Differentiation of Virulent and Biological Control *Paenibacillus larvae* Strains Associated with American Foulbrood in Bee Hives

Benjamin Grayham Saville

For the degree of PhD
The University of York
Biology

September 2011

Abstract

The detection and control of American Foulbrood (AFB) is made more challenging due to a lack of evidence regarding the virulence mechanisms employed by this honeybee pathogen. Whilst incidence of this pathogen within the UK has recently declined, ~100 colonies were identified as infected with AFB in 2011 (to end of September), so AFB should still be considered a serious threat to honeybee health. It is known that within the species many phenotypes exist, and the infection caused by the phenotypes differs greatly. This PhD thesis presents several advances towards a greater understanding of the intra-specific differences occurring within the species. Chapter 2 evaluates the use of 16S rRNA sequencing as a method of *Paenibacillus larvae* identification, as well as exploring the use of this ribosomal subunit for differentiation of the species. The sequencing of two housekeeping (*purH* and *PyrE*) genes assesses the potential of a Multi Locus Sequence Typing (MLST) method, as a means of subspecies differentiation. Chapter 3 assesses what can be inferred from use of Enterobacterial Repetitive Intergenic Consensus (ERIC) sequence fingerprinting, with regards to prior knowledge genetic differences. Lateral Flow Devices (LFDs), a commonly used diagnostic tool, are tested to ensure *P. larvae* isolates representing all 4 ERIC types are detected. In Chapter 4 an *in-vitro* honeybee rearing method is employed to observe the correlation between proteolytic activity of isolates and *in-vitro* virulence. The method is applied to a wider range of reference isolates, to observe the intra-species differences existing. The ability to produce large numbers of viable spores is explored as a potential difference existing between ERIC types I and IV. Whole genome shotgun sequencing is used in chapter 5 to perform comparative genomic studies on 4 *P. larvae* isolates also utilising 646 contigs from a previous sequencing project. The possible presence of plasmid DNA is explored, through GC content analysis. The genetic basis of a sporulation phenotypic difference is examined by BLAST analysis of orthologous genes. In Chapter 6 the findings of this thesis are discussed in more detail, and potential areas of further study are identified.

List Of Contents

List of Figures

List of Tables

Acknowledgements

I would firstly like to thank my supervisors, Thorunn Helagson and Richard Thwaites, for the support they have given throughout my PhD. They have provided me with valuable suggestions, but also allowed a great degree of freedom allowing me to develop as a researcher. I am truly grateful for the time that they have spent mentoring me and overseeing this project.

I would also like to thank my Thesis Advisory Panel (TAP) members, Peter Young and James Moir. Their critical appraisal of my work encouraged me to develop as a scientist, and improve the strength of arguments made.

I could not have kept going during the many long days in the laboratory, if it was not for the kind support of Jan Bew and Jayne Hall. They helped to keep me sane during larvae grafting. I'm especially grateful that they did not mind the influx of honeybee larvae in the incubators and occasional escaped bee around the lab.

I would like to thank the National Bee Unit (NBU) for the support they have provided; the team provided me with an expert working knowledge of AFB and taught me how to identify AFB symptoms, as well as showing me how to handle a frame full of bees. Without the help of Selwyn Wilkins, Ben Jones, Damian Cierniak and Jack Wilford it would not have been possible to carry out the exposure bioassay work. I am grateful for the time and effort they put in to helping me develop this method.

I would like to thank members of J1 and D0 lab groups for welcoming me to their lab meetings, and providing valuable feedback of the data that I presented. Their helpful suggestions provided new areas of study, which might otherwise have been overlooked. I would also like to thank members of the J1 lab group for welcoming me in to their office, despite my infrequent visits I was always made to feel welcome.

Without the help of Celina Whalley and Naveed Aziz in the Technology Facility and Ummey Hany at FERA, it would not been possible to collect the genomic data from the 4 isolates. A huge thank you to Rachel Glover at FERA who provided kind support in developing my knowledge of computational analysis using Perl, Python and Bioperl and for taking the time to explain things in a manner that I could understand.

Without the funding from The University of York and FERA seedcorn this project would not have been possible, I am very grateful for the original selection and subsequent funding of this project which allowed me to study at York. I am also grateful of the additional financial support offered by FERA, allowing the use of 454 pyrosequencing to explore the genomic relationship between isolates.

Finally I would like to say an enormous thank you to my family and friends who have been my lifeline throughout this PhD. To my parents David and Sandra Saville and my in-laws John and Diane Smith, who have provided me with both financial and emotional support. My best friend James Ruddick, who was always there to provide distraction during my times of need. And most importantly to my wife, Melanie Saville, (and bump (Joseph Benjamin Saville, born 23/01/12) for always being there for me during the peaks and troughs of this studentship, without the kind support of my devoted wife I do not think I would have had the strength to take on (and complete) this challenge. (And to all of the above, for finally realising I'm not a beekeeper!) Without the support of these people, it would not have been possible to complete this work.

Authors Declaration

I, Benjamin Saville, hereby certify that all the work contained within this thesis is a result of my own work and was written by myself.

Benjamin Grayham Saville
The University of York

Chapter 1
General Introduction

1.1. An overview of American Foulbrood

The Office International des Epizooties (OIE), the Organisation for Animal Health classified American Foulbrood (AFB) as belonging to list B, defined as 'transmissible diseases which are considered to be of socio-economic and/or public health importance within countries and which are significant in the international trade of animals and animal products' (de Graaf et al. 2006a). Despite its name, AFB is found worldwide (Matheson 1993), and it appears AFB can be found anywhere that bees are reared in artificial hives. Some studies have been carried out on feral colonies (Fries et al. 2006), but in countries with kept bees, so it is unknown if the disease originates in the feral colonies or is transferred to them. How the disease spreads is not well understood: beekeepers and bee keeping practices are believed to spread the disease (Brown et al. 2007). The epidemiology of the disease in feral colonies is not understood, neither are the origins of the disease. It is unlikely that beekeeping practices are solely responsible for the transmission of the disease, as historically the spread of AFB between countries cannot be explained by beekeeping practices as the production of honey and sourcing of hive components and bees would be performed locally. The More Economically Developed Countries (MEDCs) policy of importing rather than producing queens and honey is likely to have had an effect on the transmission of AFB. Previously transmission was likely indirect or accidental (via stowaway insects etc.) rather than direct transmission that is now possible due to the importing of items such as bees and honey (Ritter 2003).

1.1.1. AFB Symptoms

Typical symptoms of AFB are brown or black larvae that leave very viscous remains of that same colour. These remains can be drawn out to approximately

one inch using a matchstick, often referred to as the “ropiness” test. Beyond this point the remains start to dry out and reduce in size until a very small and well-adhered patch of spores (termed a “Scale”) is barely visible at the bottom of a cell. These are rarely seen as infections are usually caught early on. It is estimated that these scales can contain up to 100,000 AFB spores (Brown et al. 2007).

1.1.2. Infectious dose of AFB

AFB infects honeybee larvae rather than adult bees (Wilson 1971), and can only infect during the first 14 days of larval development. Beyond this time larvae are able to withstand extremely high doses of spores with no infection developing (Williams 2000). Spores can be transported in honey, providing a medium in which they can survive for long periods of time (Dustmann and von der Ohe 1997; Haseman 1961). If the honey is used to feed hives then any spore contamination could result in a diseased hive. Twenty spores constitute the LD_{50} of a one-day-old larva, however millions of spores are needed to infect a 4-5 day old larva (Ratnieks 1992). Until recently studies have not been carried out on honey to check for AFB spores, even though many countries collect honey samples to check its quality (Antúnez et al. 2004). In a recent Uruguayan survey 52 out of 101 samples contained AFB spores (Antúnez et al. 2004), showing that contaminated honey has the potential to be very widespread and of global importance with a high volume of international trade.

1.1.3. Probable Route of Infection

It is possible to hypothesise this model of infection; young larvae ingest the spores and once in the mid gut they germinate and grow due to the larval diet of sugars (Haydak 1970). Here the bacteria are able to thrive with the larval diet and gut conditions providing near optimum growth conditions (Yue et al. 2008). After massive proliferation the bacteria are transferred across the gut wall and are internalised within the larvae (Yue et al. 2008). In the absence of the sugary larval diet the conditions become unfavorable for vegetative cells so levels of

sporulation increase. Upon sporulation proteases are produced (Dancer and Chantawannakul 1997) digesting the larvae to support vegetative growth or provide a protective environment for the storage of spores. The proteolytic action soon digests the entire larvae and eventually all that is left is the decaying remains (Genersch et al. 2005). These remains dry to form AFB scales that adhere tightly to the bottom of the cells (Genersch et al. 2006).

1.1.4. AFB Resistance

Although resistance to AFB can occur naturally within the environment, a recent study suggests that the 'cost' of such resistance may affect survival rates, development rates or possibly productivity levels as an adult worker bee (Evans and Pettis 2005). This could at least partly explain why it is managed colonies that become infected with AFB. In the wild, resistant phenotypes would be able to become dominant, as they would be able to survive an AFB infection. New AFB resistant queens will be created, and this in turn will lead to new AFB resistant colonies. However commercial beekeepers carry out this natural process themselves by selective breeding, to produce better colonies or increase productivity. Breeding for increased productivity, may come at the cost of greater susceptibility to AFB and other diseases (Alippi 2001).

1.2. *Paenibacillus larvae*: the causative agent of AFB

Paenibacillus larvae is a gram positive elongated rod shaped bacterium (Lauro et al. 2003). The ability to form spores is one of the reasons it has such devastating effects on honeybees (*Apis mellifera*) where it causes the disease American Foulbrood (AFB) (Ashiralieva and Genersch 2006). *P. larvae* produces small spherical colonies when grown on Brain Heart Infusion agar. The colonies are cream in colour with a very smooth appearance. Some strains of *P.larvae* are able to produce a red pigment and often produce colonies ranging from cream with a red centre to almost entirely red colonies, these colonies are often slightly

smaller in size but share the same morphological characteristics in all other aspects (Heyndrickx et al. 1996).

1.3. Taxonomy of *Paenibacillus larvae*

Bacillus larvae (White 1906) was originally identified as the causative agent of AFB. Katznelson (1950) isolated bacteria from an infected hive that had initially been diagnosed as AFB, the symptoms produced however were somewhat different. Upon laboratory culturing and analysis this organism was classified as *Bacillus pulvifaciens* however, it was not determined whether *B. pulvifaciens* was a contaminant after death or the causative agent.

1.3.1. Revision of the *Bacillus* genus

These classifications remained for many years until it was realised that the genus *Bacillus* was extremely divergent. For organisms to be placed in the *Bacillus* genus they were required to be rod shaped and to produce spores under aerobic conditions. However the phenotypic characteristics of the genus ranged from alkalophiles to strict aerobes and thermophiles to halophiles. Researchers looked for new ways to classify the organisms within this genus, In 1993 newly available technologies were used to study rRNA sequences, by examining the differences in the ~1.5kb subunit they were able to show that several species believed to be from the *Bacillus* genus formed a distinct out-group (Ash et al. 1993). They proposed that this out-group should be re-classified as the genus *Paenibacillus*. The Latin name *Paenibacillus* refers to the Latin word *paene* meaning nearly or almost, this reflects the close relationship in existence between *Bacillus* and *Paenibacillus.* At this time several species including *Bacillus larvae* and *Bacillus pulvifaciens* were transferred to this new genus. On phylogenetic trees produced it was clear to see that *P.larvae* and *P.pulvifaciens* were more closely related than other strains belonging to the new genus (Heyndrickx et al. 1996; Shida et al. 1997a; Shida et al. 1997b).

Characteristic or test	*P. larvae*	*P. pulvifaciens*
Motility[1]	V	V
Ellipsoidal spores[1]	+	+
Central and subterminal spores[1]	+	+
Swollen sporangia[1]	+/-	+/-
Anaerobic growth[1]	+	+
Thiamine requirement[1]	+	-
Esculin hydrolysis[1]	+	+
Arginine dihydrolase[1]	-	+
Casein hydrolysis[1]	+	+
Citrate utilization[1]	-	-
Hydrogen sulfide production[1]	-	-
Indole production[1]	-	-
Gelatin liquefaction[1]	+	+
Nitrate reduction[1]	V	V
Urease[1]	-	-
Voges-Proskauer[1]	V	V
Acid from;		
N-Acetylglucosamine, D-glucose, glycerol, D-mannose, ribose, and D-trehalose[1]	+	+
Mannitol[1]	-	+
Salicin[1]	+	-
D-Fructose, D-galactose, methyl-D-glucoside, and D-tagatose[1]	V	V
16S Sequence Similarity[2]	~99%	

Table 1.1: Biochemical and 16S rRNA differences exhibited by strains of *Paenibacillus larvae* and *Paenibacillus pulvifaciens* (V = variable result). [1] (Heyndrickx et al. 1996) [2] (Genersch et al. 2006)

1.3.2. Subspecies Differentiation?

Heyndrickx (1996) continued research on *Paenibacillus* and showed that *P. larvae* and *P. pulvifaciens* were, as suggested, more closely related and that two species actually belonged to the same species. These were reclassified as subspecies termed *Paenibacillus larvae* subspecies *larvae* (Pll) and

Paenibacillus larvae subspecies *pulvifaciens* (Plp). This reflected relatively minor differences observed between the groupings (table 1.1) combined with the observation that subspecies *larvae* was able to cause AFB and subspecies *pulvifaciens* caused a relatively benign disease called Powdery Scale. Differentiation to subspecies level was often considered to be an obstacle in the diagnosis of AFB, due to a high degree of similarity between the subspecies (de Graaf et al. 2006b).

1.3.3. Latest Taxonomic Revision of *P. larvae*

Genersch *et al*. (2006) dropped the subspecies differentiation be dropped as new methods and experimental data suggested the previous subspecies grouping did not reflect the organisms' virulence. Most studies on the differences of *Paenibacillus* strains focused on host tolerance and disease outcomes as these have been far easier to study; studies of pathogenic mechanisms and virulence factors have been limited by a lack of phenotypic studies of the organism (Genersch et al. 2005). A combination of data from several fingerprinting techniques, 16S rRNA sequencing, and API biochemical testing kits and, importantly, pathogenicity tests were carried out showing that several Plp strains were capable of being equally pathogenic and in some cases exhibited greater pathogenicity than Pll. As a result of these findings they concluded that the subspecies differentiation should be dropped and all isolates would be labelled as *Paenibacillus larvae* (Genersch et al. 2006). However in 2007 more data emerged which contradicted this, using SDS-PAGE it was found that all isolates of *Paenibacillus larvae* gave the same patterns excluding Plp which had at least 3 protein bands that were absent from the rest of the results suggesting a significant difference in the organisms (Antúnez et al. 2007), however this study failed to perform pathogenicity tests on the isolates used. Without this data, the true relevance of those differences is not known. To the present date all isolates originally identified as *Bacillus larvae* or *Bacillus pulvifaciens* are now known as just one species *Paenibacillus larvae,* with no subspecies differentiations.

1.3.4. Isolate 6993: A non-pathogenic isolate?

A previous project (Watkins et al. 2003) identified a strain of *P. larvae* as non-pathogenic. This work focused around the creation of a bacterial bio-control strain, capable of protecting colonies from European Foulbrood (EFB) infection. Numerous tests showed that larvae fed with the bacterial spores of isolate 6993, were not susceptible to EFB infections. It was proposed that this strain should be further tested and validated as a bio-control for EFB. It was noted in some instances increased larval mortality were recorded; this was attributed to the incorrect dose of bacterial spores fed to the infant larvae. It was suggested that death was due to massive proliferation of bacteria within the larval gut but typical AFB symptoms were not displayed and none of these deaths were attributed to AFB infections. Shortly after the findings of this project, the subspecies differentiation was dropped. Further work was halted, as it would not be possible to validate a bio-control isolate that held the potential to cause AFB.

1.3.5. Other Bacterial species found in hives

Honeybee hives are not normally sterile environments; a wealth of bacterial species are routinely isolated from hive materials (Disayathanoowat et al. 2011; Kaltenpoth 2011; Martinson et al. 2011). Many of the species of bacteria are environmental isolates potentially found on plants or in soil; it is believed that these species are transported into the hive by foraging bees. Once within the hive the sugary and warm conditions provided are near optimum for most bacteria, these invading species colonise the area. Most of these species are not capable of causing disease, and are present in a commensal relationship. A study of the natural gut flora found in healthy Japanese honeybees (*Apis cerana japonica*), showed the presence of bacteria from the following genera *Bacillus*, *Sphingomonas*, *Bartonella*, *Simonsiella*, *Serratia* and *Providencia* (Yoshiyama and Kimura 2009). Whilst this represents a different country and a different species of honeybee, it still reflects the diversity of the normal gut flora. A study of honeybees affected by Colony Collapse Disorder (CCD) showed that bacteria

were present from *Lactobacillus*, Firmicutes, *Bifidobacterium*, *Bartonella*, *Gluconacetobacter*, *Simonsiella* and Gammaproteobacteria (Cox-Foster et al. 2007). Whilst this metagenomic study represents unhealthy honeybees, it is not believed that a bacterium is responsible for CCD (Higes et al. 2009; Maori et al. 2009). This highlights that honeybees around the world are likely to possess similar normal gut floras, with many similarities between the genera represented.

1.3.6. *Paenibacillus alvei:* Honeybee friend or foe?

A commonly isolated bacterium from hives is *P. alvei* (Djordjevic et al. 2000); the true relationship between honeybees and this bacterium is unknown. Whilst *P. alvei* can be isolated from healthy colonies, it is more likely to be isolated from colonies infected with AFB or EFB. *P. alvei* is unlikely to grow in the gut of healthy larva, instead becoming established in infected colonies (Bailey 1963). It has been suggested that this bacterium does not possess the ability to cause infection within infant larvae, but is a secondary infection capable of outcompeting pathogenic bacteria once the larva has succumbed to the infection (Hornitzky and Anderson 2003). *P. alvei* is capable of producing similar decaying larvae to that expected from an AFB infection (Djordjevic et al. 2000). No prevalence data exists for this bacterium, and it is not considered a notifiable disease. It is unknown what proportion of AFB or EFB infected hives contain this bacterium, and what affect its presence has on the symptoms observed. This organism is only detected if the causative agent of AFB or EFB is to be isolated for further study, as it is not an organism screened for routinely. *P. alvei* was also determined to be the cause of cellulitis in a Korean immunocompetent patient (Shin et al. 2005). Intravenous drug users or immune compromised patients were found to have bacteraemia caused by *Paenibacillus* (*Bacillus*) *alvei*, their weakened immune systems were predicted to be the reason for infection with a secondary honeybee larvae disease (Reboli et al. 1989). Despite these isolated incidents, *P. alvei* is not considered to be a human pathogen.

1.4. The Pathogenicity of *P. larvae*

Unfortunately over the years the organisms' pathogenicity has been neglected. Very little work has been carried out to discover why the organism causes such a deadly disease in young honey bee larvae. *P. larvae* has a very limited host range, no reports exist of *P. larvae* infecting other insects. It has been reported that intravenous drug users in Germany developed bacteraemia, after preparing methadone or illicit drugs with *P. larvae* contaminated honey (Rieg et al. 2010). These were isolated incidents and all subjects infected had previous medical conditions making them more susceptible to infection. *P. larvae*, like *P. alvei*, is not considered to be a human pathogen.

Few if any studies have researched the underlying mechanism controlling virulence and to date, no virulence genes or even reliable virulence markers have been explicitly identified. It was originally thought that Pll caused AFB and that Plp caused a relatively benign disease Powdery scale. The subspecies grouping whilst not implicitly defined by virulence traits was determined to be an accurate representation of virulence. Reclassification of the isolates to exclude the subspecies differentiation, due to all tested isolates exhibiting *in-vitro* pathogenicity, suggests that previously observed differences between subspecies cannot be solely linked to the virulence of the organism.

P. larvae forms large numbers of oval endospores within just a few days of infection (Yue et al. 2008). As noted in other endospore forming bacteria, the spores produced are highly resistant to heat and chemical agents and can survive in the environment for several decades (Setlow 2006). Only the endospores of Pll are capable of causing disease (Bakonyi et al. 2003). As with other devastating spore forming organisms, such as *Bacillus anthracis*, the spores remain in a dormant phase until more favourable conditions are detected.

A recent study utilised fluorescence *In-Situ* Hybridisation (FISH) (Yue et al. 2008) to observe the interactions between *P. larvae* and the honeybee digestive system during the course of an AFB infection. This was able to provide the best

overview of an AFB infection and the time course of an infection. Massive proliferation within the larval gut was shown, before *P. larvae* penetrated epithelial junctions causing a systemic, fatal, infection (Yue et al. 2008). The reason for penetration at the epithelial junctions still remains speculative due to a lack of supporting evidence.

1.4.1. The proteolytic activity of *P. larvae*

Both *B. thuringiensis* and *P. larvae* are insect larva pathogens producing similar liquefied larval remains. The symptoms caused by American Foulbrood would suggest the involvement of at least one protease working in a similar way to those found in *Bacillus thuringiensis* (Brar et al. 2007; Li and Yousten 1975; Oppert 1999). There are two schools of thought regarding the production and involvement of proteases in AFB infection. There is a consensus that proteases are involved, but there is no clear picture of the timing of production. Some suggest they are produced upon sporulation (Dancer and Chantawannakul 1997), which has been demonstrated in many *Bacillus sp.* (Andrews et al. 1985). Other research groups state that the production of proteases is more likely during the vegetative cell stage, also observed in a variety of *Bacillus sp.* (Tjalsma et al. 2004).

Dancer and Chantawannakul (1997) were the first to suggest the involvement of metalloproteases, optimum conditions for these enzymes match those provided by the larval tissues and of the larval food sources. Metalloproteases have been linked to the virulence of several species of pathogenic bacteria from the Genus *Bacillus* (Chung et al. 2006; Fricke et al. 2001; Kim et al. 2004; Li and Yousten 1975; Yu and Kroos 2000). A more recent study of metalloproteases in *P. larvae* has taken further steps towards identifying proteolytic activity as a major virulence factor. Using immunofluorescence and FISH shows that the protease is present in abundance within the vegetative cells and is released upon sporulation existing on the spore surface. The study went on to show damage to larval gut cells when the metalloprotease expressing cells were present compared to healthy gut cells in a non-infected larva (Antúnez et al. 2011). This study lacked

the production of a knockout mutant unable to express this metalloprotease. This would have proved the involvement of this particular metalloprotease, rather than simply implying its involvement.

1.4.2. Is virulence plasmid mediated?

Many other *bacilli* are able to cause disease due to the presence of virulence plasmids. *Bacillus anthracis* is able to cause anthrax a deadly disease of animals and humans mainly due to the presence of pX01 and pX02 plasmids containing a number of genes essential for survival and genes for toxin production (R Okinaka et al. 1999; Van der Auwera et al. 2005). *Bacillus thuringiensis* causes similar symptoms to *P. larvae* but in a wider range of insects due to the presence of plasmid acquired δ-endotoxin crystals genes (Van der Auwera et al. 2005). Logically it is assumed that plasmids could be key to the virulence of *P. larvae*, however to date none of the 4 plasmids discovered can be linked with the virulence of the organism. The only plasmid to studied in detail, exists predominantly in North America; pMA67 confers tetracycline resistance (Murray et al. 2007) and is believed to be a response to the US method of dosing hives with oxytetracycline at the first signs of infection. To date there is no published evidence that British isolates of *P. larvae* have been found to contain plasmid pMA67 or any other plasmids.

