The corresponding sequences were obtained from GenBank hosted at the National Center for Biotechnology Information (http://www.ncbi.nlm.nih.gov/genome/genomes/486?). To exclude putative Ba isolates, all strains were tested for specific Ba virulence factors. Each MLST gene was aligned to determine the corresponding regions of each housekeeping gene and was trimmed to shortest shared region, resulting in orthologous sequences from 298 to 394 bp, as previously described (Jolley et al., 2004). Analysis was performed with concatenated DNA sequences of all seven gene loci. For construction of a phylogenetic tree, MEGA software 7.0.14 was used (Kumar et al., 2016). Sequence alignment was performed using the ClustalW algorithm and a phylogenetic tree was constructed by using the neighbor-joining method (Saitou and Nei, 1987). The robustness of the tree was evaluated by bootstrap analysis with 1000 resamplings.

Whole Genome Sequencing and Assembly

DNA extracted from isolated bacteria was subjected to whole-genome sequencing (**Table 1**) using the MiSeq sequencer (Illumina, San Diego, USA). For this purpose, Nextera_XT (Illumina, San Diego, USA) paired-end libraries (2 × 301 bp) were prepared according to the manufacturer's protocols. Resulting reads were quality-filtered with Trimmomatic 0.32 (Bolger et al., 2014) and evaluated using Fastqc (Bahabram Informatics, Babraham Institute; UK). Spades 3.5.0 (Bankevich et al., 2012) was used to assemble processed reads. For scaffolding, the resulting contigs were aligned to reference genomes Bt 407 (CP003889) and Bw KBAB4 (CP000902) using the Mauve Genome Alignment software (Darling et al., 2004). Sequencing results and genome characteristics are summarized in **Table 1**. Automatic annotation was carried out with the IMG-ER (Integrated Microbial Genomes-Expert Review) system (Markowitz et al., 2009) and with Prokka v1.9 (Seemann, 2014) using Bt 407 (Sheppard et al., 2013) as species reference and a comprehensive toxin protein database (including Cry, Cyt, Vip, Sip proteins). The Prokka pipeline was applied using gene calling by prodigal (Hyatt et al., 2010), rRNA genes and tRNA genes identification with RNAmmer 1.2 (Lagesen et al., 2007) and Aragorn (Laslett and Canback, 2004), respectively. Additionally, signal peptides were identified with SignalP 4.0 (Petersen et al., 2011) and non-coding RNA species with an Infernal 1.1 search against the Rfam database (Eddy, 2011).

TABLE 1 | Genome statistics of the 20 selected isolates.

Strain	Genome size [bp]	Number of contigs	Mean coverage	GC [%]	Total number of genes	Protein coding genes	RNA genes	CRISPR	GenBank accession
Bt GOE1	5,359,363	30	214.5	35.1	5499	5389	109	0	LXLF00000000
Bt GOE2	5,373,416	43	122.1	35.1	5493	5383	110	0	LXLG00000000
Bt GOE3	5,347,504	60	146.1	35.1	5455	5340	114	0	LXLM00000000
Bt GOE4	6,008,382	63	110.9	34.8	6079	5967	111	0	LXLH00000000
Bt GOE5	5,772,687	123	123.7	35.1	5829	5736	92	0	LXLN00000000
Bt GOE6	5,857,591	187	75.3	35	6001	5926	74	0	LXLI00000000
Bt GOE7	5,976,466	91	124.2	35	5976	5898	77	1	LXLJ00000000
Bw GOE1	5,613,589	113	100.6	35.3	5742	5617	124	0	LXLK00000000
Bw GOE2	5,581,838	79	97.6	35.2	5722	5611	110	0	LXLO00000000
Bw GOE3	5,644,666	117	100.6	35.2	5787	5676	110	0	LXLP00000000
Bw GOE4	5,597,134	137	96.4	35.2	5733	5614	118	0	LXLQ00000000
Bw GOE5	5,589,822	151	110.4	35.3	5679	5564	114	0	LXLR00000000
Bw GOE6	5,632,050	118	98.1	35.2	5761	5643	117	0	LXLS00000000
Bw GOE7	5,772,687	126	185.2	35.2	5812	5699	112	0	LXLT00000000
Bw GOE8	5,851,749	156	113	35.2	5989	5882	106	0	LXLU00000000
Bw GOE9	5,846,892	144	142.2	35.2	5969	5885	83	0	LXLV00000000
Bw GOE10	5,823,784	143	78.1	35.2	5963	584	122	1	LXLW00000000
Bw GOE11	5,625,374	105	215.3	35.2	5776	5654	121	0	LXLX00000000
Bw GOE12	5,674,772	128	103.5	35.3	5805	5686	118	0	LXLY00000000
Bw GOE13	5,642,420	146	101.2	35.2	5763	5648	114	0	LXLL00000000

Genome Analysis of Bacterial Isolates

The genomes of all strains were screened for genes encoding Bt-specific toxins, such as crystal toxins, cytolytic toxins, vegetative insecticidal protein-toxins, and secreted insecticidal protein toxins (Cry, Cyt, Vip, and Sip toxins), and for candidate-genes for the production of antifungal secondary metabolites and chitinases. For identification of Cry-toxins, all protein sequences derived from the genome sequences were scanned using HMMSCAN implemented in HMMER v.3.1b2 (hmmer.org; Eddy, 1998) against a HMM-profile database. The profile database was generated from holotype Cry toxin sequences retrieved from the official BT Toxin Nomenclature website (http://www.btnomenclature.info/) and Uniprot (Bateman et al., 2015), respectively. To generate toxin-specific HMMs, all toxin sequences were grouped by Markov Cluster Algorithm v.12068 (Enright et al., 2002; van Dongen, 2007) according to the default options with an inflation value (−I) of 10.0. Representative members of each cluster were aligned by ClustalW v.1.83 (Thompson et al., 1994) and used as input to build the model database. After optimization of the models for sensitivity and specificity, the resulting HMMs were used for detection. Resulting group hits were further classified accordingly to the Cry Toxin Nomenclature. For the detection of Sip, Vip, and Cyt toxins, all protein sequences annotated accordingly from Uniprot (Bateman et al., 2015), National Centre for Biotechnology Information (NCBI, http://www.ncbi.nlm.nih.gov/), and the holotype-toxin list (http://www.btnomenclature.info/) were controlled for duplicates by 100% sequence identity. Only one representative was kept to create a BLASTp database, which was performed for each genome with an *e*-value of $1e^{-50}$. The analysis for secondary metabolites, i.e., polyketides, alkaloids, terpenes, phenanzines, microlides, and non-ribosomal antibiotic peptides was performed with antiSMASH 3.0 available at (https://antismash.secondarymetabolites.org/) (Weber et al., 2015). Chitinases were identified and compared to chitinase sequences of other Bcsl members. They were aligned using the ClustalW algorithm (Thompson et al., 1994) in MEGA7 (Kumar et al., 2016). To construct a phylogenetic tree, the Neighbor-Joining method (Saitou and Nei, 1987) and evolutionary distances were computed by the Maximum Composite Likelihood method (Tamura et al., 2004). InterPro (Mitchell et al., 2015) was used to group detected chitinases into chitinase families and for the prediction of domains, signatures, and active sites.

TABLE 2 | Organisms used in this study.

Abbreviation	Strain	Source
4D2	*Bacillus thuringiensis* subsp. *kurstaki* 2	BGSC
4Q1	*Bacillus thuringiensis* subsp. *israelensis* ONR60A	BGSC
Bt GOE1	*Bacillus thuringiensis* isolate	This study, personal request
Bt GOE2	*Bacillus thuringiensis* isolate	This study, personal request
Bw GOE1	*Bacillus weihenstephanensis* isolate	This study, personal request
Bt GOE3	*Bacillus thuringiensis* isolate	This study, personal request
Bw GOE2	*Bacillus weihenstephanensis* isolate	This study, personal request
Bt GOE4	*Bacillus thuringiensis* isolate	This study, personal request
Bt GOE5	*Bacillus thuringiensis* isolate	This study, personal request
Bt GOE6	*Bacillus thuringiensis* isolate	This study, personal request
Bw GOE3	*Bacillus weihenstephanensis* isolate	This study, personal request
Bw GOE4	*Bacillus weihenstephanensis* isolate	This study, personal request
Bw GOE5	*Bacillus weihenstephanensis* isolate	This study, personal request
Bw GOE6	*Bacillus weihenstephanensis* isolate	This study, personal request
Bw GOE7	*Bacillus weihenstephanensis* isolate	This study, personal request
Bw GOE8	*Bacillus weihenstephanensis* isolate	This study, personal request
Bw GOE9	*Bacillus weihenstephanensis* isolate	This study, personal request
Bw GOE10	*Bacillus weihenstephanensis* isolate	This study, personal request
Bw GOE11	*Bacillus weihenstephanensis* isolate	This study, personal request
Bw GOE12	*Bacillus weihenstephanensis* isolate	This study, personal request
Bw GOE13	*Bacillus weihenstephanensis* isolate	This study, personal request
Bt GOE7	*Bacillus thuringiensis* isolate	This study, personal request
Bt MYBT18246	*Bacillus thuringiensis* MYBT18246	Available at Schulenburg lab[a] (Masri et al., 2015)
Bt Bt18247	*Bacillus thuringiensis* Bt18247	Available at Schulenburg lab[a] (Masri et al., 2015)
Bt Bt18679	*Bacillus thuringiensis* Bt18679	Available at Schulenburg lab[a] (Masri et al., 2015)
Btt	*Bacillus thuringiensis* biovar *tenebrionis*	Available at Kurtz lab[b] (Milutinovic et al., 2014)
DH5α	*Escherichia coli* DH5α (*fhuA2 Δ(argF-lacZ)U169 phoA glnV44 Φ80Δ (lacZ)M15 gyrA96 recA1 relA1 endA1 thi-1 hsdR17)*	New England BioLabs, C2989K
Vd JR2	*Verticillium dahliae* JR2	Fradin et al., 2009
Vl 43	*Verticillium longisporum* 43	Zeise and von Tiedemann, 2002

Supplementary information on organisms used in this study.
[a]*Department of Evolutionary Ecology and Genetics, Zoological Institute, Christian-Albrechts-University of Kiel, Kiel, Germany.*
[b]*Animal Evolutionary Ecology Group, Institute for Evolution and Biodiversity, University of Münster, Münster, Germany.*

Co-cultivation Assay

To evaluate the antifungal potential of 20 phenotypically diverse *Bacillus* strains, co-cultivation assays with the two phytopathogens *V. dahliae* JR2 (Fradin et al., 2009) and *V. longisporum* 43 (Zeise and von Tiedemann, 2002) were conducted. All used strains are listed in **Table 2**. The insecticidal Bt subsp. *kurstaki* 2 and the Bt subsp. *israelensis* ONR60A strain were obtained from the *Bacillus* Genetic Stock Centre (BGSC; Columbus; USA) and served as reference strains. Three nematocidal strains (Bt MYBT18246, Bt Bt18247, and Bt Bt18679) and one insecticidal Bt strain (Btt) were used (Milutinovic et al., 2013; Masri et al., 2015). *Escherichia coli*

TABLE 3 | Diversity (represented by the Shannon index H') and richness (number of observed OTUs) of tomato-associated bacterial communities.

Sample	Richness	Diversity	Michaelis-Menten-Fit	Coverage
Rep1	740.6	5.82	1006.46	0.74
Rep2	673.0	5.72	898.02	0.75
Rep3	740.3	5.81	1033.25	0.72

DH5α served as negative control. Co-cultivation assays with two phytopathogenic fungi and selected *Bacillus* strains were performed using three different media. For standard bacterial cultivation, LB was used. The complex pectin containing SXM with limited nutrition tries to simulate the nutrient conditions which are available for a fungus in a plant. Due to the restriction of a synthetic medium it is limited in simulating the natural plant environment, where water, minerals, organic acids and ~500 µm amino acids are available in highly specific concentrations (Singh et al., 2010). The potato dextrose medium (PDM; Carl Roth GmbH, Karlsruhe, Germany with 5 g/L agar), a full medium containing potato starch and glucose, were used for optimal cultivation of *Verticillium* (Gams et al., 1998). Solid media were prepared containing 2% agar. Approximately 1×10^5 harvested *Verticillium* conidiospores (Beckman Coulter Counter Size Analyzer, Krefeld, Germany) were homogeneously inoculated per plate using sterile glass beads (2.85–3.45 mm in diameter). A hole (diameter = 0.9 cm) was excised in the middle of the plate. This *in vitro* setting was designed to study *Bacillus*/*Verticillium* interactions under controlled conditions simulating the substrate supply of a plant and animal-associated habitat.

Bacteria were inoculated in 5 ml LB rotating with 120 rpm (New Brunswick Incubator Shaker Innova 2300, Neu-Isenburg, Germany) overnight at 37°C (Bt strains) or 25°C (Bw strains). Isolates were centrifuged for 2 min with 6000 rpm and washed with 2 ml sterile water. Afterwards, 5 ml of LB, PDM, or SXM were inoculated and adjusted to an OD_{595} of 0.1. Main

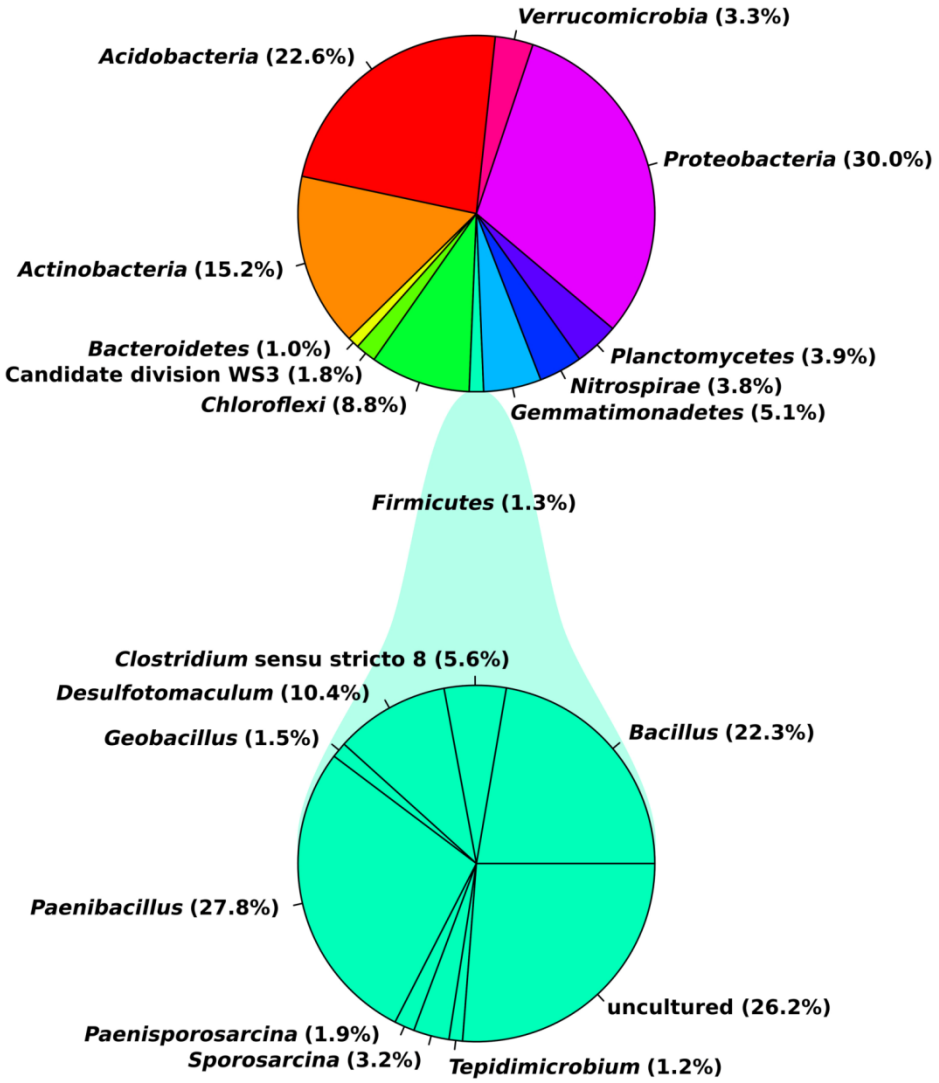

FIGURE 1 | General soil bacterial community composition at phylum level and for the *Firmicutes* at genus level. Plots were generated in R.

PDM Bw cultures were inoculated to an OD_{595} of 0.01 and grown overnight at 25°C to an OD_{595} of 1. All other main cultures were grown at 30 or 25°C rotating with 120 rpm (New Brunswick Incubator Shaker Innova 2300, Neu-Isenburg, Germany) to an OD_{595} of 1. When the optical density was reached, 60 µl of the respective bacterial cultures were added into the hole in the plates. Plates were incubated at 25°C for 7 days. Zones of fungal growth inhibition were measured and quantified. Media without bacterial cells were used as positive controls for fungal growth. Each treatment was performed in technical triplicates and biological duplicates. To analyze possible differences between the isolates, a repeated measures ANOVA (Crawley, 2007) was conducted in R due to spatial pseudoreplication.

Deposition of Isolated Strains and Genome Sequences

Sequence data were deposited in the Sequence Read Archive (SRA) of the NCBI under the accession number SRA401353. Genome sequence data were deposited in JGI and GenBank (**Table 1**). Strains are available on request (**Table 2**).

RESULTS AND DISCUSSION

Soil Bacterial Communities of Tomato Plants

Root-associated bacterial community composition and diversity were assessed by amplicon-based analyses of the V6-V8 region

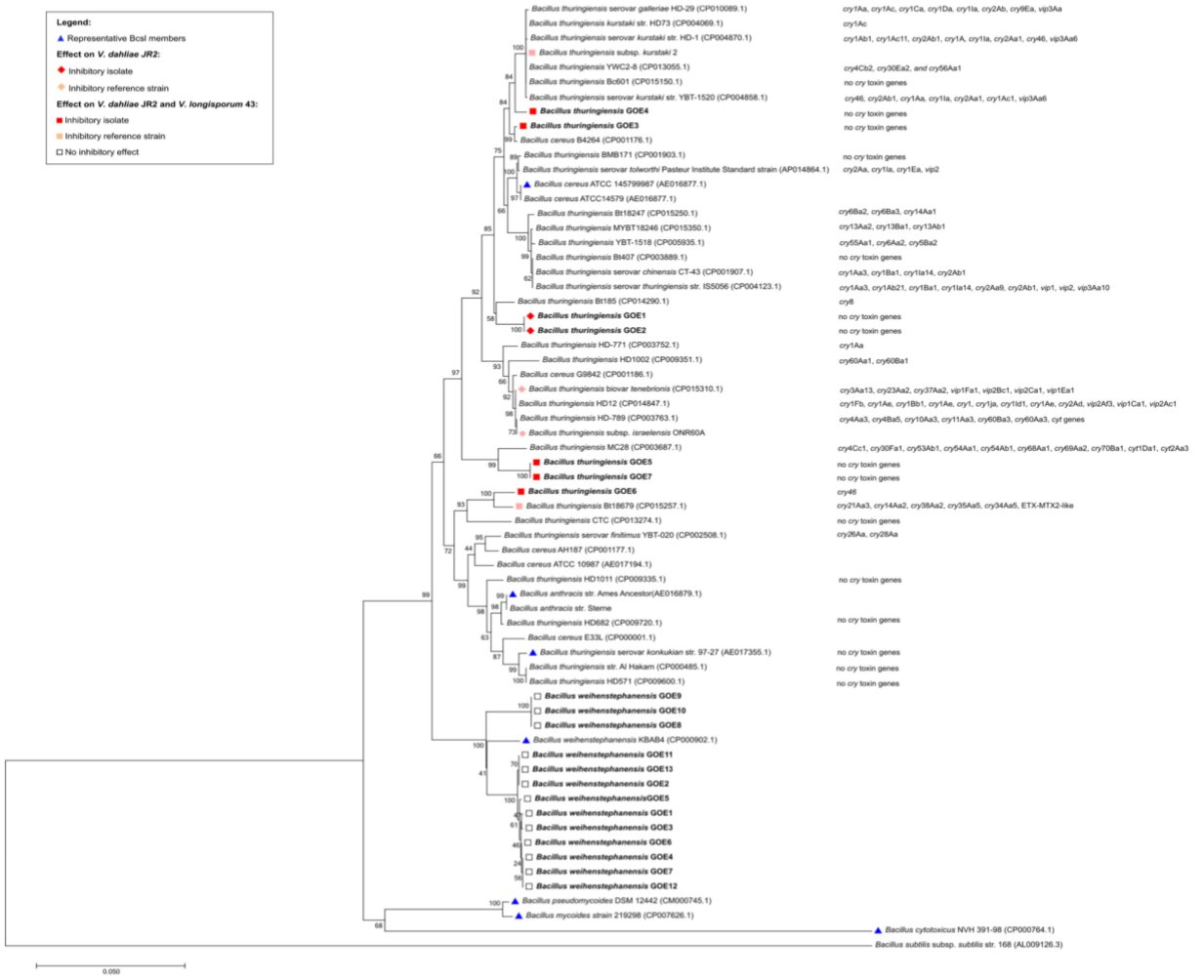

FIGURE 2 | Phylogenetic analysis of 20 Bcsl group isolates. The 20 newly isolated strains are depicted in bold. (Blue triangle) Representatives of the Bcsl group species members. Strains which have been tested and showed an effect on *V. dahliae* JR2 or both *V. dahliae* JR2 and *V. longisporum* 43 are indicated by a diamond or square, respectively. Effects on *V. dahliae* JR2 (red diamond): inhibitory isolates, (beige diamond) inhibitory reference strains. Effect on *V. dahliae* JR2 and *V. longisporum* 43 (red square): inhibitory isolate, (beige square) inhibitory reference strain, (white square) no inhibitory effect. The phylogenetic analysis was performed using MEGA7 (Kumar et al., 2016), details are described in the Materials and Methods Section.

of the bacterial 16S rRNA gene. After quality filtering, denoising, and removal of potential chimeras and non-bacterial sequences, a total of 13,596 high-quality sequences were retrieved and used for further analyses. These sequences were grouped into 3994 OTUs. Calculated rarefaction curves (Supplementary Figure 1) as well as the mean coverage at species level (**Table 3**) revealed that the majority of bacterial community was recovered by the surveying effort. Richness (number of observed OTUs) and diversity (Shannon indices) for bacterial communities ranged from to 673 to 740.6 and 5.72 to 5.82, respectively (**Table 3**). All sequences could be classified below phylum level.

