# The intimate host-pathogen relation of *Paenibacillus larvae* and the honey bee *Apis mellifera*

Tine Descamps

Promotor: prof. dr. Dirk de Graaf
Copromotor: prof. dr. Paul De Vos
Dissertation submitted in fulfillment of the requirements for the degree of
Doctor (Ph.D.) in Science: Biochemistry and Biotechnology
Academic year 2017-2018

# Members of the examination committee

**Chair**
prof. dr. Savvas Savvides

**Secretary**
prof. dr. Aurélien Carlier

**Voting members**
prof. dr. Peter Vandamme
prof. dr. Daisy Vanrompay
prof. dr. Marc Heyndrickx
dr. Stefan Roels

**Non-voting members**
prof. dr. Dirk de Graaf
prof. dr. Paul De Vos

# Dankwoord

Vier jaar geleden kreeg ik na mijn thesis er te hebben uitgevoerd, het aanbod om ook een doctoraat te starten aan het L-MEB. Vol goede moed begon ik met *Paenibacillus larvae* te werken. En hoewel ik een hevige fan van de bacterie was, bleek de liefde niet altijd wederzijds te zijn. Gelukkig kon ik steeds rekenen op mijn collega's om me verder op weg te helpen met hun kennis en goede raad, maar ook om gewoon een luisterend oor te bieden. Ook mijn promoter prof. de Graaf en mijn co-promoter prof. De Vos waren steeds beschikbaar om mijn doctoraat tot een goed einde te helpen brengen.

Mijn familie en vrienden hebben tot vervelens toe moeten luisteren naar mijn verhalen over bijen en bacterien. Ze hebben me echter steeds gesteund en zonder hen zou ik nooit zo ver geraakt zijn.

Ten slotte wil ik hier ook Niels bedanken, om me te troosten en op te peppen als er weer iets fout liep en om mee te vieren als er iets gelukt was.

# Samenvatting

Honingbijen (*Apis mellifera*) worden wereldwijd bedreigd door talrijke factoren, gaande van voedseltekorten over pesticiden tot pathogenen. De belangrijkstse onder de bacteriële pathogenen is *Paenibacillus larvae*, veroorzaker van Amerikaans Vuilbroed. Deze aangifte-plichtige ziekte tast enkel het broed aan en leidt onbehandeld tot het ten onder gaan van de ganse kolonie. De eerste meldingen van Amerikaans vuilbroed gaan terug tot in de Griekse Oudheid, waar Aristoteles gewag maakte van een broedziekte met dezelfde specifieke symptomen. Tot deze symptomen behoren het kleverig worden van de dode larvale massa, waardoor deze tot een slijmerige draad kan getrokken worden en de specifieke "vuile"geur van het broed (vandaar de naam vuilbroed).

De meest gebruikte genotyperingsmethode voor *P. larvae* is ongetwijfeld ERIC rep-PCR. Het grote voordeel van deze methode is dat ze de bacterie opsplitst in biologisch relevante groepen. Elk van de ERIC genotypes heeft een specifieke virulentie. ERIC I is de minst virulente op larvaal niveau van de vier groepen. Deze lage larvale virulentie zorgt echter dat de larve sterft wanneer ze reeds in een gesloten broedcel zit. Hierdoor herkennen de volwassen werksters de ziekte pas in een later stadium en zorgen ze voor een efficiëntere verspreiding van de sporen in de kast. ERIC genotype II is virulenter op larvaal niveau, de werksters herkennen aangetaste larven echter sneller waardoor ze deze kunnen opruimen vooraleer deze tot een kleverige massa gevuld met sporen worden gereduceerd. Dit geno-type is dus minder virulent op kolonieniveau. ERIC genotypes III en IV zijn nog virulenter op larvaal niveau dan ERIC II, maar zijn reeds decennia lang niet meer gevonden in het veld. Het grote nadeel van de ERIC genotypering is dat het ongeschikt is voor epidemiolo-gisch onderzoek. In deze thesis werd een typeringsmethode MLVA geïmplementeerd voor *P. larvae* waarbij de verschillende ERIC types konden opgedeeld worden in verscheidene groepen. De typeringsmethode maakt het mogelijk om zeer snel de ERIC genotypering af te leiden, waardoor het grote voordeel van deze typeringsmethode behouden blijft.

Een efficiënte aanpak van een ziekte is slechts mogelijk wanneer de onderliggende in-fectieweg volledig gekend is. Veel virulentiefactoren van *P. larvae* zijn momenteel nog onbekend. Een doel van deze thesis is het aanbieden van nieuwe virulentiefactoren door een onvooringenomen mutagenese. Hoewel de transposon gebaseerde mutagenese gefaald lijkt, werden toch onvooringenome mutanten met een aangetaste virulentie bekomen. Op basis van deze mutanten kon een lijst worden aangelegd met de kandidaten voor virulentiefactoren. Het bevestigen van deze kandidaten ligt buiten deze thesis.

# Abstract

Honey bees (*Apis mellifera*) are threatened worldwide by a whole spectrum of factors, going from feed shortage over pesticides to pathogens. One of the most important bacterial pathogens is *Paenibacillus larvae*, the causative agent of American Foulbrood. This notifiable disease is only infectious for the brood, but untreated it can cause the collapse of the whole colony. The first known description of American Foulbrood dates back to the ancient Greeks, where Aristotle described a brood disease with the same specific symptoms. These symptoms include the transformation of the larvae to a sticky ropy mass and the specific foul smell coming from the brood (hence the name foulbrood).

The most used genotyping method for *P. larvae* is without doubt ERIC rep-PCR. The big advantage of this method is that it clusters the bacterial strains into biologically relevant groups. Each of the ERIC genotypes has a specific virulence. ERIC I is the least virulent on larval level, but most virulent on the colony level. This is because the low larval virulence results in the dead of the larvae on a moment where the brood cell is already closed. This makes it difficult for the adult bees to recognize the dead larva. By the time they clean up the larva it is already transformed to a ropy mass and the bees get covered in infectious spores. This leads to an efficient spread of the disease throughout the colony and thus a high virulence on colony level. ERIC genotype II is more virulent on the larval level and kills the larva before the brood cell is capped. Worker bees recognize these larvae faster and are able to clean them up before they transform to the ropy mass. This results in a lower virulence at colony level. Finally ERIC genotype III and IV are highly virulent on the larval level, but have not been detected in the field for decades. Although this genotyping method is very relevant, it is unfit for epidemiological research. In this thesis we implemented the MLVA typing method for *P. larvae*. The method is able to subdivide the different ERIC groups in several subgroups, but it still enables to quickly recognize the ERIC genotype each strain belongs to. It is thus able to subdivide the ERIC strains without losing its biological relevance.

An efficient method to handle any disease is only possible when the underlying pathways used for infection are fully understood. Many virulence factors of *P. larvae* are currently still unknown. A goal of this thesis was the identification of new virulence factors using an untargeted mutagenesis. Although the transposon based mutagenesis failed, we were able to obtain untargeted mutants with a diminished virulence. Using these mutants, a list with candidate virulence factors was constructed.

# Contents

# Abbreviations

**A**             Adenine

**ABC**           Adenosine triphosphate binding cassette

**AFB**           American Foulbrood

**ADP**           Adenosine diphosphate

**AFLP**          Amplified fragment length polymorphism

**AMP**           Antimicrobial peptide

**BCCM**          Belgian coordinated collections of microorganisms

**BLAST**         Basic local alignment search tool

**C**             Cytosine

**CCD**           Colony collapse disorder

**CFU**           Colony forming unit

**Cry protein**   Crystal protein

**Dif**           Dorsal related immune factor

**DNA**           Deoxyribonucleic acid

**EMP**           Embden-Meyerhof-Parnas

**EPS**           Exo-polysaccharides

**ERIC rep-PCR**  Enterobacterial repetitive intergenic consensus repetitive element polymerase chain reaction

**EU**            European union

**FAVV-AFSCA**    Federaal agentschap voor de veiligheid van de voedselketen - Agence fédérale pour la sécurité de la chaîne alimentaire

**FISH**          Fluorescence in situ hybridization

**GMO**           Genetically modified organism

| | |
|---|---|
| **GO** | Gene ontology |
| **Ld** | Lethal dose |
| **Lt** | Lethal time |
| **MALDI-TOF MS** | Matrix assisted laser desorption/ionisation time-of-flight mass spectrometry |
| **Mbp** | Megabasepair |
| **MCS** | Multicloning site |
| **ME** | Mosaic ends |
| **MLST** | Multilocus sequence typing |
| **MLVA** | Multiple Locus Variable number of tandem repeat Analysis |
| **MUT** | Methionine ABC uptake transporter |
| **NRPS** | Non-ribosomal peptide synthethases |
| **OD** | Optical density |
| **ORF** | Open reading frame |
| **PCR** | Polymerase chain reaction |
| **PCR-DGGE** | Polymerase chain reaction denaturing gradient gel electrophoresis |
| **PFGE** | Pulsed field gel electrophoresis |
| **PGRP** | peptidoglycan recognition protein |
| **PIAS** | Protein inhibitor of activated STAT |
| **PKS** | Polyketide synthethases |
| **PM** | Peritrophic membrane |
| **PMF** | Peptide mass fingerprint |
| **proPO** | Prophenoloxidase |
| **PTS** | Phosphotransferase system |
| **qPCR** | Quantitative polymerase chain reaction |
| **RATE** | Random amplification of transposon ends |
| **rDNA** | Ribosomal deoxyribonucleic acid |
| **REFP** | Restriction endonuclease fragment patterns |
| **RFLP** | Restriction fragment length polymorphism |

| | |
|---|---|
| **RNA-Seq** | Ribonucleic acid sequencing |
| **rRNA** | Ribosomal ribonucleic acid |
| **SDS-PAGE** | Sodium dodecyl sulfate polyacrylamide gel electrophoresis |
| **SE** | Standard error |
| **SOCS** | Suppressor of cytokine signaling |
| **TgIP** | Transgenerational immune priming |
| **TLR** | Toll-like receptors |
| **TNF** | Tumor necrosis factor |
| **TRAF-2** | TNF receptor associated factor 2 |
| **UDP-Glc DH** | Uridine triphosphate glucose dehydrogenase |
| **USA** | United states of America |
| **VNTR** | Variable number of tandem repeat |

# Chapter 1

# Introduction

## 1   In a nutshell

The European honey bee *Apis mellifera* is subject to many challenges. Among these challenges are the winter mortality and colony collapse disorder (CCD), often mentioned in the media. One of the diseases that causes mortality is the so-called American Foul Brood (AFB). The etiological agent of AFB is a less studied bacterium (less than 400 references in *Web of knowledge* [1]) and now named *Paenibacillus larvae*. It forms the subject of this PhD thesis.

The bacterium is gram-positive, rod shaped with size ranging from 2.5-5 $\mu$m by 0.5-0.8 $\mu$m (Figure 1.1). The pathogen is motile due to its peritrichous flagellae and forms infectious endospores. It is highly adapted to its specific host, the honey bee larva. Very young larvae ($< 36h$ age) are most susceptible and resistance increases with age. It causes a deadly intestinal infection of the honey bee brood [2]. AFB is a notifiable disease in most countries, but differences exist in the adopted control strategy.
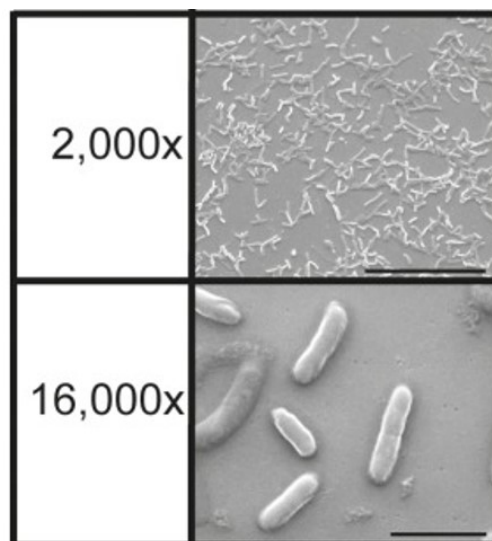


Figure 1.1: **Electron microscopy analysis of *P. larvae*.** Bars represent 50 µm (2,000 fold) and 5 µm (16,000 fold). Figure from [3].

The earliest mention of AFB is probably done by Aristotle in book IX of his *Historia Animalium* (Figure 1.3). The Greek scientist described a deadly honey bee disease which can be identified by the bad smell coming from the brood [4]. The name foulbrood was introduced in 1769 by Schirach (Figure 1.3), who described a honey bee disease with a foul smell as a characteristic symptom, hence the name foulbrood [5]. Foulbrood was believed to manifest in two different pathologies: 'Mild and curable' in unsealed brood and 'malignant and incurable' in sealed brood. In 1885, *Bacillus alvei* was originally named and regarded as the etiological agent of foulbrood [6] (Figure 1.3). In 1906 was discovered that the responsible bacteria was *Bacillus larvae* [7], which is now classified as *Paenibacillus larvae* [2].

Infection starts with the oral uptake of the highly resistant endospores, which are spread in or between colonies by adult bees that carry them or by interventions of the beekeeper [8, 9]. These spores are the only infectious form of the pathogen [10]. Once ingested, they germinate in the midgut lumen and proliferate massively. The bacteria eventually breach the epithelium and invade the hemocoel [11]. Invasion of the hemocoel co-occurs with the dead of the larvae. *P. larvae* transforms these larval remains to a ropy mass filled with spores. This ropy mass is one of the most characteristic symptoms and a matchstick test is often used as a field test in diagnosis (Figure 1.2). Untreated, the disease causes the collapse of the entire colony since more and more larvae get infected and die, causing a lack of progeny [12, 13].
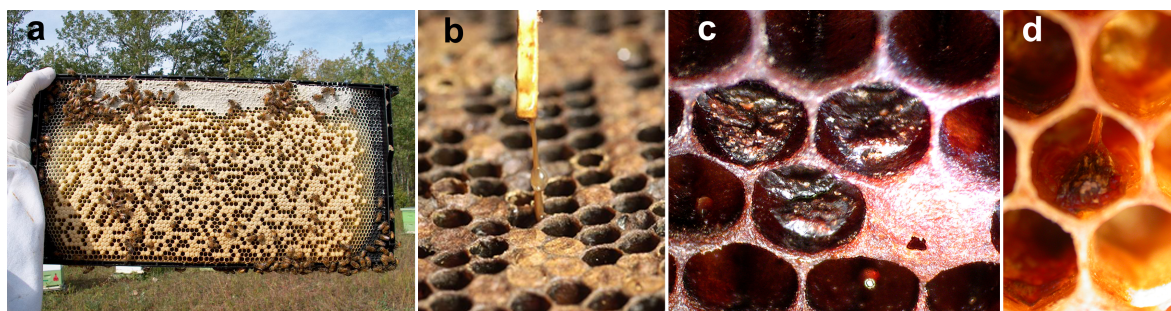


Figure 1.2: **Recognizing American Foulbrood** *Top*: American Foulbrood can be recognized in the field by some characteristic symptoms. From left to right: the hailshot pattern, where brood cells cleaned by the workers are present randomly among sealed brood, is characteristic for different honey bee diseases and should alert the beekeeper. Using the matchstick test the glue-like ropy remains can be pulled from the brood cell, this is the most characteristic symptom of AFB. The remains dry and form hard scales. Pupae dying of AFB protrude their tongue from the head, this symptom is rarely observed since most deaths occur during the late larval stage. (Figure from [14]). *Bottom*: Movie showing the different symptoms characteristic for American Foulbrood.

# 2 Taxonomy

## 2.1 Classification

In 1906 the American scientist White noted that there are two different bacterial brood diseases covered by the term foulbrood [7], being American Foulbrood (the malignant pathology) and European Foulbrood (the mild pathology). Using Kochs postulates, he proved that the infectious agent of American Foulbrood was *Bacillus larvae*. The generic assignment was based on the rod-shaped morphology of the vegetative form and the ability to form endospores. On the contrary, European Foulbrood is caused by *Melisococcus plutonius*, with *B. alvei* as a secondary invader [15, 16](Figure 1.3). Early research on European Foulbrood was mainly conducted by the British scientist Bailey. Both diseases have a worldwide prevalence, the terms American and European do not refer to the geographical origin of each disease, but to the geographical area where research on these pathologies was performed.



**c.a. 350 BC** — Probably earliest reporting of American Foulbrood by Aristotle

**1769** — Foulbrood: honey bee disease characterized by foul smell of brood

**1885** — Etiological agent identified as *Bacillus alvei*

**1906** — Foulbrood splits:
 - American Foulbrood (AFB) caused by *Bacillus larvae*
 - European Foulbrood (EFB) caused by *Melisococcus plutonium* (1983)

**1994** — Reclassification of *B. larvae and B. pulvifaciens* to the genus *Paenibacillus* (Ash *et al.,* 1993)

**1996** — *Paenibacillus larvae* and *Paenibacillus pulvifaciens* reclassified to the same species with differentiation at subspecies level
➜ *P. larvae* subsp. *larvae* and *P. larvae* subsp. *pulvifaciens*

**2006** — Both subspecies are united into one species and hence classified as *P. larvae*

Figure 1.3: **Overview American Foulbrood.**Historical and chronological overview of reporting American Foulbrood and the classification of its bacterial agent *Paenibacillus larvae.*

In 1950, 'powdery scale disease' was described as a new bacterial honey bee disease (Figure 1.3). It was characterized by the formation of powdery scales from the larval remains, in contrast to AFB which forms hard scales. The isolated bacterium resembled *B. larvae*. However, based on the differences in size of the vegetative cells, colony morphology and growth conditions it was considered a different species named *Bacillus pulvifaciens* [17].

The introduction of molecular methods caused different rounds of reclassification. Using 16S rRNA gene sequence analysis, the genus *Bacillus* was split up in different genera [18]. One of these was the genus *Paenibacillus*, to which both *B. larvae* and *B. pulvifaciens* were assigned. They were renamed to *Paenibacillus larvae* and *Paenibacillus pulvifaciens* respectively [19, 20] (Figure 1.3).
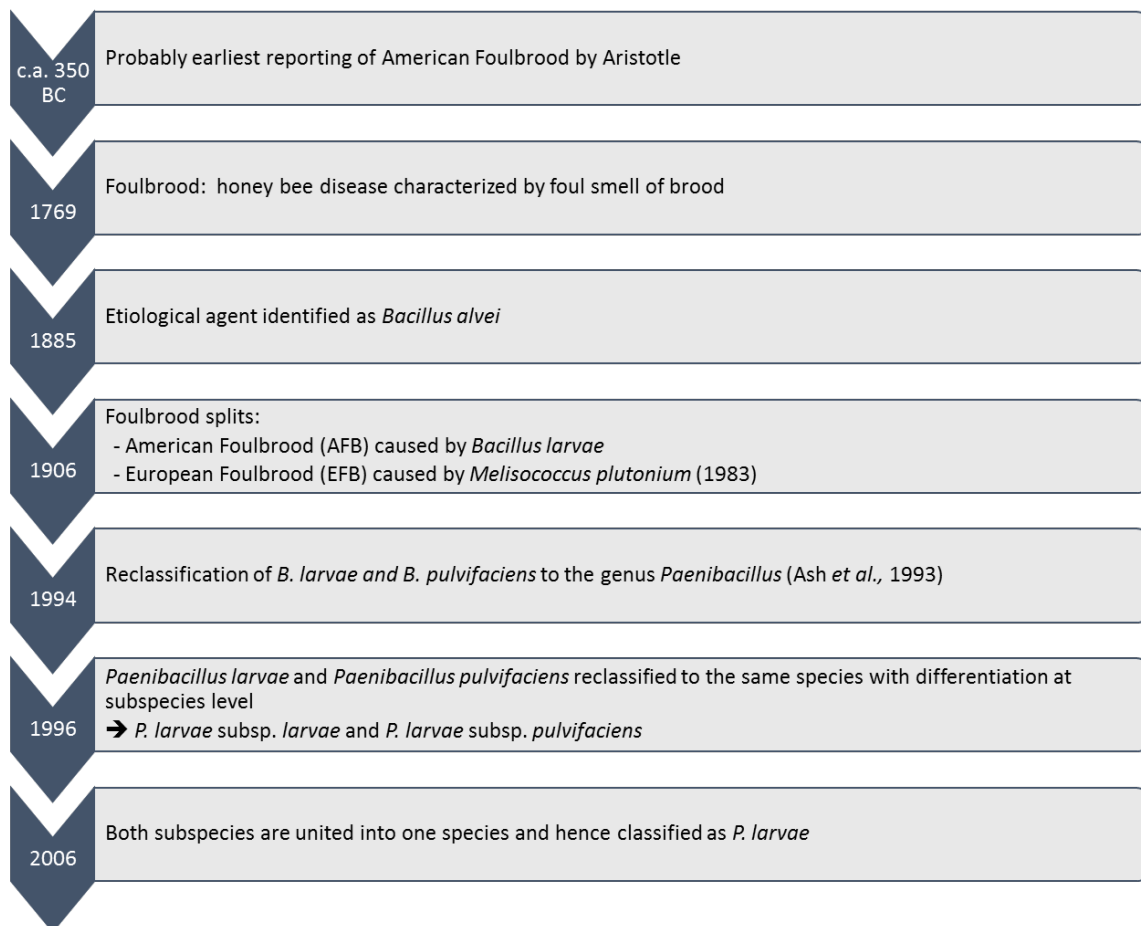
In 1996, both *P. larvae* and *P. pulvifaciens* were reclassified under one species *P. larvae* [21](Figure 1.3). This decision was based on the high similarity of the rDNA restriction patterns and DNA-DNA binding studies. A differentiation between the two species was still made at a subspecies level, resulting in the species *P. larvae* that contains two subspecies, namely *P. larvae* subsp. *larvae* and *P. larvae* subsp. *pulvifaciens*. This subspecies division was supported by the differences in colony morphology, biochemical features and pathology. *P. larvae* subsp. *pulvifaciens* showed a strong orange colony pigmentation, which was not found in the *P. larvae* subsp. *larvae* strains. However in 1994, 1 year before the validation of the generic reclassification [20], it was reported that orange-pigmented colonies were isolated from diseased brood showing symptoms of AFB [22]. In 2004, the orange pigmentation was discarded as a typical *P. larvae* subsp. *pulvifaciens* characteristic, since a genotype of *P. larvae* subsp. *larvae* was isolated which also showed this morphology [23]. Differentiation based on biochemical properties was discarded since the acidification of salicin and mannitol had been reported to be a variable property of *P. larvae* subsp. *larvae* field strains [24–26]. A further motivation for revision of the subspecies differentiation was the identical 16S rRNA gene sequences of the *P. larvae* subsp. *larvae* and *P. larvae* subsp. *pulvifaciens* type strains DSM 7030$^T$ and DSM 3615$^T$ [27].

The subspecies were subsequently discarded taxonomically united at the species level [2] (Figure 1.3). This at present final classification was the result of a polyphasic approach in which type, reference and field strains from Germany, Finland and Sweden were included. Among the field strains were pigmented and non-pigmented variants of both subspecies of *P. larvae*. Strain differentiation was shown by performing rep-PCR with Enterobacterial Repetitive Intergenic Consensus (ERIC) primers [28, 29]. This typing method revealed that each subspecies contained two different genotypes, resulting in a total of 4 ERIC-genotypes within *P. larvae* [2]. The results of SDS-PAGE and pulsed field-gel electrophoresis profiling strongly supported the weak subspecies differentiation as they have been published. However, there was still the issue of difference in pathology. To investigate this pathology, Genersch and colleagues performed exposure bioassays [2]. Both *P. larvae* subsp. *larvae* and *P. larvae* subsp. *pulvifaciens* proved pathogenic and caused the same disease symptoms namely transformation of the dead larvae to a ropy mass which dried to a hard scale. This result contrasted the earlier reports of the formation of a powdery scale by *P. larvae* subsp. *pulvifaciens*. Although no difference in pathology could be observed, there is a clear difference in virulence (section 3.4).

## 2.2 The genus *Paenibacillus*

The name of the genus expresses its close relationship to *Bacillus*, since *paene* is Latin for almost. The creation of the genus was done in 1993 [19] and *P. polymyxa* was assigned as the type species [30]. Most species that were reclassified into the genus were former *Bacillus* species, although some species belonged to the genus *Clostridium* [31, 32].

Nowadays, genus *Paenibacillus* harbors a wide variety of bacteria, with the majority of them belonging to non-pathogenic soil bacteria. Of these species, an important part are known to be associated with plant roots promoting growth. Among these so-called rhizobacteria interactions with economically important plants as maize [33], pumpkin [34] and rice [35] are observed. The bacteria can promote plant growth directly, by expression of auxine phytohormones [36, 37], augmenting the accessibility of phosphorous [38] or nitrogen fixation [39]. Indirectly, *Paenibacillus* species can express biocidal substances and can trigger the induced systemic resistance in plants resulting in a higher resistance to phytopathogens including bacteria [40], fungi [41], nematodes [42] and viruses [43].

Besides *P. larvae* two other insect pathogens within the genus are known. *P. popillae* and *P. lentimorbus* are the causative agents of milky disease. They have a broader host range than *P. larvae* and kill larvae of beetles [44] and lepidopterans [45], but are not known to affect honey bees. Some important pests are susceptible for *P. popillae* with its application as a microbial control agent in the US [46]. The species expresses chitinases and crystal (Cry) proteins as virulence factors [47]. Biotechnological engineering with *Paenibacillus* strains is currently targeting to combine the insecticidal activities of Cry proteins with the plant growth promoting activities of *P. polymyxa*. In this context it is important to note that Cry proteins have not been identified in *P. larvae* although they are one of the best known insecticidal compounds.

Many *Paenibacillus* species have to compete with other organisms to colonize their niche, a wide variety of antibiotics and fungicides are expressed by members of this genus and several of these compounds have a broad host range [48]. Between the different members of the genus variety in expression of these antimicrobial compounds is present, but this variety is also observed *intra* specific. Strains of the type species *P. polymyxa* have a diversity of antimicrobial gene clusters resulting in a different response towards the plant pathogen *Phytophthora capsici* [38, 49, 50]. These antimicrobial compounds of are often explored in the context of biocontrol against plant pathogens, however they are also considered for use against human pathogens or as food preservatives [51]. *P. polymyxa* compounds have been explored for their medical usage, their exo-polysaccharides (EPS) have anti-oxidant and anti-tumour properties [52, 53] and one bacteriocin, polymyxin B, has been produced on an industrial scale as the active compound in ointments and antibiotic creams [54]. Nonmedical applications of *Paenibacillus* are explored for bioremediation (*P. jamilae, P. macerans, P. polymyxa, P. validus* and *P. elgii*) and the laundry industry (*P. terrae, P. tarimensis*) [51].

## 2.3   Typing of *P. larvae* strains

Typing concerns differentiation at the infra (sub)specific level i.e. strain differentiation. It can be divided in phenotyping and genotyping. Phenotyping techniques may include cell and colony morphology studies, biochemical profiling and protein/proteome studies. Genotyping techniques are based on the genomic profiles. Among all techniques Rep-PCR using ERIC-primers is routinely used for typing in *Paenibacillus larvae*.

### 2.3.1   Phenotyping in *Paenibacillus*: a short overview

As described earlier orange colony pigmentation has long been considered a specific trait to assign strains to the former *P. larvae* subsp. *pulvifaciens*. The isolation of a *P. larvae* subsp. *larvae* strains with orange pigmentation showed its weakness for strain allocation applications [23].

In sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) of whole cell proteins the protein fingerprints are used for strain differentiation and contributed as such to reclassification of *P. larvae* and *P. pulvifaciens* as subspecies of *P. larvae* [21]. The method separates proteins based on their molecular weight. SDS is added to the proteins, which causes protein denaturation of the negatively charged proteins.

Pyrolysis mass spectrometry contributed to the reclassification of of *P. larvae* and *P. pulvifaciens* as subspecies of *P. larvae* [21]. For this technique colonies are smeared onto iron-nickel foils [55]. The samples are oven-dried and transferred to a pyrolysis mass spectrometer. The pyrolysate is generated in a vacuum by heating the sample and enters a expansion chamber. It diffuses down a molecular beam tube to the ionization chamber. In this chamber the pyrolysate is ionized, the ionized fragments are subsequently accelerated and directed into a quadrupole mass filter. They get separated based on their mass-to-charge ratio and thus separated based on their molecular weight.

Matrix-assisted laser desorption-ionization time-of-flight mass spectrometry (MALDI-TOF MS) is another phenotypic method that can be considered as the modern version of SDS-PAGE and uses the differences in the proteome for subtyping. In this high-throughput method the bacterial proteins are extracted from single colonies. The pellet is resuspended in 70% formic acid and mixed with acetonitrile. Drying of this suspension results in the crystalization of the sample and matrix. After centrifugation the supernatans can be spotted on a MALDI target plate. The sample air-dried, overlaid with a saturation solution and dried again at ambient temperature. The sample is subsequently ionized by the laser beam and the charged ions are sent to a flight tube with a detector at its end. The time-of-flight (TOF) of these ions is correlated to their weight by their charge, which results in a separation based on their molecular weight (Figure 1.4). In 2014 this technique has been implemented for *P. larvae* [56] and is thus one of the most recent typing techniques. With MALDI-TOF MS strains can be clustered within their routinely used ERIC-type. Within the ERIC clusters, clusters could be identified which contained closely related strains. This suggests the usability of the method for epidemiological studies or source tracking, but this has not been tested yet.
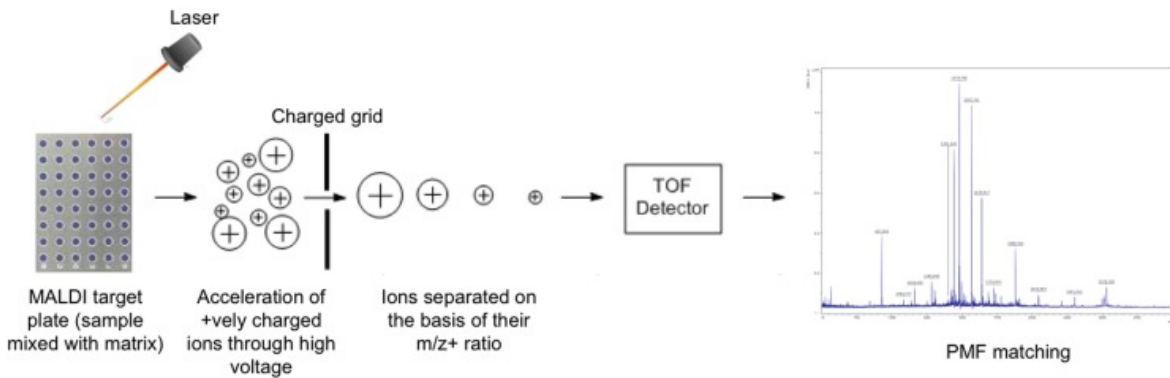
Figure 1.4: **MALDI-TOF MS principle.** The laser ionizes the crystalized sample and matrix. The positively charged ions are accelerated and seperated by their size in a flight tube with a TOF detector at the end. This results in a unique peptide mass fingerprint (PMF) for each type. Figure from [57].

### 2.3.2 Genotyping

A first molecular technique, Restriction Endonuclease Fragmentation Patterning (REFP) [58], was used in 1994 to type *Bacillus larvae* strains. Restriction endonucleases are used to digest the genomic DNA and comparative analysis of the separated fragment on a polyacrylamidegel is peformed. Based on *Cfo*I restriction patterning, five clonal formerly named *B. larvae* types among 20 geographically diverse isolates were differentiated.

Restriction fragment length polymorphism (RFLP) analysis is a second restriction enzyme based technique that has been used to genotype *P. larvae* [59]. The difference between this technique and REFP is that in RFLP an amplicon is subjected to enzymatic restriction instead of whole genomic DNA. In case of *P. larvae* amplified 16S rRNA genes were subjected to*Hae*III digestion. This allowed in 2002 to distinguish between *P. larvae* subsp. *larvae* and *P. larvae* subsp. *pulvifaciens*.

Amplified fragment length polymorphism (AFLP), another genotyping technique that combines PCR and restriction enzymes, has been used in several studies of *P. larvae* [21, 26]. In this genotyping technique restricted genomic DNA with *Eco*RI and *Mse*I is used as template for selective amplification. Adaptors are ligated to the sticky ends and the obtained fragments are then amplified with selected extended primers (for bacterial genome the extension is usually one nucleotide) e.g. C or A. The amplicons can be separated using e.g. capillary sequencer. This method may reveal strain specific AFLP fingerprints that can be comparatively analyzed via suited software.

In the most recent reclassification step, pulsed field gel electrophoresis (PFGE) typing [2] was used in support to unite the subspecies into one single species (*P. larvae*). Previously, PFGS had been used applied to classify 44 *P. larvae* isolates in twelve distinct PFGE groups [60]. Therefore, whole genome DNA (almost intact) is captured in agarose plugs. The DNA is then digested with a restriction enzyme, for *P. larvae Not*I has been used. Subsequently, the large fragments are separated over the gel with switching directions of the electric field.

Another PCR-based method is PCR-denaturating gradient gel electrophoresis (PCR-DGGE) in which an amplicon is separated over a polyacrylamide gel with a denaturating gradient.

The technique is based on the difference in denaturating abillity due to the small differences in the base pair sequence. A 436 bp fragment of the 16S rRNA gene was amplified in *P. larvae*, the denaturating gel had a gradient of urea and formamide from 45% to 70% [61]. The technique allowed to type the isolates according to their former subspecies classification.

In 1998, repetitive element PCR fingerprinting using BOX primers identified three subtypes [62]. Six years later, it was demonstrated that rep-PCR with Enterobacterial Repetitive Intergenic Consensus (ERIC) primers was able to differentiate the former subspecies of *P. larvae*. Consequently, the technique was used as a subspecies specific PCR protocol [59]. Gensch and Otten typed isolates with rep-PCR using BOX A1R and MBO REP1 primers [63] and revealed the existence of at least four genotypes (*AB, Ab, ab and αB*).

In the last round of reclassification, the rep-PCR using ERIC primers was used. Although this method did not reveal decisive differentiation as applied in 1998 [64], in the study of Gensch and colleagues the former subspecies of *P. larvae* could be divided into two genotypes each (Figure 1.5). Until now, ERIC-PCR is still the most important subtyping technique for *P. larvae* strains. The reason for this is that the ERIC-types coincide with virulence characteristics, being thus biological meaningful (section 3.2).

It has long been assumed that the geographical distribution of ERIC II strains was restricted to Europe, while ERIC I strains were present in worldwide. However, recently ERIC II strains have been isolated in Asia, North America and Australia, showing the worldwide distribution of this genotype [65]. ERIC III and IV strains have not been identified in the field in decades and are only available as historical isolates in culture collections [13].
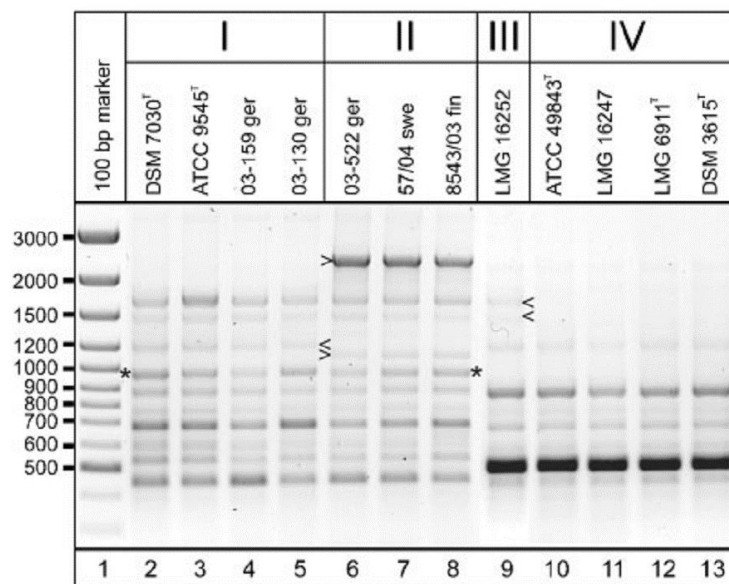


Figure 1.5: **ERIC-typing of *Paenibacillus larvae* strains**. Using ERIC-primers, four subtypes can be found. Two of these (ERIC I and II) belong to the former *P.larvae* subsp. *larvae*, the other two (ERIC III and IV) belong to the former subspecies *P. larvae* subsp. *pulvifaciens*. The asterix indicates specific bands for ERIC I and II. The arrowheads indicate specific bands for differentiating the ERIC-genotypes. Figure from [2].

Recently, the sequencing based method Multilocus Sequence typing (MLST) was implemented for typing *P. larvae* [65]. In MLST typing the sequence of short regions of six or seven household genes are compared. At each given locus, each found allele is allocated a unique number. Combining the numbers gives the MLST scheme (Figure 1.6). For the typing of *P. larvae* seven household genes are targeted (*clpC, ftsA, glpF, glpT, Natrans, rpoB* and *sigF*). The obtained scheme was able to cluster the strains to their ERIC-group and to extend these ERIC genotypes by identifiyng several types within them. In total 21 different types are found using 294 isolates.

In this thesis, the implementation of Multiple Locus Variable number of tandem repeats Analysis (MLVA) as a genotyping tool for *P. larvae* is described (Part II, chapter 3).
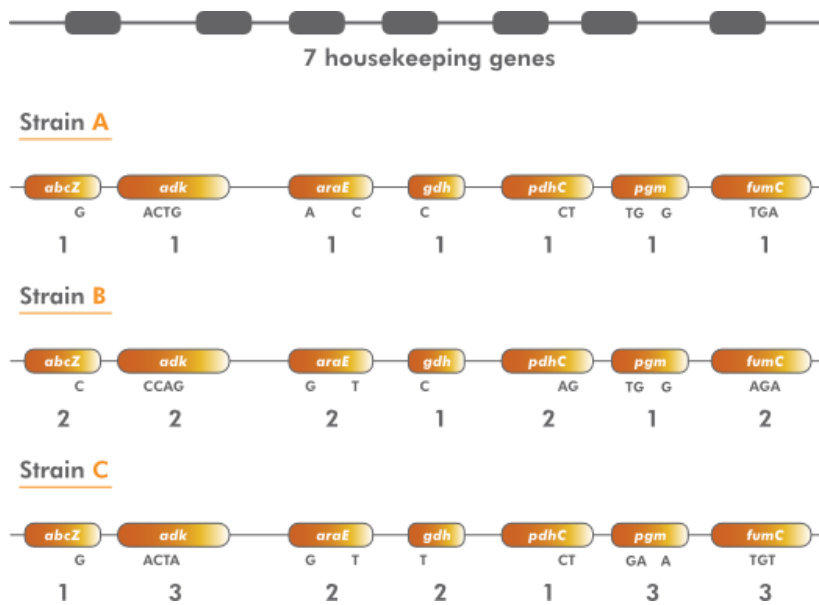


Figure 1.6: **MLST principle.** Different strains have differences in the sequences of their housekeeping genes. In MLST the loci of a set of housekeeping genes is compared, different alleles get a unique number. Figure from [66].

# 3 Pathology

It was mentioned that *Paenibacillus* contains pathogenic as well as non-pathogenic species. *P. larvae* is a pathogenic species, which means that all known strains of this species are able to cause disease. But important differences in virulence have been observed amongst strains. In contrast to pathogenicity, virulence is a quantitative character. Virulence does not only depend on the micro-organism, but also on the host. This host-pathogen interaction is influenced by the immunological response of the host [67, 68].

## 3.1 Definitions

**Pathogenicity**: this is a qualitative trait, referring to the inherent, genetic capacity of a microorganism to cause disease, mediated by specific virulence factors. Whether or not the organism becomes pathogenic, is the result of the specific host-pathogen interactions.

**Virulence**: this is a quantitative trait, representing the extent of the pathology caused by a microorganism. Virulence expresses the interaction between a pathogen and its host. In view of the significant influence of the immunological condition of the host, the definition of virulence has been re-adressed recently[67, 68].

Virulence is usually correlated to the pathogen's capacity to multiply in the host [67], for example by the growth rate of the bacteria. It can also be affected by host and environmental factors, such as the transmission route or life-stage infected. For example, a pathogen may be virulent when infecting one life-stage and non-virulent when infecting another one [68]. Virulence is therefore dependent on the nature of the infection.

Virulence is also a relative trait, referring to the differences in the degree of pathology caused by strains of the same pathogen, or differences in the efficiency in which different strains can cause symptoms [69]. A pathogen strain that requires a low dose to produce disease symptoms would be more virulent than a strain that requires a high dose to produce the same symptoms.

Since virulence is a quantitative measure, methods have been developed to quantify the relative contributions of different virulence factors to a phenotype [70].