1.5. Detection of American Foulbrood outbreaks

In the UK AFB has had a variable incidence rate over the last ten years, with a peak of ~275 infected colonies in 2002 numbers have since declined (excluding a brief rise in 2009), and at present ~100 colonies have been identified with AFB infections for 2011 (https://secure.fera.defra.gov.uk/beebase/public/BeeDiseases/trendDiseaseChart.cfm?id=10). At the first sign of infection, inspectors are informed and they undertake a visit to the apiary. The inspectors perform a visual inspection of the suspect colonies. After many years of experience in the field, inspectors can successfully diagnose a colony based solely upon visual inspection. To confirm

the infection a Lateral Flow Device (LFD) is used. If a positive result has been recorded, a sample is sent on to the lab at the National Bee Unit for further confirmation. A simple spore stain visualising the spores under a light microscope provides confirmation of an AFB positive colony.
Bacteria can also be isolated from infected hive material. Once isolated further methods are required to confirm that the isolated bacterium is *P. larvae*, as other spore forming bacteria such as *Bacillus sp.* can also be isolated in this way. 16S rRNA sequencing seems to be the most reliable method so far.

Although now rejected as a valid infra-specific classification of the species, the previous subspecies differentiation can provide useful information on intra-specific variation amongst bacterial isolates. It has been shown that this differentiation does not reflect virulence but it has been shown that isolates identified as subspecies *pulvifaciens* are a divergent group of isolates within the species (Genersch et al. 2006).

Currently the most efficient method of laboratory-based detection for bacterial and viral pathogens is considered to be PCR, and increasingly, Real-Time PCR. This proves problematic with species such as *P. larvae* due to high levels of sequence similarity both intra-species and within the genus (Heyndrickx et al. 1996; Shida et al. 1997a; Shida et al. 1997b). Finding unique sequence data to design the primers and probe to can be very difficult. Several sets of published primers are available, however published data suggests they require further testing to ensure correct strain identification. A primer set is able to detect *P. larvae* isolates, although there is no data to show that the primers have been tested against former Plp, the closest relation tested appears to be *P.alevi* (Govan et al. 1999). A set of primers able to identify subspecies larvae isolates (Alippi et al. 2004) and a *P. larvae* nested PCR (Lauro et al. 2003) were more rigorously tested against a wider range of bacterial species. Potential is shown here for the ability to create a PCR based testing method but as it is not yet fully understood which isolates are pathogenic, it remains unknown what such an assay should target.

The goal for PCR based *P. larvae* identification is to create an assay to detect pathogenic strains of *P. larvae;* this would require a pathogen-specific sequence that is unique to pathogenic strains. This requires primers that target the genes directly responsible for virulence, or another marker consistently linked to virulence. To date, no such markers are identified, and indeed intra-species differences in pathogenicity are not understood.

1.6. Control of American Foulbrood

Under the Bee Diseases and Pests Control (England) Order, AFB is a notifiable disease and the National Bee Unit at FERA carries out monitoring and control of the disease (Brown et al. 2007). The current method of control in the UK is destruction by incineration which has been an effective method of control since it was introduced in the 1940s bringing down the yearly incidence rate from several thousand to around 60 in 2006 (Brown et al. 2007). The main problem with disinfecting the hives is penetration of the spores into the wooden structures; this renders techniques such as scorching with a blowtorch or the use of most chemical disinfectants useless, as they are simply unable to penetrate the wooden fibres (Dobbelaere et al. 2001).

1.6.1. Control of AFB using Antibiotics

Some countries such as the USA, favour the use of antibiotics as a method of control. As a result of the prolonged use of antibiotics, through a series of horizontal transfers *Paenibacillus larvae* has acquired the resistance plasmid pMA67 (Murray et al. 2007). The resistance to oxytetracyline is spreading and the use of antibiotics as a method of control is becoming even less effective (Murray et al. 2007). The use of antibiotics from the tetracycline family also causes other problems including larval mortality and retarded growth (Peng et al. 1992). The use of the antibiotic Tylosin has also been proven to be far more effective and far less lethal to the brood, but oxytetracycline is still the only antibiotic approved for prevention or treatment of AFB in the USA (Peng et al. 1996).

1.6.2. Chemical control of AFB

Ethylene Oxide (ETO) has been used in some countries as a method of sterilising hive equipment post AFB infection and has shown varying results (Takahashi et al. 2001). Some studies report no, or very few, recurring infections after fumigation, where as other reports say recurring infections were predominantly found (Takahashi et al. 2001). The problems lie with the fumigation procedures. Unless a fully sealed non-penetrable container is used then the process is rendered useless. Makeshift polyethylene sealed fumigation containers are unreliable as ETO is able to pass through the polyethylene. Successful trials have usually occurred in a laboratory environment using a container such as an autoclave for fumigation. ETO does not penetrate biological material very well, so spores within capped cells and in honey are often unharmed by this process and able to cause a repeat infection. ETO is toxic and expensive; this coupled with the varying results has ruled it out as a viable control measure.

1.6.3. Control of AFB by irradiation of hives

Irradiation of hives and hive parts by Gamma radiation showed a 99.9994% chance that all spores had been left non-viable (Takahashi et al. 2001). No ill effects were observed in the honey subsequently produced. This process is already commercially available in many places for the food industry, so protocols could easily be adapted to suit the requirements (Takahashi et al. 2001). It has been suggested that hives could be subjected to irradiation twice a year to remain free of many harmful pathogens. Public perception seems to be the limiting factor here, and honey produced could be avoided by companies and the general public which would make bee keepers reluctant to use this method (Takahashi et al. 2001).

1.6.4. Use of biological controls against AFB infections

The use of a biological control for *P. larvae* has been suggested and studies have been carried out to look at the effectiveness of bacterial strains as control agents. *B. cereus, B. megaterium, B. licheniformis, B. pumilus, B. subtilus and B. laterosporus* were studied and found to inhibit Pll with varying levels of success (Alippi and Reynaldi 2006). There are significant barriers to the approval of such a method, so it is unlikely that in the near future this will be pursued.

1.7. Sequencing the *P. larvae* genome

As a significant honeybee pathogen, sequencing of *P. larvae* (Strain BRL-230010) began in 2003 (Qin et al. 2006). This project was completed in 2006 but a complete genome was never produced. Instead the genome data exists as a series of 646 contigs with an average length of 6817bp (Acesssion: AARF00000000). The contigs were combined to produce 349 scaffolds, the largest of which is 137,418 bp (Qin et al. 2006). With total scaffold coverage of 4,015,777bp in a genome estimated to be approximately 4-5Mb, it is highly unlikely that these assemblies cover the entire genome.

The sequencing of the *P. larvae* genome was augmented with further Sanger sequence data and Illumina reads (Chan et al. 2011). This was completed in February 2011; this sequencing project has produced a series of 353 contigs with an average length of 12.3 Kb. This sequencing attempt covers 4,352,378bp; once more the combined length of contigs that could be assembled equals the approximate genome size. Further work is required to complete these genomes, and without the availability of a complete genome to act as a reference for future assemblies.

1.8. Can intra-species differentiations reflect the virulence exhibited by an isolate?

Although the taxonomy of *P. larvae* has undergone several revisions since its original description in 1906, it was not until 2006 when it was experimentally shown that the taxonomy of the organism did not reflect virulence (Genersch et al. 2006). Whilst it has been shown that isolates previously defined as subspecies *pulvifaciens* have the potential to cause AFB, it is unknown if strain 6993 is a non-pathogenic isolate as previously suggested. This thesis aims to determine if strain 6993 should be considered to be a pathogenic isolate and as a result develop a better understanding of the intra-specific differences existing between *P. larvae* isolates and how these relate to virulence.

Whilst it has been shown that isolates exhibit a great level of sequence identity (Genersch et al. 2006), high levels of phenotypic differentiation are also evident (Heyndrickx et al. 1996). Despite the suggestion that all isolates of *P. larvae* are pathogenic, it seems logical to assume non-pathogenic isolates exist. Isolates originally identified as *B. pulvifaciens* were shown to cause non-AFB related symptoms. This describes experiments aimed at differentiating isolates of *P. larvae* based upon observed phenotypic differences. The observed differences were compared to the pathogenicity of the isolates, for an improved understanding of the intra-specific groupings produced.

Recent advances in genome sequencing allow for large-scale genetic comparisons to be made from a single sequencing run. By utilising this technology it will be possible to establish if detected phenotypic differences are related to genetic differences or are related to gene expression.

1.9. Overview of Chapters

Chapter 2 - Intra-species differentiation by means of 16S rRNA sequencing and development of a Multi Locus Sequence Typing Scheme

It has been shown that isolates exhibit ~99% sequence similarity of the 16S rRNA subunit but phylogenetic study of this region has not been fully explored as a method of discrimination between isolates. Multi Locus Sequencing Typing (MLST) has been successfully used to differentiate closely related bacterial species in instances where 16S rRNA phylogenies were unable. Similar schemes have been widely documented within the genus *Bacillus*, but have yet to be used for the differentiation of *P. larvae* isolates.

Chapter 3 - Intra-species differentiation by means of PCR based fingerprinting, and its impact on UK detection

DNA fingerprinting often exhibits higher levels of variation than that observed in coding areas of the genome. These methods target the highly variable inter-genic regions; previously it has been shown that *P. larvae* isolates can be clustered into 4 distinct groups using Enterobacterial Repetitive Intergenic Consensus (ERIC) sequence fingerprinting. Whilst not related to the virulence of *P. larvae* it was believed that the ERIC groupings reflected the differences in virulence. Lateral Flow devices (LFDs) used for primary identification in the UK were based upon these groupings. Here we investigate if LFDs are capable of detecting all AFB outbreaks within the UK.

Chapter 4 - *In-vitro* comparison of virulence displayed by *Paenibacillus larvae* isolates

Whilst data surrounding the epidemiology of isolates does exist, full details of the level of pathogenicity of individual isolates are unknown. Data exists recalling the various symptoms seen at time of isolation; the time course of the infection or the severity of the resulting infection was never documented. This

resulting in isolates previously being classified as non-pathogenic *P. pulvifaciens* or pathogenic *P. larvae*. Whilst very few isolates of *P. pulvifaciens* were isolated, the pathogenicity of these isolates has been questioned. This methodology will allow direct comparisons to be made between isolates exhibiting different phenotypic properties.

Chapter 5 - Comparative genomics of *Paenibacillus larvae* isolates

Recent advances in sequencing technology allow for the acquisition of large quantities of genomic data from a single sequencing run. Whilst unlikely that a complete *P. larvae* genome will be produced, the wealth of genomic data will allow for comparative genomics to assess the genomics differences in existence between isolates. Whilst a complete genome of *P. larvae* does not exist, a wealth of genomic data from a known isolate is publically available. Use of this previous sequencing attempt as a reference allows for a further comparative analysis to be made against a previously published isolates.

Chapter 2
Intra-species differentiation by means of 16S rRNA sequencing and development of a Multi Locus Sequence Typing Scheme

2.1. Summary

- 16S rRNA sequencing is the preferential method of species identification. It is quick and does not provide ambiguous identification as achieved through use of combined biochemical screening and colony morphology or Fatty Acid Profiling (FAP).
- A maximum sequence divergence of <2.5% was observed in the 16S rRNA subunit of 26 *P. larvae* isolates.
- A Multi Locus Sequence Typing (MLST) approach using *purH* and *PyrE* genes revealed average sequence similarity of 99.71% and 98.7% respectively. This showed that, on average, *purH* exhibited greater sequence identity between isolates than observed in the 16S rRNA subunit.
- 16S rRNA and MLST phylogenies were able to group isolates according to previous subspecies differentiation; therefore suggesting a genetic difference exists between the subspecies.
- It remains unpredictable what phenotypes are represented by the subspecies differentiation, so it is unknown what the intra-species grouping displayed by 16S rRNA or MLST phylogenies reflects.

2.2. Introduction

Over the ~100 years since its first isolation the taxonomy of *Paenibacillus larvae* has been revised multiple times (Ash et al. 1993; Genersch et al. 2006; Heyndrickx et al. 1996; Katznelson 1950; Nakamura 1984; White 1906). Prior to 2006, isolates were separated in two subspecies (Heyndrickx et al. 1996). Subspecies *larvae* were believed to be responsible for all AFB infections, whereas subspecies *pulvifaciens* were believed to be responsible for a benign disease called Powdery Scale (Katznelson 1950; Nakamura 1984). In 2006 it was shown that subspecies *larvae* and *pulvifaciens* were genetically closer related than previously assumed (Genersch et al. 2006). Several strains of (non-pathogenic) subspecies *pulvifaciens* were able to cause AFB and in some instances the resultant infection was more severe than (pathogenic) subspecies *larvae* strains (Genersch et al. 2006). It was suggested the subspecies differentiation was dropped and all strains regardless of subspecies differentiation were grouped as *Paenibacillus larvae* (Genersch et al. 2006).

Whilst identification of AFB infections in honeybee colonies had been improved by the creation of field based test kits (Vita Europe Ltd.), less progress was made with laboratory-based identification of the causative agent. Bacterial identification was originally due to phenotypic differences and observed colony morphology, no specific media exists for the isolation of *P. larvae*, and many bacterial species are able to produce identical colonies. Many molecular diagnostic methods were developed for the identification of *P. larvae* (Alippi et al. 2004; Govan et al. 1999; Lauro et al. 2003), however since the removal of the subspecies differentiation, the validity of these testing methods is unknown. Several methods were developed to exclude isolates previously identified as subspecies *pulvifaciens* (Alippi et al. 2004; Lauro et al. 2003) and one assay did not include subspecies *pulvifaciens* in the validation process (Govan et al. 1999). An accurate method was required for the identification of isolated *Paenibacillus* strains, testing using fatty acid profiling (FAP) (Stead 1988) was only able to identify isolates as belonging to the *Paenibacillus* genus (data not shown).

2.2.1. 16S rRNA Identification of *P. larvae*

The use of DNA sequence as a method of exploring the phylogenetic relationship between organisms was first realised in 1965 (Zuckerkandl and Pauling 1965). Sequence similarity existing between 16S sequences from different *Bacillus sp.* was soon noted (Dubnau et al. 1965). Due to the constant revisions in bacterial taxonomy, identification of bacterial species by means of biochemical and morphological analysis, became increasingly difficult. Each taxonomic revision altered the parameters by which each species was identified. Identification became laborious and often-ambiguous results could result in incorrect identification. In the field of molecular microbiology sequencing of the 16S ribosomal RNA subunit has become common practice (Olsen and Woese 1993). Use of the 16S subunit is considered an accurate representation of bacterial evolution as no change in function has been noted and it is present in all bacterial species (Janda and Abbott 2007). The 1,500bp subunit is considered a good length for computational analysis (Patel 2001). Comparisons of 16s rRNA sequences have revealed that isolates, to be considered of the same species, have a minimum sequence identity of 97% (Janda and Abbott 2007). As a result, 16S rRNA sequencing is considered the gold standard molecular microbiological identification tool.

2.2.2. 16S rRNA as a means of intra-species discrimination

Studies of the 16S subunit in *B. cereus* show very low levels of nucleotide variation existed between *B. cereus, B. anthracis and B. thuringiensis* (Sacchi et al. 2002). This suggested that sequencing of the 16S subunit would be unable to differentiate within a species, as required by this project. Unlike *B. cereus*, the 16S subunit of *P. larvae* has not been widely used for the discrimination of isolates. It has been suggested that *P. larvae* isolates share more than 97% 16S sequence homogeny (Genersch et al. 2006). The frequency and position of base substitutions in the *P. larvae* 16S sequence has not been documented and therefore the resulting phylogeny produced is also unknown.

2.2.3. Application of Multi Locus Sequence Typing (MLST) to intra-species differentiation

It has been suggested that isolates possessing 97.5% 16S sequence identity, potentially share as little as 60% genome sequence identity (Stackebrandt and Goebel 1994). Whilst this incorporates non-coding sections of the bacterial genomes, it suggests that a greater degree of genetic variability will also be observed in sections of the genome coding for other genes. *B. cereus, B. anthracis* and *B. thuringiensis* have been shown to possess such little genetic variability, that they should be considered the same species (Helgason et al. 2000). The phenotypic characteristics of the three species are considered to be highly diverse, despite the lack of genomic variability. Multi locus sequence-typing (MLST) has been used to differentiate between *B. cereus*, *B. anthracis* and *B. thuringiensis* with great levels of accuracy (Helgason et al. 2004). MLST involves the sequencing of several housekeeping genes, these genes are considered to be less conserved than the 16S rRNA subunit. The greater degree of genetic variability present across multiple genes can be successfully used to differentiate between species with very low genetic variability (Helgason et al. 2004). Several suitable genes were identified from similar schemes that had been used within the *Bacillus* genus, allowing for the differentiation of isolates that 16S rRNA sequencing failed to differentiate (Helgason et al. 2004; Marston et al. 2006; Sorokin et al. 2006). It is predicted that 16S rRNA variability between isolates of *P. larvae* will be very low, despite the diverse phenotypic traits of the various strains being well documented. It is believed that a similar scheme developed for *P. larvae*, could have similar results to the schemes developed for the *Bacillus* genus.

2.3. Aims

1. To assess the use of 16S rRNA subunit sequencing as a method of identifying *P. larvae*.
2. To assess the use of 16S sequencing for the intra-species discrimination of *P. larvae*.
3. To develop an MLST scheme for the intra-species discrimination of *P. larvae* isolates

2.4. Materials and Methods.

2.4.1. Bacterial Isolates

A panel of 39 bacterial isolates were chosen this included reference isolates from the LMG culture collection belonging to *Paenibacillus larvae* subspecies *larvae*, *Paenibacillus larvae* subspecies *pulvifaciens* and *Paenibacillus alvei* (table 2.3). Additional 16S sequences for other species were obtained from the NCBI database. All isolates were cultured at 34°C on BHI-T media (described in appendix 1); DNA was extracted using a standard Promega gram positive bacterial extraction method and stored at -20°C (described in appendix 2).

2.4.2. Loci – primer design

Many sets of primers have been published which amplify the whole or a partial section of the 16S rRNA amplicon. UFP1 and URP1 (table 2.2) were selected due to amplification of the whole 1.5kb subunit. This primer pair was widely used by FERA for the identification of environmental samples, whilst untested on *P. larvae* they had successfully identified many isolates on which conventional methods had failed.

The MLST scheme developed here for *P. larvae* is based upon several similar schemes used to differentiate *B. cereus, B. anthracis* and *B. thuringiensis*

(Helgason et al. 2004; Marston et al. 2006; Sorokin et al. 2006). tBLAST searches were performed on the array of genes used in *Bacillus* MLST schemes, looking for orthologous sequences within the 646 BCM P. larvae contigs. Two potential target genes were identified (table 2.1).

Gene	Function / Product
PyrE	Orotate phosphoribosyltransferase
purH	Final enzymes in the de novo purine biosynthesis pathway

Table 2.1: genes selected for initial MLST sequencing scheme, chosen due to essential nature and availability of gene sequence within *P. larvae* genome data.

Nucleotide sequences were imported into text files, and primers were manually designed for gene amplification (table 2.2).

Target	F Primer	R Primer	Annealing temp (°C)	Amplicon size (bp)
16S	UFP1 5'- AGT TTG ATC CTG GCT CAG -3' *	URP1 5'- GGT TAC CTT GTT ACG ACT T -3' *	53.1	1500
pyrE	pyrE-F 5'- TCT TCG GGG TTA TCG AT -3' **	pyrE-R 5'- CAA GAG CAG CAA TGG AGA ACA -3' **	57.6	650
purH	purH-F 5'- AAC CCG CAT ACC ATC CGC T -3' **	purH-R 5'- GAA TGC GGC TCT GCA ACT G -3' **	58.8	730

Table 2.2: Details of the primers used for direct sequencing of 16S rRNA subunit and PyrE and purH MLST genes (all primers used at a final concentration of 0.3pmol). * FERA, Un-published primers. ** MLST primers designed within this project.

2.4.3. PCR and sequencing

PCR amplification of both the 16S subunit and the MLST genes used a standard 25µl bacterial PCR mastermix (documented in appendix 5). The following thermal profile was used for amplification 2 min at 95°C followed by 35 cycles of 35s at 95°C, 1Min at annealing temp (Table 2.2), 1Min at 72°C. This was followed by a final extension of 7Min at 72°C. PCR purification used a standard bench top centrifuge protocol from Qiagen (Appendix 6). Macrogen, Korea carried out sequencing. All samples submitted met the submission guidelines detailed.

2.4.4. Data analysis

Upon receiving sequence data, the electropherograms were studied using APe (http://biologylabs.utah.edu/jorgensen/wayned/ape/). The quality thresholds for base calling were adjusted. Trimmed sequences meeting the quality criteria were used in subsequent analysis.

Sequence data was imported into BioEdit (http://www.mbio.ncsu.edu/bioedit/bioedit.html), sequence data obtained from the reverse primer was reverse complemented. The sequences were now assembled using the CAP3 (Huang 1999) plugin for BioEdit using standard parameters. Assembled sequences were analysed against the NCBI nucleotide database to ensure that the processed sequence files still returned the correct BLAST identity.

FASTA sequences were imported into eBioX (http://www.ebioinformatics.org/ebiox/) before aligning using standard clustalW parameters. Bootstrap neighbour joining trees were constructed using ClustalX2 (Larkin et al. 2007) and were exported in phylip format. Phylip trees were produced based upon 10,000 bootstrap trials.

Aligned sequences were exported as fasta files and the alignment statistics were calculated using the –Alistat function of the HMMer software package (http://hmmer.janelia.org/).

Identification	PROTECT	NCPPB	Depositor	Isolated By	Geographic location	Biological Origin
Paenibacillus alevi	6558	LMG13255	Caledonian University	N. Logan	N/A	N/A
Paenibacillus alevi	6559	LMG13260	Caledonian University	J. Buissiere	Lyon, France	N/A
Paenibacillus alevi	6562	LMG17052	ATCC	N/A	N/A	N/A
Paenibacillus larvae ssp. Larvae	6576	LMG16250	NRRL	N. Logan	N/A	N/A
Paenibacillus larvae ssp. Larvae	6254	LMG9820	ATCC	E. Holst	N/A	Foulbrood of honeybees
Paenibacillus larvae ssp. Larvae	6255	LMG14425	ATCC	H. Shimanuki	Ohio, USA	Diseased honeybee larvae
Paenibacillus larvae ssp. Larvae	6256	LMG14426	ATCC	H. Shimanuki	Virginia, USA	Diseased honeybee larvae
Paenibacillus larvae ssp. Larvae	6257	LMG16147	Apicultural Research Institute, Czech Republic	V. Drobnikova	Czech Republic	Foulbrood of honeybees
Paenibacillus larvae ssp. Larvae	6259	LMG16241	NRRL	N. Logan	N/A	Diseased Honeycomb
Paenibacillus larvae ssp. Larvae	6260	LMG16250	NRRL	N. Logan	N/A	N/A
Paenibacillus larvae ssp. Pulvifaciens	6261	LMG16251	NRRL	N. Logan	N/A	Powdery Scale
Paenibacillus larvae ssp. Pulvifaciens	6262	LMG14428	ATCC	G. Skyring	N/A	N/A
Paenibacillus larvae ssp. Pulvifaciens	6263	LMG15974				
Paenibacillus larvae ssp. Pulvifaciens	6264	LMG16247	NRRL	H. Katznelson	N/A	Powdery Scale
Paenibacillus larvae ssp. Pulvifaciens	6265	LMG16249	NRRL	N. Logan	N/A	N/A

Table 2.3.

Paenibacillus larvae ssp. Pulvifaciens	6266	LMG16252	NRRL	N. Logan	N/A	Dead Honeybee larvae
Paenibacillus larvae ssp. Larvae	6678	LMG16241	NRRL	N. Logan	N/A	Diseased Honeycomb
Paenibacillus sp.	6815	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6817	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6818	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6819	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6820	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6821	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6832	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6833	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6834	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6835	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6836	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6837	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6870	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6873	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6911	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6912	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6914	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6929	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6930	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6931	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus sp.	6932	N/A		FERA	UK Environmental	LFD +ve Colony
Paenibacillus larvae ssp. pulvifaciens	6993	LMG14427	ATCC	N/A	N/A	N/A

Table 2.3: Detailed history of isolates used within this thesis. All known (relevant) information is documented. Unfortunately details available for several reference isolates were not complete.