Eleven abundant phyla (>1% of all sequences across all samples) were present in each soil sample and accounted for more than 96% of all bacterial sequences analyzed (**Figure 1**). *Proteobacteria* (30.0%), *Acidobacteria* (22.6%) and *Actinobacteria* (15.2%) were the most abundant (>1% abundance

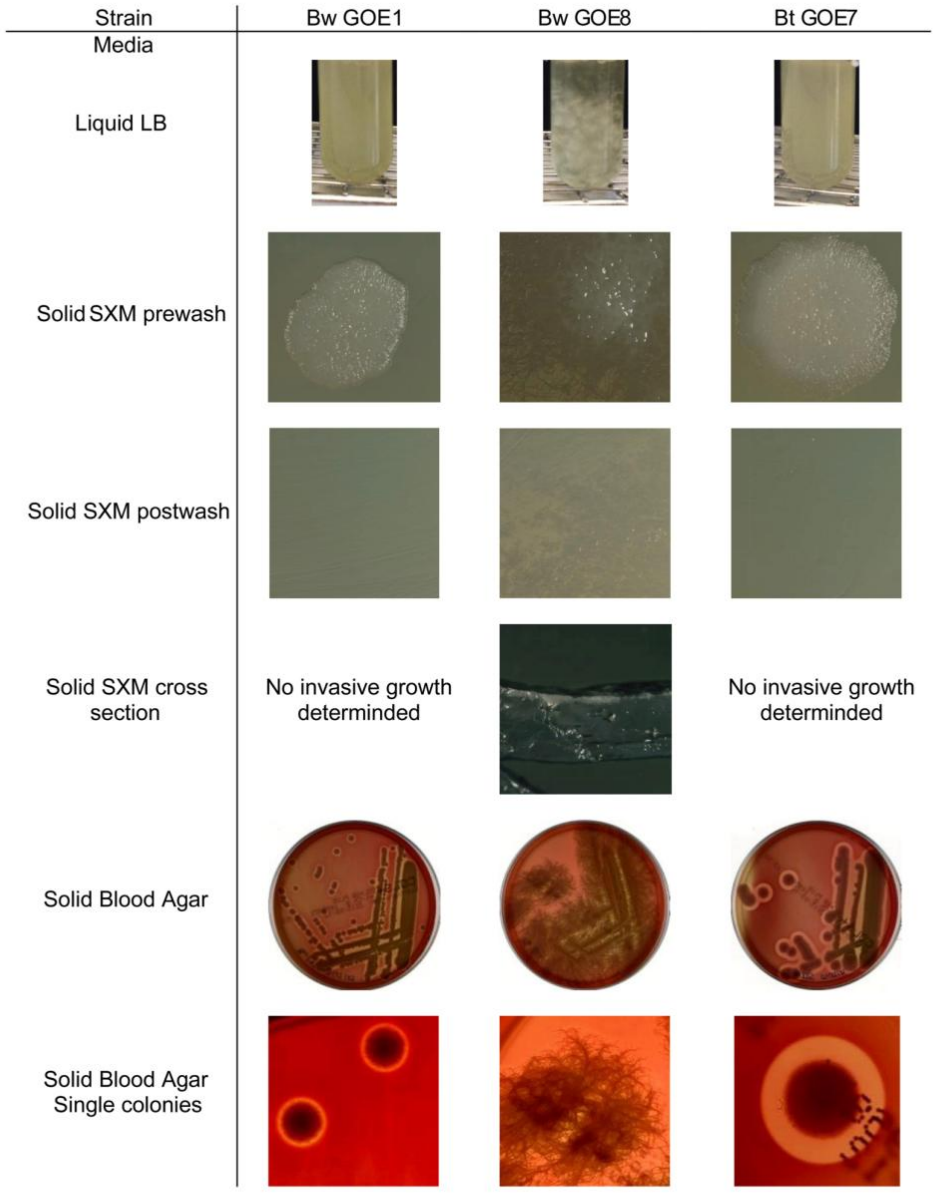

FIGURE 3 | Representative colony morphology phenotypes of tomato isolates. The growth of three different strains, Bw GOE1, Bw GOE8, and Bt GOE7, was tested on several media. (Top row) Strains in liquid LB media, (2nd, 3rd, and 4th row) close up of single colonies on solid SXM media, pre/after scraping and washing with spatula and sterile water, and cross section of the agar. (5th and bottom row) Growth on solid blood agar and close up of single colonies. Strain abbreviations are explained in **Table 2**.

of all sequences) bacterial phyla. The *Proteobacteria* were mainly represented by *Beta-* and *Gammaproteobacteria*. Other prominent phyla were *Chloroflexi* (8.8%), *Verrucomicrobia* (3.3%) and *Firmicutes* (1.3%). Although their relative abundance is variable, these phyla have been previously found in different soils (Shange et al., 2012) and in the rhizosphere of different plant species (Romero et al., 2014; Pii et al., 2016). Li J. G. et al. (2014) investigated bacterial communities in roots and rhizosphere soils of healthy and diseased tomato plants and found that *Proteobacteria* was the most abundant phylum across all samples, followed by *Actinobacteria* and *Bacteroidetes*. This is in line with a study of Bulgarelli et al. (2015) on bacterial communities in bulk soil, rhizosphere, and roots of barley. In contrast to the above-mentioned findings, *Verrucomicrobia* (24%), *Acidobacteria* (23%) and *Proteobacteria* (17%) were the major taxonomic groups in rhizosphere bacterial communities of tomato (Romero et al., 2014).

In our study, observed genera included *Blastocatella* (2.6%), *Massilia* (1.64%) *Gaiella* (2.0%), *Nitrospira* (1.76%), *Paenibacillus* (0.34%), *Bacillus* (0.27%), *Bradyrhizobium* (0.17%), *Burkholderia* (0.04%), and *Pseudomonas* (0.006%; Supplementary Table 2). Most sequences of the *Firmicutes* were assigned to the two genera *Paenibacillus* (27.8%) and *Bacillus* (22.3%). Several of the observed genera including *Bacillus*, *Burkholderia*, and *Pseudomonas* are reported as the most significant phosphate-solubilizing bacteria (Bhattacharyya and Jha, 2012). Moreover, members of the genera *Bacillus*, *Paenibacillus*, *Bradyrhizobium, Pseudomonas*, and *Burkholderia* are known for their plant growth-promoting functions and/or their use as biocontrol agents against different phytopathogens and pests (Bhattacharyya and Jha, 2012; Glick, 2012). In a previous investigation on cultivable bacteria associated with tomato leaves, *Bacillus* (*Firmicutes*) showed strong *in vitro* antifungal activity against three important pathogens of tomato (*Botrytis cinerea*, *Fulvia fulva*, and *Alternaria solani*; Enya et al., 2007).

Classification of Bacterial Isolates

The whole Bcsl taxonomy is recently discussed based on the different methods that are used for strain classification (Okinaka and Keim, 2016). Depending on the specific focus of the study, pathogenic properties or taxonomic features are used for classification resulting in an inconclusive taxonomy. We tried to address this question by performing MLST analysis as described by Priest et al. (2004) in combination with the analysis of phenotypic, biochemical and pathogenic characteristics. Our strain collection comprised 267 *Bacillus* strains. Twenty isolates with diverse colony morphologies were selected for further analysis.

For the genome analysis of strains, the DNA was isolated and sequenced. Accession numbers and sequencing metadata are summarized in **Table 1**. The genome sizes ranged from 5.3 to 6.0 MB with a mean GC-content of 35%, which is typical for members of the Bcsl group (http://www.ncbi.nlm.nih.gov/genome/genomes/486?). The total number of genes varied from 5456 to 6080 and the number of identified RNA genes from 157 to 194. As the taxonomy of the Bcsl group is complex, the classical 16S rRNA phylogeny used for sequence based species assignments was ineffective for differentiation within this group. Based on MLST, 7 *B. thuringiensis* (Bt GOE1-7) and 13 *B. weihenstephanensis* (Bw GOE1-13) strains were identified (**Table 1**, **Figure 2**). None of the isolates clustered in close proximity to the human pathogen Ba or to any of the other known human pathogenic strains from the Bcsl species group. The analysis of similarity using the known Ba toxin genes or capsule genes such as *lef*, *cya*, *pagA*, and *capA-capC* showed that none of these genes are present in the new isolates. Additionally, there were no similarities to the four species-specific prophages (lambda01-lambda04), which are used for the PCR-based identification of Ba strains (Kolstø et al., 2009). All seven Bt isolates encode genes for the PlcR regulator with 100% sequence identity to Bt-reference sequences whereas no Bw isolate contained a *plcR* gene (Supplementary Table 3).

The morphology of Bcsl members was previously described in general as similar: the colony shape is irregular with undulate or curled margins (De Vos et al., 2009). Moreover, the colonies are flat to raised and opaque while Bm and Bp exhibit a mycoid to rhizoid growth with hairy-looking adherent colonies (De Vos

TABLE 4 | Hemolytic activity assay.

Bacterial Strain	Hemolytic activity on blood agar plates
E. coli DH5α	–
4D2*	++
4Q1*	++
Bt GOE1	++
Bt GOE2	++
Bw GOE1	+
Bt GOE3	++
Bw GOE2	+
Bt GOE4	++
Bt GOE5	++
Bt GOE6	++
Bw GOE3	+
Bw GOE4	+
Bw GOE5	+
Bw GOE6	+
Bw GOE7	+
Bw GOE8	nd
Bw GOE9	nd
Bw GOE10	nd
Bw GOE11	+
Bw GOE12	+
Bw GOE13	+
Bt GOE7	++
Bt MYBT18246**	–
Bt Bt18247**	+
Bt Bt18679**	++
Btt*	++

*Symbols refer to the radius of the zone of clearance (mm): –, absence of zone; presence of zone of clearance: + = 0.5–2 mm; ++ ≥ 2 mm; *insecticidal strains; **nematocidal strains; nd, not determined due to growth structure.*

et al., 2009). The morphological analysis of 20 isolates revealed two different growth-types in liquid as well as on solid media (**Figure 3**). The first growth type is represented by 7 Bt and 10 Bw isolates which formed circular to weakly irregular colonies with entire or undulate edges. The surface texture was ground glass to granular. The Bt isolates exhibited larger single colony sizes compared to Bw strains. These 17 strains were hemolytic positive on blood agar (**Table 4**, **Figure 3**). In liquid LB medium, the bacterial solutions were white yellowish and grew equally cloudy. In contrast, three Bw isolates (Bw GOE8-10) exhibited a Bm-like colony shape, with rhizoid growth on solid medium. The cells grew adherent and covered rapidly the whole agar plate as it is known for Bm which is in contrast to our MLST analysis (**Figure 3**; Di Franco et al., 2002). Moreover, the isolates displayed an invasive growth into the media plates (**Figure 3**). To the best of our knowledge, this has not been described previously for Bw strains. In addition, hemolytic activity could not be determined. In liquid medium, the strains produced an aggregation of clumps comparable to Bm (Di Franco et al., 2002). All Bw isolates showed mesophilic growth at 30°C instead of the described psychrophilic growth (Di Franco et al., 2002).

To determine the pathogenic properties of the isolates, we screened for virulence factors specific for the Bt species group (Cry, Cyt Sip, and Vip toxins). Within the MLST tree (**Figure 2**), the Bt isolates clustered with *B. thuringiensis* reference strains, with present or absent Cry toxins. In contrast, all Bw isolates clustered with *B. weihenstephanensis* references (**Figure 2**). The strains Bt Bc601, Bt BMB171, Bt CTC, Bt HD1011, Bt HD682, Bt str. Al Hakam, and Bt HD571 contained no homologs to any known Cry, Cyt, or Vip toxins (**Figure 2**). Notably, even in the type strain Bt serovar *konkukian* str. 97-27, which was isolated from a necrotic human wound, no full-length hits to known Cry, Cyt, or Vip toxins were identified (Han et al., 2006). Only Bt GOE6 contained a gene with similarity to Cry46. In all other isolates, no homologs to any known Cry, Cyt, Sip, or Vip toxins were detected using BLASTp and HMM models. Moreover, proteins sharing a domain with Cry6 and Cry22 were detected in all Bt and Bw isolates. Modeling and detection of known as well as for new Cry toxins is not trivial based on their variable structures. Cry toxins can be sub-grouped into three classes: Three-domain toxins, Bin-toxins and Mtx-toxins (de Maagd et al., 2003). Additionally, some Cry toxins

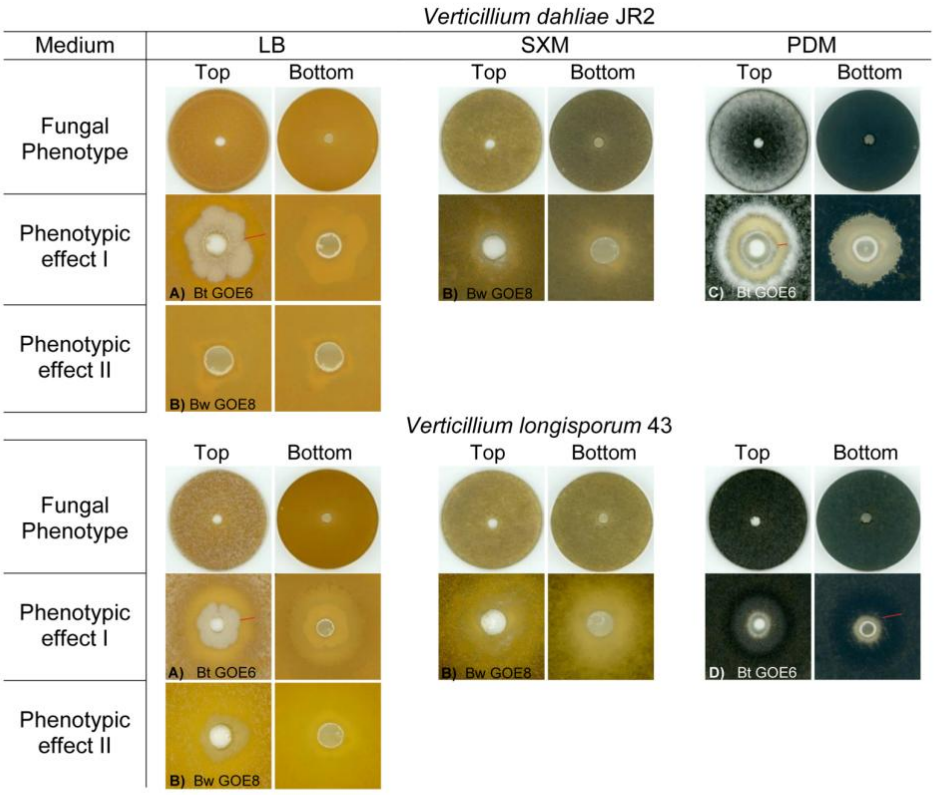

FIGURE 4 | Phenotypic effects of selected bacterial strains against *Verticillia*. Fungal phenotypes without bacteria are displayed in the 1st and 4th row. Specific interaction effects of *V. dahliae* JR2 or *V. longisporum* 43 and antagonistic bacterial strains on different media are depicted: **(A)** Bacteria produce biofilm with or without inhibition of the fungus, **(B)** Bacteria produce hyphae-like structures where the growth of the fungus is reduced, **(C)** White air mycelia is build by the fungus around the bacteria which inhibits *V. dahliae* JR2, **(D)** *V. longisporum* 43 showed a strong melanization surrounding the bacterial application site visible as a dark ring. Inhibition zone and melanization zone of fungi are depicted by a red bar. Strain abbreviations are explained in **Table 2**.

are phylogenetically unrelated and unique such as Cry6, Cry22, Cry34, Cry37, Cry55, and Cry46 (Palma et al., 2014). It is possible that our strains encode novel Cry toxins. The proteins sharing a domain with Cry6 and Cry22 and might represent new toxin candidates. However, Cry toxins are often encoded on mobile elements located on plasmids (González et al., 1982). Cry toxin-encoding Bt plasmids may be instable and thus can get lost within cultivation (Masri et al., 2015; Sheppard et al., 2016).

The inconsistent combination of taxonomic relevant features supports the hypothesis of Okinaka and Keim (2016) that the Bcsl species group is complex and that exclusive phenotypic and biochemical characterization may be misleading. The authors suggest that the whole Bcsl complex should be considered as a single species. Investigation of only the phenotypic characteristics of the strains would lead to a mis-classification of Bt isolates, as no known Cry toxins were detected. Moreover, the invasive growing Bw strains would be mis-classified to the species Bm or Bp. However, there are examples for the challenging taxonomic classification of Bcsl members. Bt serovar, *navarrensis*, *bolivia*, and *vazensis* have been classified as Bt based on their ability to produce Cry toxins. They showed the typical psychrotolerant growth and encode the *cspA* gene signature, which is species specific for Bw (Soufiane and Côté, 2009). We classified our isolates as 7 Bt and 10 Bw based on our MLST study combined with phenotypic features (**Figure 2**). The three special Bw isolates (Bw GOE8-Bw GOE10) combined several morphological as well as genomic characteristics of the species Bm, Ba, and Bw indicating that they might represent a distinct subspecies within the species Bw (**Figure 2**). In conclusion, our findings indicate that a combination of MLST analysis and phenotypic, biochemical, and pathogenic classification is the only possibility to distinguish between members of the Bcsl complex.

In vitro Antagonistic Properties of Root-Associated *Bacilli* toward *Verticillium*

We further investigated the antagonistic potential of new isolated *Bacilli* strains against two phytopathogenic *Verticillium* species with different host ranges. In total, 20 *Bacillus* isolates and 6 Bt reference strains were tested for their ability to suppress the haploid tomato pathogen *V. dahliae* JR2 (Fradin et al., 2009) or the diploid rapeseed pathogen *V. longisporum* 43 (Zeise and von Tiedemann, 2002; Tran et al., 2013) on three different media (**Table 2**). Not all Bt reference strains were able to suppress *V. dahliae* JR2 and showed a broad range of different interaction effects against *V. longisporum* 43 indicating that antifungal activity is a property of a particular Bt strain and not of the whole species (Supplementary Table 4, Supplementary Figures 2, 3). The co-cultivation assays of *Bacillus* isolates revealed specific phenotypic effects for the antagonistic bacteria as well as for the phytopathogens (**Figure 4**). The different bacterial isolates varied in their ability to inhibit the mycelia growth of either *V. dahliae* JR2 or *V. longisporum* 43. A significant *in vitro* antagonism was observed for all Bt isolates against *V. dahliae* JR2 on PDM, excluding Bt GOE6 compared to the tested Bt control strains (Bt Bt18247, Bt MYBT18246, 4Q1) which exhibited no inhibitory effect (**Figure 5A**, Supplementary Table 5). A clear inhibition zone without *Verticillium* mycelium, sometimes with a slight formation of microsclerotia, was detected (Supplementary Figure 2). Additionally, *V. dahliae* JR2 showed an altered phenotype compared to the fungal control by the building of a strong white air mycelium around the bacteria (**Figure 4C**). The antagonistic effect of Bt varied but was consistent in all biological replicates and had a mean ranging from 1.5 to 8.6 mm (**Figure 5**). In contrast to all other tested Bt isolates, which exhibited an antagonistic effect on PDM, isolate Bt GOE6 showed a significant antagonism against *V. dahliae* JR2 only on LB (**Figure 4A**, **Figure 5**, Supplementary Table 5). Notably, none of the strains belonging to Bw showed an inhibitory effect on *V. dahliae* JR2 or *V. longisporum* 43 (Supplementary Table 4). In comparison, a weaker antagonistic effect of bacteria on *V. longisporum* 43 was recorded. All strains which had an effect on *V. dahliae* JR2 lead to an altered phenotype of *V. longisporum* 43 on PDM. The bacteria induced a stronger melanization in *V. longisporum* 43, which

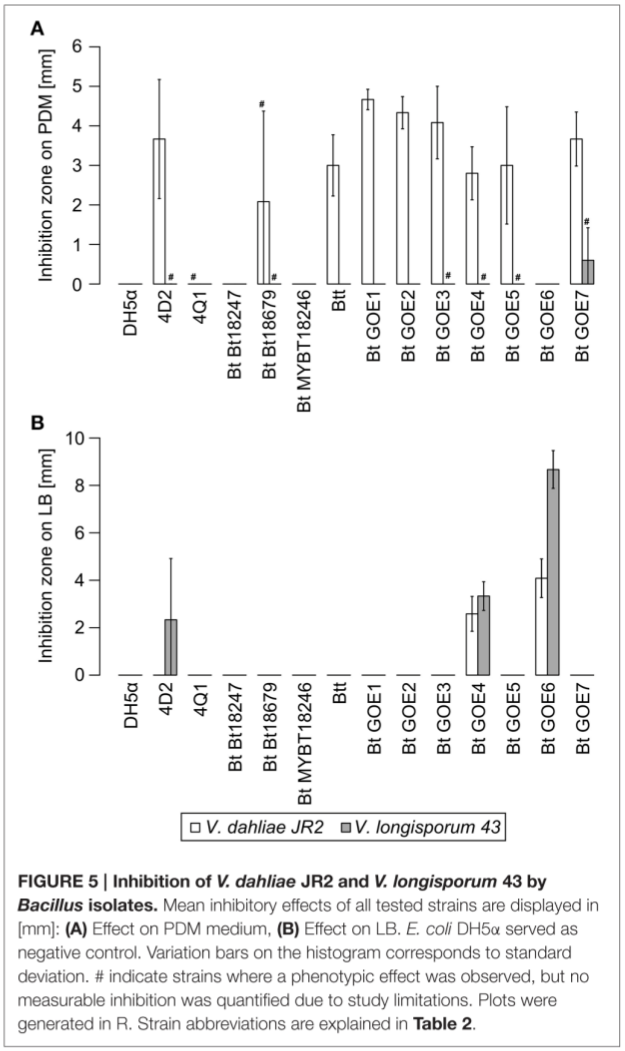

FIGURE 5 | Inhibition of *V. dahliae* JR2 and *V. longisporum* 43 by *Bacillus* isolates. Mean inhibitory effects of all tested strains are displayed in [mm]: **(A)** Effect on PDM medium, **(B)** Effect on LB. *E. coli* DH5α served as negative control. Variation bars on the histogram corresponds to standard deviation. # indicate strains where a phenotypic effect was observed, but no measurable inhibition was quantified due to study limitations. Plots were generated in R. Strain abbreviations are explained in **Table 2**.

could be observed on the back of the plate as a black ring around the bacteria (**Figure 4D**). Only for strain Bt GOE7 an inhibitory effect with high variance in replicates was observed compared to all other tested strains (**Figure 5**, Supplementary Table 5). On LB, Bt GOE4 and Bt GOE6 exhibited a significant inhibitory effect against *V. longisporum* 43 compared to all other tested strains (**Figure 5**, Supplementary Table 5). Additionally, 4D2 showed an inhibitory effect aginst *V. longisporum* 43 as well but with an increased variance in biological treatments. Moreover, Bt GOE6 had a significant higher inhibitory effect against *V. longisporum* 43 compared to 4D2 and Bt GOE4. Significant inhibitory effects against both phytopathogens, *V. dahliae* JR2 and *V. longisporum* 43, were only observed for Bt GOE4 and Bt GOE6 on LB (**Figure 5**, Supplementary Table 5). However, two additional phenotypic effects of bacteria were observed: (a) biofilm production of bacteria with varying inhibition effects on fungi and (b) building of hyphae-like growth structures of bacteria leading to a suppression or growth reduction of both phytopathogens. Effect (a) was detected for all Bt isolates, which build moderate biofilms. Bw isolates conferred only a weak biofilm production with no suppression effect, excluding Bw GOE8-GOE10. Effect (b) was observed for isolates Bw GOE8-GOE10 on LB and SXM (**Figure 4B**, Supplementary Table 4). The hyphae-like structures or rhizoid growth enables the bacterium to invade the same habitat as *Verticillium*. This might strengthen the competitive access of *Bacilli* for nutrients, leading to suppression or reduction of fungal growth to different extents. As the isolated Bt strains suppressed *V. dahliae* JR2 growth, this indicates that tomato-associated Bt strains are able to inhibit phytopathogens. In contrast, the effect on the foreign plant pathogen *V. longisporum* 43 was in general weaker or even not detectable. We assume that different host-range specificities play an important role in *Verticillium* species, resulting in two different life strategies. *Verticillium dahliae* is haploid

TABLE 5 | Identified secondary metabolite gene clusters of bacterial strains.