## 3.2 Virulence

As stated in the previous section, the ERIC-genotypes differ in virulence. ERIC-type I needs the longest time to kill its larval host with an $Lt_{50}$ (time to kill 50% of the infected larvae) of 5 days and $Lt_{100}$ of 12 days. ERIC-type II has an $Lt_{50}$ of 4 and $Lt_{100}$ of 7 days, while ERIC-type III-IV have a $Lt_{50}$ of 2 days and a $Lt_{100}$ of 7 days [2, 71] (Figure 1.7).

Because of the social nature of honey bees, a negative correlation between the individual insect-level virulence and the colony-level virulence is present [72]. Larval brood cells are open in the larval stages, they are capped after 6 days when metamorphosis to the pupal stage starts. This means that for ERICII-IV almost all larvae are killed before capping. The vast majority of the infected larvae are recognized and removed by the nursing bees before the degradation to the ropy stage or scale formation. For brood killed by ERIC I strains, this is not the case. Most infected individuals die after capping, which makes it
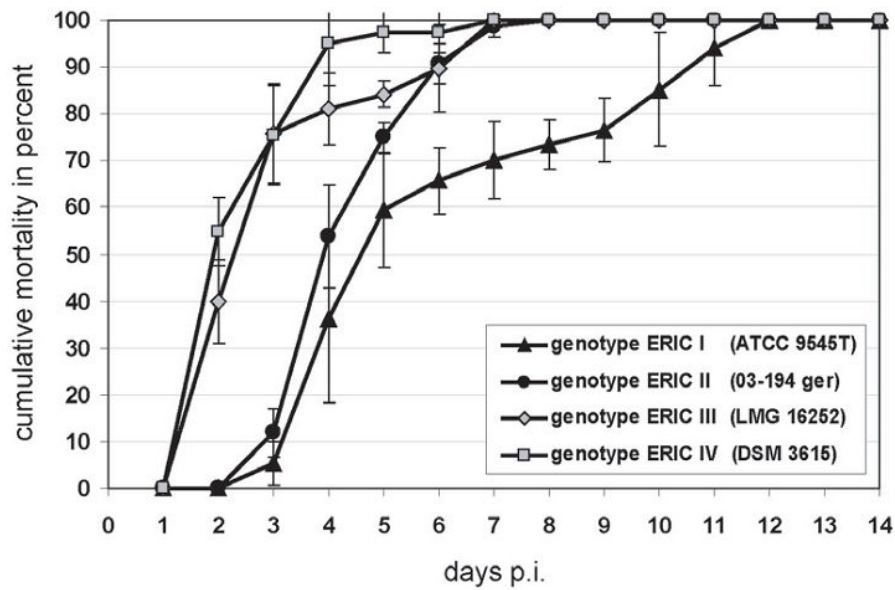
Figure 1.7: **Virulence of different ERIC-genotypes.** Honey bee larvae have been infected with the a dose representing the $LC_{50}$ of a ERIC-strain. The cumulative mortality and associated standard deviation is plotted in function of time post infection (p.i.). Figure from [4].

more difficult to be recognized by the adult bees. The brood can be transformed to the ropy stage and eventually scale formation. When hygienic bees detect these individuals, they get covered in spores when removing the scales, resulting in a more efficient spread of the disease in the colony. Thus although ERIC I is less virulent on the larval level, due to the late detection by the nursing bees it is more virulent on the colony level. This negative relationship between larval and colony level virulence might be the reason why ERIC III and ERIC IV strains are not detected in the field. These extreme fast-killing strains might allow a too-efficient hygienic response and thereby impair their own transmission [72].

It has to be noted that although *P. larvae* outbreaks are rare in Belgium, a screening revealed that 10 % of the Belgian apiaries are infected with *P. larvae* spores [73]. It is currently unknown what causes this discrepancy.

## 3.3 Transmission

Transmission of AFB happens on two levels, within-colony and between-colony. The between-colony transmission can again be subdivided in vertical and horizontal transmission [13]. Horizontal transmission occurs when healthy colonies rob diseased ones or when drifting bees carrying spores are taken up by the healthy colony. Vertical transmission occurs by swarming of infected colonies. Between-colony transmission is also often due to the beekeeper using contaminated material.

## 3.4 Infection process

Several important steps triggered the unraveling of the infection process and the involved virulence factors. The ERIC-genotyping allowed comparative genomics and proteomics between ERIC I and ERIC II, which differ in their virulence [74, 75]. Another step was the sequencing of the genomes of *P. larvae* ERIC I strains BRL 230010 and DSM 25719 and ERIC II strain DSM 25430 [76–78], which allowed the investigation of the genomic potential of the bacterium (section 3.5). A third step, the successful electroporation and transformation of *Paenibacillus larvae* [79, 80] allowed the implementation of site-specific mutagenesis using the TargeTron system (Sigma-Aldrich) by which several virulence factors were functionally confirmed (Table 1.1)[3, 81–84].

Table 1.1: **Confirmed virulence factors.** Although several possible virulence factors have been identified in the genome, only a few have been functionally confirmed. The table gives an overview of all confirmed virulence factors, their role and to which ERIC-type they are associated.

| Factor | Function | ERIC type | reference |
| --- | --- | --- | --- |
| Paenilamicin | Antimicrobial component eliminating bacterial competitors in the gut. | ERIC II | [84] |
| *Pl*CBP49 | Chitin-degrading enzyme necessary for disrupting the peritrophic matrix. | ERIC I and ERIC II | [81] |
| Plx1 | AB toxin involved in attack of the midgut epithelium. | ERIC I | [82] |
| Plx2 | AB toxin involved in attack of the midgut epithelium. | ERIC I | [82] |
| SplA | S-layer protein mediating adhesion to midgut epithelium cells. | ERIC II | [3] |

In 2008 the infection was visualized over time using the Fluorescence in Situ Hybridization (FISH) assay [11] where a 16S rRNA-targeted oligonucleotide probe was used to follow the infection route of *P. larvae* ERIC I strains ATCC 9545 and field isolate 01-440 and ERIC II field isolates 03-200 and 04-309. When ingested, the *P. larvae* spores start to germinate within 24h post infection. The germination is not place-specific in the midgut, since vegetative bacteria were found in the anterior as well as in the posterior part. After the germination stage, massive proliferation can be observed (Figure 1.9). In order to do this, *P. larvae* probably gets rid of the bacterial competitors in the gut by expressing antibacterial compounds [85], of which the group of non-ribosomal peptide antibiotics has been investigated well [83, 84, 86–88] (Figure 1.8). Paenilamicin has been confirmed as an antimicrobial compound and virulence factor of *P. larvae* ERIC II strain DSM 25430. During this phase no alterations of the larval epithelium is observed. In general, more bacteria seemed present in larvae infected with ERIC II field isolates compared to the ERIC I strains.
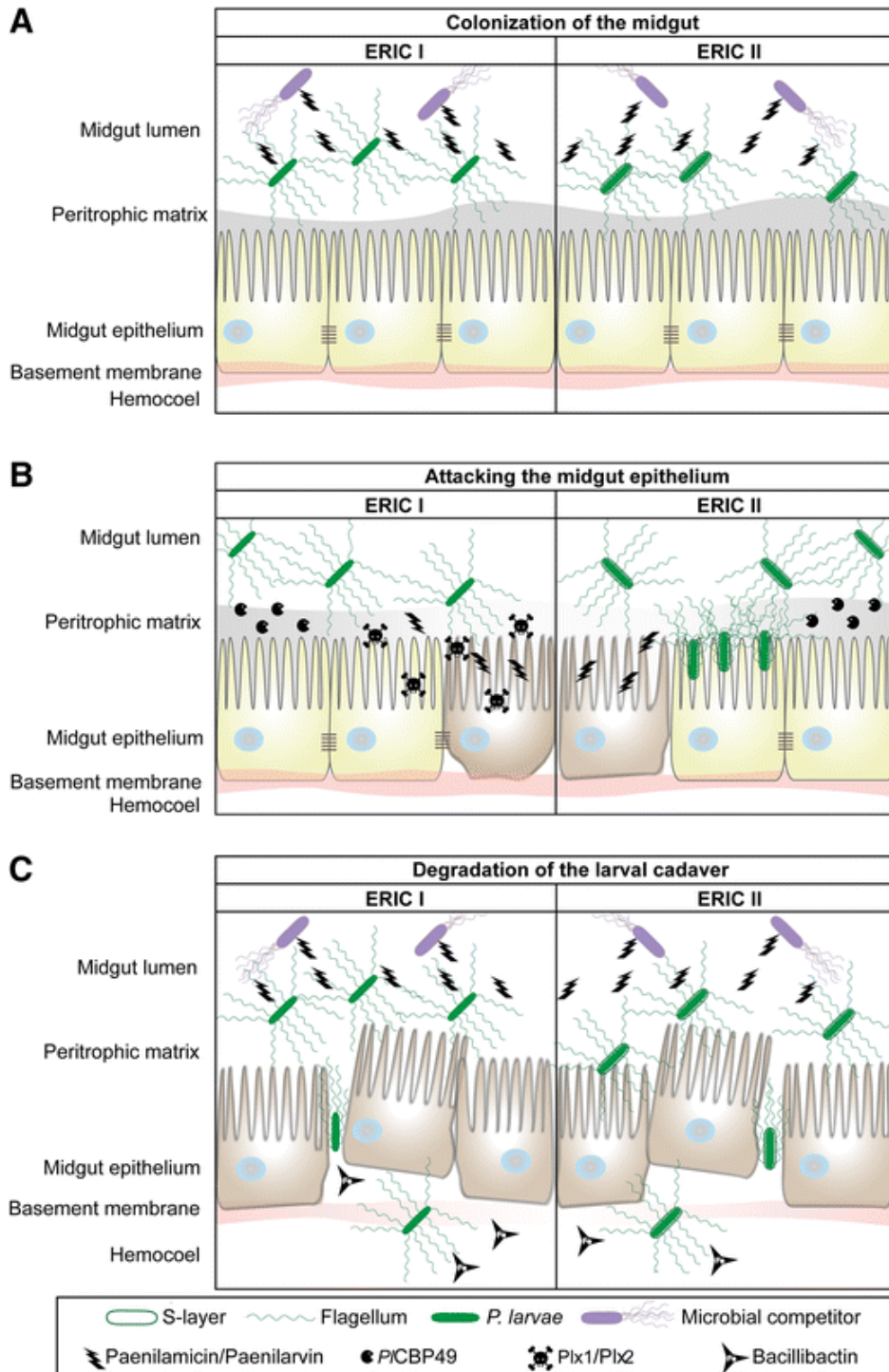
Figure 1.8: **Proposed virulence route of *P. larvae* ERIC I and ERIC II strains**. A: The first stage of infection is characterized by the massive proliferation. To colonize the midgut, antibiotics as Paenilamicins and Paenilarvins are expressed. B: The next stage is the degradation of the peritrophic membrane using *Pl*CBP49 and the ERIC-type specific attack of the epithelium. C: The paracellular route is followed when breaching the epithelium. Colonization by secundary invaders is prevented by expressing antimicrobial compounds as Bacillibactin. Figure from [1].

13
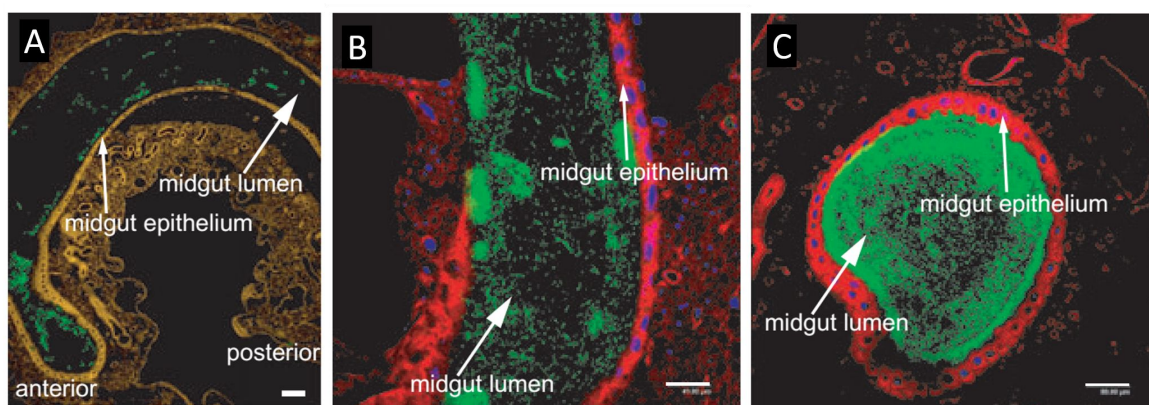
Figure 1.9: **Proliferation stage of *P. larvae*.**The FISH analysis used a green fluorescent probe to visualize *P. larvae* and a yellow (A) or red (B and C) fluorescent probe to visualize eukaryotic cells. The bars represent $40\mu$m (B) or $80\mu$m (A and C). A and B: Longitudinal section of a live larva 2 days p.i. C: A cross-section of a live larva 3 days p.i.. The figures show the massive proliferation of *P. larvae*. Figure from [11].

The midgut epithelium is protected by the chitine rich peritrophic matrix, which has to be broken down by the bacteria to reach the epithelium. Chitin-degrading enzymes are necessary for this, the functional role of one chitin-degrading enzyme *Pl*CBP49 has already been confirmed as a key virulence factor [81] (Figure 1.8).

In 1973 and 1998 electron microscopy studies suggested that the midgut lumen was penetrated by phagocytosis [89, 90]. The FISH analysis, however, revealed that the penetration route is paracellular, by destroying cell-cell junctions [11] (Figure 1.10). *P. larvae* was detected between the epithelial and the smooth muscle cell layer, which suggests that the bacteria are able to degrade the basement membrane and/or interfere with the cell matrix adhesion. Once arrived in the hemocoel, the bacteria are distributed intercellular, never intracellularly. *P. larvae* attacks the midgut lumen randomly, but at discrete points along the entire organ, although some preference for the stomodeal valve and the valve ring was observed. The stomodeal valve is the posterior part of the foregut which projects into the midgut, while the valve ring is the band of midgut cells encircling the stomodeal valve. The invasion of the hemacoel seems to be a sudden process, in which the bacterial mass poures into the hemacoel through local penetrations. Toxins are used to attack the epithelium (Figure 1.8) as confirmed in ERIC I strain ATCC 9545. In ERIC II strain DSM 25430, the functional role of a S-layer protein has been confirmed [3]. The protein is involved in the attachment of *P. larvae* to the midgut epithelium (Figure 1.8).

It was long assumed that sporulation was the last step in the infection process [89, 91]. However, using phase-contrast microscopy [11] it was found that sporulation is already present at early times of infection. This is similar to some members of the closely related *Bacillus* spp., where sporulation is observed during the entire infection process [92, 93]. To protect colonization of the dead larva by secondary invaders, *P. larvae* expresses antibacterial compounds to protect this niche [1] (Figure 1.8).
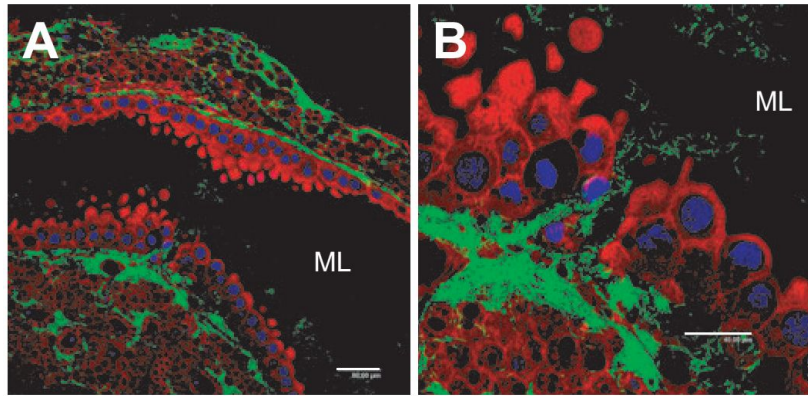
Figure 1.10: **Breaching of the gut epithelium by *P. larvae***The FISH analysis used a green fluorescent probe to visualize *P. larvae* and red fluorescent probe to visualize eukaryotic cells, nuclei are stained blue. The bars represent $80\mu$m. A: Section of a dead larva 4 days p.i.. Penetration of the epithelium happens at random locations. B: Magnification of A. *P. larvae* follows the paracellular route, it is detected between cells but not within cells. Figure from [11].

## 3.5   Genomic potential

### 3.5.1   Sequencing and processing

The genome sequencing of *P. larvae* meant an important step in understanding the pathogenic and metabolic potential of the bacterium. A first sequencing was performed on ERIC I strain BRL 230010 in 2006 by whole genome shotgun sequencing, however at a low coverage (5-6x) [77]. It was re-sequenced to a coverage of 182x in 2011 [76]. *De novo* assembly was performed, but was relatively fragmented. The final synteny of the contigs was performed based on the genome of *P. JDR2*, which shares 93% identity at 16S rDNA loci. Despite the difference in assembly size (the genome of *P. larvae* is 4.4 Mbp, while that of *P. JDR2* is 7.2 Mbp), the gene-level synteny was conserved. Some regions between both genomes were divergent, these regions of *P. larvae* might harbor species-specific genes which could be important for survival in its unique niche the honey bee larvae. Within these regions, an ORF with homology to synthesases of plipastatin (an antibiotic inhibiting phospholipase A2) and the hemolysin surfactin was detected [76]. Another ORF encodes homologs of a putative iron-siderophore ABC transporter, in many bacterial infections iron uptake is an important process.

The sequencing of an ERIC I (DSM 25719) and an ERIC II (DSM 25430) strain by Djukic and colleagues in 2014 was an important addition in understanding the genomic potential of the species [78]. The study was able to close the ERIC II DSM 25430 genome, which has a size of 4,57 Mbp. The genome of the ERIC I strain DSM 25719 could still not be closed and is now available in 7 contigs. The failure to close this latter genome is due to the presence of long genomic repeats [76, 78]. In both strains, a plasmid of 9.7 kb is present. These plasmids, called pPLA1_10 and pPLA2_10 differ in only 49 bp. It was known since 2004 that ERIC II strains harbored a plasmid [23], but this was never found in ERIC I strains. The genome sequencing study proved that not all ERIC I strains contain a plasmid, which means that the found plasmid was strain-specific but not genotype-specific [78].

The annotation step was performed using GLIMMER and BLAST searches against *Bacillus* and *Streptococcus* identifying 3568 gene models. Of these models, 35% were confirmed with shotgun proteomics [76]. The different classes will be described in the following subsections.

### 3.5.2   Metabolism

Since *P. larvae* grows preferably in aerobic conditions, it is not surprising to find genes involved in the oxygen-dependent respiration in both DSM25719 and DSM25430 [78]. Several strains of the species are able to use nitrate as an alternative electron acceptor in anaerobic conditions, including the DSM 25719 and DSM 25430. Several genes are found to encode for a putative respiratory nitrate reductase, but the complete general nitrite reductase genes could only be found in the ERIC II genome [78].

All sequenced *P. larvae* strains contain genes of the Embden-Meyerhof-Parnas (EMP) and the oxidative pentose-phosphate pathway [78]. Using these pathways, D-glucose and D-fructose can be metabolized. These simple sugars are the dominant sugars in honey, which is fed to larvae in combination with royal jelly and pollen. The ability to metabolize these sugars suggests that the bacterium lives on the incoming larval diet when it proliferates in the midgut during the early stage of infection. The main sugar in hemolymph is the disaccharide trehalose. It can be metabolized by *P. larvae* when the pathogen breaches the epithelium since it contains a putative trehalose specific II$^c$ component of a phosphotransferase system (PTS).

### 3.5.3   Flagellar system

*P. larvae* is a motile bacterium due to the presence of peritrichous flagellae [13]. Comparison of the genome of BRL 230010 with other flagellated bacteria as *Bacillus subtilis* and *Escherichia coli* led to the identification of almost all orthologs for genes involved in the flagellar system [76] (Figure 1.11).

Of the 41 associated genes, several are coding for motor/switch, flagellar rings, rod hook and filament proteins. Others are coding for proteins involved in the regulation or chaperones. However, the orthologs coding for the flagellar hook formation are missing, which is a structure that acts as a joint as well as motor for each individual flagellum. The missing orthologs are FlgD, a monomer scaffolding protein and FliK, which acts as a checkpoint for the correct length of the flagellae. However, this does not necessarily mean that *P. larvae* BRL 230010 developed a unique way to proceed with hook assembly, it could just be the result of the detection method used. As with *B. subtilis* and *E. coli*, the assembly and regulatory components for the base of the flagellum are coded by *Fli* and *Mot* gene families [76].
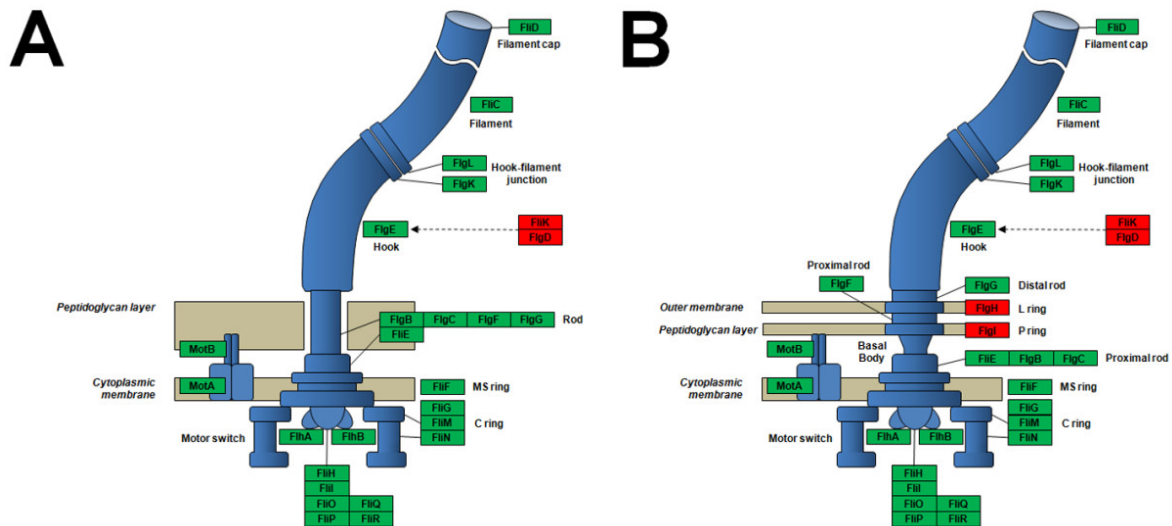
Figure 1.11: **The flagelar system.** Comparison of flagellar system of gram positive (A) and gram negative (B) bacteria. Proteins encoded by *P. larvae* are highlighted in green. Figure from [76].

### 3.5.4   Toxins

Since *P. larvae* is a pathogen, it is not surprising to find orthologues of known toxins in the genomes of all sequenced strains. Toxins can be used to colonize the host and escape the immune system [78].

The genome sequencing of BRL 230010 was able to identify a domain of the *Clostridium perfringens* epsilon toxin [76], this toxin is known to form pores on host cells, leading to cell death. A more thorough sequencing in 2014 of another ERIC I strain DSM 25719 identified the presence of this domain in 4 toxin encoding loci [78]. These loci are gene homologues of the clostridial epsilon toxin ETX/ *Bacillus* mosquitocidal toxin MTX2. More precisely, the amino acid sequences revealed 31% identity to the ETX toxin of *C. perfringens* and 29% identity to the *Lactobacillus sphaericus* toxin Mtx2. The functionality of these toxins is however still unclear. A homologue of this Mtx2 toxin has also been identified in ERIC II strain DSM 25430, it is however unlikely to be functional because of a transposon insertion destroying the start codon [78].

All sequenced *P. larvae* strains possess the classical binary toxins, in which the first part is the membrane component that makes the host cell permeable to the enzymatic component. The existence of these so-called AB toxins was already suggested before evidence in the genome was found [11]. Putative AB toxin gene fragments have subsequently been identified in ERIC I strain ATCC 9545 using subtractive suppression hybridization [74]. The sequencing of Djukic in 2014 identified seven toxin encoding loci in the genome of ERIC I DSM 25719 [78], of which five were coded for potentially functional toxins (Plx1-5). Two of these toxins have been confirmed as virulence factors (section 3.4) in ATCC 9545. Plx4 and Plx5 were identified as coding for binary AB toxins, with separate ORFs for the A and B domain. The B domains are located upstream of the A domains. The A domains showed similarity to ADP ribosyltransferases encoded by *Bacillus* spp., while the B domains show similarity to the clostridial binary $C^2$ toxin B domain, which is the membrane domain

[76, 78]. The unique combination of these two domains suggests that *P. larvae* developed species-specific toxins [78].

The observed homology with several genes of *C. botulinum* suggests horizontal gene transfer. This food-borne pathogen is frequently isolated from wax and honey where it comes into contact with *P. larvae*. However, a big part of the found loci encode for non-functional toxins. This lack of function is the result of mutations splitting the gene in several ORFs, introducing stop-codons or interruptions by transposon insertions [78]. A possible explanation is that *C. botulinum* is a vertebrate pathogen and his toxins can be harmless for an insect host. Since expressing these toxins would not result in an increase in fitness for *P. larvae*, there is no evolutionary pressure to correct mutations in these loci.

Toxin-encoding loci were also identified in the ERIC II strain DSM 25430. However, they all seemed to code for an non-functional toxin [78]. Again mutations resulting in this lack of function are due to insertions of transposons and deletions of gene fragments. This lack of functional toxins is surprising, since ERIC II is known to be more virulent on the larval level than ERIC I. The genotype must have different strategies to kill the larval host.

### 3.5.5 Hemolysins

Once the bacterium breaches the gut epithelium, it circulates in the hemolymph. All sequenced *P. larvae* strains have orthologues of hemolysins [76, 78], these proteins are able to lyse the insect blood cells (hemocytes) either by disrupting the host cell membranes or by pore formation. In ERIC I strains BRL 230010 and DSM 25719 four hemolysin domains were predicted, in contrast to two domains predicted in the ERIC II strain DSM 25430. The found hemolysins can be divided into two classes based on their similarity to known proteins. The first class comprises three hemolysins of ERIC I and one of ERIC II, showing similarity with the hemolysin III-related protein family. Hemolysin III is an important virulence factor in *Bacillus cereus*, which lyses the target cell by pore-formation [94]. The other class clusters within the pore-forming TlyA protein family.

An important difficulty for bacterial pathogens is acquiring sufficient amounts of iron [78]. Since iron is restricted in the extracellular fluid, a tactic of pathogens is extracting this metal from host proteins. By lysing host cells, these proteins become available for secreted siderophores. These iron chelators are typically produced and secreted upon iron deficiency. Two ABC transport systems are encoded in ERIC I DSM 25719, while only one is encoded in ERIC II DSM 25430 [78]. One of these shows similarity to the *isd* system of *Staphylococci*, which binds hemoproteins, removes the heme molecule and transports it into the bacterial cytoplasm. Besides these ABC transport systems, both strains DSM 25719 and DSM 25430 contain genes encoding for a ferrous iron transport cluster [78].

### 3.5.6 Proteases

Proteases have already been detected and identified in larval remains in 1997 [95]. Proteases are however not only important to break down the larval remains, but are also known to be virulence factors [96]. A total of 159 full or truncated proteases have been identified in the ERIC I genome of DSM 25719, and 128 in the ERIC II genome of DSM 25430 [78].

The penetration through the gut epithelium follows the paracellular route, this means the active breakdown of the cell-cell junctions. In order to do this, the bacterium has to excrete proteases. An active search for genes encoding these protein domains identified several possible virulence factors. Among these is the *Clp* protease complex, which has been linked to virulence factor production in other gram positive bacteria [97]. Other identified proteins may be able to cleave the transmembrane regions of substrate proteins. This might be of importance since tight junctions between epithelial cells are composed of claudin and occludin proteins, which have transmembrane domains [98].

The extracellular matrix has collagen as its main structural component. Collagenases, hyaluronidases and proteases are secreted by pathogens to break down this barrier [99]. Loci of two different collagenase encoding gene families have been identified in all sequenced strains [78].

Other important components of the extracellular matrix are glycosaminoglycans. In order to break these down, *P. larvae* could use enzymes belonging to the polysaccharide lyase family 8. A putative member of this family has been identified in both genomes. This group comprises a variety of secreted bacterial lyase enzymes f.e. hyaluronidases, which are able to degrade hyaluronan, chondroitin and chondroitin sulfates [100].

During infection *P. larvae* has to break down the peritrophic matrix in order to reach the gut epithelium (section 3.4). This matrix is a chitin and glycoprotein layer that borders the larval midgut, protecting it from food particles, digestive enzymes and of course pathogens [101]. To break this down, *P. larvae* could make use of a possible chitinase. A region encoding a chitinase A N-terminal domain was found in DSM 25719 and DSM 25430 [78]. Again, sequence similarity was found with the putative *C. botulinum* chitodextrinase [102]. The existence of a chitin-binding protein as a virulence factor has been confirmed for the ERIC-type II strain DSM 25430 [3] (section 3.4). A key virulence factor of *P. larvae* is *Pl*CBP49, which is a modular chitin-degrading protein. Its function has been confirmed in both ERIC I strain ATCC 9545 and ERIC II strain DSM 25430 [81]. The enzyme contains a module belonging to the auxiliary activity 10 family of lytic polysaccharide monooxygenases which degrades chitin using a metal ion-dependent, oxidative mechanism.

Another type of protease that could play a role in infection are the enhancin-like proteases. Enhancin is a viral protein used to disrupt the peritrophic membrane. In contrast to the above described chitinases, this protein targets insect intestinal mucin [103, 104]. An enhancin-like protein with 20 to 30% amino acid similarity has been identified in *B. thuringiensis* and has a similar mode of action [105]. The orthologues identified in *P. larvae* strains DSM 25719 and DSM 25430 seem however dysfunctional due to transposon insertions or frameshift mutations [78].

One third of the known proteolytic enzymes are serine proteases. These ubiquitous enzymes play a role in a wide range of biological processes such as digestion, fertilization, development, complement activation, pathogenesis, apoptosis, secondary metabolism, ... [106, 107]. It is thus not surprising to find loci encoding for serine proteases in the genome of all sequenced *P. larvae* strains [78].

Pathogens as *B. thuringiensis* are known to produce proteases that help to escape the host immune system [108]. It has been suggested that such a mechanism is present in *P. larvae*, since it is not affected by the intensive immune response of its host [78, 109, 110] (section 5). In the genome of the ERIC II strain DSM 25430, a gene encoding for a metalloprotease

belonging to the M6 peptidase family has been identified [78]. Its orthologue in ERIC I strain DSM 25719 is probably dysfunctional because of a frameshift mutation. The protein shows 41% amino acid similarity to the immune inhibitor precursor A of *B. thuringiensis*. This peptidase cleaves antibacterial peptides produced by the insect host [108]. The orthologue in *P. larvae* could have a similar function.

### 3.5.7  Antimicrobial compounds

*P. larvae* has always been isolated from larvae in pure culture, this suggests the bacterium is capable of excreting antibacterial compounds to protect its niche. It has been demonstrated that *P. larvae* can acquire antibiotic resistance by means of plasmids [111], this has implications to the use of antibacterial compounds as treatment (section 6).

Polyketide synthetases (PKS) and non-ribosomal peptide synthetases (NRPS) are metabolite producers. Metabolites often have a secondary role in self-defense or aggression including antibiotic and antifungal activities [112, 113]. In the sequenced *P. larvae* genomes clusters encoding for NRPS or PKS have been identified [78]. Paenilamicin is a NRPS/PKS hybrid antimicrobial compound that is a confirmed virulence factor of *P. larvae* [84]. The compound shows not only strong antibacterial and antifungal activities, but also cytotoxic activities. Although a knock-out of the gen does not result in a complete loss of virulence, the time course of the disease is influenced. One NRPS/PKS hybrid cluster encodes for a putative iturin family lipopetide antibiotic called Paenilarvins, which shows strong antifungal activity. This cluster has been proven not to encode for a virulence factor using exposure assays with knockout mutants [114]. A NRPS cluster putatively encodes for a gramicidin S synthetase, which is an antibacterial compounds that disrupts the cell membrane of its target microorganism. The last cluster with a similarity to known clusters possibly encodes for a bacillibactin like siderophore, an iron binding compound.

# 4  Worldwide impact AFB

Honey bees are direct producers of honey, wax and propolis, but their mere value is as pollinators. Bees are pollinators of several important agricultural crops. Using pollinators for fertilization has a direct impact in the yield and quality of crops. Fruit farmers often rent honey bee pollinators to carry out these services. Although worldwide the population of managed honey bees is increasing, this is not the case in the US and Europe, where bees suffer significant losses due to CCD and winter loss. In 2008 the number of managed colonies in the US was estimated at 2.3 million [115]. The value of pollinators is estimated at 3.66 billion USD in 2009 [116].

The impact of AFB on honey bees is difficult to establish. Only estimated prevalences of the disease are available, in Europe this has been estimated by the ANSES lab (the European reference lab for honey bee diseases) within the EPILOBEE project [117]. This estimated prevalence is $< 12\%$ over all member states, but significant differences exist between member states. The main difficulty in this estimation is the high number of amateur and hobbyist beekeepers, which are often not registered.

In most countries control measurements have to be taken when AFB is detected. These measures have tremendous effects for the beekeeper. Although the beekeeper can sometimes maintain the adult bees when no clinical signs are present, all honey, wax, propolis and other material will be destroyed. The biggest economic loss is however the impact on the pollination services. Honey bees are important pollinators and are often rented to fruit farmers to carry out the pollinator services. When AFB is detected, a restrictive perimeter is erected where no bees can be transported out or in. This necessary measurement has an important impact not only on the beekeeper who's hives are contaminated, but also for every other beekeeper and fruit farmer within this perimeter. Control measures that are taken by the governments have of course also a cost. One cost-benefit study has been conducted in Queensland (New-Zealand) in 1999 to address whether this cost of control measurements actually results in an economical benefit [118]. The cost-benefit analysis compared several scenarios where no control measures were taken with the current measurements taken as base level. The current measurements include registration of beekeepers/apiaries, reporting of cases of AFB, quarantining of AFB infected apiaries, destruction of infected bees and materials, compulsory testing of honey for AFB spores, certification of equipment for interstate transfer, prohibitions on the use of antibiotics to treat AFB and compensation schemes. The estimated total net benefit of beekeeping in Queensland was about $10.4 million at the time of the study, of which honey production was accounted for $8.4 million. The prevalence of AFB was estimated at 2% for this region. The investigators showed that without government program and an increase of AFB with 10% would result in an increased cost of $0.331 million, a raise of 20% would increase the cost with $4.713 and 30% is estimated at an increased cost $7.303. The main cost (98%) is attributed to the decline in honey production. This shows that without restrictive measurements the net cost of AFB would increase when the prevalence of the disease increases. A major impairment of this study is that the pollination services conducted by the honey bees are not taken into account. Since pollination services are regarded as the major contribution of honey bees to the economy, the cost-benefit analysis was severely biased.

# 5 The host *Apis mellifera*

The European honey bee (*Apis mellifera*) belongs to the family of the Apidae and the order of the Hymenoptera. The honey bee was one of the first domesticated animals and is highly appreciated for its honey production and pollination services. Several subspecies are present within *A. mellifera*. The subspecies frequently used in beekeeping are *A. mellifera* subsp. *carnica* and *A. mellifera* subsp. *ligustica*.

## 5.1 Casts and life cycle

The honey bee is an eusocial insect, where thousands of individuals live densely together in one colony. Three casts can be differentiated: queen, workers and drones. The queen and workers develop from fertilized eggs and are female, while unfertilized eggs develop to the male drone [119]. Each cast has its specific development, lifespan, morphology and task in the colony. A larva that hatches from a fertilized egg can develop in either queen as worker, this differentiation in cast depends on the larval food fed to them by nurse bees (young worker bees). During the first two instars all larvae are fed with royal jelly. From the third instar on workers are fed with a mixture of pollen, honey and royal jelly, while queen larvae are continuously fed the surplus of royal jelly [120–122]. Queens are responsible for reproduction, they lay up to 2000 eggs per day. The full development from egg to queen is 16 days, 3 days as egg, 6 days as larva (with 5 instars) and 7 days as pre-pupa and pupa. Five to six days after emergence, the queen is sexually mature [123], oogenesis occurs 3-7 days after mating [124]. Queens can live upto 8 years, however most die within 3 years [125]. Unfertilized eggs will develop to drones, which are thus haploid [119]. The egg hatches after 3 days, followed by 6 days as larva, the pre-pupal and pupal stage takes 15 days [126]. With a total period of 24 days from egg to drone, this cast takes the longest time to develop. The sole purpose of drones is to fertilize the queen, they are only present in the colony during the summer period. Mating is lethal, drones who are present in the colony after the mating period are expelled by the workers and killed or left to die of starvation. All other tasks in the colony are performed by worker bees. Worker bees develop within 21 days, 3 days as egg, 6 as larva and 12 days as pre-pupa and pupa (Figure 1.12)[126] . Larvae go through five instars in six days. Summer bees live on average 15-38 days, while winter bees survive for more than 140 days. Dependent on their age, workers perform different tasks within the colony. The youngest bees clean the nest and nurse the larvae, middle-aged bees build combs, process food and defend the nest while the oldest bees forage for food and water.

| D1 | D2 | D3 | D4 | D5 | D6 | D7 | D8 | D9 | D10 | D11 | D12 | D13 | D14 | D15 | D16 | D17 | D18 | D19 | D20 | D21 |
|----|----|----|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| egg | | | L1 | L2 | L3 | L4 | L5 | | PP | | pupae | | | | | | | | | |

Figure 1.12: **Developmental stages of the worker bee.** The development from egg to adult worker bee takes 21 days. Three days as egg, six as larva and 2 as pre-pupa and 10 as pupa. Figure adapted from [127].

## 5.2  Immunity

*P. larvae* is a major pathogen of the honey bee. Only larvae <36h old are susceptible. Upon infection an immune response is activated. Several studies towards the respons during the first 72h of infection have been done thoroughly [105, 109, 110, 128]. Next section is a general introduction to the immune pathways present in the honey bee. This is followed by the description of the specific immune reaction towards *P. larvae*.

### 5.2.1  Immune pathways

A review in 2003 by Hoffmann and colleagues thoroughly summarizes the insights in the existing immune pathways of model organism *Drosophila*, which formed the base of immunological research in insects [129]. Upon sequencing of the honey bee genome, the involved genes in *A. mellifera* were identified using comparative genomics [128].

Honey bees are social insects. This social behavior serves as an extra immune defense, but has also consequences for the strategies applied by pathogens. In a honey bee colony, thousands of bees live together at a very high density. The presence of stored resources and a homeostatic nest environment makes honey bees attractive targets for pathogens [130]. The social strategies developed by insects are grooming, nest hygiene and hygienic behavior. An example of this behavior are hygienic workers, who detect and remove infected larvae from the healthy brood [131].

The insects also possess individual defense mechanisms. The first defense is their hard exterior, which is protected by a layer of antimicrobial secretions. Gut pathogens have to overcome a very hostile gut environment [128]. The last morphological defense is the epithelium and the peritrophic matrix.

Insects not only have morphological defense mechanisms against pathogens, but also possess an effective cellular and humoral immunity. The humoral immunity results in the secretion of antimicrobial peptides, phagocytosis, melanization and the enzymatic degradation of pathogens [128, 129]. There are four immune pathways present in insects and for most pathway members an representative is present in honey bees and in *Drosophila*.

Toll pathway
The Toll pathway plays a dual role in insects, it is involved in immune response as well as in development [128, 129]. This pathway is activated by Gram-positive or fungal infections. The activation of the pathway starts when extracellular recognition factors initiate protease cascades that lead to the activation of the Toll ligand Speatzle. To these recognition factors belong peptidoglycan recognition protein (PGRP)-SA and Gram-negative binding protein (GNBP) 1 and 3. While GNBP3 is responsible for yeast recognition, GNBP1 and PGRP-SA recognize Gram-positive bacteria. Upon activation, Speatzle will bind to the Toll receptor (Figure 1.13). When activated the Toll receptor undergoes conformational and binds to MyD88. This interaction results subsequently in the recruitment of Tube and Pelle which form a complex with MyD88 through death-domain mediated interactions. The MyD88-Tube-Pelle complex leads to the phosphorylation and degradation of the I$\kappa$B factor Cactus. Degradation of this inhibitor results in translation of the NF-$\kappa$B transcription factor Dorsal (Dorsal-related immune factor Dif in *Drosophila*). Dorsal interacts with

Pelle, Tube and Cactus and will translocate to the nucleus upon pathway activation. The intracellular components Tollip, Pellin, Cactin and TNF receptor associated factor-2 are believed to aid the main players of the pathway. Dorsal will induce the expression of several effector encoding genes. Effectors of the Toll-pathway in honey bees include antimicrobial peptides, the melanizing agent phenoloxidase and three lysozymes.