2.5. Results

2.5.1. Species identification using 16S rRNA sequencing

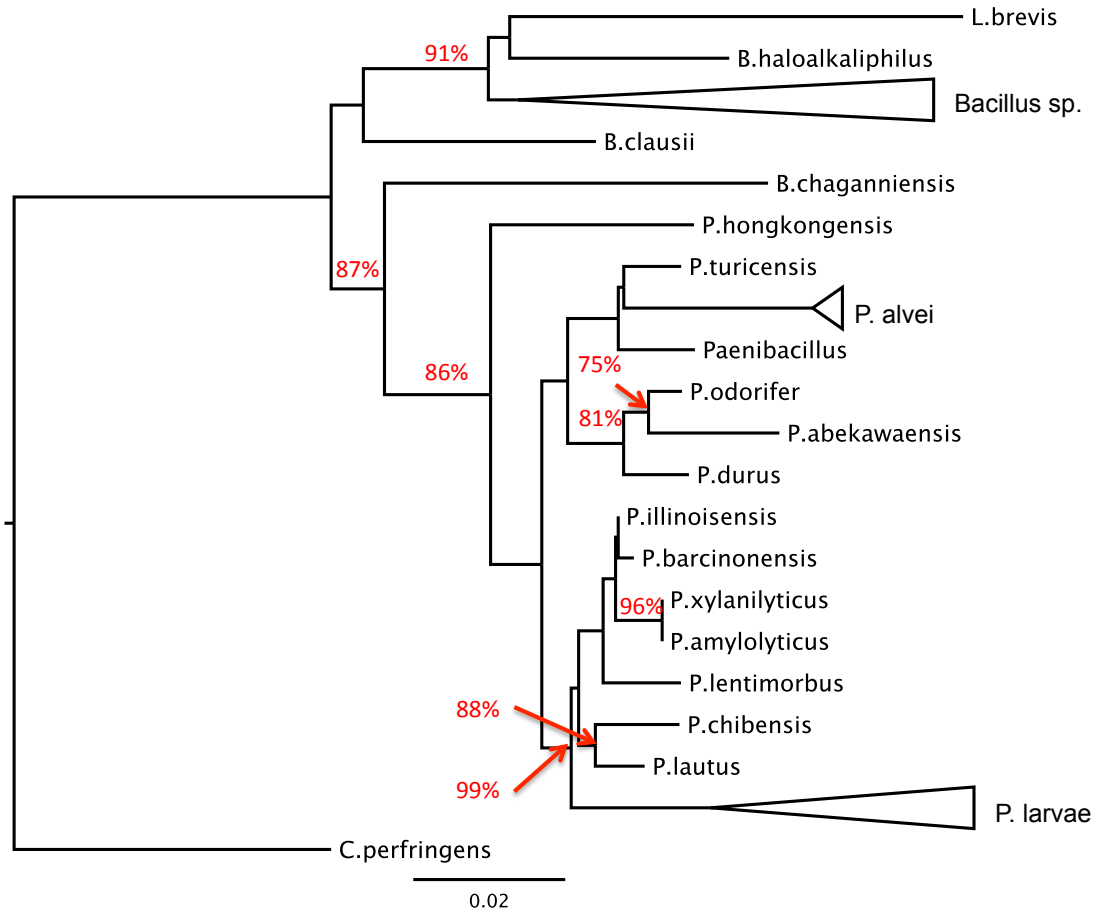


Figure 2.1: Neighbour-joining bootstrap analysis of the 16S rRNA phylogenetic relationship between species belonging to the *Bacillales* order. Tree rooted against type strain of *Clostridium perfringens.* Bootstrap support over 70% shown in red.

Whilst the 16S subunit is highly conserved amongst *P. larvae* isolates, a greater level of variation can be seen between *P. larvae* isolates and other bacterial species (figure 2.1). By expanding the range of species included in the alignment it was possible to calculate the similarity that exists between the 16S sequences of these isolates. This alignment revealed that *P. larvae* isolates share a minimum of 89.9% sequence identity with isolates identified as *P. alvei*. It was shown that the average minimum

sequence identity with *P. alvei* was 90.51%. Whilst the initial sequence identity figure for the *Bacillus* genus was 83.17%, this increased to 84.32% when *Bacillus haloalkaliphilus* was removed from the alignment. As expected *Clostridium perfringens* shared the least sequence identity at 82.3%.

Production of a bootstrapped tree provided a visual representation of the alignment statistics. With *Paenibacillus* isolates forming a separate clade to isolates belonging to more distant genera. Even within a larger sample set the *P. larvae* isolates formed an outgroup from the *Paenibacillus* clade. With most *Paenibacillus* species being of soil / environmental origins, it is clear that *P. larvae* has followed a different evolutionary pathway.

2.5.2. Intra-species discrimination using 16S rRNA sequencing

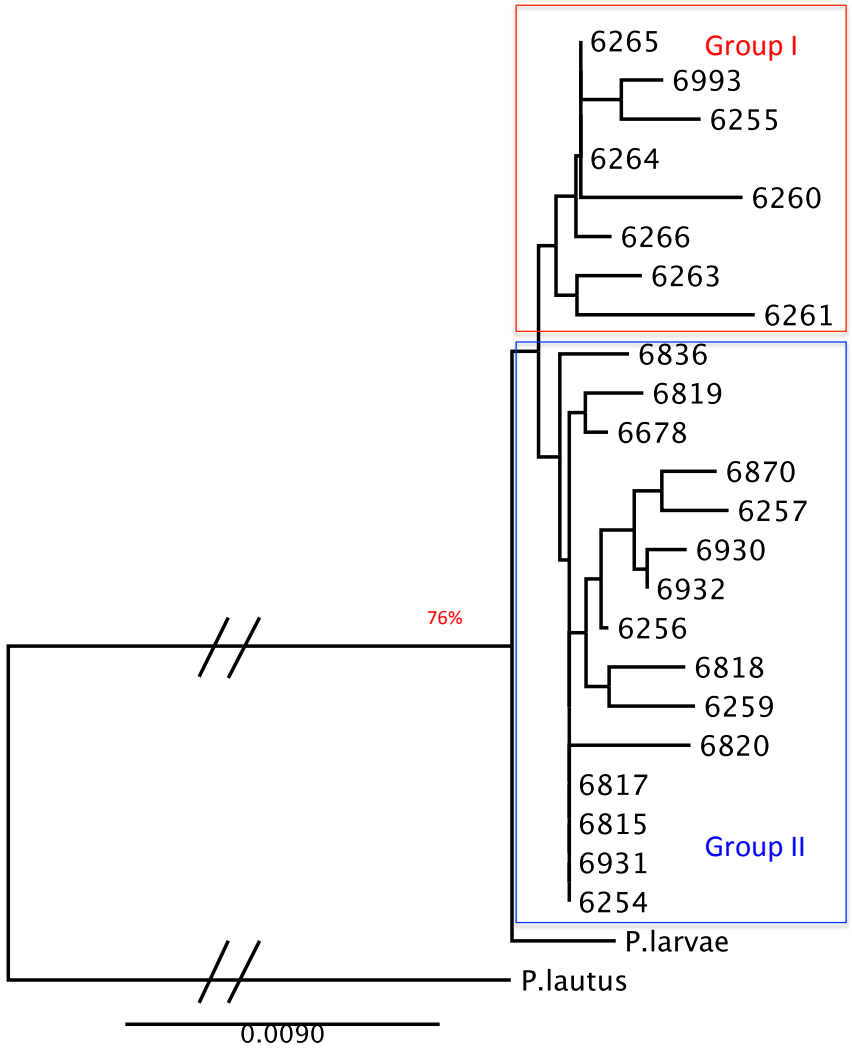


Figure 2.2: Neighbour-joining bootstrap analysis of the 16S rRNA phylogenetic relationship between *P. larvae* isolates, tree shown rooted against type strain of *Paenibacillus lautus*. Group I contains isolates formerly identified as subspecies *pulvifaciens*, Group II contains isolates formerly identified as subspecies *larvae*. Bootstrap support over 70% shown in red. Branches dissected by parallel diagonal lines have been reduced to a third of the original length.

The amplified fragment of the 16S rRNA subunit was 1.46kb, trimming of sequence data not meeting the quality threshold set reduced this to 1.42kb. Analysis of the aligned sequences revealed that *P. larvae* isolates had an average of 86.72% identity

with *P. lautus*. Re-aligning the sequences without *P. lautus* allowed for better analysis of the identity between *P. larvae* isolates (data not shown). Isolates were shown to have an average of 98.84% sequence identity (range over 99.9% to 97.5%).

A bootstrap tree of the alignment highlighted the potential for 16S rRNA sequencing to be used to differentiate subspecies *pulvifaciens* from subspecies *larvae* (figure 2.2). Group 1 contains all reference isolates belonging to subspecies *pulvifaciens* with the exception of 6262. Also contained within the clade are strains 6255, 6260 and 6261 all were previously labelled as subspecies *larvae*. Group 2 contains isolates that were previously labelled as subspecies *larvae*.

2.5.3. Developing a Multi Locus Sequence Typing Method for *P. larvae*

Genes were selected from previous successful MLST schemes used in *Bacillus cereus*, and primers were design as detailed in 2.5.2. Initially 5 primer pairs were designed, two of these primer pairs failed to amplify the target sequence despite optimising the PCR reaction whilst one primer pair amplified multiple bands. The primer pairs for purH and *PyrE* amplified the gene sequence successfully; prior to re-designing the primer pairs for the extra genes, it was decided to assess the variability present in the selected genes.

2.5.4. Genetic diversity within the *purH* gene

Sequence identity between isolates of the purH gene was much lower between *P. larvae* and *B. cereus* than had been found between the organisms 16S subunit. The average sequence identity between the *P. larvae* isolates and *B. cereus* was 62.69%. The increase in sequence divergence between genera was not displayed intra-species, with an average sequence identity of 99.71%. The maximum sequence identity was 100% whilst the minimum of 95.1% was lower than displayed in the 16S sequences, this level of identity was only displayed between isolates 6264 and 6816.

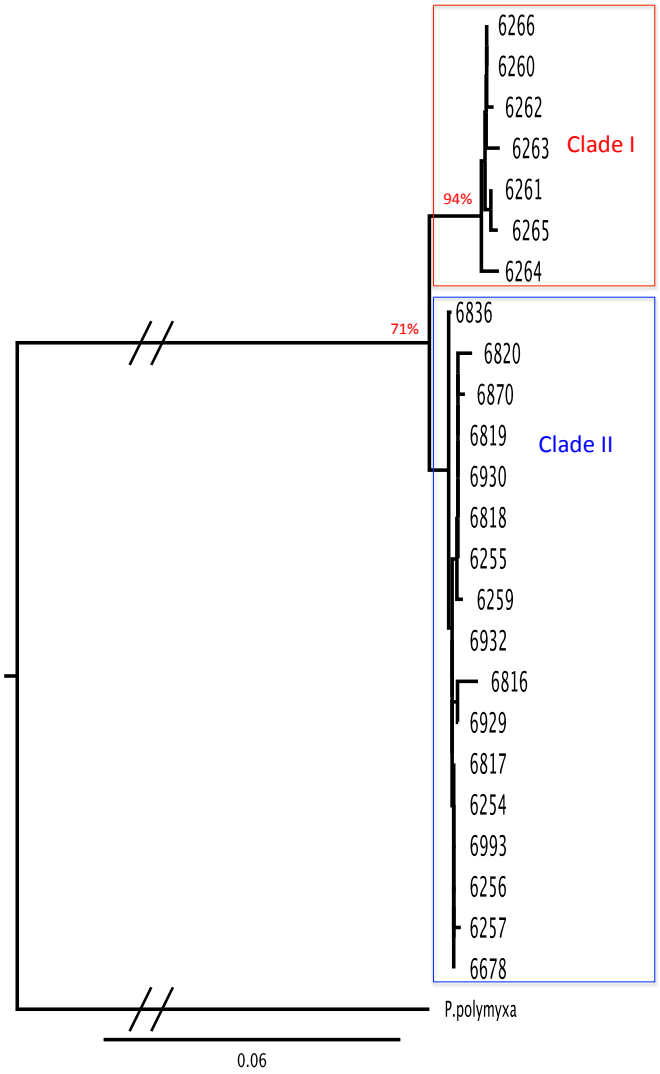


Figure 2.3: Neighbour-joining bootstrap analysis of the *purH* phylogenetic relationship between *P. larvae* isolates, tree shown rooted against type strain of *Paenibacillus polymyxa.* Clade I contains isolates formerly identified as subspecies *pulvifaciens*, Clade II contains isolates formerly identified as subspecies *larvae*. Bootstrap support over 70% shown in red. Branches dissected by parallel diagonal lines have been reduced to a third of the original length.

The bootstrap analysis produced a tree with two distinct clades (figure 2.3); the production of this grouping matched the previously defined subspecies differentiation with subspecies *pulvifaciens* appearing in clade 1 and subspecies larvae in clade 2. 6993 was the only subspecies *pulvifaciens* isolate to appear in the incorrect clade. Whilst 6260 and 6261 were the subspecies *larvae* isolates appearing in the wrong clade.

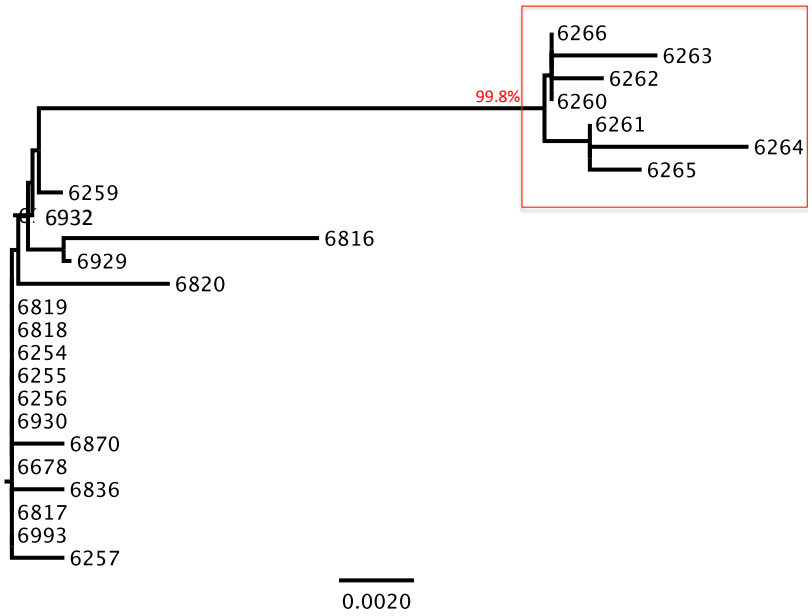


Figure 2.4: Neighbour-joining bootstrap analysis of the *purH* phylogenetic relationship between *P. larvae* isolates. The red box contains isolates formerly identified as subspecies *pulvifaciens*. Bootstrap support over 70% shown in red.

Despite this project successfully sequencing 24 isolates of *P. larvae* all containing this highly similar section of DNA from the purH gene. No significant matches could be found in the NCBI database with over 70% sequence identity, as a result the rooted tree produced has a large root with compresses the 24 sequenced isolates. The unrooted bootstrap tree (figure 2.4) allows for the observation of the phylogenetic differences existing between the 24 strains without the compression caused by rooting the tree.

2.5.5. Genetic diversity within the *PyrE* gene

Sequence identity between isolates of the PyrE gene was lower between *P. larvae* and *B. cereus* than had been found between the organisms 16S subunit and the *purH* gene. The average sequence identity between the *P. larvae* isolates and *B. cereus* was 60.93%. This lowest level of sequence identity was not reflected between *P. larvae* isolates, with a higher level of intra-species similarity than displayed in the purH gene. The average identity between *P. larvae* isolates was 99.01%, with the lowest level of identity being 98.7% and the greatest level of identity being 100%.

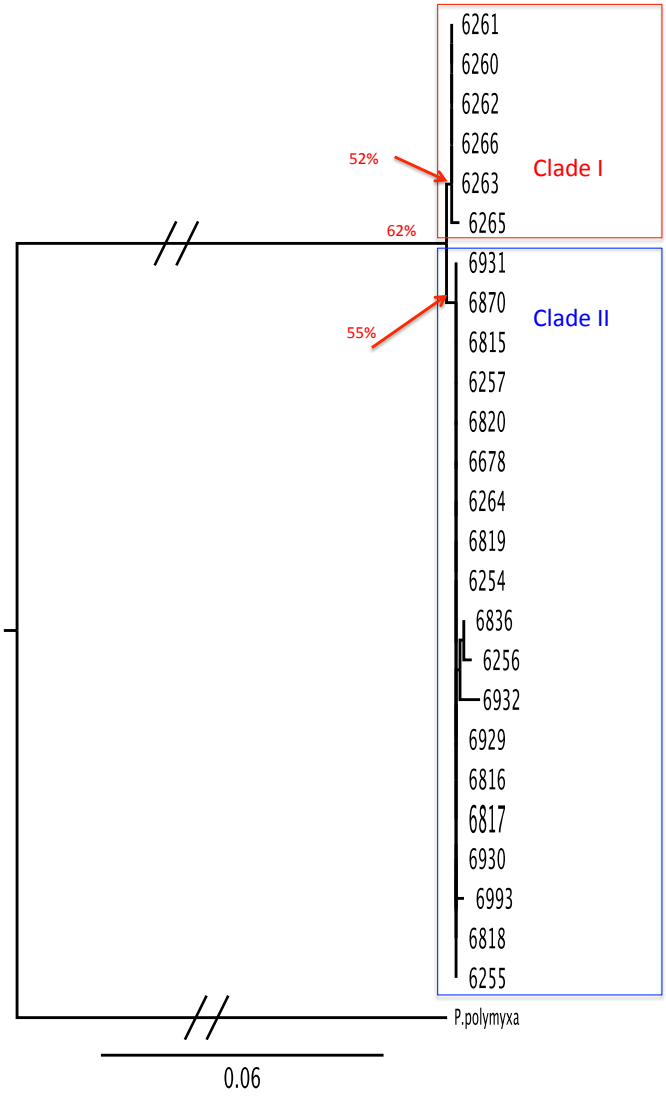


Figure 2.5: Neighbour-joining bootstrap analysis of the *PyrE* phylogenetic relationship between *P. larvae* isolates, tree shown rooted against type strain of *Paenibacillus polymyxa.* Clade I contains isolates formerly identified as subspecies *pulvifaciens*, Clade II contains isolates formerly identified as subspecies *larvae*. Bootstrap support over 70% shown in red. Branches dissected by parallel diagonal lines have been reduced to a third of the original length.

The bootstrap analysis produced a tree with two distinct clades (figure 2.5), as seen for *purH* this matched the previously defined subspecies, with subspecies *pulvifaciens* appearing in clade 1 and subspecies *larvae* in clade 2. Isolates 6993 and 6264 previously defined as subspecies *pulvifaciens*, appeared in the incorrect clade. Isolates 6260 and 6261 were again grouped incorrectly with the *pulvifaciens* isolates despite being defined as subspecies larvae.

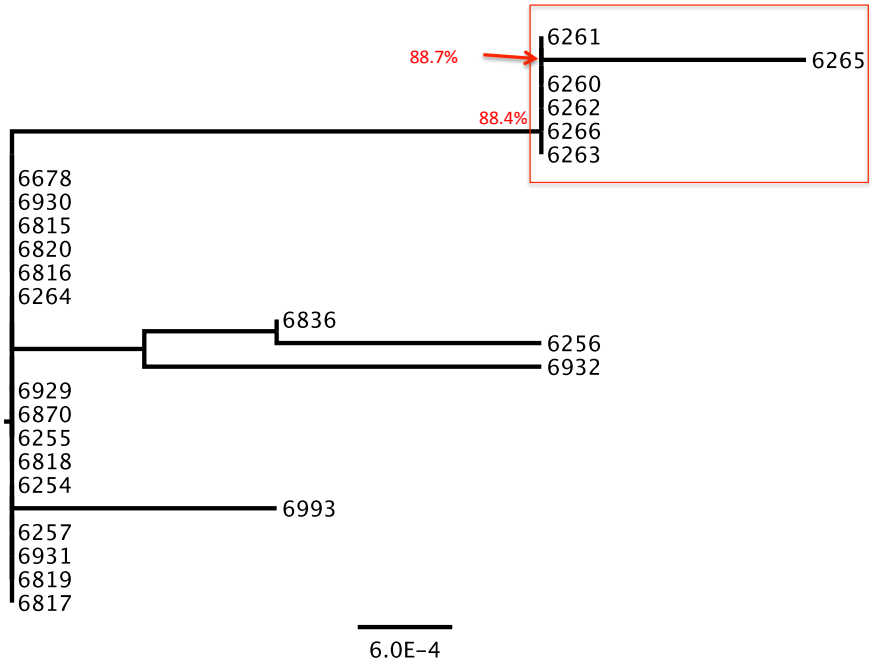


Figure 2.6: Neighbour-joining bootstrap analysis of the *PyrE* phylogenetic relationship between *P. larvae* isolates. The red box contains isolates formerly identified as subspecies *pulvifaciens*. Bootstrap support over 70% shown in red.

Despite this project successfully sequencing 25 isolates of *P. larvae* all containing this highly similar section of DNA from the purH gene. No significant matches could be found in the NCBI database with over 75% sequence identity, as a result the rooted tree produced has a large root with compresses the 25 sequenced isolates. The unrooted bootstrap tree (figure 2.6) allows for the observation of the phylogenetic differences existing between the 25 strains without the compression caused by rooting the tree.

2.6. Discussion

2.6.1. Selection of *P. larvae* isolates

In this chapter the utilisation of 16S rRNA subunit sequencing as a method of species identification and intra-species differentiation was assessed. To successfully represent the species a selection of reference isolates from the LMG culture collection were

added to the growing collection of UK environmental isolates held at FERA (table 2.3). This collection of isolates would ensure that work carried out was relevant within the UK, but also comparable with the wider research community, as many of the LMG isolates had previously been documented in several published articles (Forsgren et al. 2010; Heyndrickx et al. 1996; Loncaric et al. 2008; Tomkies et al. 2009).

Following their original isolation in 2007, the UK environmental isolates were identified using fatty acid profiling. This technique was common practice for cultured environmental isolates at FERA. This analysis was able to identify the isolates down to genus level, confirming that isolates were from the *Paenibacillus* genus. Unfortunately this method was unable to identify to species level. 16S rRNA sequencing confirmed that the environmental isolates (table 2.3) were *Paenibacillus larvae* isolates. It is recommended that the database be amended to reflect the identification to species level.

2.6.2. 16S identification of *P. larvae* isolates

The sequencing of the 16S subunit has been used for many years for the identification of bacterial isolates (Clarridge III 2004; Janda and Abbott 2007; Petti et al. 2005). Sequencing was more time and cost effective than the alternative plating and biochemical identification methods (Katznelson 1950; Nakamura 1984; White 1906). The use of morphological and biochemical data had formed the basis of earlier taxonomic groupings along with symptoms observed in the hive. The use of previous taxonomic groupings as an accurate means of determining the virulence of an isolate was shown to be unfounded (Ashiralieva and Genersch 2006).