Identified secondary metabolite gene cluster	NRPS		RiPPs*						Terpene	Others	Total
	Siderophore (Petrobactin, Bacillibactin)	Not further categorized	Bacteriocins	Microcins	Lanthipeptide	Lassopeptide	Linaridin	Ladderane			
BACTERIAL STRAINS											
Bt GOE1	2	3	3	1	1	–	–	–	1	1	12
Bt GOE2	2	4	3	1	1	–	–	–	–	1	12
Bw GOE1	2	1	2	1	–	1	–	–	1	–	8
Bt GOE3	2	2	3	1	–	–	–	–	1	1	10
Bw GOE2	2	1	2	1	–	1	–	–	1	–	8
Bt GOE4	2	3	4	1	4	–	–	–	1	2	17
Bt GOE5	2	3	5	1	–	1	–	–	1	1	14
Bt GOE6	2	3	3	1	–	–	–	–	1	1	11
Bw GOE3	2	1	2	1	–	1	–	–	1	–	8
Bw GOE4	2	1	2	1	–	1	–	1	1	–	9
Bw GOE5	2	1	4	1	–	1	–	–	1	–	10
Bw GOE6	2	1	2	1	–	1	–	–	1	–	8
Bw GOE7	2	1	2	1	–	1	–	1	1	–	9
Bw GOE8	2	1	3	1	1	2	–	–	1	1	12
Bw GOE9	2	1	3	1	1	1	–	–	1	1	11
Bw GOE10	2	1	3	1	1	2	–	–	1	1	12
Bw GOE11	2	1	3	1	–	1	–	1	1	–	10
Bw GOE12	2	1	2	1	–	1	–	1	1	–	9
Bw GOE13	2	1	3	1	–	1	–	1	1	–	10
Bt GOE7	2	2	4	1	–	1	1	–	1	2	14
Bt MYBT18246	2	2	6	8	1	–	–	–	1	1	21
Bt Bt18247	2	2	5	–	–	–	–	–	1	1	11
Bt Bt18679	2	3	4	–	–	–	–	–	1	–	10
Btt	2	4	3	–	–	–	–	–	1	1	11
Bt serovar *kurstaki* HD-1	2	5	3	8	1	–	–	–	1	1	21
Bw KBAB4	2	1	3	8	–	–	–	–	1	1	16

**(RiPPs): Ribosomally synthesized and post-translationally modified peptides.*

and comprises strains with a broad host range (Inderbitzin et al., 2011) whereas *V. longisporum* is amphidiploid (hybrid of *V. dahliae* and an unknown haploid *Verticillium* species) with narrow host range infecting mainly *Brassicacae* (Depotter et al., 2016). This could lead to long-term adaptation of *V. longisporum* resulting in an increased fitness, as *V. longisporum* has to persist in soil waiting for its specific host plant. *Verticillium dahliae* is able to colonize a broad range of host plants and thus has to cope with different environmental conditions. We suggest that this results in a reduced adaptation potential for the different types of host-residing pathogens. An opportunity to enter the next potential host are spores or the formation of resting structures, such as microsclerotia, which are formed under stress conditions and can easily spread by wind or survive in the ground.

On SXM, only minor and inconsistent effects were observed around the bacterial application site and not further quantified. Instead of an inhibition, a reduced formation of fungal mycelium was observed (Supplementary Figures 2, 3). Nonetheless, the ability of a bacterial strain to inhibit fungal growth differed depending on the medium. Using different phytopathogenic interaction partner had also an impact on phenotypic effects of bacteria (Supplementary Table 4). Generally, the inhibitory effect on *V. dahliae* JR2 was always more prominent than the effect on *V. longisporum* 43 (**Figure 4**, Supplementary Figures 2, 3). In particular, on PDM, the fungus is able to sense the bacteria or antifungal substances. This leads to a strong physiological reaction of *V. dahliae* JR2 resulting in an increased production of air-mycelium in distance to the bacterial application site. Only in individual cases, a reduced number of microsclerotia was detected (Supplementary Figure 2; Bt GOE5, Bt GOE7). For *V. longisporum* 43, an increased production of microsclerotia was observed indicating an induced stress response by bacteria. However, SXM is also a rich medium containing the complex heteropolysaccharide pectin, which is more difficult to use for bacteria. We suggest that this reduce or enhance the growth efficiency of bacteria or fungi, respectively. In contrast, LB medium is favored by bacteria, which could explain the biofilm production, whereas both fungi built minor mycelium (**Figure 4**).

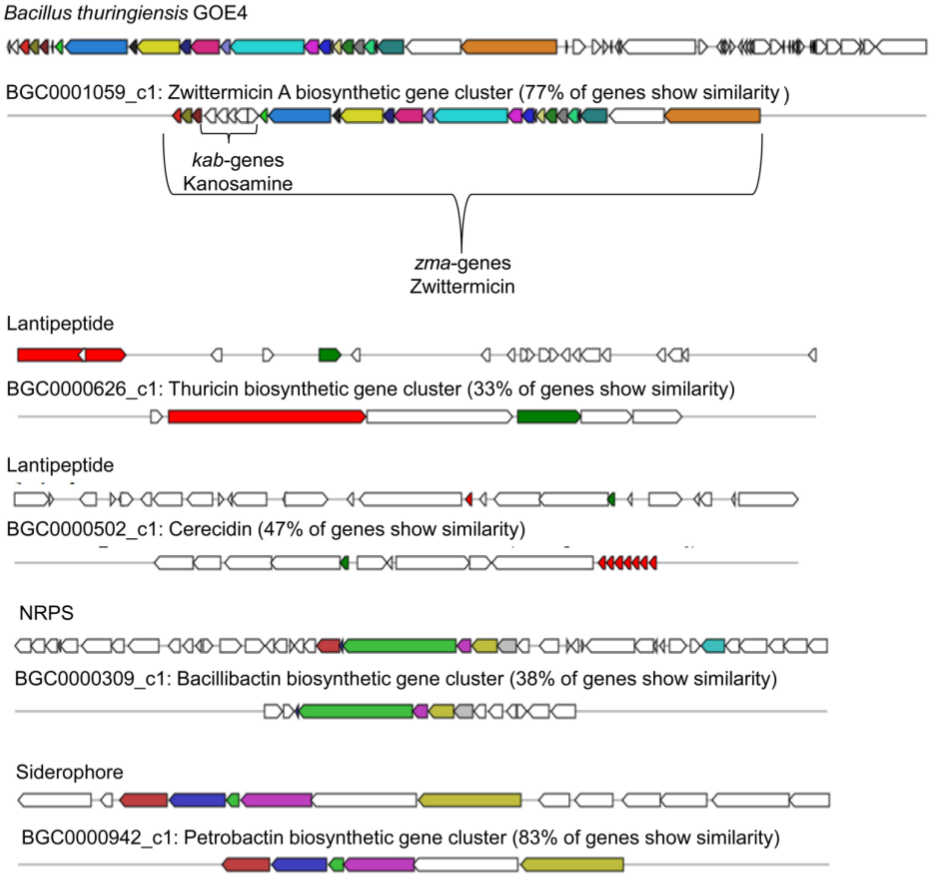

FIGURE 6 | Biosynthetic gene clusters of Bt GOE4. The analysis was performed with AntiSMASH3.0 (Weber et al., 2015). Same colored genes display similarity to published and known gene clusters. A minimum of 30% cutoff was set for visualization of gene clusters.

Secondary Metabolites of Root-Associated *Bacilli*

The secondary metabolite gene prediction tool (AntiSMASH 3.0; Weber et al., 2015) was used to perform a comparative genome analysis and identify candidates contributing to the observed antifungal effects. In total, 304 gene clusters including 101 NRPS gene clusters and 203 RiPP gene clusters were identified (**Table 5**). In each genome, secondary metabolite gene clusters were detected, including genes assigned to the production of siderophores and terpenes. The gene clusters assigned to siderophore synthesis showed similarity (80–83%, BGC0000942) to a petrobactin biosynthetic gene cluster and a bacillibactin cluster (38–46%, BGC0000309.1). Both are catecholate siderophores, which are produced under iron-limited conditions, and present in Ba and other Bcsl members. Li B. et al. (2014) reported that the transcription of bacillibactin genes was upregulated when strains were confronted with fungal pathogens. Bacillibactin-deficient mutants exhibited no or reduced antagonistic effects. Although the suppression could not be observed in each treatment, the findings indicate that bacillibactin plays a supporting role in the suppression of the two different *Verticillium* species (Li B. et al., 2014).

It was suggested that petrobactin could contribute to virulence of Ba (Lee et al., 2011), but antifungal activity was not discussed. A number of plants including tomato are known to produce terpenes for deterring or attracting herbivores, parasites, and predators (Lange et al., 2000; Martin et al., 2003). Terpene assigned gene clusters were identified in all strains and shared only low similarity (11–17%, BGC0000916.1) with molybdenum cofactor biosynthesis genes. Bt GOE4 showed the highest number of detected secondary metabolites and antimicrobial peptides with similarity to known gene clusters with antifungal activity (**Figure 6**). One gene cluster putatively producing the antifungal zwittermicin A (BGC0001059) was detected. It belongs to the class of antibiotics of type I PKS gene clusters, which is known to suppress plant diseases (Handelsman et al., 1990). Zwittermicin A is a linear aminopolyol, which was first identified in *B. cereus* UW85 (Silo-Suh et al., 1994). The cluster was only identified in Bt GOE4 and in the reference genome Bt serovar *kurstaki* HD-1. In Bt GOE4, the cluster differs by lacking the five *kab* (*kabR; kabA-kabD*) genes, which are important for the kanosamine production. Kanosamine is not important for the synthesis of zwittermicin A, but it is also known to exhibit fungicidal activities (Janiak and Milewski, 2001). In these strains, a thuricin-like cluster (66% similarity, BGC0000626.1) and a thuricin H biosynthetic gene cluster (90% similarity, BGC0000600.1) were detected. Thuricin H is described as member of a small subclass of bacteriocins, which act on related bacterial strains and not on fungi (Mathur et al., 2015). A cerecidin cluster was detected (47%, BGC0000502) in Bt GOE4. Cerecidin is a novel antibiotic with high activity against a broad range of Gram-positive bacteria (Wang et al., 2014). In addition, a putative paenilamicin (BGC0001033.1) producing type I PKS gene cluster was exclusively identified in Bt18679 and Bw KBAB4 with shared similarity of 35%. Paenilamicins are known to act antibacterial and antifungal (Müller et al., 2014). As this gene cluster was not detected in all suppressing isolates, it is not exclusively the paenilamicins which lead to the inhibitory effect in our experiments.

Gene clusters encoding ribosomally synthesized and post-translationally modified peptides (RiPPs) such as bacteriocins, microcins, lanthipeptides, lassopeptides, ladderane, and linaride were identified as well, but could not be further characterized due to weak similarity to known clusters (**Table 5**). Only in case of the microcin bacitracin (BGC0000310.1), genes with similarities of 44% have been identified in the strains Bt subsp. *kurstaki* HD-1 and Bw KBAB4. Bacitracin is not known for

TABLE 6 | Identified chitinases in bacterial strains.

Bacterial strain	Chtitinase	Length in bp	Length in aa
Bt GOE1	ChiA	1083	360
Bt GOE1	ChiB	2025	674
Bt GOE2	ChiA	1083	360
Bt GOE2	ChiB	2025	674
Bt GOE3	ChiA	1083	360
Bt GOE3	ChiB	2025	674
Bt GOE4	ChiA	1083	360
Bt GOE4	ChiB	2025	674
Bt GOE5	ChiA	1083	360
Bt GOE5	ChiB	2025	674
Bt GOE6	ChiA	1083	360
Bt GOE6	ChiB	2025	674
Bw GOE8	ChiC-like	945	314
Bw GOE9	ChiC-like	945	314
Bw GOE10	ChiC-like	945	314
Bt GOE7	ChiA	1083	360
Bt GOE7	ChiB	2025	674
Bt MYBT18246	ChiA	1083	360
Bt MYBT18246	ChiB	2025	674
Bt Bt18247	ChiA	1083	360
Bt Bt18247	ChiB	2025	674
Bt Bt18679	ChiA	1083	360
Bt Bt18679	ChiB	2025	674
Btt	ChiA	1083	360
Btt	ChiB	2025	674
Bw KBAB4 ChiA	ChiA	1083	360
Bw KBAB4 ChiB	ChiB	2025	674
Pseudomonas aeruginosa	ChiC	1463	483
Bt serovar *colmeri* strain 15A3	ChiA	1083	360
Bt serovar *colmeri* strain 15A3 Chitinase	ChiB	2076	688
Bt serovar *kurstaki* Chi255	Chi255	2710	376
Bc strain 28-9 Chitinase CW	ChiCW	2450	674
Btt DSM-2803	ChiA	2331	676
Bt *kenyae* LBIT-82	ChiA	2331	676
Bt strain SBS-Bt5	ChiB	2331	676
Bc AH621	ChiC	912	303
Bc AH621	ChiB	2067	688
Bc AH621	ChiA	1083	360

a fungicidal effect, but rather for the effect on spore size and crystal production (García-Patrone, 1985). Gene clusters with the potential to encode antifungal compounds, such as zwittermicin A, are present in some strains. Remarkably, the genes are not shared by all isolates acting as fungal antagonists. In contrast, bacillibactin producing gene clusters were identified in each isolate independently of the ability to suppress the growth of *V. dahliae* JR2 or *V. longisporum* 43. We think that NRPS/PKS synthesized secondary metabolites or RiPPS could contribute to the observed growth inhibition. We strongly suggest that different complex mechanisms and substances act in combination and are important for an efficient suppression of the fungi investigated.

Chitinases of Root-Associated *Bacilli*

To elucidate the second class of fungicidal substances, we focused on chitinases as the mycolytic activity of many *Bacilli* is known (Swiontek Brzezinska et al., 2014). Genome analysis revealed the presence of three different classes of chitinases shared by the genomes of the strains with antifungal activities. Notably, genes encoding a chitinolytic polypeptide with 674 (ChiA) and 360 (ChiB) amino acids were identified in all inhibiting isolates (**Table 6**, **Figure 7**) but not in Bw isolates, with exception of the invasive-growing Bw strains (Bw GOE8-10). These isolates contained a gene encoding a chitinase-like enzyme of 314 amino acids. Comparison of the amino acid sequences of all detected chitinases with reference chitinases with known antifungal activity lead to a classification of three chitinase groups: Group I: ChiB; Group II: ChiA and Group III: ChiC (**Figure 7B**). Identified ChiA chitinases showed sequence similarity to known mycolytic chitinases of *Bt* subsp. *tenebrionis* DSM-2803 (de la Fuente-Salcido et al., 2016), *B. cereus* strain 28-9 (Huang et al., 2005) and Bt subsp. *kurstaki* (Driss et al., 2005; **Figure 7A**). They grouped into glycoside hydrolase family 18 with a catalytic domain and an active site. ChiB chitinases clustered to reference chitinases, including ChiB of Bc AH621 and the mycolytic chitinase chiA of Bt serovar *colmeri* (Liu et al., 2010). They also grouped into glycoside hydrolase family 18, but in addition to the catalytic domain and the active site, the enzymes contained a chitin insertion domain and a carbohydrate-binding type2 domain, which are separated by a fibronectin type III domain.

Our experiments demonstrate that the classification of ChiA from Bt serovar *colmeri* should be reconsidered, as the deduced amino acid size of the protein and domain analysis revealed that this chitinase belongs to the ChiB group of chitinases (Group I). The chitinases of the invasive Bw strains were also classified as glycoside hydrolase family 18, containing a catalytic domain. They were most similar to a chitinase of *B. cereus* AH621 and *Pseudomonas aeruginosa* (**Figure 7**). In contrast, the Bw chitinases harbored no chitinase insertion domain, fibronectin type III domain, carbohydrate-binding type-2 domain or a chitinase active site. We suggest that we identified potential

FIGURE 7 | Analysis and comparison of chitinases encoded by the genomes of the isolates. (A) Phylogenetic tree of all identified chitinases (bold black) compared to chitinases of reference strains, with known antifungal activity (blue) target organism is depicted in brackets. The phylogenetic analysis was performed using in MEGA7 (Kumar et al., 2016), details are described in the Materials and Methods Section. **(B)** Domain structures of three identified chitinase types. InterPro (Mitchell et al., 2015) was used for functional analysis of chitinases, classification into families and for the prediction of domains and active sites. Identified structures are colored: (blue) glycoside hydrolase family 18 with catalytic domain, (red) chitinase insertion domain, fibronectin type III domain, (yellow) carbohydrate-binding type-2 domain, (black) chitinases active site.

mycolytic chitinases in our Bt isolates whereas we can only hypothesize which of the identified groups of chitinases (chiA or chiB) are able to suppress *Verticillium*. Strain Bt GOE6 exhibited an antifungal effect on LB only but not on PDM. In addition, the genomes of the Bt control strains MYBT18246 and Bt18247 encode two chitinases but showed no effect on the fungal pathogens at all. These results suggest that the growth period and abiotic factors such as pH, temperature, and presence of metal ions, carbon, or nitrogen-sources influence the expression and/or the activity of chitinases. The identified chitinase in the invasive Bw strains exhibited less similarities to known sequences and thus could be a new type of chitinase. To our knowledge, this is the first time that several Bt isolates comprising multiple chitinases, showed the potential to inhibit *V. dahliae in vitro*. As chitinases are an effective biocontrol substance against a number of phytopathogenic fungi (Bhattacharya et al., 2007), future *in planta* experiments are needed to elucidate the full potential of chitinases of Bt isolates.

CONCLUSION

In the present study, we evaluated the antifungal potential of 20 phenotypically diverse *Bacillus* isolates toward *V. dahliae* and *V. longisporum*. A classification of new isolates based on a combination of morphological, pathogenic, and taxonomic properties revealed 7 *B. thuringiensis* and 13 *B. weihenstephanensis* strains. The dual cultivation assays showed a correlation between taxonomy and antagonistic activities. All *B. thuringiensis* strains exhibited an *in vitro* antifungal effect against *V. dahliae* while only limited antagonism was observed against *V. longisporum*. Additionally, three *B. weihenstephanensis* isolates showing an invasive growth-type competed with both phytopathogenic fungi. The relation of the rhizoid growth and the mechanism of competition of *B. weihenstephanensis* strains have not been described previously and thus represent a fascinating new research topic.

The genome analysis of the 20 *Bacillus* strains revealed that strains with antifungal activity shared genes assigned to bacillibactin production and mycolytic chitinases, which are thus the most promising candidates for encoding the antifungal effect.

The hereby produced genomic and physiological data provide an excellent foundation for the identification, purification, and characterization of the active antifungal substances of Bt. Nonetheless, future *in planta* experiments are necessary to determine the efficacy of these strains in controlling plant pathogens.

AUTHOR CONTRIBUTIONS

The study was designed and conceived by HL, RH, RD, and JH. Soil bacterial communities were analyzed by JH and FW. Bacterial isolation, PCR, and bacterial identification were performed by AK, JH, EB, and HL. Genome sequencing, data processing, and genome assembly were done by JH and AP. Genome annotation, data submission, and data interpretation were performed by JH, SDV, and HL. Antagonistic assays were undertaken by AK, RH, and JH. Data Interpretation and manuscript preparation were done by JH, FW, RH, and HL. The whole project was supervised by RH, SB-S, GB, KN, RD, and HL. All authors interpreted the results, read and approved the final manuscript.

FUNDING

We thank the German Research Foundation (DFG-SPP1399, Grants LI 1690/2-1 and DFG BR1502/15-1) for financial support.

ACKNOWLEDGMENTS

We would like to thank Jan Teer for performing MLST PCR experiments. We would like to thank Nicole Scheiter for technical assistance. We acknowledge support by the German Research Foundation and the Open Access Publication Funds of the Göttingen University.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: http://journal.frontiersin.org/article/10.3389/fmicb.2016.02171/full#supplementary-material

REFERENCES

Ahemad, M., and Kibret, M. (2014). Mechanisms and applications of plant growth promoting rhizobacteria: current perspective. *J. King Saud Univ. Sci.* 26, 1–20. doi: 10.1016/j.jksus.2013.05.001

Aleti, G., Sessitsch, A., and Brader, G. (2015). Genome mining: prediction of lipopeptides and polyketides from *Bacillus* and related *Firmicutes*. *Comput. Struct. Biotechnol. J.* 13, 192–203. doi: 10.1016/j.csbj.2015.03.003

Angelopoulou, D. J., Naska, E. J., Paplomatas, E. J., and Tjamos, S. E. (2014). Biological control agents (BCAs) of verticillium wilt: influence of application rates and delivery method on plant protection, triggering of host defence mechanisms and rhizosphere populations of BCAs. *Plant Pathol.* 63, 1062–1069. doi: 10.1111/ppa.12198

Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comput. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021

Bateman, A., Martin, M. J., O'Donovan, C., Magrane, M., Apweiler, R., Alpi, E., et al. (2015). UniProt: a hub for protein information. *Nucleic Acids Res.* 43, D204–D212. doi: 10.1093/nar/gku989

Bhattacharya, D., Nagpure, A., and Gupta, R. K. (2007). Bacterial Chitinases: properties and potential. *Crit. Rev. Biotechnol.* 27, 21–28. doi: 10.1080/07388550601168223

Bhattacharyya, P. N., and Jha, D. K. (2012). Plant growth-promoting rhizobacteria (PGpr): emergence in agriculture. *World J. Microbiol. Biotechnol.* 28, 1327–1350. doi: 10.1007/s11274-011-0979-9

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Bone, E. J., and Ellar, D. J. (1989). Transformation of *Bacillus thuringiensis* by electroporation. *FEMS Microbiol. Lett.* 49, 171–177. doi: 10.1111/j.1574-6968.1989.tb03039.x

Bragg, L., Stone, G., Imelfort, M., Hugenholtz, P., and Tyson, G. W. (2012). Fast, accurate error-correction of amplicon pyrosequences using Acacia. *Nat. Methods* 9, 425–426. doi: 10.1038/nmeth.1990

Bulgarelli, D., Garrido-Oter, R., Münch, P. C., Weiman, A., Dröge, J., Pan, Y., et al. (2015). Structure and function of the bacterial root microbiota in wild and domesticated barley. *Cell Host Microbe* 17, 392–403. doi: 10.1016/j.chom.2015.01.011

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinformatics* 10:421. doi: 10.1186/1471-2105-10-421

Caporaso, J. G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Crawley, M. J. (2007). *The R Book*. West Sussex: John Wiley and Sons.