In the *Drosophila* genome six families of Spaetzle related molecules are identified, they function as ligands for the nine Toll receptors [132, 133]. In the genome of *Apis mellifera* two plausible *Spaetzle* orthologues are identified. Two homologues of dorsal were found in the honey bee genome, neither of which were orthologous of Dif, the Dorsal-related immune factor in *Drosophila*. For Tollip, Pellino, Cactin and TRAF-2 representatives are identified in the genome of honey bee.

Imd pathway
In contrast to the dual role of the Toll pathway, the Imd pathway is only involved in immunity [134, 135]. The pathway is activated by binding of peptidoglycan to PGRP-LC, which recruits a signaling complex consisting Imd, dFadd and Dredd (Figure 1.13). Ubiquitination by E3-ligase inhibitor of apoptosis (Iap2), which associates with UEV1a, Bendless (Ubc13) and Effete (Ubc5) activates Dredd. Upon activation Dredd cleaves Imd, subsequently Iap2 ubiquinates Imd. This leads to the recruitment and activation of the complex Tab2/Tak1, which phosphorylates and activates the IKK complex. This complex leads to the phosphorylation and activation of the transcription factor Relish.

JNK pathway
The receptor Imd not only induces the Imd pathway, but also the JNK pathway (Figure 1.13). This pathway seems to result in both positive and negative feedback for the expression of the antimicrobioal peptides. The pathway is a canonical MAPK signaling module, it acts through a kinase phosphorylation cascade. Orthologues for each of the major components of this pathway are identified in the honey bee genome.

JAK/STAT pathway
The final humoral immune pathway is the JAK/STAT pathway. Initiation of this pathway happens through cytokine-like molecules in blood cells (Figure 1.13) [136]. In *Drosophila* the ligand Upd, an extracellular glycosylated protein, activates the pathway by binding to the receptor Domeless. This binding causes the phosphorylation of the receptor associated JAk tyrosine kinase and Domeless itself. Phosphorylation of Domeless results in the recruitment of and phosphorylation of STAT92E. STAT92E dimerises and translocates to the nucleus where it induces expression of the effector genes. In honey bees the cytokine receptor Domeless, the JAK tyrosine kinase Hopscotch and the STAT92E transcription factor are identified as the main components of the pathway. SOCS (suppressor of cytokine signaling) and PIAS (protein inhibitor of activated STAT) act as negative pathway regulators. Remarkably, no orthologue of Upd was identified.
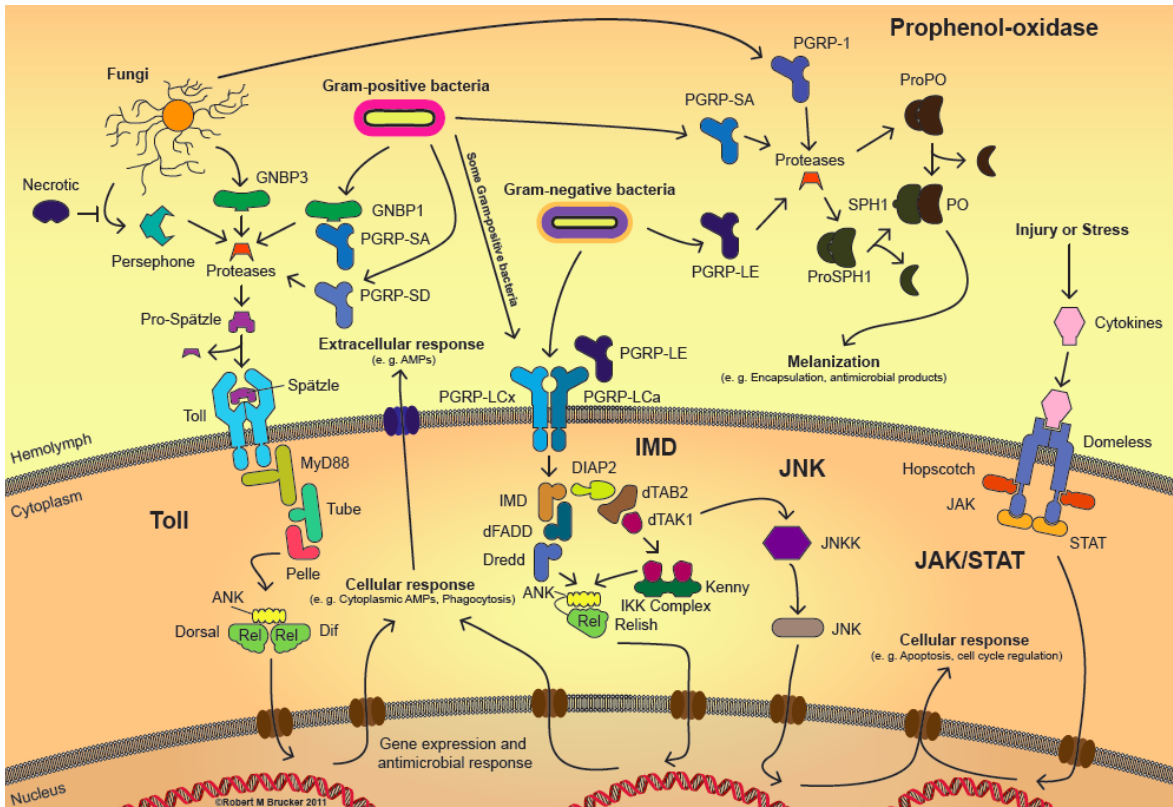
Figure 1.13: **Innate immune pathways in insects.** The names of the members are from the *Drosophila* pathway.From left to right: Toll pathway, IMD pathway, JNK pathway and JAK/STAT pathway. The **Toll** pathway is activated by Gram-positive bacteria and fungi. The associated recognition factors activate Speatzle. Speatzle binds and the Toll receptor inducing a conformational change. Intracellular signaling through the formed MyD88-Tube-Pelle complex leads to the degradation of Cactus (ANK) and release of transcription factor Dorsal. The **IMD** pathway is activated by Gram-positive and Gram-negative bacteria through the actions of the membrane-bound recognition factor PGRP-LC. PGRP-LC activation results in the construction of the Imd-dFadd-Dredd complex. Dredd gets activated by Iap2. Dredd cleaves Imd, which gets ubiquinated by Iap2. Subsequent recruitement of Tab2/Tak1 activates IKK. IKK phosphorylates and activates transcription factor Relish. The **JNK** pathway is activated through the Imd pathway. It consists of a phosphorylation cascade activating JNKK and JNK. Finally the **JAK/STAT** pathway is activated when Upd binds to the receptor Domeless. The associated factor JAK phosphorylates itself and Domeless, which results in the recruitement and phosphorylation of STAT92E. STAT92E dimerises and translocates to the nucleus. Figure from [137].

Immune effectors

A total of six immune effectors are present in the honey bee. Abaecin and Apidaecin are proline-rich antimicrobial peptides. Apidaecin consists of a conserved N-terminus which is followed by several exons [138]. The antimicrobial peptide (AMP) has also been identified in the venom, possibly playing a role in social immunity [139]. Using splice variation, a specific response to pathogens can be generated [128, 138, 139].

Defensin1 and Defensin2 are conventional defensins and Apisimin and Hymenoptaecin are distinct from all other recognized antimicrobial peptides [128, 140].

Diversity immune factors

When in 2006 the immune pathway orthologues of *A. mellifera* were identified in the genome, it became clear that the honey bee has a smaller variety of several immune factors [128]. PGRPs are major players in pathogen recognition. However in the honey bee genome they seem to be less diverse compared to other insects. Only 4 PGRPs are present, compared to 13 in *Drosophila* and 7 in *Anopheles* [128]. Several other recognition classes area reduced in honey bees these include the $\beta$-glucan recognition proteins, galectines and fibrinogen-related proteins [128]. The serine protease gene family in insects participates in the regulation of the immune cascade. In honey bees, an overall reduction of in size of this family is noted. This is also the case for the serine protease inhibitors [128].

A size reduction compared to *Drosophila* and *Anopheles* is also noted for the Toll-related receptor genes. Next a reduction in the number of immune effectors was noted, honey bees seem to express only six immune effectors in contrast to nine in *Anopheles* and 20 in *Drosophila*. This is also the case for the gene encoding for Prophenoloxidase (proPO), honey bees possess only one of these genes, while *Drosophila* possesses three and *Anopheles* nine.

From the above comparison, it is obvious that the pallet of immune effectors is less diverse for the honey bee compared to the model organisms *Drosophila* and *Anopheles*. A possible explanation for this is the social nature of honey bees. Social insects are hygienic, they remove organisms from their nests and secrete antimicrobial molecules that can reduce the survival of pathogens in their colonies. Larvae are raised in individual cells and fed with royal jelly (in later stages combined with pollen) which contains strong antibacterial substances. This social immunity is a protection that is absent in both *Drosophila* and *Anopheles*.

### 5.2.2 Immunological response induced by AFB

Multiple studies have been conducted to map the immunological response of the host against *P. larvae*. The first study was conducted in 2004 [109] and followed the immune response to infection of several larval instars until 3 days post-infection. Infection of first and second instars showed the upregulation of abaecin from 24h post infection. The author argues that this is probably the moment of germination of *P. larvae* spores, triggering this innate response. The peptide is known to be effective against Gram-positive bacteria [141]. A second antimicrobial peptide, Defensin, also seemed to be responsive to *P. larvae* infection.

A later study in 2006 followed larvae again until 24h post infection [128]. In this study, an upregulation of proPO activator (but not of proPO) was noted. Also the immune response

proteins PGRP-LC, Toll, Myd88, Speatzle, Kenny, Lysozyme 1 and Cactus 1 were higher expressed. These results broadened the results found in the earlier study.

A study in 2009 used proteomics to follow infected larvae until 5 days after infection [105]. An upregulation of 33 out of 179 proteins were noted. These included mitochondrial metabolic enzymes, feed proteins, chaperones and heat-shock proteins. Of the immune proteins, a clear upregulation was noted for Lysozyme 3 and Hymenoptaecine. The upregulation of Apidaecine and Defensine could not be confirmed, ProPO did show an upregulation, confirming the previous results.

Finally an RNA-Seq study was conducted in 2013 [110]. Again the expression in larvae until 3 days post infection was followed. A total of 75 upregulated genes and 6 downregulated genes were found. Among these significantly upregulated genes were the AMPs *hymenoptaecin*, *apidaecin* and *defensin 1*. However, the previously found upregulation of *lysozymes* and *proPO* were not confirmed.

An RNA-Seq study 72h post infection showed an upregulation of genes encoding antimicrobial peptides, two peritrophic-matrix domains (Pfam 01607), and all genes with the *Osiris* domain (Pfam DUF1676). The authors argue that this indicates that the unknown functions of these genes are broadly impacted by AFB. Another remark is that developmental asynchrony between treatments and control groups is a probable confounder [110]. This study thus confirmed the expression of *defensin1* found in the earlier study of 2004. The expression of *abaecin* seemed upregulated, but could not be statistically confirmed. The authors further explored this hypothesis by a qPCR study at 48h and 72h post infection. Indeed, it seemed that the energy depletion and tissue damage had delayed some developmental programs during the first two larval instars. This also confirmed the results of Chan and colleagues in 2009 [105], who did a proteomic study towards host response to *P. larvae* infection during the first 5 days of infection. This latter study also found an upregulation of *lysozyme* and *phenoloxidase*, which were not differentially expressed in the RNA-Seq analysis. This is probably due to the different time points taken, but also to the fact that proteomics include post-translational changes, which are not detected with genomic studies.

Although an extensive immune response is clearly present, no reports are known from larvae who overcome the disease [78]. This means that *P. larvae* must possess a strategy to circumvent and/or overwhelm the immune response (section 3.5).

## 5.3   Other challenges

Honey bees are confronted with several important challenges. Factors as starvation, queen loss and a variety of pests and pathogens (viruses, bacteria, fungi and metazoans) are current threats [142]. Co-infection of multiple parasites is often observed and many of the pathogens follow common transmission routes. Before the introduction of the parasitic mite *Varroa destructor* American Foulbrood was the most severe disease of the European honey bee [13].
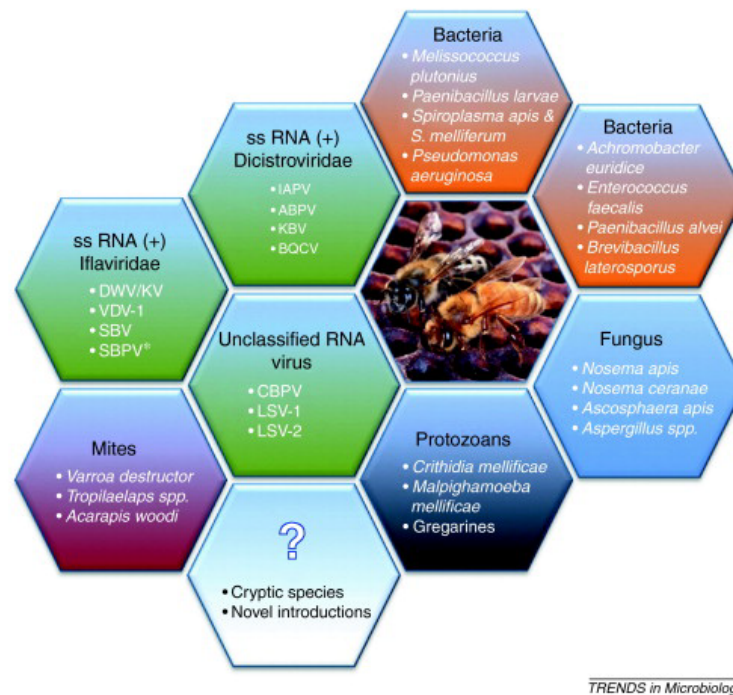
Figure 1.14: **Honey bee pathogens.** Honey bees are challenged by a wide variety of pathogens. Figure from [142].

### 5.3.1 Pathogens

Mite parasitism poses a major threat to honey bees worldwide. The most important among these parasites are *Varroa destructor*, *Acarapis woodi* and *Tropilaelaps clareae* [143, 144]. *V. destructor* is an ectoparasite causing major damage to honey bees in every part of the world. The mite originated from Asia where it infects its primary host the Asian honey bee *Apis cerana*. When infecting the European honey bee *A. mellifera* the immune system of the bee diminishes [145], making transmission of viruses who use these mites as a vector easier [146].

Honey bee viruses are often associated with weak and dying colonies [147, 148]. Several viruses are able to adapt the morphology, physiology and behavior of their honey bee host. As mentioned before, the ectoparasite *V. destructor* is often the vector used in transmission of these viruses [146].

American and European foulbrood are the most important bacterial diseases of the honey bee. European Foulbrood is considered to be a milder disease than American Foulbrood, mostly because the bacterium cannot produce spores. The disease only affects the brood, untreated it can cause collapse of the whole colony.

Three *Microsporidiae* infect the honey bee, being *Nosema apis* [149], *Nosema ceranae* [150] and the recently identified *Nosema neumanni* [151]. *N. apis* and *N. cerana* infect the epithelial cells of the adult midgut where they proliferate and produce infectious spores [152]. The parasites reduce the feed absorption and result in a higher probability of colony collapse [153]. Another fungus *Ascosphaera apis* infects the brood and is the etiological agent of the disease chalkbrood [154]. A typical symptom of the disease is the presence of "mummified" larvae, which are dead dried larvae covered by the fungus.

### 5.3.2 Colony Collapse Disorder and winter mortality

Colony collapse disorder (CCD) is a so-called disappearing disease where the loss of adult workers from the colony and the lack of dead bees within or near the colony is a specific symptom [148, 155]. Winter mortality is the cause of major losses in our regions. Before the introduction of *Varroa* a loss of 10% of colonies over the winter was considered normal. Since the introduction of the parasite global losses of 20% are considered normal, with reports of abnormal high losses over winter [155]. Winter mortality is clearly correlated with *Varroa* and *Nosema*.

Both losses are suspected to be associated with an combination of factors which lower the resistance of the honey bee colony. To these factors belong not only pathogens but also a lack of nutrition and presence of pesticides [147].

# 6 Controlling AFB

As stated before, *P. larvae* produces endospores which are the only infectious form of the pathogen. These spores are very resistant, they are unaffected by antibiotics, survive temperatures of minimal 80°C as well as freezing and can stay dormant for decades. This makes the treatment and eradication of American Foulbrood very difficult. AFB is a notifiable disease in may countries and control measures are often regulated by law [13].

In the US, *P. larvae* infections are allowed to be treated with antibiotics despite it is well known that the use of antibiotics eventually leads to resistance. Tetracycline was used as a treatment for decades, but resistance is conferred by the plasmid pMA67 [111]. This resistance motivates the search for new treatment methods, including alternative antibiotics and more resistant honey bees.

In European countries, treatment with antibiotics are forbidden. This results in the application of more destructive methods as burning and sulfur treatment or the less destructive shook-swarm method [13]. However, this practice is costly to the beekeeping community since it results in loss of hive materials, productive hives and produces significant economic loss [156].

Because of the above reasons, a lot of research is done to alternative treatment approaches. In the following sections all treatments are discussed in detail.

## 6.1 Shook-swarm method

Although swarming is considered a vertical transmission route of the disease (section 3.3), it can also be considered as a treatment method [13]. For successful vertical transmission, it is necessary that the adult bees carry spores on them, since the brood is left behind. These spores have to remain within the colony and infect the new brood to ensure the survival of the pathogen within the colony. When the swarm is able to get rid of these spores e.g. by grooming before the new brood gets infected, the colony will be disease-free. It was demonstrated in 2006 [157] that the initial infection level of the colony determines if swarming can be used to cure the colony. Swarms from infected, although not yet obviously diseased colonies (determined by $<$ 20 cfu/bee) were suitable to be cured by the shook-swarm method (Figure 1.15). The spore count of these colonies decreased drastically within 2 months post swarming and were nearly zero over 13 months, meaning that no larvae were successfully infected in these colonies and no new spores were produced. When clinically diseased colonies swarm, the spore load also decreases within 5 weeks, but increases again after this period. Therefore, swarming of clinically diseased colonies may present an efficient vertical transmission route, whereas swarming (naturally or artificially by shook-swarm) of non-clinically diseased colonies is a control measure [13]. This method is often used in countries that do not allow antibiotic treatment in combination with the destruction of clinically diseased colonies.
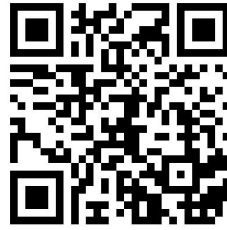
Figure 1.15: **The shook swarm method.** Movie showing the different steps in the shook swarm method.

## 6.2 Antibiotic treatments

As stated before, for more than 50 years oxytetracycline was used successfully in the treatment of *P. larvae* in the USA. However, in the '90 s more and more resistant strains were noted. This longevity of an antibiotic is remarkable, and is probably the result of the strict handling of foulbrood outbreaks. AFB is closely monitored by inspection agencies and suspected colonies lead in most countries to removal of both bees and equipment. Thanks to these measures, resistance to the antibiotic became widespread only after a long period of time. Resistance was obtained by the horizontal transfer of the plasmid pMA67 [111]. To be able to treat tetracycline-resistant *P. larvae*, the macrolide Tylosin is used now [158]. Several research groups are investigating honey bee gut bacteria able to inhibit growth of vegetative *P. larvae*. The antibiotics used by these bacteria can be promising in the treatment of American Foulbrood. Treatment methods using natural antibiotic substances are currently under development. An example of this is the use of essential oils. This could be promising, since it could combine a treatment for *Varroa* and *P. larvae*.

Antibiotic treatments (synthetic or natural) have some important negative characteristics. None of the known treatments affect spores, moreover they can disrupt the normal microbial ecology of the honeybee gut micro-organisms, especially in long-term use [159]. Honey bees are also a producer of honey, a product for human consumption. Antibiotics are known to leave residual amounts in honey [160]. This is the main reason why antibiotic use in honey bee colonies is forbidden in the European Union. Finally, antibiotic resistance is an inherent result of antibiotic treatment and control measures should always be taken to avoid the risk of generating resistant bacteria. As new antibiotics become available, their effectiveness can be prolonged by establishing strict guidelines for the rotation of different antibiotic treatments [161].

## 6.3 Antagonistic bacteria

Commensal lactic acid bacteria isolated from the crop of adult bees, have been found to inhibit the growth of *P. larvae*. However, none of the antimicrobial compounds have any effect on the highly resistant spores.

Another problem is that *P. larvae* controls its niche through the excretion of antimicrobial substances [13, 78]. The use of antagonistic bacteria *in vivo* could thus be difficult.

## 6.4  Bacteriophage treatment

A potential treatment method is bacteriophage therapy. Phages are capable of infecting and lysing a specific target, for bacteriophages these targets are bacteria [162]. Phage therapy has been used to treat a variety of diseases, but bacteriophage therapy of AFB is still in its development phase [156]. When phages can be used against *P. larvae*, this would be beneficial for economic and environmental health issues. Because of the specificity of phages, there would be no risk for bees, humans or other beneficial bacteria. The need for destructive measurements or antibiotic treatment would become obsolete.

Possible bacteriophages are isolated from environmental relevant samples as soil around the beehive, samples with bee-related ingredients, hives etc. Phages belonging to the *Siphoviridae* with a severe host specificity towards *P. larvae* have recently been isolated [156], but more research is needed in this field.

However, a negative point for bacteriophage treatment is the possibility of microorganisms to become resistant to their phage by mutations in their receptors or by becoming lysogenic. This risk can be reduced by using a cocktail of different phages. Another important restriction is that (again) the spores are not affected by this therapy.

## 6.5  Resistance acquired by honey bees

Resistance to the pathogen can also be acquired by the honey bee. Researchers try this by breeding for hygienic behavior and by trans-generational immune priming [163].

Breeding resistant honey bee colonies focuses on hygienic behavior. This method is actually inspired by the susceptibility of the Asian honey bee *Apis cerana* [164]. By *in vitro* rearing and infection, it was proven that very young larvae of this honey bee are susceptible for *P. larvae* and resistance grows with age. This is similar to *Apis mellifera*. But to obtain a same level of mortality, a higher spore dose has to be fed to *A. cerana* larvae compared to the *A. mellifera* larvae of the same age. Remarkably clinical signs of American Foulbrood have never been found during hive inspections. An important reason for this is that more than 80% of the infected larvae are detected by workers and removed before the capped stage. This hygienic behavior is a major aspect of the immunity against *P. larvae*.

Trans-generational immune priming (TgIP) is the transmission of maternal immune experience to the progeny [163]. Honey bees could benefit from TgIP since they are likely to face the same pathogen pressure as the queen. The queen can influence the immunity of the direct offspring, and thus increase the resistance to the infection upon immunological encounter with the pathogen. TgIP has been demonstrated when queens were injected with heat-killed *P. larvae* cells. TgIP resulted in a increase of the survival rate of 20% in AFB challenged offspring. However, the existence of this mechanism under natural conditions still has to be evaluated. Moreover, it is unknown whether the priming effect is a consequence of a general immune activation in the queen (e.g. through antimicrobial peptides) or of a targeted pathogen-dependent response (e.g. through haemocytes). Another possibility is the involvement of epigenetic factors.

# Chapter 2

# Aims and scope

> In order to combat a disease to the best advantage it is clear that its cause must be known, as well as the means by which the infection is transmitted and the environmental conditions which are favorable for the breaking out of an epidemic
>
> George F. White, 1906

George F. White, who first isolated *Paenibacillus larvae*, stressed the importance of obtaining knowledge about a disease-causing bacterium in order to combat it. More than a century later, this message is still prevalent.

Knowledge about the epidemiological traits of a bacterial infection is of major importance to develop accurate prevention and control strategies. Outbreaks are often caused by related isolates, arising from a common source. Understanding clonal relatedness between microbial isolates is thus essential for epidemiological surveillance. It allows to determine the source of infections, confirm or rule out outbreaks, recognize particularly virulent strains and evaluate effectiveness of control measures.

American Foulbrood is known to be a highly infectious disease. Honey bees are social insects living in a colony which may contain up to 50,000 individuals during the summer. The larvae are raised by the adult bees in this colony. Although the larvae themselves are not in direct contact, brood diseases can be spread within the nest through the activity of nursing bees. The adults spread the disease within the nest by feeding larvae the infectious spores. It is unclear how these spores end up in the hive, but feeding of honey is known to be a high risk factor. When the infection is not detected in an early stage, the colony will manifest disease symptoms and will have to be destroyed to prevent contamination of neighboring apiaries. However, by the time an outbreak is detected, typically multiple apiaries are already infected. Transmission between apiaries can be the result of drifting and robbing bees, but more often the beekeeper practices are responsible (using and trading contaminated material).
Discriminatory typing of the species can assist epidemiological work on the disease, help tracing the source of infection, aid in controlling it and help to assess whether the measures

taken prevent a further spread of this specific outbreak. Sensitive epidemiological information is however lacking for this disease, merely because to unavailability of a suitable typing tool.

A **first aim** in this doctoral research was to develop and implement a sensitive typing tool for *P. larvae*. As discussed in the introduction, several typing systems have been developed for *P. larvae* of which only MLST and MALDI-TOF MS are suited for high-throughput epidemiological research. However, both methods require expensive material, which is not available for many labs performing AFB diagnostics. An optimal typing method should meet several criteria: 1) the method must be sensitive, thus having a high differentiating power, 2) it should be intra- and interlaboratory reproducible and 3) all strains of the species have to be typeable with the method, 4) it must be fast and cheap to allow high-throughput analysis.

Multiple locus variable number of tandem repeat analysis (MLVA) is a tool that has been used for typing in other bacteria. The method can be designed to have a high discriminatory power and epidemiological studies using MLVA-typing has been conducted. In chapter 3 MLVA has been implemented for *P. larvae* and the feasibility to perform epidemiological studies using these MLVA-types has been addressed.

The availability of this typing tool could facilitate epidemiological research. A better understanding of the epidemiology of *P. larvae* could contribute to increase the effectiveness of surveillance systems and has provided significant clues to control strategies.

To be able to treat a disease it is of crucial importance to understand not only the epidemiology of an infection, but also the fundamental strategy the bacteria uses to infect and eventually kill its host. Understanding the infection route includes an overview of the involved virulence factors.

As described in the introduction the infection of *P. larvae* starts when the spores are fed to the larvae. They germinate and proliferate in the gut, where they successfully compete with other bacteria for the niche. In this step the role of antimicrobial compounds have been confirmed. In a second phase the peritrophic matrix, which protect the midgut epithelium, is broken down. The chitinase *Pl*CBP49 has a major role in the breakdown of the chitin-rich peritrophic matrix. In a third phase the gut epithelium is attacked and penetrated through the paracellular route. In this step toxins Plx1 and Plx2 are known to play a role. Finally the larva is digested to a ropy mass, probably through the activity of proteases.

Previous studies successfully identified the above described virulence factors based on a targeted approach, where the targets were chosen by comparative genomics or suppression subtractive hybridization. Although these efforts significantly improved the knowledge about the disease, some gaps are still present.

In this doctoral research the **second aim** is to identify factors in an untargeted way, allowing the identification of factors unique for *P. larvae*. The mutagenesis can be performed by using transposons. These mobile elements are known to integrate the host genome in an quasi-random way. Transposon-mediated mutagenesis has been implemented in other bacteria to identify virulence factors. Three different transposon methods were tried as discussed in chapter 4. First temperature-sensitive transposon mutagenesis was conducted, subsequently *in vitro* transposition was performed using the overactive Himar1 transposase and finally EZ-Tn5 transposome mutagenesis was implemented. The virulence

of the obtained mutants has to be tested using an *in vitro* infection assay. Although *in vitro* rearing of larvae results in intercasts and ignores the social immunity of honey bees, it allows a standardized approach to assess the virulence of the mutants. Using this setting confounders as age and ingested spore load can be controlled. Moreover, since AFB is a notifiable disease, the possibility to perform natural infections is almost impossible. Since a affected growth rate could indirectly result in attenuated virulence, growth curves should be created for each obtained mutant. By creating untargeted mutants, we try to fill the gaps that are still present in the infection route and provide a complete picture of the different strategies used by *P. larvae* to successfully infect its host *Apis mellifera*.

# Chapter 3

# Multiple Locus Variable number of tandem repeat Analysis: A molecular genotyping tool for *Paenibacillus larvae*

**Multiple Locus Variable number of tandem repeat Analysis: A molecular genotyping tool for *Paenibacillus larvae***

Authors:

Tine Descamps[1]*, Lina De Smet[1], Pieter Stragier[2], Paul De Vos[2], Dirk C. de Graaf[1]

*For correspondence: E-mail: tine.descamps@ugent.be

tel: +32 92648734

Address: Krijgslaan 281 S2, 9000, Ghent, Belgium

Affiliations:

[1] Laboratory of Molecular Entomology and Bee Pathology, Faculty of Sciences, Ghent University, Ghent, Belgium

[2] Laboratory of Microbiology, Faculty of Sciences, Ghent University, Ghent, Belgium

Descamps T. wrote the paper and conducted all analysis and experiments. De Smet L. provided guidance in the analysis and conduction of the experiments and assisted in the writing process. Stragier P. provided guidance in the design of the MLVA and assisted in the writing process. De Vos P. and de Graaf D.C. were co-promotor and promotor. They conceived the study, gave guidance during the experiments and assisted in the writing of the publication.

# 1 Abstract

American Foulbrood, caused by *Paenibacillus larvae*, is the most severe bacterial disease of honey bees (*Apis mellifera*). In order to perform genotyping of *P. larvae* in an epidemiological context, there is a need of a fast and cheap method with a high resolution. Here we propose Multiple Locus Variable number of tandem repeat Analysis (MLVA). MLVA has been used for typing a collection of 209 *P. larvae* strains from which 23 different MLVA types could be identified. Moreover, the developed methodology not only permits the identification of the four ERIC genotypes, but allows also a discriminatory subdivision of the most dominant ERIC type I and ERIC type II genotypes. A biogeographical study has been conducted showing a significant correlation between MLVA genotype and the geographic region where it was isolated.

# 2 Introduction

The Gram-positive, rod-shaped, endospore-forming bacterium *Paenibacillus larvae* is the ethiological agent of American Foulbrood (AFB) [2], a deadly brood disease of the honey bee (*Apis mellifera*). Very young larvae, below 36h of age, are most susceptible. Transmission occurs by the oral uptake of the highly resistant spores that are spread inside the colony or between colonies by adult bees that carry them or by interventions of the beekeeper [8, 9]. AFB is classified as a notifiable disease in most countries and depending on the local control strategy, sick colonies are destroyed by burning, decontaminated by the shaking method or treated with antibiotics. Hence, *P. larvae* is responsible for considerable economic losses in the beekeeping sector worldwide. Historically, it was thought that American Foulbrood and 'powdery scale disease' were caused by different species, *Bacillus larvae* and *Bacillus pulvifaciens* respectively [17]. After different rounds of reclassification [18, 19], they were classified as the same species, but split at subspecies level [21]. In 2006 was demonstrated that both should be classified as the species *Paenibacillus larvae* without separation at the subspecies level [2]. Over the years, different techniques have been developed for genotyping of *P. larvae*, as there are Restriction Endonuclease Fragment Patterns [58, 59], rep-PCR [62, 63], Pulsed-Field Gel Electrophoresis [2, 60], Amplified Fragment Length Polymorphism [26], Denaturing Gradient Gel Electrophoresis [61] and most recently Multi Locus Sequence Typing (MLST) [65]. Of these techniques, only rep-PCR has commonly been used for genotyping *P. larvae*. Enterobacterial Repetitive Intergenic Consensus (ERIC) rep-PCR of the bacteria revealed four genotypes [2, 63]. Both genotypes ERIC I and ERIC II have a worldwide distribution [56, 65]. Genotypes ERIC III and IV have not been identified in field for decades, but exist as few isolates in culture collections [13]. ERIC genotyping has the advantage that it splits the species into biological relevant groups which differ in colony morphology and virulence at individual and colony level [2, 23, 71]. The sequencing of the genomes of *P. larvae* ERIC I and ERIC II strains [76–78] provided an important step towards the development of better molecular typing methods. The recently developed MLST method [65] ratifies and extends the ERIC typing scheme. Although this MLST is a very useful method for epidemiology and source tracking, its use in diagnostic labs is hampered by its expensive and labor-intensive sequencing step . We developed an alternative, equivalent technique called Multi Locus Variable number of tandem repeat Analysis (MLVA) [165] for genotyping *P. larvae*. MLVA is a proven highly discriminatory subtyping method used for many pathogens, such as *Mycobacterium tuberculosis* [166], *Yersinia pestis* [167], *Staphylococcus aureus* [168], *Erwinia amylovora* [169] and *Campylobacter jejuni* [170]. The method has also been used to link genotype information with geographical information to study how bacteria (*Yersinia pestis*, *Burkholderia pseudomallei*, *Xanthomonas citri pv. citri*, *Bacillus anthracis* and *Erwinia amylovora*) behave within smaller geographic areas or single outbreaks [169, 171–174]. The typing method uses the Variation in Number of Tandem Repeats (VNTR) on different loci among the genome to classify the strains [165]. These variations are caused by slipped strand mispairing [175]. VNTRs have been described as fast molecular clocks [165]. Indeed, since dynamics of VNTRs depend on repeat copy number [176], different VNTRs show different clock speeds [177, 178]. This makes the method suitable for phylogenetic and evolutionary studies. In MLVA, different VNTR loci are combined, allowing inspection of phylogenetic relationships among different bacterial strains. Typically 5 to 6 loci are combined in one multiplex PCR, which can be analyzed using electrophoresis. In this paper we describe the development of a MLVA-

based genotyping protocol for *P. larvae*, and the subsequent typing of a collection of 209 *P. larvae* strains, demonstrating that MLVA has great potential for genotyping *P. larvae* in an epidemiological context.

# 3   Results and Discussion

### 3.0.1   Identification of tandem repeat regions

We screened three publically available *P. larvae* genomes (BRL 230010 [77] and its up-dated sequence B3650 [76], DSM 25430 and DSM 25719 [78]) for tandem repeats using an online software package (Tandem Repeats Database at https://tandem.bu.edu/cgi-bin/trdb/trdb.exe; [179]. BRL 230010 and DSM 25719 belong to the ERIC I genotype, whereas DSM 25430 is an ERIC type II. In order to permit VNTR analysis on an agarose gel, the search criteria for tandem repeats were set at a size between 15 to 120 bp with 80% sequence match. This resulted in the finding of 40 different tandem repeat loci (Supplementary material: Table S1).

The search for suitable VNTRs was continued by a two-step procedure. Candidate loci were selected based on different criteria in silico, and their respective primers were designed. Three criteria were used: (1) the locus had a different copy number in at least 2 genomes or (2) it had a pattern size between 15 and 30 or (3) it had a copy number of more than 5 units. In a second step the differentiating power of these loci was tested on 13 different isolates. These isolates included the four different ERIC genoytpe and were isolated at different locations. ERIC types III (LMG 16252) and IV (LMG 16247) were represented once and ERIC II twice. The first selction criterium resulted in eight candidate loci, of which four were differentiating enough to be retained for implementation in the multiplex PCR: i.e. VNTR A, VNTR C, VNTR F and VNTR G (Table 3.3). An optimal MLVA procedure includes five to six VNTRs and therefore the search for more VTNRs was needed. The second selection criterium rendered six additional loci. Of these six loci, 2 were retained after testing in the 13 strains: VNTR B and VNTR D (Table 3.3). Finally, loci with more than five repeat units were screened. This gave two loci to be tested, one of which was retained: VNTR E (Table 3.3). Thus, in total seven loci were found with enough differentiating power for implementation in the multiplex PCR.

## 3.1   Multiplex PCR

Optimization of the multiplex PCR was realised by establishing the optimal concentration of $MgCl_2$, primers and template DNA. In the final PCR two loci were omitted because they created non-target amplicons (VNTR F) or no amplicons during multiplexing (VNTR G). The obtained MLVA profiles were analyzed using a 3% High Resolution Agarose Gel (Sigma-Aldrich, USA) (Figure 1). The screening of the 13 strains created a total of 9 different MLVA patterns. The screening was repeated three times and gave reproducible results.

## 3.2   MLVA genotyping

The Laboratory of Molecular Entomology and Bee Pathology has access to a large collection of *P. larvae* strains available from the bacteria collection of the Belgian Co-ordinated Collections of Micro-organisms (BCCM/LMG) and from its own working collection. The strains that were retained for the present study originated from a nationwide honey screening in Belgium [73] (116 isolates) or from clinical outbreaks of AFB in Belgian apiaries between 2013-2015 (74 isolates), and were extended with 3 additional reference strains (BRL 230010 [77], LMG 16252, LMG 16247; both obtained from BCCM/LMG), 6 Austrian ERIC II isolates [180] and 10 Italian ERIC II strains [181]. Using these 209 different isolates, 14 new MLVA types were discovered resulting in 23 MLVA patterns (Figure 3.1, Table 3.1).
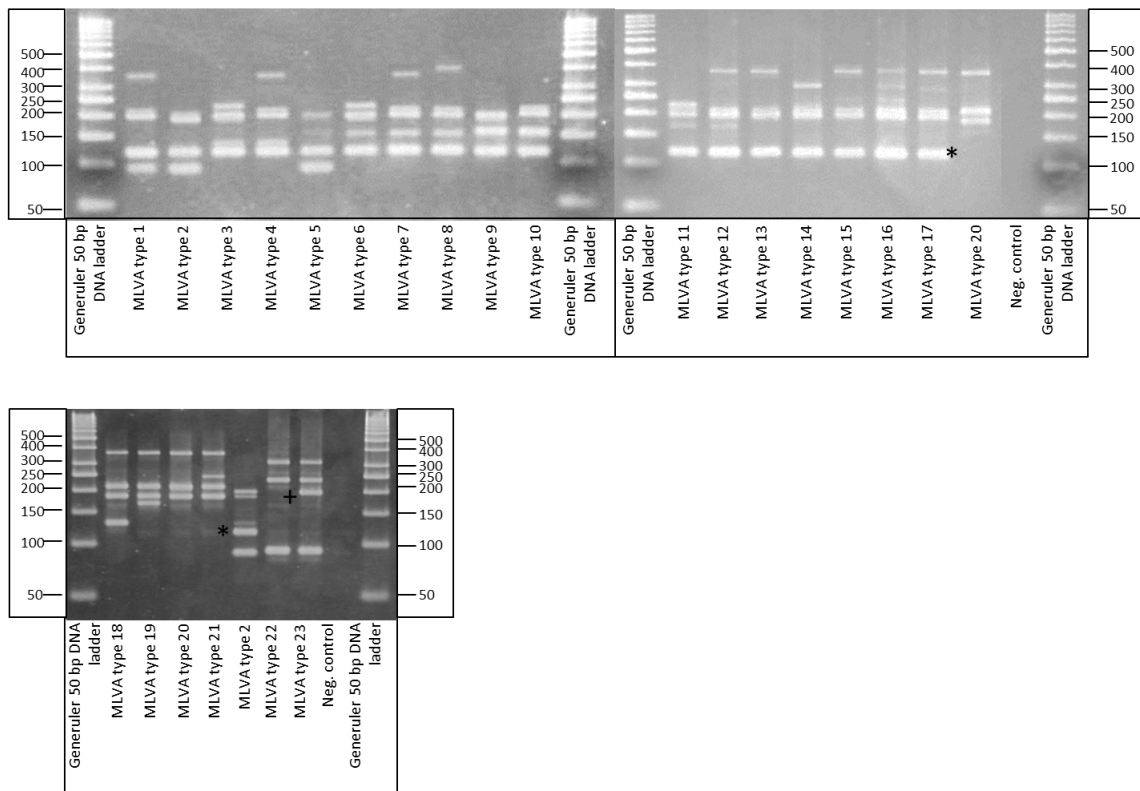


Figure 3.1: **MLVA types.** The tested collection of *Paenibacillus larvae* strains contained 23 different MLVA types. The ERIC I types (MLVA type 1–17) all have a band of 120 bp (*), which is absent in all other ERIC types. ERIC III (LMG 16252, MLVA type 22) could be differentiated from ERIC type IV (LMG 16247, MLVA type 23) by the presence of a 190 bp band (+). ERIC types II and IV do not show an VNTR B amplicon, ERIC type III does not show an VNTR E amplicon.

A representative isolate for each pattern was chosen and its amplicons were sequenced. The resulting sequences confirmed that the primers targetted their respective tandem repeat in each MLVA pattern and that no off-targets were amplified. The copy size of the tandem repeats could be established for each locus. These were consistent with the size of the amplicon on the gel (after substracting the flanking sequences) (Table 3.3, Figure 3.1).

In Table 3.1 the typical VNTR code is given for the different MLVA patterns. In order to facilitate the analysis, we attributed a MLVA-type number (MLVA 1 – 23) to each unique VNTR combination (Table 3.1).