The use of 16S sequencing as an identification tool was shown to be very effective, it was able to reliably differentiate *P. larvae* from other bacterial species within the *Firmicute* phylum (figure 2.1). 16S sequencing was shown to be able to correctly distinguish *P. larvae* isolates from the *Clostridium* genus due to a >17.7% level of sequence divergence. Of more importance was the ability of 16S sequencing to distinguish *P. larvae* from other species within the genus and to successfully

distinguish *P. larvae* from other bacterial genera and species commonly from in hive material. Species from the *Bacillus* genus are commonly isolated from hive material and environmental samples; here it was shown that on average 12.8% sequence divergence existed between *P. larvae* isolates and strains from the *Bacillus* genus. It was shown that an average 10.8% sequence divergence existed between *P. larvae* and available 16S rRNA sequences from the *Paenibacillus* genus. Crucially it was able to distinguish *P. larvae* from *P. alvei*, a bacterium often considered to cause secondary infections within hives infected with AFB or EFB (Hornitzky and Anderson 2003). 16S rRNA sequencing highlighted a 9.5% sequence divergence, and was capable of distinguishing the species. The use of the 16S rRNA subunit was shown to be capable of identification to species level, an improvement on the FAP methodology commonly used at FERA. It is recommended that identification of suspected *P. larvae* isolates be performed using 16S rRNA analysis.

Whilst possible to view the alignment as a phylogenetic tree, this merely serves as confirmation of the levels of identity existing between the 16S sequences. This does show, as expected, a closer relation between *P. larvae* and other species within the *Paenibacillus* genus. Unexpected was the presence of several members of the *Paenibacillus* genus which showed a closer relationship to *P. larvae* than *P. alvei*. This was unexpected due to *P. alvei* being commonly isolated from hive material. This was previously observed (Shida et al. 1997a; Shida et al. 1997b), however no explanation is given and it remains unclear why the bacterial species found within the same environment have a more diverse 16S rRNA subunit than species from more diverse environmental backgrounds.

2.6.3. Intra-species differentiation based on 16S RNA sequencing

It has previously been shown that the 16S rRNA subunit is the sequence of choice by molecular microbiologist due high levels of similarity and no change in function over time (Janda and Abbott 2007). Differences exhibited by *P. larvae* isolates were <3%, the value often associated with isolates belonging to the same species.

Phylogenetic analysis of the *P. larvae* 16S alignment rooted against *Bacillus cereus* highlighted the high level of identity between sequences. The clades produced by this neighbour-joining bootstrapping method were able to separate isolates previously identified as different subspecies (figure 2.2). With the exception of three isolates 6255, 6260 and 6261, these isolates previously identified as subspecies *larvae* were shown here to group with the subspecies *pulvifaciens* isolates. This grouping shown in the phylogeny produced suggested the potential for 16S sequencing to group the isolates, however a larger selection of subspecies *pulvifaciens* isolates would be required to confirm this. The anomalous isolates may have been miss-identified at point of original isolation, however without a more detailed history of these isolates this would be difficult to explore, many of the characteristics used to differentiate the subspecies produced variable results, therefore accurate identification to the subspecies level was considered to be difficult (de Graaf et al. 2006b).

2.6.4. Evaluating the potential use of a Multi Locus Sequence Typing method for the intra-species differentiation of *P. larvae*

A successful Multi Locus Sequence Typing scheme required the sequencing, concatenation and alignment of seven highly conserved core genes. Whilst these genes are highly conserved, they were predicted to be less conserved than the 16S rRNA subunit. Therefore the study of these genes has a greater potential to differentiate between strains. Initially five genes were chosen; the primers designed for three of these genes failed to amplify the target sequence or were not specific to the target sequence. It was decided to continue the scheme using purH and PyrE in the first instance, assessing the potential of an MLST scheme at intra-species differentiation.

The aligned sequences for the purH and PyrE genes highlighted a difference 3.25% and 0.99% respectively. As expected purH showed a greater level of sequence variation than displayed in the 16S rRNA sequences, however the level of sequence variation was far lower than expected. The alignment produced for the PyrE

sequences showed that the level of sequence identity was similar to that of the 16S sequences and in fact displayed 0.17% more sequence similarity. DNA hybridization has shown that isolates sharing 97% 16S rRNA sequence similarity may share <60% genome identity (Stackebrandt and Goebel 1994). It was predicted that as previously seen in *Bacillus cereus*, *anthracis* and *thuringiensis*, a greater level of sequence divergence would be present in these genes (Helgason et al. 2004; Marston et al. 2006; Sorokin et al. 2006). The sequence similarity calculated suggests that these genes are more suited to MLST schemes in the *Bacillus* genus for which they were designed, and are less applicable to the *Paenibacillus* genus.

Production of a bootstrap tree (figures 2.3 and 2.4) showed that purH and PyrE were, as seen in the 16 rRNA phylogeny, able to differentiate based upon the previous subspecies differentiation. Anomalies were still present though, unlike the 16S phylogeny, strain 6993 was grouped with the wrong subspecies in both of the gene phylogenies and strain 6264 was miss-grouped using the PyrE alignment. As seen in the 16S phylogenies strains 6260 and 6261 were grouped with the subspecies *pulvifaciens* isolates, strain 6255 was correctly grouped using these gene alignments. If the previously used subspecies differentiation were to be used, the data presented here would suggest that strains 6260 and 6261 were miss-identified at the point of original isolation. Whilst this method had correctly grouped several isolates (based upon the previous taxonomic grouping), it was unknown what relevance this grouping had on the taxonomy or phenotypic characteristics of the species.

Production of un-rooted bootstrap trees for both purH and pyrE showed the high level of similarity existing between sequenced isolates. No suitable root was found in the NCBI database, however this sequencing project was able to produce 25 sequences that exhibited very low variation. It is worth noting that the closest NCBI BLAST hits were that of the corresponding gene. It is worth noting that *Paenibacillus* sequences are very limited and often only exist singularly per species, further sequencing of a larger group of isolates would be able to identify the true relationship based upon these genes.

2.6.5. Can potential genotypes be identified?

Phylogenies produced from 3 separate coding regions of *P. larvae*, were largely in agreement that 2 groups of isolates could be formed. The separate grouping of isolates identified as subspecies *larvae* and *pulvifaciens* has previously been noted (Shida et al. 1997a; Shida et al. 1997b), however previously this has only been shown in 16S rRNA. Here this relationship is also shown to existing outside of the ribosomal subunit and in housekeeping genes. Based upon this data the previous subspecies differentiation does appear to reflect the genetic differences. Targeted sequencing approaches are likely to result in the same outcome, with the overall grouping reflecting the previous subspecies differentiation. As this grouping has been shown not to accurately reflect the virulence of *P. larvae*, it is suggested that a different approach is used.

2.7. Conclusion

The use of 16S rRNA sequencing as a method of identification was shown to be accurate. This method was fast and robust at correctly identifying isolates recovered from hive material. Alignments showed that a significant difference existed between the *P. larvae* 16S sequences and 16S sequences from bacteria within the genera or from other *Bacillus* species.

The utilisation of these methods for intra-species differentiation was never realised. Whilst the sequencing of the 16S subunit and purH and PyrE genes had initially given promising results, it was decided that whilst these methods held the potential for the grouping of isolates. It was shown that the use of 16S rRNA sequencing or purH and PyrE sequencing reflected the previous subspecies differentiation of the organism. However the true relevance of this taxonomic grouping in relation to the organism was unknown, as isolates belonging to both subspecies were shown to be pathogenic (Genersch et al. 2006).

Three separate sequence based methodologies showed evidence that strains 6260 and 6261 shared a greater level of sequence similarity with isolates belonging to subspecies *pulvifaciens*. If the subspecies differentiation was still considered accurate, this data would have led to further investigation potentially leading to a revised description of the isolates. As the relevance of the subspecies differentiation remains unknown at present time, the grouping of these isolates was not further explored. With an improved knowledge of intra-species differences, this could be revisited at a later date.

After the sequencing of purH and PyrE genes, the revised taxonomic grouping of *P. larvae* became widely accepted; isolates were no longer referred to with a subspecies differentiation. It was decided to halt the MLST scheme at this point, whilst the method had shown promising results and the addition of subsequent genes was likely to improve the level of differentiation observed. Since the subspecies differentiation was dropped (Genersch et al., 2006), the true relevance of this differentiation is unknown. Whilst it has been shown that a selection of subspecies *pulvifaciens* isolates were able to cause infection, it has not been shown that all isolates of subspecies *pulvifaciens* possess this ability. Nor has it been proven that all isolates previously identified as subspecies *larvae*, are capable of causing infection.

This suggested that a greater understanding of the virulence of isolates was required before the true relevance of the previous subspecies groupings, reflected in these methods, would be fully understood. Continuation of this methodology no longer met the overall aims of the project and it was decided to adopt different approaches.

2.8. Acknowledgements

I thank Pauline Guiglini for the isolation of UK environmental isolates prior to the commencement of my PhD and I thank Dr. Simon Weller for his supervision during the first 2 months of this work.

Chapter 3

Intra-species differentiation by means of PCR based fingerprinting, and its impact on UK detection

3.1. Summary

- Isolates were identified as belonging to the 4 previously identified *P. larvae* ERIC types.
- ERIC type I isolates are monophyletic on a 16S rRNA phylogeny, despite no known link between ERIC repeats and the ribosomal subunit.
- Due to a small sample size, it is not possible to accurately identify the relationship between ERIC type II-IV and the 16S rRNA phylogeny.
- The AFB LFD kit was shown to only detect AFB outbreaks belonging to ERIC type I isolates. The epidemiology or occurrence of ERIC type II-IV infections is not widely understood. So the impact of this on detection of AFB is not fully known.

3.2. Introduction

Discrimination of *Paenibacillus larvae* isolates used biochemical and morphological identification methods for around 80 years (Katznelson 1950; Nakamura 1984; White 1906). Use of PCR and genetic identification methods began to be used for isolates of *P. larvae* in 1993 (Ash et al.). *P. larvae* was subject to many taxonomic revisions, the most important revision was in 2006 when it was shown that isolates believed to be non pathogenic, possessed the ability to cause AFB (Genersch et al. 2006). This revision was widely accepted

despite the presence of many phenotypically diverse isolates (Heyndrickx et al. 1996).

In chapter 2, the use of 16S rRNA and housekeeping gene sequencing was explored as a means of intra-species differentiation. It was shown that these methods differentiated isolates based upon previous subspecies differentiations. Since the realization that this differentiation did not reflect virulence (Genersch et al. 2006), the groupings produced further highlight the genetic difference that exists between the subspecies. Despite the adoption of the revised taxonomy of *P. larvae* by the research community by late 2007, it remained unlikely that the true intra-species relationships had been fully explored. Previously isolates were considered to possess a significant degree of phenotypic diversity, warranting the taxonomic positioning within two separate subspecies (Heyndrickx et al. 1996). Unfortunately very few examples of *Paenibacillus pulvifaciens* were isolated, so only a small data set exists. It remains important to fully understand why, at time of original isolation, these isolates were considered non-pathogenic. A better understanding of the phenotypes and genotypes present within *P. larvae* will lead to a better understanding of the infection pathway. Further sequencing of housekeeping genes was likely to display the same phylogenetic groupings. PCR based fingerprinting techniques had been used successfully in the research of *P. larvae*, phenotypic and genotypic differences had been shown to exist between isolates grouped using these methodologies (Alippi and Aguilar 1998; Genersch and Otten 2003).

3.2.1. ERIC Fingerprinting

Enterobacterial Repetitive Intergenic Consensus (ERIC) sequences, were originally identified in *Escherichia coli, Salmonella typhimurium*, *Yersinia pseudotubercuiosis, Kiebsiella pneumoniae and Vibrio cholera* (Hulton et al. 1991). Although initially discovered in members of the *Enterobacteriacae*, these repetitive elements found throughout the genome were shown to be well conserved amongst Bacteria (Versalovic et al. 1991). The use of these repetitive elements as a meaning of PCR fingerprinting was soon exploited. Whilst initially

used on *Enterobacteriacae,* this method has also been applied to more diverse species of bacteria (De Arruda et al. 2003; Khan et al. 2002). ERIC PCR fingerprinting has also been applied to the identification of yeasts (Hierro et al. 2004).

3.2.2. Why is ERIC fingerprinting more discriminate than direct sequencing of genes?

ERIC sequences are composed of a 126bp unit found to be well conserved within the *Enterobacteriacae.* These sequences were shown to exist in flanking regions within ~50bp of a transcribed gene or within intergenic regions of mRNA operons (Hulton et al. 1991). Up to 21 ERIC repeats were found in Enterobacterial species, whilst well conserved, the positioning of these elements on the chromosome varied between species (Hulton et al. 1991). Unlike standard PCR amplification, primers are designed to extend from this region; amplifying a series of bands of varied length representing the genome sequence between these ERIC repeats (De Bruijn 1992). The number of repeats within a bacterial genome combined with the PCR cycle used determines the number and length of bands produced. Unlike the use of direct sequencing in chapter 2 targeted to a specific ribosomal subunit or housekeeping gene, ERIC fingerprinting is targeted to repetitive elements existing at a variety of distances within the genome. The frequency and location of these elements within the genome is varied between species (Hulton et al. 1991), allowing discrimination to this level. Whilst direct sequencing relies on genetic mutations occurring in highly conserved genes, ERIC PCR relies on the conservation of a section of non-coding DNA. Due to its non-coding nature, variation is more likely to exist within closely related species. It has been suggested that primers designed to target these ERIC sequences in *Enterobacteriacae* do not necessarily target the same sequences within other bacterial species (Gillings and Holley 1997), however the target sequence is not important providing replicable fingerprints can be achieved.

3.2.3. Why is ERIC typing of these isolates important?

ERIC primers are able to amplify several bands from strains of *P. larvae*; upon analysis of the patterns produced it is possible to group the strains in to 4 distinct ERIC types. It was thought that ERIC types I and II were pathogenic and that ERIC types III and IV were non-pathogenic (Genersch et al. 2006). Whilst the subspecies differentiation was not based upon ERIC typing, It was shown that isolates belonging to ERIC type I were exclusively subspecies *larvae* and isolates belonging to ERIC III and IV were exclusively subspecies *pulvifaciens*, ERIC type II isolates were identified as predominantly belonging to subspecies *larvae* (Genersch et al. 2006). This has been one of the most widely used research methods in the field of AFB research and has become almost the gold standard method for comparing strains from around the world. The epidemiology of AFB infections was better realised upon ERIC typing of isolates, it was shown that ERIC type I infections were predominant across the world, with <15 type II outbreaks in Europe. Type III or IV outbreaks of AFB have never been knowingly isolated from AFB infected hives; all isolates in existence were isolated as suspected 'Powdery Scale' infections (Katznelson 1950).

Whilst based upon the length of genomic sequence between the repetitive units, phenotypic differences have been shown to exist between ERIC groups. Differences occur between isolates labeled as type I and II and types III and IV, with the most predominant difference being the ability to cause disease. ERIC types I and II possessing the ability to cause a systemic hive infection, compared to ERIC types III and IV isolates only able to cause acute infections (Genersch 2009).

Only clinically diseased hives are routinely studied for the presence of AFB, it is currently unknown what percentages of colonies contain sub-clinical AFB infections. In a small Argentinian study, 35% of colonies not displaying clinical AFB symptoms, contained AFB spores in extracted honey (Iurlina and Fritz 2005). Obtaining a true representation of the epidemiology of the four ERIC

types is unlikely, as in sub-clinical infections *P. larvae* is likely to be outcompeted by other bacterial species.

3.2.4. Other fingerprinting methodologies

Other PCR based fingerprinting technologies have been developed, primarily used amongst bacterial isolates are Repetitive Extragenic Palindromic (REP) sequences (Stern et al. 1984) and BOX (Martin et al. 1992). Comparative analysis of fingerprinting technologies has not determined a preferential method (Hermans et al. 1995), these methods continue to co-exist and are utilised by various research groups.

3.2.5. Can these methodologies be applied to AFB research?

Combining ERIC typing with other fingerprinting methods such as BOX or REP, adds another level of grouping to the already grouped isolates. This method has been able to identify a set of ERIC/BOX isolates that were at the time unique to that region of South America (Alippi and Aguilar 1998). Whilst this method proved useful in the bio-geographical grouping of the isolates, it was not related to the organisms' phenotype. Many isolates in this bio-geographical study were not fully characterised, phenotypic differences between isolates were not explored, Whilst it has been shown that a combination of fingerprinting methods can add a greater resolution to the discrimination of isolates, the true relevance of this discrimination remains unknown.

3.2.6. Lateral flow immunological assays

Lateral Flow Devices (LFDs) were originally developed for the on-site detection of plant pathogens (Danks and Barker 2000). Later this technology was adapted

to suit a wider range of pathogens, importantly was the development of a LFD test for the on-site confirmation of AFB infections (Vita Europe Ltd.). Use of these test kits allow quick preliminary confirmation of the visual diagnosis made. Antibodies were developed for the LFD whilst the causative agent was considered *Paenibacillus larvae* subspecies *larvae*. ERIC typing was used to identify isolates belonging to ERIC types I and IV, a target unique to ERIC I was used in the development of the AFB antibody (Un-published data, FERA). As a result it is believed that LFD kits will be unable to detect ERIC type IV outbreaks, the ability of the LFD kits to detect ERIC types II and III remains unknown.

3.3. Aims

- To ensure the sample set of *P. larvae* isolates used within this thesis represent all 4 previously identified ERIC types
- To assess any congruence existing between ERIC types and 16S rRNA or MLST phylogenies
- To determine if LFD kits are capable of detecting AFB outbreaks belonging to all 4 ERIC types

3.4. Materials and Methods

3.4.1 Bacterial Isolates

27 previous isolated *Paenibacillus larvae* strains were used (table 3.3), these comprised of 15 LMG culture collection reference isolates and a further 12 UK environmental strains isolated in 2007. For full details of isolates used see Chapter 2.

3.4.2. Culture Methods and DNA extraction

Bacterial isolates were cultured (Appendix 1) from frozen storage vials (Appendix 2). DNA extractions were performed (Appendix 3) on all of these isolates and they were appropriately stored (appendix 4).

3.4.3. PCR

The fingerprinting technique used within this chapter used the PCR reagents and methodologies that are detailed in Appendix 5. Any differences to afore mentioned protocols are detailed below.

Specific ERIC PCR conditions

ERIC primers (ERIC1R: 5′-ATG TAA GCT CCT GGG GAT TCA C-3′ and ERIC2: 5′-AAG TAA GTG ACT GGG GTG AGC G-3′) (De Bruijn 1992) were used to produce the *P. larvae* fingerprints. Standard PCR reagents were used (Appendix 5). $MgCl_2$ concentration was trialed and the optimum was observed to be a final concentration of 5 mM per reaction. Primers were used at a final concentration of 2µM per reaction. 0.75U of Taq was used per reaction.

Reactions were amplified using an initial 2 minute denaturing step at 95°C. Followed by 35 cycles of 1 minute denaturing at 95°C, 1 minute annealing at 53°C and 2.5 minutes extension at 72°C. This was followed by a 10-minute final extension at 72°C. PCR products were visualised on a 1.5% agarose gel stained with ethidium bromide.

3.4.4. ERIC Typing of isolates

Within the strains held at FERA were representative LMG reference strains, or isolates from other research groups that had previously been ERIC typed (Table 3.3). By comparing the banding patterns produced by previously typed strains, it

was possible to assign isolates to ERIC types I, III and IV. No type II isolates held had been previously typed, only 4 band patterns were produced so isolated not belonging to types I, III and IV were assigned to ERIC type II.

3.4.5. Use of LFDs to identify *P. larvae*

Lateral Flow Devices (LFDs) are small rapid detection kits available for detecting a variety of infectious organisms. The sample to be tested is placed in the sample jar provided; this contains a buffer and a series of metal ball bearings. The jar is shaken to lyse the sample, 1-3 drops of this suspension are placed on to the LFD. Contained within the plastic housing is a membrane, this membrane draws in the liquid by capillary action. Two lines of the membrane are impregnated with antibodies. The first is a control line confirming that this test kit is functional; the test line contains target specific antibodies. A positive test is indicated by the presence of both control and test lines (table 3.1).

Control Line	Test Line	Outcome
-	-	Faulty kit
-	+	Potential fault (Repeat)
+	-	Negative Result
+	+	Positive Result

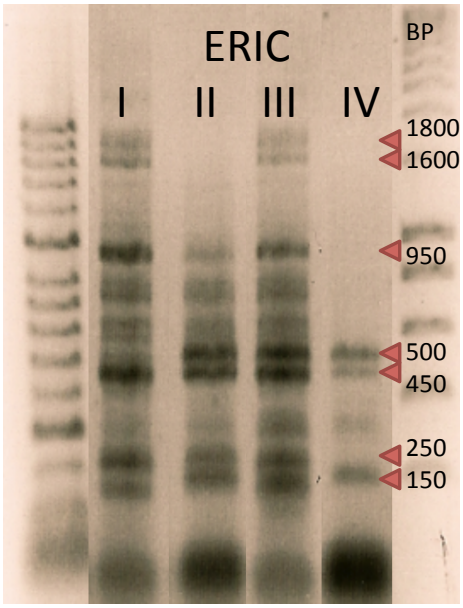
Table 3.1: Possible results obtained from an LFD test kit and the meaning of the result obtained

3.4.6. Phylogenetic analysis

Phylogenetic analysis was performed as described in chapter 2.4. 16S sequences were aligned using ClustalW alignment algorithm and bootstrap neighbour-joining phylogenetic trees were created to visualise the relationship between isolates.

3.5. Results

3.5.1. ERIC typing of *P. larvae* isolates



ERIC ID	Bands Present (bp)
I	150, **250**, 450, **950**, **1600**, **1800**
II	150, **250**, 450, **500**
III	150, **250**, 450, **500**, **950**, **1600**, **1800**
IV	150, 450, **500**

Figure 3.1 / Table 3.2: The four banding patterns produced by P. larvae isolates. Lanes 2-5 contain ERIC types I-IV respectively. Bands produced are listed in table 3.2. Key bands used for typing are shown in bold print.

ERIC profiling of the isolates produced 4 distinct band patterns. The availability of isolates belonging to pre-determined ERIC groups provided a key (figure 3.1) by which further samples could be typed. The band patterns produced displayed high levels of similarity, typing of the isolates was performed by the presence or absence of bands at 250, 500, 950, 1600 and 1800bp (table 3.2).

25 isolates were successfully ERIC typed (table 3.3), of which 14 were LMG reference isolates and the remaining 11 were UK environmental isolates. 18 of the isolates were identified as ERIC type I representing 68% of the sample subset. ERIC types III and IV were less well represented each with 3 isolates, 12% of the sample subset. Only 2 isolates were identified as ERIC type II 8% of the sample subset. Isolates identified as subspecies larvae were shown to belong to ERIC type I, with the exceptions of strain 6260 (ERIC II) and strain 6261

(ERIC III). All other isolates identified as belonging to ERIC types II-IV were previously identified as subspecies pulvifaciens.