Darling, A. C., Mau, B., Blattner, F. R., and Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. *Genome Res.* 14, 1394–1403. doi: 10.1101/gr.2289704

de la Fuente-Salcido, N. M., Casados-Vázquez, L. E., García-Pérez, A. P., Barboza-Pérez, U. E., Bideshi, D. K., Salcedo-Hernández, R., et al. (2016). The endochitinase ChiA Btt of *Bacillus tenebrionis* subsp. tenebrionis DSM-2803 and its potential use to control the phytopathogen *Colletotrichum gloeosporioides*. *Microbiologyopen* 5, 819–829. doi: 10.1002/mbo3.372

de Maagd, R. A., Bravo, A., Berry, C., Crickmore, N., and Schnepf, H. E. (2003). Structure, diversity, and evolution of protein toxins from spore-forming Entomopathogenic Bacteria. *Annu. Rev. Genet* 37, 409–433. doi: 10.1146/annurev.genet.37.110801.143042

Depotter, J. R., Deketelaere, S., Inderbitzin, P., Tiedemann, A. V., Höfte, M., Subbarao, K. V., et al. (2016). *Verticillium longisporum*, the invisible threat to oilseed rape and other brassicaceous plant hosts. *Mol. Plant Pathol.* 17, 1004–1016. doi: 10.1111/mpp.12350

de Serres, J. F., and Hollaender, A. (1982). *Chemical Mutagens. Principles and Methods for Their Detection*. New York, NY: Plenum Press.

De Vos, P., Garrity, G. M., Jones, D., Krieg, N. R., Ludwig, W., Rainey, F. A., et al. (2009). *Bergey's Manual of Systematic Bacteriology Volume 3: The Firmicutes*. New York, NY: Springer.

Di Franco, C., Beccari, E., Santini, T., Pisaneschi, G., and Tecce, G. (2002). Colony shape as a genetic trait in the pattern-forming *Bacillus mycoides*. *BMC Microbiol.* 2:33. doi: 10.1186/1471-2180-2-33

Dixon, G. R., and Pegg, G. F. (1971). Changes in amino-acid content of tomato xylem sap following infection with strains of *Verticillium albo-atrum*. *Ann. Bot.* 36, 147–154.

Driss, F., Kallassy-Awad, M., Zouari, N., and Jaoua, S. (2005). Molecular characterization of a novel chitinase from *Bacillus thuringiensis* subsp. *kurstaki*. *J. Appl. Microbiol.* 99, 945–953. doi: 10.1111/j.1365-2672.2005.02639.x

Eddy, S. (1998). Profile hidden Markov models. *Bioinformatics* 14, 755–763. doi: 10.1093/bioinformatics/14.9.755

Eddy, S. R. (2011). Accelerated profile HMM searches. *PLoS Comput. Biol.* 7:e1002195. doi: 10.1371/journal.pcbi.1002195

Edgar, R. C. (2010). Search and clustering orders of magnitude faster than BLAST. *Bioinformatics* 26, 2460–2461. doi: 10.1093/bioinformatics/btq461

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinformatics* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Enright, A. J., Van Dongen, S., and Ouzounis, C. A. (2002). An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res.* 30, 1575–1584. doi: 10.1093/nar/30.7.1575

Enya, J., Koitabashi, M., Shinohara, H., Yoshida, S., Tsukiboshi, T., Negishi, H., et al. (2007). Phylogenetic diversities of dominant culturable *Bacillus*, *Pseudomonas* and Pantoea species on tomato leaves and their possibility as biological control agents. *J. Phytopathol.* 155, 446–453. doi: 10.1111/j.1439-0434.2007.01256.x

Fradin, E. F., and Thomma, B. P. H. J. (2006). Physiology and molecular aspects of *Verticillium* wilt diseases caused by *V. dahliae* and *V. albo-atrum*. *Mol. Plant Pathol.* 7, 71–86. doi: 10.1111/j.1364-3703.2006.00323.x

Fradin, E. F., Zhang, Z., Juarez Ayala, J. C., Castroverde, C. D. M., Nazar, R. N., Robb, J., et al. (2009). Genetic dissection of *Verticillium* wilt resistance mediated by Tomato Ve1. *Plant Physiol.* 150, 320–332. doi: 10.1104/pp.109.136762

Frank, M. N. (2003). Development of alternative strategies for management of soilborne pathogens currently controlled with Methylbromide. *Annu. Rev. Biochem.* 72, 291–336. doi: 10.1146/annurev.phyto.41.052002.095514

Fravel, D. R. (2005). Commercialization and implementation of biocontrol. *Annu. Rev. Phytopathol.* 43, 337–359. doi: 10.1146/annurev.phyto.43.032904.092924

Gams, W., Hoekstra, E. S., and Aptroot, A. (1998). *CBS Course of Mycology*. Delft: Centraalbureau voor Schimmelcultures Baarn, Delft.

García-Patrone, M. (1985). Bacitracin increases size of parasporal crystals and spores in *Bacillus thuringiensis*. *Mol. Cell. Biochem.* 68, 131–137. doi: 10.1007/BF00219377

Glick, B. R. (2012). Plant growth-promoting bacteria: mechanisms and applications. *Scientifica* 2012, 1–15. doi: 10.6064/2012/963401

González, J. M. Jr., Brown, B. J., and Carlton, B. C. (1982). Transfer of *Bacillus thuringiensis* plasmids coding for delta-endotoxin among strains of *B. thuringiensis* and *B. cereus*. *Proc. Natl. Acad. Sci. U.S.A.* 79, 6951–6955. doi: 10.1073/pnas.79.22.6951

Haas, D., and Défago, G. (2005). Biological control of soil-borne pathogens by fluorescent pseudomonads. *Nat. Rev. Microbiol.* 3, 307–319. doi: 10.1038/nrmicro1129

Han, C. S., Xie, G., Challacombe, J. F., Altherr, M. R., Bhotika, S. S., Brown, N., et al. (2006). Pathogenomic sequence analysis of *Bacillus cereus* and *Bacillus thuringiensis* isolates closely related to *Bacillus anthracis*. *J. Bacteriol.* 188, 3382–3390. doi: 10.1128/JB.188.9.3382-3390.2006

Handelsman, J., Raffel, S., Mester, E. H., Wunderlich, L., and Grau, C. R. (1990). Biological control of damping-off of alfalfa seedlings with *Bacillus cereus* UW85. *Appl. Environ. Microbiol.* 56, 713–718.

Helgason, E., Okstad, O. A., Caugant, D. A., Johansen, H. A., Fouet, A., Mock, M., et al. (2000). *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*–one species on the basis of genetic evidence. *Appl. Environ. Microbiol.* 66, 2627–2630. doi: 10.1128/AEM.66.6.2627-2630.2000

Huang, C. J., Wang, T. K., Chung, S. C., and Chen, C. Y. (2005). Identification of an antifungal chitinase from a potential biocontrol agent, *Bacillus cereus* 28-9. *J. Biochem. Mol. Biol.* 38, 82–88. doi: 10.5483/bmbrep.2005.38.1.082

Hyatt, D., Chen, G. L., Locascio, P. F., Land, M. L., Larimer, F. W., and Hauser, L. J. (2010). Prodigal: prokaryotic gene recognition and translation initiation site identification. *BMC Bioinformatics* 11:119. doi: 10.1186/1471-2105-11-119

Inderbitzin, P., Bostock, R. M., Davis, R. M., Usami, T., Platt, H. W., and Subbarao, K. V. (2011). Phylogenetics and taxonomy of the fungal vascular wilt pathogen *Verticillium*, with the descriptions of five new species. *PLoS ONE* 6:e28341. doi: 10.1371/journal.pone.0028341

Janiak, A. M., and Milewski, S. (2001). Mechanism of antifungal action of kanosamine. *Med. Mycol.* 39, 401–408. doi: 10.1080/mmy.39.5.401.408

Jolley, K. A., Chan, M. S., and Maiden, M. C. (2004). mlstdbNet - distributed multi-locus sequence typing (MLST) databases. *BMC Bioinformatics* 5:86. doi: 10.1186/1471-2105-5-86

Jones, J. D., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286

Katan, J. (1981). Solar heating (Solarization) control of soilborne pests. *Ann. Rev. Phytopathol.* 19, 211–236. doi: 10.1146/annurev.py.19.090181.001235

Klindworth, A., Pruesse, E., Schweer, T., Peplies, J., Quast, C., Horn, M., et al. (2013). Evaluation of general 16S ribosomal RNA gene PCR primers for classical and next-generation sequencing-based diversity studies. *Nucleic Acids Res.* 41, 1–11. doi: 10.1093/nar/gks808

Kolstø, A. B., Tourasse, N. J., and Økstad, O. A. (2009). What sets *Bacillus anthracis* apart from other Bacillus species? *Annu. Rev. Microbiol.* 63, 451–476. doi: 10.1146/annurev.micro.091208.073255

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 7, 1870–1874. doi: 10.1093/molbev/msw054

Lagesen, K., Hallin, P., Rødland, E. A., Stærfeldt, H. H., Rognes, T., and Ussery, D. W. (2007). RNAmmer: consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* 35, 3100–3108. doi: 10.1093/nar/gkm160

Lange, B. M., Rujan, T., Martin, W., and Croteau, R. (2000). Isoprenoid biosynthesis: the evolution of two ancient and distinct pathways

across genomes. *Proc. Natl. Acad. Sci. U.S.A.* 97, 13172–13177. doi: 10.1073/pnas.240454797

Laslett, D., and Canback, B. (2004). ARAGORN, a program to detect tRNA genes and tmRNA genes in nucleotide sequences. *Nucleic Acids Res.* 32, 11–16. doi: 10.1093/nar/gkh152

Lechner, S., Mayr, R., Francis, K. P., Prüss, B. M., Kaplan, T., Wiessner-Gunkel, E., et al. (1998). *Bacillus weihenstephanensis* sp. nov. is a new psychrotolerant species of the *Bacillus cereus* group. *Int. J. Syst. Bacteriol.* 48(Pt 4), 1373–1382.

Lee, J. Y., Passalacqua, K. D., Hanna, P. C., and Sherman, D. H. (2011). Regulation of Petrobactin and Bacillibactin biosynthesis in *Bacillus anthracis* under iron and oxygen variation. *PLoS ONE* 6:e20777. doi: 10.1371/journal.pone.0020777

Li, B., Li, Q., Xu, Z., Zhang, N., Shen, Q., Zhang, R., et al. (2014). Responses of beneficial *Bacillus amyloliquefaciens* SQR9 to different soilborne fungal pathogens through the alteration of antifungal compounds production. *Front. Microbiol.* 5:636. doi: 10.3389/fmicb.2014.00636

Li, J. G., Di Ren, G., Jia, Z. J., and Dong, Y. H. (2014). Composition and activity of rhizosphere microbial communities associated with healthy and diseased greenhouse tomatoes. *Plant Soil* 380, 337–347. doi: 10.1007/s11104-014-2097-6

Liu, D., Cai, J., Xie, C. C., Liu, C., and Chen, Y.-H. (2010). Purification and partial characterization of a 36-kDa chitinase from *Bacillus thuringiensis* subsp. *colmeri,* and its biocontrol potential. *Enzyme Microb. Technol.* 46, 252–256. doi: 10.1016/j.enzmictec.2009.10.007

Luo, X., Xie, C., Dong, J., Yang, X., and Sui, A. (2014). Interactions between *Verticillium dahliae* and its host: vegetative growth, pathogenicity, plant immunity. *Appl. Microbiol. Biotechnol.* 98, 6921–6932. doi: 10.1007/s00253-014-5863-8

Markowitz, V. M., Mavromatis, K., Ivanova, N. N., Chen, I. M., Chu, K., and Kyrpides, N. C. (2009). IMG ER: a system for microbial genome annotation expert review and curation. *Bioinformatics* 25, 2271–2278. doi: 10.1093/bioinformatics/btp393

Martin, D. M., Gershenzon, J., and Bohlmann, J. (2003). Induction of volatile terpene biosynthesis and diurnal emission by methyl jasmonate in foliage of Norway spruce. *Plant Physiol.* 132, 1586–1599. doi: 10.1104/pp.103.021196

Martin, M. (2011). Cutadapt removes adapter sequcences from high-throughput sequencing reads. *EMBnet. J.* 17, 5–7. doi: 10.14806/ej.17.1.200

Masri, L., Branca, A., Sheppard, A. E., Papkou, A., Laehnemann, D., Guenther, P. S., et al. (2015). Host–pathogen coevolution: the selective advantage of *Bacillus thuringiensis* virulence and its cry toxin genes. *PLos Biol.* 13:e1002169. doi: 10.1371/journal.pbio.1002169

Mathur, H., Rea, M. C., Cotter, P. D., Hill, C., and Ross, R. P. (2015). The sactibiotic subclass of bacteriocins: an update. *Curr. Protein Pept. Sci.* 16, 549–558. doi: 10.2174/1389203716666150515124831

Milutinovic, B., Fritzlar, S., and Kurtz, J. (2014). Increased survival in the red flour beetle after oral priming with bacteria-conditioned media. *J. Innate Immun.* 6, 306–314. doi: 10.1159/000355211

Milutinovic, B., Stolpe, C., Peu, R., Armitage, S. A. O., Kurtz, J., and Milutinovic, B. (2013). The red flour beetle as a model for bacterial oral infections. *PLoS ONE* 8:e64638. doi: 10.1371/journal.pone.0064638

Mitchell, A., Chang, H. Y., Daugherty, L., Fraser, M., Hunter, S., Lopez, R., et al. (2015). The interpro protein families database: the classification resource after 15 years. *Nucleic Acids Res.* 43, D213–D221. doi: 10.1093/nar/gku1243

Müller, S., Garcia-Gonzalez, E., Mainz, A., Hertlein, G., Heid, N. C., Mösker, E., et al. (2014). Paenilamicin: structure and biosynthesis of a hybrid nonribosomal peptide/polyketide antibiotic from the bee pathogen *Paenibacillus larvae*. *Angew Chem. Int. Ed. Engl.* 53, 10821–10825. doi: 10.1002/anie.201404572

Nübel, U., Engelen, B., Felske, A., Snaidr, J., Wieshuber, A., Amann, R. I., et al. (1996). Sequence heterogeneities of genes encoding 16S rRNA in *Paenibacillus polymyxa* detected by temperature gradient gel electrophoresis. *Appl. Environ. Microbiol.* 178, 5636–5643.

Okinaka, R. T., and Keim, P. (2016). The phylogeny of *Bacillus cereus sensu lato*. *Microbiol. Spectr.* 4, 1–12. doi: 10.1128/microbiolspec.tbs-0012-2012

Ongena, M., and Jacques, P. (2008). *Bacillus* lipopeptides: versatile weapons for plant disease biocontrol. *Trends Microbiol.* 16, 115–125. doi: 10.1016/j.tim.2007.12.009

Palma, L., Muñoz, D., Berry, C., Murillo, J., and Caballero, P. (2014). *Bacillus thuringiensis* toxins: an overview of their biocidal activity. *Toxins* 6, 3296–3325. doi: 10.3390/toxins6123296

Patel, K. D., Bhanshali, F. C., and Ingle, S. S. (2011). Diversity and characterization of *Bacillus thuringiensis* isolates from alluvial soil of mahi river basin, India. *J. Adv. Dev. Res.* 2, 14–20.

Pegg, G. F., and Brady, B. L. (2002). *Verticillium Wilts*. Wallingford, CT; Oxfordshire: CABI Publishing.

Petersen, T. N., Brunak, S., von Heijne, G., and Nielsen, H. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nat. Methods* 8, 785–786. doi: 10.1038/nmeth.1701

Pii, Y., Borruso, L., Brusetti, L., Crecchio, C., Cesco, S., and Mimmo, T. (2016). The interaction between iron nutrition, plant species and soil type shapes the rhizosphere microbiome. *Plant Physiol. Biochem.* 99, 39–48. doi: 10.1016/j.plaphy.2015.12.002

Priest, F. G., Barker, M., Baillie, L. W., Holmes, E. C., and Maiden, M. C. (2004). Population structure and evolution of the *Bacillus cereus* Group. *Society* 186, 7959–7970. doi: 10.1128/jb.186.23.7959-7970.2004

Priest, F. G., Kaji, D. A., Rosato, Y. B., and Canhos, V. P. (1994). Characterization of *Bacillus thuringiensis* and related bacteria by ribosomal RNA gene restriction fragment length polymorphisms. *Microbiology* 140, 1015–1022. doi: 10.1099/13500872-140-5-1015

R Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: http://www.R-project.org/

Rocha, L. O., Tralamazza, S. M., Reis, G. M., Rabinovitch, L., Barbosa, C. B., and Corrêa, B. (2014). Multi-method approach for characterizing the interaction between *Fusarium verticillioides* and *Bacillus thuringiensis* subsp. *kurstaki*. *PLoS ONE* 9:e92189. doi: 10.1371/journal.pone.0092189

Romero, F. M., Marina, M., and Pieckenstain, F. L. (2014). The communities of tomato (*Solanum lycopersicum* L.) leaf endophytic bacteria, analyzed by 16S-ribosomal RNA gene pyrosequencing. *FEMS Microbiol. Lett.* 351, 187–194. doi: 10.1111/1574-6968.12377

Saitou, N., and Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* 4, 406–425.

Sambrook, J., and Russell, D. W. (2001). *Molecular Cloning: A Laboratory Manua. 3rd Edn.*, New York, NY: Cold Spring Harbor Laboratory Press.

Scherling, C., Ulrich, K., Ewald, D., and Weckwerth, W. (2009). A metabolic signature of the beneficial interaction of the endophyte *Paenibacillus* sp. isolate and *in vitro*-grown poplar plants revealed by metabolomics. *Mol. Plant. Microbe. Interact.* 22, 1032–1037. doi: 10.1094/MPMI-22-8-1032

Schnathorst, W. C. (1981). "Life cycle and epidemiology of *Verticillium,*" in *Fungal Wilt Diseases of Plants*, eds M. E. Mace, A. A. Bell, and C. H. Beckman (San Francisco, CA: Academic Press), 81–111.

Schnepf, E., Crickmore, N., Van Rie, J., Lereclus, D., Baum, J., Feitelson, J., et al. (1998). *Bacillus thuringiensis* and its pesticidal crystal proteins. *Microbiol. Mol. Biol. Rev.* 62, 775–806.

Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153

Shange, R. S., Ankumah, R. O., Ibekwe, A. M., Zabawa, R., and Dowd, S. E. (2012). Distinct soil bacterial communities revealed under a diversely managed agroecosystem. *PLoS ONE* 7:e40338. doi: 10.1371/journal.pone.0040338

Sheppard, A. E., Nakad, R., Saebelfeld, M., Masche, A. C., Dierking, K., and Schulenburg, H. (2016). High instability of a nematicidal cry toxin plasmid in *Bacillus thuringiensis*. *J. Invertebr. Pathol.* 133, 34–40. doi: 10.1016/j.jip.2015.11.009

Sheppard, A. E., Poehlein, A., Rosenstiel, P., Liesegang, H., and Schulenburg, H. (2013). Complete genome sequence of *Bacillus thuringiensis* strain 407 cry. *Genome Announc.* 1, 158–112. doi: 10.1128/genomeA.00158-12

Silo-Suh, L. A., Lethbridge, B. J., Raffel, S. J., He, H., Clardy, J., and Handelsman, J. (1994). Biological activities of two fungistatic antibiotics produced by *Bacillus cereus* UW85. *Appl. Environ. Microbiol.* 60, 2023–2030.

Singh, S., Braus-Stromeyer, S. A., Timpner, C., Tran, V. T., Lohaus, G., Reusche, M., et al. (2010). Silencing of Vlaro2 for chorismate synthase revealed that the phytopathogen *Verticillium longisporum* induces the cross-pathway control in the xylem. *Appl. Microbiol. Biotechnol.* 85, 1961–1976. doi: 10.1007/s00253-009-2269-0

Soufiane, B., and Côté, J. C. (2009). Discrimination among *Bacillus thuringiensis* H serotypes, serovars and strains based on 16S rRNA, gyrB and aroE gene sequence analyses. *Antonie Van Leeuwenhoek* 95, 33–45. doi: 10.1007/s10482-008-9285-4

Swiontek Brzezinska, M., Jankiewicz, U., Burkowska, A., and Walczak, M. (2014). Chitinolytic microorganisms and their possible application in environmental protection. *Curr. Microbiol.* 68, 71–81. doi: 10.1007/s00284-013-0440-4

Tamura, K., Nei, M., and Kumar, S. (2004). Prospects for inferring very large phylogenies by using the neighbor-joining method. *Proc. Natl. Acad. Sci. U.S.A.* 101, 11030–11035. doi: 10.1073/pnas.0404206101

Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994). CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* 22, 4673–4680. doi: 10.1093/nar/22.22.4673

Tran, V. T., Braus-Stromeyer, S. A., Timpner, C., and Braus, G. H. (2013). Molecular diagnosis to discriminate pathogen and apathogen species of the hybrid *Verticillium longisporum* on the oilseed crop *Brassica napus*. *Appl. Microbiol. Biotechnol.* 97, 4467–4483. doi: 10.1007/s00253-012-4530-1

van Dongen, S. (2007). *Graph Clustering by Flow Simulation*. Dissertation, Amsterdam: Centre for Mathematics and Copmuter Science (CWI).

Wang, J., Zhang, L., Teng, K., Sun, S., Sun, Z., and Zhong, J. (2014). Cerecidins, novel lantibiotics from *Bacillus cereus* with potent antimicrobial activity. *Appl. Environ. Microbiol.* 80, p2633–p2643. doi: 10.1128/AEM.03751-13

Weber, T., Blin, K., Duddela, S., Krug, D., Kim, H. U., Bruccoleri, R., et al. (2015). antiSMASH 3.0-a comprehensive resource for the genome mining of biosynthetic gene clusters. *Nucleic Acids Res.* 43, W237–W243. doi: 10.1093/nar/gkv437

Wemheuer, B., Güllert, S., Billerbeck, S., Giebel, H. A., Voget, S., Simon, M., et al. (2014). Impact of a phytoplankton bloom on the diversity of the active bacterial community in the southern North Sea as revealed by metatranscriptomic approaches. *FEMS Microbiol. Ecol.* 87, 378–389. doi: 10.1111/1574-6941.12230

Wemheuer, B., Wemheuer, F., and Daniel, R. (2012). RNA-Based assessment of diversity and composition of active archaeal communities in the German Bight. *Archaea* 2012, 1–8. doi: 10.1155/2012/695826

Zeise, K., and von Tiedemann, A. (2002). Host specialization among vegetative compatibility groups of *Verticillium dahliae* in relation to *Verticillium longisporum*. *J. Phytopathol.* 150, 112–119. doi: 10.1046/j.1439-0434.2002.00730.x

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Supplementary information

Supplementary information for this manuscript can be found at the Frontiers website under the following address:
http://journal.frontiersin.org/article/10.3389/fmicb.2016.02171/full#supplementary-material

Additionally, supplementary figures and tables are provided along with the electronic version of this thesis (on DVD), under the following paths:

Supplementary Material:

Supplementary Material (S1, S3-S4 Table) Supplementary Material/ChapterII.1/Data Sheets/Data Sheet 1.docx

S2 Table Supplementary Material/ChapterII.5/Data Sheets/Data Sheet 2.xlsx

S5 Table Supplementary Material/ChapterII.5/Data Sheets/Data Sheet 3.xlsx

Supplementary Figures:

S1 Figure Supplementary Material/ChapterII.5/Figures/Image 1.tif

S2 Figure Supplementary Material/ChapterII.5/Figures/Image 2.tif

S3 Figure Supplementary Material/ChapterII.5/Figures/Image 3.tif

CHAPTER III DISCUSSION

III.1 General discussion

Bacillus thuringiensis is a ubiquitous organism and can be isolated worldwide from all known aerobic habitats including soil, aquatic habitats phylloplane, dust , insects, and feces of arid birds (Burges and Hurst, 1977; Donovan et al., 1988; Iriarte et al., 2000; Martin et al., 1989; Poopathi et al., 2014; Schnepf et al., 1998; Smith and Couche, 1991)**.** The importance of *B. thuringiensis* for man is correlated to the species specific pathogenic feature, the production of insecticidal or nematicidal toxins which find greatly application as biocontrol agents in the agriculture and in the industry as either pure formulations but also in the development of various transgenic crops (Schnepf et al., 1998).