Table 3.1: **MLVA types.** Of each MLVA type the VNTR-code, ERIC genotype and prevalence in our dataset is given. Seventeen MLVA types belonged to ERIC type I strains, four belong to ERIC type II. ERIC type III (LMG 16252) and IV (LMG 16247) were represented by only one MLVA type. Most MLVA types show a very low prevalence, only three types have a prevalence higher than 10%.

| MLVA | VNTR-code | ERIC genotype | Prevalence (%) |
|---|---|---|---|
| 1 | 2-6-4-3-5 | I | 1.0 |
| 2 | 4-6-3-3-1 | I | 30.1 |
| 3 | 4-6-3-3-3 | I | 1.0 |
| 4 | 4-6-4-3-5 | I | 1.4 |
| 5 | 5-6-3-3-1 | I | 1.0 |
| 6 | 5-6-3-3-3 | I | 3.8 |
| 7 | 5-6-4-3-5 | I | 4.3 |
| 8 | 5-6-4-3-6 | I | 1.0 |
| 9 | 6-6-3-3-2 | I | 2.3 |
| 10 | 6-6-4-3-2 | I | 11.0 |
| 11 | 6-6-4-3-3 | I | 1.4 |
| 12 | 6-6-4-3-5 | I | 1.9 |
| 13 | 7-6-4-3-5 | I | 17.7 |
| 14 | 10-6-4-3-4 | I | 3.8 |
| 15 | 10-6-4-3-5 | I | 4.8 |
| 16 | 13-6-4-3-4 | I | 0.5 |
| 17 | 13-6-4-3-5 | I | 3.3 |
| 18 | 4-0-3-7-5 | II | 0.5 |
| 19 | 6-0-3-7-5 | II | 1.0 |
| 20 | 7-0-3-7-5 | II | 6.2 |
| 21 | 8-0-3-7-5 | II | 0.5 |
| 22 | 2-6-5-11-0 | III | 0.5 |
| 23 | 2-0-5-11-0 | IV | 0.5 |

## 3.3   VNTR characteristics

VNTR A was found to be the most diverse locus, splitting up in 8 different lengths of tandem repeats. The number of tandem repeats ranged from 2 to 13. The second most divers VNTR was VNTR E, followed by VNTR C, D and B (Table 3.2 and Figure 3.2). The discriminatory power of the MLVA is due to the complementary discriminating power of each VNTR locus. The MLVA typing method based on these VNTR loci is powerful enough to find new MLVA types for strains that are not present in the tested dataset.
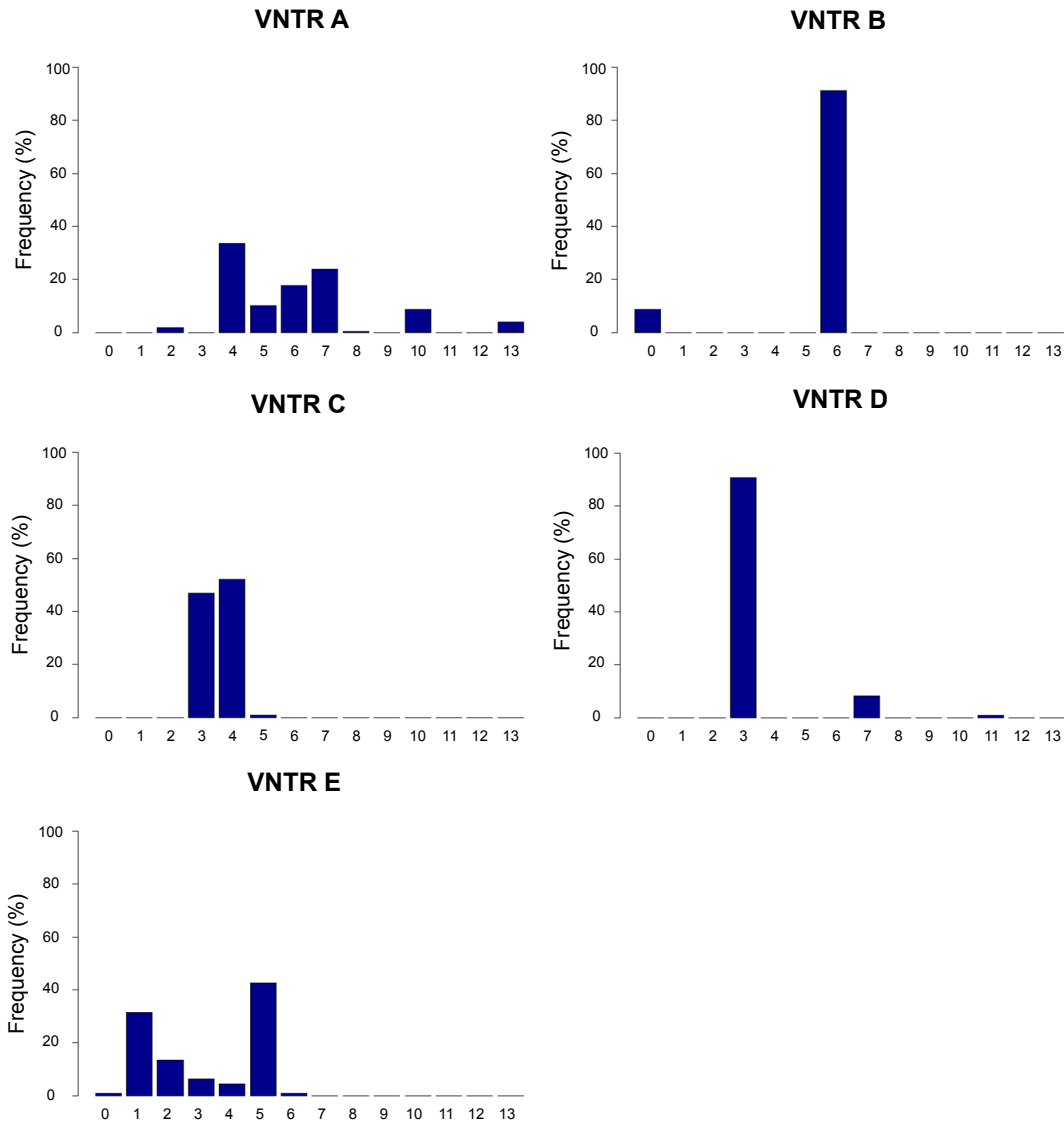
Figure 3.2: **Diversity of VNTR loci.** For each VNTR locus the tandem repeat (TR) copy numbers are given with their respective prevalence in our dataset. VNTR A: 8 alleles (2, 4, 5, 6, 7, 8, 10, 13 TR), VNTR B: 1 allele (6 TR) and no amplicon in $8.2\%$ of the strains (the ERIC II strains and ERIC IV (LMG 16247)), VNTR C: 3 alleles (3, 4, 5 TR), VNTR D: 3 alleles (3, 7, 11 TR), VNTR E: 6 alleles (1, 2, 3, 4, 5, 6) and no amplicon in $0.5\%$ of the strains (ERIC III strain LMG 16252).

Table 3.2: **VNTR characteristics of the full dataset (193 isolates).** The length of the tandem repeat, the number of different alleles (No. alleles), the number of tandem repeats in each allele (No. repeat copies), the maximum and the minimum number of repeats (max rep, min rep) are given. The Shannon Index and Simpson Index were both used to measure the diversity.

| VNTR | Length TR | No. alleles | No. repeat copies | Min rep | Max rep | Shannon | Simpson |
|---|---|---|---|---|---|---|---|
| VNTR A | 19 | 8 | 2, 4, 5, 6, 7, 8, 10, 13 | 2 | 13 | 1.68 | 0.78 |
| VNTR B | 21 | 1 | 6 | 6 | 6 | 0.29 | 0.16 |
| VNTR C | 24 | 3 | 3, 4, 5 | 3 | 5 | 0.74 | 0.51 |
| VNTR D | 24 | 3 | 3, 7, 11 | 3 | 11 | 0.34 | 0.17 |
| VNTR E | 68 | 6 | 1, 2, 3, 4, 5, 6 | 1 | 6 | 1.39 | 0.70 |

## 3.4 Evolutionary history

The five loci had all different molecular clock speeds, which makes it possible to look at the evolutionary history of the MLVA types (Figure 3.3). The first node in the tree clustered the types 22 and 23 apart. To these MLVA types belong the ERIC type III (LMG 16252) and type IV (LMG 16247) strains respectively. These are the only strains of the tested collection that correspond to the former *P. larvae pulvifaciens* subspecies. The second node clustered all MLVA types that contained ERIC type I strains apart from the types that contained the ERIC type II strains.

## 3.5 ERIC genotyping

All strains were also typed using ERIC rep-PCR, which confirmed the presence of 189 ERIC type I strains, 18 ERIC type II strains, 1 ERIC type III (LMG 16252)and 1 ERIC type IV strain (LMG 16247). Within ERIC type I, 17 MLVA genotypes could be discriminated, and ERIC type II 4 MLVA genotypes were identified. Within ERIC type III (LMG 16252) and ERIC type IV (LMG 16247), each time a single MLVA type could be attributed (Table 3.1). The subdivision of ERIC type I and ERIC type II in multiple subtypes is in the line of the findings of Morrissey [65] who found by MLST 16 subtypes in 173 ERIC type I strains and 3 subtypes in 92 ERIC II strains. Both typing schemes have thus on average the same resolution. The present MLVA genotyping protocol also permits to differentiate between the ERIC genotypes by the presence of some discriminatory bands. Indeed, a band of 120 base pairs is always present in ERIC type I, but not in the ERIC type II. ERIC type III (LMG 16252) and type IV (LMG 16247) have a similar pattern, however, ERIC type III has always a band of 200 base pairs, which is absent in ERIC type IV (Figure 3.1). The VNTR B primers do not create an amplicon in ERIC II and IV strains, while the VNTR E primers do not create an amplicon in ERIC II strains.
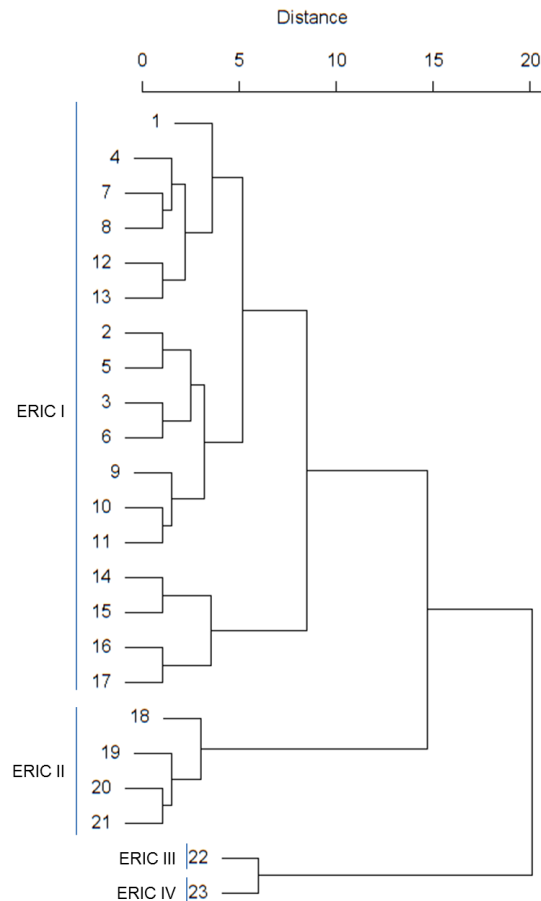
Figure 3.3: **Evolutionary history of MLVA types.** The evolutionary history of the MLVA types is based on hierarchical clustering. The distances were calculated using the sum of absolute differences in VNTR repeat number. The grouping reflected the ERIC-typing method

## 3.6 Prevalence and biogeography

The prevalence of the MLVA types within the complete collection differed significantly (Table 3.1). Only three types had a prevalence of more than 10.0%: MLVA 2 (30.1%), MLVA 13 (17.7%) and MLVA 10 (11.0%). Within the ERIC II strains, MLVA 20 was predominant (14/18 isolates), this type contained isolates from Belgium, Italy and Austria. Almost half of the MLVA types (10 out of 23) were represented by 1 or 2 isolates, corresponding to a prevalence of 0.5-1%.

As the geographic spread of the apiaries where the isolates originated from (honey or brood samples) were mostly known (169 strains), we were able to provide the distribution per province (Figure 3.4). A chi square test using the geographic region (province) of the apiaries and the MLVA type was conducted and gave a very significant p-value of <0.001 (chi-square 153.33, df = 144). We saw a much greater diversity in the subset of strains that originated from the honey screen compared to those from the clinical outbreaks (Shannon diversity index of 2.11 and 1.36, respectively). In fact, in the latter mostly 1 or 2 MLVA types were found per province, this could be due to the rapid spread of the spores in the vicinity
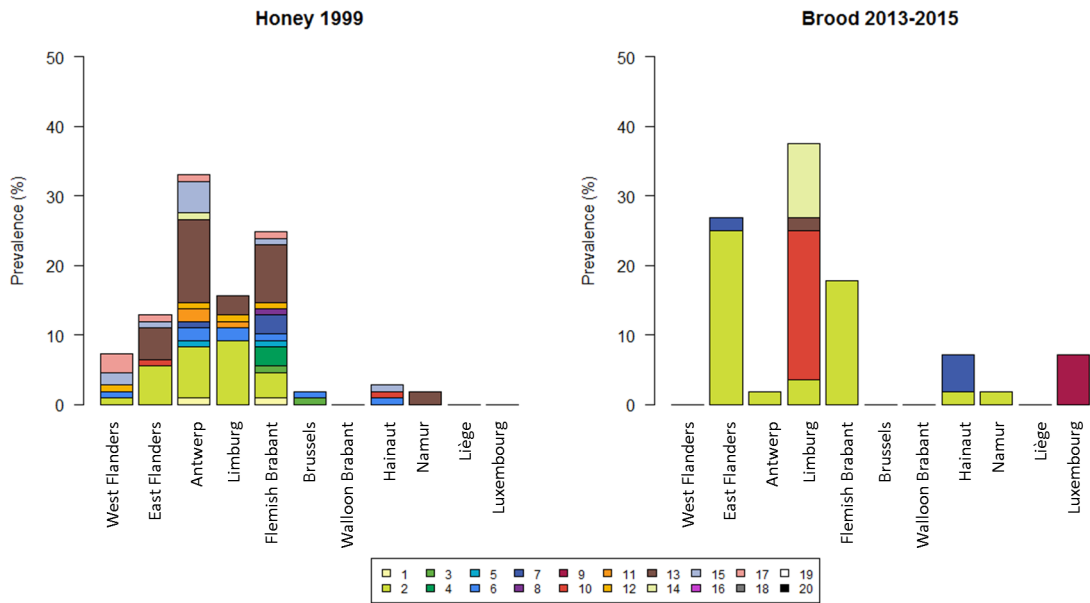
Figure 3.4: **Distribution of MLVA-types.** The geographical location of each isolate from honey in 1999 and diseased brood in 2013–2015 was recorded. A significant difference in distribution over the provinces was observed. A much higher diversity in MLVA-types was found in the honey samples of 99 than in the brood samples of 2013-2015.

of a clinical outbreak as described in de Graaf et al. (2001) [73]. Remarkably, the MLVA type 2 which was most abundant in clinical cases in East Flanders, Antwerp, Limburg and Flemish Brabant and was likewise abundant in honey samples taken almost two decades earlier. It is not unraveled yet whether this is due to failure to eliminate the contamination with this MLVA type completely or to its biological characteristics (virulence). Another MLVA type (MLVA 13) that was abundant in honey samples of the same provinces, was almost entirely absent in the clinical cases of the period 2015-2013. In the brood samples of the province Limburg MLVA 10 represents more than half of the isolates, while it is not present in any of the clinical cases of the other provinces. Finally, it was investigated if classification of the isolates using a dissimilarity matrix resulted in the same MLVA types classified in the same group. Agglomerative hierarchical clustering was used as the algorithm. From the resulting tree (Figure 3.5), we can conclude that indeed strains with the same MLVA type are on average more similar to each other based on non-genetic traits (regional code, province, beekeeper, isolation year, isolation source) than to other MLVA-types. The significant correlation with the region is also visual.

# 4   Conclusions

In this paper, the development of a new genotyping method for *P. larvae* is presented. Using this method, we could subdivide ERIC type I into 17 different genotypes and ERIC type II into 4 different genotypes. ERIC type III (LMG 16252) and IV (LMG 16247) were represented each time by only a single MLVA genotype. Moreover, in the present MLVA genotyping protocol amplicons indicative for the ERIC genotype were included, making it possible to
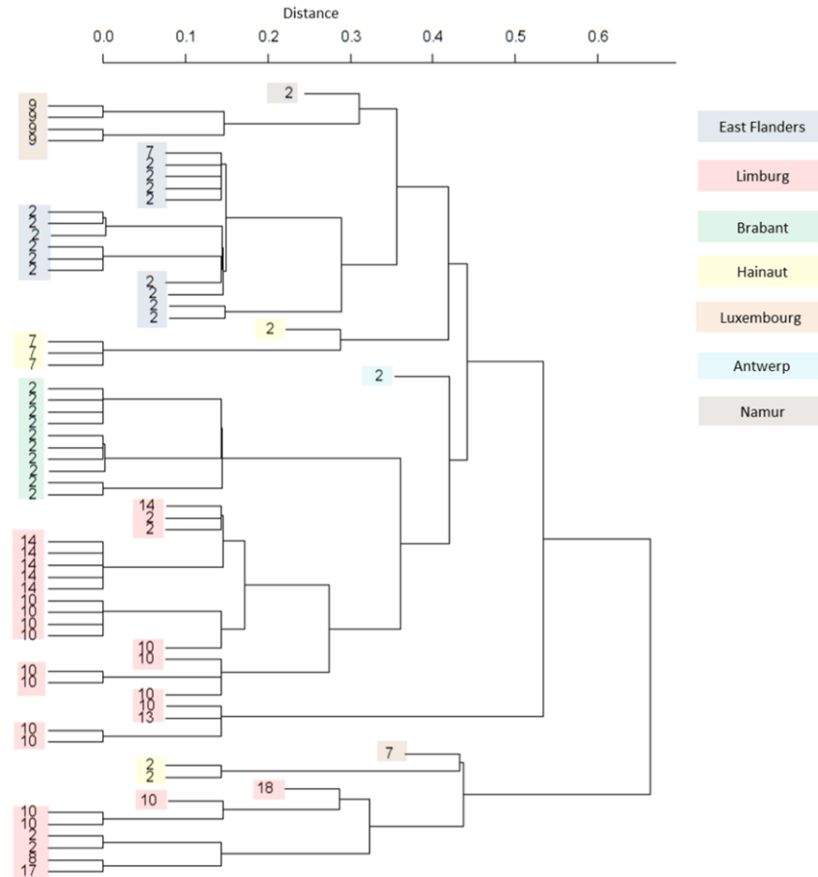
Figure 3.5: **Agnes hierarchical clustering.** Agnes hierarchical clustering based on the dissimilarities of non-genetic traits clustered the strains by MLVA type and province. The correlation between these two parameters was visual.

extract information of the ERIC genotype from the MLVA pattern. This study clearly shows the usefulness of the MLVA method in epidemiological and biogeographical studies. MLVA genotyping has the unique combination of desirable properties for a genotyping method. It is very fast, highly specific, cheap and not labor intensive, which makes it an almost perfect method to implement in bee health extension laboratories. The method is usable for epidemiological, phylogenetic and evolutionary studies.

# 5 Experimental procedures

## 5.1 Dataset

In the library 116 isolates were collected in the context of the honey screening in 1999 [73]. This study was conducted anonymously throughout Belgium, with the postal code of the area where the apiary was located as geographical reference. One honey sample was collected per apiary. The library also consisted of isolates collected from diagnostic AFB

cases in Belgium, including a set of isolates from clinical outbreaks from 2013-2015. From the latter the postal code and beekeeper was recorded. Finally, the set of field isolates was completed with 6 Austrian ERIC type II isolates [180] , 10 Italian ERIC type II isolates [181], ERIC type III (LMG 16525), and ERIC type IV (LMG 16247) (both from the BCCM/LMG bacteria collection) reference strains, and a sequenced ERIC type I strain (BRL 230010) [76, 77] .

## 5.2    Preparation dataset

Bacterial genomic DNA was prepared using InstaGene matrix (Bio-Rad, USA) according to the protocol described by Genersch and Otten [63]. To confirm that the isolates belong to *P. larvae* a 16S rDNA PCR was performed as described by Dobbelaere [25]. Of each strain the ERIC genotype has been determined according to the protocol of Genersch and Otten (2003).

## 5.3    VNTR prediction

The publically available genomes of BRL 230010 [76, 77], DSM25430 [78], and DSM 25719 [78] were used as input in the Tandem Repeats database (Gelfand et al., 2006). As selection criteria we set the alignment parameters to match = 2, mismatch = 5, indels = 7, the minimal alignment score = 80, the pattern size <= 120 bp. From the obtained list of possible targets (Table S1), primers (IDT, Belgium) were designed for loci that (1) had a different copy number in at least 2 genomes or (2) had a pattern size between 15 and 30 or (3) had a copy number of more than 5 units. The primers were constructed targeting the flanking regions of the tandem repeat locus and to have an annealing temperature of 52 °C. The differentiating power of the tandem repeat loci were tested using 13 *P. larvae* isolates. Ten of the thirteen isolates were randomly selected from the dataset, and were extended with R-20833 (ERIC type II), LMG 16252 (ERIC type III) and LMG 16247 (ERIC type IV). A locus suitable to include in the multiplex PCR should generate a specific amplicon for each isolate and had to make discriminate between strains possible using agarose gel electrophoresis.

## 5.4    Construction multiplex PCR

Seven loci were initially combined in a multiplex PCR. Two loci were omitted, because the first one generated off target amplicons and the second failed to generate amplicons. The multiplex PCR was optimized by testing a DNA-concentration gradient (20-120 ng DNA), $MgCl_2$ gradient (1 − 3 nM) and variable combinations of primer concentrations (0.2 − 4 µM). The final multiplex PCR used the HotStarTaq Plus DNA Polymerase kit (Qiagen, Germany) and consisted of 1x PCR buffer, 2.5 mM MgCl2, 1 x Q-solution, 400 µM dNTPs, 0.8 µM of each VNTR A primer (IDT, Belgium), 0.4 µM of each VNTR B − D primer and 4 µM of each VNTR E primer (Table 3.3). The total reaction volume was 25 µl and 100 ng total DNA is used as template. The PCR-program was as followed: 5' 95°C, 30 x (1' 94°C, 1' 52°C, 1' 72°C) and 10' 72°C. The multiplex PCR was repeated 3 times on the 13 isolates and proved to give reproducible results.

Table 3.3: **VNTR loci.** Five VNTR loci (depicted with *) were used to combine in the multiplex PCR. Of each VNTR locus the length, forward primer and reverse primer are given. The distance (in base pairs) between the primer annealing site and the tandem repeat was mentioned as length 5' and length 3'

| VNTR | Length TR | Forward Primer | Length 5' | Reverse primer | Length 3' |
|---|---|---|---|---|---|
| VNTR A* | 19 | GAGGGATATACCCCACCTCTTT | 5 | GGGGAAGTATGATCCCGAAG | 17 |
| VNTR B* | 21 | CCGGAATAATCCGCTTATGA | 22 | ATCACCAGAGTTGGCGATTC | 3 |
| VNTR C* | 24 | TGGTTTAGGAACCGGTGTTG | 47 | CACATTAAAGCCTGTGCAGGTA | 38 |
| VNTR D* | 24 | ATCATGGCGGTTGGGATG | 2 | CACAGGCTCGACAACCACTA | 13 |
| VNTR E* | 68 | TGTTCAATTTTGATTGTTTTGTTCA | 73 | TATATGGCGGTCGGCTTAAT | 2 |
| VNTR F | 48 | TACCCCAATCTGCCTTGTTG | 70 | CATGCTCCTGCGTGGTATAA | 41 |
| VNTR G | 18 | GTCATTACGGCCCAGGTG | 20 | TGAGGCTGCAAAGACAGATG | 22 |

## 5.5   Data scoring and description profiles

The length and number of tandem repeats was determined by analyzing the agarose gel and confirmed by sequencing the differing amplicons of the 23 MLVA types. Sequencing was done by Sanger sequencing performed by GATC-biotech (Germany). The number of repeat units was rounded to the next highest integer number. If more than one amplicon for a specific VNTR was detected and the difference in length between these amplicons was higher than the repeats length (stutter peaks), the amplicon with the most intense band was used to assign the repeat number. The Shannon and Simpson diversity index was calculated using R [182]. Using agglomerative hierarchical clustering (Agnes), evolutionary analysis was performed. The clustering method used the sum of the absolute distances to calculate the dissimilarities between observations.

## 5.6   Analysis correlations

All statistical analysis are performed using R [182]. Correlations were analyzed using a Pearson's Chi-square test with the geographical location (province) and MLVA-type as variables. To perform a reliable test, MLVA-types with less than 5 isolates were removed, provinces containing less than 5 isolates were joined with neighboring provinces. A Pearson's chi square test was performed on the full dataset.

## 5.7   Agnes classification

In a final stage of the epidemiological analysis, an Agnes classification was performed. This is a hierarchical classification and can be used to show how the different MLVA classes cluster together when the MLVA-type is not taken into consideration. Since our variables are almost exclusively categorical variables, a dissimilarity matrix had to be constructed first. This numerical dissimilarity matrix was then be used as input in Agnes.

# 6   Acknowledgment

# Chapter 4

# Untargeted mutagenesis contributes to a better understanding of the virulent behavior of *Paenibacillus larvae*

The results of this chapter are published in the Journal of Applied Microbiology as:

**Unbiased random mutagenesis contributes to a better understanding of the virulent behavior of *Paenibacillus larvae***

Authors:
Tine Descamps[1], Lina De Smet[1], Paul De Vos[2], Dirk C. de Graaf[1]
*For correspondence: E-mail: dirk.degraaf@ugent.be
tel: +32 92648734
Address: Krijgslaan 281 S2, 9000, Ghent, Belgium
Affiliations:

[1] Laboratory of Molecular Entomology and Bee Pathology, Faculty of Sciences, Ghent University, Ghent, Belgium [2] Laboratory of Microbiology, Faculty of Sciences, Ghent University, Ghent, Belgium

Descamps T. conducted all analysis and experiments and wrote the paper. De Smet L. guided all experiments and analysis and assisted in the writing of the paper. De Vos P. and de Graaf D.C. were co-promotor and promotor. They conceived the study, gave guidance during the experiments and assisted in the writing process.

Additional information excluded from the publication:

- Experiments preceeding Tn5 transposome mutagenesis.

- Exact definition of the mutations affecting the different loci

- Discussion of difficulties encountered mutagenesis.

- More detailed discussion of the biological context of the mutated genes.

# 1   Abstract

American Foulbrood, caused by the Gram-positive bacteria *Paenibacillus larvae*, is one of the most severe bacterial diseases of the European honey bee. The bacterium has been known for long, but only the last decade the mechanisms used by the pathogen to cause disease in its host are starting to unravel. In this study the knowledge of this virulent behavior is expanded and several possible virulence factors are suggested.

Identification of possible virulence factors has been done by random mutagenesis to ensure untargeted approach. A library of mutants was tested for a significant difference in virulence using *in vitro* exposure assays. Affected loci were characterized and their potential to contribute in virulence of the pathogen was assessed.

The identified mutated loci *ywqD* and *cirD* are involved in pathways affecting virulence factors in other bacteria. They suggested the possible involvement of biofilm formation (through EPS), the antibiotic circularin A in the infection strategy of *P. larvae*. Another mutated locus *gbpA* encodes the virulence factor GbpA in other bacteria, a similar role in *P. larvae* is possible.

The study identified new possible virulence factors for *P. larvae* genotype ERIC I in an untargeted way. This contributes to the knowledge and understanding of the possible mechanisms used by this pathogen to colonize and kill its host.

# 2   Introduction

The Gram-positive, rod-shaped, endospore-forming bacterium *Paenibacillus larvae* is the etiological agent of American Foulbrood (AFB) [2], a deadly brood disease of the honey bee (*Apis mellifera*). AFB is classified as a notifiable disease in most countries and depending on the local control strategy affected colonies are destroyed by burning, decontaminated by the shaking method or treated with antibiotics. Hence, *P. larvae* is responsible for considerable economic losses in the beekeeping sector worldwide.

Recently a thorough review of the biopathology of *P. larvae* has been published [1], which describes the pathogenesis of this brood disease step by step. Very young larvae, below 36 h of age, are most susceptible and resistance to the disease grows with age. The infection starts with the oral uptake of highly resistant spores, the only infectious form of the bacteria [10]. These endospores germinate in the larval midgut lumen and massively proliferate before breaching the epithelium [11]. *P. larvae* eliminates bacterial competitors in the larval gut by releasing non-ribosomal peptides [83, 84, 86–88]. Subsequently, the peritrophic membrane (PM) that protects the underlying gut epithelium is attacked by chitin-degrading enzymes, one of which (*Pl*CBP49) was recognized as key virulence factor [81]. Several toxins known to attack the gut epithelium are likewise considered to be virulence factors of *P. larvae*. Two toxins Plx1 and Plx2 are confirmed as AB toxins present in ERIC genotype I strains [82]. Another toxin, called C3larvin has been found in both ERIC type I and ERIC type II strains, but its biological role in the pathogenesis has not been established yet [183]. In ERIC type II strains, the functional role of an S-layer protein has been discovered [3]. The protein seems to be involved in the attachment of *P. larvae* to the midgut epithelium.

In the last decade, several molecular tools became available that seemed crucial to further explore the virulence mechanisms of the bacterium. First, the genomes of representative *P. larvae* ERIC type I [76–78] and ERIC type II [78] strains were sequenced and annotated. Second, successful electroporation and transformation of *P. larvae* [79, 80] enabled site-specific mutagenesis using the TargeTron system (Sigma-Aldrich), a procedure that allowed to functionally confirm several virulence factors [3, 81–84]. Further discovery of the virulence of *P. larvae* will contribute to better understand the process of pathogenesis of this disease causing bacterium.

In this chapter we describe a search for virulence factors of *P. larvae* that is based on the construction of a library of mutants. In this chapter the experimental procedures are discussed before the results, since several mutagenesis strategies were attempted before kanamycin-resistant mutants were obtained. These included temperature-sensitive transposon mutagenesis using plasmid pID408 [184], *in vitro* transposition with Himar1 transposase [185, 186] and finally Tn5-based transposome mutagenesis (epicenter, Madison, USA). The obtained *P. larvae* mutants were tested for atypical virulence using *in vitro* infection assays. Finally, the disrupted genes of the mutants with atypical virulence were identified using whole genome sequencing.

# 3 Experimental procedures

## 3.1 *P. larvae* strain

*P. larvae* BRL230010 was used throughout all experiments, this strain was isolated from a clinical affected AFB colony in Berkely, CA, USA. This strain is genotyped as ERIC I as described in [2] and has been sequenced in 2006 and updated in 2011 [76, 77]. BRL230010 was routinely grown on MYPGP plates or in liquid MYPGP cultures containing containing $0,01\%$ m/v nalidixic acid and $0,02\%$ m/v pipemidic acid. For selection of mutants kanamycin (30 µg/ml), erythromycin (10 µg/ml) and/or chloramphenicol (15 µg/ml) were added.

## 3.2 Optimization electroporation

The mutagenesis techniques were being performed using electroporation. Two other *P. larvae* strains were already electroporated by other research groups [79, 80]. Neither of these protocols proved successful to electroporate our strain. To be able to mutagenize BRL230010, the electroporation protocol was optimized using plasmid pMG14 (BCCM). First the protocol to prepare electrocompetent *P. larvae* cells was adapted. We started from the protocol of Murray and colleagues of 2008 [80].

For this electroporation to be successful the protocol to make electrocompetent cells had to be adapted. The long lag phase during growth of *P. larvae* was a severe burden to obtain a high enough transformation rate suitable to allow transposon mutagenesis. It has been proven that the growth phase is of crucial importance for the electrocompetence of bacterial cells [187]. Adding larval fluids helped to shorten the lag phase significantly.

A pre-culture of 3ml MYPGP was inoculated with 1 colony of *P. larvae* and grown for 16h at 37°C under constant shaking. One ml of this culture was used to inoculate 150 ml MYPGP + $3\%$ larval fluids. Adding the larval fluids improved the growth rate of the bacteria significantly [188], making the culture grow to an $OD_{600}$ of $2.5$ in about 6h under constant shaking at 180rpm. All larvae used were collected from the experimental beekeeping facility Honeybee Valley (Ghent University, Belgium). The larval fluids were extracted by centrifuging larvae at $4°C$, $7000$ $g$ during 30 minutes. To avoid activation of prophenoloxidase and subsequent melanization, phenylthiourea (Hopkins and Williams Ltd) [$100\mu$g/ml phosphate buffered saline (PBS)(Sigma-Aldrich)] was added to a final concentration of $10\mu$g/ml. The fluids were filter sterilized with a $0.2\mu$m filter (Whatman). Subsequently the protocol of Murray and colleagues [80] was followed to make the cells electrocompetent.

*P. larvae* cells were electroporated with a range of voltages from 1kV/cm-2.5kV/cm. One ml MYPGP was added immediately, followed by 16h recuperation at $37°C$ under constant shaking at 180 rpm. Selection of mutants was performed on MYPGP plates with $30$ µg/ml kanamycin. After 3 days of incubation at $37°C$, successful mutants could be selected. Electroporation was defined as be successful when the plasmid could be extracted from the electroporated cells.

## 3.3   Temperature sensitive transposon mutagenesis

The first method that was used for mutagenesis was temperature sensitive transposon mutagenesis. This method was based on the protocol of Mei and colleagues [184] who successfully transformed and identified virulence factors in *Staphylococcus aureus*. Plasmid pID408 was constructed by these authors for mutagenesis. It contains an temperature-sensitive replicon and a chloramphenicol resistance gene that allows replication and selection in *S. aureus* at 32°C. For *P. larvae* the recuperation phase was tested at different temperatures (32°C - 35°C). The transposon Tn917 present in the plasmid contains an erythromycin resistance gene. Successfully recuperated *P. larvae* cells are subsequently selected on erythromycin and chloramphenicol containing MYPGP plates at the same temperature used for recuperation. After 3 days colonies containing the pID408 plasmid should be visible. To generate the mutants, bacteria from the master plate are picked and transferred to liquid MYPGP media containing only erythromycin and grown at non-permissive temperature (37°C). Stable transformants should be erythromycin resistant and chloramphenicol sensitive.

## 3.4   *In vitro* transposition

*In vitro* mutagenesis using the overactive Himar1 transposase is a second method suitable for transforming bacteria in a untargeted way. In this method the transposition is performed outside the cell. The DNA fragment containing the transposon is then electroporated in the cell where homologous recombination inserts the transposon in the genome.

### 3.4.1   Purification of Mariner Transposase

The first step in the transposition is the purification of the transposase. The transcription of the transposase is controlled by a IPTG inducible promotor. IPTG was added to 0.2 mM when the culture had grown to an $OD_{600}$ of 0.7. After induction growth of the culture was extended for 2h. The culture was chilled in an ice bath for 5 minutes. The cells were pelleted (15' 1000xG) at 4 °C and resuspended in 5 ml resuspension buffer (20 mM Tris-Cl pH 7.6, 25% sucrose, 2 mM $MgCl_2$, 1 mM DTT, and a 1 % protease inhibitor cocktail (Sigma, contains inhibitors for serine, metallo-, aspartic, and cysteine proteases). The downstream steps are performed on 1 ml of this solution. The rest can be divided into 1 ml aliquots and stored at -80 °C. The cells in the 1 ml aliquot were lysed by adding 0.25 mg lysozyme and incubation at room temperature during 5 minutes. One ml of detergent buffer (20 mM Tris-Cl pH 7.6, 4 mM EDTA, 200 mM NaCl, 1% deoxycholate, 1 % NP40, 1 mM DTT, 1 % protease inhibitor cocktail) was added, followed by incubation during 15 min at room temp. Subsequently $MgCl_2$ was added to 10 mM final concentration, followed by 60 $\mu$g DNaseI. The solution was mixed gently until it was no longer viscous and incubated for 20 min at room temp. The mariner transposase is present in inclusion bodies, thus the suspension was pelleted and the supernatans discarded. The pellet was washed with 1 ml of 0.5% NP-40, 1 mM EDTA at 4°C and twice in 1 ml 6 M deionized urea. Resuspension of the pellet was performed in 0.5 ml column buffer (4 M guanidine-HCl, 20 mM Tris-Cl pH 7.6, 50 mM NaCl, 5 mM DTT, 1 % protease inhibitor cocktail). Insoluble debris was removed by pelleting for 2 minutes, the supernatans containing the mariner transposase

was transferred to a DEAE Sephacel column (8 ml) equilibrated with buffer at 4 °C. Mariner transposase was eluted from the column with 16 ml of column buffer. Subsequently the transposase is dialyzed into 1 L of 50 mM Tris-Cl pH7.5, 100 mM NaCl, 10 mM MgCl$_2$, and 1 mM DTT at 4 °C. The dialysis buffer was changed two times over a period of 12-20 hrs. Any insoluble material was pelleted after which the solution was filter sterilized with a 0.2 $\mu$m syringe filter. The protein concentration was determined with a spectrophotometer.

### 3.4.2  *In vitro* transformation

In a second step the transformation is carried out *in vitro*. In a microfuge tube kept on ice, 500 ng of the transposon donor plasmid pJFP2 is mixed with 1.5 $\mu$g of digested *P. larvae* chromosomal DNA (previous digested with *Eco*RI and/or *Xho*I). Two $\mu$l 10x restriction enzyme buffer 4 (New England Biolabs), 0.2 $\mu$l 10 mg/ml BSA (New England Biolabs) 0.01 $\mu$g of purified transposase was added. The total volume was brought to 20 $\mu$l by adding H$_2$O. The solution was incubated at 30°C for 1h, followed by heat-inactivation (10' 70°C). Subsequently, 17.5 $\mu$l was precipitated with 0.5 $\mu$l glycogen, 2 $\mu$l 3 M NaOAc and 50 $\mu$l ethanol. The rest of the solution was used to check the success of the transformation using a southern blot. A UDP-labeled probe targetting the kanamycin resistance gene in the transposon was used. The DNA pellets were suspended in 14 $\mu$l 0.1xTE buffer. To this suspension 2 $\mu$l 10x T4 DNA polymerase buffer, 2 $\mu$l 1 mM dNTPs, 1 $\mu$l of a 1:10 dilution of 10 mg/ml BSA and 1 $\mu$l of T4 DNA polymerase (1-3 units/$\mu$l) was added. Polymerisation was performed during 20 minutes at room temperature, the reaction was stopped by heat-inactivation. To this solution 4 $\mu$l 10x ligase buffer (NEB), 1 $\mu$l of T4 DNA ligase (NEB) and H_2O (to 20 $\mu$l) was added. Ligation was conducted at room temperature during 1h. After purifying the final suspension with PCR purification kit (Qiagen), electroporation can be performed.