Identification	FERA Ref.	NCPPB	ERIC	Previous ERIC type*
Paenibacillus larvae ssp. Larvae	6254	LMG9820	I	I
Paenibacillus larvae ssp. Larvae	6255	LMG14425	I	N/A
Paenibacillus larvae ssp. Larvae	6256	LMG14426	I	N/A
Paenibacillus larvae ssp. Larvae	6257	LMG16147	I	N/A
Paenibacillus larvae ssp. Larvae	6259	LMG16241	I	N/A
Paenibacillus larvae ssp. Larvae	6260	LMG16250	II	N/A
Paenibacillus larvae ssp. Larvae	6261	LMG16251	III	N/A
Paenibacillus larvae ssp. Pulvifaciens	6263	LMG15974	III	N/A
Paenibacillus larvae ssp. Pulvifaciens	6264	LMG16247	IV	IV
Paenibacillus larvae ssp. Pulvifaciens	6265	LMG16249	II	N/A
Paenibacillus larvae ssp. Pulvifaciens	6993	LMG14427	IV	IV**
Paenibacillus larvae ssp. Pulvifaciens	6266	LMG16252	III	III
Paenibacillus larvae ssp. Pulvifaciens	7029	LMG14427	IV	IV**
Paenibacillus larvae ssp. Larvae	6678	LMG16241	I	N/A
Paenibacillus sp.	6815	N/A	I	N/A
Paenibacillus sp.	6816	N/A	I	N/A
Paenibacillus sp.	6817	N/A	I	N/A
Paenibacillus sp.	6818	N/A	I	N/A
Paenibacillus sp.	6819	N/A	I	N/A
Paenibacillus sp.	6820	N/A	I	N/A
Paenibacillus sp.	6836	N/A	I	N/A
Paenibacillus sp.	6870	N/A	I	N/A
Paenibacillus sp.	6930	N/A	I	N/A
Paenibacillus sp.	6931	N/A	I	N/A
Paenibacillus sp.	6932	N/A	I	N/A

Table 3.3: Strains held in the collection at FERA are listed showing LMG reference status where appropriate. Strains were typed to 4 ERIC groups, also shown are the previous ERIC types assigned by other research groups. *(Genersch et al. 2006) **(Watkins et al. 2003)

3.5.2. Comparison between assigned ERIC types and a 16S rRNA phylogeny

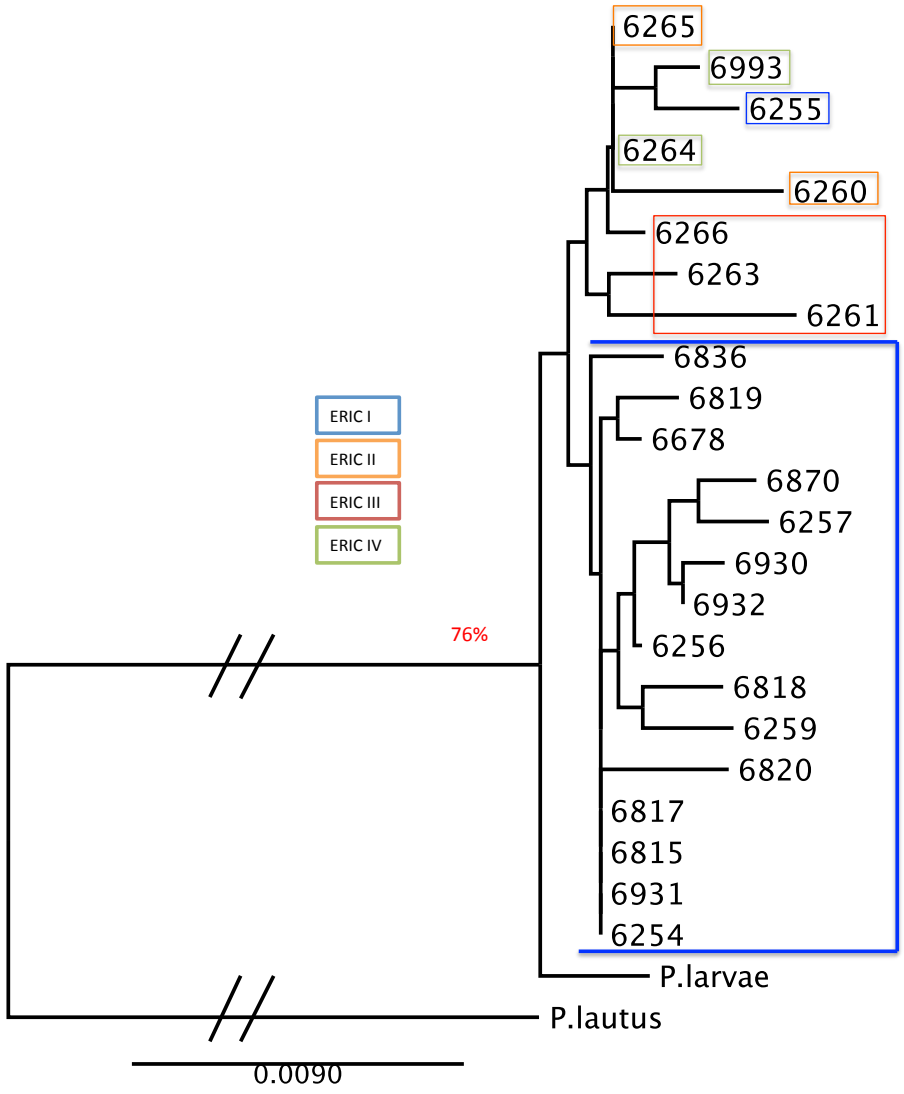


Figure 3.2: Bootstrap Neighbour joining analysis of the phylogenetic relationship between *P. larvae* isolates. Coloured boxes highlight the ERIC groups assigned to the isolates, showing some congruence with the 16S phylogeny produced. Branches dissected by parallel diagonal lines have been reduced to a third of the original length.

A degree of congruence was observed between the 16S phylogeny produced and the assigned ERIC types. Whilst isolates representing types II - IV showed no phylogenetic separation within the 16S phylogeny. ERIC type I isolates were

shown to group together within the phylogeny. Only isolate 6255 belonging to ERIC type I was shown to group separately from the other ERIC type I isolates.

3.5.3. Assessing the ability of LFDs to detect all ERIC types of *P. larvae*

Strain Number	Details	ERIC Type	LFD Result
6254	Type Strain	I	+
6678	Type I Replicate	I	+
6260	PII strain	II	-
6265	PIp Strain	II	-
6261	PII Strain	III	-
6266	PIp Strain	III	-
6993	Bio-control strain	IV	-
6264	Type IV Replicate	IV	-

Table 3.4: Cultured *P. larvae* isolates representing all 4 ERIC types tested by LFD to assess it's capability of detecting AFB outbreaks belonging to all ERIC types.

It was observed that isolates belonging to ERIC type I were able to produce a positive LFD test result (table 3.3). Isolates belonging to ERIC types II - IV were unable to produce a positive test result. This showed that the LFD kits were unable to detect all isolates of *P. larvae* following the revised taxonomy, and further suggests that they are unable to detect isolates from type II that had been previously identified as subspecies *larvae*.

3.6. Discussion

Genetic fingerprinting has been used in the identification and taxonomy of bacteria for many years. The field of AFB research has used many different fingerprinting techniques to achieve the differentiation of isolates (Alippi and Aguilar 1998; Genersch et al. 2006). Whilst fingerprinting was never directly used to achieve the subspecies differentiation of *P. larvae* it was thought to reflect this and the pathogenicity of the organism (3.2.3). It was only in 2006 that this was shown to be incorrect (Genersch et al. 2006).

3.6.1. ERIC typing *P. larvae* isolates

A potential weakness of PCR based fingerprinting methodologies is the reproduction of consistent banding patterns between laboratories. The collection of reference isolates held several strains which had previously been identified to ERIC types I, III and IV (Table 3.3). The band patterns produced by these strains were used as a key to identify the ERIC type of all other strains. No reference isolate was held belonging to ERIC type II, so it was not possible to identify these isolates in this way. Comparison to published ERIC type II patterns allowed the identification of the fourth banding pattern, confirming it to be ERIC type II. It is believed that this method has correctly identified all four ERIC types of the *P. larvae* isolates.

As AFB infections worldwide have been shown to be caused by isolates belonging to ERIC I (Antunez et al. 2009; Genersch et al. 2006), ERIC type I was widely considered to be the most pathogenic subset. Type II were shown to have varied levels of pathogenicity, and types III and IV were believed to be non-pathogenic (Genersch et al. 2006). In this study, it was confirmed by ERIC typing that the majority of isolates belonged to ERIC I, further suggesting that within the UK, and possibly worldwide, that ERIC I isolates are the most prevalent. All environmental samples were expected to be ERIC type I isolates, as LFDs were designed to only detect isolates belonging to this group (3.2.6).

Due to the selection of ERIC type I isolates by UK detection methods, it is not possible to accurately comment on the prevalence of ERIC type II-IV isolates. Colonies colonised with type II-IV isolates are not detected by LFD (table 3.4), identification of infected colonies, would rely on the presence and/or detection of symptoms. UK bee inspectors were asked to alert us to the presence of colonies displaying AFB symptoms and producing a negative LFD test. During the 3 years of this project, no colonies presenting AFB symptoms but producing a negative LFD test were identified, suggesting that ERIC II-IV infections remain subclinical or symptomless negating the possibility of detection by inspectors. This could also indicate that all AFB isolates from the UK belong to ERIC type

I, however a large proportion of colonies would have to be sampled to prove or disprove this.

The ERIC typing of isolates showed that strains previously identified as subspecies *larvae* predominantly belonged to ERIC type I, and ERIC type II on rare occasions (Genersch et al. 2006). Isolates 6260 and 6261 were previously identified as subspecies *larvae* at point of isolation (Chapter 2.4.1). These isolates were identified as ERIC types II and III respectively. ERIC type II commonly contains subspecies *larvae* isolates, however exceptions have been noted. ERIC III is believed to contain only subspecies *pulvifaciens* isolates (Genersch et al. 2006). Phylogenetic analysis of these isolates also highlighted a closer genetic relation to samples identified as subspecies *pulvifaciens* (Chapter 2.5). The observation of these isolates belonging to these ERIC types further suggests that these isolates were previously incorrectly identified to subspecies level.

The assay was developed to detect changes in inter-genic repetitive sequences within the *enterobacteriacae*, it remains unknown what the target sequence is within *P. larvae*. It was shown that amplification favoured gram-negative bacteria, but ERIC-like patterns could be formed from a diverse range of bacteria (Versalovic et al. 1991). Upon increasing the annealing temperature, and therefore primer specificity, no bands are amplified from non-enterobacterical species suggesting primers are binding to sites with partial identity (Gillings and Holley 1997). ERIC typing successfully groups isolates based on inter-genic sequences into 4 ERIC types. The relationship between this genetic difference and the disease phenotype is also unknown. Bioinformatic analysis of the genome to highlight the location of the target sequences was not possible due to the incomplete nature of the genome.

Whilst the ERIC typing of isolates allowed for a broader understanding of where these isolates fit within the global community of isolates, at this time it was not possible to use the groupings achieved as a method of differentiating isolates. Even though 4 ERIC groups were identified, it is not possible to predict the phenotype of isolates based on these groupings. A variety of other fingerprinting

methods have previously been successful at grouping isolates, by combining several fingerprinting methods it has been possible to highlight differences that exist within a particular fingerprinting type. So far the use of this methodology has only highlighted bio-geographical differences between isolates of *P. larvae* (Alippi and Aguilar 1998). The use of other methods such as MBO REP and BOX PCR would likely highlight further differences between isolates, however without prior knowledge of how the species should be grouped it would remain unknown what these intra-species groupings showed.

3.6.2. Congruence displayed between ERIC types and 16S rRNA phylogeny

The correlation between the 16S phylogeny and the observed ERIC types, has not been previously explored. The 16S rRNA subunit is the focus of phylogenetic studies of bacteria due to its highly conserved nature. ERIC typing is based upon amplifying sections of bacterial genome between ERIC repeated sequences. Within *Enterobacterial* species these 126bp repeats are located on polycistronic operons in intergenic regions or flanking a transcribed gene (Hulton et al. 1991). It is believed that lower annealing temperatures allow for the priming of areas with less similarity (Gillings and Holley 1997). This suggests that ERIC repeats most likely do not exist as discussed by Hulton (1991) in non-*Enterobacterial* species. It is unknown if a link exists between ERIC-like repeats and 16S rRNA, however a connection is unlikely. It would seem more likely that genetically ERIC types II-IV are more similar to each other than to ERIC type I. Comparison to the phylogenies produced by purH and PyrE genes, also show the same grouping (Chapter 2.5) which adds further conviction to this theory. Isolate 6255 is an anomaly, whilst this was identified as ERIC type I it did not group with the other ERIC type I isolates.

The availability of more isolates belonging to ERIC types II-IV would allow a more systematic analysis of the relationship between the 16S rRNA subunit and the fingerprinting data. The low sample numbers may explain the low level of congruence between these isolates in the phylogeny. The tree building method is

highlighting the closest match within the alignment; with a lack of isolates belonging to these groups it is possible that the closest genetic match is in fact an anomaly. It is highly unlikely that further isolates belonging to ERIC types II-IV will be isolated from hive material. UK detection methods were designed to identify ERIC type I isolates, so any sampling based upon LFD detection will fail to isolates strains belonging to ERIC types II-IV. It is unknown what symptoms ERIC types II-IV infections produce; it is possible that these infections exist sub-clinically within colonies. To assess the epidemiology of ERIC type I-IV strains it would be necessary to sample colonies randomly, regardless of identification of symptoms, from a high percentage of the total population. However prior to the commencement of such a study, the presence of ERIC types II-IV in UK hives would need to be confirmed. It remains unlikely that the true significance of the link between the 16S phylogeny and the ERIC types will be fully realised.

3.6.3. Assessing the suitability of LFD kits for the detection of AFB outbreaks

The LFD results show that the primary method used in the UK for confirming AFB infection method only gave positive result for isolates belonging to ERIC type I. It failed across all isolates tested to detect isolates identified as belonging to ERIC types II-IV. Due to the previous taxonomic groupings and the prevalence of ERIC type I isolates, the antibodies present in the LFD were raised against a target specific to ERIC type I isolates. At the time of design, ERIC I isolates were believed to represent the virulent phenotype of the species.

The true impact of this assay only identifying isolates belonging to ERIC I, is not known. Whilst it is known that this assay does not detect all isolates, it is known that the most prevalent ERIC type is I (table 3.3). It is believed that isolates belonging to ERIC types II-IV exist as subclinical or symptomless infections within colonies. The overall impact of these infections is unknown. Three years of careful monitoring of colonies for AFB or suspicious symptoms coupled with a negative LFD test showed that inspectors did not face unexplained diseased

colonies. So even upon the exclusion of LFD results, no colonies were identified as potential ERIC II-IV AFB infections, so no environmental ERIC type II-IV isolates where isolated. It is still unknown how prevalent ERIC type II-IV strains are within UK colonies, or within colonies worldwide.

In Sweden two colonies displaying AFB symptoms were shown to be LFD negative, these were later identified as belonging to ERIC type II. This was the only reported outbreak of AFB, where the LFD was not able to confirm the infection (pers comm. Dr Richard Thwaites, FERA).

Without a much larger random sampling plan, with a LFD independent isolation technique, it is not possible to comment on the numbers of colonies with ERIC type II-IV infections. ERIC type I infections have been shown to cause larval mortality slower than ERIC type II-IV infections (Genersch et al. 2006). This slower rate of mortality allows for infected cells to be capped over by worker bees, this allows the infection to fully progress until only AFB scales remain. It has been suggested that ERIC type II-IV infections are capable of killing larvae prior to cell capping, this results in the removal of the dead larvae before AFB scales are formed limiting the spread of the disease. Due to this it is predicted that ERIC I infections are more likely to infect an entire colony opposed to ERIC II-IV infections that are more likely to cause acute infections limited to isolated cells. Overall the findings of inspectors within the UK suggest that colonies displaying AFB symptoms are infected with ERIC type I isolates. It is unknown if asymptomatic colonies exist within the UK, and unknown what ERIC type would be responsible. Whilst acute ERIC II-IV infections may occur in hives, the affect of these infections on the overall health of the colony is unknown. Without the occurrence of ERIC II-IV outbreak during the past 3 years, it is unlikely that re-designing the LFD antibodies would result in the detection of any further AFB infections.

3.7. Conclusion

The use of ERIC fingerprinting has been widely documented within AFB, as previously noted strains were grouped in to 4 ERIC types (Genersch et al. 2006). As expected UK outbreaks of AFB were found to belong to ERIC type I. Outbreaks around the world have also been shown to predominantly belong to ERIC type I, however other ERIC types have been isolated from infected colonies (Antunez et al. 2009; Genersch et al. 2006). Once more the lack of detailed information surrounding the isolates combined with the revision of the phenotypic traits associated with the isolates meant that whilst it was possible to achieve intra-species groupings the relevance of the groupings was not fully understood. Without further knowledge surrounding the pathogenicity factors or the differing levels of virulence that exist between isolates, it was not possible to utilise the intra-species groupings achieved. Whilst previously these groupings were believed to reflect the virulence of the organism, taxonomic revision has shown this not to be the case.

It seems that congruence exists between 16S rRNA phylogenies and the ERIC types of isolates (Figure 3.2). This relationship is not understood, as the 16S rRNA subunit is not linked to these inter-genic amplified regions of the genome. The data shown suggests that ERIC type I isolates are monophyletic within a 16S rRNA phylogeny. It is possible to comment on the relationship between ERIC types II-IV and the 16S phylogeny as the true relationship between these isolates is unlikely to be shown with minimal sample size. Chapter 2 also highlighted a link between housekeeping genes and ERIC type I strains, thus suggesting that genetically ERIC types II-IV differ from ERIC type I (of which the type strain belongs to). A larger sample size would be required to determine the genetic relationship existing between ERIC types II-IV, at present it is not possible to accurately separate these groups based upon genetic data.

Whole genome analysis of *P. larvae* highlighting the location and significance of these inter-genic regions would provide an insight into the potential future use of this method for the intra-species differentiation of isolates. Genome analysis

would allow for exploration of the link between 16S and the ERIC types to be better understood. Whilst many methods are available for the intra-species differentiation of these isolates, the true meaning of these differentiations will not be known without a more detailed knowledge of the virulence of the organism.

Isolates previously identified as belonging to subspecies *larvae*, classically associated with AFB outbreaks were shown to belong to ERIC group I. However isolates 6260 and 6261 both isolates originally classified as subspecies *larvae* were shown to belong to ERIC types II and III. Combined with the phylogenetic data collected in a previous experiment, it is likely that these isolates were incorrectly identified at the time of isolation. The use of the previous taxonomic subspecies standards has been completely dropped, so the relevance of identifying previous groupings serves merely as an insight to the origins of the isolate. Whilst it is possible to suggest that these isolates are incorrectly identified, without the use of the previous subspecies differentiation, the production of revised descriptions of the isolates would be meaningless.

Lateral flow devices are widely used within the UK as a secondary diagnostic tool; these devices are also used worldwide for the identification of AFB infections. This experiment showed that whilst the LFDs would detect the predominant ERIC type of AFB infections. They did not reflect the latest taxonomic revision and would not identify outbreaks belonging to ERIC types II-IV. The epidemiology of ERIC type II-IV infections remains unknown, so the impact of not detecting these isolates is not known.

Within the duration of this project no outbreaks were discovered in the UK that belonged to ERIC types II-IV. It therefore seems unreasonable to deem the LFD unfit for purpose, but it should be combined with prior knowledge of AFB infections to ensure ERIC type II-IV outbreaks are correctly identified. A large-scale LFD independent random sampling method, would be required in order to assess the epidemiology of the different ERIC types found within colonies. However if, as previously suggested, ERIC II-IV infections are more likely to exist as acute cell localised infections with limited spread within a colony,

detection of these ERIC types would be unlikely even if infected colonies were sampled.

It is suggested that prior to the re-designing of the antibodies used in the LFD kit, ERIC types II-IV are first isolated from UK colonies. A better understanding of the colony level infections caused by these ERIC types is also required, whilst it is known that infections are more likely to be acute, the effect of these acute infections on the colony is unknown. As experienced bee inspectors perform primary identification, it is unlikely that current testing methods are failing to detect symptomatic AFB infections. It is likely that asymptomatic infections are missed, but it is unknown if these infections lead to a symptomatic AFB infection or would continue to be asymptomatic. Therefore it is unknown if subclinical AFB infections should be detected in order to control the spread of AFB.

3.8. Acknowledgements

I would like to than Catherine Ho for her assistance in developing a working ERIC PCR protocol, capable of producing band patterns as previously seen.

Chapter 4

In-vitro comparison of virulence displayed by *Paenibacillus larvae* isolates

4.1. Summary

- No correlation found between proteolytic activity observed on milk agar and *in-vitro* virulence of isolates
- Strain 6993, suggested to be a non-pathogenic isolate, is highly virulent *in-vitro*
- Production of large quantities of endospores believed to be the phenotypic difference responsible for the difference in virulence observed between strains 6254/6678 and 6993.
- Reducing the initial spore dose of 6993 resulted in similar larval mortality to that produced by strains 6254/6678. Suggesting sporulation was responsible for the difference in *in-vitro* virulence.

4.2. Introduction

AFB is a significant disease of UK apiaries. Although incidence has reduced, this nevertheless remains a destructive disease, whose underlying epidemiology is not well understood (Alippi et al. 2007; Genersch 2009; Genersch et al. 2005). In chapter 2 it was shown that minimal genetic differences existed between isolates of *P. larvae*, however based upon the genes studied it was possible to identify two potential genotypes. Chapter 3 explored the use of ERIC fingerprinting as a means of intra-species differentiation, 4 ERIC types can be identified within *P. larvae*. It is known that ERIC I isolates are responsible for the majority of outbreaks worldwide (Antúnez et al. 2009; Genersch et al. 2006), ERIC IV

isolates whilst capable of causing infection *in-vitro* have never been identified as the causative agent behind a colony wide AFB infection (Genersch et al. 2006). Based upon the genotypes seen in chapter 2, it appears that ERIC types I-III share more sequence homogeny than the more diverse ERIC IV isolates. Detection using LFD kits showed a difference between ERIC type I, II and III, with the device only able to identify bacteria belonging to ERIC type I.

Pathogenicity is the key phenotype, yet the LFD tests and strain assays are based on a limited subset of the diversity within the species. Since the adoption of the 2006 taxonomic revision, to drop the subspecies differentiation based upon *in-vitro* pathogenicity (Genersch et al. 2006), previous intra-species differentiations are not representative of the pathogenic sub-set. To understand fully the epidemiology, distribution and outbreak characteristics, we need a quantitative assessment of pathogenicity among isolates.

Many reference isolates are available of *P. larvae* of both subspecies *larvae* and subspecies *pulvifaciens*, however the information regarding the origins or pathogenicity of the isolates is virtually none existent. The type strain of *P. larvae* 6254 has very little data regarding location or the disease produced. It was apparent that differences observed in the laboratory or in genetic data would have little relevance to the virulence of the organism, as this was essentially unknown. Due to the revised species definition several strains that had been identified as subspecies *pulvifaciens* had no reliable background information, as many of these strains had now been shown to cause infection. The inability to cause infection was considered the main factor in the taxonomic definition of subspecies *pulvifaciens*. Tracking down the original depositors of the strains was not an option, as often this information was missing. The majority of isolates were deposited over 50 years ago; it was unlikely after this time that the contact information was still correct. To reliably understand the differing virulence between isolates, new data would have to be collected.

4.2.1 Assessing the virulence of *P. larvae* isolates

The intra-species differences in virulence have been more widely studied in bacteria capable of causing human infection. The normal human intestinal gut flora contains *E. coli* (Eckburg 2005), despite its presence one strain of *E. coli* (0157:H7) is considered potentially life threatening (Tarr 1995). Despite the vast differences in symptoms produced by *E. coli* infections, all isolates remain grouped under one species differentiated as serotypes of the species. This project aimed to determine whether a differentiation could be made between pathogenic and non-pathogenic strains of *Paenibacillus larvae*. It was apparent that methodologies used to differentiate strains was inadequate and did not reflect the pathogenicity of all *P. larvae* isolates. The route of infection in *P. larvae* is not fully understood. It has been shown that only spores are capable of causing infection (Genersch 2010a; Genersch 2010b), after germination and massive proliferation in the larval gut (Yue et al. 2008) the infection becomes systemic. It is not understood how the bacterium is able to travel across the larval gut wall, however it has been suggested that the bacterium exploits a weak section of the gut (Yue et al. 2008). Once the infection has become systemic the bacterium begins to digest the larval remains until only bacterial spores remain in the form of AFB scales (Alippi et al. 2005). Whilst the mode of infection is not understood two key factors can be identified from the infection pathway; bacterial spores and proteolytic enzymes.