The species *B. thuringiensis* is a member of the *Bacillus cereus sensu lato* (Bcsl) group which comprises seven species groups either pathogens, opportunists and environmental pathogens (Argolo-Filho and Loguercio, 2014; Priest et al., 1994). Two species have significant impact on humans and animals such as the etiological agent *B. anthracis* and the food-poisoning *B. cereus* (Helgason et al., 2000). The group includes four additional members, *B. mycoides, B. pseudomycoides, B. weihenstephanensis* and *B. cytotoxicus***,** which are minor investigated but also suspicious in being harmful for humans (Guinebretière et al., 2013; Lechner et al., 1998; Nakamura, 1998; Nakamura and Jackson, 1995). Hence, the classification of isolates of this group is highly important and was extensively studied by applying various molecular typing methods (Liu et al., 2015). The phylogenetic classification of *B. thuringiensis* in the Bcsl group is, however, challenging due to the close genetic relation (Okinaka et al., 2016). Intensive research has been done to clarify the role of *B. thuringiensis* in natural habitats and its primary ecological niche. As a result members of the species have been described as obligate pathogens, opportunist pathogens, soil saprophytes, plant endophytes, and environmental pathogens. However, the question of the "natural habitat" or the primary lifestyle of *B. thuringiensis* is still under discussion (Argolo-Filho and Loguercio, 2014; Ruan et al., 2015). How *B. thuringiensis* has evolved and its diverse and host specific set of virulence factors including Cry, Cyt, Vip, and Sip toxins and why the gene equipment is as well found in habitats where no host organism is available is questionable. Moreover, it is unclear which additional genomic features contribute to the virulence and if a combined mechanism of toxins and virulence factors are

advantageous within the fast evolution and adaptation of the multi-host environmental microparasite *B. thuringiensis*.

The first part of this thesis was to generate high quality reference genomes of nematicidal *B. thuringiensis* strains used in diverse co-evolution experiments. In particular, an in depth comparative genome analysis was performed for the identification and annotation of the key virulence factors and fitness factors contribute to a successful infection of nematodes. In addition, genome sequence wide modifications such as mC5 and mC4 methylation pattern as well as the metabolic spectrum were determined. Mobile elements such as extra-chromosomal elements, bacteriophages, transposases, and insertion elements which influence the genome plasticity and therefore are a driving force for the evolution of the environmental pathogen *B. thuringiensis* were determined (Chapter II.1 and Chapter II.2). To characterize relevant strains with regard to their prophage equipment, a combinatory approach of next generation sequencing (NGS) and basic induction experiments was established, to pin point problematic prophages and difficult production strains (Chapter II.3). In addition, the evolution of nematicidal *B. thuringiensis* strains were determined in host-parasite co-evolution experiments, with *Caenorhabditis elegans* as host. Finally, the "Red Queen" hypothesis was investigated with both organisms addressing general genetic traits under selective pressure in co-evolution (Chapter II.4).

The knowledge of the biology of *B. thuringiensis* is biased due to its role as invertebrate pathogen. The second aim of this thesis was to investigate the biology of plant associated Bcsl group species, especially *B. thuringiensis* (Chapter II.5). The focus of the research was the interactions with wilt causing phytopathogenic *Verticillia* that compete within the same habitat for host-plant resources. Therefore, a strain collection was established with natural wiltype strains enriched for Bcsl members sampled from *Solanum lycopersicum* (tomato) as host plant. The phylogeny of isolates, with regard to the non-trivial classification of Bcsl group members, was investigated. Finally, the anti-fungal potential of wiltype *B. thuringiensis* and *B. weihenstephanensis* strains against were analyzed by using two phytopathogenic fungi *Verticillium dahliae* JR2 and *Vertillium longisporum* 43 which are either natural pathogens of *Solanum lycopersicum* or *Brassicacea* (oil seed rape).

III.2 Whole genome and comparative analysis of *Bacillus thuringiensis*

In this thesis, the three nematicidal strains *B. thuringiensis* MYBT18246, *B. thuringiensis* MYBT18247, and *B. thuringiensis* MYBT18679 as well as one insecticidal *B. thuringiensis* strain Btt were sequenced, assembled and whole genome sequences analyzed with a focus on virulence factors and finally comparative analysis were performed to determine the pan-core genome and resulting strain specific factors (Chapter II.1, Chapter II.2 and Chapter II.4). Each of the nematicidal strains had a highly specific genome size including a variable number of extra-chromosomal elements Table 1.

Table 1. Genome information of fully sequenced *B. thuringiensis* strains.

B. thuringiensis strain	Genome Size [Mbp]	Chromosome Size [Mbp]	Extra-chromosomal elements	Range of extra-chromosomal elements [kb]	Toxins
Bt MYBT18246	**6.75**	**5.87**	**11**	**6-150**	**Cry13Aa2, Cry13Ba1, Cry13Ab1, Cry-like, 2Bin-like, Vip-like**
Bt MYBT18247	**6.14**	**5.61**	**6**	**12-175**	**Cry6Ba2, Cry6Ba3, Cry21-like**
Bt MYBT18679	**6**	**5.48**	**9**	**4-188**	**Cry21Aa3, Cry14Aa2, Cry38Aa2, Cry35Aa5, Cry34Aa5, ETX-MTX2-like**
Btt	**6.3**	**5.67**	**6**	**15-250**	**Cry3Aa13, Cry23Aa2, Cry37Aa2, Vip1Fa1, Vip2Bc1, Bi2Ca1, Vip1Ea1, 2 TTX-MTX2-like**
Bt serovar *konkukian* str. 97-27	5.3	5.24	1	77	no known Cry, Cyt, Vip or Sip homologs
BtYBT-1518	6.67	6	6	17-240	Cry55Aa1, Cry6Aa2 and Cry5Ba2
Bt str. Al Hakam	5.31	5.26	1	55	no known Cry, Cyt, Vip or Sip homologs
Bt BMB171	5.64	5.33	1	312	acrystalliferous mutant strain
Bt serovar *finitimus* YBT-020	5.68	5.36	2	18-139	Cry26Aa and Cry28Aa
Bt serovar *chinensis* CT-43	6.15	5.49	10	8-281	Cry1Aa3, Cry1Ba1, Cry1Ia14, Cry2Aa9, and Cry2Ab1, Vip3Aa10
Bt HD-771	6.43	5.89	8	9-171	Cry1Aa
Bt HD-789	6.33	5.5	6	6-349	Cry4Aa3, Cry4Ba5, Cry10Aa3, Cry11Aa3, Cry60Ba3, and Cry60Aa3, plus 3 Cyt toxin
Bt MC28	6.69	5.41	7	7-429	Cry4Cc1, Cry30Fa1, Cry53Ab1, Cry54Aa1, Cry54Ab1, Cry68Aa1, Cry69Aa1, Cry69Aa2, Cry70Ba1, Cyt1Da1, and Cyt2Aa3
Bt Bt407	6.13	5.5	9	2-501	acrystalliferous
Bt serovar *kurstaki* str. HD73	5.90	5.65	7	7-77	Cry1Ac
Bt serovar *thuringiensis* str. IS5056	6.77	5.49	14	6-328	Cry1Aa3, Cry1Ab21, Cry1Ba1, Cry1Ia14, Cry2Aa9, Cry2Ab1, Vip1, Vip2, and Vip3Aa10
Btserovar *kurstaki* str. YBT-1520	6.58	5.6	11	2-422	Cry1Aa, Cry1Ab, Cry1Ac and Cry2
Bt serovar *kurstaki* str. HD-1	6.76	5.63	13	2-431	Cry1Ab1, Cry1Ac, Cry2Ab, Cry1A, Cry1Ia, Cry2Aa, Cry46-like, Vip3Aa

B. thuringiensis strain	Genome Size [Mbp]	Chromosome Size [Mbp]	Extra-chromosomal elements	Range of extra-chromosomal elements [kb]	Toxins
Bt XL6	5.70	5.31	1	39	not investigated
Bt serovar *galleriae* 4G5	6.74	5.7	10	5-426	Cry1Aa, Cry1Ac, Cry1Ca, Cry1Da, Cry1Ia, Cry2Ab, Cry9Ea and Vip3Aa
Bt strain HD1011	6.09	5.23	4	69-385	no known cry, cyt, or vip homologs
Bt strain HD571	5.31	5.26	1	55	no known cry, cyt, or vip homologs
Bt strain HD682	5.29	5.21	3	7-56	no known cry, cyt, or vip homologs
Bt HD1002	6.57	5.49	7	6-359	Cry60Ba1, Cry60Aa1
Bt serovar *morrisoni* strain BGSC 4AA1	6.17	5.65	6	4-232	Cry3Aa, Cry15Aa
Bt strain YC-10	6.78	5.68	9	7-761	Cry1Aa, Cry1Ac, Cry1Ia, Cry2Aa, Cry2Ab and CryB1
Bt strain:HS18-1	6.40	5.29	9	7-509	Cry4Cb1, Cry50Aa1, Cry69Ab1, Cry30Ga, Cry30Ea, Cry70Aa, Cry71Aa, Cry72Aa, Cry56Aa and Cry54Ba
Bt subsp. *indiana* strain HD521	6.19	5.43	6	7-314	three Cry7
Bt YWC2-8	6.22	5.67	6	8-250	Cry4Cb2, Cry30Ea2, and Cry56Aa1
Bt strain CTC	5.35	5.33	1	25	no known Cry, Cyt, Vip or Sip homologs
Bt serovar *tolworthi* Pasteur Institute Standard strain	6.87	5.9	8	7-437	Vip2, Cry2Aa, Cry1Ia, and Cry1Ea
Bt Bt185	6.39	5.24	8	7-635	Cry8
Bt strain HD12	6.49	5.78	6	17-345	Cry1Fb, Cry1Ae, Cry1Bb, Cry1Ae, Cry1, Cry1ja, Cry1Id, Cry1e, Cry2Ad, Vip3Af, Vip1Ca1, Vip2Ac1, ETX_MTX-like
Bt Bc601	6.11	5.63	6	8-171	not investigated
Bt serovar *alesti* strain BGSC 4C1	5.81	5.4	6	8-267	Cry1Ae, Cry1Gb, Cry2Ab, Cry1M, Vip1, Vip2, Vip3Aa
Bt KNU-07	6.15	5.34	2	293-514	not investigated

*Status 15.12.2016. Toxin information was taken from publications (black), from toxin screening within this thesis (blue), or not investigated (red).

In general, all finished *B. thuringiensis* genomes are highly variable in their genomes sizes and vary from 5.3 Mbp to 6.8 Mbp. Each genome includes a single chromosome with a size ranging from 5.21 Mbp to 6 Mbp and a number of extra-chromosomal elements (1 to 14 only finished genomes were considered; https://www.ncbi.nlm.nih.gov/genome/genomes/486?, accessed 13.12.2016; Table 1). Lawrence *et al.* in 2005 proposed that pathogens have common genome strategies which include continually rearranged genes, gene gain and gene loss in pathogen populations through ongoing adaptations to the host (Lawrence, 2005). This is also true for *B. thuringiensis* strains which show a high variability in the genome (Table 1). The chromosome structure of all nematicidal strains includes various mobile elements (phages, plasmids, and insertion elements) as potential vectors for HGT. *Bacillus thuringiensis* MYBT18246 has the biggest genome of our investigated strains with 6.75 Mbp, organized on one circular chromosome (5.87 Mbp) and eleven plasmids ranging from 6 kb to 150 kb. *Bacillus thuringiensis* MYBT18679 has the smallest genome, however it comprises a higher number of extra-chromosomal elements compared to *B. thuringiensis* MYBT18247. The variable genome sizes of *B. thuringiensis* strains (Table 1), including a highly flexible number of exchangeable extra-chromosomal elements supports the hypothesis that the genome evolution of *B. thuringiensis* includes a well balanced system of gene loss and gene acquisition which is an excellent way for the adaptation to multiple ecological niches (Lawrence, 2005). Moreover, no significant genome reduction is observed within this species which is often described for organism which live at constant environmental conditions such as mycoplasma (Loguercio and Argôlo-Filho, 2015; Razin et al., 1998). In each of the nematicidal strains more than three different toxins were identified which contribute to their pathogenicity (Table 1). *Bacillus thuringiensis* MYBT18247 comprises three nematicidal *cry* toxin genes on three different plasmids. Notably, two genes of nematicidal Cry6 subfamily are located on two different plasmids: p172778 (Cry6Ba2) and p113275 (Cry6Ba3). Besides, a putative Cry21-like nematicidal toxin variant is located on a small cryptic plasmid p15092. However, the strain encodes two Cry6Ba toxins which belong to a currently small group of phylogenetically distinct Cry toxins for which a mode of action is not investigated yet (Figure 6) (Wei et al., 2003). This toxin belongs to a group of unique Cry toxins and could be promising in the creation of new transgenic crop pyramids also with regard to delay evolution of resistance in

nematodes (Carrière et al., 2015). *Bacillus thuringiensis* MYBT18679 encodes six toxins which cover each of the four known Cry toxin homology groups (Figure 6). Plasmid p22591 encodes the nematicidal three-domain toxins *cry*14Aa2 and *cry*21Aa3 toxin genes. In addition, plasmid p15831 encodes *cry*34Aa5 (unrelated Cry toxin), *cry*35Aa5 (Bin-group toxin), and *cry*38Aa2 (Mtx-group toxin). The *B. thuringiensis* MYBT18679 genome comprises toxins against two different hosts, *Nematodes* and *Coleoptera* (Figure 4). It is known that Cry34 and Cry35 act as binary toxin, where Cry 35 is only active in the presence of Cry34 and Cry34 alone shows a reduced biological activity against *Coleoptera* (Kelker et al., 2014). However, Cry34 itself is not related to other Cry toxins. Additionally, we identified the Mtx-group toxin Cry38 which is located together with Cry34/Cry35 in one operon. This operon organization has been observed in *B. thuringiensis* EG5899 but the biological activity of Cry38 and its role in toxicity is unknown (Baum et al., 2004; Rupar et al., 2000). Sheppard *et al.* 2016 could show that the *B. thuringiensis* MYBT18679 plasmid p22591 carrying the nematicidal toxins Cry21 and Cry14 frequently get lost in a bacterial population that is not challenged by a nematicidal host (Sheppard et al., 2016). This has not been observed for the smaller plasmid p15831 carrying the insecticidal toxins. This difference in plasmid stability could indicate that the strain *B. thuringiensis* MYBT18679 has recently undergone a host switch to nematodes and that the gained plasmid p22591 is not completely optimized for the strain. This could be the first example of how a host switch could take place but where the adaptation or the evolution is not finalized yet. *Bacillus thuringiensis* MYBT18246 comprises seven different toxins where notably three nematicidal three-domain toxin *cry*13 gene variants (Cry13Aa2, Cry13Ba1, Cry13Ab1) are located in the chromosome. Mostly, toxin genes in *B. thuringiensis* strains are located on plasmids indicating the importance of conjugative mobilized virulence factors (Palma et al., 2014). The importance of plasmids is reflected by four additional putative toxin genes of different toxin classes that reside on the plasmids p120510 (Vip-like toxin), p120416 (Cry-like toxin) and p109822 (two Bin-like toxins). *Bacillus thuringiensis* MYBT18246 is the only of our sequenced strains that encode a protein annotated as putative vegetative insecticidal toxin which point to insects as alternative host. However, in this thesis we could show that the three nematicidal Cry13 toxins are encoded by the chromosome and notably each gene is associated with a single prophage region. This could reflect a non-plasmid way of gene-mobilization and

thus indicate an incidence of lysogenic conversion (Brüssow et al., 2004) in *B. thuringiensis*. The different toxin variants could be the result of adaptive evolution where the virulence factor was under positive selection by different host switch events which would result in several rounds of adaptation and counter-adaptation.

In this study we investigated the genomes for δ-endotoxins as well as for additional potential virulence factors (Chapter II.1 &Chapter II.2). The research was focused on different types of toxins such as hemolysins and non-haemolytic enterotoxins which have been found to be encoded in the genomes of all nematicidal strains. These toxins have been described as important for the nutrient supply during infection by suppressing host immunity (Tran et al., 2013). Thuringiensin or β–exotoxins type II were not detected in any of the nematicidal strains. Thus these strains match the constraint for WHO recommended insect control agents. Besides, the genomes revealed the presence of many enzymes having toxicity effects against insects and nematodes such as phospholipase C, chitinases, camelysins, collagenases, metalloproteases, bacillolysins, and N-acyl homoserine lactonases. Especially, chitinases enable the degradation of chitin in the midgut peritrophic membrane of many insects (Sampson and Gooday, 1998). Proteases such as camelysins, collagenases, and phospholipases were identified and may play an important role in either activating protoxins (Nisnevitch et al., 2010) destruction of the intestine of *C. elegans* (Peng et al., 2016) or in hydrolyzingphospholipids of host cell membranes (Hergenrother and Martin, 1997). The immune inhibitor A metalloprotease, bacillolysins and an N-acyl homoserine lactonase are suspicious for boosting the nematicidal or insecticidal activity, as well as overcoming the host immune system by cleaving host antibacterial peptides (Fedhila et al., 2002; Park et al., 2008; Raymond et al., 2010). Each strain comprise potential gene clusters for the production of secondary metabolites also used for intra-species specific competition which could be helpful to suppress the host gut microbiome which is part of the immune response and bacterial secondary metabolites such as bacteriocins and microcins are able to suppress other pathogenic microbes. In total, these findings are not uncommon for genomes of the whole *B. thuringiensis* species group. Taking together, the armory of virulence factors comprised in the nematicidal strains clearly underpins the fast evolution and adaptation to a broad spectrum of ecological niches and the variety of potential hosts therein. Combined with the absence of known pathogenicity

and virulence factors against humans, the strains represent interesting subjects for future research on their biocontrol potential.

III.3 Key players in the Evolution of nematicidal *Bacillus thuringiensis*

The question how *B. thuringiensis* evolves and which genetic mechanisms are the driving forces for niche adaptation and what induces the evolution of new toxins is still not fully understood. There are two major mechanisms which have an impact on the evolution of new pathogens including (i) indel driven genetic variation by long term adaptation and (ii) fast genetic recombination of genes acquired by horizontal gene transfer (Arber, 2014; Gogarten et al., 2002). The former includes local changes in the nucleotide sequence, such as insertions, deletions and substitutions where the selection of positive sequence modification is linked to the selective success of the encoding cell lines. Additionally, recombination events such as gene duplications, deletions, and inversion can lead to genetic variations. The latter gains a high number of opportunities if it includes foreign DNA acquired for instance by, transformation, trans-conjugation, and/or transduction. All these molecular mechanism are act in combination and are under strong natural selection (Arber, 2014). Moreover, mobile elements can confer additional genomic flexibility by transposon, and insertion element driven migration of genes between chromosomes, plasmids, bacteriophages, genomic islands, group I intron and group II intron elements (Werren, 2011).

We could show in this study that the sequenced nematicidal strains are well equipped for genomic adaptations by mobile elements as well as for the exchange of genetic material between different strains. Table 1 shows the remarkable plasmid gene pool of the sequenced strains and for representative *B. thuringiensis* strains. The number of extra-chromosomal elements in the species *B. thuringiensis* is highly variable but all encode a variety of functions such as virulence and self-transfer capabilities for the distribution on species and inter-species level (Van der Auwera et al., 2007). Additionally, in the chromosome of each strain a high number of insertion sequences, transposases, transposons, and phages were identified (Chapter II.1 & Chapter II.2). The comparison of all closed *B. thuringiensis* genomes (Table 1) revealed that most of the insecticidal and nematicidal toxins reside on specific mobile extra-chromosomal elements. At a first glance, it looks that the number of plasmids is positively correlating

with the number of identified toxins which is a misleading conclusion. A more detailed analysis reveals that the toxin genes are distributed only on one to three specific plasmids in a strain. The other plasmids apparently do not contribute to the toxin-equipment. However they carry a various number of genes contributing to genome fluidity including transposases, insertion sequences, transcriptional regulators and recombinases. Taking the variety and the different number of plasmids in *B. thuringiensis* strains into account, it gets obvious that plasmids are important exchangeable vehicles that are able to confer selective pathogenically and hostile environmentally relevant advantages for the adaption to new ecological niches and the evolution of new pathogenic weapons. Besides, the well known Cry Cyt, Vip, and Sip toxins most of the potential beneficial factors harbored on plasmids have been rarely studied and have not fully elucidated yet. However, bacteriocins, toxin-antitoxin systems, and mobile elements have been identified on transposable plasmids which support the advantage of plasmids in *B. thuringiensis* (Driss et al., 2011; Fguira et al., 2014; Liu et al., 2008).

Some of the mechanisms to maintain the plasmid based extra-chromosomal genetic information is not very effective and thus the genes can be also easily lost in a population (Sheppard et al., 2016). In *B. thuringiensis* MYBT18246 an alternative mechanism for maintenance and acquisition of virulence factors has evolved. We could show that potential virulence factors are either located on extra-chromosomal elements, often located adjacent to mobile elements allowing molecular genome flexibility but also in the chromosome. Three nematicidal Cry toxin variants are located in the chromosome in *B. thuringiensis* MYBT18246 adjacent to three prophage regions, respectively. Phages are known to contribute to the evolution of bacterial pathogens by lysogenic conversion (Brüssow et al., 2004). The location of phage regions in close proximity to virulence factors in the chromosome of *B. thuringiensis* MYBT18246 could be a strong hint that phages are also important key factors in horizontal gene transfer in *B. thuringiensis* and consequently contribute to evolution. However, evolution is controlled as well by limiting mechanisms. Plasmid carriage has been identified as limiting factor for bacteria-phage co-evolution (Harrison et al., 2015b). Bacteriophages as well influence plasmid dynamics such as plasmid gain, loss or maintenance for instance by methylation guided restriction systems (Harrison et al., 2015a). Moreover, bacterial restriction modification systems, which in many cases are

encoded by prophages, and the microbial immunity system (CRISPR/Cas) limit horizontal gene transfer in bacteria (Arber, 2014). In none of the nematicidal strains CRISPR repeats or *cas* genes were identified. In contrast specific restriction modifications such as N^6-methyladenine (m6A) motifs, N^4-methylcytosin (m4C) motifs, and putative motifs, where the identification of the methylation was not possible, were observed in all strains. Moreover, all strains revealed different modification and restriction pattern indicating the implementation of strains specific pattern guided evolutionary barriers for gene acquisition. Other *B. thuringiensis* strains as well as the majority of Bcsl group members seem to lack CRISPR/Cas systems as well (http://crispr.i2bc.paris-saclay.fr/crispr/ accessed 18.12.2016). In only single *B. thuringiensis* strains CRISPRs were detected and in the majority, only questionable structures are predicted, which indicates that in the *B. thuringiensis* species group and also for the whole Bcsl group the bacterial immunity by restriction is more important than the CRISPR/Cas system. This fits also to the high number of phages recognized in the *B. thuringiensis* chromosome.