## 3.5   Custom-made EZ-Tn5 transposome

A third way that was tried to create a library of transposon mutants, is using the EZ-Tn5 custom-made transposome system (epicenter, Madison, USA). The EZ-Tn5 transposon contains a R6K$\lambda$ori, which makes replication possible in pir+ *E. coli* cells. A kanamycin resistance gene was cloned into the transposon to allow selection of transposon mutants. The kanamycin resistance gene from pMG14 (Belgian Co-ordinated Collections of Micro-organisms) was amplified by PCR using primers KanF and KanR (IDT, Leuven, Belgium) (Figure 4.1, Table 4.1). Primer KanF harbors a *Eco*RI restriction site and KanR binds immediately after a *Eco*RI site in pMG14. The PCR reaction mixture contained 5 x PCR buffer, 5 x Q-solution, 8 $\mu$M of each primer KanF and KanR and 2$u$ HotStarTaq Plus DNA polymerase (Qiagen, Hilden, Germany). The total reaction volume was 50 $\mu$l. The PCR-program started with 5' 95°C initial heating, 40 x (15" 95 °C denaturation, 1' 54 °C annealing and 1' 72 °C elongation) and 10' 72 °C final elongation. Subsequently the amplicon was digested with *Eco*RI and cloned into the multiple cloning site (MCS) of the EZ-Tn 5 backbone vector pMOD-3 (Figure 4.1). The transposon was amplified by PCR using the manufacturers recommendations (Figure 4.1, Table 4.1).
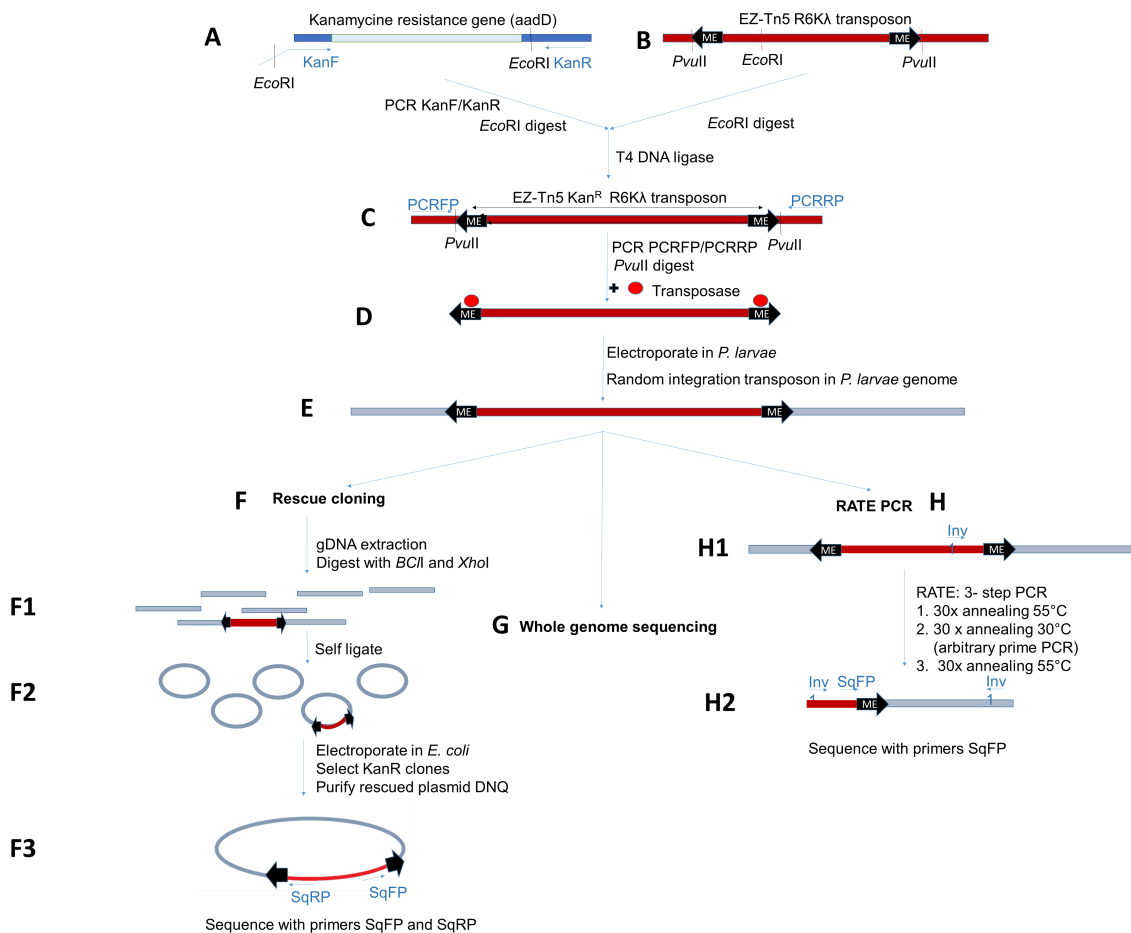
Figure 4.1: **Flowchart of methodology.** The flowchart summarizes all steps to create the mutant library and the steps taken to identify the mutated genes. A: Kanamycin resistance gene of plasmid pMG14 was picked up by PCR. The amplicon was digested by *Eco*RI. B: The EZ-Tn5 R6K$\lambda$ transposon (Epicenter, Madison, USA) contains a *Eco*RI restriction site in its MCS. The transposon contains also a R6K$\lambda$ origin of replication which makes rescue cloning (step F) possible. C: The digested amplicon of A is ligated into the MCS of B creating a EZ-Tn5Kan$^R$R6K$\lambda$ transposon. This transposon picked up by PCR and the ends are subsequently trimmed by *Pvu*II. The EZ-Tn5 transposase is added. D: In absence of MgCl$_2$ the transposase binds to the mosaic ends (ME) of the transposon and form a stable complex. E: This complex is electroporated into electrocompetent *P. larvae* cells. The availability of MgCl$_2$ activates the transposase, which results in the random insertion of the transposon in the *P. larvae* genome. F: A first method to identify the insertion locus is Rescue Cloning. F1: Rescue cloning starts with the isolation and digestion of the genomic DNA of the mutant. Double digestion by *Bcl*I and *Xho*I is performed, these enzymes have no restriction site in the transposon. The ends are blunted by T4 DNA polymerase. F2: Self ligation of the gDNA fragments is performed by adding T4 DNA ligase. The circularized gDNA fragments are electroporated into pir+ *E. coli*. Since the transposon contains a R6K$\lambda$ ori the created plasmid in F2 can replicate in pir+ *E. coli* cells. The transposon also contains a kanamycin resistance marker, which makes selection possible. F3: Using the appropriate sequencing primers, the locus of insertion can be identified. G: The locus of insertion can also directly be identified by whole genome sequencing of the extracted gDNA. H: Random Amplification of Transposon Ends (RATE) PCR is a third method suited for identifying the locus of insertion. H1: RATE is a three-step PCR consisting of a stringent temperature step, a low temperature step and again a stringent temperature step. Only one primer is used in RATE. H2: By using a nested sequencing primer, the locus of insertion can be identified.

Table 4.1: **Primer sequences.** Sequence of all primers used in the Tn5 transposome protocol. Restriction enzyme recognition sites are underlined.

| Name | Sequence |
|------|----------|
| KanF | 5' TAAGCA<u>GAATTC</u>AGGAGAAGTTAATAAA 3' |
| KanR | 5' TAAACAGCCCTTCCCGTAGA 3' |
| PCRFP | 5' ATTCAGGCTGCGCAACTGT 3' |
| PCRRP | 5' GTCAGTGAGCGAGGAAGCGGAAG 3' |
| SqFP | 5' GCCAACGACTACGCACTAGCCAAC 3' |
| SqRP | 5' GAGCCAATATGCGAGAACACCCGAGAA 3' |
| Inv1 | 5' ATGGCTCATAACACCCCTTGTATTA 3' |

Digestion with *Pvu*II trimmed the sequences flanking the mosaic ends (ME). The final transposome complex was formed by adding the Tn5 transposase in absence of $MgCl_2$.

## 3.6    Infection assay

The mutants were screened for a difference in virulence using exposure assays. The mutants were triggered to sporulate using a *in lab* optimized method, since only the spores of *P. larvae* are infectious. De Smet and colleagues proved that adding larval fluids resulted in the expression of known sporulation genes [188]. It was reasoned that spiking MYPGP agar slants with $5\%$ larval fluids would result in sporulation of the bacteria. One ml of over night *P. larvae* culture in liquid BHI was added to the spiked slants. These slants were incubated at $37°$C during 10 days. The washing steps described by Poppinga and colleagues [79] were followed to obtain the final spore stock. The concentration of each stock was obtained by heating the spore suspension for 10 minutes at $80°$C, followed by serial dilution plating.

*In vitro* infection assays were used to screen for mutants with atypical virulence. We were bound to the seasonality of honey bees to implement our infection assays (these are only feasible between end May and half August, depending on the weather conditions). Rearing of the larvae sets a lot of stress on the honey bee hive and queen, which resulted in the observation of queens who stopped laying eggs, were killed or colonies who showed extremely aggressive behavior. The infection assay was performed according to previously published assays [71] (Figure 4.2).

All mutants were tested in a preliminary virulence screening using 10 larvae per group and 2000 spores/ml feed. Mutants with atypical virulence (either more or less virulent) were rescreened in triplicate with 30 larvae per group and 500 spores/ml. This corresponds to the dose that the wild type BRL230010 strain needs to kill on average $80\%$ of all larvae ($LD_{80}$). The brood was monitored during 15 days. Dead larvae or pupae were counted daily, crushed in PBS and plated on MYPGP plates to confirm the presence of vegetative *P. larvae*. In each experiment control groups were taken into account, a positive control group infected with the wild type strain and a non-infected negative control group.
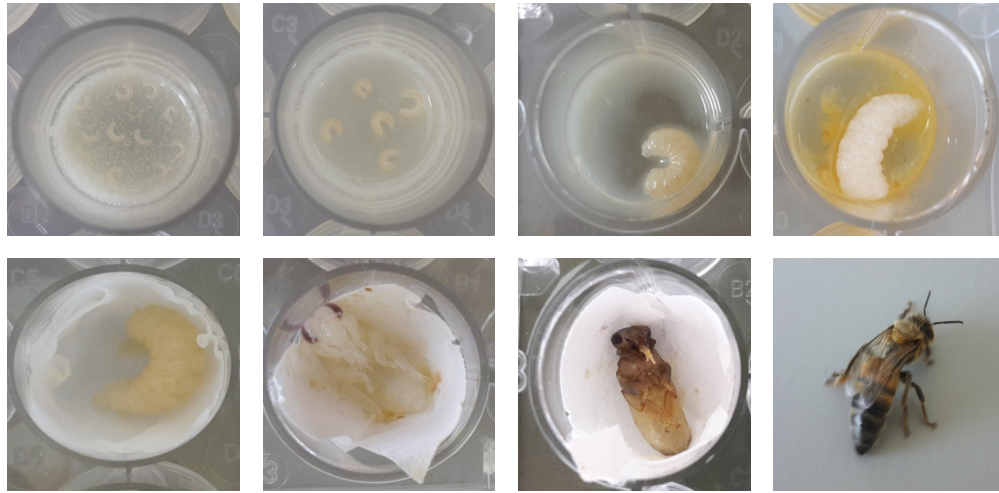
Figure 4.2: ***In vitro* rearing of larvae.** An infection assay was performed using in vitro reared larvae.Upper row: First instar larvae are grouped per ten in a 24-well plate and fed ad libitum with feed containing spores (except for the negative control). After 24h they get non-spore containing feed. During their development, the larvae are transferred daily to fresh feed and are distributed over several wells. The end of the larval phase is characterized by defecation. Bottom row: After defecation the larvae are transferred to pupal wells. In these wells healthy larvae transform to the prepupal phase, pupal phases and finally are born as an adult honey bee.

## 3.7   Statistical analysis

The mortality rate is calculated as the ratio of total number of larvae that died from AFB compared to the number of exposed larvae. An unpaired Student's t-test was performed at the $5\%$ significance level to statistically confirm the difference in mortality rate using R [182]. In this test the null hypothesis is defined as no difference in mortality rate due to AFB between the mutant and the wild type BRL 230010. Larvae which died from other causes than AFB, were subtracted from the total number of larvae.

A survival analysis using Cox proportional hazard models was performed to statistically compare the complete mortality curves with the mortality curve of the wild type strain [189, 190].

When the atypical virulence of the mutant was statistically confirmed, rescue cloning, RATE and finally whole genome sequencing were performed to identify the disrupted gene.

## 3.8   Rescue cloning

To perform rescue cloning (Figure 4.1) a high amount of genomic DNA is necessary. Extraction of this DNA was done by scraping the bacteria from 5 full grown MYPGP + 30 $\mu$g/ml kanamycin plates and suspending them in 5 ml PBS. The bacteria were pelleted by centrifugation at 12000 g during 10 minutes. The pellet was suspended 400 $\mu$l lysis buffer G (Stratec, Birkenfeld, Germany) and 10 $\mu$g/$\mu$l proteinase K was added. The suspension was incubated at $52°$C overnight. The lysed material was pelleted by 5 minutes centrifugation at 12000 g. A phenol-chloroform purification of the supernatans was performed twice. In

between RNAse A was used to degrade the RNA. DNA was precipitated using ethanol and resuspended in 50 $\mu$l MQ water. This resulted in a concentration of on average 1 $\mu$g/$\mu$l DNA.

The genomic DNA was digested with the restriction enzymes *BCl*I and *Xho*I which do not cut the transposon (Figure 4.1). The digest was purified using the GeneJet PCR purification kit (Thermo Scientific, Waltham, USA) and blunt end were obtained using T4 DNA polymerase (Invitrogen, Waltham, USA). The ends were allowed to self ligate with T4 DNA ligase (Promega, Madison, USA) (Figure 4.1).

The ligate was again purified with the PCR purification kit and eluted in a final volume of $10\mu$l. The whole ligate was subsequently used to electroporate EC 116-pir transformax *E. coli* cells (Epicenter, Madison, USA) (Figure 4.1). These cells express the pir gene, needed to recognize the R6K$\lambda$ori which is present in the transposon. Successful rescue clones were selected using LB-plates with $50\mu$g/ml kanamycin (Figure 4.1) and subsequently sequenced using the SqFP and SqRP primers (Figure 4.1,table 4.1). These primers bind at the extreme ends of the transposon and are directed to the outside of the transposon.

## 3.9   RATE PCR

Random amplification of transposon ends (RATE) PCR [191] is a three-step single prime PCR (Figure 4.1). The PCR reaction mixture contained 1 x PCR buffer, 1.5 mM MgCl$_2$, 400 $\mu$M dNTPs, 4 $\mu$M of each primer Inv1 and 1 u HotStarTaq Plus DNA polymerase (Qiagen, Hilden, Germany). The total reaction volume was 25 $\mu$l. In this PCR three steps each consisting of 30 cycles is performed. The first and the third step are performed at a stringent annealing temperature (55°C). At this temperature the single primer will only bind at its specific targeted location. The second step is performed at a low annealing temperature (30°C), which allows the aspecific binding of the single primer. This will thus allow nonspecific amplification of the products generated in the first step. The third step which uses again the stringent temperature will amplify both the specific and non-specific products generated in the previous steps. The amplicons of the RATE reaction were sequenced using the SqRP.

## 3.10   Whole genome sequencing

The mutants with atypical virulence were sequenced by whole genome sequencing. Genomic DNA extraction was performed by Invisorb DNA extraction kit (Stratec, Birkenfeld, Germany). Sequencing was performed by the Oklahoma medical research foundation NGS core (Oklahoma City, USA). The paired-end reads were generated by MiSeq Illumina sequencing. The read length was set at 150 and a 50x coverage was obtained.

The reads were analyzed in the CLC genomics workbench 9.5 (Qiagen, Hilden, Germany). They were mapped to the fully sequenced an annotated ERIC type I strain DSM 25719. Both the available contigs of BRL 230010 sequenced by Qin and colleagues in 2009 [77] and updated by Chan and colleagues in 2011 [76] were also mapped against this reference. Mapping the obtained reads against the transposon sequence allowed to check whether the transposon was present in the genome and which genes were disrupted. Finally structural variances analysis was used to compare all mutations present in the mutant genomes

(insertions, deletions, replacements, inversions and translocations) and to identify genes affected by mutagenesis.

## 3.11  Growth characteristics

A difference in virulence can be an indirect effect from a change in growth characteristics. Growth curves were constructed in triplicate to compare whether the growth of the mutants were influenced by the transposon insertion. Each mutant was cultivated in 3 ml MYPGP. The optical density at a wavelength of 600 nm ($OD_{600}$) was measured and 100 ml MYPGP was spiked to obtain a final OD of 0.01. Serial dilution of the inoculate confirmed that each culture was spiked with the same amount (2.7 E9 cfu) of colony forming units (cfu). The $OD_{600}$ of the cultures was subsequently monitored every hour for 20 h. Every 4 hours the presence of contaminants was checked by plating. By modeling the exponential phase of the curves, the difference in growth rate can be statistically analyzed.

# 4  Results

## 4.1  Optimization electroporation

As mentioned in the methods section, the addition of 5% larval fluids significantly reduced the lag phase. The cells were harvested at $OD_{600}$ of 0.3 and subsequently washed according to the steps described in the protocol of Murray and colleagues ([80]). Several voltages were tested on these electrocompetent cells. Electroporation was most successful at a voltage of 1.5 kV. This voltage was used throughout the rest of the experiments.

## 4.2  Temperature sensitive transposon mutagenesis

After successful electroporation with the shuttle plasmid, temperature dependent transposon mutagenesis was attempted. Plasmid pID408 was used throughout these experiments. It carries a transposon Tn917 and a temperature sensitive transposase and has been used previously to mutagenize *Bacillus spp.* This electroporation failed. *P. larvae* was not able to recover at the temperature of 35°C that was needed with this temperature sensitive transposase.

## 4.3  *In vitro* transposition

Subsequently *in vitro* transposition with Himar1 mariner transposase and pJFP1 was performed. For this transposition method the transposase was obtained from prof. Kitten who also provided the protocol they used for *in vitro* transposition in *Bacillus spp.*. The *in vitro* transposition was checked with a Southern blot and a labeled probe against the transposon. Only the original plasmid pJFP1 was detected, so the transposase Himar1 was not active.

## 4.4   EZ-Tn5 Transposome

Electroporation using EZ-Tn5 transposome mutagenesis resulted in a library of 158 kanamycin resistant mutants. Successful sporulation was only obtained for 89 mutants. A prescreening of the virulence of these 89 mutants resulted in 11 to be tested in detail and of which 7 were statistically confirmed being less virulent. The other four mutants had no significant difference in virulence compared to the wild type. The result of the identification of these mutants is described further. The name given to each mutant is deducted from the altered gene that has the highest probability to be responsible for the observed phenotype (section 4.5).

Seven mutants have a significant difference at the $5\%$ significance level in percentage of dead larvae when administered the $LD_{80}$ of the wild type BRL230010, even after correcting for multiple testing (Figure 4.3,Table 4.2). The wild type has a mean mortality of 0.771 (sd = 0.25). Mutant BRL230010$cirD$:: results in the highest decline in virulence, with a mean mortality of 0.159 (sd = 0.138, p-value = 0.004). Followed by BRL230010$ywqD$:: which has a mean mortality of 0.175 (sd = 0.189, p-value = 0.011). BRL230010$\triangle X$, BRL230010$IN(gbpA)$, BRL230010::$\triangle ERIC1\_1c21190$ and BRL230010$lysC$:: with a mean mortality of respectively 0.238 (sd = 0.150, p-value = 0.009), 0.254 (sd = 0.125, p-value = 0.008), 0.288 (sd = 0.185, p-value = 0.023) and 0.307 (sd = 0.197, p-value = 0.032. Finally BRL230010::$\triangle dnaK$ showed the least difference in virulence, with a mortality of 0.404 (sd = 0.143, p-value = 0.039).
The mortality curves for these mutants were statistically compared with the wild type using Cox proportional hazard regression analysis. All mutants have a significant difference in mortality rate compared with the wild type (Figure 4.3,Table 4.2 ). The log hazard rate relative to BRL230010 decreases for mutant BRL230010$cirD$:: with -1.634 (se = 0.238, p-value = 6.2E-12), BRL230010$ywqD$:: with -1.563 (se = 0.209, p-value = 1.0E-12). For the mutants BRL230010$\triangle X$, BRL230010$IN(gbpA)$, BRL230010::$\triangle ERIC1\_1c21190$ and BRL230010$lysC$:: a decrease of respectively -1.185 (se = 0.306, p-value = 7.5E-9), -1.162 (se = 0.313, p-value = 3.2E-10), 1.114 (se = 0.328, p-value = 7.1E-10) and -1.079 (sd = 0.340, p-value = 0.5.3E-9). Finally BRL230010::$\triangle dnaK$ has the lowest decrease in log hazard rate with- 0.673 (se = 0.510, p-value = 4.2E-5) was determined.

Table 4.2: **Difference in virulence.** 1) T-test were performed to statistically confirm the difference in virulence of each mutant with the wild type BRL230010. In this table the mean mortality, standard deviation (sd) and p-value of the t-test is given. 2) A Cox proportional hazard model was fit to compare the mortality curves. The table gives the coefficient difference in log hazard rate compared to the wild type ($\triangle$log(hazard rate)), its standard error (se) and p-value.

| Name | T-test | | | Cox proportional hazard | | |
|---|---|---|---|---|---|---|
| | Mean mortality | sd | p-value | $\triangle$log(hazard rate) | se | p- value |
| BRL230010 | 0.771 | 0.250 | | | | |
| BRL230010$cirD$:: | 0.159 | 0.138 | 0.004 | -1.634 | 0.238 | 6.2E-12 |
| BRL230010$ywqD$:: | 0.175 | 0.189 | 0.011 | -1.563 | 0.219 | 1.0E-12 |
| BRL230010$\triangle X$ | 0.238 | 0.150 | 0.009 | -1.162 | 0.185 | 3.2E-10 |
| BRL230010$IN(gbpA)$ | 0.254 | 0.125 | 0.008 | -1.185 | 0.205 | 7.5E-9 |
| BRL230010::$\triangle ERIC1\_1c21190$ | 0.288 | 0.185 | 0.023 | -1.114 | 0.181 | 7.1E-10 |
| BRL230010$lysC$:: | 0.307 | 0.197 | 0.032 | -1.079 | 0.185 | 5.3E-9 |
| BRL230010::$\triangle dnaK$ | 0.404 | 0.143 | 0.039 | -0.673 | 0.164 | 4.2E-5 |
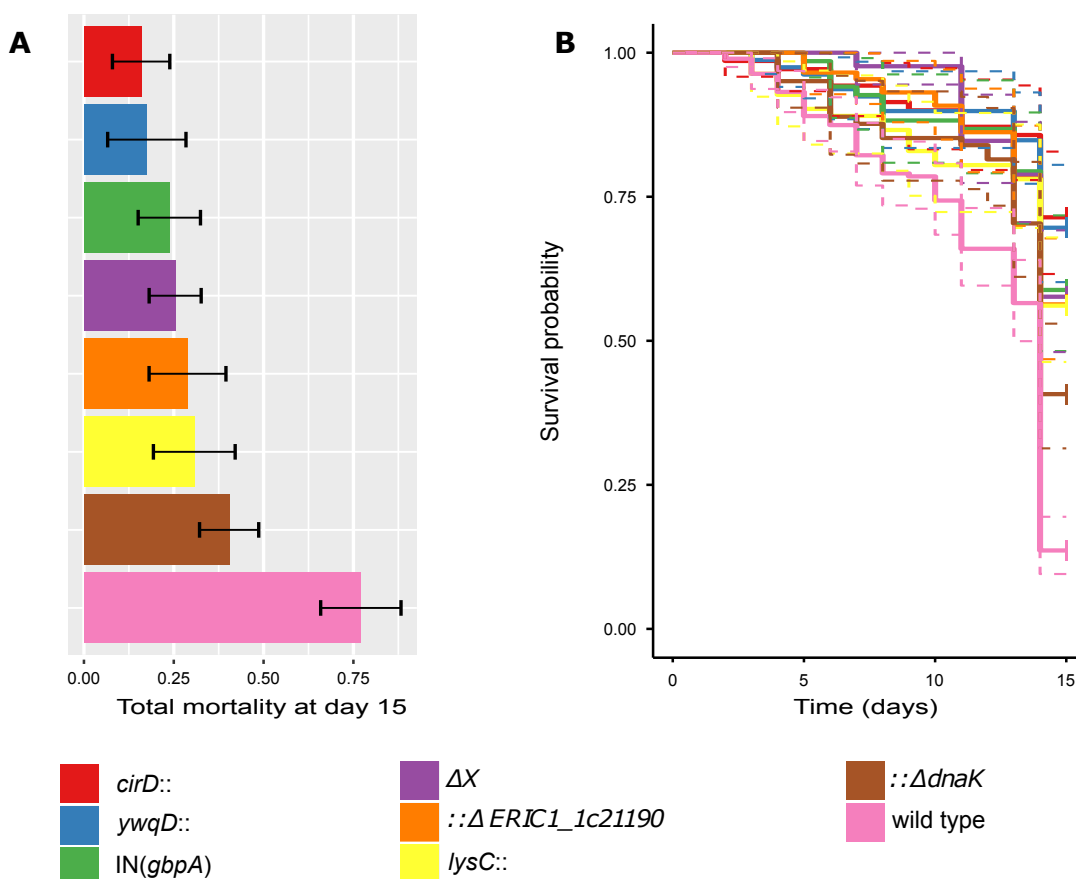
Figure 4.3: **Mortality** A) All selected mutants have a significant lower percentage of total amount of larvae that were dead at day 15 at the $LD_{80}$ (dose needed to kill $80\%$ of larvae) of the wild type BRL230010. Error bars represent standard deviation. B) The Kaplan-meier curves of the mortality rate show that all mutants have a significant slower mortality rate than the wild type. The $95\%$ confidence interval is displayed using dotted lines.

## 4.5   Identification

Rescue cloning was the first method used to identify the disrupted genes, however no successful clones were obtained. RATE-PCR also failed in locating the transposon. Finally whole genome sequencing of the mutants was performed. To check whether the transposon was present in the genome, the reads were mapped against the transposon sequence. The transposon could however not be identified in any of the 7 genomes. Subsequently the reads were mapped to the reference *P. larvae* genome and a structural variance analysis was performed. This resulted in the identification of several structural alterations in the genomes from the different mutants compared to the wild type genome of BRL230010. Every affected locus in the mutant genomes was analyzed. Alterations in non-coding loci were discarded. All but one mutant were affected in at most 9 coding loci, which makes a thorough discussion and selection of probable candidate virulence factors possible. One mutant was however altered in to many coding regions to make a credible selection and is referred to as BRL230010$\Delta X$.

The nomenclature of the mutants is according the international standards, in which insertions are defined by the symbol " :: ", deletions by "Δ" and inversions by "$IN$". A recombination is essentially a combination of an insertion and deletion and is accordingly referred to with the symbol " :: Δ". This nomenclature can be found at $http$ : $//jb.asm.org/site/misc/journal - ita\_nom.xhtml$.

BRL230010::Δ*dnaK* has a total of 9 mutations in coding regions. The mutant has 5 replacements: 21 basepairs in *dnaK*, 14 basepairs in *wecB*, 10 basepairs in *porC*, 51 basepairs in *sMC* and 214 bp in *prkA*. It also has 4 complex mutations, which are a combination of inversions and insertions: insertions of 590084 bp over a region of 99 bp with a breakpoint at the extreme end of *dacB* and in *spmA*, insertions of 140783 bp over a region of 101 bp in *ERIC1_2c03180* which encodes for a ATPase and insertions of 140783 bp over a region of 101 bp in *UxaC* and 99 bp in *ERIC1_2c07220* which encodes for an oligoendopeptidase of the pepF/M3 family.

BRL230010*ywqD*:: has also mutations in 9 coding regions. Six of the genes have insertions which all are all due to tandem duplications: *ywqD*, *MetN*, *ERIC1_1c07870* which encodes for an ABC efflux transporter-like protein, *nuoC*, *pdhA* and finally *ERIC1_1c28620* which encodes for a hypothetical protein. Using the InterPro online program (EMBL-EBI, $http$ : $//www.ebi.ac.uk/interpro/$) the hypothetical protein was analyzed. The program is able to identify the protein family membership, domains and repeats and GO term prediction. Five transmembrane regions were found in the protein, but no protein family or GO term prediction was possible. A possible function for this hypothetical protein could not be suggested. Besides these 6 insertions, the mutant has also a replacement in *rfbG* of 39 bp, a deletion of 497 bp partially affecting *sdaC* and removing hypothetical protein *ERIC1_2c02780*. InterPro only found signal peptide domains in the hypothetical protein, but no function could be predicted. Finally a inversion of 30 bp is present in *smc*.

BRL230010*lysC*:: is again affected in 9 coding regions. Insertions affect 4 regions, as in the mutant BRL230010*ywqD*:: these insertions are due to tandem duplications. The affected genes are *ERIC1_1c11620* and *ERIC1_2c02840* both encoding for a hypothetical protein, *ERIC1_1c32770* encoding for a putative membrane protein and *lysC*. InterPro found for the ERIC1_1c11620 hypothetical protein a structural domain with a 3-layer alpha/beta/alpha structure that contains mixed beta-sheets and represents a structural domain found in several acyl-CoA acyltransferase enzymes. For hypothetical protein ERIC1_2c02840 InterPro found a non-cytoplasmatic domain and a signal peptide domain, but could not identify a protein family nor a possible function. Three replacements can be found in the mutant: 46 bp in *norM*, 207 bp in *ydcC*, 20 bp in *serC* and 28 bp in hypothetical protein *ERIC1_2c02970*. A signal peptide domain could be identified in this hypothetical protein, but again no function could be attributed. Finally two complex mutations are present with a inversion of 99 bp and insertion of 510249 bp, one affecting the transposon encoding gene *trpA* and the other affecting *ERIC1_2c06910* with encoding for a two component transcriptional regulator of the AraC family.

BRL230010*cirD*:: has six mutations in coding regions. One is a deletion of 1388 removing *tnp16*. Four coding regions are affected by replacements: 10 bp in *rplY*, 13 bp in *gyrA*, 28 bp in *msmX1* and 16 bp in *ERIC1_1c17630* which encodes for a hypothetical protein. InterPro identified a signal peptide region and a non-cytoplasmatic region in this protein but was again not able to identify a possible function or protein family. Finally a complex mutation of insertions of a total of 42222 bp affects a 99 bp region at the extreme end of *cirD*.

BRL230010$IN(gbpA)$ has only two mutations affecting coding regions. An inversion of 14121 bp has a breakpoint interrupting *gbpA*, a complex mutation affects *ERIC1_1c01290* which encodes for a transposase.

Finally BRL230010::$\Delta ERIC1\_1c21190$ has mutations in two loci encoding for a hypothetical protein, a replacement of 38 bp in *ERIC1_1c21190* and a complex mutation of multiple deletions over 200 bp affecting *ERIC1_1c15570)*. According the InterPro ERIC1_1c21190 represents a type 2 phosphatidic acid phosphatase enzyme, while ERIC1_1c15570 belongs to the family of cysteine protease Prp.

## 4.6 Growth Characteristics

A difference in virulence can also be due to a difference in growth characteristics. To see whether the mutations led to a difference in growth, the growth rate of each one was compared with the wild type. Modeling the exponential phase of the growth curves with the wild type BRL230010 as reference gave following model where t = time in hours from start exponential phase, T = Matrix of mutant parameter estimates (Table 4.3). The model has an adjusted R$^2$ value of 0.712 and a p-value of $< 2.2E - 16$.

$$OD_{600} \sim 0.159 + 0.037t + T$$

The results from the model are also reflected in the growth curves (Figure 4.4). The slowest growing mutant is BRL230010::$\Delta ERIC1\_1c21190$, which has a parameter estimate of -0.284 (SE = 0.024, p-value = $< 2E - 13$). Mutants BRL230010*lysC*:: and BRL230010::$\Delta dnaK$ grow only slightly faster and have a parameter estimate of -0.278 (SE =0.024, p-value = $< 2E - 16$) and $-0.265$ (SE = 0.024, p-value = $< 2E - 16$) respectively. They are followed by BRL230010*cirD*:: which has a parameter estimate of $-0.214$ (SE = 0.024, p-value = $< 2E - 16$). The other three have a growth rate that is very close to each other. BRL230010$\Delta X$ has a parameter estimate of $-0.179$ (SE = 0.024, p-value = $9.51E - 13$), followed by BRL230010*ywqD*:: with a parameter estimate of $-0.175$ (SE = 0.024, p-value = $2.32E - 12$). BRL230010$IN(gbpA)$ is the fastest growing mutant, with a parameter estimate of $-0.172$ (SE = 0.026, p-value = $1.46E - 10$).

Table 4.3: **Parameter estimates growth model.** The parameter estimates of the growth model are given with their standard error and p-value.

| Parameter | Estimate | Standard Error | p-value |
|---|---|---|---|
| Intercept | 0.159 | 0.020 | 1.14E-14 |
| Time (t, in hours) | 0.037 | 0.0015 | <2E-16 |
| BRL230010::$\Delta dnaK$ | -0.265 | 0.024 | <2E-16 |
| BRL230010$\Delta X$ | -0.179 | 0.024 | 9.51E-13 |
| BRL230010*ywqD*:: | -0.175 | 0.024 | 2.32E-12 |
| BRL230010*lysC*:: | -0.278 | 0.024 | <2E-16 |
| BRL230010*cirD*:: | -0.214 | 0.024 | <2E-16 |
| BRL230010$IN(gbpA)$ | -0.172 | 0.026 | 1.46E-10 |
| BRL230010::$\Delta ERIC1\_1c21190$ | -0.284 | 0.024 | <2E-16 |
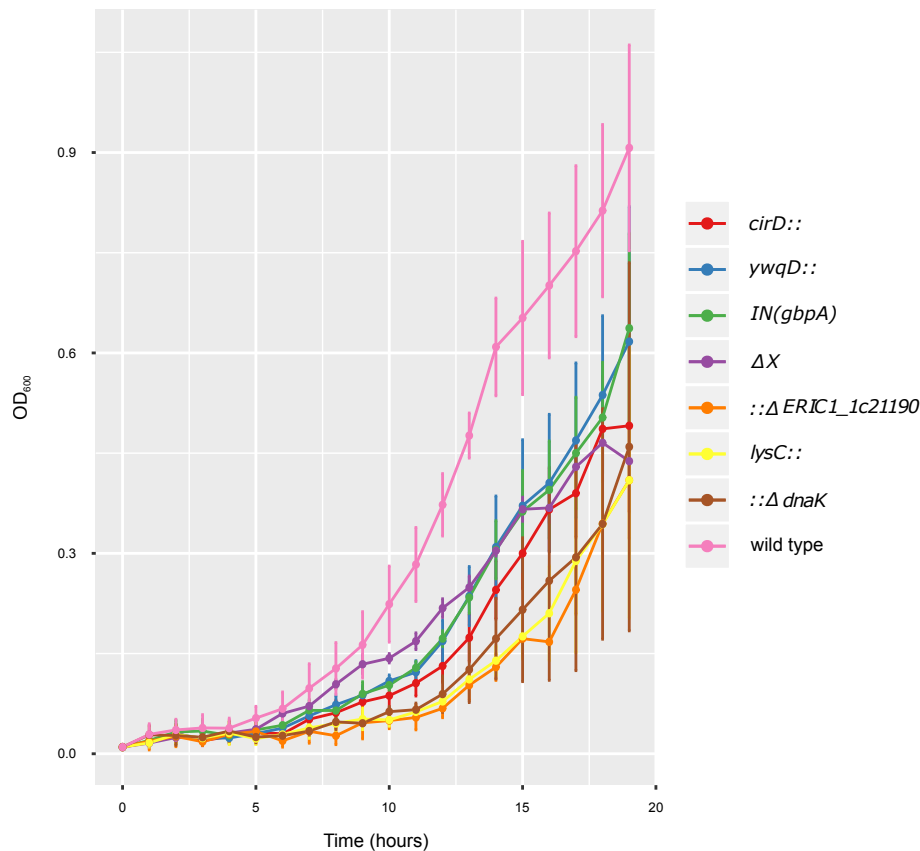
Figure 4.4: **Growth curves.** The evolution of the $OD_{600}$ in function of time represents the growth rate of the different mutants. All have a significant slower growth rate than the wild type strain. Error bars represent the standard deviation at each time point.

# 5  Discussion

The results of the whole genome sequencing showed that the EZ-Tn5kan$^R$R6K$\lambda$ transposon was not present in the genome of any of our mutants. However a library of 158 kanamycin resistant mutant *P. larvae* colonies was created of which 7 showed a significant attenuation of both virulence and growth rate. It is possible that the detected mutations are due to an unstable insertion of the EZ-Tn5kan$^R$R6K$\lambda$ transposon or to spontaneous mutations caused by the presence of kanamycin. The recovery time of 16h is exeptionally long, which has also been reported by Murray in the initial protocol for electroporation of *P. larvae* [80]. This recuperation time might give an overestimate of transformation efficiency sinc colonies arising on the plates could be due to reproduction of the transformed cells during recovery. After recovery, the cells have to grow during 3 days on selective media, due to the slow growth rate of *P. larvae*. This can result in non-transformed cells obtaining kanamycin resistance due to the presence of spontaneous mutations in one or several genes.

Of 6 mutants, a thorough discussion of the affected genes and their possible function follows. The 7th mutant was altered in too many regions to make a credible selection possible. A summary of all findings is given in Suppl. Table S2. These affected genes could play a role in virulence, but their function should be confirmed in a follow-up study.

## 5.1 Virulence associated factors

Whole genome analysis of the mutants pointed to affected genes that were known to encode virulence factors in other bacteria and are therefore regarded as likely cause of the observed attenuation in virulence.

Mutant BRL230010*cirD*:: has only one affected gene – *cirD*- that is known to influence virulence. This gene has been suggested to encode for an enzyme involved in the processing and/or circularization of the circularin A prepeptide [192]. Circularin A is a bacteriocin produced by *Clostridium beijerinckii* [193] and is able to kill lactococci, enterococci and some *Lactobacillus* and *Clostridium* strains. Attenuated virulence could result from less successful or slower colonization of the gut due to a lack of expression of this bacteriocin (Figure 4.5).
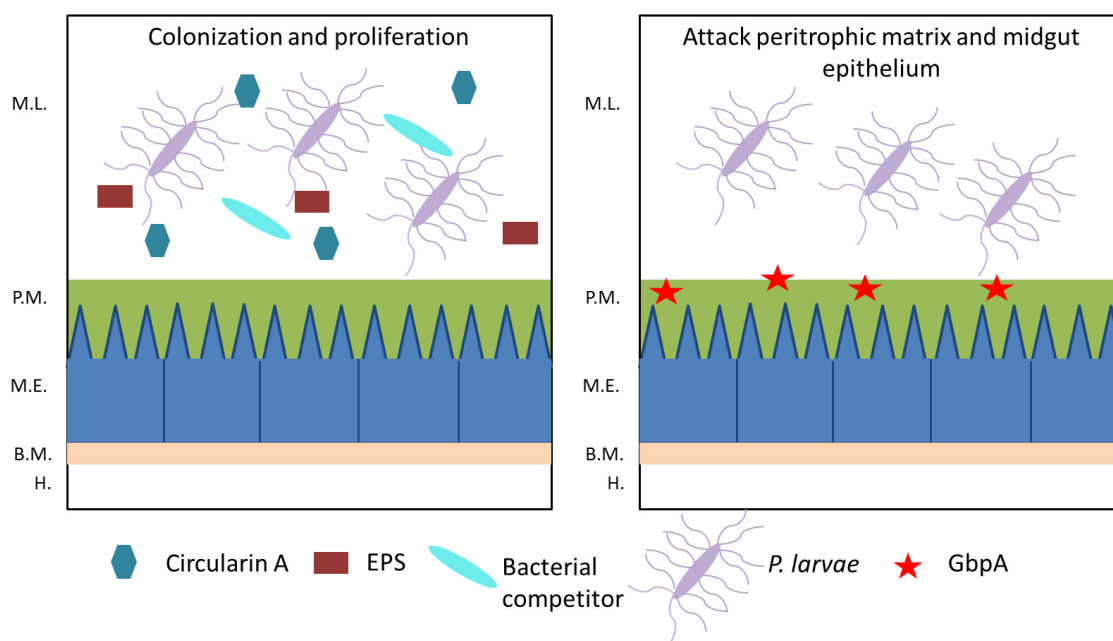


Figure 4.5: **Possible virulence factors.** Different mutated genes code for or are associated with the synthesis of virulence factors known in other bacteria. These factors include the bacteriocin Circularin A, the biofilm component EPS and the chitin and mucin binding protein GbpA. Circularin A is probably expressed during the colonization phase, where they kill competitive bacteria. Possibly biofilm formation is also a process during this colonization and EPS is expressed during this phase. In the next phase of infection, *P. larvae* has to break down the peritrophic membrane and attack the epithelial cells. GbpA is probably expressed during this phase since it contains a chitin and mucin binding domain. Chitin and mucin-like proteins are two main components of the peritrophic matrix. Abbreviations: M.L. = midgut lumen, P.M. = peritrophic membrane, M.E. = midgut epithelium, B.M. = basale membrane, H. = hemocoel.

During the infection process of *P. larvae* the gut epithelium is attacked. In ERIC II *P. larvae* strains the attachement to the epithelial cells is conducted by the S-layer protein [3]. This protein is however not functionally present in ERIC I genotypes. Mutant BRL230010*IN*(*gbpA*) had an inversion disrupting gene *gbpA*. GbpA is a known virulence factor of *Vibrio cholerae*, the protein is a common adherence factor for chitin and mucin [194]. It is possible that

GbpA has a similar role as the S-layer protein. During infection *P. larvae* has to break down the protective peritrophic matrix in order to reach the gut epithelium. The peritrophic matrix contains high levels of chitin and mucin-like proteins [195]. In *P. larvae Pl*CBP49 is responsible for degrading the peritrophic membrane, this is another possible function for GbpA (Figure 4.5).

Mutant BRL230010*ywqD*:: has one affected gene that is known to contribute to virulence in other bacteria. The affected protein is YwqD, a protein-tyrosine kinase which activates YwqF, a UDP-glucose dehydrogenase (UDP-Glc DH). Both are encoded by genes of the *ywqCDEF* operon. UDP-Glc DH helps in the production of UDP-glucuronate (UDP-GlcA) which is a precursor for the synthesis of acidic type exopolymeric substances (EPS) in *B. subtilis*. These EPS play an important role as structural components of the extracellular matrix during biofilm development. Because of this reason they contribute to virulence and tolerance to physiological stress in *B. subtilis*, a similar role in *P. larvae* could be possible (Figure 4.5). Since the *ywqCDEF* operon controls the EPS synthesis, altered activity of the encoded proteins is known to affect virulence [196].