4.2.2. Proteases as a virulence factor?

Bacterial proteases are often over looked as virulence factors, as they are considered to be core to the survival of the bacterium under any growth conditions. It has been suggested that proteases should be considered virulence factors if shown to aid the infection of a host (Lantz 1997). Studies of the *Bacillus* genus have shown that *B. subtilis*, often considered a non-pathogenic bacterium, contains 24 proteases genes compared to the pathogenic species *B. cereus* containing 49-61 and *B. thuringiensis* containing 52 (Han et al. 2006). Whilst no direct link exists between the additional protease genes and the

virulence of the organism, it is reasonable to assume that these genes were acquired to provide an advantage to the infectious bacterium. Comparisons can be made between *B. thuringiensis* infections and *P. larvae* infections as symptoms produced are indicative of proteolytic activity.

Although no protease has been proven to be implicitly involved in the virulence of *P. larvae*, many have been suggested as potential virulence factors (Dancer and Chantawannakul 1997; Holst and Sturtevant 1940). Commonly implied is the involvement of a metalloprotease (Dancer and Chantawannakul 1997; Hrabak and Martinek 2007; Jarosz and Glinski 1990). Metalloproteases have been suggested as a potential virulence factor in *B. anthracis* (Chung et al. 2006), a closely related organism to *B. thuringiensis*. Genersch (2004) identified a metalloprotease only present in subspecies *larvae*. Whilst subsequent *in-vitro* pathogenicity testing revealed isolates not possessing this metalloprotease were able to cause infection (Ashiralieva and Genersch 2006), it does highlight the potential for phenotypic differences to be observed.

One of the earliest identification methods for AFB was the observation of proteolytic activity on milk (Alippi 2001). The presence of proteases produced a zone of clearing, which was believed to indicate a positive AFB test result. With the symptoms displayed and the lack of any other suggested virulence factors, identifying differences in the proteolytic activity of isolates was the logical starting point for the differentiation of isolates based upon phenotypic properties.

4.2.3. Sporulation essential for virulence of *P. larvae*

P. larvae is only capable of causing infection as a spore (Genersch 2010a), so the ability to form adequate numbers of viable spores is arguably a critical virulence factor. It was previously believed that *P. larvae* once in the larval mid gut existed solely as vegetative cells (Dancer and Chantawannakul 1997). The existence of giant whips, considered an indication of sporulation, at all time points during infection suggests that *P. larvae* sporulates throughout (Yue et al. 2008). It has been suggested that proteases are produced by *P. larvae* at the point of

sporulation (Dancer and Chantawannakul 1997), this has been noted in other spore-forming bacteria (Hoch 1993).

Research into the sporulation of *P. larvae* had suggested it was a notoriously difficult organism to induce sporulation in, however previous experimental work in this project had not exhibited this problem. The pathogenic background of the majority of isolates held, suggested that isolates all posed the ability to sporulate in sufficient numbers to cause infection. Previous experimental work did not require specific spore counts for individual isolates.

4.2.4. Origins of the *in-vitro* Honeybee rearing methodology

Early beekeeping involved the capture of a swarm of bees in a suitable hive like container. During the 1600s scientists and beekeepers were able to observe bees much more closely due to improving hive designs. As the understanding of honeybee lifecycles improved, beekeepers were able to further improve husbandry techniques effectively shaping the way that beekeeping is carried out today out. As the importance of breeding and husbandry became clear, methods were developed that allowed controlled breeding within their apiaries, effectively allowing the selection of queens that would produce the best offspring.

This method involved grafting larvae into special cell cups that were then adhered to a specially adapted frame. This was then inserted into the centre of a brood box, whilst the queen of that hive was excluded from the chamber. Workers would then continue the process and raise new queens in the absence of their own queen. Once the cells are capped over, they can be moved to a special smaller hive called a nuc. Workers continue to raise the queen until emergence. At this point controlled mating can be carried out using pre-selected drones. This produces a mated queen ready to be inserted into a colony at which point she would begin to lay.

This method has been adapted to enable rearing of honeybee larvae in a sterile environment. Understanding the key milestones of the honeybee lifecycle (figure 4.1) and dietary requirements of the larvae it is possible to successfully rear the larvae *in-vitro*.

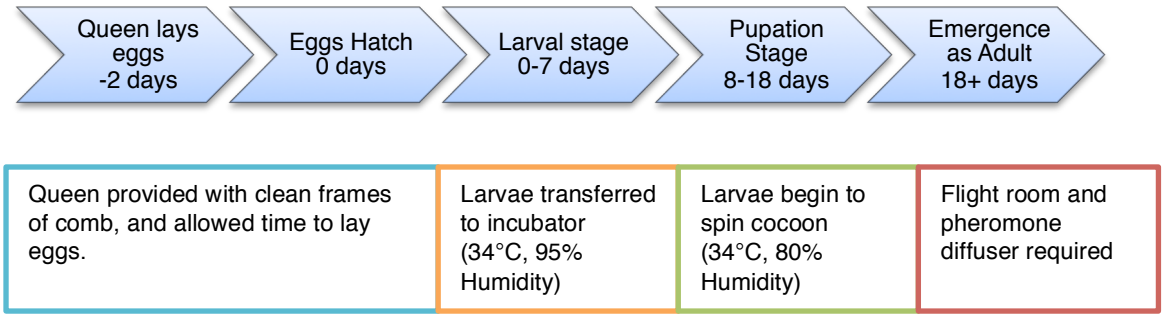


Figure 4.1: Honeybee larval development and how this impacts *in-vitro* rearing.

4.2.5. Original *in-vitro* Honeybee rearing method

A consortium of European laboratories lead by the French National Institute for Agricultural Research (INRA) have developed a standard method defining optimum age and size of brood to be grafted into artificial cell cups (Appendix 7). By maintaining the optimum environmental conditions, whilst providing the appropriate diet honeybee larvae reared this way can reach maturity. The method was developed for pesticides testing to evaluate the impact on honeybee brood. This standardised test allows comparison of results from different laboratories.

Genersch (Ashiralieva and Genersch 2006; Genersch et al. 2005; Genersch et al. 2006) adapted this method for the study of larval mortality after AFB infection. The methodology allowed the infections to be observed at various time points prior to death; this was carried out in a sterile environment under highly controlled conditions. This method was able to standardise the pathogenicity testing, so that results could be compared. It did this by removing the differences between colonies and environmental conditions at different locations. It was possible to observe AFB infections under controlled conditions; as such direct comparisons could be made between isolates.

4.3. Aims

- To determine if a correlation exists between the level of proteolytic activity observed and the in-vitro virulence of isolates.
- To collect more quantifiable pathogenicity data from isolates.
- To determine if the type-strain isolate of *P. larvae* (6254) is an accurate representative of the pathogenic sub-type.
- To determine if strain 6993 is non-pathogenic as previously stated.

4.4. Materials and Methods

4.4.1. Bacterial Isolates

Initially the relationship of proteolytic activity and virulence was studied using 5 isolates (table 4.1). The experiment was later expanded to compare the *in-vitro* virulence of 11 reference isolates (figure 4.3). Full details of these isolates can be found in chapter 2.4.1.

4.4.2. Growth of Bacterial Isolates

Bacterial isolates were grown from frozen storage beads when required on BHI-T or TMYGP growth media at 34°C (appendix 1).

4.4.3. Protease Screening

Milk agar was made to a 5% concentration using Oxoid agar No.1. The milk solution was heated to 72°C using a steamer, avoiding the denaturing of the proteins. This milk solution was added to molten agar at <50°C. Milk agar was inoculated using a sterile spike inoculated with a broth *P. larvae* culture. In the presence of proteases, a zone of clearing is produced around the site of

inoculation. If no zone of clearing was present this indicated a lack of proteolytic activity. Data produced was statistically analysed using XLSTAT 2012, ANOVA testing was performed to test the null hypothesis that inoculation with *P. larvae* spores is not responsible for producing zones of clearing; indicating protease production.

4.4.4. Inducing sporulation

BHI-T broth was used for the accumulation of bacterial spores. This was incubated at 34°C and shaken on an orbital shaker at 170rpm. Spore counts were made after 72 hours. The following media were tested in order to induce sporulation;

- BHI / BHI-T
- TMYGP
- Corn steep liquor
- Corn Steep solids
- Blood agar slants
- Blood agar plates / BHI & BHI-T plates

No significant increase in sporulation was detected; it was decided to use BHI-T liquid media for the sporulation of *P. larvae* isolates.

4.4.5. Enumeration of spores

1 ml of spore culture was taken and spun in a centrifuge at 13,000rpm for 1 minute. The pellet was then re-suspended in dH_2O before heating to 80°C for 10 minutes to kill all vegetative cells. Serial dilutions were made of the heated spore suspension; these dilutions were plated on BHI-T agar and incubated at 34°C for 72 hours. The number of colonies present after this time allowed for the calculation of the number of viable spores present in the spore suspension.

4.4.6. Exposure Assay

The in-vitro exposure assay (Appendix 8) was used to measure the difference in virulence exhibited by the strains producing differing levels of proteolytic activity (table 4.2). Survival analysis was performed using XLSTAT2012, this allowed the comparison of data providing statistical representation of the differences exhibited after treatment with *P. larvae* strains. The same method was also employed to test a wider number of strains (figure 4.3). In 2010 the method was adapted to test one strain of bacteria at a variety of pre-determined spore loads. Experimental set up was identical to previous experiments, with differing spore counts replacing different strains of bacteria. XLSTAT was used to calculate the significance of the difference existing between different replicates of the same spore dose.

The experiment used 48 well plates, 2 groups of 24 larvae per plate. Due to the experimental set up each larva was considered a replicate, therefore each bacterial isolate or control contained 24 replicates.

4.4.7. Evaluating differences in virulence between ERIC types I and IV

This experiment was set to run from June – September 2010, testing 4 strains of *P. larvae* along with a negative control.

Treatment	ERIC Type
Negative Control (dH_2O)	N/A
6254	I
6678	I
6264	IV
6993	IV

Table 4.1: details of the strains chosen to explore the differences in virulence existing between ERIC type I and IV.

Strains 6254 and 6993 were chosen due to their previous phenotypic observations, with 6254 being labeled pathogenic and 6993 labeled non-

pathogenic. Strains 6264 and 6678 both displayed moderate levels of virulence representing the majority of isolates tested.

Strain 6254 is the type strain of pathogenic *P. larvae* isolates so was included as a point of reference. Previous work had highlighted strain 6993 as a potential bio-control isolate for use against EFB, this strain was also an example of subspecies *pulvifaciens* and was thought to be non-pathogenic. Strain 6678 is an ERIC type I replicate and strain 6264 is an ERIC IV replicate. The use of these isolates allows comparison both between the different ERIC types and amongst the same ERIC types to look at the variation of results.

4.5. Results

4.5.1. Measuring Proteolytic activity

Strain	LMG Ref	Average Zone of Clearing (mm) *[Minimum / Maximum]*		Previous taxonomic Grouping
6932	N/A	9.5	*[8 / 11]*	N/A
6255	14425	8	*[8 / 8]*	Subspecies *larvae*
6256	14426	11	*[11 / 11]*	Subspecies *larvae*
6930	N/A	12	*[12 / 12]*	N/A
6993	14427	0	*[0 / 0]*	Subspecies *pulvifaciens*

Table 4.2: The observed proteolytic activity of 5 *P. larvae* isolates, average zone of clearing produced is shown with the minimum and maximum values also included. The previous subspecies identification of the isolates is also listed.

Source	Degrees of Freedom	Sum of squares	Mean squares	F	P Value
Model	5	291.750	58.350	77.800	< 0.0001
Error	6	4.500	0.750		
Corrected Total	11	296.250			

Table 4.3: ANOVA analysis of the variance within the data set

The milk agar assay displayed varied results, isolates 6255 and 6256 did produce zones of clearing, and similar results were achieved from the UK environmental

isolates 6932 and 6930. The proposed bio-control isolate 6993 produced no zone of clearing. ANOVA analysis of the data (table 4.3) produced showed that the difference between treatments was significantly different across the 5 groups $F^{5,6}$ = 77.8, p= <0.0001. Post-Hoc Dunnett (two sided) analysis of the differences between treatments and the control category (H2O) showed that the greatest zone of clearing was caused by isolate 6930 (12mm). The same analysis showed isolates 6256, 6932 and 6255 (ranked in order below isolate 6930) to cause a significant zone of clearing (p= <0.0001) at a 95% confidence interval.
Isolate 6993 was shown not to cause a significant difference in the zone of clearing produced (p=1) at a 95% confidence interval.

4.5.2. Correlation between observed proteolytic activity and *in-vitro* virulence

In-vitro pathogenicity testing revealed that the correlation expected between proteolytic activity and virulence was not seen (Figure 4.2). Strains producing larger amounts of proteolytic activity (6255, 6256, 6930 and 6932), exhibited very low levels of virulence in-vitro. The maximum number of deceased larvae for these isolates was seven, whilst the minimum was four after two hundred and forty hours. The bio-control isolate (6093) was shown to display high levels of virulence, despite exhibiting low levels of proteolytic activity. Both replicates of this isolate killed 100% of the test larvae within one hundred and ninety two hours.

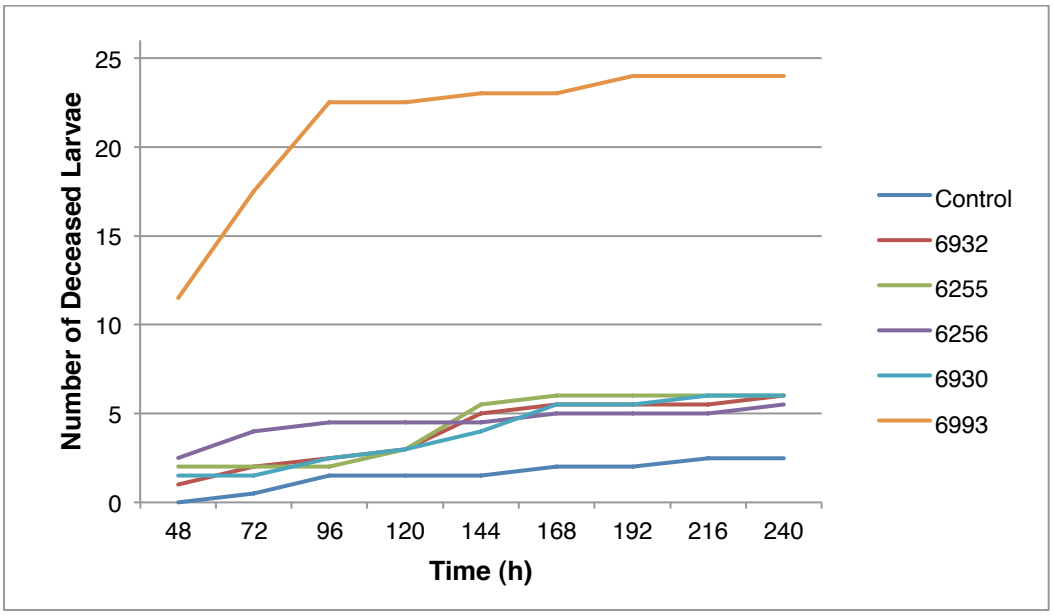


Figure 4.2: Time series plot of larval mortality when infected with isolates of P. larvae. Each sample comprised of 24 honeybee larvae, and the experiment was performed twice on separate occasions, average results are shown. Survival analysis showed that all isolated tested were significantly different to the control.

Treatment	Mean Time of Death (hours)
6932	210
6255	154
6256	205
6930	192
6993	72
Control	207

Table 4.4: The average time taken by the tested treatments to cause mortality in the test larva. Calculated by Kaplan-meier survival analysis.

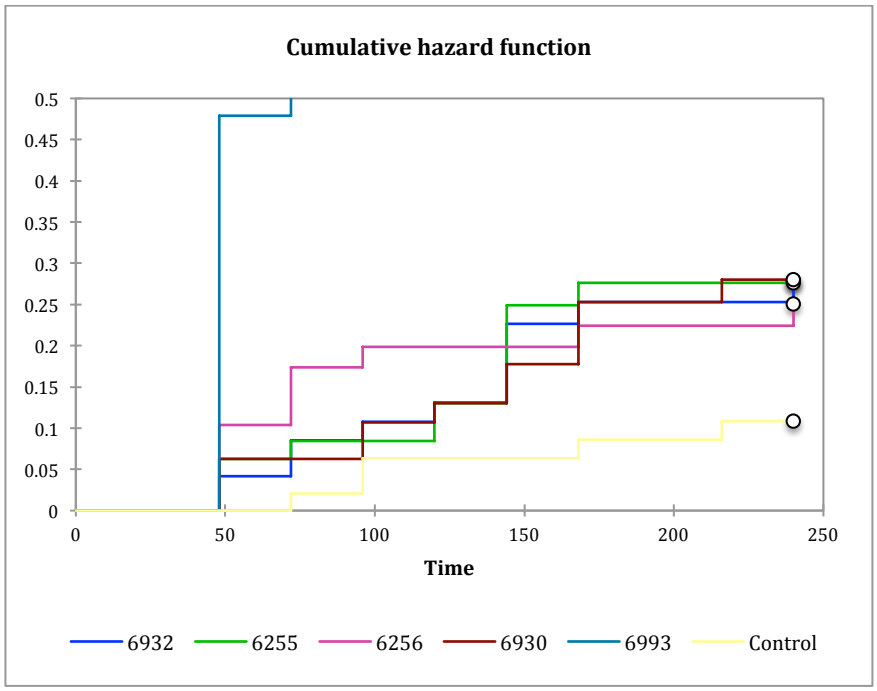


Figure 4.3: The cumulative hazard function calculated using Nelson-aalen survival analysis. This represents the hazard associated with each respective treatment.

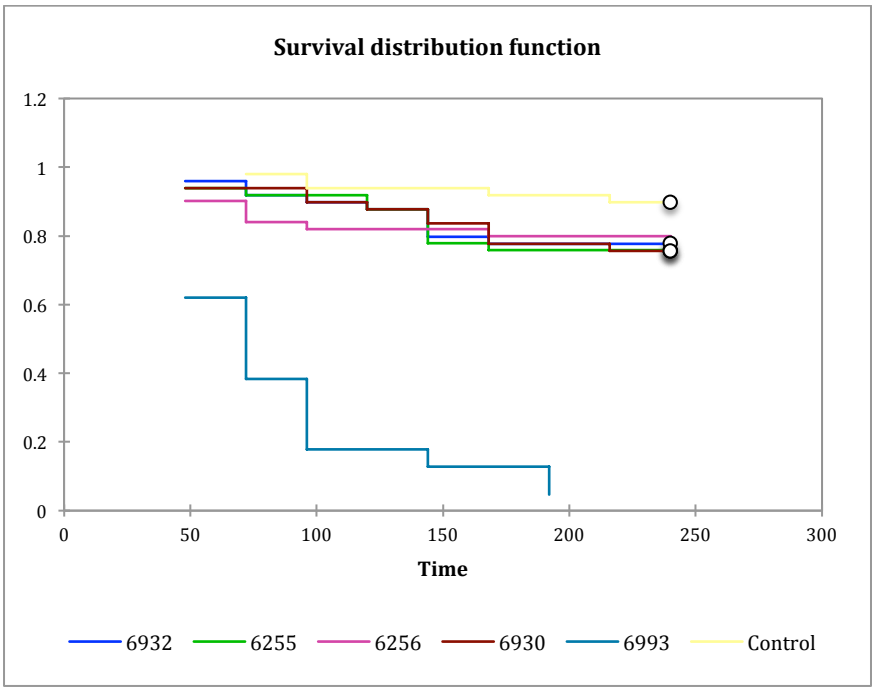


Figure 4.4: The survival distribution function as calculated by Nelson-aalen survival analysis. This represents the likelihood of survival by larvae infected with the respective treatments.

Survival analysis confirms what was displayed in the time series chart, that treatment with isolate 6993 poses the greatest hazard to the survival of honeybee

larva. This can be clearly seen in the mean time to death as well as the cumulative hazard function and the survival distribution function (tables 4.3 and 4.4). The remaining treatments all proved to be less likely to cause larval mortality, this was again reflected in the mean time to death, cumulative hazard function and survival distribution function. The control was shown by both survival distribution function and cumulative hazard function to have the least affect on the health of larva. However the mean time to death of control larvae was calculated to be lower than that of treatment 6932, this can be explained after inspecting the raw data. Treatment with 6932 caused deaths at later time points than the control where the majority of deaths happened closer to the beginning of the experiment.

4.5.3. Comparative *in-vitro* virulence testing of reference isolates

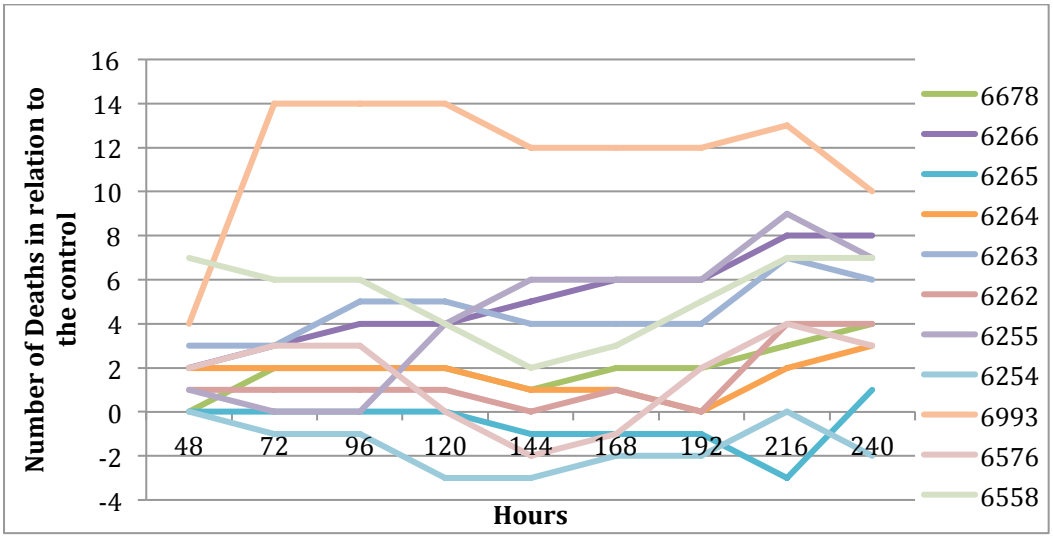


Figure 4.5: Time series plot of larval mortality from all reference isolates, data normalised against control mortality to allow comparisons between grafting sub-sets.

The data collected showed that most strains failed to diverge more than +/-5 larvae away from the mortality rate of the control cells (figure 4.5). Several strains displayed lower levels of mortality than the control; one of these isolates was strain 6254. This isolate is considered the pathogenic type strain of the species. The majority of isolates tested displayed between 3 and 8 more deaths than the control at 240 hours. Once more the replicate bio-control isolate 6993 exhibited the highest level of virulence.

4.5.4. Differences in sporulation

Initial experiments were carried out on a select few strains of *P. larvae*, these experiments looked at the levels of sporulation exhibited by the strains in question. No significant differences were observed at this time, all strains tested produced relatively similar numbers of viable spores. Earlier pathogenicity testing had sought to cause infection, spore doses were not being tested and it was planned to provide an excessive amount of spores in order to maximize the symptoms observed. All strains were grown to an optical density that was known

in several isolates to produce an excessive amount of spores. The cells were harvested and heat shocked to produce a spore suspension.