III.4 Phages in *Bacillus thuringiensis* and *Bacillus licheniformis* and their role in evolution

The lack of CRISPR/Cas systems in *B. thuringiensis* as well as in *B. licheniformis* could be an explanation for the high number (up to 16) of potential prophages in these organisms (Chapter II.1 & Chapter II.2). Notably, in all sequenced nematicidal strains prophage regions or prophage-like regions were detected either in the chromosome as well as circular-state phages. Also in plasmids potential integrative phages were identified which seems to be common for *B. thuringiensis* (Kanda et al., 1989). The number of identified prophage or prophage-like regions varied in the nematicidal strain in the chromosome from 8 to 16 regions, supporting the importance of phages for the genome plasticity of these organisms. In particular, chromosome sequence alignment comparison with other *B. thuringiensis* revealed that differences are frequently insertions, deletions, DNA rearrangements. Notably, most prominent differences resulted from prophage regions indicating a dominant role of prophages for strain differentiation. Due to the modular nature of prophages those differences are mostly represented by distinct modules of phages such as head and tail structures which are

evolutionary conserved and not always represent completely new prophage regions. Moreover, prophages are transient vehicles in the bacterial chromosome and under decay which leads to defective phages or phage remnants (Brüssow et al., 2004). A broad spectrum of phages contributing to genome architectural changes has been observed for the Bcsl group and also for the *B. thuringiensis* species. For example the phage phIS3501 integrates into the chromosomal *hylII* gene in *B. thuringiensis* sv. *israelensis* ATCC35646, encoding haemolysin II, and thereby disrupts the gene and leads to loss of function (Moumen et al., 2012). The transducing phage TP-13 is able to convert sporulation in combination with crystal production to *B. thuringiensis* (Perlak et al., 1979). In addition, the temperate phage TP-21, identified in *B. thuringiensis* sv. *kurstaki* HD-1, is the only specialized transducing phage for Bcsl group with an additional plasmidal prophage state (Walter and Aronson, 1991). Furthermore, in *B. thuringiensis* MYBT18246 we identified 16 putative prophage-like regions. The high number of identified prophage regions is interesting because as it is known phages are under constant decay and only small amounts are fixed and do not accumulate to large numbers (Brüssow et al., 2004). However, this strain comprises an increased number of prophage-like regions, while three regions had nematicidal toxins located in close proximity, suggesting a selective advantage for the bacteria in the ecological niche. The determination of the pan-core genome of *B. thuringiensis* MYBT18246 using *B. thuringiensis* YBT-1518, *B. thuringiensis* CT-43, and *B. thuringiensis* Bt407 as reference revealed a core of 4,298 genes of shared gene families as well as 1242 genes specific for *B. thuringiensis* MYBT18246. The number of singletons is a bit "higher" compared to what is statistically proposed by (Fang et al., 2011). Singletons are basically important for supplementary biochemical pathways, selective advantages such as adaptation to different habitats, and antibiotic resistance (Vernikos et al., 2015). In *B. thuringiensis* MYBT18246 singletons are attributed to mobile elements and especially to prophage-like regions and hypothetical genes. Considering the toxic activity of Cry toxins adjacent to prophage regions, which could be described as "morons", could provide an advantage to the host and indirect to the prophage by the propagation of the host bacterium (Brüssow et al., 2004). Other species groups of the genus *Bacillus*, including *B. subtilis*, *B. pumilus*, *B. anthracis*, and *B. licheniformis*, harbor prophages in their genomes. As mentioned prophages have an impact on the evolution of strains due to lysogenic conversion. Moreover, the induction of phages

which results in bacterial lysis is of great interest in the industry for strain design as well as for the stability of production processes (Hertel et al., 2015; Ho et al., 2016; Mahony et al., 2016). To address the activities of prophages in genomes a combinatory approach using genome- and next generation sequencing technology and wet lab activity tests for the activity evaluation of prophage regions was established (Chapter II.3). The approach was applied and evaluated using *B. licheniformis* DSM13, which comprises seven prophage-like regions. Deletion mutants where some prophage loci were deleted have been used as a reference system. In particular, we could show that three prophages-loci produced phage particles and are active after induction which could impact the usage in industrial production. Interestingly the phages competed for cellular resources since the particle production of the two less active loci dramatically increased after deletion of the more active phage-loci. The phage particles of the different loci have been visualized with TEM which revealed that in deed different particle types correlated each to a single genome locus.

III.5 Evolution of *Bacillus thuringiensis* in a host

To investigate how pathogenicity evolves under different selection regimes nematicidal *B. thuringiensis* strains and *Caenorhabditis elegans* have been used in a host-parasite co-evolution experiment. In total, 28 infection cycles have been followed using five distinct treatments favoring either host or pathogen evolution or both. The treatments were: (i) host control, (ii), host one-sided adaptation (iii) host-pathogen co-evolution, (iv) parasite one-sided adaptation, (v) pathogen control (Figure 8, Chapter II.4).

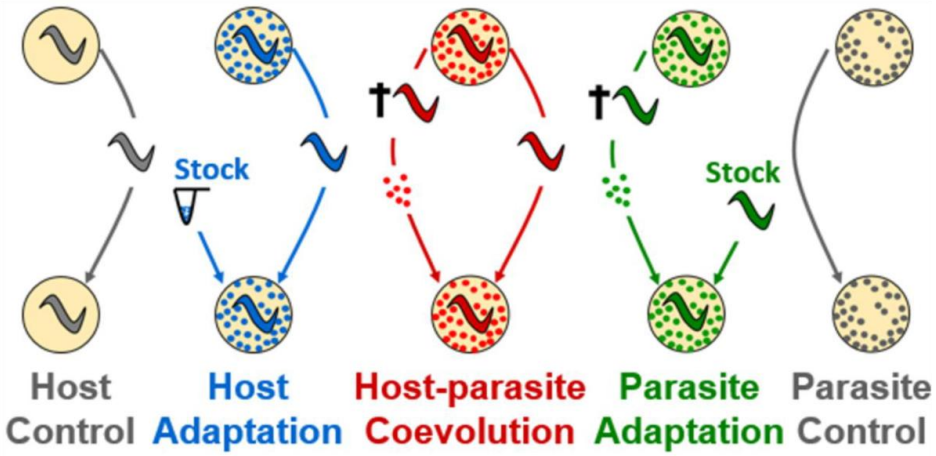

Figure 8. The five evolution treatments: (i) host control (grey) adapting to general laboratory conditions in the absence of the pathogen, (ii) host one-sided adaptation (blue) where the host adapted to the non-evolving, ancestral pathogen taken from a frozen stock culture at each transfer, (iii) host–pathogen co-evolution (red) during which both antagonists were continuously forced to co-evolve to each other, (iv) pathogen one-sided adaptation (green) where the pathogen adapted to the non-evolving, ancestral host population taken from a frozen stock culture at each transfer; and (v) pathogen control (grey) adapting to general laboratory conditions in the absence of the host (Masri et al. 2015).

Especially, the maintenance of the host specific toxin genes has been shown to be under selective pressure. *B. thuringiensis* MYBT18679 was dominant to its competitors within a mixed infection and exhibited the most virulence against the host *C. elegans*. The phenotype has been pin-pointed at two specific toxin genes (Cry21Aa3 and Cry14Aa2 located on the plasmid p22591). The high virulence was positively correlated with elevated copy numbers of the plasmid encoding the nematicidal toxins indicating a selective evolutional advantage for the nematicidal *B. thuringiensis* MYBT18679. Moreover, the maintenance of virulence factors under selection on a high level indicates that continuous co-evolution is distinct from unidirectional selection and that bacterial genomes are shaped by life history characteristics. Notably, the high virulence of *B. thuringiensis* MYBT18679 was fixed in all replicate populations in the co-evolution treatments but only for some in one-sided adaptation treatments highlighting the importance of evolutionary interaction with the host. Infection dynamics were observed in one-sided adaptation treatments, where the more virulent *B. thuringiensis* MYBT18679 was not dominant and in contrast the avirulent strain *B. thuringiensis* BT-22 showed up to same extents. The less virulent strains including *B. thuringiensis* MYBT18246 and *B. thuringiensis* MYBT18247 seem to have a different strategy to compete and persist in a host, profiting from the high virulence of *B. thuringiensis* MYBT18679 in the first infection phase. This is cost effective for those

strains, because they act in the second phase when the host is already weakened. Moreover, it could be a hint that those less virulent strains are better adapted to the specific host because killing of the host is contra productive for long-term persistence (Brzuszkiewicz et al., 2011). In contrast, the high virulent *B. thuringiensis* MYBT18679 kills fast and due to the easily loss and increase of the virulence plasmid in a population could indicate that this plasmid was acquired a short while ago. However, considering that all these nematicidal strains are ingested as spore and enter the vegetative cycle in the host gut, we cannot explain the whole interplay of different *B. thuringiensis* strain in the host-gut microbiome.

We conclude that co-evolution favors bacterial virulence while a persistent pathogenic lifestyle was selected by one-side adaptation (high infection load) or adaptation in the absence of a host (environmental persistence through biofilm formation). Lastly, the here reported study (Chapter II.4.) presents evidence that more than one infection and host-exploitation strategies have evolved in *B. thuringiensis* and that for the maintenance or for the evolution of high virulence a continuous co-evolution with a target host is essential. This indicates a wide spectrum of complex interaction possibilities of *B. thuringiensis* with various host-taxa or maybe even with kingdoms.

III.6 Insights into the Ecology and biological potential of *Bacillus thuringeinsis*

Bacillus thuringiensis is a cosmopolite which can be isolated worldwide from soil, plant and animal associated aerobic habitats. Obviously, the production of spores enables the invasion of soil, aquatic habitats, phylloplane, dust, insects, and feces of arid birds (Burges and Hurst, 1977; Donovan et al., 1988; Iriarte et al., 2000; Martin et al., 1989; Poopathi et al., 2014; Schnepf et al., 1998; Smith and Couche, 1991) (Figure 2). The question if the species can be assigned to a primary ecological niche and lifestyle is a challenge due to the ability to survive and grow in multiple ecological niches such as in a variety of different hosts including insects, nematodes, mites, ticks or in a variety of environments outside of a host, such as in soil, on leaf surfaces or in plant tissues (Argolo-Filho and Loguercio, 2014; Cangelosi et al., 2004; Erban et al., 2009; Guillem and Porcar, 2012; Hendriksen and Hansen, 2002; Schnepf et al., 1998). However different lifestyles have been proposed including: (i) *B. thuringiensis* as insect pathogen

and (ii) *B. thuringiensis* as saprophytic organism (iii) *B. thuringiensis* as environmental multi-host pathogen (Argolo-Filho and Loguercio, 2014; Loguercio and Argôlo-Filho, 2015). How *B. thuringiensis* genomes reflect this enormous flexibility is a fascinating question as well as how such an armory of host-specific toxins against a broad range of host organism living in diverse habitats has been evolved. Consequently, how *B. thuringiensis* is able to adapt to or switch between several ecological niches, within and as well outside of a host organism?

Within this thesis, the objective of one study was the determination and the comparison of the biological potential of natural sampled isolates from the genus *Bacillus* including *B. thuringiensis* and other Bcsl group members such as *B. weihenstephanensis* (Chapter II.5). For this purpose, a *Bacillus* strain collection was generated including approximately 260 new isolates sampled from tomato root-associated soil as primary ecological niche by using an enrichment method for *B. thuringiensis*. The strain collection was used as a first snap shot to investigate whether the classical isolates, which have been generated during the last eight decades to achieve biocontrol strains, are representative for strains of the species *B. thuringiensis*. The abundance of the genera within the root-associated bacterial community composition of tomato plants was determined based on the bacterial 16S rRNA amplicon sequencing. Moreover, twenty *Bacillus* isolates were selected for their phenotypic growth diversity on solid media and their genomes were sequenced. All isolates were tested for their anti-fungal effect against two different *Verticillium* species, against *V. dahliae* JR2 as natural tomato plant pathogen and *V. longisporum* 43 as foreign host plant pathogen, was determined. A phylogenetic analysis revealed that the 20 isolates cluster in 7 *B. thuringiensis* and 13 *B. weihenstephanensis* strains. The dual cultivation assays showed a correlation between taxonomy and antagonistic activities. Strikingly, all *B. thuringiensis* strains exhibited an *in vitro* antifungal effect against tomato pathogen *V. dahliae,* while only limited antagonism was observed against the foreign phytopathogenic *V. longisporum*. Three *B. weihenstephanensis* isolates showing an invasive growth-type competed with both phytopathogenic fungi. The genome analysis of the 20 *Bacillus* strains revealed that strains with antifungal activity shared genes assigned to bacillibactin production and mycolytic chitinases, which are thus the most promising candidates for encoding the antifungal effect.

In conclusion, it could be shown that *B. thuringiensis* strains have anti-fungal effects on *Verticillia* and the produced genomic and physiological data provides the foundation for the identification and characterization of the active antifungal substance of *B. thuringiensis*. Moreover, the biocontrol efficiency must be determined in *in planta* experiments and must be also tested for other phytopathogens or also other kingdoms. For the first time a competing effect of *B. weihenstephanensis* strains as well as a rhizoid growth has been reported within this study. This indicates that members of the species *B. weihenstephanensis* as well may have the potential to be used as future biocontrol agents. However, this has to be investigated in future studies.

III.7 The *Bacillus cereus sensu lato* group and the importance of the taxonomy of the environmental pathogen *Bacillus thuringiensis*

In 2015 and 2016 the issue of the problematic *Bacillus cereus sensu lato* group phylogeny was highly reviewed showing the importance in this field of research (Liu et al., 2015; Okinaka et al., 2016; Wang and Ash, 2015). Obviously, it is of vital interest for the biotechnological application to differentiate between human safe members of the species complex such as *B. thuringiensis* from the lethal ones including *B. anthracis*. The increasing speed of technology development and the lowering costs of next-generation sequencing (NGS) results in a tremendous output of genomic data. As already described above, the experimental phylogenetic classification of new, as well as of previously classified, *B. thuringiensis* strains or of other Bcsl species strains is laborious and challenging (Chapter II.1-Chapter II.5). For the typing of new isolates we combined molecular MLST analysis as described by Priest *et al.* (Priest et al., 2004) and confirmed the assignments by the analysis of phenotypic, biochemical and pathogenic characteristics to avoid misclassifications. We could show that *B. thuringiensis* MYBT18246 and *B. thuringiensis* MYBT18247 which are less virulent compared to *B. thuringiensis* MYBT18679 cluster together in a phylogenetic subgroup and seem to be closer related (Chapter II.2, Chapter II.5). *Bacillus thuringiensis* MYBT18679 clusters into a different subgroup comprising also *B. cereus* and *B. anthracis* strains, supporting the hypothesis that *B. thuringiensis* MYBT18679 newly gained the virulence plasmid for adapting to a new ecological niche, indicating a host switch. However, insufficient 16S rRNA analysis for phylogenetic classification of strains are still

commonly in use (Li et al., 2015; Liu et al., 2014b, 2014c). Researchers combine this insufficient genetic marker (insufficient in this species complex) with the defining phenotypic or pathogenic feature of the species group without taking the genetic background into account. There are multiple examples where it is shown that the pure pathogenic features are misleading (Li et al., 2015; Soufiane et al., 2013; Soufiane and Côté, 2010). Considering the plethora of extra-chromosomal elements of *B. thuringiensis* and those can be easily spread over the species groups or get lost in populations, the pure pathogenicity should not be the basis for phylogeny (Figure 9). For instance, the problem can be defined for the genome of *B. cereus* biovar *anthracis* (Klee et al., 2010). The genome comprises both virulence plasmids pXO1 and pXO2 from *B. anthracis* and causes anthrax but contains a chromosome that sorts to *B. cereus* by the routinely used molecular phylogenetic markers. Should the strain still be classified as *B. cereus* or reclassified as *B. anthracis?* Klee et al. (Klee et al., 2010) decided to take simply all strain features into account for a correct classification leading to this strain description "*B. cereus* var. *anthracis"*. There are many more examples: (ii) changes a *B. thuringiensis* into a *B. cereus* if it loses the *cry*-toxin plasmid? Should it be reclassified as *B. cereus*? (Sheppard et al., 2013) (iii) should a natural *B. thuringiensis* strains which produce crystal toxins but grow at low temperatures such as the psychrophil *B. weihenstephanensis* species be considered as *B. weihenstephanensis* (Soufiane et al., 2013)*?*. We could show that our isolates exhibiting mycoid growth cluster more closely to *B. weihenstephanensis* than to *B. mycoides* or *B. pseudomycoides*. Moreover, the optimal growth range of those strains was lowered compared to *B. thuringiensis* but was not psychrophil as described for *B. weihenstephanensis* strains. Additionally, hemolysis was not observed for mycoid growing strains which are a feature of *B. anthracis* (Chapter II.5). Based on this combination of descriptive "intermediate" features the isolates may represent a new "intermediate" species that nevertheless cluster clearly distinct based on genetic features (MLST). This classification is supported by the study of Liu *et al* 2015. Liu *et al.* 2015 suggested in their phylogenetic analysis of the Bcsl group, based on whole-genome sequence-based Genome BLAST Distance Phylogeny (GBDP) approach, four additional new species members, namely *B. toyonensis* (Jiménez et al., 2013), *B. gaemokensis* (Jung et al., 2010), *B. manliponensis* (Jung et al., 2011) and *B. bingmayongensis* (Liu et al., 2014a, 2015) (Figure 9).

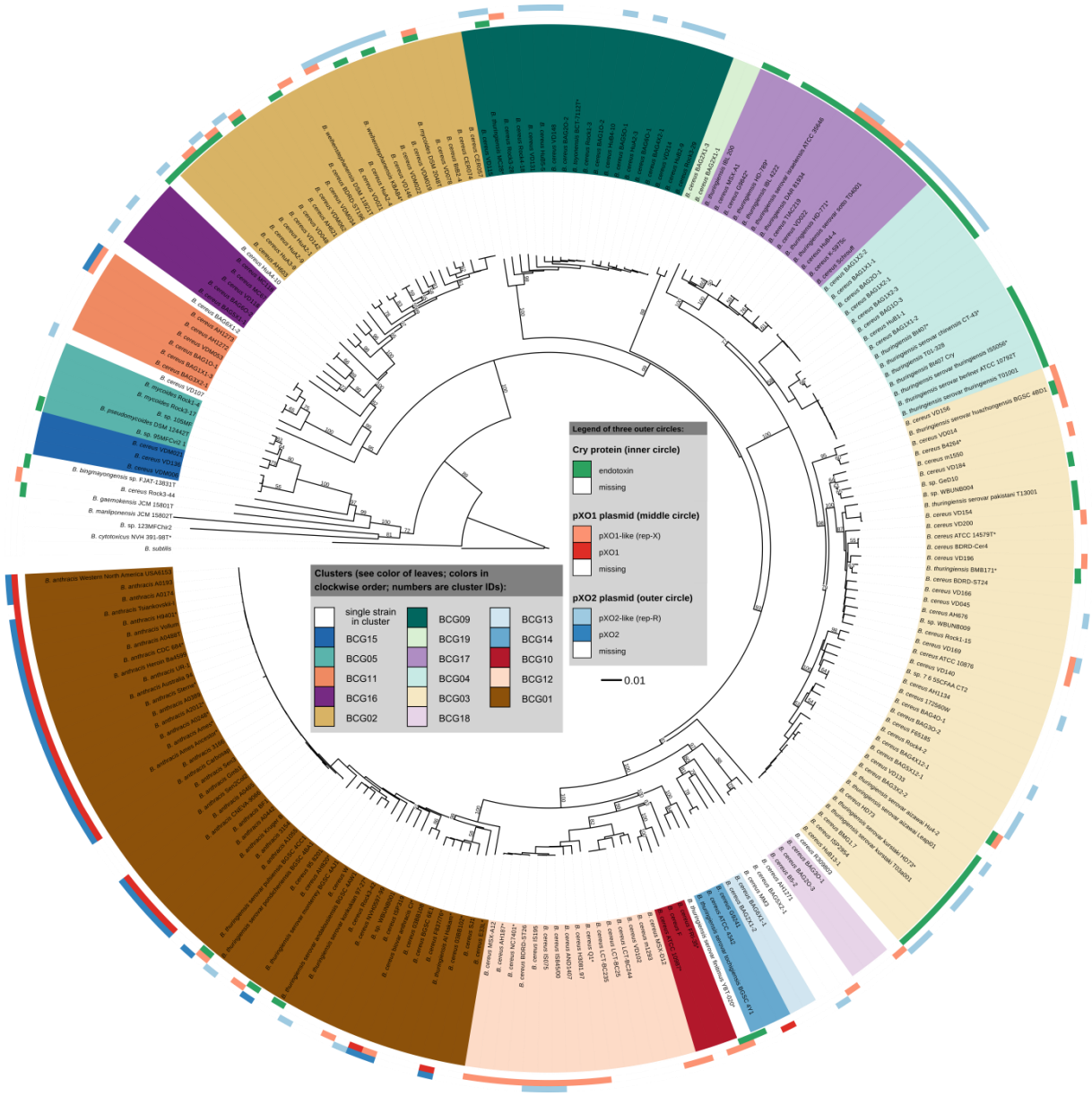

Figure 9. Whole-genome phylogeny of 224 bacteria of the *B. cereus* group inferred using the latest GBDP version and rooted with *B. subtilis* ATCC 6051. Numbers above branches are greedy-with-trimming pseudo-bootstrap33 support values from 100 replicates if larger than 50%. Leaves are colored according to their affiliation to clusters (i.e., *Bacillus cereus* groups, BCG). The three outer circles show whether or not the (i) cry locus, (ii) pXO1(-like) plasmid and/or the (iii) pXO2(-like) plasmid is found. Type strains are printed in bold font as well as marked by an asterisk (*). The tree was inferred using FastME48 and visualized using iTOL57. The figure was taken from (Liu et al., 2015).

As depicted in Figure 9, there are many strains that are classified to one species based on the genetic background but share other species specific features with a different Bscl species group member (virulence plasmids of *B. anthracis*, Cry-toxin plasmids of *B. thuringiensis*). Notably, several strains of the *B. cereus* group such as the designated *B. cereus* bv. *anthracis* carries the plasmids pXO1 and pXO2 and belong to a single chromosomal clade distinct from the *B. anthracis* clade (Antonation et al., 2016). Moreover, those strains seem to exhibit a unique genomic variation based on core chromosomal and plasmid SNP data and their evolution seems to be mostly driven by mutations arising in the context of a clonal lifestyle resulting in an more open pan-genome which contradicts the idea of *B. anthracis* as a monophyletic species. However, genetic rearrangements were not considered and could strongly influence the results as we have shown that the members of the Bcsl comprise many mobile elements. In summary, Figure 9 shows the quandary of the Bcsl phylogeny which shows directly the contradiction of a genomic versus a pathogenic classification. Notably, all published reviews regarding exactly this topic (Liu et al., 2015; Okinaka et al., 2016; Wang and Ash, 2015) are again based on different classification methods. The overall question one must ask is: What exactly should be the GOLD-standard for classification? The presence of virulence factors for a specific species group? Or should the overall genetic background should be the basis for a classification?