## 5.2  Survival associated factors

Some of the identified factors are genes stronly associated with survival of the bacteria. Affecting these genes can therefore result in the observed attenuation in virulence. Disruption of *lysC*, which encodes an enzyme involved in DAP biosynthesis, results in a decreased fitness among several bacteria. In *Salmonella enterica*, *Shigella flexneri* and *Legionella pneumophila* intact *lysC* is required for survival in the host [197]. In these organisms DAP is part of the cell wall peptidoglycan and not only involved as a part of the lysine biosynthesis. Most *Paenibacillus spp.* including *P. larvae* have the DAP type peptidoglycan [198]. Since BRL230010*lysC*:: is not completely avirulent, LysC is probably not completely nonfunctional. Another gene affected in this bacteria is *serC*, encoding an enzyme of the serine biosynthesis pathway. In the Gram-negative *Burkholderia pseudomallei* disruption of the gene resulted in a significant decrease of virulence due to a decreased fitness compared with wild-type [199], this is probably due to a diminished fitness.

BRL230010::Δ*dnaK* has a replacement affecting the gene *dnaK*. DnaK is a chaperone protein that functions in the heat-shock system. In Staphylococcus aureus mutants a nonfunctional DnaK system resulted in a significant decrease in survival [200], due to an altered transcriptional regulation. It is possible that the affected *dnaK* in *P. larvae* results indirectly to a lower virulence and kanamycin resistance by differential expression of involved genes. Another affected locus is a three gene operon formed by *dacB*, *spmA* and *spmB*. DacB encodes for the penicillin-binding protein 5. Mutations affecting one of these genes resulted in *B. subtilis* in lower heat resistant spores [201]. However, an attenuated heat resistance of the spores in our experiments was not observed as the spores of BRL230010::Δ*dnaK* were heated at 80 °C during 10 minutes to determine the spore suspension concentration. Penicillin-binding-protein 5 is also an important building block of the cell wall in Gram-positive bacteria. Deficiency of the Dac proteins interferes with the architecture of the cell wall and led to lower survival of *Streptococcus pneumonia* [202]. The gene *sMC* is also affected in this mutant. sMC forms together with ScpAB an important chromosome-organizing machine in bacteria[203, 204]. In *Pseudomonas aeruginosa* it was proven that the inactivation of sMC induces defects in competitive growth and colony

formation [205]. A final affected survival associated gene is *uxaC*. UxaC proteins are part of the Ashwell pathway which provides compounds for the pentose phosphate pathway and citrate cycle. The UxaC protein is thus important for energy metabolism [206]. A deficient UxaC protein could result in the observed attenuated growth rate.

Mutant BRL230010*ywqD*:: has an insertion in *MetN*. This gene encodes for an ATPase and is part of the Methioning ABC (ATP binding cassette) Uptake Transporter (MUT) family [207]. Methionine synthesis is essential for survival of many pathogens during infection and methionine biosynthetic genes are required for the full fitness and associated virulence of several pathogens as *Brucella melitensis* [208], *Haemophilus parasuis* [209] and *Salmonella enterica* [210]. The affected gene *RfbG* encodes CDP-D-glucose-$4, 6$-dehydratase. In *Azotobacter vinelandii* disruption of this gene has been shown to result in drastic reduction of growth rate [211]. *ERIC1_3c0050* encodes for another sMC protein, as mentioned earlier these class of condensines are known to play a role in virulence due to attenuated growth [205].

MsmX is an ATPase required to energize different saccharide transporters and is affected in BRL230010*cirD* ::. A deletion mutant in *B. subtilis* has been shown to exhibit a severe decrease in growth when only nonlinear arabinooligosaccharides were present as carbon source [212]. When other sugars were present, no significant difference in growth compared to the wild type was observed.

## 5.3 Antibiotic resistance factors

The observed kanamycin resistance of BRL230010*ywqD*:: could be due a gain-of-function mutation in *ERIC1_1c07870*, which encodes for a ABC efflux transporter-like protein. Efflux pumps are important players in antibiotic resistance [213].

*NorM* encodes a multidrug transporter [214]. The alterations of this gene in BRL230010*lysC*:: could be responsible for the kanamycin resistance phenotype as has been shown for *Vibrio parahaemolyticus* [215]. Disruption of *rplY*, as observed in BRL230010*cirD* ::, is known to increase resistance against aminoglycosides including kanamycin [216]. Another affected gene in this mutant is *gyrA*. A non-functional *gyrA* gene has been associated with induced virulence in Gram-negative bacteria due to the elevated expression of the type III secretion system genes which are involved in virulence [217]. Gram-positive bacteria as *P. larvae* do not contain this type of secretion system. In these Gram-positive bacteria a non-functional *gyrA* gene is associated with resistance to quinolones [217]. *P. larvae* is however a bacterium known to be resistant to quinolones as nalidixic acid.

## 5.4 Other factors

The other affected genes have no known influence on virulence or survival in other bacteria is known. In Gram-negative bacteria the *wecB* gene is an important contributor to virulence due to its involvement in the enterobacterial common antigen synthesis. However, this antigen is not synthesized in Gram-positive bacteria including *P. larvae*, where it fulfills a role as a UDP-N-acetylglucosamine 2-epimerase [218]. *PorC* is part of the PorCDAB tetrameric protein and encodes for a distinct pyruvate-flavodoxin oxidoreductase which mediates electron transport to NADP [219]. *PrkA* encodes in *Bacillus subtilis* for a serine

protein kinase [220]. In *Mesorhizobium alhagi* the enzyme plays a role in stress resistance [221], but a role in antibiotic stress was not tested. *ERIC1_2c03180* encodes for an ATPase of the AAA family. Different genes encoding for a protein of this enzyme family are present in the genome. *NuoC* is a subunit of the prokaryotic membrane-bound proton-translocating NADH-quinone oxidoreductase (NDH-1) [222]. SdaC is a secondary carrier for threonine and serine, belonging to the *sdaCB* operon [223]. The protein is known to be a specific inner membrane receptor for the antibiotic Colicin V [224]. In *Bacillus subtilis* the gene *ydcC* contributes to sporulation [225]. Disruption of this gene reduced the sporulation efficiency. Sporulation efficiency was not checked in our strain, so it is unknown if in *P. larvae* disruption also leads to less efficient sporulation.

BRL230010*cirD* :: contains a deletion removing *tnp16* encoding for a transposase that controls the insertion sequence element IS256 in transposon Tn4001. Transposons can become active as a response to stress [226] and can induce mutations. BRL230010$IN(gbpA)$ also contains a mutation in *ERIC1_1c01290* which encodes for a transposase. Hypothetical proteins could also be of importance as possible new *P. larvae* ERIC type I specific virulence factors. Herewith they are good candidates for site-specific mutagenesis to confirm their possible function as virulence factors. The hypothetical protein ERIC1_2c02780 could not be attributed a possible function by InterPro. *ERIC1_2c06910* encodes for a two component transcriptional regulator of the AraC family which contains more than 100 proteins. The hypothetical protein ERIC1_1c11620 contains a structural domain found in several acyl-CoA acyltransferases. *ERIC1_1c32770* encodes for a putative membrane protein. Hypothetical protein ERIC1_2c02840 contained a non-cytoplasmatic domain and a signal peptide domain and hypothetical protein ERIC1_2c02970 contained a signal peptide. The locus *ERIC1_1c17630* contains a signal peptide and a non-cytoplasmatic region as identified by InterPro, but no function could be suggested.

BRL230010::$\Delta ERIC1\_1c21190$ was found to carry only two mutations in total, which both affecting hypothetical proteins. Using InterPro a possible function and gene family was suggested. ERIC1_1c15570 was attributed to the cysteine protease Prp family. This family cleaves the N terminus extension of ribosomal subunit L27 in bacteria [227]. ERIC1_1c21190 was attributed by InterPro to the type 2 phosphatidic acid phosphatases (PAP2) enzymes. To this family belongs also the antibiotic bacitracin transport permeases found in *Bacillus licheniformis*. A possible role in the observed kanamycin resistance is thus possible for this gene.

# 6   Acknowledgement

# Chapter 5

# General discussion

## 1 Experimental results

In this doctoral research, two different topics concerning *Paenibacillus larvae* were covered. The first topic considered the development of a new genotyping tool, while the second aimed at providing more insight in the virulence mechanism used by the pathogen. In this section a general discussion concerning these topics is given. First we focus on the scientific relevance of the results, subsequently the implementation of this thesis in controlling AFB is discussed.

### 1.1 Genotyping

Currently the most important genotyping technique for *P. larvae* is without doubt ERIC-rep PCR [2]. The reason this genotyping method is used so frequently is the biological relevance of its groups. The ERIC genotyping method divides the species *P. larvae* in four groups (ERICI-IV) [2, 71]. Two of these genotypes (ERIC I and II) can be detected in the field, these genotypes differ in virulence both on the individual as on the colony level [72].

In a hive, nursing bees express hygienic behavior, which means they are able to recognize and remove diseased larvae. ERIC II is very virulent on the individual level and larvae die before their cells are capped. Since these brood cells are not closed with wax yet, nursing bees recognize the dead larvae fast and are able to remove them before millions of spores are formed. This limits the spread of spores in the hive and results in a lower virulence on colony level. In contrast ERIC I is less virulent on the individual level, which results that most larvae die after their brood cells have been capped. Recognition of these dead larvae by nursing bees is more difficult, and by the time the dead larvae are detected and removed millions the ropy stage and hard scales containing millions of spores has been formed. The nursing bees are covered with these spores when removing the larval remains and spread them very efficiently thorough the entire hive. A hive infected with an ERIC type I strain will more often show the typical clinical signs, like the hail shot pattern and a positive match stick test, which will make it easier to detect by the beekeeper or the veterinarian. Because of these reasons, most detections of AFB outbreaks are caused by ERIC I, while ERIC II strains are more often detected in honey instead of in diseased brood. The last two ERIC genotypes (ERIC III and IV) are not detected in the field anymore. These

genotypes are very virulent on the individual level, making detection by nursing bees very easy. If these genotypes still exist in nature, this high virulence and efficient detection by the nursing bees could explain why we never detect it in clinical outbreaks [72].

ERIC genotyping is not only important due to its biological relevance, but also because of its historical relevance. Historically ERIC I and II formed a seperate species (*Bacillus larvae*) [7] and ERIC III and IV formed the species *Bacillus pulvifaciens* [17]. These pathogens were believed to cause different diseases, AFB and powdery scale disease respectively. Later these species were reclassified under the genus *Paenibacillus* [19] and subsequently under the same species *Paenibacillus larvae* but with subspecies differentiation [21]. Only in the last reclassification round in 2006 this subspecies differentiation was discarded.

The ERIC genotyping technique has however some important limitations, it is only able to separate the species in four groups, making it unsuitable for epidemiological purposes. Other genotyping techniques exist that separate *P. larvae* in multiple groups, however these are often not suitable to use in a routine diagnostic lab. One of the main goals of this thesis was to develop a genotyping technique that is sensitive, fast and cheap, but also contains the advantages of the ERIC-genotyping technique. Multiple Locus Variable number of tandem repeats Analysis (MLVA) was developed to meet these requirements. MLVA is a technique has been implemented several bacteria. An overview of these bacteria can be found in a summarizing table (Table S3).

Since we developed MLVA in such a way that the ERIC-groups were still easy to recognize, the method will be biased by design. We purposely included VNTR loci that differed between ERIC-groups, but to our knowledge were conservative within their ERIC-genotype. Other biases in the technique are due to our requirement to be able to separate the amplicons on agarose gel. Another option is to separate the amplicons on capillary electrophoresis, which makes a broader range of amplicons possible. Moreover with capillary electrophoresis, fluorescent labeling of the different primers is possible, making a more sensitive analysis of the obtained pattern able. We wanted however to develop a technique that could be used in routine diagnostics worldwide, thus also in laboratories with only basic equipment. With this in mind the choice for agarose gel based electrophoresis makes sense. A last bias is the almost exclusive use of Belgian strains in the development of the technique. This could make the method overfitting for Belgian or European strains, but not pose this specificity towards strains from other regions. In a later phase these strains were extrapolated with Italian ERIC II strains on which the method proved to be equally discriminatory. When in the future the technique proves not to be discriminatory enough towards exotic strains, adding a sixth VNTR locus should solve this problem.

During the thesis other research groups also focused on the development of new typing methods, namely MALDI-TOF MS [56] and MLST [65]. Similarly to MLVA, these techniques allow high-throughput analysis of strains and are very sensitive. MALDI-TOF MS was shown to succesfully discriminate the ERIC genotypes and was able to cluster strains within them. However the use of the method in epidemiology remains to be tested [56]. As with MLVA, MLST is able to show the evolutionary distances between the strains and subdivide the ERIC genotype groups [65]. The MLST and MLVA method provide more or less the same discriminatory power, with MLST dividing 173 ERIC I isolates in 16 subgroups and 92 ERIC II isolates in 3 subgroups, while MLVA divided 189 ERIC I isolates in 17 subgroups and 18 ERIC II strains in 4 subgroups. MLVA is has the advantage of being

more suitable for use in routine diagnostics, since it does not require high-end techniques or equipment such as sequencing.

## 1.2   Virulence mechanism

As mentioned in the above section, different strains of *P. larvae* exhibit different virulence. Although this has been known for more than a decade, the underlying molecular mechanisms remain mostly obscure [1, 2]. Some research groups have focused on the difference in virulence between ERIC I and ERIC II to identify virulence factors, while others used comparative genomics as a starting point. The aim in this thesis was to identify virulence factors by using an untargeted approach. To accomplish this several mutagenesis techniques were used, these included temperature sensitive transposon mutagenesis using plasmid pID408 and transposon Tn917, *in vitro* mutagenesis using the Himar1 transposase and EZ-Tn5 transposon mutagenesis. Although successfull virulence-impaired mutants were created after implementing the latter technique, the transposon was not detected in these strains. It remains unknown whether these mutant strains were caused by the transposon insertion and subsequent removal (so-called unstable transposition) or were caused by spontanuous mutagenesis induced by electroporation-stress and kanamycine-stress. When the EZ-TN5 transposon would have caused the mutations, a fingerprint would have been left. Tn5 is part of the "cut-and-paste" transposons, where the transposon is cleaved from its original position and pasted into a new target site [228]. This process typically leads to the introduction of tandem duplications. Although tandem duplications have been found in our results, they seem too long to be the result of transposon activity. Kanamycin-induced mutagenesis has been used as a mutation technique for *Bacillus subtilis* [229]. These kind of mutations belong to the class of stress-induced mutations and are of importance in adaptive evolution [230]. In the electroporation protocol a recuperation phase of 16h is followed by a 3 day incubation on selective media. Although this selective growth of 3 days is common used in other mutagenesis protocols of *P. larvae* [79, 80], it significantly increases the risk of obtaining these stress-induced mutants.

In the search for virulence factors several biases can be detected. The first and foremost bias is the use of our mutagenesis technique. Other techniques could be used for this purpose and would undoubtly give other results, however all methods would suffer from biases. For example UV mutagenesis is a technique that can be used for untargetting mutagenesis, but is also biased towards mutating CT to TT [231]. Other methods using chemical compounds suffer from similar biases. A second bias in our setup was the choice of the strain BRL 230010 to be mutated. At the start of the thesis, this choice was motivated because it was the only *P. larvae* strain that was sequenced [77, 105]. In the meantime other strains have been sequenced to higher depth and sequencing cost has decreased to the point that whole-genome sequencing can be done routinely. BRL 230010 is a ERIC type I genotype (MLVA type 2) and was first isolated from a clinical AFB colony in the USA. The strain was donated to the laboratory by Dr. Queenie Chan (University of British Columbia, Canada) and we started from this original aliquot in order to minimize subcloning and to conserve the original virulent strain as good as possible.

As mentioned before there is a difference in virulence between ERIC I and ERIC II strains, it has already been shown that both type of strains code and express different virulence

factors. The choice for this ERIC I strain has as a logical consequence that ERIC II-specific virulence factors won't be detected.

In the previous section was discussed that the ERIC I group includes different MLVA genotypes. It is not unlikely that there are differences in virulence factors between these strains, thus by limiting us to the BRL 230010 strains, we exclude by design any other virulence factors present within the ERIC I genotype. When obtaining the results of the conducted experiment we can observe a bias towards the selection of growth impairing factors. By selecting only those mutated strains that do not have a changed growth profile to implement in the infection assays, these factors would be ignored.

Due to our design, no mutations impairing spore formation or spore germination would be detected. Spores had to be formed in order to perform the infection assay, since they are the only virulent form of the pathogen. The spores had to be able to germinate, since we had to be able to count the concentration of spores after heating the suspension and dilution plating.

During the experiment infection assays were used to detect the virulence-impaired mutants. Although we selected our larvae from different hives to minimize selection bias, this will still have an influence on our results. First we were restricted to the use of *Apis mellifera carnica* larvae, other subspecies could have a different resistance towards the pathogen and could thus give rise to other results. It is also of importance to mention that *in vitro* rearing of larvae results in the breeding of so-called intercast bees. These bees fall between the worker cast and the queen and have characteristics of both casts. Moreover by using *in vitro* rearing, an important part of the honey bee immunity, namely the social behavior, is not taken into account. This social immunity level incorporates the hygienic behavior, for *Apis ceranae* it is known that this hygienic behavior causes the resistance against AFB. These constraints of the infection assay will have influenced the results of our experiment.

Infection with AFB starts when spores are fed to the larvae. In the gut the spores germinate and proliferate. In this step of the infection, the pathogen has to compete with other bacteria for colonization of the gut. In earlier research, several antibiotics were already identified [83–88]. Circularin A was identified in our research as a bacteriocin that could be involved in this step. The antibiotic belongs to the class V circular bacteriocins and has been previously identified in *Clostridium beijerinckii* [193]. In this bacterium the peptide is active against a broad range of Gram-positive bacteria. Another possible virulence factor, EPS, is involved in biofilm formation in other bacteria and influences virulence [196]. If *P. larvae* would be able to form biofilms, it would not only be an advantage for colonizing its niche in the gut, but it would also be able to protect vegetative cells outside this niche. Biofilms make it possible for vegetative cells to adhere to solid surfaces. The next step in the infection process is the attack of the peritrophic matrix and eventually its destruction. The peritrophic matrix is rich of chitin and mucin-like proteins. Previous research identified *Pl*CBP49 as the responsible chitin-degrading protein. GbpA has been identified as a possible virulence factor in *P. larvae*. It is possibly expressed during this phase since it contains a chitin and mucin binding domain. Chitin and mucin-like proteins are two main components of the peritrophic matrix. The last part of the infection process involves the attachement and breaching of the gut epithelium through the paracellular route. In ERIC II genotypes the attachement to the epithelial cells is conducted by the S-layer protein [3]. This protein is however not present in ERIC I genotypes. GbpA provides in *Vibrio cholerae* binding to the epithelial cells [194].It could thus be possible that GbpA is not involved in the

previous step, but instead fulfills a similar role as the S-layer protein in ERIC I genotypes. The toxins Plx1 and Plx2 have been known to attack the epithelial cells in ERIC I genotype *P. larvae* [82] Our research did not suggest other candidates for this role.

Fundamental research on virulence factors can also result in applications in other fields. For example several EPS from *Paenibacillus spp.* are used in pharmacy, cosmetics and bioremediation. The identification of EPS in *P. larvae* could therefore prove to be useful in these fields.

# 2 Controlling AFB

American Foulbrood is a notifiable disease in Belgium and all reported outbreaks are archived by the Federal Agency for the Safety of the Food Chain (FAVV-AFSCA). The controlling methods of the disease are thus strictly regulated based on so-called "Royal Decisions". An important adaptation in the controlling strategy of AFB was published in the Royal Decision of March 7 2007. Before this date, an eradication strategy was followed to control AFB. However, de Graaf and colleagues showed in 2001 that $10\%$ of the Belgian colonies were contaminated with spores of AFB, although only about $0.1\%$ of these developed clinical signs. Following the eradication strategy in which all contaminated bees and material were destroyed, resulted thus in the unnecessary loss of strong honey bee colonies.

In November 2006 the FAVV-AFSCA formulated an advice where they stated that based on scientific research [232, 233] honey bee colonies could be treated using the shook-swarm method under certain conditions. This advice resulted in the Royal Decision of March 7 2007 where honey bee colonies contaminated with AFB were eradicated when they showed clinical signs but preserved by using the shook-swarm method when no clinical signs were present. The protocol of FAVV-AFSCA defines now that if samples sent by the beekeeper are positive for isolation (colonies on MYPGP plate containing $0,01\%$ m/v nalidixic acid and $0,02\%$ m/v pipemidic acid after heat treatment of the sample (10' 80°C) and incubation at 37°C for 3 days) and identification based on catalase-test and Gram-staining the hive labeled as suspected for AFB. A veterinarian will inspect the hives for symptoms of the disease and take samples for confirmation by PCR. When the presence of *P. larvae* is confirmed, the hive is considered as positive for AFB. If not, the hive is suspected and under surveillance. If no clinical symptoms have been observed by the veterinarian, the hive will not be put under surveillance.

Even though this Royal decision results in maintaining a significant amount of the honey bee colonies, AFB still causes an economical burden not only for beekeepers, but also for fruit farmers. Honey bees are not only important as producers of honey and propolis, but are also important pollinators. It is standard practice to transport honey bee colonies to fruit farms, since pollination by insects results in a higher yield and higher quality of fruits. When AFB is detected a perimeter is set around the diseased colony and transport of honey bees in or out of this perimeter is forbidden. The development of efficient controlling methods should prove to be be economical beneficial.

AFB is considered as a rare disease, with an incidence of $0-4$ outbreaks per year. However in August 2013 a first AFB outbreak after a period of tree years was detected during the screening of the FAVV in the Epilobee project. In 2014 and 2015 more than 50 AFB outbreaks

were detected (36 and 16 respectively) as result of the screening of honey bee colonies during the Epilobee project (Fig.5.1) [234].
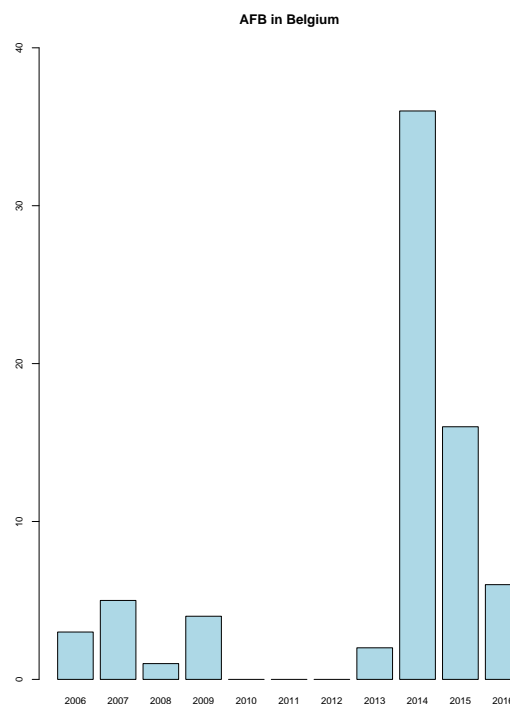
**AFB in Belgium**



Figure 5.1: **AFB outbreaks in Belgium.** The yearly incidence of AFB was 0-4 before 2014. In 2014 an absolute peak was reached with 36 outbreaks. This number declined to 16 in 2015 and 6 in 2016. Information from [234].

Samples of these outbreaks were compared using MLVA, which made a sensitive subdivision possible. Several outbreaks were caused by the same MLVA type which probably have the same origin of infection. The project showed the effectiveness and advantages of using this technique in routine diagnostics. Implementing MLVA would allow a more efficient screening of the beekeepers surrounding an outbreak. By screening beehives near the outbreak for the presence of the correct MLVA strain it would be possible to eliminate only those strains that are causing the outbreak. This would result in a shorter duration of the maintained perimeter and a more effective treatment. Moreover using this technology in an epidemiological context allows tracing the source of the infection and eliminating it. Special attention has to be given to honey in this context. Honey is often given back to bees by the beekeepers. Our study of Flemish honey showed that it contained a much wider pallet of MLVA-types than diseased brood. These types could be low or non-virulent on the honey bee colony level, allowing them to co-exist with their host. The detected strains could however also be spores still present from an earlier infection by the use of contaminated material and thus pose a possible risk of causing a new outbreak. Honey is a product that is often imported in Belgium. A lot of this honey is contaminated with *P. larvae* spores, it would thus be beneficial to monitor the import of these foreign strains using MLVA.

As mentioned earlier there is a big discrepancy between the amount of honey bee colonies that contain *P. larvae* spores and the incidence of the disease. One possible explanation

for this discrepancy has been given in the above paragraph and suggests the possibility of low or non-virulent strains on colony level. As seen with the ERIC genotypes this does not mean that these strains also have to be low virulent on larval level. To the contrary, one hypothesis for the lack of ERIC III and IV detection in the field is the very high virulence of these strains at larval level [72]. These strains would be recognized very efficiently by the hygienic worker bees, preventing a colony-wide spread of the disease. Another possibility is that the spores found in the non-clinical hives are dormant and need an external trigger before they germinate and proliferate. To investigate this it would be interesting to know whether there is proliferation of *P. larvae* cells in non-clinical hives and whether vegetative cells can be detected.

If there is a difference in virulence between different MLVA types, this should be reflected in their virulence factors. The differences between the virulence factors can be large (absence or non-functional factor in avirulent strains) or very subtle (SNPs causing minor differences in virulence).

Another possible explanation of this discrepancy can be found at the level of the honey bee host. It is possible that the non-clinical colonies are more resistant, either because of better hygiene or because of a better immune system, which would allow them to co-exist with the pathogen.

A last influential factor is the beekeeper himself. Some beekeeping methods significantly increase the probability of having an AFB outbreak. One standard example is feeding honey. Honey can be infected with *P. larvae* spores, feeding this infected honey will augment the spore titer in the beehive which can cause AFB. Other bad practices are not properly cleaning material before use or exchanging non-cleaned material, resulting in a spread of the infectious spores.

Gaining more knowledge about the virulence strategies used by the pathogen could result in the development of new reinstatement methods (f.e. antibiotics). However, since honey bees are producers of human food (honey and propolis), the use of antibiotics is forbidden in the EU. In the USA the antibiotic tetracycline has been used routinely to treat AFB. However, *P. larvae* has gained resistance against this antibiotic. Resistance will always pose a problem when using antibiotics, moreover it has to be taken into account that they do not affect the spores of the pathogen. It is arguable if treating the disease should be preferred over the efficient combination of the shook-swarm method and eradication.

# Chapter 6

# Perspectives

From previous genotyping methods it was already known that the ERIC typing grouped several genetically different strains, which was confirmed by the MLVA tool. It would be interesting to know if having a different MLVA type also results in biological differences. It has been suggested that variations in tandem repeats could result in the small biological differences that could be beneficial for the pathogen, since this flexibility can be used to escape the immune system. Moreover the closely related bacterium *P. polymyxa* is known to have strong variety of antimicrobial-encoding gene clusters among its strains. If this is also the case for *P. larvae*, this could influence its capability to colonize the honey bee gut. Variation in a pathogenicity can be checked by using *in vitro* infection assays. Differences in growth rate or in expression of factors can be used by comparative RNAseq on f.e. liquid cultures. When differences in pathogenicity or virulence are found, it would explain why not all hives containing *P. larvae* spores develop AFB. It would also allow to group the strains based on their pathogenicity and thus form pathotypes.

Differences in pathogenicity can be due to differences in virulence factors. Using comparative whole genome sequencing it should be possible to detect these variations. Another example of a difference in pathogenicity is the variation in immune answer of the honey bee. It is possible that the honey bee will mount an augmented immunological response to some strains. If the difference is due to a humoral response, *in vitro* infected larvae can be analyzed by RNAseq or qPCR. It should be noted that the difference can also be due to the hygienic behavior of the adult bees (some strains are recognized faster than others). This behavior is more difficult to check for AFB, since this is a notifiable disease. To compare hygienic behavior, infected hives should be placed in absolute isolation in order to prevent any contamination outside the lab, which is practically challenging.

The possible virulence factors detected in our research should be confirmed by using site-directed mutagenesis followed by complementation. Site-directed mutagenesis has been conducted previously to identify virulence factors of *P. larvae*, this system is based on the Targetron system developed by Sigma. The targetron contains a bacterial mobile group II intron. Group II introns are mobile ribozymes capable of creating a intron lariat that can integrate in the DNA at a specific place [235]. The intron lariat is created by precursor RNA that self-splices. The invasion in the genomic DNA site is done by reverse splicing. The intron lariat recognizes a specific DNA sequence by the presence of three hairpin loops in its structure. By changing the sequence in these hairpins the intron can be retargeted

to another sequence in the DNA. Once the candidate virulence factor is disrupted, the virulence of the mutant should be assessed using an infection assay. If the mutant is affected in its virulence it can be concluded that the gene is responsible for the mutant phenotype. A final confirmation should be done by complementing the mutant with the affected gene and affirming that it regains its wild-type virulence. Complementing can be performed by electroporating a plasmid containing the wild-type gene and its promotor in the mutant. This plasmid should also contain an antibiotic resistance factor under a constitutive promotor to make selection of successful transformants possible.

This confirmation can give insight in the possible role in virulence of GbpA, antibiotic Circularin A or biofilm formation through EPS. If biofilm formation plays a role, it would be interesting to investigate how *P. larvae* performs quorum sensing. Quorum sensing in Gram-positive bacteria is often performed by peptides [236]. If biofilm formation is true, the peptides could be extracted from the extracellular matrix and analyzed by MALDI-TOF MS. Reverse translation could give insight in which genes are responsible for quorum sensing in this bacterium.

Even when these factors are confirmed, a lot of factors are still missing. Since the genus *Paenibacillus* contains yet another insect pathogen *P. popilliae* which also has the larvae as its host, virulence factors of this species are probably also present in *P. larvae* f.e. cry proteins. Comparative genomics with non-pathogenic species of the genus could also reveal new virulence factors.

This PhD thesis has focused on both genotyping and virulence factors of the honey bee pathogen *P. larvae*. In the above chapter several tracks are proposed to further investigate the bacterium. It is clear that still a lot is to be discovered to fully understand the relationship between this interesting pathogen and its honey bee host.

# Appendices

# Appendix A

# Supplementary material

**Table S1 Tandem repeats of published genomes.** With the *Tandem repeats database* tool (Gelfand et al., 2006), 40 different tandem repeat loci could be found in the 4 published genomes of *Paenibacillus larvae*. Of each locus the pattern size as found by the tool is given. For each genome the indices where the tandem repeat locus can be found is given together with the respective copy number.

| Locus | pattern size | B3650 indices | B3650 copy nr | DSM25430 indices | DSM25430 copy nr | BRL230010 indices | BRL230010 copy nr | DSM25719 indices | DSM25719 copy nr |
|---|---|---|---|---|---|---|---|---|---|
| 1 (VNTR F) | 96 | 14006-14228 | 2,32 | 3748698-3749016 | 3,32 | 31114305-3114552 | 2,58 | | |
| | 48 | 14059-14207 | 3,09 | 3748751-3748995 | 5,09 | | | | |
| | 48 | | | 3748830-3748978 | 3,14 | | | | |
| 2 | 42 | 266373-266523 | 3,59 | 213152-213302 | 3,59 | 273771-273921 | 3,59 | 2416266-2416416 | 3,59 |
| | 43 | | | 213059-213182 | 2,93 | | | | |
| 3 | 42 | 341012-341107 | 2,28 | 2329305-2329400 | 2,28 | | | | |
| 4 | 27 | 461373-461470 | 3,62 | | | 3871250-3871347 | 3,62 | 3295951-3296048 | 3,62 |
| 5 | 36 | 535762-535841 | 2,22 | 827383-827464 | 2,27 | | | 1822792-1822871 | 2,22 |
| 6 | 6 | 599877-599981 | 17,5 | | | 3966747-3933851 | 17,5 | 3990523-3990627 | 17,5 |
| 7 | 40 | 901066-901150 | 2,12 | | | 1887921-1888005 | 2,12 | 3848292-3848376 | 2,12 |
| 8 (VNTR A) | 19 | 932188-932289 | 5,36 | 399514-399615 | 5,36 | 424455-424613 | 8,36 | 2867714-2867815 | 5,36 |
| 9 | 36 | 954413-954526 | 3,16 | 2630403-2630516 | 3,16 | 3901081-3901203 | 3,43 | 1342002-1342115 | 3,16 |
| 10 | 28 | 1174279-1174339 | 2,18 | | | 447740-447803 | 2,28 | 1624897-1624960 | 2,28 |
| 11 | 84 | 1258490-1258706 | 2,65 | 2327560-237776 | 2,65 | 2652233-2652428 | 2,36 | 4528637-4528853 | 2,65 |
| | 42 | 1258511-1258706 | 4,73 | 2327581-2327776 | 4,73 | 2652234-2652449 | 5,28 | 4528658-4528853 | 4,73 |
| | 83 | | | | | 2652234-2652450 | 2,65 | | |
| | 42 | | | | | 2652006-2652101 | 2,28 | | |
| 12 | 36 | 1500754-1500824 | 1,97 | | | 2292269-2292339 | 1,97 | 702952-703022 | 1,97 |
| 13 | 30 | 1549326-1549385 | 2 | 2680828-2680887 | 2 | 1006296-1006355 | 2 | 3316794-3316853 | 2 |
| 14 | 66 | 1560818-1561054 | 3,47 | | | | | 590200-590436 | 3,47 |
| 15 | 48 | 1699220-1699312 | 1,93 | 2799503-2799688 | 3,87 | 3155825-3156013 | 3,87 | 3436237-3436425 | 3,93 |
| 16 | 51 | 1743570-1743675 | 2,07 | 3418713-3418826 | 2,23 | 1667616-1667721 | 2,23 | 1749166-1749279 | 2,07 |
| 17 (VNTR G) | 18 | 1791546-1792111 | 32,94 | 690030-690627 | 35,22 | 3672843-3673699 | 50,1 | 4112219-4112694 | 27 |
| | 36 | 1791546-1792118 | 16,67 | | | 16331-16588 / 3673013-3673712 | 14,8 / 2,16 | | |

| | Count | Range 1 | Val 1 | Range 2 | Val 2 | Range 3 | Val 3 | Range 4 | Val 4 |
|---|---|---|---|---|---|---|---|---|---|
| | 99 | 1791555-1792118 | 5,88 | | | | | 4112228-4112694 | 4,65 |
| | 9 | | | 689987-690622 | 70,61 | | | | |
| | 89 | | | 689994-690622 | 7,17 | | | | |
| | 81 | | | 690005-690624 | 7,59 | | | | |
| | 80 | | | 690120-690751 | 7,96 | | | | |
| | 45 | | | 690238-690726 | 11,56 | 16373-16559 | 4,14 | | |
| | 27 | | | | | 16356-16577 | 8,23 | | |
| 18 | 78 | 2047981-2048193 | 2,73 | 2798732-2799023 | 3,74 | 3155057-3155348 | 3,74 | 3436902-3437192 | 2,48 |
| | 39 | 2048006-2048191 | 4,76 | 2798732-2799022 | 7,46 | 3155057-3155347 | 7,46 | 3436903-3437193 | 7,46 |
| | 117 | | | 2798732-2799023 | 2,49 | 3155057-3155348 | 2,49 | 3436902-3437192 | 2,48 |
| 19 | 47 | 2113412-2113519 | 2,29 | | | | | | |
| 21 | 36 | 2348185-2348255 | 1,97 | | | | | | |
| 22 | 48 | 2348975-2349138 | 3,41 | | | 2533998-2534161 | 3,41 | 4190293-4190456 | 3,41 |
| | 27 | 2348990-2349141 | 6,07 | | | 2534094-2534146 | 1,96 | 4190308-4190459 | 6,07 |
| 23 | 63 | 2357340-2357475 | 2,15 | | | 456088-456223 | 2,15 | 4354854-4354989 | 2,15 |
| 24 (VNTR C) | 48 | 2595946-2596108 | 3,39 | 3429304-3429458 | 3,17 | 1704713-1704842 | 2,7 | 425682-425811 | 2,7 |
| | 24 | 2596009-2596088 | 5,33 | | | 1704715-1704842 | 5,33 | 425684-425811 | 5,33 |
| 25 | 15 | 2886203-2886252 | 3,33 | 833249-833328 | 5,25 | 3050585-3050634 | 3,33 | 1097173-1097222 | 3,33 |
| | 31 | | | 833249-833318 | 2,29 | | | | |
| 26 | 45 | 3005137-3005355 | 4,84 | | | | | 3133078-3133296 | 4,84 |
| 27 | 42 | 3403722-3403806 | 2,07 | | | | | 319288-3192968 | 2,07 |
| 28 | 18 | | | 109356-109415 | 3,33 | | | | |
| 29 (VNTR E) | 68 | | | 1647396-1647833 | 6,46 | | | | |
| 30 | 120 | | | 2907751-2908124 | 3,12 | 681789-682163 | 3,12 | 1187984-1188358 | 3,12 |
| 31 | 29 | | | 3332560-3332625 | 2,27 | | | | |
| 32 | 96 | | | 3393455-3393634 | 1,87 | | | 2163806-2164111 | 3,18 |
| | 69 | | | 3393572-3393705 | 1,94 | | | 1724662-1724790 | 1,86 |
| | 48 | | | | | | | 2163846-2164090 | 5,09 |
| 33 | 45 | | | 3475347-3475790 | 9,84 | | | | |
| 34 | 45 | | | 3487920-3488015 | 2,13 | | | | |

| | | | | | |
|---|---|---|---|---|---|
| 35 (VNTR B) | 42 | | 2683910-2683997 | 2,09 | |
| | 21 | | 2683932-2683997 | 3,14 | |
| 36 | 36 | | 3065163-3065244 | 2,27 | |
| 37 | 31 | | 3407625-3407686 | 2,03 | |
| 38 | 69 | | 3415180-3415308 | 1,86 | |
| 39 | 18 | | 3137833-3138128 | 16,44 | 457398-4574227 | 13,44 |
| | 36 | | | | | 8,16 |
| 40 (VNTR D) | 24 | 3429304-3429458 | 6,33 | | |

**Table S2: Summary table virulence factors.** For every mutant the affected proteins are given and mutation, their function, whether they are known to be associated in virulence, AB resistance and growth (with mentioning the associated bacterium + Gram positive or negative) and the reference to the literature.