This experiment highlighted that there was a vast difference in sporulation between ERIC I and ERIC IV strains, to the order of 10^3 spores. The difference in levels of sporulation was hard to correct due to the low levels produced by ERIC I isolates. The maximum number of spores produced was 200 ml^{-1}, it was calculated that experiments required at least 20,000 spores. Upon scaling up the volume of the ERIC I sporulation cultures, showed that over 15ml the amount of viable spores decrease, whilst volumes over 50ml produced no viable spores. This study suggests the optimum volume for sporulation is <10ml, however >25% of the inoculated broths produced no viable spores.

4.5.5. Evaluating the virulence of strain 6993 at lower spore doses

In 2009 all strains were grown to a uniformed optical density. From this it was possible to calculate the number of spores fed to the larvae in this experiment. Previous tests had shown that this optical density produced sufficient spores to cause infection.

Strain	Viable Spores at OD 0.60 (ml^{-1})	Spore dose per *in-vitro* well
6254	70	44
6993	22,000	13,750

Table 4.5: Enumeration of viable spores ml^{-1}, at the optical density strains were grown to for previous in-vitro testing.

It was possible to calculate the dose of spores fed to the larvae in the 2009 experiment. This showed that larvae infected with isolate 6254 had received 44 spores each, whilst larvae inoculated with strain 6993 had received 13,750 spores each.

Due to the large quantity of spores produced by strain 6993, it was possible to observe the virulence of the isolate at varying doses. Virulence data was collected from the varied spore doses (Figures 4.6 and 4.7). It was shown that

similar levels of virulence were displayed by strain 6993 when the initial spore dose was lowered to a similar value to that of 6254. Results from plate one showed the expected dose curve, with an increase in larval mortality correlating with the increased spore dose. Results from plate 2 showed a greater number of larval deaths overall, and all infected larvae produced very similar levels of mortality with 20 spores showing higher levels of mortality than 200 spores. The control values on both plates showed very similar levels of mortality. Comparison of the survival distribution function plots was able to provide a measure of the significance of the difference existing between the treatments on plate 1 and 2. In the instance of the control and the 2000 spore dose, the difference observed between plates was statistically insignificant with a P-value of >0.01. The differences observed between plates at 20 and 200 spore doses were calculated to be significant, with the difference at a dose of 20 spores being the most statistically significant.

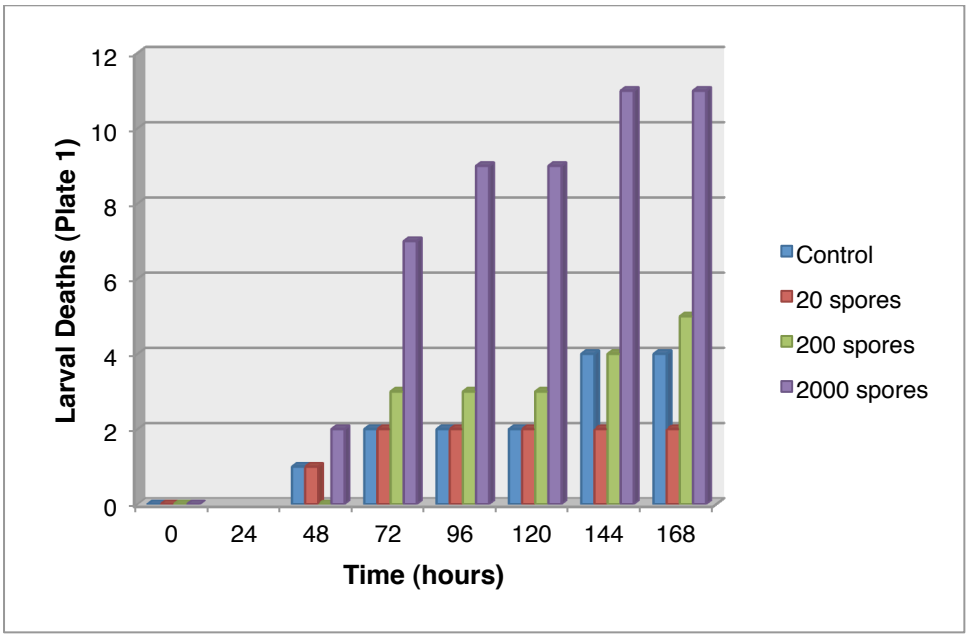


Figure 4.6: Larval mortality at 24-hour intervals. Each group consisted of 12 larvae and data shown on this graph represents the first repetition.

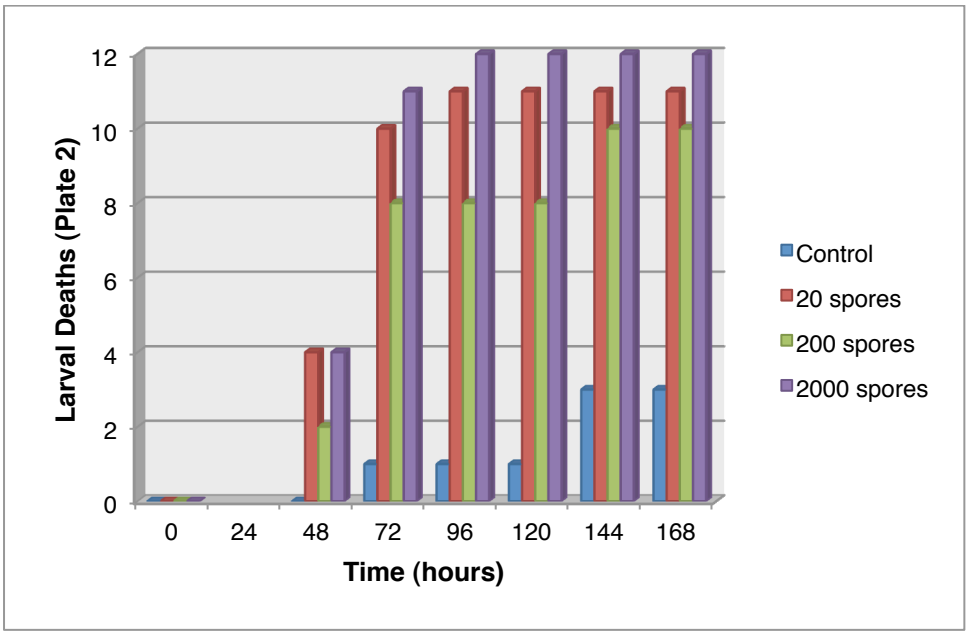


Figure 4.7: Larval mortality at 24-hour intervals. Each group consisted of 12 larvae and data shown on this graph represents the second repetition.

4.6. Discussion

4.6.1. Evaluation of the laboratory *in-vitro* honeybee rearing method

Development of the method demonstrated that *in-vitro* testing could be used successfully to follow larval development. Several larvae tested were able to develop well into the pupation stage at which time they took on a more adult-like form. The method was however terminated before any of these pupae had the chance to emerge as adult bees, however it is likely that adult bees could be produced using this method.

Most AFB infections take place very early on in the life cycle of the larvae, however some strains can take up to 8-12 days to show symptoms. After this test it was decided that 10 days was the optimum amount of time for this experiment to run, at around this stage larvae begin to defecate which leaves them in un-hygienic conditions without the worker bees to clean out the cells. It was decided that it was impractical to move them in to a new clean environment at this point as that introduces yet more variables and more manual handling in to the already

complex procedure. Background mortality rates were sufficiently low (<10% - Appendix 8) that 4 plates per day (192 larvae) could be grafted per day for meaningful results.

4.6.2. Is proteolytic activity a good proxy for determining the virulence of isolates?

Analysis of isolates on milk agar showed that several strains classically defined as subspecies *larvae* produced larger zones of clearing than those classically defined as *pulvifaciens*. Whilst the UK environmental isolates were never defined to a subspecies level, it was known that these isolates came from colonies displaying typical AFB symptoms. Proteolytic activity has always been associated with AFB infections (Dancer and Chantawannakul 1997; Holst and Sturtevant 1940). Despite several attempts to identify the protease responsible for the symptoms; no research has ever directly linked proteolytic activity with the virulence of the organism.

Based upon the assumed involvement of a proteolytic enzyme in the virulence of *P. larvae*, the results obtained were to be expected. With isolates from a known pathogenic background exhibiting the most proteolytic action and isolates defined as non-pathogenic displaying no proteolytic action. Upon testing the virulence of these isolates in-vitro, a negative correlation was observed. Pathogenicity testing showed that the isolates believed to be pathogenic, displayed low levels of virulence and isolates termed non-pathogenic displayed very high levels of virulence.

This showed that detectable differences in the proteolytic activity produced by these isolates, could not be used to predict the in-vitro virulence. The results here are unable to confirm the implication of proteases in the virulence of *P. larvae*, as it is likely that growth media used with the laboratory does not cause the expression of the same genes that growth within a larval gut would.

4.6.3. Comparative *in-vitro* virulence testing on a wider number of isolates

Decisions had to be made about the future of this experiment; otherwise further testing would have to be put on hold until the 2010 beekeeping season. It was at this time that it was decided to collect as much data as possible before the end of the 2009 season, leaving any and all analysis until the quieter winter period.

The *in-vitro* assay showed variability among strains in virulence. A degree of variability was expected between isolates, though previously all strains of *P. larvae* were shown capable of causing increased larval mortality (Ashiralieva and Genersch 2006; Genersch et al. 2005; Genersch et al. 2006). Differences had been previously noted in the time taken to reach larval mortality (Ashiralieva and Genersch 2006; Genersch et al. 2005; Genersch et al. 2006), although no significant change was noted in the percentage of larval mortality. It is worth noting that the methods employed by Genersch (Ashiralieva and Genersch 2006; Genersch et al. 2005; Genersch et al. 2006) differ from the methodology used within this thesis, Genersch used controlled doses of bacterial spores and after the initial feed containing bacterial spores larvae were transferred to a new sterile environment to observed disease progression. Whilst this does mimic hygienic behavior exhibited by worker bees, it was believed that the increased handling of the larvae would likely cause increased larval mortality. The difference existing between the methodologies has not yet been compared.

This experiment allowed all strains to be compared equally, without having to interpret various sources of data surrounding the epidemiology of the various outbreaks and unlike those sources of data; the credibility of this experiment did not need to be considered during the analysis. Once again it was shown that strain 6254 did not display the expected levels of virulence. AFB infections should take around 8 days before larval death is observed (Genersch 2010a; Genersch et al. 2005), but this experiment ran for 10 days and no increase in larval death was noticed throughout. By contrast, strain 6993 displayed high levels of virulence despite the strain being suggested as a potential bio-control to

prevent EFB in honeybees (Watkins et al. 2003). *In-vitro* testing had shown that use of this isolate for control of EFB was not possible due to the hypervirulence observed (figure 4.1-3). The reason for this increased level of virulence was unknown, however it was previously noticed that this strain caused increased levels of larval mortality (Watkins et al. 2003). This was attributed to an increased initial spore load, and was not considered problematic. It is not advisable to continue the work regarding this isolate as a bio-control, as it is highly likely any colonies dosed with these spores will result in significant levels of larval mortality. Over 96% of larvae infected with this isolate were deceased within the 10 days of the experiment.

4.6.4. Confirming *P. larvae* was responsible for larval mortality

Ideally a Koch's postulates type experiment would have been performed whereby the bacteria recovered from the deceased larvae was identified and used to re-infect a further group of larvae. This was not deemed possible, whilst it was possible to recover *P. larvae* from all deaths associated to strain 6993, it was not possible to confirm beyond reasonable doubt that the same strain of bacteria had been recovered. Deceased larvae often measured <2mm x 1mm, a solution could not be found to ensure that the recovered bacteria originated within the larval gut. Methods were trialed whereby larvae were cleaned, to ensure that bacterial spores did not exist on the surface of the larvae, due to the fragile nature of the decaying larvae it was not possible to adequately clean the external surface of the larvae to remove all bacterial spores. As a large control set was set up alongside all subsets of isolates tested, this ensured that under non-infected conditions lower levels of larval mortality were seen.

4.6.5. Evaluating the observed difference in *in-vitro* virulence between ERIC type I and ERIC type IV isolates

Whilst it would have been possible to collect repeat data from all of the strains tested in 2009, at this time the differences between strains 6993 (considered non-pathogenic) and 6254 (considered pathogenic) were of greater interest. Whilst recent research had suggested that ERIC IV isolates had the potential to be more pathogenic than ERIC I isolates. The suggestion of isolate 6993 as a bio-control for use in honeybees, suggested that the levels of virulence displayed within this project were higher than those previously observed. ERIC I isolates, of which 6254 is the type strain, have been associated with the vast majority of AFB outbreaks around the world (Antunez et al. 2009; Ashiralieva and Genersch 2006). Within the UK all environmental samples tested, proved to be ERIC type I (Chapter 3.5.1). It was not understood why this difference in virulence existed *in-vitro*.

It has been shown that *in-vitro* studies of AFB infections have a negative correlation with in-situ testing (Rauch et al. 2009). This suggested that higher rates of larval mortality observed *in-vitro* were not observed in-situ, hygienic behavior of workers was deemed to be responsible for this difference. Work carried out *in-vitro* excludes adult bees to limited contamination between replicates. In a colony larvae that die prior to cell capping are removed and the cells are cleaned. AFB infections must progress to the production of AFB scales for the disease to be spread by workers (Lindström 2008; Lindström et al. 2008). This suggests that rapid *in-vitro* larval mortality is more likely to occur as an acute infection within a colony, opposed to slower rates of *in-vitro* mortality capable of causing a systemic colony infection. Whilst it had been shown that isolates displaying more rapid larval mortality were less virulent within a hive. Strain 6254 was not able to increase larval mortality over the control cells this could not be explained.

Strains 6678 and 6264 were chosen as replicates, both isolates displayed low levels of in-vitro virulence. The additional study of these isolates, allowed for the comparison of differences both between ERIC types and of the differing virulence displayed within an ERIC type.

4.6.6. Differences in sporulation

Data collected in 2009 agreed with previously published data, showing that ERIC IV strains are able to exhibit far greater levels of virulence in an in-vitro setting (Rauch et al. 2009). As such the observed differences in rate of larval mortality were accepted as this appeared to concur with similar research (Genersch et al. 2005; Genersch et al. 2006). It was believed that the length of the *in-vitro* experiment was at least in part responsible for the lower levels of larval mortality exhibited by ERIC I infections, as it has been noted that the average time till larval mortality is approximately 8-12 days for an ERIC I infection compared to 3-4 days for an ERIC IV infection (Genersch et al. 2005; Genersch et al. 2006).

Early experimental data suggested that whilst there were distinct morphological differences between isolates, all of these isolates were able to produce spores. This was expected, as these strains had all been collected from hives showing various symptoms. Given that *P. larvae* is only infective as a spore (Bakonyi et al. 2003), then it was logical to assume that all strains had this capability. Previously it had been shown that a larger proportion of ERIC type IV spores were capable of germination (Forsgren et al. 2008), but also shown that a larger number of ERIC type IV spores were required to cause infection (Ashiralieva and Genersch 2006; Genersch 2009; Genersch et al. 2005).

Pathogenicity testing during 2010 required specific doses of spores, until this point it had been assumed that whilst there would be differences in the numbers of spores produced by different strains, these levels would be comparable. An optical density was chosen for previous testing, known to provide an excess of viable spores. Initial testing simply focused on causing infection; as a result the

dose of spores was intentionally kept high, this was comparable with other research on AFB *in-vitro* pathogenicity (Genersch et al. 2005). For the testing during 2010, it was decided that the dose of spores needed to be controlled in order to make the data more comparable between strains. The inability to produce ERIC type I spores led to major delays in the *in-vitro* exposure assays being employed by this project. No practical solutions could be thought of at this time. As the inconsistency of the results produced by ERIC I spore cultures meant that no workable solution could be realised within the scope of this project.

It is necessary to re-visit the lack of correlation between proteolytic activity and *in-vitro* virulence in light of the differences in sporulation. It has been repeatedly suggested that proteases are produced upon sporulation (Dancer and Chantawannakul 1997). It is likely that differences in sporulation affected the results produced by this experiment. However strain 6993 was shown to produce large quantities of spores but lacked proteolytic activity. It is therefore predicted that whilst proteolytic activity may be a virulence factor, it is likely not essential for the infection of honeybee larvae. As previously mentioned it was not possible to increase the number of spores produced by ERIC type I isolates, so was not possible to explore the possibility of increased proteolytic activity at higher spore concentrations. Whilst causing greatly increased larval mortality, strain 6993 did not produce any of the typical AFB symptoms expected with the larva not decaying to form a ropey mass. The increased *in-vitro* virulence produced by ERIC type IV isolate 6993 would suggest that proteolytic enzymes are responsible for symptoms produced and are not essential for host invasion.

4.6.7. Sporulation responsible for the differences in virulence between strains 6254/6678 and 6993

This experiment was able to first calculate the initial spore dose fed to larvae infected with strain 6254 in 2009 (Table 4.3). Whilst this number of spores was capable of causing infection in larvae of the age tested (Woodrow and Holst 1942). The gross difference in initial doses was predicted to be the reason for the observed differences in virulence. In previous exposure assays, controlled doses

of spores had been used (Genersch et al. 2006; Rauch et al. 2009). The methodologies responsible for producing large quantities of ERIC type I spores (Genersch et al. 2005) were trialed, but in all instances the calculated number of viable spores was shown to be lower than growth in standard BHI-T broth media. Previously spore doses fed to larvae had been estimated (Genersch et al. 2005) or were calculated but variable between isolates in order to ensure infection (Genersch et al. 2006; Yue et al. 2008). It was possible to reduce the initial spore dose of 6993 to similar levels of strain 6254 used in 2009. The reduction of spore doses was able to reduce the level of virulence (Figure 4.4), suggesting that the increased virulence was due to a reduction in the time taken to proliferate within the larval gut. It has been shown that the rupturing of the gut, and the infection becoming systemic likely causes larval mortality (Yue et al. 2008).

This experiment was set up in parallel, 2 repeats were simultaneously run. Unfortunately the two repeats for this experiment produced differing results. Plate one produced the results expected, with all doses given resulting in a low level of larval mortality (figure 4.4) than that observed in 2009 (figure 4.3). Also seen here was that doses of 20 and 200 spores were not able to cause a greater level of mortality than a ddH_2O control. The experiment ended with the control wells having a greater number of dead larvae than the wells dosed with 20 spores, deaths caused by the control and 20 spore doses are likely unrelated to the inoculum added to the food. Death was more likely due to other complications (appendix 8). As expected the 200 and 2000 spore dose caused greater level of mortality, final numbers of deceased larvae were 5 and 11 respectively. These results agree with the suggestion that larval death is caused by massive proliferation of the bacteria within the larval gut, as a 10 fold increase in initial population will result in a decrease in time taken to reach the population required to cause larval death.

Plate 2 displayed similar results, but the numbers of dead larvae for 20 and 200 spores were far closer to that of 2000 spores (figure 4.5). The number of control deaths was comparable between plates, suggesting that human error was not to blame for the increase in larval mortality. The larvae selected for this experiment came from a very limited section of comb, with the larvae used on plate one

being selected from side A and plate two from side B. The susceptibility of larvae to *P. larvae* infections deceases with the increasing age of the larvae (Brødsgaard et al. 1998). A difference of 12-24 hours in the age of the larvae used, would be very difficult to visually detect in the appearance of the larvae. A change in age of this size could dramatically increase the initial number of spores required to cause infection. Knowledge of laying patterns followed by queen bees, suggests that a difference of up to 24 hours could be found between the larvae on differing sides of the comb (Camazine 1991). It is suggested that the observed difference between the results obtained from the replicate plates are a direct result of larval age. It is believed that the difference in age between larvae on sides A and B of the comb was approximately 12 hours, it is suggested that larvae used on plate 2 were younger than those on plate 1, resulting in the increased larval mortality observed at lower spore doses. Statistical analysis of the results also suggests that larval age was likely responsible for the differences observed, as it was shown that the difference between the controls and 2000 spore doses were insignificant. The 20 and 200 spores doses were calculated to be significant, with the difference existing between 20 spore doses being the most significant. The significant differences seen at low spore doses concurs with the theory that larval age was likely responsible for the differences.

4.7. Conclusions

Proteolytic activity was believed to be a significant virulence factor in AFB infections, it was therefore assumed that a correlation would be detected between proteolytic activity and observed *in-vitro* virulence. Testing of this hypothesis showed that this was not the case, only 4 isolates of *P. larvae* showed proteolytic activity on the assay used. Crucially the isolate believed to be non pathogenic (6993) produced no proteolytic activity (table 4.2). *In-vitro* virulence testing of these isolates revealed that it was strain 6993 that caused increased levels of larval mortality. This was confirmed by repeating the experiment; this suggested that proteolytic activity observed in the lab was not a marker for predicting virulence.

It was shown that differences exist in the virulence of *P. larvae* isolates (Figure 4.3). Significantly, ERIC type I isolates formerly identified as subspecies *larvae* were shown not to be capable of causing a significant increase in larval mortality. Whilst it has been suggested that strains belonging to ERIC IV (subspecies *pulvifaciens*) were capable of being more virulent than other strains, this was not fully supported by this experiment. Strain 6993 displayed the increased levels of virulence which had previously been reported, however 6264 showed no increased in larval mortality throughout the course of the experiment. Larval mortality was comparable between the majorities of bacterial isolates, only strain 6993 produced results that significantly differed from the mean.

The enumeration of the viable spores produced by these isolates, showed vast differences in the production of viable spores. It is thought that these differences led to differing initial spore doses (table 4.3), whilst all the calculated doses are capable of causing infection in larvae of that age (Woodrow and Holst 1942), the difference in initial spore load of an ERIC type IV isolate was as high as 314 times that of an ERIC type I isolate. The virulence of *P. larvae* is not fully understood, but it has been shown that massive proliferation and the eventual bursting of the larval gut are likely the cause of mortality (Yue et al. 2008). By reducing the number of 6993 spores it was shown that this strain is not hypervirulent, rather it is capable of producing larger quantities of viable spores under laboratory conditions. At a dose of 20 spores per larvae, strain 6993 was not capable of causing greater larval mortality than the control (Figure 4.4). This experiment was not replicable within this project, and the results from the repeat showed that a dose of 20 spores was capable of causing significant numbers of larval deaths (figure 4.5). Due to the understanding of queen behavior, it is believed that the difference in larval deaths results from a difference in larval age. It has been shown that susceptibility to AFB infections decreases with age (Brødsgaard et al. 1998). It is proposed that the larvae differed in age by approximately 12 hours, causing the increased susceptibility shown on plate 2. Further testing with an increased number of replicates would be able to confirm this, along with more accurate identification of larval age. Unfortunately this was not possible within the time frame of this project.

Whilst production of viable spores is suggested as the reasoning behind a difference in virulence between isolates 6993 and 6254 / 6678. This does not explain the lower levels of virulence displayed by strain 6264, this isolate was capable of producing similar numbers of viable spores as strain 6993 but did not display increased virulence. It is therefore proposed that 3 phenotypes have been identified:

1. Pathogenic / poor sporulation (ERIC I – Strains 6254 and 6678)
2. Pathogenic / good sporulation (ERIC IV - Strain 6993)
3. Non-pathogenic / good sporulation (ERIC IV - Strain 6264)

Whilst it is possible to suggest the difference between strains 6254 / 6678 and 6993 is a result of differing levels of viable spores. It is not possible to suggest a reason for the difference in virulence displayed by strains 6993 and 6264. A different approach is required to better understand the differences between the phenotypes identified.

It has been shown that increased virulence observed in an in-vitro environment, does not result in increased virulence within a hive. This suggests that whilst it was shown 6993 was capable of killing larvae within 4 days; within a hive environment this would not result in a systemic infection. Rather a localised infection of just a few cells, this localised infection would occur before the cells are capped. Resulting in the triggering of increased hygienic behavior by worker bees, the infected larvae would be removed from the hive prior to the disease forming the typical AFB scale. This would result in no disease transmission within the hive. The natural affect of an ERIC type IV infection on a colony of bees is unknown, whilst it is unlikely that these infections would be noticed and identified as AFB. If a significant proportion of the larvae were fed infected food, this would result in the death of many new larvae. Depending on the magnitude of this loss, the colony may be weakened and more susceptible to other infections.