In conclusion, the plethora of different molecular and bioinformatic methods in combination with a lack of wet-lab information (phenotype determination, metabolic properties of strains) lead to an inefficient taxonomy for the whole Bcsl taxonomy. This is problematic considering the importance of Bcsl members for humans. A GOLD-standard must be set for the species group and a reclassification must be performed for all strains available in databases. Moreover, 16S analysis should be avoided and not accepted for a phylogenetic classification in sequence databases or journals. In future, the available sequence data will increase and methods for taxonomic classification should be alignment free by using Genome-Blast Distance Phylogeny (GBDP) and Feature Frequency Profile (FFP). An additional advantage of these methods is that also unique genetic material (plasmids) is considered for the analysis which definitely plays a crucial role in taxonomy as well as in virulence. However, considering the need for a fast and cheap strain typing in laboratories MLST or MLSA represent still a valuable approach compared to whole genome sequencing. Finally for security reason the pure

taxonomic classification has to be complemented by a phenotypic investigation of pathology factors if new isolates shall be used in research or biotechnology.

CHAPTER IV SUMMARY AND CONCLUSION

IV.1 Summary

Two different research topics of *B. thuringiensis* have been investigated in this thesis and are interdisciplinary linked: Genomics of *B. thuringiensis* and the Ecological importance of *B. thuringiensis*. Genomics allows following many evolutionary processes and key events in the evolution of bacteria and can be tracked down to the molecular level. But how these mechanisms influence the natural genetic variation and phenotypic traits in bacteria under a selective pressure is under debate. Our knowledge about the interplay of genetics and niche adaptation, host switches and evolution of organisms, especially with regard to interaction partners is still limited (Figure 10). The different molecular mechanisms of evolution as well as the full adaptive potential of *B. thuringiensis* are poorly understood. *Bacillus thuringiensis* is a good model organism for studying evolution because of its inhabitation of various ecological niches and specific hosts. Moreover, the armory of toxins produced by *B. thuringiensis*, their industrial usage and the opportunity to investigate complex lifestyle switches makes the organism to a perfect study model. Despite the increasing number of studies on the evolution of pathogens only a glimpse is known about single evolutionary mechanisms which partially contribute to "rapid evolution". Additionally, how a given selection regime determines the various opportunities of *B. thuringiensis* as a species is unknown. The aim of this thesis was to gain insights in the emerging of successful new *B. thuringiensis* strains from the complex and diverse evolutionary puzzle of this species group by investigating the genomic equipment and their mechanisms behind.

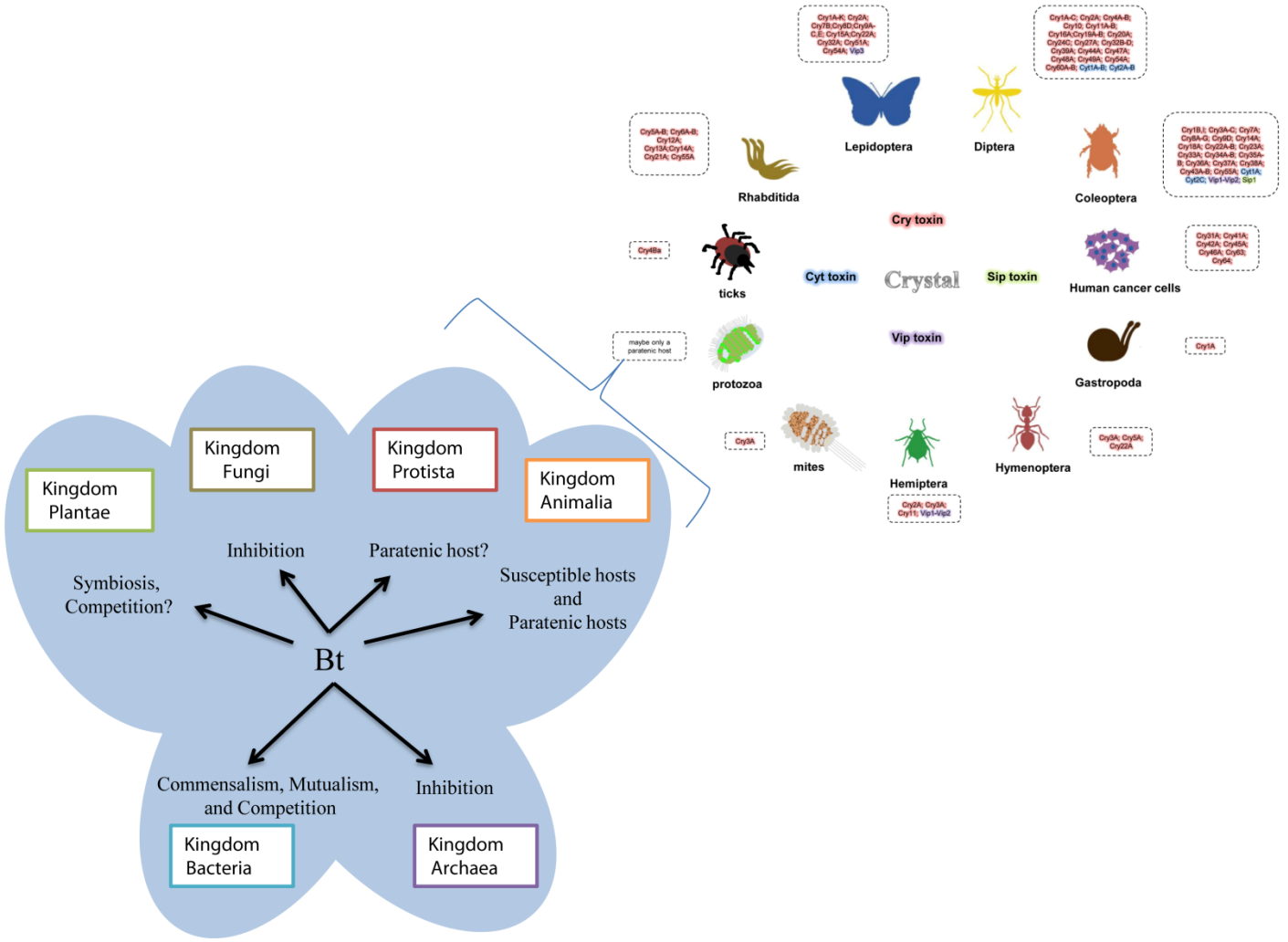

Figure 10. Possible interaction scenarios of *B. thuringiensis* with the six kingdoms: Plantae, Fungi, Protista, Animalia, Archaea, and Bacteria.

In summary, the species *B. thuringiensis* is versatile, especially with respect to their genetics (Chapter II.1, Chapter II.2). Whole genomes of nematicidal *B. thuringiensis* strains were investigated with a focus on virulence factors, fitness factors, methylation pattern and metabolic properties. Moreover, mobile elements such as bacteriophages, IS elements and transposases were determined and characterized. Two strains, *B. thuringiensis* MYBT18246 and *B. thuringiensis* MYBT18247 encoded a plethora of virulence-, fitness factors and mobile elements including nematicidal toxins. In *B. thuringiensis* MYBT18246 three different nematicidal toxin genes (*cry*13Aa2, *cry*13Ba1 and *cry*13Ab1) were identified on the chromosome. In addition, four additional putative toxin genes were identified with low sequence similarity to other known toxins on plasmids. In *B. thuringiensis* MYBT18247 three *cry* toxin genes (*cry*6Ba2, *cry*6Ba3 and *cry*21-like) were identified on plasmids. In both strains additional virulence factors were identified which might contribute to host pathogenicity. In Chapter II.2, in *B. thuringiensis* MYBT18247 different toxins such as hemolysins and non-haemolytic enterotoxins were identified. Beside theses virulence factors as well fitness factors such as enzymes like chitinases, camelysins, collagenases, phospholipases, immune inhibitor A metalloproteases, bacillolysins and N-acyl homosoerine lactonase have been detected. All those fitness factors strengthen the pathogen in *infectio* and thus ultimately increase the pathogenicity *B. thuringiensis* against specific hosts. In addition factors targeting at competing bacteria especially, microcins, biosynthetic clusters including bacteriocins, siderophores, NRPS and terpene cluster which have been identified und thus support the hypothesis of a complex interplay of *B. thuringiensis* with hosts but also with other bacteria or other kingdoms. Comparative genomic analysis of *B. thuringiensis* strains, including *B. thuringiensis* MYBT18246 and *B. thuringiensis* MYBT18247 revealed that prophages are highly abundant in this species group. Moreover the analysis in Chapter II.1 showed that the *B. thuringiensis* MYBT18246 genome comprised a large number of mobile elements involved in genome plasticity including eleven plasmids and sixteen chromosomal prophages. The close proximity of the chromosomal nematocidal *cry* toxin genes to three distinct prophages regions indicated a contribution of phages in defining the host range of this strain. In addition, plasmids and prophages are important HGT elements indicating the importance of genetic exchange as driving force for the evolution of this pathogen and for pathogen/host adaptation.

The study described in Chapter II.3 presented a new next generation sequencing (NGS) method in combination with classical TEM- and gel-based methods for an in depth bacteriophages analysis. The method showed that in a host-genome multiple prophage infections are sequentially active and compete for the translational capacities of their host-bacteria, especially in *Bacilli*. In the test organism *B. licheniformis* DSM13 seven prophage- regions were identified. Induction experiments (Mitomycin C treatment) revealed the activity of three prophages and with the help of the sensitive NGS approach, read mappings enabled the assignment of active prophages to their genomic locations. The NGS approach can be used for strain optimization in biotechnological relevant strains, such as *B. licheniformis* DSM13, where a sufficient and especially stable production is needed.

Co-evolution experiments described in Chapter II.4 was the first which investigated and dissects the phenotypic and genomic consequences of experimental co-evolution and one-sided adaptation of a host (*C. elegans*) and its parasite (*B. thuringiensis*). In particular, there is a selective advantage of nematicidal toxin genes correlated with their high copy number during the process of adaptation to the host und co-evolution conditions. In contrast, under one-sided adaptation the selective advantage of a high infection load correlated to the loss of killing activity by *B. thuringiensis* genotypes. In some populations it resulted in the complete extinction of pathogenic *B. thuringiensis* strains. This supports the hypothesis of wide-spread co-evolutionary interactions of *B. thuringiensis* and various host taxa under natural conditions where virulence has to be maintained (Chapter II.4) (Figure 10). Minor studies of *B. thuringiensis* interaction with other kingdoms than the well known interaction with Animalia include the kingdom of Protista (Manasherob et al., 1998), of Plantae, Bacteria and Archaea. In plants an endophytic lifestyle was observed, inhibition was detected against other bacteria as well as against Archaea (Praça Botelho et al., 2012; Salazar-Marroquín et al., 2016; Yudina et al., 2004) demonstrating different ways of interaction and the uninvestigated full potential of *B. thuringiensis*.

The ecological interaction types between biological partners such as competition, amensalism, exploitation, neutralism, commensalism, mutualism, or symbiosis in complex environments shared between animals, plants, fungi and *B. thuringiensis* has been rarely studied in detail. Despite the increasing studies on insecticidal and nematocidal toxins of *B. thuringiensis* because of their use in agriculture, the antifungal

potential has been observed but not intensively studied. Considering the evolutionary advantage of plant associated bacteria protecting their host organisms the second aim of this thesis was to investigate if there is an anti-fungal potential of plant associated *B. thuringiensis* strains and to determine their genomic relation to Bcsl group species (Chapter II.5). This was investigated using *Solanum lycopersicum* as host plant and two species of wilt causing phytopathogenic *Verticillia.*

The study described in Chapter II.5 resulted in the expected identification of new isolated wildtype strains of *B. thuringiensis* and as well as isolates of another Bcsl group species *B. weihenstephanensi*. Twenty genomes of selected isolates were sequenced and the phylogeny was determined. The phylogenetic classification of new isolates was confirmed based on a combination of morphological, pathogenic, and physiological properties. The 20 isolates were assigned to 7 *B. thuringiensis* and 13 *B. weihenstephanensis* strains. The anti-fungal potential of the strains were tested against two phytopathogenic fungi, namely *Verticillium dahliae* JR2 and *Vertillium longisporum* 43, which are either natural pathogens of *Solanum lycopersicum* or *Brassicacea* (oil seed rape), respectively. A dual cultivation assay revealed a correlation between taxonomy, morphology and antagonistic activities. All *B. thuringiensis* strains exhibited an *in vitro* antifungal effect against *V. dahliae* while only limited antagonism was observed against *V. longisporum*. Additionally, three *B. weihenstephanensis* isolates showed an invasive growth-type and competed apparently with both phytopathogenic fungi. The relation of the rhizoid growth and the mechanism of competition of *B. weihenstephanensis* strains have not been described previously. Genome analysis of all strains with antifungal activity encodes genes assigned to bacillibactin production and mycolytic chitinases. In summary, *B. thuringiensis* is not a standalone entity and for understanding the whole evolutionary process and the importance of *B. thuringiensis* in nature an overall holistic perception must be considered (Figure 10).The results provided in this thesis contribute to our understanding of the complex interplay of the environmental pathogen *B. thuringiensis* with the environment including multiple host organisms and the genetics contributing to this lifestyle.

IV.2 Conclusion

In conclusion, the evolution of *B. thuringiensis* is a multi-mechanism process influenced by variety of different parameters such as genomic plasticity, natural conditions and interaction partners including hosts, competitors or commensals.

- The genomes of *B. thuringiensis* contain approximately 15% to 20% genetic material encode by genomic elements assigned to genome plasticity.
- *cry*-toxins encoded by plasmids have been identified as actively selected features under different selection regimes. The effect of the genes has been confirmed by molecular reconstruction in mutant strains.
- For the first time prophages have been identified as likely candidates for the mobilization of chromosomally encoded cry-toxin in *B. thuringiensis.*

The second research topic demonstrated antifungal potential of *B. thuringiensis* wildtype strains against phytopathogenic *Verticillium* species and produced genomic and physiological data for the identification, purification and characterization of new active antifungal substances.

- Exclusively, all *B. thuringiensis* isolates from the rhizosphere of *Solanum lycopersicum* inhibit the growth of the tested *Verticillium* species in a *cry*-toxin independent manner.
- For the first time it has been shown that isolates of *B. weihenstephanensis* have an impact on the growth of *Verticillium*. The isolates clustered in a distinct phylogenetic subgroup of *B. weihenstephanensis*.

IV.3 Outlook

In future, the different fields of research must be linked and especially a new standard must be set for a re-classification of the whole Bcsl group. Sequencing methods for taxonomic classification should be probably are alignment free such as Genome-Blast Distance Phylogeny (GBDP) and Feature Frequency Profile (FFP). An additional advantage of these methods is that also unique genetic material (plasmids) is considered for the analysis which definitely plays a crucial role in taxonomy. Secondly, due to the high amount of sequencing data and that many new putative virulence factors can be identified in co-evolution experiments a new problem arose due to the lack basic biological research because for most of the interesting candidate genes an in depth function or basic function is unknown. The existence of various toxins and unknown factors with lack of host-specificity is a problem that must be solved for drawing a comprehensive picture of *B. thuringiensis* ecology related to pathogenicity and relationships between microbes and their environment. The results provided in this thesis contribute to our understanding of the complex interplay of the environmental pathogen *B. thuringiensis* with the environment including multiple host organisms and the genetics contributing to this lifestyle.

CHAPTER VGENERAL REFERENCES

V.1 References

Antonation, K. S., Grützmacher, K., Dupke, S., Mabon, P., Zimmermann, F., Lankester, F., et al. (2016). *Bacillus cereus* Biovar Anthracis Causing Anthrax in Sub-Saharan Africa—Chromosomal Monophyly and Broad Geographic Distribution. *PLoS Negl. Trop. Dis.* 10, e0004923. doi:10.1371/journal.pntd.0004923.

Arber, W. (2014). Horizontal Gene Transfer among Bacteria and Its Role in Biological Evolution. *Life* 4, 217–224. doi:10.3390/life4020217.

Argolo-Filho, R. C., and Loguercio, L. L. (2014). *Bacillus thuringiensis* is an environmental pathogen and host-specificity has developed as an adaptation to human-generated ecological niches. *Insects* 5, 62–91. doi:10.3390/insects5010062.

Van der Auwera, G. A., Timmery, S., Hoton, F., and Mahillon, J. (2007). Plasmid exchanges among members of the *Bacillus cereus* group in foodstuffs. *Int. J. Food Microbiol.* 113, 164–172. doi:10.1016/j.ijfoodmicro.2006.06.030.

Baum, J. A., Chu, C. R., Rupar, M., Brown, G. R., Donovan, W. P., Huesing, J. E., et al. (2004). Binary toxins from *Bacillus thuringiensis* active against the western corn rootworm, *Diabrotica virgifera virgifera* LeConte. *Appl. Environ. Microbiol.* 70, 4889–4898. doi:10.1128/AEM.70.8.4889-4898.2004.

Brüssow, H., Canchaya, C., and Hardt, W.-D. (2004). Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. *Microbiol. Mol. Biol. Rev.* 68, 560–602. doi:10.1128/MMBR.68.3.560–602.2004.

Brzuszkiewicz, E., Thürmer, A., Jörg, S., Leimbach, A., Liesegang, H., Meyer, F.-D., et al. (2011). Genome sequence analyses of two isolates from the recent *Escherichia coli* outbreak in Germany reveal the emergence of a new pathotype: Entero-Aggregative-Haemorrhagic *Escherichia coli* (EAHEC). *Arch. Microbiol.* 193, 883–891. doi:10.1007/s00203-011-0725-6.

Burges, D., and Hurst, J. A. (1977). Ecology of *Bacillus thuringiensis* in storage moths. *J. Invertebr. Pathol.* 30, 131–139. doi:10.1016/0022-2011(77)90210-5.

Cangelosi, G. A., Freitag, N. E., and Buckley, M. R. (2004). From outside to inside: Environmental Microorganisms as Human Pathogens. Washington, DC, USA.

Carrière, Y., Crickmore, N., and Tabashnik, B. E. (2015). Optimizing pyramided transgenic Bt crops for sustainable pest management. *Nat. Biotechnol.* 33, 161–168. doi:10.1038/nbt.3099.

Donovan, W. P., Gonzalez, J. M., Gilbert, M. P., and Dankocsik, C. (1988). Isolation and characterization of EG2158, a new strain of *Bacillus thuringiensis* toxic to coleopteran larvae, and nucleotide sequence of the toxin gene. *Mol. Gen. Genet.* 214, 365–72.

Driss, F., Tounsi, S., and Jaoua, S. (2011). Relationship between plasmid loss and gene expression in *Bacillus thuringiensis*. *Curr. Microbiol.* 62, 1287–1293. doi:10.1007/s00284-010-9857-1.

Erban, T., Nesvorna, M., Erbanova, M., and Hubert, J. (2009). *Bacillus thuringiensis* var. *tenebrionis* control of synanthropic mites (Acari: Acaridida) under laboratory conditions. *Exp. Appl. Acarol.* 49, 339–346. doi:10.1007/s10493-009-9265-z.

Fang, Y., Li, Z., Liu, J., Shu, C., Wang, X., Zhang, X., et al. (2011). A pangenomic study of *Bacillus thuringiensis*. *J. Genet. Genomics* 38, 567–576. doi:10.1016/j.jgg.2011.11.001.

Fedhila, S., Nel, P., and Lereclus, D. (2002). The InhA2 metalloprotease of *Bacillus thuringiensis* strain 407 is required for pathogenicity in insects infected via the oral route. *J. Bacteriol.* 184, 3296–3304. doi:10.1128/JB.184.12.3296-3304.2002.

Fguira, I. Ben, Fourati, Z., Kamoun, F., Tounsi, S., and Jaoua, S. (2014). Isolation of the *Bacillus thuringiensis* plasmid carrying Bacthuricin F4 coding genes and evidence of its conjugative transfer. *J. Infect. Dev. Ctries.* 8, 4–9. doi:10.3855/jidc.3552.

Gogarten, J. P., Doolittle, W. F., and Lawrence, J. G. (2002). Prokaryotic evolution in light of gene transfer. *Mol. Biol. Evol.* 19, 2226–2238. doi:10.1093/oxfordjournals.molbev.a004046.

Guillem, M., and Porcar, M. (2012). Ecological Mysteries: is *Bacillus thuringiensis* a Real Insect Pathogen? *Bt Res.* 3, 1–2. doi:10.5376/bt.2012.03.0001.

Guinebretière, M. H., Auger, S., Galleron, N., Contzen, M., de Sarrau, B., de Buyser, M. L., et al. (2013). *Bacillus cytotoxicus* sp. *nov*. is a novel thermotolerant species of the *Bacillus cereus* group occasionally associated with food poisoning. *Int. J. Syst. Evol. Microbiol.* 63, 31–40. doi:10.1099/ijs.0.030627-0.

Harrison, E., Truman, J., Wright, R., Spiers, A. J., Paterson, S., Brockhurst, M. A., et al. (2015a). Plasmid carriage can limit bacteria – phage coevolution. *Biol. Lett.* 11, 20150361. doi:10.1098/rsbl.2015.0361.

Harrison, E., Wood, J. A., Dytham, C., Pitchford, J. W., Truman, J., Spiers, A., et al. (2015b). Bacteriophages Limit the Existence Conditions for Conjugative Plasmids.

MBio 6, 1–9. doi:10.1128/mBio.00586-15.Editor.

Helgason, E., Okstad, O. A., Caugant, D. A., Johansen, H. A., Fouet, A., Mock, M., et al. (2000). *Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*--one species on the basis of genetic evidence. *Appl. Environ. Microbiol.* 66, 2627–30.

Hendriksen, N. B., and Hansen, B. M. (2002). Long-term survival and germination of *Bacillus thuringiensis* var. *kurstaki* in a field trial. *Can. J. Microbiol.* 48, 256–61.

Hergenrother, P. J., and Martin, S. F. (1997). Determination of the Kinetic Parameters for Phospholipase C (*Bacillus cereus*) on Different Phospholipid Substrates Using a Chromogenic Assay Based on the Quantitation of Inorganic Phosphate. *Anal. Biochem.* 251, 45–49. doi:10.1006/abio.1997.2251.

Hertel, R., Rodríguez, D. P., Hollensteiner, J., Dietrich, S., Leimbach, A., Hoppert, M., et al. (2015). Genome-Based Identification of Active Prophage Regions by Next Generation Sequencing in *Bacillus licheniformis* DSM13. *PLoS One.* 10, e0120759. doi:10.1371/journal.pone.0120759.

Ho, C.-H., Stanton-Cook, M., Beatson, S. A., Bansal, N., and Turner, M. S. (2016). Stability of active prophages in industrial *Lactococcus lactis* strains in the presence of heat, acid, osmotic, oxidative and antibiotic stressors. *Int. J. Food Microbiol.* 220, 26–32. doi:10.1016/j.ijfoodmicro.2015.12.012.

Iriarte, J., Porcar, M., Lecadet, M.-M., and Caballero, P. (2000). Isolation and Characterization of *Bacillus thuringiensis* Strains from Aquatic Environments in Spain. *Curr. Microbiol.* 40, 402–408. doi:10.1007/s002840010078.