| Mutant | Affected protein (class mutation) | Function | Survival associated | Virulence associated | AB resistance associated | Reference |
|---|---|---|---|---|---|---|
| BRL230010::Δ*dnaK* | DnaK (replacement: 21bp) | DnaK is a chaperone protein that function in the heat-shock system. | Yes *S. aureus* (Gram +) | | | Singh *et al.* 2012 |
| | SpmA/dacB (Inversion + insertion: 99bp + 590084bp) | DacB and SpmA form together with SpmB a three gene operon. DacB encodes for the penicillin-binding protein 5, which is an important part of the cell wall in Gram-positive bacteria. | Yes *S. pneumonia* (Gram +) | | | Popham *et al.* 1995, Abdullah *et al.* 2014 |
| | SMC (replacement: 51bp) | sMC/ScpAB complex is a condensine and promotes chromosome segregation | Yes *P. aeruginoasa* (Gram -) | | | Zhao *et al.* 2016 |
| | UxaC (Inversion + insertion : 101bp + 140783bp) | Part of the Ashwell pathway which converts glucuronate into 2-keto-3-deoxygluconate (KDG) | Yes | | | Haine *et al.* 2005 |
| | WecB (replacement: 14 bp) | UDP-N-acetylglucosamine 2-epimerase | | | | Klein *et al.* 2012 |
| | PorC (replacement: 10bp) | Pyruvate-flavodoxin oxidoreductase which mediates electron transport to NADP | | | | Hughes *et al.* 1998 |
| | PrkA (replacement 270 bp) | Serine protein kinase | | | | Fischer *et al.* 1996 |
| | ERIC1_2c03180 (inversion + insertion : 101bp + 140783 bp) | ATPase of the AAA family | | | | |
| | Locus: 3046817-3046828 (replacement: 12 bp) | Non-coding | | | | |
| | Locus: 3391520-3391733 (replacement: 214 bp) | Non-coding | | | | |
| | Locus: 591971-591979 (replacement: 9bp) | Non-coding | | | | |
| BRL230010*ywqD::* | MetN (insertion:1987914bp) | MetN is a ATPase and is part of the Methioning ABC (ATP binding cassette) Uptake | Yes *B. melitensis* (Gram -) | | | Merlin *et al.* 2002 , Lestrate *et al.* 2000, Ejim *et al.* 2004 |

| | | Transporter (MUT) family | S. enterica (Gram -) | | |
|---|---|---|---|---|---|
| | YwqD (insertion: 3362371bp) | YwqD is a protein-tyrosine kinase which activates YwqF, a UDP-glucose dehydrogenase (UDP-Glc DH). They influence the synthesis of acidic type exopolymeric substances (EPS) | | Yes *B. subtilis* (Gram +) | Marvasi *et al.* 2010 |
| | RfbG (replacement: 39bp) | CDP-D-glucose-4,6-dehydratase | Yes *A. vinelandii* (Gram -) | | Gavini *et al.* 1997 |
| | ERIC1_3c0050 (inversion: 30bp) | Encodes for chromosome partition protein sMC. sMC/ScpAB complex is a condensine and promotes chromosome segregation | Yes *P. aeruginoasa* (Gram -) | | Zhao *et al.* 2016 |
| | ERIC1_1c07870 (insertion: 1698996bp ) | ABC efflux transporter-like protein | | Yes | Webber and Piddock 2003 |
| | NuoC (insertion: 1698996bp) | Subunit of the proton-translocating NADH-quinone oxidoreductase (NDH-1) | | | Castro-Guerrero *et al.* 2010 |
| | SdaC (deletion: 497bp) | Secondary carrier for threonine and serine, belonging to the *sdaCB* operon. Specific inner membrane receptor for the antibiotic Colicin V | | | Gerard *et al.* 2005 |
| | ERIC1_1c28620 (insertion: 3362371bp) | Hypothetical protein containing signal peptide domains | | | |
| | Locus 618482-618492 (replacement 11bp) | Non-coding | | | |
| | Locus 939041-939056 (replacement: 16bp) | Non-coding | | | |
| | Locus 968297-969299 (replacement: 3bp) | Non-coding | | | |
| | Locus 1268080-1268081 (replacement: 2bp) | Non-coding | | | |
| | Locus 2631022-2631075 (deletion: 54bp) | Non-coding | | | |
| BRL230010 *lysC::* | LysC (insertion: 111739bp) | Enzyme involved in DAP biosynthesis | Yes *S. enterica* (Gram -) | | Fields *et al.* 1986, Benton *et al.* 2004 |

| | | | S. aureus (Gram +) | | |
|---|---|---|---|---|---|
| | SerC (replacement: 20bp) | A phosphoserine aminotransferase involved in serine biosynthesis | Yes B. pseudomallei (Gram -) | | Rodrigues et al. 2006 |
| | NorM (replacement: 46bp) | Multidrug resistance protein. Multidrug efflux transporter | | Yes | Brown et al. 1999, Morita et al. 1998 |
| | YdcC (replacement: 207bp) | Required for efficient sporulation | | | Feucht et al. 2003 |
| | TrpA (inversion+insertion: 99bp + 510249bp) | Transposon | | | |
| | ERIC1_2c06910 (inversion + insertion: 99 bp +510249 bp) | Two component transcriptional regulator belonging to AraC family | | | |
| | ERIC1_1c11620 (insertion: 1978715 bp) | Hypothetical protein with structural domain found in several acyl-CoA acyltransferases | | | |
| | ERIC1_1C32770 (insertion: 1978715 bp) | Putative membrane protein | | | |
| | ERIC1_2c02840 (insertion: 111739 bp) | Hypothetical protein | | | |
| | Locus: 1220566-1220665 (deletion: 99bp) | Non-coding | | | |
| BRL230010cirD:: | CirD (multiple insertions: 42222 bp) | Enzyme involved in the processing and/or circularization of the circularin A prepeptide | Yes C. beijerinckii (Gram +) | | Maqueda et al. 2008, Kemperman et al. 2003 |
| | MsmX1 (replacement: 28 bp) | ATPase required to energize different saccharide transporters | Yes B. subtilis (Gram +) | | Ferreira and de Sa-Nogueira 2010 |
| | RplY (replacement: 10 bp) | Component of 50S ribosomal subunit | | Yes | El'Garch et al. 2007 |
| | GyrA (replacement: 13 bp) | Type II topoisomerase. | | Yes | Beceiro et al. 2013 |
| | Tnp16 (deletion: 1388 bp) | Transposase | | | Tenaillon et al. 2004 |
| | ERIC1_1c17630 | hypothetical protein with | | | |

| | | | | |
|---|---|---|---|---|
| | (replacement: 16 bp) | signal peptide and non-cytoplasmatic domain | | |
| | Locus: 842153-842156 (replacement: 4bp) | Non-coding | | |
| | Locus: 1159844-1159867 (replacement: 29 bp) | Non-coding | | |
| BRL230010IN(gbpA) | GbpA (Inversion: 14121 bp) | A common adherence factor for chitin and mucin | Yes V. cholera (Gram -) | Bhowmick et al. 2008 |
| | ERIC1_1c01290 (Multiple insertions: 3379464 bp) | Transposase | | |
| | Locus 842325-842345 (replacement: 21 bp) | Non-coding | | |
| BRL230010::ΔERIC1_1c21190 | ERIC1_1c21190 (replacement: 38 bp) | Hypothetical protein similar to type 2 phosphatidic acid phosphatases (PAP2) enzymes | | |
| | ERIC1_1c15570 (multiple deletions: 200 bp) | Hypothetical protein similar to the cysteine protease Prp family | | |
| | Locus 1159850-1159871 (replacement: 22bp) | Non-coding | | |

**Table S3: Summarizing table MLVA** The table gives a list of bacteria for which MLVA has been implemented.

| Bacterium | year | reference |
|---|---|---|
| *Yersinia pestis* | 2004 | [237] |
| *Bordetella pertussis* | 2004 | [238] |
| *Streptococcus pneunomiae* | 2005 | [239] |
| *Coxiella burnetii* | 2006 | [240] |
| *Brucella* | 2006 | [241] |
| *Bacillus anthracis* | 2006 | [242] |
| *Clostridium difficile* | 2006 | [243] |
| *Streptococcus uberis* | 2006 | [244] |
| *Legionella pneumophila* | 2007 | [245] |
| *Escherichia coli* | 2007 | [246] |
| *Shigella* | 2007 | [246] |
| *Burkholderia pseudomallei* | 2007 | [172] |
| *Yersinia enterocolitica* | 2007 | [247] |
| *Pseudomonas aeruginosa* | 2007 | [248] |
| *Lactobacillus casei* | 2007 | [249] |
| *Staphylococcus aureus* | 2007 | [250] |
| *Listeria monocytogenes* | 2008 | [251] |
| *Staphylococcus epidermidis* | 2008 | [252] |
| *Enterococcus faecium* | 2008 | [253] |
| *Mycobacterium bovis* | 2008 | [254] |
| *Xanthomonas citri* | 2009 | [173] |
| *Mycobacterium tyberculosis* | 2008 | [255] |
| *Vibrio parahaemolyticus* | 2009 | [256] |
| *Mycobacterium leprae* | 2009 | [257] |
| *Streptococcus agalactiae* | 2010 | [258] |
| *Leptospira* | 2010 | [259] |
| *Vibrio cholerae* | 2010 | [260] |
| *Acinetobacter baumannii* | 2011 | [261] |
| *Salmonella enterica* | 2011 | [262] |
| *Francisella noatunensis* | 2011 | [263] |
| *Chlamydia trachomatis* | 2011 | [264] |
| *Helicobacter pylori* | 2011 | [265] |
| *Ehrichia ruminantium* | 2012 | [266] |
| *Chlamydia felis* | 2012 | [267] |
| *Ralstonia solanacearum* | 2013 | [268] |
| *Yersinia pseudotuberculosis* | 2013 | [269] |
| *Xanthomonas arboricola* | 2014 | [270] |
| *Edwardsiella piscicida* | 2014 | [271] |
| *Propionibacterium acnes* | 2015 | [272] |
| *Pseudomonas syringae* | 2015 | [273] |
| *Shigella sonnei* | 2017 | [274] |
| *Ralstonia pseudosolanacearum* | 2017 | [275] |

# Appendix B

# Curriculum vitae

## 1 Publications

### 1.1 Published

- Descamps T., De Smet L., De Vos P., de Graaf D.C., "Unbiased random mutagenesis contributes to a better understanding of the virulent behavior of *Paenibacillus larvae.*" *J Appl Microbiol* Epub (2017)

- Descamps T., De Smet L., Stragier P., De Vos P., de Graaf D.C., "Multiple Locus Variable number of tandem repeat Analysis: A molecular genotyping tool for *Paenibacillus larvae.*" *Microb Biotechnol* **9**,6 (2016)

- Chemurot M., Brunain M., Akol A.M., Descamps T., de Graaf D.C., "First detection of *Paenibacillus larvae* the causative agent of American Foulbrood in a Ugandan honeybee colony." *Springerplus* **5**,1 (2016)

- Chemurot M., Akol A.M., Masembe C., de Smet L., Descamps T., de Graaf D.C., "Factors influencing the prevalence and infestation levels of *Varroa destructor* in honeybee colonies in two highland agro-ecological zones of Uganda." *Exp Appl Acarol.* **68**,4 (2016)

- Ravoet J., Schwarz R.S., Descamps T., Yañez O., Tozkar C.O., Martin-Hernandez R., Bartolomé C., De Smet L., Higes M., Wenseleers T., Schmid-Hempel R., Neumann P., Kadowaki T., Evans J.D., de Graaf D.C., "Differential diagnosis of the honey bee trypanosomatids *Crithidia mellificae* and *Lotmaria passim.*" *J Invertebr Pathol* **130**, (2015)

## 2 Conferences

- Descamps T., De Smet L., De Vos P., de Graaf D.C., "Transposon-mutagenesis based identification of virulence factors of *Paenibacillus larvae.*" *Seventh World Congress on Microbiology* (2016) [Poster]

- Descamps T., De Smet L., Stragier P., De Vos P., de Graaf D.C., "Multiple Locus Variable number of tandem repeat Analysis: A molecular genotyping tool for *Paenibacillus larvae*." *Seventh European Conference of Apidology* (2016) [Oral Presentation]

- Descamps T., Ravoet J., De Smet L., de Graaf D.C., "*Crithidia mellificae*, cultivation of field isolates and influence on the host immune system." *Sixth European Conference of Apidology* (2014) [Poster]

# 3 Other presentations

- Descamps T., De Smet L., De Vos P., de Graaf D.C., "Overzicht van het lopende doctoraal onderzoek aan de Universiteit Gent (L-MEB): Amerikaans vuilbroed." *Wetenschap voor imkers* (2015)

# 4 Courses and training

## 4.1 Transferable skills

- Advanced academic English: Conference skills - Effective slide design

- Introduction day for new PhD's: Networking course

- Job market for young researchers: CV en sollicitatie cursus

## 4.2 Specialized courses

- Analysis of high-dimensional data

- Statistical genomics

- Epidemiology and randomized clinical studies

## 4.3 Training

- Training biosafety

# 5 Training students

## 5.1 Bachelorproject

- Students: Anneleen Remmerie, Céline Pierard, Melanie Van Cauwenberghe. *Validatie van 'Lateral Flow Device' voor de identificatie van Paenibacillus larvae, de verwekker van Amerikaans vuilbroed bij de honingbij.* (2014)

## 5.2  Master I project

- Student: Klaas Van de Loock. *Pathotyping of Paenibacillus larvae, the aetiological agent of American foulbrood (AFB) in honey bees.* (2015)

## 5.3  Master II thesis

- Student: Klaas Van de Loock. *Paenibacillus larvae: Identificatie van Virulentiefactoren en Immuunrespons bij de Gastheer Apis mellifera*(2016)

- Student: Stien Beirinckx. *Identification of virulence factors of the bee pathogen Paenibacillus larvae.*(2015)

# 6  Other laboratory tasks

- Diagnostic tasks of the laboratory in absence of lab technician Marleen Brunain.

# Bibliography

[1] J. Ebeling, H. Knispel, G. Hertlein, A. Fünfhaus, and E. Genersch, "Biology of *Paenibacillus larvae*, a deadly pathogen of honey bee larvae." *Appl Microbiol Biotechnol* **100**, 7387 (2016).

[2] E. Genersch, E. Forsgren, J. Pentikäinen, A. Ashiralieva, and S. Rauch, "Reclassification of *Paenibacillus larvae* subsp. *pulvifaciens* and *Paenibacillus larvae* subsp. *larvae* as *Paenibacillus larvae* without subspecies differentiation." *Int J Syst Evol Microbiol* **56**, 501 (2006).

[3] L. Poppinga, B. Janesch, G. Fünfhaus, A. an Sekot, E. Garcia-Gonzales, G. Hertlein, C. Schäffer, and E. Genersch, "Identification and functional analysis of the S-layer protein SplA of *Paenibacillus larvae*, the causative agent of American foulbrood of honey bees." *PLoS Path* **8**, e1002716 (2012).

[4] A. Ashiralieva and E. Genersch, "Reclassification, genotypes and virulence of *Paenibacillus larvae*, the ethiological agent of American foulbrood in honey bees - a review." *Apidologie* **37**, 411 (2006).

[5] G. A. Schirach, "Histoire des abeilles." (1769), 56 pp.

[6] F. R. Cheshire and W. W. Cheyne, "The pathogenic history and history under cultivation of a new bacillus (*B. alvei*) the cause of a disease of bee hives hitherto known as foul brood." *J R Microsc Soc V* **581** (1885).

[7] G. F. White, "The bacteria of the apiary with special reference to bee diseases." *USDA, Bureau of Entomology, technical series* **14**, 1 (1906).

[8] A. Lindstrom, "Distribution of *Paenibacillus larvae* spores among adult honey bees *Apis mellifera* and the relationship with clinical symptoms of American foulbrood." *Microbiol Ecol* **56**, 253 (2008).

[9] A. Lindstrom, S. Korpela, and I. Fries, "The distribution of *Paenibacillus larvae* spores in adult bees and honey and larval mortality, following the addition of American foulbrood diseased brood or spore-contaminated honey in honey bee *Apis mellifera* colonies." *J Invertebr Pathol* **99**, 82 (2008).

[10] H. L. A. Tarr, "Studies on American foulbrood of bees. i. the relative pathogenicity of vegetative cells and endospores of *Bacillus larvae* for the brood of the bee." *Ann Appl Biol* **24**, 377 (1937).

[11] D. Yue, M. Nordhoff, L. H. Wieler, and E. Genersch, "Fluorescence in situ hybridization (FISH) analysis of the interactions between honeybee larvae and *Paenibacillus larvae*, the causative agent of American foulbrood in honeybees *Apis mellifera*." *Environ Microbiol* **10**, 1612 (2008).

[12] E. Genersch, "*Paenibacillus larvae* and American foulbrood – long since known and still surprising." *J Verbr Lebensm* **3**, 429 (2008).

[13] E. Genersch, "American foulbrood in honeybees and its causative agent, *Paenibacillus larvae*." *J Invertebr Pathol* **103, suppl.1**, S10 (2010).

[14] D. C. de Graaf, A. M. Alippi, M. Brown, J. D. Evans, M. Feldlaufer, A. Gregorc, M. Hornitzky, S. F. Pernal, D. M. T. Schuch, D. Titera, V. Tomkies, and W. Ritter, "Diagnosis of American foulbrood in honey bees: a synthesis and proposed analytical protocols." *Lett Appl Microbiol* **43**, 583 (2006).

[15] L. Bailey, "The cause of European foulbrood." *Bee World* **38**, 89 (1957).

[16] L. Bailey, "*Melisococcus pluton*, the cause of European foulbrood of honeybees (*Apis* spp.)." *J Appl Bacteriol* **55**, 65 (1983).

[17] H. Katznelson, "*Bacillus pulvifaciens* (n. sp.), an organism associated with powdery scale of honey bee larvae." *J Bacteriol* **59**, 153 (1950).

[18] C. Ash, J. A. E. Farrow, S. Wallbanks, and M. D. Collins, "Phylogenetic heterogeneity of the genus *Bacillus* revealed by comparative-analysis of small-subunit-ribosomal RNA sequences." *Lett Appl Microbiol* **13**, 202 (1991).

[19] C. Ash, F. G. Priest, and M. D. Collins, "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*." *A Van Leeuw J Microb* **64**, 253 (1993).

[20] Validation List no 52, "Validation of the publication of new names and new combinations previously effectively published outside the ijsb." *Int J Syst Bacteriol* **45**, 197 (1995).

[21] M. Heyndrickx, K. Vandemeulebroecke, B. Hoste, P. Janssen, K. Kersters, P. De Vos, N. A. Logan, N. Ali, and R. C. Berkeley, "Reclassification of *Paenibacillus* formerly *Bacillus* pulvifaciens (Nakamura 1984) Ash et al. 1994, a later subjective synonym of *Paenibacillus* formerly *Bacillus* larvae (White 1906) Ash et al. 1994, as a subspecies of *P. larvae*, with emended descriptions of *P. larvae* as *P. larvae* subsp. larvae and *P. larvae* subsp. pulvifaciens." *Int J Syst Bacteriol* **46**, 270 (1996).

[22] V. Drobnilova, V. Richter, J. Häusler, and I. Pytelova, "Characterization of *Bacillus larvae* and related bacilli by chromatography of cell fatty acids." *J Apic Res* **33**, 69 (1994).

[23] S. Neuendorf, K. Hedtke, G. Tangen, and E. Genersch, "Biochemical characterization of different genotypes of *Paenibacillus larvae* subsp. *larvae*, a honey bee bacterial pathogen." *Microbiology* **150** (2004).

[24] E. Carpana, L. Marocci, and L. Gelmini, "Evaluation of the API50CHB system for the identification and biochemical characterization of *Bacillus larvae.*" *Apidologie* **26**, 11 (1995).

[25] W. Dobbelaere, D. C. de Graaf, J. E. Peeters, and F. J. Jacobs, "Development of a fast and reliable diagnostic method for American foulbrood disease (*Paenibacillus larvae* subsp. *larvae*) using a 16S rRNA gene based PCR." *Apidologie* **32**, 363 (2001).

[26] D. C. de Graaf, P. De Vos, M. Heyndrickx, S. Van Trappen, N. Peiren, and F. J. Jacobs, "Identification of *Paenibacillus larvae* to the subspecies level: an obstacle for afb diagnosis." *J Invertebr Pathol* **91**, 115 (2006).

[27] J. Kilwinski, M. Peters, A. Ashiralieva, and E. Genersch, "Proposal to reclassify *Paenibacillus larvae* subsp. *pulvifaciens* DSM 3615 (ATCC 49843) as *Paenibacillus larvae* subsp. *larvae*. Results of a comparative biochemical and genetic study." *Vet Microbiol* **140**, 31 (2004).

[28] J. Versalovic, M. Schneider, F. J. de Bruijn, and J. R. Lupski, "Genomic fingerprinting of bacteria using repetitive sequence-based polymerase chain reaction." *Methods Mol Cell Biol* **5**, 25 (1994).

[29] A. M. Alippi, A. C. López, and O. M. Aguilar, "A PCR-based method that permits specific detection of *Paenibacillus larvae* subsp. *larvae*, the cause of American foulbrood in honeybees, at subspecies level." *Lett Appl Microbiol* **39**, 25 (2004).

[30] H. G. Trüper, "The type species of the genus *Paenibacillus* Ash et al. 1994 is *Paenibacillus polymyxa* Opinion 77," *Int J Syst Evol Microbiol* **55**, 513 (2005).

[31] O. Shida, H. Takagi, K. Kadowaki, I. Nakamura, and K. Komagata, "Transfer of *Bacillus alginolyticus, Bacillus chondroitinus, Bacillus curdlanolyticus, Bacillus glucanolyticus, Bacillus kobensis*, and *Bacillus thiaminolyticus* to the genus *Paenibacillus* and emended description of the genus *Paenibacillus.*" *Int J Syst Bacteriol* **47**, 289 (1997).

[32] M. D. Collins, P. A. Lawson, A. Willems, J. J. Cordoba, J. Fernandez-Garayzabal, P. Garcia, J. Cai, H. Hippe, and J. A. E. Farrow, "The phylogeny of the genus *Clostridium*: proposal of five new genera and eleven new species combinations." *Int J Syst Bacteriol* **44**, 812 (1994).

[33] T. Sheela and P. Usharani, "Influence of plant growth promoting rhizobacteria (PGPR) on the growth of maize (*Zea mays L.*)." *Gold Res Thoughts* **3**, 629 (2013).

[34] M. Fürnkranz, E. Adam, H. Müller, M. Grube, H. Huss, J. Winkler, and G. Berg, "Promotion of growth, health and stress tolerance of Styrian oil pumpkins by bacterial endophytes." *Eur J Plant Pathol* **134**, 509 (2012).

[35] R. de Souza, J. Meyer, R. Schoenfeld, P. B. da Costa, and L. M. P. Passaglia, "Characterization of plant growth-promoting bacteria associated with rice cropped in iron-stressed soils." *Ann Microbiol* **65**, 951 (2014).

[36] D. Duca, J. Lorv, C. L. Patten, D. Rose, and B. R. Glick, "Indole-3-acetic acid in plantmicrobe interactions." *A Van Leeuw J Microb* **106**, 85 (2014).

[37] C. L. Patten, A. J. C. Blakney, and T. J. D. Coulson, "Activity, distribution and function of indole-3-acetic acid biosynthetic pathways in bacteria." *Crit Rev Microbiol* **39**, 395 (2013).

[38] J. Xie, H. Shi, Z. Du, T. Wang, X. Liu, and S. Chen, "Comparative genomic and functional analysis reveal conservation of plant growth promoting traits in *Paenibacillus polymyxa* and its closely related species." *Sci Rep* **6**, 21329 (2016).

[39] J. B. Xie, Z. Du, L. Bai, C. Tian, Y. Zhang, J. Y. Xie, T. Wang, X. Liu, X. Chen, Q. Cheng, S. Chen, and J. Li, "Comparative genomic analysis of N2-fixing and non-N2-fixing *Paenibacillus spp.*: organization, evolution and expression of the nitrogen fixation genes. ," *PLoS Genet* **10**, e1004231 (2014).

[40] Q. T. Phi, Y. M. Park, K. J. Seul, C. M. Ryu, S. H. Park, J. G. Kim, and S. Y. Ghim, "Assessment of root-associated *Paenibacillus polymyxa* groups on growth promotion and induced systemic resistance in pepper." *J Microbiol Biotechnol* **20**, 1605 (2010).

[41] D. F. Antonopoulos, S. E. Tjamos, P. P. Antoniou, P. Rafeletos, and E. C. Tjamos, "Effect of *Paenibacillus alvei*, strain k165, on the germination of *Verticillium dahliae* microsclerotia in planta." *Biol Control.* **46**, 166 (2008).

[42] Z. Khan, S. Son, J. Akhtar, N. Gautam, and Y. Kim, "Plant growth-promoting rhizobacterium (*Paenibacillus polymyxa*) induced systemic resistance in tomato (*Lycopersicon esculentum*) against root-knot nematode (*Meloidogyne incognita*)." *Indian J Agric Sci.* **82**, 603 (2012).

[43] S. Kumar, P. S. Chauhan, L. Agrawal, R. Raj, A. Srivastava, S. Gupta, S. K. Mishra, S. Yadav, P. C. Singh, S. K. Raj, and C. S. Nautiyal, "*Paenibacillus lentimorbus* inoculation enhances tobacco growth and extenuates the virulence of cucumber mosaic virus." *PLoS One* **11**, e0149980 (2016).

[44] A. Sharma, D. R. Thakur, S. Kanwar, and V. K. Chandla, "Diversity of entomopathogenic bacteria associated with the white grub, *Brahmina coriacea.*" *J Pest Sci* **86**, 261 (2013).

[45] S. Neung, X. H. Nguyen, K. W. Naing, Y. S. Lee, and K. Y. Kim, "Insecticidal potential of *Paenibacillus elgii* HOA73 and its combination with organic sulfur pesticide on diamondback moth, *Plutella xylostella.*" *J Kor Soc Appl Biol Chem* **57**, 181 (2014).

[46] M. G. Klein, "Pest management of soil-inhabiting insects with microorganisms." *Agric Ecosyst Environ.* **24**, 337 (1988).

[47] J. Zhang, T. C. Hodgman, L. Krieger, W. Schnetter, and H. U. Schairer, "Cloning and analysis of the first cry gene from *Bacillus popilliae.*" *J Bacteriol* **179**, 4336 (1997).

[48] R. H. Lorentz, S. Ártico, A. B. Da Silveira, A. Einsfeld, and G. Corção, "Evaluation of antimicrobial activity in *Paenibacillus spp.* strains isolated from natural environment." *Lett Appl Microbiol* **43**, 541 (2006).

[49] S. G. Kim, Z. Khan, Y. H. Jeon, and Y. H. Kim, "Inhibitory effect of *Paenibacillus polymyxa* GBR-462 on *Phytophthora capsici* causing phytophthora blight in chili pepper." *J Phytopathol* **157**, 329 (2009).

[50] A. W. Eastman, D. E. Heinrichs, and Z. C. Yuan, "Comparative and genetic analysis of the four sequenced *Paenibacillus polymyxa* genomes reveals a diverse metabolism and conservation of genes relevant to plantgrowth promotion and competitiveness." *BMC Genom* **15**, 851 (2014).

[51] E. N. Grady, J. MacDonald, L. Liu, A. Richman, and Z. Yuan, "Current knowledge and perspectives of *Paenibacillus*: a review." *Microb Cell Fact* **15**, 203 (2016).

[52] T. W. Liang and S. L. Wang, "Recent advances in exopolysaccharides from *Paenibacillus spp.*: production, isolation, structure, and bioactivities." *Mar Drugs* **13**, 1847 (2015).

[53] J. Liu, J. Luo, H. Ye, and X. Zeng, "Preparation, antioxidant and antitumor activities *in vitro* of different derivatives of levan from endophytic bacterium *Paenibacillus polymyxa* EJS-3." *Food Chem Toxicol* **50**, 767 (2012).

[54] T. Velkov, P. E. Thompson, R. L. Nation, and J. Li, "Structure-activity relationships of polymyxin antibiotics." *J Med Chem* **53**, 1898 (2010).

[55] R. Goodacre and D. B. Kell, "Pyrolysis mass spectrometry and its applications in biotechnology." *Curr opin biotechnol* **7**, 20 (1996).

[56] M. A. Schäfer, E. Genersch, A. Fünfhaus, L. Poppinga, N. Formella, B. Bettin, and A. Karger, "Rapid identification of differentially virulent genotypes of *Paenibacillus larvae*, the causative organism of American foulbrood of honey bees, by whole cell MALDI-TOF mass spectrometry." *Vet Microbiol* **170**, 291 (2014).

[57] N. Singhal, M. Kumar, P. K. Kanaujia, and J. Virdi, "MALDI-TOF mass spectrometry: an emerging technology for microbial identification and diagnosis." *Front Microbiol* **6**, 791 (2015).

[58] S. Djordjevic, M. Hoshon, and M. Hornitsky, "DNA restriction-endonuclease profiles and typing of geographically diverse isolates of *Bacillus-Larvae*." *J Apic Res* **33**, 95 (1994).

[59] A. M. Alippi, A. C. Lopez, and O. M. Aguilar, "Differentiation of *Paenibacillus larvae* subsp. *larvae*, the cause of American foulbrood of honey bees, by using PCR and restriction fragment analysis of genes encoding 16S rRNA." *Appl Environ Microbiol* **68**, 3655 (2002).

[60] X. Y. Wu, J. Chin, A. Ghalayini, and M. Hornitzky, "Pulsed-field gel electrophoresis typing and oxytetracycline sensitivity of *Paenibacillus larvae* subsp. *larvae* isolates of Australian origin and those recovered from honey imported from Argentina." *J Apic Res* **44**, 87 (2005).

[61] K. Antunez, C. Piccini, S. Castro-Sowinski, A. S. Rosado, L. Seldin, and P. Zunino, "Phenotypic and genotypic characterization of *Paenibacillus larvae* isolates." *Vet Microbiol* **124**, 178 (2007).

[62] A. M. Alippi and O. M. Aguilar, "Characterization of isolates of *Paenibacillus larvae* subsp. *larvae* from diverse geographical origin by the polymerase chain reaction and box primers." *J Invertebr Pathol* , 21 (1998).

[63] E. Genersch and C. Otten, "The use of repetitive element PCR fingerprinting (rep-PCR) for genetic subtyping of germen field isolates of *Paenibacillus larvae* subsp *larvae*." *Apidologie* **34**, 195 (2003).

[64] A. M. Alippi and O. M. Aguilar, "Unique DNA fingerprint patterns of *Paenibacillus larvae* subsp. *larvae* strains." *J Apic Res* **37**, 273 (1998).

[65] B. J. Morrissey, T. Helgason, L. Poppinga, A. Funfhaus, E. Genersch,  and G. E. Budge, "Biogeography of *Paenibacillus larvae*, the causative agent of American foulbrood, using a new multilocus sequence typing scheme." *Environ Microbiol* **17**, 1414 (2015).

[66] A. Maths, "Multilocus sequence typing (mlst) analysis," http://www.applied-maths.com/applications/mlst.

[67] A. Casadevall and L. A. Pirofski, "Host-pathogen interactions: Attributes of virulence." *J Infect Dis* **184**, 337 (2001).

[68] A. Casadevall and L. A. Pirofski, "Host-pathogen interactions: Redefining the basic concepts of pathogenicity and virulence." *Infect Immun* **67**, 3703 (1999).

[69] L. Pirofski and A. Casadevall, "Q and A What is a pathogen? A question that begs the point." *BMC Biol* **10** (2012).

[70] E. E. McClelland, P. Bernhardt,  and A. Casadevall, "Estimating the relative contributions of virulence factors for pathogenic microbes." *Infect Immun* **74**, 1500 (2006).

[71] E. Genersch, A. Ashiralieva,  and F. I., "Strain - and genotype-specific differences in virulence of *Paenibacillus larvae* subsp. *larvae*, a bacterial pathogen causing American foulbrood disease in honeybees." *Appl Environ Microbiol* **71**, 7551 (2005).

[72] S. Rauch, A. Ashiralieva, K. Hedtke,  and E. Genersch, "Negative correlation between individual-insect-level virulence and colony-level virulence of *Paenibacillus larvae*, the ethiological agent of American foulbrood of honey bees." *App Environ Microbiol* **75**, 3344 (2009).

[73] D. C. de Graaf, D. Vanderkerckhove, W. Dobbelaere, J. Peeters,  and F. J. Jacobs, "Influence of the proximity of American foulbrood cases and apiculture management on the prevalence of *Paenibacillus larvae* spores in Belgian honey." *Apidologie* **32**, 587 (2001).

[74] A. Fünfhaus, A. Ashiralieva, R. Borriss,  and E. Genersch, "Use of suppression subtrative hybridization to identify genetic differences between differentially virulent genotypes of *Paenibacillus larvae*, the etiological agent of American foulbrood of honey bees." *Environ Microbiol Rep* **1**, 240 (2009).

[75] A. Funfhaus and E. Genersch, "Proteome analysis of *Paenibacillus larvae* reveals the existence of a putative S-layer protein." *Environ Micriobio Rep* **4**, 194 (2012).

[76] Q. W. T. Chan, R. S. Cornman, I. Birol, N. Y. Liao, S. K. Chan, T. R. Doking, S. D. Jackman, G. A. Taylor, S. J. M. Jones, D. C. de Graaf, J. D. Evans,  and L. J. Foster, "Updated genome assembly and annotation of *Paenibacillus larvae*, the agent of American foulbrood disease of honey bees." *BMC Genomics* **12** (2011).

[77] X. Qin, J. D. Evans, K. A. Aronstein, K. D. Murray,  and G. M. Weinstock, "Genome sequence of the honey bee pathogens *Paenibacillus larvae* and *Ascosphaera apis*." *Insect Mol Biol* **15**, 715 (2006).

[78] M. Djukic, E. Brzuszkiewicz, A. Fünfhaus, J. Voss, K. Gollnow, L. Poppinga, H. Liesegang, E. Garcia-Gonzalez, E. Genersch,  and R. Daniel, "How to kill the honey bee larva: genomic potential and virulence mechanisms of *Paenibacillus larvae*." *PLoS One* **9**, e90914 (214).

[79] L. Poppinga and E. Genersch, "Heterologous expression of green fluorescent protein in *Paenibacillus larvae*, the causative agent of American foulbrood of honey bees." *J Appl Microbiol* **112**, 430 (2012).

[80] K. D. Murray and K. A. Aronstein, "Transformation of the Gram-positive honey bee pathogen, *Paenibacillus larvae*, by electroporation." *J Microbiol Methods* **75**, 325 (2008).

[81] E. Garcia-Gonzalez, L. Poppinga, A. Fünfhaus, G. Hertlein, K. Hedtke, A. Jakubowska,  and E. Genersch, "*Paenibacillus larvae* chitin-degrading protein *Pl*CBP49 is a key virulence factor inn American foulbrood in honey bees." *PLoS Path* **10**, e1004284 (2014).

[82] A. Fünfhaus, L. Poppinga,  and E. Genersch, "Identification and characterization of two novel toxins expressed by the lethal honey bee pathogen *Paenibacillus larvae*, the causative agent of American foulbrood." *Environ Microbiol* **15**, 2951 (2013).

[83] E. Garcia-Gonzalez, S. Müller, P. Ensle, R. D. Süssmuth,  and E. Genersch, "Elucidation of sevadicin, a novel nonribosomal peptide secondary metabolite produced by the honey bee pathogenic bacterium *Paenibacillus larvae*." *Environ Microbiol* **16**, 1297 (2014).

[84] E. Garcia-Gonzalez, S. Müller, G. Hertlein, N. C. Heid, R. D. Süssmuth,  and E. Genersch, "Biological effects of paenilamicin, a secondary metabolite antibiotic produced by the honey bee pathogenic bacterium *Paenibacillus larvae*." *Microbiol Open* **3**, 642 (2014).

[85] E. C. Holst, "An antibiotic from a bee pathogen." *Science* **102**, 593 (1945).

[86] G. Hertlein, S. Müller, E. Garcia-Gonzalez, L. Poppinga, R. Süssmuth,  and E. Genersch, "Production of the catechol type siderophore bacillibactin by the honey bee pathogen *Paenibacillus larvae*." *PloS One* **9**, e108272 (2014).

[87] S. Müller, E. Garcia-Gonzalez, A. Mainz, G. Hertlein, N. C. Heid, E. Mösker, H. van den Elst, H. S. Overkleeft, E. Genersch,  and R. D. Süssmuth, "Paenilamicin - structure and biosynthesis of a hybrid non-ribosomal peptide/polyketide antibiotic from the bee pathogen *Paenibacillus larvae*." *Angew Chem Int Ed Eng* **53**, 10547 (2014).

[88] S. Sood, H. Steinmetz, B. Beims, K. I. Mohr, M. Stadler, M. Djukic, W. von der Ohe, M. Steinert, R. Daniel, and R. Müller, "Paenilarvins: uterin family lipopeptides from the honey bee pathogen *Paenibacillus larvae*." *Chem BioChem* **15**, 1947 (2014).

[89] E. W. Davidson, "Ultrastructure of American foulbrood disease pathogenesis in larvae of the worker honey bee *Apis mellifera*." *J Invertebr Pathol* **21**, 53 (1973).

[90] A. Gregorc and I. D. Bowen, "Histopathological and histochemical changes in honeybee larvae *Apis mellifera L.* after infection with *Bacillus larvae*, the causative agent of American foulbrood disease." *Cell Biol Int* **22**, 137 (1998).

[91] P. Chantawannakul and B. N. Dancer, "American foulbrood in honey bees." *Bee World* **82**, 168 (2001).

[92] T. T. Hoa, L. H. Duc, R. Isticato, L. Baccigalupi, E. Ricca, P. H. Van, and S. M. Cutting, "Fate and dissemination of *Bacillus subtilis spores* in a murine model." *Appl Environ Microbiol* **67**, 3819 (2001).

[93] G. Casula and S. M. Cutting, "*Bacillus* probiotics: spore germination in the gastrointestinal tract." *Appl Environ Microbiol* **68**, 2344 (2002).

[94] G. E. Baida and N. P. Kuzmin, "Cloning and primary structure of a new hemolysin gene from *Bacillus cereus*." *Biochem Biophys Acta* **1264**, 151 (1995).

[95] B. N. Dancer and P. Chantawannakul, "The proteases of American foulbrood scales." *J Invertebr Pathol* **70**, 79 (1997).

[96] K. Antúnez, M. Anido, G. Schlapp, J. D. Evans, and P. Zunino, "Characterization of secreted proteases of *Paenibacillus larvae*, potential virulence factors involved in honeybee larval infection." *J Invertebr Pathol* **102**, 129 (2009).

[97] D. Frees, S. N. Qazi, P. J. Hill, and H. Ingmer, "Alternative roles of ClpX and ClpP in *Staphylococcus aureus* stress tolerance and virulence." *Mol Microbiol* **48**, 1565 (2003).

[98] S. Tsukita, Y. Yamazaki, T. Katsuno, A. Tamura, and S. Tsukita, "Tight junction-based epithelial microenvironment and cell proliferation." *Oncogene* **27**, 6930 (2008).

[99] B. Alberts, A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter, *Molecular biology of the cell. 4th Edition* (Garland Science, 2002) Chap. The extracellular matrix of animals.

[100] R. Stern and M. J. Jedrzejas, "Hyaluronidases: Their genomics, structures and mechanisms of action." *Chem Rev* **106**, 818 (2006).

[101] D. Hegedus, M. Erlandson, C. Gillott, and U. Toprak, "New insights into peritrophic matrix synthesis, architecture and function." *Annu Rev Entomol* **54**, 285 (2009).

[102] M. Sebaihia, M. W. Peck, N. P. Minton, N. R. Thomson, M. T. Holden, W. J. Mitchell, A. T. Carter, S. D. Bentley, D. R. Mason, L. Crossman, C. J. Paul, A. Ivens, M. H. Wells-Bennik, I. J. Davis, A. M. Cerdeño-Tárraga, C. Churcher, M. A. Quail, T. Chillingworth, T. Feltwell, A. Fraser, I. Goodhead, Z. Hance, K. Jagels, N. Larke, M. Maddison,

S. Moule, K. Mungall, H. Norbertczak, E. Rabbinowitsch, M. Sanders, M. Simmonds, B. White, S. Whithead, and J. Parkhill, "Genome sequence of a proteolytic (Group I) *Clostridium botulinum* strain Hall A and comparative analysis of the clostridial genomes." *Genome Res* **17**, 1082 (2007).

[103] T. Yamamoto and Y. Tanada, "Phospholipid, an enhancing component in the synergistic factor of a granulosis virus of the armyworm *Pseudaletia unipuncta*." *J Invertebr Pathol* **31**, 48 (1978).

[104] Y. Zhu, T. Hukuhara, and K. Tamura, "Location of a synergistic factor in the capsule of a granulosis virus of te armyworm *Pseudaletia unipuncta*." *J Invertebr Pathol* **54**, 49 (1989).

[105] Q. W. T. Chan, A. P. Melathopoulos, S. F. Pernal, and L. J. Foster, "The innate immune and systemic response in honey bees to a bacterial pathogen, *Paenibacillus larvae*." *BMC Genomics* **10** (2009).

[106] L. Hedstrom, "Serine protease mechanism and specificity." *Chem Rev* **102**, 4501 (2002).

[107] A. J. Barrett, "Bioinformatics of proteases in the merops database." *Curr Opin Drug Discov Devel* **7**, 334 (2004).

[108] A. Lövgren, M. Zhang, A. Engström, G. Dalhammar, and R. Landén, "Molecular characterization of immune inhibitor a, a secreted virulence protease from *Bacillus thuringiensis*." *Mol Microbiol* **4**, 2137 (1990).

[109] J. D. Evans, "Transcriptional immune responses by honey bee larvae during invasion by the bacterial pathogen, *Paenibacillus larvae*." *J Invertebr Pathol* **85**, 105 (2004).

[110] R. S. Cornman, D. Lopez, and J. D. Evans, "Transcriptional response of honey bee larvae infected with the bacterial pathogen *Paenibacillus larvae*." *PloS One* **8** (2013).

[111] K. D. Murray, K. A. Aronstein, and de Leon J. H., "Analysis of pMA67, a predicted rolling-circle replicating, mobilizable, tetracycline-resistance plasmid from the honey bee pathogen, *Paenibacillus larvae*." *Plasmid* **58**, 89 (2007).

[112] C. Khosla, R. S. Gokhale, J. R. Jacobsen, and D. E. Cane, "Tolerance and specificity of polyketide synthases." *Annu Rev Biochem* **68**, 219 (1999).

[113] L. Du and L. Lou, "PKS and NRPS release mechanisms." *Nat Prod Rep* **27**, 255 (2010).

[114] G. Hertlein, M. Seiffert, S. Gensel, E. Garcia-Gonzalez, J. Ebeling, R. Skobalj, A. Kuthning, R. D. Süssmuth, and E. Genersch, "Biological role of paenilarvins, iturin-like lipopeptide secondary metabolites produced by the honey bee pathogen *Paenibacillus larvae*." *PloS One* **11** (2016).

[115] N. W. Calderone, "Insect pollinated crops, insect pollinators and US agriculture: trend analysis of aggregate data for the period 1992-2009." *PloS One* **7**, e37235 (2012).

[116] J. E. Losey and M. Vaughan, "The economic value of ecological services provided by insects." *Bioscience* **56**, 311 (2006).