Jiménez, G., Urdiain, M., Cifuentes, A., López-López, A., Blanch, A. R., Tamames, J., et al. (2013). Description of *Bacillus toyonensis* sp. *nov.*, a novel species of the *Bacillus cereus* group, and pairwise genome comparisons of the species of the group by means of ANI calculations. *Syst. Appl. Microbiol.* 36, 383–91. doi:10.1016/j.syapm.2013.04.008.

Jung, M. Y., Kim, J. S., Paek, W. K., Lim, J., Lee, H., Kim, P. Il, et al. (2011). *Bacillus manliponensis* sp. nov., a new member of the *Bacillus cereus* group isolated from foreshore tidal flat sediment. *J. Microbiol.* 49, 1027–1032. doi:10.1007/s12275-011-1049-6.

Jung, M. Y., Paek, W. K., Park, I. S., Han, J. R., Sin, Y., Paek, J., et al. (2010). *Bacillus gaemokensis* sp. *nov.*, isolated from foreshore tidal flat sediment from the Yellow Sea. *J. Microbiol.* 48, 867–871. doi:10.1007/s12275-010-0148-0.

Kanda, K., Tan, Y., and Aizawa, K. (1989). A Novel Phage Genome Integrated into a Plasmid in *Bacillus thuringiensis* Strain AF1O1. *J. Gen. Microbiol.* 135, 3035–3041. doi:10.1099/00221287-135-11-3035.

Kelker, M. S., Berry, C., Evans, S. L., Pai, R., McCaskill, D. G., Wang, N. X., et al. (2014). Structural and Biophysical Characterization of *Bacillus thuringiensis* Insecticidal Proteins Cry34Ab1 and Cry35Ab1. *PLoS One.* 9, e112555. doi:10.1371/journal.pone.0112555.

Klee, S. R., Brzuszkiewicz, E. B., Nattermann, H., Brüggemann, H., Dupke, S., Wollherr, A., et al. (2010). The genome of a *Bacillus* isolate causing anthrax in chimpanzees combines chromosomal properties of *B. cereus* with *B. anthracis* virulence plasmids. *PLoS One.* 5. e10986. doi:10.1371/journal.pone.0010986.

Lawrence, J. G. (2005). Common themes in the genome strategies of pathogens. *Curr. Opin. Genet. Dev.* 15, 584–588. doi:10.1016/j.gde.2005.09.007.

Lechner, S., Mayr, R., Francis, K. P., Prüss, B. M., Kaplan, T., Wiessner-Gunkel, E., et al. (1998). *Bacillus weihenstephanensis* sp. *nov.* is a new psychrotolerant species of the *Bacillus cereus* group. *Int. J. Syst. Bacteriol.* 48 Pt, 1373–1382. doi:10.1099/00207713-48-4-1373.

Li, Q., Xu, L. Z., Zou, T., Ai, P., Huang, G. H., Li, P., et al. (2015). Complete genome sequence of *Bacillus thuringiensis* strain HD521. *Stand. Genomic Sci.* 10, 62. doi:10.1186/s40793-015-0058-1.

Liu, B., Liu, G. H., Hu, G. P., Cetin, S., Lin, N. Q., Tang, J. Y., et al. (2014a). *Bacillus bingmayongensis* sp. *nov.*, isolated from the pit soil of Emperor Qin's Terra-cotta warriors in China. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* 105, 501–510. doi:10.1007/s10482-013-0102-3.

Liu, X., Zhou, R., Fu, G., Zhang, W., Min, Y., Tian, Y., et al. (2014). Draft Genome Sequence of *Bacillus thuringiensis* NBIN-866 with High Nematocidal Activity. *Genome Announc* 2, 1998–1999. doi:10.1128/genomeA.00429-14.Copyright.

Liu, X., Zhu, S., Ye, W., Ruan, L., Yu, Z., Zhao, C., et al. (2008). Genetic characterization of two putative toxin-antitoxin systems on cryptic plasmids from *Bacillus thuringiensis* strain YBT-1520. *J. Microbiol. Biotechnol.* 18, 1630–1633.

Liu, Y., Lai, Q., Göker, M., Meier-Kolthoff, J. P., Wang, M., Sun, Y., et al. (2015). Genomic insights into the taxonomic status of the *Bacillus cereus* group. *Sci. Rep.* 5, 14082. doi:10.1038/srep14082.

Liu, Y. Y., Ye, W., Zheng, J., Fang, L., Peng, D., Ruan, L., et al. (2014c). High-quality draft genome sequence of nematocidal *Bacillus thuringiensis* Sbt003. *Stand. Genomic Sci.* 9, 624–631. doi:10.4056/sig.

Loguercio, L. L., and Argôlo-Filho, R. C. (2015). Anthropogenic action shapes the evolutionary ecology of *Bacillus thuringiensis*: Response to Ruan et al. *Trends Microbiol.* 23, 519–520. doi:10.1016/j.tim.2015.06.002.

Mahony, J., McDonnell, B., Casey, E., and van Sinderen, D. (2016). Phage-Host Interactions of Cheese-Making Lactic Acid Bacteria. *Annu. Rev. Food Sci. Technol.* 7, 267–285. doi:10.1146/annurev-food-041715-033322.

Manasherob, R., Ben-Dov, E., Zaritsky, A., and Barak, Z. (1998). Germination , Growth , and Sporulation of *Bacillus thuringiensis* subsp. *israelensis* in Excreted Food Vacuoles of the Protozoan Tetrahymena pyriformis. *Appl. Environ. Microbiol.* 64, 1750–1758.

Martin, P. A. W., Traverst, R. S., Travers, R. S., and Traverst, R. S. (1989). Worldwide Abundance and Distribution of *Bacillus thuringiensis* Isolates Worldwide Abundance and Distribution of *Bacillus thuringiensis* Isolates. *Appl. Environ. Microbiol.* 55, 2437–2442.

Moumen, B., Nguen-The, C., and Sorokin, A. (2012). Sequence Analysis of Inducible Prophage phIS3501 Integrated into the Haemolysin II Gene of *Bacillus thuringiensis* var *israelensis* ATCC35646. *Genet. Res. Int.* 2012, e543286. doi:10.1155/2012/543286.

Nakamura, L. K. (1998). *Bacillus pseudomycoides* sp. *nov. Int. J. Syst. Bacteriol.* 48 Pt 3, 1031–5. doi:10.1099/00207713-48-3-1031.

Nakamura, L. K., and Jackson, M. A. (1995). Clarification of the taxonomy of *Bacillus mycoides*. *Int. J. Syst. Bacteriol.* 45, 46–49.

Nisnevitch, M., Sigawi, S., Cahan, R., and Nitzan, Y. (2010). Isolation, Characterization and Biological Role of Camelysin from *Bacillus thuringiensis* Subsp. *israelensis*. *Curr. Microbiol.* 61, 176–183. doi:10.1007/s00284-010-9593-6.

Okinaka, T. R., Keim, P., Okinaka, R. T., and Keim, P. (2016). The Phylogeny of *Bacillus cereus sensu lato*. *Microbiol. Spectr.* 4, TBS-0012-2012. doi:10.1128/microbiolspec.TBS-0012-2012.

Palma, L., Muñoz, D., Berry, C., Murillo, J., and Caballero, P. (2014). *Bacillus thuringiensis* Toxins: An Overview of Their Biocidal Activity. *Toxins.* 6, 3296–

3325. doi:10.3390/toxins6123296.

Park, S. J., Park, S. Y., Ryu, C. M., Park, S. H., and Lee, J. K. (2008). The role of AiiA quorum quenching enzyme from *Bacillus thuringensis*, on the rhizosphere competence. *J. Microbiol. Biotechnol.* 18, 1518–1521.

Peng, D., Lin, J., Huang, Q., Zheng, W., Liu, G., Zheng, J., et al. (2016). A novel metalloproteinase virulence factor is involved in *Bacillus thuringiensis* pathogenesis in nematodes and insects. *Environ. Microbiol.* 18, 846–862. doi:10.1111/1462-2920.13069.

Perlak, F. J., Mendelsohn, C. L., and Thorne, C. B. (1979). Converting bacteriophage for sporulation and crystal formation in *Bacillus thuringiensis*. *J. Bacteriol.* 140, 699–706.

Poopathi, S., Thirugnanasambantham, K., Mani, C., Ragul, K., and Sundarapandian, S. M. (2014). Isolation of mosquitocidal bacteria (*Bacillus thuringiensis*, *B. sphaericus* and *B. cereus*) from excreta of arid birds. *Indian J. Exp. Biol.* 52, 739–47.

Praça Botelho, L., Gomes Mendes Menezes, C. A., Cabral, G., Martins, É. S., Sujii, E., Monnerat, R. G., et al. (2012). Endophytic Colonization by Brazilian Strains of *Bacillus thuringiensis* on Cabbage Seedlings Grown in Vitro. *Bt Res.* 3, 11–19. doi:10.5376/bt.2012.03.0003.

Priest, F. G., Barker, M., Baillie, L. W. J., Holmes, E. C., and Maiden, M. C. J. (2004). Population Structure and Evolution of the *Bacillus cereus* Group. *Society* 186, 7959–7970. doi:10.1128/JB.186.23.7959.

Priest, F. G., Kaji, D. A., Rosato, Y. B., and Canhos, V. P. (1994). Characterization of *Bacillus thuringiensis* and related bacteria by ribosomal RNA gene restriction fragment length polymorphisms. *Microbiology* 140, 1015–1022. doi:10.1099/13500872-140-5-1015.

Raymond, B., Johnston, P. R., Nielsen-LeRoux, C., Lereclus, D., and Crickmore, N. (2010). *Bacillus thuringiensis*: An impotent pathogen? *Trends Microbiol.* 18, 189–194. doi:10.1016/j.tim.2010.02.006.

Razin, S., Yogev, D., and Naot, Y. (1998). Molecular biology and pathogenicity of Mycoplasmas. *Microbiol. Mol. Biol. Rev.* 62, 1094–156. doi:1092-2172/98.

Ruan, L., Crickmore, N., Peng, D., and Sun, M. (2015). Are nematodes a missing link in the confounded ecology of the entomopathogen *Bacillus thuringiensis*? *Trends*

Microbiol. 23, 341–346. doi:10.1016/j.tim.2015.02.011.

Rupar, M. J., Donovan, W. P., Tan, Y., and Slaney, A. (2000). *Bacillus thuringiensis* CryET29 compositions toxic to coleopteran insects and *Ctenocephalides* spp. U.S. patent 6093695.

Salazar-Marroquín, E. L., Galán-Wong, L. J., Moreno-Medina, V. R., Reyes-López, M. Á., and Pereyra-Alférez, B. (2016). Bacteriocins synthesized by *Bacillus thuringiensis*: generalities and potential appplications. *Rev. Med. Microbiol.* 27, 95–101. doi:10.1097/MRM.0000000000000076.

Sampson, M. N., and Gooday, G. W. (1998). Involvement of chitinases of *Bacillus thuringiensis* during pathogenesis in insects. *Microbiology* 144, 2189–2194. doi:10.1099/00221287-144-8-2189.

Schnepf, E., Crickmore, N., Rie, J. Van, Lereclus, D., Baum, J., Feitelson, J., et al. (1998). *Bacillus thuringiensis* and Its Pesticidal Crystal Proteins. *Microbiol. Mol. Biol. Rev.* 62, 775–806.

Sheppard, A. E., Nakad, R., Saebelfeld, M., Masche, A. C., Dierking, K., and Schulenburg, H. (2016). High instability of a nematicidal Cry toxin plasmid in *Bacillus thuringiensis*. *J. Invertebr. Pathol.* 133, 34–40. doi:10.1016/j.jip.2015.11.009.

Sheppard, A. E., Poehlein, A., Rosenstiel, P., Liesegang, H., and Schulenburg, H. (2013). Complete Genome Sequence of *Bacillus thuringiensis* Strain 407 Cry. *Genome Announc* 1, 158–12. doi:10.1128/genomeA.00158-12.

Smith, R. a., and Couche, G. A. (1991). The phylloplane as a source of *Bacillus thuringiensis* variants. *Appl. Environ. Microbiol.* 57, 311–315.

Soufiane, B., Baizet, M., and Côté, J. C. (2013). Multilocus sequence analysis of *Bacillus thuringiensis* serovars *navarrensis*, *bolivia* and *vazensis* and *Bacillus weihenstephanensis* reveals a common phylogeny. *Antonie van Leeuwenhoek, Int. J. Gen. Mol. Microbiol.* 103, 195–205. doi:10.1007/s10482-012-9800-5.

Soufiane, B., and Côté, J. C. (2010). *Bacillus thuringiensis* serovars *bolivia*, *vazensis* and *navarrensis* meet the description of *Bacillus weihenstephanensis*. *Curr. Microbiol.* 60, 343–349. doi:10.1007/s00284-009-9547-z.

Tran, S.-L., Guillemet, E., Lereclus, D., and Ramarao, N. (2013). Iron regulates *Bacillus thuringiensis* haemolysin hlyII gene expression during insect infection. *J. Invertebr. Pathol.* 113, 205–208. doi:10.1016/j.jip.2013.04.001.

Vernikos, G., Medini, D., Riley, D. R., and Tettelin, H. (2015). Ten years of pan-genome analyses. *Curr. Opin. Microbiol.* 23, 148–154. doi:10.1016/j.mib.2014.11.016.

Walter, T. M., and Aronson, A. I. (1991). Transduction of certain genes by an autonomously replicating *Bacillus thuringiensis* phage. *Appl. Environ. Microbiol.* 57, 1000–1005.

Wang, A., and Ash, G. J. (2015). Whole Genome Phylogeny of *Bacillus* by Feature Frequency Profiles (FFP). *Sci. Rep.* 5, 13644. doi:10.1038/srep13644.

Wei, J.-Z., Hale, K., Carta, L., Platzer, E., Wong, C., Fang, S.-C., et al. (2003). *Bacillus thuringiensis* crystal proteins that target nematodes. *Proc. Natl. Acad. Sci. U. S. A.* 100, 2760–2765. doi:10.1073/pnas.0538072100.

Werren, J. H. (2011). Selfish genetic elements, genetic conflict, and evolutionary innovation. *Proc. Natl. Acad. Sci. U. S. A.* 108, 10863–10870. doi:10.1073/pnas.1102343108.

Yudina, T. G., Bryukhanov, A. L., and Netrusov, A. I. (2004). Susceptibility of Archaea to the Antibiotic Effect of the Parasporal Inclusion Proteins from Different *Bacillus thuringiensis* subspecies. *Microbiology* 73, 19–23. doi:10.1023/B:MICI.0000016362.55398.82.

CHAPTER VI APPENDIX

VI.1 Acknowledgement

Als erstes möchte ich Prof. Rolf Daniel für die Übernahme des Referates und die Ermöglichung meiner Arbeit in seiner Abteilung bedanken. PD Dr. Michael Hoppert möchte ich für die Übernahme des Co-Referates danken. Die vielen konstruktiven Gespräche sowie Planungsmeetings an denen ihr teilgenommen habt, haben mir immer geholfen die Arbeit und auch die Publikationen fertig zu stellen. Ich möchte mich daher für euer beider herzliches Interesse an meiner Arbeit bedanken.
Desweiteren gilt mein Dank den Mitgliedern der Prüfungskommission, Jun.-Prof. Heimel, PD Dr. Kramer, Prof. Braus und Prof. Vidal, für ihre Bereitschaft diese Aufgabe zu übernehmen.

Meiner ganzen Familie möchte ich für jegliche Unterstützung danken. Insbesondere aber meinen Eltern, ihr seid mir immer eine große emotionale Stütze, egal wie „schwierig“ ich manchmal bin!
Ich danke meinem Freund Niko, für seine bedingungslose Liebe und das du immer für mich da bist und auf mich aufpasst.

Ich möchte mich auch bei meinen Leidensgenossen Astrid, Katja, Chris und Stefania bedanken, besonders für eure Freundschaft. Ihr habt mich immer immer immer immer wieder aufgebaut und zum Lachen gebracht. ☺

Im speziellen möchte ich mich bei meinem Doktorvater Dr. Heiko L. bedanken, der mich und meine Arbeit den ganzen Weg hinweg begleitet hat. Ich danke dir für deine intensive Betreuung, gute Gespräche und den täglichen Rückhalt. Ich finde wir sind einfach ein super Team, auch wenn du mich, (und ich dich wahrscheinlich auch) am Ende einige graue Haare gekostet hast.
Ganz besonders möchte ich mich bei Anja bedanken. Wobei ein einfaches „Danke“ nicht ausreicht! Du standest mir, fachlich und auch menschlich, immer zur Seite, hast mich betreut, supported, und an mich geglaubt wenn ich gedanklich schon wieder auf dem Dach war. VIELEN DANK.

Franzi und Bernd möchte ich für den uneingeschränkten Rückhalt, Support und nette Kooperationen danken.
Ein weiterer Dank geht an die vielen Kooperationspartner. Rebekka Harting möchte ich für eine wundervolle Kooperation danken. Dein Lachen sowie dein wissenschaftlicher Enthusiasmus sind einfach ansteckend.
Dem ganzen SPP1399 Cluster möchte für jahrelange gute Kooperationen danken, aber insbesondere Prof. Hinrich Schulenburg, Prof. Joachim Kurtz und Dr. Rebecca Schulte-Iserlohe möchte ich für zahlreiche konstruktive Gespräche und Feedback sowie bereitgestellte Daten danken.

Sascha, Robert und Cynthia möchte ich für die Hilfe auf jeglicher Ebene danken, wissenschaftlich wie auch menschlich.
Additionally, I want to thank Amelie who suffered with me the “end-phase” and always made me smile when I used to freak out like HULK. ☺
Zuletzt, möchte ich dem ganzen G2L-Team danken, oben wie auch unten. Dominik, Kristin, Richard, Simone, Hennes, Dirk, Amelie, Randi und Ming-Ji, Sonja Volland, Ela, Heiko N., Andrea, Florian, Daniela und Petra für immer nette Gespräche.
Den technischen Assistentinnen, Frauke, Melanie, insbesondere Mechthild und Kathleen, dafür dass sie mir immer mit Rat und Tat, Tipps und Tricks zur Seite standen.
Dem Bioinformatik-Support-Team, Tarek & Ansgar möchte ich danken, da ihr immer gesprungen seid wenn mein PC wieder ein Eigenleben entwickelt hat, oder ICH wieder inkompatibel mit dem Rest der Technik.

VI.2 Thesis Declaration

Declaration of plagiarism

I hereby confirm that I have written this doctoral thesis independently. I have not used other sources or facilities others than the ones mentioned in the chapters. The contributions of the authors are given proceeding the respective manuscripts. Moreover, I have not used unauthorized assistance and have not submitted this thesis previously in any form for another degree at any institution or university.

_____

Göttingen,

VI.3 Additional publications

1. **The green impact: bacterioplankton response towards a phytoplankton spring bloom in the southern North Sea assessed by comparative metagenomic and metatranscriptomic approaches.**

   Bernd Wemheuer, Franziska Wemheuer, **Jacqueline Hollensteiner**, Frauke-Dorothee Meyer, Sonja Voget, and Rolf Daniel; *Frontiers in Microbiology* (2015), 6:805

2. **Draft genome sequence of the endophyte *Bacillus pumilus* GM3FR, isolated from aerial plant tissues of *Festuca rubra L.***

   Jacqueline Hollensteiner, Anja Poehlein, Rolf Daniel, Heiko Liesegang, Stefan Vidal, and Franziska Wemheuer, *Genome Announcement* submitted, January, 2017

3. **Draft genome sequence of *Pseudomonas putida* GM4FR isolated from the aerial plant tissues of *Festuca rubra L.***

   Franziska Wemheuer, **Jacqueline Hollensteiner**, Anja Poehlein, Sandra Granzow, Rolf Daniel, Stefan Vidal and Bernd Wemheuer, *Genome Announcement* submitted, January, 2017

VI.4 Curriculum Vitae

Name: Jacqueline Hollensteiner

Date of Birth: 17th July, 1987

Place of Birth: Detmold, Germany

Education

Since 01/2013	Georg-August-University Göttingen, Germany **PhD thesis at the Department of Genomic and Applied Microbiology at the Institute of Microbiology and Genetics** "Evolutionary Genomics of *Bacillus thuringiensis*"
08/2010-09/2012	Georg-August-University Göttingen, Germany **Master of Science in Biochemistry and Microbiology** "Comparative and functional plasmid genomics on nematocidal *Bacillus thuringiensis"*
08/2007-08/2010	Georg-August-University Göttingen, Germany **Bachelor of Science in Biology** "The defective Phage PBLB as reparatur mechanism of *Bacillus licheniformis*" (Der defekte Phage PBLB als Reparaturmechanismus von *Bacillus licheniformis*)
1998-2007	Secondary school Leopoldinum in Detmold
1994-1998	Primary school in Heiligenkirchen, Germany

Practical experiences

02/2016	Supervision of Lab Rotation Stefani Diaz
01/2016	Supervision of Bachelor Thesis "Genetische Zugänglichkeit und Antibiotika Resistenzen von Isolaten der Spezies *Bacillus thuringiensis* und *Bacillus weihenstephanensis"* (by Thore Christian Schulz)
06/2015	Supervision of Bachelor Thesis "Phylogenetische Analyse von *Bacillus thuringiensis* Stämmen mit *Solanum lycopersicum* assoziierter Lebensweise" (by Jan Teer)
06/2014	Supervision of Bachelor Thesis "Improving the handling of selected strains of *Bacillus thuringiensis* by determination of the resistance-spectrum and optimization of the genetic accessibility" by (by Simon Schiwek)
06/2014	Supervision of Lab Rotation Anna Kolarzyk
04/2014	Supervision of Lab Rotation Simon Schiwek
03/2014	Supervision of Lab Rotation Michael Keuler
01/2013	Supervision of Bachelor Thesis "Charakterisierung und Analysie induzierbarer Prophagen" (by Jan Christoph Müller)
11/2012	Supervision of Lab Rotation of Jan Christoph Müller
11/2012-12/2012	Georg-August-University Göttingen, Germany Research Assistance

Scientific Meetings and Method Courses

05/2016	SPP1399 5th *Bacillus*-Invertebrate-Meeting, Münster
03/2016	VAAM 2016 - Annual Conference 2016 of the Association for General and Applied Microbiology, Jena, Germany
10/2015	SPP1399 *Bacillus*-Invertebrate-Meeting
09/2015	ProkaGENOMICS 2015 - 6th European Conference on Prokaryotic and Fungal Genomics in Göttingen, Germany
10/2014	SPP1399 4th *Bacillus*-Invertebrate-Meeting
08/2014	SIP2014 Mainz International congress on invertebrate pathology an microbial control 47th Annual Meeting of the Society for Invertebrate pathology
03/2014	SPP1399 *Bacillus*-Invertebrate-Meeting, Osnabrück
11/2013	SPP1399 *Bacillus*-Invertebrate-Meeting, Göttingen
17/6-21/06/2013	Lab Rotation- Practical course "Department of Evolutionary Ecology and Genetics at the Zoological Institute CAU Kiel
07/2013	SPP1399 *Bacillus*-Invertebrate-Meeting, Münster
03/2013	SPP1399 3rd Bacillus-Invertebrate-Meeting, Giessen

DVD