[117] M. Laurent, P. Hendrikx, M. Ribiere-Chabert, and M. Chauzat, "A pan-European epidemiological study on honeybee colony losses 2012-2014." (2016).

[118] I. Jarratt and M. A. Franco-Dixon, "Cost-Benefit Analysis of American Foulbrood (AFB) disease management options in Queensland - Preliminary results. Proceedings 43rd Annual conference of the Australian Agricultural and Resource Economics Society." *Proceedings 43rd Annual conference of the Australian Agricultural and Resource Economics Society* (1999).

[119] F. L. W. Ratnieks and L. Keller, "Queen control of egg fertilization in the honey bee," *Behav ecol sociobiol* **44**, 57 (1998).

[120] N. Weaver, "Rearing of honeybee larvae on royal jelly in the laboratory," *Science* **121**, 509 (1955).

[121] N. Weaver, "Effect of larval age on dimorphic differentiation of the female honeybee." *Ann Ent Soc Am* **50**, 283 (1957).

[122] Y. S. Peng and S. C. Jay, "Larval rearing by worker bees lacking their mandibular glands. I. Rearing by small numbers of workers bees," *Can J Zool* **109**, 1175 (1977).

[123] F. Ruttner, "The mating of the honey bee." *Bee World* **37**, 3 (1956).

[124] R. E. Snodgrass, *Anatomy and Physiology of the Honey Bee.* (McGraw-Hill, 1925).

[125] T. D. Seeley, "Life history strategy of the honey bee *Apis mellifera.*" *Oecologia* **32**, 109 (1978).

[126] M. L. Winston, *The Biology of the Honey Bee.* (University Press, 1987).

[127] H. P. Hendriksma, S. Hartel, and I. Steffan-Dewenter, "Honey bee risk assessment: new approaches for *in vitro* rearing and data analysis." *Methods Ecol Evol* **2**, 509 (2011).

[128] J. D. Evans, K. A. Aronstein, Y. P. Chen, C. Hetru, J. L. Imler, H. Jiang, M. Kanost, G. J. Thompson, Z. Zou, and D. Hultmark, "Immune pathways and defence mechanisms in honey bee *Apis mellifera.*" *Insect Mol Biol* **15**, 645 (2006).

[129] J. A. Hoffmann, "The immune response of *Drosophila.*" *Nature* **426**, 33 (2003).

[130] P. Schmid-Hempel, *Parasites in Social Insects.* (Princeton University Press, Princeton, NJ, 1998).

[131] M. Spivak and G. S. Reuter, "Resistance to American foulbrood disease by honey bee colonies *Apis mellifera* bred for hygienic behavior." *Apidologie* **32**, 555 (2001).

[132] J. S. Parker, K. Mizuguchi, and N. J. Gay, "A family of proteins related to spatzle, the toll receptor ligand, are encoded in the *Drosophila* genome." *Proteins* **45**, 71 (2001).

[133] S. Valanne, J. H. Wang, and M. Rämet, "The *Drosophila* toll signaling pathway." *J Immunol* **186**, 649 (2011).

[134] D. Hultmark, "*Drosophila* immunity: Paths and patterns." *Curr Opin Immunol* **15**, 12 (2003).

[135] H. Myllymaki, S. Valanne, and M. Ramet, "The *Drosophila* imd signaling pathway." *J Immunol* **8**, 3455 (2014).

[136] S. Bina and M. Zeidler, *JAK-STAT pathway in disease* (Landes Bioscience, 2009) Chap. JAK/STAT pathway signalling in *Drosophila melanogaster*.

[137] Bordenstein Lab, "Innate immunity pathways and responses," http://bordensteinlab.vanderbilt.edu/IIID/test/pathway.php (2011).

[138] K. Casteels-Josson, T. Capaci, P. Casteels, and P. Tempst, "Apidaecin multipeptide precursor structure: a putative mechanism for amplification of the insect antibacterial response." *Embo J* **12**, 1569 (1993).

[139] M. Van Vaerenbergh, D. Cardoen, E. M. Formesyn, M. Brunain, G. Van Driessche, S. Blank, E. Spillner, P. Verleyen, T. Wenseleers, L. Schoofs, B. Devreese, and D. C. de Graaf, "Extending the honey bee venome with the antimicrobial peptide apidaecin and a protein resembling wasp antigen 5." *Insect Mol Biol* **22**, 199 (2013).

[140] K. Casteels-Josson, W. Zhang, T. Capaci, P. Casteels, and P. Tempst, "Acute transcriptional response of the honeybee peptide-antibiotics gene repertoire and required post-translational conversion of the precursor structures." *J Biol Chem* **269**, 28569 (1994).

[141] P. Casteels, C. Ampe, L. Riviere, J. Van Damme, C. Elicone, M. Fleming, F. Jacobs, and P. Tempst, "Isolation and characterization of abaecin, a major antibacterial response peptide in the honeybee (*Apis mellifera*)." *Europ J Biochem* **187**, 381 (1990).

[142] J. D. Evans and R. S. Schwarz, "Bees brought to their knees: microbes affecting honey bee health." *Trends Microbiol* **19**, 614 (2011).

[143] G. C. Eickwort, *Mites: ecological and evolutionary analyses of life-history patterns* (New York: Springer, 1994) Chap. Evolution and life-history patterns of Mites associated with honey bees, pp. 218–251.

[144] D. Sammataro, U. Gerson, and G. Needham, "Parasitic mites of honey bees: Life history, implications and impact," *Annu Rev Entomol* **45**, 519 (2000).

[145] X. L. Yang and D. L. Cox-Foster, "Impact of an ectoparasite on the immunity and pathology of an invertebrate: Evidence for host immunosuppression and viral amplification." *P Natl Acad Sci USA* **102**, 7470 (2005).

[146] Y. P. Chen, J. S. Pettis, J. D. Evans, M. Kramer, and M. F. Feldlaufer, "Transmission of kashmir bee virus by the ectoparasitic mite *Varroa destructor*." *Apidologie* **35**, 441 (2004).

[147] E. Genersch and M. Aubert, "Emerging and re-emerging viruses of the honey bee (*Apis mellifera L.*)." *Vet Res* **41** (2010).

[148] D. Vanengelsdorp, J. D. Evans, C. Saegerman, C. Mullin, E. Haubruge, B. K. Nguyen, M. Frazier, J. Frazier, D. Cox-Foster, Y. Chen, R. Underwood, D. R. Tarpy, and J. S. Pettis, "Colony collapse disorder: a descriptive study." *PloS One* **4**, e6481 (2009).

[149] E. Zander, "Tierische parasiten als krankheitserreger bei der biene," *Leipziger Bienenztg.* **24**, 147 (1909).

[150] I. Fries, F. Feng, A. da Silva, S. B. Slemenda, and N. J. Pieniazek, "*Nosema ceranae* n. sp. (microspora, nosematidae), morphological and molecular characterization of a microsporidian parasite of the Asian honey bee *Apis cerana* (hymenoptera, apidae)." *Eur J Protistol* **32**, 356 (1996).

[151] M. Chemurot, L. De Smet, M. Brunain, R. De Rycke, and D. C. de Graaf, "*Nosema neumanni* n. sp. (Microsporidia, Nosematidae), a new microsporidian parasite of honeybees, *Apis mellifera* in uganda." *Eur J Protistol* **61**, 13 (2017).

[152] E. Forsgren and I. Fries, "Comparative virulence of *Nosema ceranae* and *Nosema apis* in individual European honey bees." *Veterinary parasitology* **170**, 212 (2010).

[153] R. Martin-Hernandez, C. Botias, L. Barrios, A. Martinez-Salvador, A. Meana, C. Mayack, and M. Higes, "Comparison of the energetic stress associated with experimental *Nosema ceranae* and *Nosema apis* infection of honey bees (*Apis mellifera*)." *Parasitol Res* **109**, 605 (2011).

[154] A. Mehr, W. Wilson, and R. Sackett, "Persistence of chalkbrood (*Ascosphaera apis*) in some North American honey-bee colonies one year after infection." *Apiacta* **13** (1978).

[155] D. Vanengelsdorp, J. Hayes, R. M. Underwood, and J. Pettis, "A survey of honey bee colony losses in the US, fall 2007 to spring 2008," *PloS One* **3** (2008).

[156] D. G. Yost, P. Tsourkas, and P. S. Amy, "Experimental bacteriophage treatment of honeybees (*Apis mellifera*) infected with *Paenibacillus larvae*, the causative agent of American foulbrood disease." *Bacteriophage* **6**, e1122698 (2016).

[157] I. Fries, A. Lindström, and S. Korpela, "Vertical transmission of American foulbrood (*Paenibacillus larvae*) in honey bees (*Apis mellifera*)." *Vet Microbiol* **114**, 269 (2006).

[158] N. Woodford and D. M. Livermore, "Infections caused by Gram-positive bacteria: a review of the global challenge." *J Infect* **59**, S4 (2009).

[159] T. C. Olofsson and A. Vásquez, "Detection and identification of a novel lactic acid bacterial flora within the honey stomach of the honeybee *Apis mellifera*." *Curr Microbiol* **57**, 356 (2008).

[160] D. Ortelli, P. Edder, and C. Corvi, "Analysis of chloramphenicol residues in honey by liquid chromatography-tandem mass spectrometry." *Chromatographia* **59**, 61 (2004).

[161] J. D. Evans, "Diverse origins of tetracycline resistance on the honey bee bacterial pathogen *Paenibacillus larvae*." *J Invertebr Pathol* **83**, 46 (2003).

[162] R. M. Carlton, "Phage therapy: past history and future prospects." *Arch Immunol Ther Exp* **47**, 267 (1991).

[163] J. H. López, W. Schuehly, K. Crailsheim, and U. Riessberger-Gallé, "Transgenerational immune priming in honeybees." *Proc R Soc B* **281**, 20140454 (2014).

[164] Y. W. Chen, C. H. Wang, J. An, and H. Kai-Kuang, "Susceptibility of the Asian honey bee, *Apis cerana*, to American foulbrood, *Paenibacillus larvae larvae.*" *J Apicul Res* **34**, 169 (2000).

[165] A. van Belkum, "The role of short sequence repeats in epidemiologic typing." *Curr Opin Microbiol* **2**, 306 (1999).

[166] E. Mazars, S. Lesjean, A. L. Banuls, M. Gilbert, V. Vincent, B. Gicquel, M. Tibayrenc, C. Locht, and P. Supply, "High-resolution minisatellite-based typing as a portable approach to global analysis of *Mycobacterium tuberculosis* molecular epidemiology." *PNAS* **98**, 1414 (2001).

[167] A. M. Klevytska, L. B. Price, J. M. Schupp, P. L. Worsham, J. Wong, and P. Keim, "Identification and characterization of variable-number tandem repeats in the *Yersinia pestis* genome." *J Clin Microbiol* **39**, 3179 (2001).

[168] N. Malachowa, A. Sabat, M. Gniadkowski, J. Kryszton-Russjan, J. Empel, J. Miedzobrodzki, K. Kosowska-Shick, P. C. Appelbaum, and W. Hryniewicz, "Comparison of multiple-locus variable-number tandem-repeat analysis with pulsed-field gel electrophoresis, spa typing, and multilocus sequence typing for clonal characterization of *Staphylococcus aureus* isolates." *J Clin Microbiol* **43**, 3095 (2005).

[169] A. Buhlmann, T. Dreo, F. Rezzonico, J. F. Pothier, T. H. Smits, M. Ravnikar, J. E. Frey, and B. Duffy, "Phylogeography and population structure of the biologically invasive phytopathogen *Erwinia amylovora* inferred using minisatellites." *Environ Microbiol* **16**, 2112 (2014).

[170] P. Techaruvichit, H. Takahashi, M. Vesaratchavest, S. Keeratipibul, T. Kuda, and B. Kimura, "Development of multiple-locus variable-number tandem-repeat analysis for molecular subtyping of *Campylobacter jejuni* by using capillary electrophoresis." *Appl Environ Microbiol* **81**, 4318 (2015).

[171] J. M. Girard, D. M. Wagner, A. J. Vogler, C. Keys, C. J. Allender, L. C. Drickamer, and P. Keim, "Differential plague-transmission dynamics determine *Yersinia pestis* population genetic structure on local, regional and global scales." *PNAS* **101**, 8408 (2004).

[172] J. M. U'Ren, H. Hornstra, T. Pearson, J. M. Schupp, B. Leadem, S. Georgia, R. W. Sermswan, and P. Keim, "Fine-scale genetic diversity among *Burkholderia pseudomallei* soil isolates in northeast thailand." *App Environ Microbiol* **73**, 6678– (2007).

[173] T. N. Bui, C. Verniere, P. Jarne, P. Brisse, F. Guerin, S. Boutry, L. Gagnevin, and O. Pruvost, "From local surveys to global surveillance: three high-throughput genotyping methods for epidemiological monitoring of *Xanthomonas citri* pv. *citri* pathotypes." *Appl Environ Microbiol* **75**, 1173 (2009).

[174] C. W. Stratilo and D. E. Bader, "Genetic diversity among *Bacillus anthracis* soil isolates at fine geographic scales." *Appl Environ Microbiol* **78**, 6433– (2012).

[175] J. Torres-Cruz and M. W. van der Woude, "Slipped strand mispairing can function as a phase variation mechanism in *Escherichia coli.*" *J Bacteriol* **185**, 6990 (2003).

[176] A. J. Vogler, C. Keys, Y. Nemoto, R. E. Colman, Z. Jay, and P. Keim, "Effect of repeat copy number on variable-number tandem repeat mutations in *Escherichia coli* O157:H7." *J Bacteriol* **188**, 4253 (2006).

[177] B. A. Lindstedt, "Multiple-locus variable number tandem repeats analysis for genetic fingerprinting of pathogenic bacteria." *Electrophoresis* **26**, 2567 (2005).

[178] A. van Belkum, "Tracing isolates of bacterial species by multilocus variable number of tandem repeat analysis (MLVA)." *FEMS Immunol Med Microbiol* **49**, 22 (2007).

[179] Y. Gelfand, A. Rodriguez, and G. Benson, "TRDB- the tandem repeats database." *Nucleic Acids Res* **35**, D1 (2006).

[180] I. Loncaric, I.and Derakshifar, J. T. Oberlerchner, H. Köglberger, and R. Moos-beckhofer, "Genetic diversity among isolates of *Paenibacillus larvae* from austria." *Invertebr Pathol* **100**, 44 (2009).

[181] S. Bassi, G. Formato, M. Milito, K. Trevisiol, C. Salogni, and E. Carra, "Phenotypic characterization and ERIC-PCR based genotyping of *Paenibacillus larvae* isolates recovered from anerican foulbrood outbreaks in honey bees from italy." *Vet Quarterly* **35**, 27 (2015).

[182] R Core Team, "R: a language and environment for statistical computing." https://www.R-project.org/ (2015).

[183] D. Krska, R. Ravulapalli, R. J. Fieldhouse, M. R. Lugo, and A. R. Merrill, "C3larvin toxin, an ADP-ribosyltransferase from *Paenibacillus larvae.*" *J Biol Chem* **290**, 1639 (2015).

[184] J. M. Mei, F. Nourbakhsh, C. W. Ford, and D. W. Holden, "Identification of *Staphylococcus aureus* virulence genes in a murine model of bacteraemia using signature-tagged mutagenesis." *Mol Microbiol* **26**, 399 (1997).

[185] D. J. Lampe, B. J. Akerley, E. J. Rubin, J. J. Mekalanos, and H. M. Robertson, "Hyperactive transposase mutants of the *Himar1* mariner transpon." *Porc Natl Acad Sci USA* **96**, 11428 (1999).

[186] S. Paik, L. Senty, S. Das, J. C. Noe, C. L. Munro, and T. Kitten, "Identification of virulence determinants for endocarditis in *Streptococcus sanguinis* by signature-tagged mutagenesis," *Infect Immun* **73**, 6064 (2005).

[187] M. Szostkova, D. Horakova, and M. Nemec, "The influence of the growth phase of enteric bacteria on electrotransformation with plasmid DNA," *Folia Microbiol* **44**, 177 (1999).

[188] L. De Smet, D. De Koker, A. K. Hawley, L. J. Foster, P. De Vos, and D. C. de Graaf, "Effect of bodily fluids from honey bee *Apis mellifera* larvae on growth and genome-wide transcriptional response of the causal agent of American foulbrood disease *Paenibacillus larvae*." *PLoS One* **9**, e89175 (2014).

[189] T. M. Therneau and P. M. Grabsch, *Modeling Survival Data: Extending the Cox Model.* (New York: Springer, 2000).

[190] A. Kassambara and M. Kosinki, "survminer: Drawing survival curves using 'ggplot2'. ed. 0.4.0., r.p.v." (2017).

[191] A. V. Karlyshev, M. J. Pallen, and B. W. Wren, "Single-primer PCR procedure for rapid identification of transposon insertion sites." *Biotechniques* **28**, 1078 (2000).

[192] M. Maqueda, M. Sanchez-Hidalgo, M. Fernandez, M. Montalban-Lopez, E. Valdivia, and M. Martinez-Bueno, "Genetic features of circular bacteriocins produced by Gram-positive bacteria," *Fems Microbiol Rev* **32**, 2 (2008).

[193] R. Kemperman, M. Jonker, A. Nauta, O. P. Kuipers, and J. Kok, "Functional analysis of the gene cluster involved in production of the bacteriocin circularin a by *Clostridium beijerinckii* ATCC 25752." *Appl Environ Microb* **69**, 5839 (2003).

[194] R. Bhowmick, A. Ghosal, B. Das, H. Koley, D. R. Saha, S. Ganguly, R. K. Nandy, R. K. Bhadra, and N. S. Chatterjee, "Intestinal adherence of *Vibrio cholerae* involves a coordinated interaction between colonization factor GbpA and mucin." *Infect Immun* **76**, 4968 (2008).

[195] L. Shao, M. Devenport, and M. Jacobs-Lorena, "The peritrophic matrix of hematophagous insects." *Arch Insect Biochem* **47**, 119 (2001).

[196] M. Marvasi, P. T. Visscher, and L. C. Martinez, "Exopolymeric substances (EPS) from *Bacillus subtilis*: polymers and genes encoding their synthesis." *Fems Microbiol Lett* **313**, 1 (2010).

[197] P. I. Fields, R. V. Swanson, C. G. Haidaris, and F. Heffron, "Mutants of *Salmonella-Typhimurium* that cannot survive within the macrophage are avirulent." *P Natl Acad Sci USA* **83**, 5189 (1986).

[198] P. De Vos, G. Garrity, D. Jones, N. R. Krieg, W. Ludwig, F. A. Rainey, K. Schleifer, and W. Whitman, *Genus I: Paenibacillus. Bergey's Manual of Systematic Bacteriology: volume 3: The Fermicutes.* (Springer Science and Business Media, 2011).

[199] F. Rodrigues, M. Sarkar-Tyson, S. V. Harding, S. H. Sim, H. H. Chua, C. H. Lin, X. Han, R. K. M. Karuturi, K. Sung, K. Yu, W. Chen, T. P. Atkins, R. W. Titball, and P. Tan, "Global map of growth-regulated gene expression in *Burkholderia pseudomallei*, the causative agent of melioidosis." *J Bacteriol* **188**, 8178 (2006).

[200] V. K. Singh, M. Syring, A. Singh, K. Singhal, A. Dalecki, and T. Johansson, "An insight into the significance of the DnaK heat shock system in *Staphylococcus aureus*." *Int J Med Microbiol* **302**, 242 (2012).

[201] D. L. Popham, B. Illadesaguiar, and P. Setlow, "The *Bacillus subtilis* Dacb gene, encoding penicillin-binding protein-5-asterisk, is part of a 3-gene operon required for proper spore cortex synthesis and spore core dehydration." *J Bacteriol* **177**, 4721 (1995).

[202] M. R. Abdullah, J. Gutierrez-Fernandez, T. Pribyl, N. Gisch, M. Saleh, M. Rohde, L. Petruschka, G. Burchhardt, D. Schwudke, J. A. Hermoso, and S. Hammerschmidt, "Structure of the pneumococcal L, D-carboxypeptidase DacB and pathophysiological effects of disabled cell wall hydrolases DacA and DacB." *Mol Microbiol* **93**, 1183 (2014).

[203] A. Minnen, L. Attaiech, M. Thon, S. Gruber, and J. W. Veening, "SMC is recruited to oriC by ParB and promotes chromosome segregation in *Streptococcus pneumoniae.*" *Mol Microbiol* **81**, 676 (2011).

[204] A. Minnen, F. Burmann, L. Wilhelm, A. Anchimiuk, M. L. Diebold-Durand, and S. Gruber, "Control of Smc coiled coil architecture by the ATPase heads facilitates targeting to chromosomal parb/pars and release onto flanking DNA." *Cell Rep* **14**, 2003 (2016).

[205] H. Zhao, A. L. Clevenger, J. W. Ritchey, H. I. Zgurskaya, and V. V. Rybenkov, "*Pseudomonas aeruginosa* condensins support opposite differentiation states." *J Bacteriol* **198**, 2936 (2016).

[206] V. Haine, A. Sinon, F. Van Steen, S. Rousseau, M. Dozot, P. Lestrate, C. Lambert, J. J. Letesson, and X. De Bolle, "Systematic targeted mutagenesis of *Brucella melitensis* 16M reveals a major role for GntR regulators in the control of virulence." *Infect Immun* **73**, 5578 (2005).

[207] C. Merlin, G. Gardiner, S. Durand, and M. Masters, "The *Escherichia coli* metD locus encodes an ABC transporter which includes Abc (MetN), YaeE (MetI), and YaeC (MetQ). ," *J Bacteriol* **184**, 5513 (2002).

[208] P. Lestrate, R. M. Delrue, I. Danese, C. Didembourg, B. Taminiau, P. Mertens, X. De Bolle, A. Tibor, C. M. Tang, and J. J. Letesson, "Identification and characterization of *in vivo* attenuated mutants of *Brucella melitensis.*" *Mol Microbiol* **38**, 543 (2000).

[209] C. E. Hill, D. S. Metcalf, and J. I. MacInnes, "A search for virulence genes of *Haemophilus parasuis* using differential display RT-PCR." *Vet Microbiol* **96**, 189 (2003).

[210] L. J. Ejim, V. M. D'Costa, N. H. Elowe, J. C. Loredo-Osti, D. Malo, and G. D. Wright, "Cystathionine p-lyase is important for virulence of *Salmonella enterica* serovar *typhimurium.*" *Infect Immun* **72**, 3310 (2004).

[211] N. Gavini, B. S. Hausman, L. Pulakat, R. P. Schreiner, and J. A. Williamson, "Identification and mutational analysis of rfbG, the gene encoding CDP-D-glucose-4,6-dehydratase, isolated from free living soil bacterium *Azotobacter vinelandii.*" *Biochem Bioph Res Co* **240**, 153 (1997).

[212] M. J. Ferreira and I. de Sa-Nogueira, "A multitask ATPase serving different ABC-type sugar importers in *Bacillus subtilis*." *J Bacteriol* **192**, 5312 (2010).

[213] M. A. Webber and L. J. V. Piddock, "The importance of efflux pumps in bacterial antibiotic resistance." *J Antimicrob Chemoth* **51**, 9 (2003).

[214] M. H. Brown, I. T. Paulsen, and R. A. Skurray, "The multidrug efflux protein NorM is a prototype of a new family of transporters." *Mol Microbiol* **31**, 394 (1999).

[215] Y. Morita, K. Kodama, S. Shiota, T. Mine, A. Kataoka, T. Mizushima, and T. Tsuchiya, "NorM, a putative multidrug efflux protein, of *Vibrio parahaemolyticus* and its homolog in *Escherichia coli*." *Antimicrob Agents Ch* **42**, 1778 (1998).

[216] F. El'Garch, K. Jeannot, D. Hoccluet, C. Llanes-Barakat, and P. Plesiat, "Cumulative effects of several nonenzymatic mechanisms on the resistance of *Pseudomonas aeruginosa* to aminoglycosides." *Antimicrob Agents Ch* **51**, 1016 (2007).

[217] A. Beceiro, M. Tomas, and G. Bou, "Antimicrobial resistance and virulence: a successful or deleterious association in the bacterial world?" *Clin Microbiol Rev* **26**, 185 (2013).

[218] K. A. Klein, H. S. Fukuto, M. Pelletier, G. Romanov, J. P. Grabenstein, L. E. Palmer, R. Ernst, and J. B. Bliska, "A transposon site hybridization screen identifies galU and wecBC as important for survival of *Yersinia pestis* in murine macrophages." *J Bacteriol* **194**, 653 (2012).

[219] N. J. Hughes, C. L. Clayton, P. A. Chalk, and D. J. Kelly, "*Helicobacter pylori* porCDAB and oorDABC genes encode distinct pyruvate : flavodoxin and 2-oxoglutarate: acceptor oxidoreductases which mediate electron transport to NADP." *J Bacteriol* **180**, 1119 (1998).

[220] C. Fischer, C. Geourjon, C. Bourson, and J. Deutscher, "Cloning and characterization of the *Bacillus subtilis* prkA gene encoding a novel serine protein kinase." *Gene* **168**, 55 (1996).

[221] X. D. Liu, Y. T. Luo, Z. F. Li, and G. H. Wei, "Functional analysis of PrkA - a putative serine protein kinase from *Mesorhizobium alhagi* CCNWXJ12-2-in stress resistance." *BMC Microbiol* **16**, 227 (2016).

[222] N. Castro-Guerrero, P. K. Sinha, J. Torres-Bacete, A. Matsuno-Yagi, and T. Yagi, "Pivotal roles of three conserved carboxyl residues of the NuoC (30k) segment in the structural integrity of proton-translocating NADH-quinone oxidoreductase from *Escherichia coli*." *Biochemistry-Us* **49**, 10072 (2010).

[223] Z. Q. Shao, R. T. Lin, and E. B. Newman, "Sequencing and characterization of the Sdac gene and identification of the Sdacb operon in *Escherichia coli* k12." *Eur J Biochem* **222**, 901 (1994).

[224] F. Gerard, N. Pradel, and L. F. Wu, "Bactericidal activity of colicin V is mediated by an inner membrane protein, SdaC, of *Escherichia coli*." *J Bacteriol* **187**, 1945 (2005).

[225] A. Feucht, L. Evans, and J. Errington, "Identification of sporulation genes by genome-wide analysis of the sigma(e) regulon of *Bacillus subtilis*." *Microbiol-Sgm* **149**, 3023 (2003).

[226] O. Tenaillon, E. Denamur, and I. Matic, "Evolutionary significance of stress-induced mutagenesis in bacteria." *Trends Microbiol* **12**, 264 (2004).

[227] E. A. Wall, J. H. Caufield, C. E. Lyons, K. A. Manning, T. Dokland, and G. E. Christie, "Specific N-terminal cleavage of ribosomal protein L27 in *Staphylococcus aureus* and related bacteria. ," *Mol Microbiol* **95**, 258 (2015).

[228] W. S. Reznikoff, "Tn5 as a model for understanding DNA transposition." *Mol Microbiol* **47**, 1199 (2003).

[229] H. Taber and G. M. Halfenger, "Multiple-aminoglycoside-resistant mutants of *Bacillus subtilis* deficient in accumulation of kanamycin." *Antimicrob agents ch* **9**, 251 (1976).

[230] I. Bjedov, O. Tenaillon, B. Gérard, V. Souza, E. Denamur, M. Radman, F. Taddei, and I. Matic, "Stress-induced mutagenesis in bacteria." *Science* **300**, 1404 (2003).

[231] D. E. Brash, "UV signature mutations," *Photochem Photobiol* **91**, 15 (2015).

[232] M. L. Del Hoyo, M. Basualdo, A. Lorenzo, M. A. Palacio, E. M. Rodriguez, and E. Bedascarrasbure, "Effect of shaking honey bee colonies affected by American foulbrood on *Paenibacillus larvae larvae* spore loads," *J Apicultural Res* **40**, 65 (2001).

[233] H. Hansen and C. J. Brodsgaard, "Control of American foulbrood by the shaking method." *Apiacta* **38**, 140 (2003).

[234] FAVV, "Dierengezondheid - bijen," http://www.favv.be/bijenteelt/dierengezondheid/ (2017).

[235] A. M. Lambowitz and S. Zimmerly, "Group II introns: Mobile ribozymes that invade DNA," *Cold Spring Harb Perspect Biol* **3**, aDD3616 (2011).

[236] G. M. Dunny and B. A. Leonard, "Cell-cell communication in Gram-positive bacteria." *Annu Rev Microbiol* **51**, 527 (1997).

[237] C. Pourcel, F. André-Mazeaud, H. Neubauer, F. Ramisse, and G. Vergnaud, "Tandem repeats analysis for the high resolution phylogenetic analysis of *Yersinia pestis*." *BMC Microbiol* **8**, 4 (2004).

[238] L. M. Schouls, H. G. van der Heide, L. Vauterin, P. Vauterin, and F. R. Mooi, "Multiple-locus variable-number tandem repeat analysis of dutch *Bordetella pertussis* strains reveals rapid genetic changes with clonal expansion during the late 1990s." *J Bacteriol* **186**, 5496 (2004).

[239] J. L. Koeck, B. M. Njanpop-Lafourcade, S. Cade, E. Varon, L. Sangare, S. Valjevac, G. Vergnaud, and C. Pourcel, "Evaluation and selection of tandem repeat loci for *Streptococcus pneumoniae* MLVA strain typing." *BMC Microbiol* **16**, 66 (2005).

[240] S. Svraka, R. Toman, L. Skultety, K. Slaba, and W. L. Homan, "Establishment of a genotyping scheme for *Coxiella burnetii*." *FEMS Microbiol Lett* **254**, 268 (2006).

[241] P. Le Flèche, I. Jacques, M. Grayon, S. Al Dahouk, P. Bouchon, F. Denoeud, K. Nöckler, H. Neubauer, L. A. Guilloteau, and G. Vergnaud, "Evaluation and selection of tandem repeat loci for a *Brucella* MLVA typing assay." *BMC Microbiol* **9** (2006).

[242] F. Lista, G. Faggioni, S. Valjevac, A. Ciammaruconi, J. Vaissaire, C. le Doujet, O. Gorgé, R. De Santis, A. Carattoli, A. Ciervo, A. Fasanella, F. Orsini, R. D'Amelio, C. Pourcel, A. Cassone, and G. Vergnaud, "Genotyping of *Bacillus anthracis* strains based on automated capillary 25-loci multiple locus variable-number tandem repeats analysis." *BMC Microbiol* **6**, 6 (2006).

[243] J. W. Marsh, M. M. O'Leary, K. A. Shutt, A. W. Pasculle, S. Johnson, D. N. Gerding, C. A. Muto, and L. H. Harrison, "Multilocus variable-number tandem-repeat analysis for investigation of *Clostridium difficile* transmission in hospitals." *J Clin Microbiol* **44**, 2558 (2006).

[244] F. B. Gilbert, A. Fromageau, J. Lamoureux, and B. Poutrel, "Evaluation of tandem repeats for MLVA typing of *Streptococcus uberis* isolated from bovine mastitis." *BMC Vet Res* **17**, 2 (2006).

[245] C. Pourcel, P. Visca, B. Afshar, S. D'Arezzo, G. Vergnaud, and N. K. Fry, "Identification of variable-number tandem-repeat (VNTR) sequences in *Legionella pneumophila* and development of an optimized multiple-locus VNTR analysis typing scheme." *J Clin Microbiol* **45**, 1190 (2007).

[246] B. A. Lindstedt, L. T. Brandal, L. Aas, T. Vardund, and G. Kapperud, "Study of polymorphic variable-number of tandem repeats loci in the ecor collection and in a set of pathogenic *Escherichia coli* and *Shigella* isolates for use in a genotyping assay." *J Microbiol Methods* **69**, 197 (2007).

[247] R. Gierczyński, A. Golubov, H. Neubauer, J. N. Pham, and A. Rakin, "Development of multiple-locus variable-number tandem-repeat analysis for *Yersinia enterocolitica* subsp. *palearctica* and its application to bioserogroup 4/o3 subtyping." *J Clin Microbiol* **45**, 2508 (2007).

[248] H. Vu-Thien, G. Corbineau, K. Hormigos, B. Fauroux, H. Corvol, A. Clément, G. Vergnaud, and C. Pourcel, "Multiple-locus variable-number tandem-repeat analysis for longitudinal survey of sources of *Pseudomonas aeruginosa* infection in cystic fibrosis patients." *J Clin Microbiol* **45**, 3175 (2007).

[249] L. Diancourt, V. Passet, C. Chervaux, P. Garault, T. Smokvina, and S. Brisse, "Multilocus sequence typing of *Lactobacillus casei* reveals a clonal population structure with low levels of homologous recombination." *Appl Environ Microbiol* **73**, 6601 (2007).

[250] S. Védy, E. Garnotel, J. L. Koeck, F. Simon, S. Molinier, and A. Puidupin, "Interest of new molecular typing method in the study of hospital transmitted *Staphylococcus aureus* population," *Pathol Biol* **55**, 382 (2007).

[251] B. A. Lindstedt, W. Tham, M. L. Danielsson-Tham, T. Vardund, S. Helmersson, and G. Kapperud, "Multiple-locus variable-number tandem-repeats analysis of *Listeria monocytogenes* using multicolour capillary electrophoresis and comparison with pulsed-field gel electrophoresis typing." *J Microbiol Methods* **72**, 141 (2008).

[252] P. Francois, A. Hochmann, A. Huyghe, E. J. Bonetti, G. Renzi, S. Harbarth, C. Klingenberg, D. Pittet, and J. Schrenzel, "Rapid and high-throughput genotyping of *Staphylococcus epidermidis* isolates by automated multilocus variable-number of tandem repeats: a tool for real-time epidemiology." *J Microbiol Methods* **72**, 296 (2008).

[253] J. Top, R. Willems, S. van der Velden, M. Asbroek, and M. Bonten, "Emergence of clonal complex 17 *Enterococcus faecium* in the netherlands." *J Clin Microbiol* **46**, 214 (2008).

[254] L. R. Martinez, B. Harris, W. C. t. Black, R. M. Meyer, P. J. Brennan, V. D. Vissa, and R. L. Jones, "Genotyping North American animal *Mycobacterium bovis* isolates using multilocus variable number tandem repeat analysis." *J Vet Diagn Invest* **20**, 707 (2008).

[255] B. Lu, C. Pourcel, J. H. Liu, H. Y. Dong, Y. Y. Zhang, Y. Jiang, Z. G. Liu, X. Q. Zhao, and K. L. Wan, "Establishment of the standard operation program on multiple loci variable numbers tandem repeats analysis typing on *Mycobacterium tuberculosis*," *Zhonghua Liu Xing Bing Xue Za Zhi* **29**, 919 (2008).

[256] E. Harth-Chu, R. T. Espejo, R. Christen, C. A. Guzmán, and M. G. Höfle, "Multiple-locus variable-number tandem-repeat analysis for clonal identification of *Vibrio parahaemolyticus* isolates by using capillary electrophoresis." *Appl Environ Microbiol* **75**, 4079 (2009).

[257] M. Kimura, R. M. Sakamuri, N. A. Groathouse, B. L. Rivoire, D. Gingrich, S. Krueger-Koplin, S. N. Cho, P. J. Brennan, and V. Vissa, "Rapid variable-number tandem-repeat genotyping for *Mycobacterium leprae* clinical specimens." *J Clin Microbiol* **47**, 1757 (2009).

[258] A. Radtke, B. A. Lindstedt, J. E. Afset, and K. Bergh, "Rapid multiple-locus variant-repeat assay (MLVA) for genotyping of *Streptococcus agalactiae*." *J Clin Microbiol* **48**, 2505 (2010).

[259] P. Bourhy, L. Collet, S. Clément, M. Huerre, P. Ave, C. Giry, F. Pettinelli, and M. Picardeau, "Isolation and characterization of new *Leptospira* genotypes from patients in Mayotte (Indian Ocean)." *PLoS Negl Trop Dis* **4**, e724 (2010).

[260] C. S. Teh, K. H. Chua, and K. L. Thong, "Multiple-locus variable-number tandem repeat analysis of *Vibrio cholerae* in comparison with pulsed field gel electrophoresis and virulotyping." *J Biomed Biotechnol* **817190** (2010).

[261] C. Pourcel, F. Minandri, Y. Hauck, S. D'Arezzo, F. Imperi, G. Vergnaud, and P. Visca, "Identification of variable-number tandem-repeat (VNTR) sequences in *Acinetobacter baumannii* and interlaboratory validation of an optimized multiple-locus VNTR analysis typing scheme." *J Clin Microbiol* **49**, 539 (2011).

[262] S. L. Kruy, H. van Cuyck, and J. L. Koeck, "Multilocus variable number tandem repeat analysis for *Salmonella enterica* subspecies." *Eur J Clin Microbiol Infect Dis* **30**, 465 (2011).

[263] O. J. Brevik, K. F. Ottem, and A. Nylund, "Multiple-locus, variable number of tandem repeat analysis (MLVA) of the fish-pathogen *Francisella noatunensis*." *BMC Vet Res* **7**, 5 (2011).

[264] Y. Wang, R. J. Skilton, L. T. Cutcliffe, E. Andrews, I. N. Clarke, and P. Marsh, "Evaluation of a high resolution genotyping method for *Chlamydia trachomatis* using routine clinical samples." *PLoS One* **6**, e16971 (2011).

[265] C. Guo, Y. Liao, Y. Li, J. Duan, Y. Guo, Y. Wu, Y. Cui, H. Sun, J. Zhang, B. Chen, Q. Zou, and G. Guo, "Genotyping analysis of *Helicobacter pylori* using multiple-locus variable-number tandem-repeats analysis in five regions of China and Japan." *BMC Microbiol* **11**, 197 (2011).

[266] R. Nakao, L. J. Morrison, L. Zhou, J. W. Magona, F. Jongejan, and C. Sugimoto, "Development of multiple-locus variable-number tandem-repeat analysis for rapid genotyping of *Ehrlichia ruminantium* and its application to infected *Amblyomma variegatum* collected in heartwater endemic areas in Uganda." *Parasitology* **139**, 69 (2012).

[267] K. Laroucau, A. Di Francesco, F. Vorimore, S. Thierry, J. L. Pingret, C. Bertin, H. Willems, G. Bölske, and R. Harley, "Multilocus variable-number tandem-repeat analysis scheme for *Chlamydia felis* genotyping: comparison with multilocus sequence typing." *J Clin Microbiol* **50**, 1860 (2012).

[268] C. A. N'guessan, S. Brisse, A. C. Le Roux-Nio, S. Poussier, D. Koné, and E. Wicker, "Development of variable number of tandem repeats typing schemes for *Ralstonia solanacearum*, the agent of bacterial wilt, banana moko disease and potato brown rot." *J Microbiol Methods* **92**, 366 (2013).

[269] J. Halkilahti, K. Haukka, and A. Siitonen, "Genotyping of outbreak-associated and sporadic *Yersinia pseudotuberculosis* strains by novel multilocus variable-number tandem repeat analysis (MLVA)." *J Microbiol Methods* **95**, 245 (2013).

[270] S. Cesbron, J. Pothier, S. Gironde, M. A. Jacques, and C. Manceau, "Development of multilocus variable-number tandem repeat analysis (MLVA) for *Xanthomonas arboricola* pathovars." *J Microbiol Methods* **100**, 84 (2014).

[271] T. Abayneh, D. J. Colquhoun, D. Austin, and H. Sorum, "Multilocus variable number tandem repeat analysis of *Edwardsiella piscicida* isolates pathogenic to fish." *J Fish Dis* **37**, 941 (2014).

[272] Y. Hauck, C. Soler, P. Gérôme, R. Vong, C. Macnab, G. Appere, G. Vergnaud, and C. Pourcel, "A novel multiple locus variable number of tandem repeat (VNTR) analysis (MLVA) method for *Propionibacterium acnes*." *Infect Genet Evol* **33**, 233 (2015).

[273] S. Ciarroni, L. Gallipoli, M. C. Taratufolo, M. I. Butler, R. T. Poulter, C. Pourcel, G. Vergnaud, G. M. Balestra, and A. Mazzaglia, "Development of a multiple loci variable number of tandem repeats analysis (MLVA) to unravel the intra-pathovar structure of *Pseudomonas syringae pv. actinidiae* populations worldwide." *PLoS One* **10**, e0135310 (2015).

[274] B. Bakhshi, B. Bayat, and A. R. Lari, "Multiple- locus variable-number tandem-repeat analysis (MLVA) of *Shigella sonnei* isolates of 2012 outbreak I. R. Iran." *Microb Pathog* **102**, 69 (2017).

[275] J. Guinard, A. Latreille, F. Guérin, S. Poussier, and E. Wicker, "New multilocus variable-number tandem-repeat analysis (MLVA) scheme for fine-scale monitoring and microevolution-related study of *Ralstonia pseudosolanacearum* phylotype I populations." *Appl Environ Microbiol* **83**, e03095 (2017).