It is suggested that whilst proteolytic enzymes are likely responsible for the typical symptoms observed during an AFB infection, they are not essential for infection and causing larval mortality. This agrees with the lack of AFB symptoms observed in this experiment and previously reported. It is therefore suggested that further studies in to the pathogenicity of *P. larvae* do not focus solely on the involvement of proteolytic enzymes.

4.8. Acknowledgements

I would like to thank Selwyn Wilkins, Ben Jones, Damian Cierniak and Jack Wilford for their expert help with setting up the *in-vitro* exposure assay, and for obtaining brood comb for me containing larvae of the correct age. I would also like to thank James Chisholm, who had previous experience with the *in-vitro* rearing method and 'showed me the ropes'.

Chapter 5

Comparative genomics of *Paenibacillus larvae* isolates

5.1. Summary

- Large amounts of genomic data collected from 4 *P. larvae* isolates belonging to ERIC types I and IV.
- Type strain 6254 formed a better primary assembly, than seen in previous attempts to sequence the *P. larvae* genome.
- No assembled contigs are believed to be from plasmids, all assembled data is assumed to be *P. larvae* chromosomal data.
- 89 out of 120 *B. subtilis* sporulation genes were found to have orthologs existing within *P. larvae*. 16 sporulation genes contained residue substitutions. The genotypes shown reflected the sporulation phenotypes previously described (chapter 4).
- The 16 sporulation genes have been linked to the initiation of sporulation or spore coat formation. So likely that these mutations allow ERIC type IV isolates to sporulate in laboratory media.
- Unknown what genetic difference exists between strains 6993 and 6264, responsible for the differences in virulence observed.

5.2. Introduction

It is known that a wide variety of *Paenibacillus larvae* strains exist, it is possible to group these strains using phylogenetic analysis or ERIC fingerprinting. However, it remains unknown what these groupings reflect. Whilst it was previously believed that the groupings reflected the pathogenicity of *P. larvae*, this was shown not to be the case (Genersch et al. 2006).

It has been shown that significant phenotypic differences exist between the strains grouped as ERIC types I and IV (Chapter 4). *In-vitro* exposure to these strains revealed an ERIC IV isolate that was capable of causing larval mortality at a greatly increased rate over other isolates. It was believed that a 'hyper-virulent' isolate had been identified; further investigation revealed that differences existed in the number of spores produced by isolates belonging to ERIC types I and IV. When the spore dose of the 'hyper-virulent' isolate was reduced to a comparable level to that of an ERIC type I isolate, the level of larval mortality observed was greatly reduced (chapter 4). It was concluded that the difference in sporulation was likely responsible for the differing levels of *in-vitro* pathogenicity.

It was determined that a whole genome shotgun-sequencing approach was the best methodology to explore the genetic differences existing between isolates identified as ERIC types I and IV. This method is not target to specific regions of DNA like the use of 16S rRNA and MLST sequencing (Chapter 2). Intra-species differences had been identified previously in chapters 2 and 3, but the relevance of these groupings in relation to the virulence of the organism was unknown. The genetic basis of differences in sporulation shown to exist in chapter 4 were unknown, therefore the untargeted sequencing approach was selected as the most appropriate method for the study of this difference.

5.2.1. Use of next-generation sequencing.

During the course of this project pyrosequencing became a more affordable and widely available technology. Advancements were also made increasing the length of the sequence reads produced by this method. A single ten-hour run can produce up to 600 megabases of genomic data, allowing for large coverage of bacterial genomes in a single run. Completion of a bacterial genome remained unlikely even at maximum coverage. Producing a complete *P. larvae* genome is the ultimate goal from use of this technology, however the time and money required was likely to be outside the remit of this project.

The use of next-generation sequencing to uncover the genetic basis of phenotypic differences has not been widely explored. The majority of sequencing projects apply bioinformatic analysis of the data produced to 'suggest' possible genetic differences that may cause a phenotypic difference. This methodology has been applied to human pathogens, where the phenotypic characteristics (the disease caused) have been previously documented (Stabler et al. 2009). Due to the lack of previously sequenced isolates, identification of mutations based on test subject-reference comparisons cannot confirm the correlation existing between genotypes and phenotypes (Hall 2007). With only one partially sequenced genome (at time of planning), any genome comparisons would require genetic information from multiple isolates. With the production of large quantities of genomic data, there is a possibility to 'get lost' mining for differences existing between isolates (Vinatzer and Yan 2007). We aim to avoid this by assessing genetic differences based upon observed phenotypic characteristics, as the relevance of other genetic differences would be unknown in relation to the virulence of the organism.

5.2.2. Previous Attempts to sequence the genome

Despite several genome sequencing projects, the first of them commencing in 2005, the amount of genomic data available for *P. larvae* is very limited. This data does not seem to represent the entire genome and the level of coverage at the basic assembly stage is relatively low. The only *P. larvae* genome data available during analysis was a series of 646 contigs assembled for Sanger reads in 2006 (Qin et al. 2006), this project was aborted shortly after this first assembly.

A genome project recently concluded in British Columbia, Canada (Chan et al. 2011). Data from this project was made public in February 2011. This provided a more recent set of contigs on GenBank (Accession: ADZY00000000), however as this data was still incomplete, it was decided to use the original BCM contigs as a point of reference.

5.2.3. Virulence Plasmids

Plasmids are often associated with the virulence of closely related *Bacillus sp.* (Van der Auwera et al. 2005; Wang et al. 2008). Therefore whilst *P. larvae* virulence has never been proven to be plasmid mediated, the potential for a plasmid encoding virulence determinants has always remained a possibility. A complete *P. larvae* plasmid genome has been produced, however no plasmids have been recovered from UK isolates (data not shown). It is thought that the plasmid encodes for Oxytetracycline (OTC) resistance (Alippi et al. 2007; Murray et al. 2007), AFB infections in the UK are not treated with OTC, and so prevalence of the plasmid would offer no genetic advantage to the bacterium.

Plasmids, unlike genetic mutations that result in the alteration of a single nucleotide, involve the incorporation of large amounts of DNA from a different source. It has been shown that bacteria with a lower GC content contain plasmids with a more diverse GC content (van Passel et al. 2006). A large number of bacterial species were found to contain plasmids of a lower GC ratio than that of chromosomal DNA (Bohlin et al. 2008), so identification of plasmid DNA from 454 sequence data should be relatively simple.

5.2.4. Control of Sporulation: Genetic or Environmental?

Due to the vast differences in the number of viable spores produced, it was predicted that differences would also exist in the genes controlling sporulation. Previously it has been shown that sporulation mutant strains, lacking vital genes often lack the ability to produce spores (Brehm et al. 1973; Fortnagel and Freese 1968; Hoch 1970; Klofat et al. 1969). It is highly likely that significant changes existing within sporulation genes will have a significant effect on the ability of those isolates to produce viable spores. Whilst the regulation of sporulation within *P. larvae* has not been widely studied, it has been extensively studied in *Bacillus subtilis* (Catalano et al. 2001; Errington 1993; Ozin et al. 2001; Takamatsu et al. 2000).

It remains likely that the regulation of sporulation in *P. larvae* is also influenced by differing environmental conditions (Sonenshein 2000). It is believed that temperature, oxygen concentration, population density, by-products of metabolism and pH are all interrelated in the control of sporulation within many *Bacillus sp.* (Knaysi 1945). It is predicted that any genetic differences observed in sporulation genes, will likely be responsible for the activation of the sporulation process.

It is known that ERIC type I isolates are responsible for the majority of AFB outbreaks (Genersch 2010a). During the latter stages of the infection, the larval remains dry and adhere to the bottom of the cell as an AFB scale (Alippi et al. 2002; Dancer and Chantawannakul 1997). Previous study of AFB scales has revealed up to millions of spores (Ashiralieva and Genersch 2006; Dancer and Chantawannakul 1997; Genersch et al. 2006), so it is logical to assume that under the right conditions ERIC I isolates are able to produce large quantities of viable spores. Phylogenetic analysis shows a close relationship exists between *P. larvae* and *B. subtilis* (Chapter 1), so it is likely that the regulation of sporulation is under the control of orthologous genes.

5.3. Aims

- To produce a large amount of *P. larvae* genomic sequence data from four isolates.
- To assemble the data in to larger contiguous sequences
- To confirm the absence of plasmid DNA from the sequenced contigs
- To use comparative analysis of the data produced to highlight differences in spore genes existing between isolates 6254 / 6678 and 6993 / 6264.

5.4. Materials and Methods

5.4.1. Strain growth and DNA extraction

A large quantity of DNA is required for the pyrosequencing process. Rather than utilising a different extraction method, the starting quantity of culture was increased to a volume which would provide a calculated excess of DNA. Fresh cultures were taken from the frozen PROTECT storage beads, these were first grown on BHI-T agar to ensure the correct colony morphology was observed. 100ml BHI-T broth cultures were inoculated with a single colony picked from the agar. Broth cultures were incubated at 34°C and shaken at 175rpm in the orbital incubator for 72 hours. DNA extraction was performed using the Power Microbial Maxi DNA isolation kit (Mobio cat #12226-25). The concentration of DNA required for 454 sequencing was a total volume of 50µl at a concentration of 100ng/µl. All extracted DNA contained over 1200ng/µl of DNA, these were diluted appropriately.

5.4.2. DNA Library production

Preperation of the 454 libraries was carried out as detailed in the Roche GS FLX Titanium Rapid library kit (cat#05608228001). The isolates were MID tagged for post-run sequence recovery (table 5.1).

MID Tag	Isolate	LMG Reference
1	6254	9820
2	6264	16247
3	6678	16241
4	6993	14427

Table 5.1: Details of the strains sequenced and the MID tag assigned to those strains.

Fragment length was calculated using the Agilent Bioanalyzer 2100, confirming that fragments of a suitable length had been produced during the nebulizing stage. A custom QPCR assay (developed by the University of York) was used to calculate the concentration of library required for use during the emPCR stage.

5.4.3. Sequence generation

The pico titre plate was loaded at FERA and run on the pyrosequencer on site. This procedure followed all manufacturers guidelines; no alterations were made to this process.

5.4.4. Genome sequence assembly

Sequence assembly was performed using the NEWBLER assembly software (Roche). A variety of other sequence assembly methods were trialed, in this instance the Newbler assembly produced the best output. Most assembly statistics were provided directly by the newbler output files, coverage was calculated as follows (total number of bases for strain/assembled length).

5.4.5. GC content analysis

GC content analysis was performed using the InfoSeq feature from the EMBOSS software package (European Bioinformatics Institute (EBI), http://www.ebi.ac.uk/Tools/emboss/). This was run via the eBioX software package (Erik Lagercrantz, http://www.ebioinformatics.org/ebiox/). This package produced an output file containing contig names and GC content values, these were used for further analysis.

5.4.6. BLAST Analysis

All contigs were imported in to a single Fasta file, utilising a pre-written perl script these were translated into all 6 possible reading frames. Standalone BLAST+ executables (Altschul et al. 1990) (ftp://ftp.ncbi.nlm.nih.gov/blast/executables/blast+/LATEST/) was used to perform the BLAST analysis. A BLAST library was constructed using the translated contig sequences, a search was performed for 120 sporulation genes found in *B. subtilis* (online database, http://www.rhul.ac.uk/biological-sciences/AcademicStaff/Cutting/Spo/spo2.html).

Bioperl (http://www.bioperl.org) was used to extract data from the resulting output file. Hits were selected to be manually checked if they had an E-value

below 0.0001 and query coverage >70%. These BLAST matches were manually checked to ensure suitability.

5.4.7. Sequence alignment details

A concatenated sequence was produced from 23 partial gene sequences. Details of the residue differences in the 2983 Amino Acid sequence, were extracted using InfoAlign in the EMBOSS software package. This produced a PDF file documenting the residue differences and highlighted residue substitutions that were similar and more distant.

5.4.8. Phylogenetic Tree building

Tree building was performed using Clustal X (Larkin et al. 2007), the tree was created using the UPGMA clustering algorithm. All default settings were maintained for this analysis.

5.5. Results

5.5.1. Genome Sequence Assembly

Isolate	6254	6264	6678	6993
Total Number of Bases	175698956	45100012	46311468	43529897
Total Number of Reads	591394	149661	158881	149553
Assembled Length	4638858	4307168	4230245	4382896
Coverage based on read depth	37.88	10.47	10.95	9.93
Peak Depth	38	9	10	8
Estimated genome size	4.6 MB	5.0 MB	4.6 MB	8.0 MB
Aligned Reads (%)	98.14	97.3	96.36	97.08
Inferred Read Error (%)	0.87	0.87	0.89	0.91
Number of Contigs	199	455	384	600
Number of Contigs >1.5kb	134	292	251	353
Average Contig Length (bp)	23310.84	9466.3	11016.26	7304.83
Minimum Contig Length (bp)	134	101	102	103
Maximum Contig Length (bp)	239719	79204	141267	169517
Average Number of reads/contig	2984.39	334.59	412.24	253.27
Average Mismatch of Contigs (%)	0.83	0.95	0.97	1.1

Table 5.2: Assembly statistics highlighting the number of contigs produced, the length of those contigs and the number of reads per contig. Details of the errors within the assembly process are also shown.

Over 200 megabases of DNA were produced from a single titanium pico titre plate. The collected sequence data was separated by MID tags and assembled in to contigs (table 5.2). Strain 6254 that filled a half plate produced the most sequence data. Due to the larger amounts of data produced this was assembled in to 199 contigs. The other three isolates were allocated a third of the remaining half of the plate. Strain 6678 produced the most sequence data from the remaining isolates, with 6264 producing approximately 12 megabases less of DNA and strain 6993 produced 15 megabases less of DNA. As expected the number of contigs produced increased as the amount of sequence data collected decreased. It was also seen that increasing the amount of genetic data collected increased the number of reads making up each contig.

5.5.2. GC Content Analysis

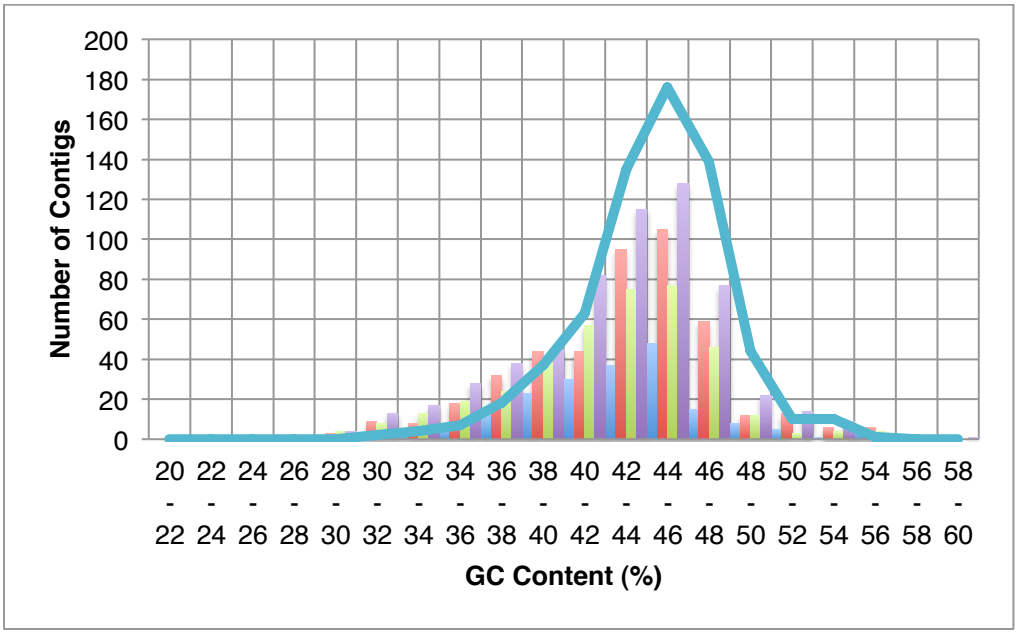


Figure 5.1: A histogram showing the distribution of contigs across a range of % GC content. The line plotted shows the contigs produced by the Baylor College of Medicine, in a previous sequencing attempt. (Blue = 6254, Red = 6264, Green = 6678 and Purple = 6993)

Contigs were analysed for GC content (figure 5.1). It was shown that the GC content peaked at 44-46% for all four of the sequenced isolates. The distribution of contigs produced the expected curve, with a slight bias towards the lower GC

content. Very few contigs were produced with over 48% GC content. The 646 contigs produced by the Baylor College of Medicine were also analysed for GC content and plotted against the distribution data collected. It was found that the distribution of contigs based on GC content, matched the distribution seen from the 646 BCM contigs. BLAST analysis revealed no sequence similarity with *Paenibacillus larvae* plasmid pMA67 (data not shown); this was despite the presence of several contigs of a similar GC content to the plasmid. It was also seen that contigs displaying the most extreme differences in GC content were relatively small sequences with very low read coverage.

5.5.3. Sporulation Gene Copy Number

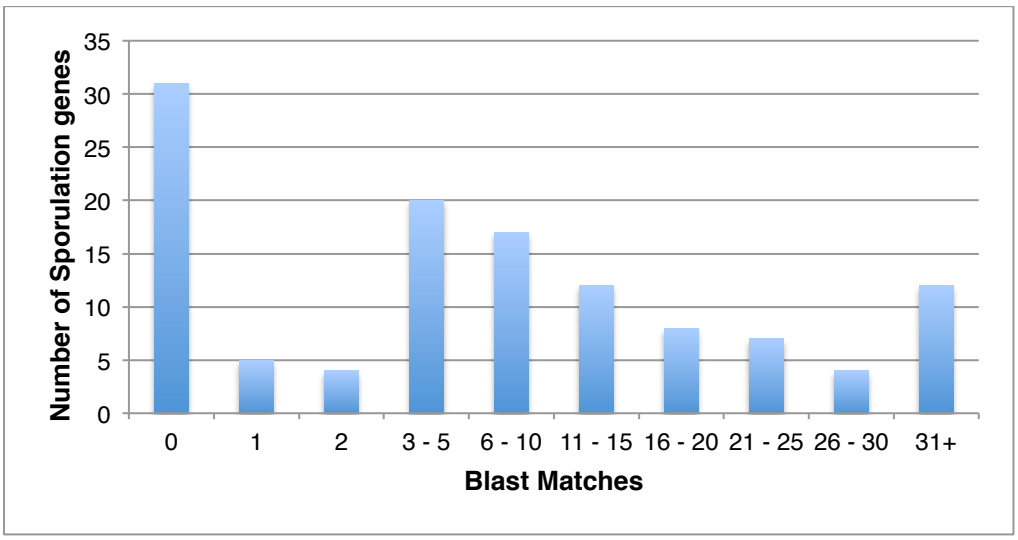


Figure 5.2: The Number of BLAST hits returned for each sporulation gene.

A BLAST search of sporulation genes found in *B. subtilis* revealed that 31 of the gene sequences had no significant BLAST matches in the contigs. The remaining 89 genes did have matches within the sequenced contigs. The number of significant matches found ranged from a single match to over 31 (Figure 5.2). It was observed that matches were partial gene sequences, rather than complete gene matches. Genes that had more BLAST hits returned more partial gene sequences, however these did not assemble to produce complete gene sequences. Only one of the genes searched for produced multiple BLAST hits per contig,

inferring that part of the gene within *P. larvae* is significantly different to that of *B. subtilis*.

5.5.4. Alignment of partial sporulation gene amino acid sequences

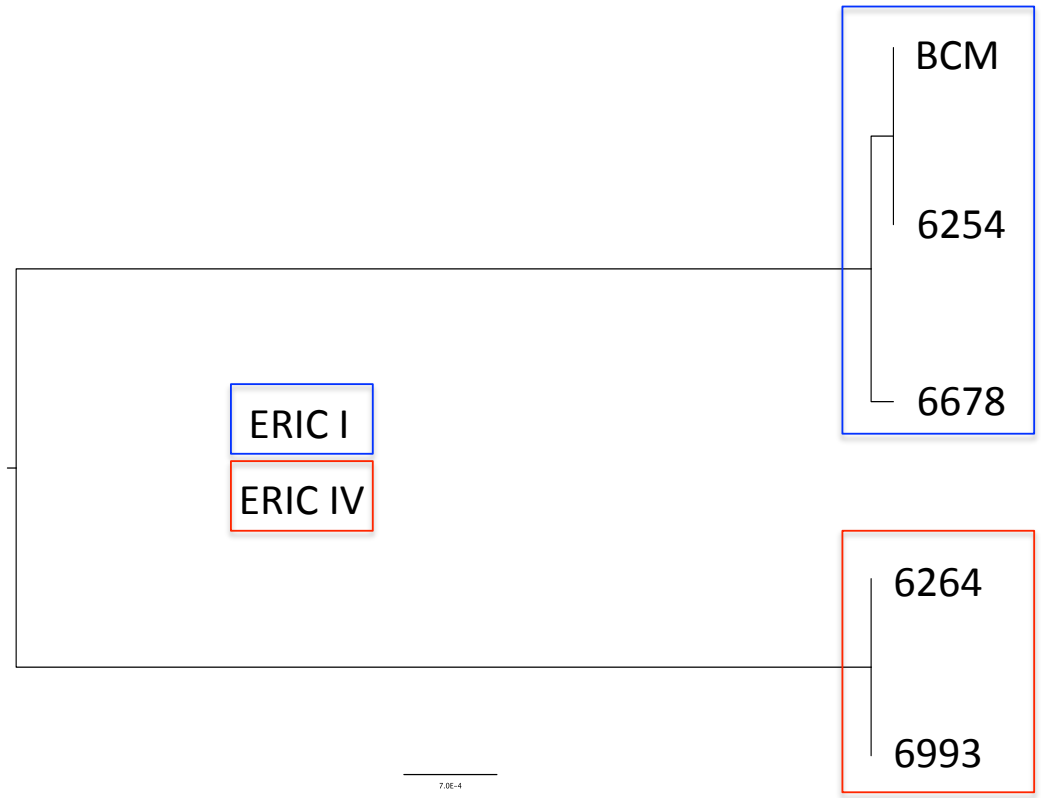


Figure 5.3: UPGMA tree showing the relationship between 23 concatenated partial sporulation gene sequences, in which substitutions were found.

From all of the BLAST hits that met the criteria used whilst parsing the BLAST data in Bioperl (E-value below 0.0001 and query coverage >70%). Partial gene sequences, from 16 genes, were found to contain substitutions or insertions. Other partial sequences were identified which varied greatly, however no homologous sequences were present in sequence data from the other strains. Constructing a tree of the concatenated sequence clearly shows the differences that existed between the ERIC types tested (Figure 5.3). 6264 and 6993 were shown to be monophyletic and BCM and 6254 shown to cluster together. 6678

clustered closest to 6254 and BCM, but grouped separately due to a single residue substitution.

Number of Partial Sequences	23
Total Residues	2983
Number of Substitutions	39
-Similar Residue Substitutions	19
- Less Similar Residue Substitutions	20
Number of Additions	1

Table 5.3: Details of the observed differences existing within the concatenated sporulation gene sequences.

It was observed that 38 residues in the 23 partial gene sequences were different in ERIC type I and IV isolates (Table 5.3). A further 1 substitution existed between strain 6678 and all other isolates tested. ERIC type IV isolates were also found to have an additional residue. Of the 39 substitutions 19 were found to be substitutions for similar residues. The remaining 20 substitutions were less similar residues. The maximum number of substitutions found in a single partial gene sequence was 5 and the minimum was a single substitution.

5.5.5. Genes in which residue substitutions were present

Genes with similar substitutions	*gdh, kinB, kinC, sigE, sigF, sigK, spo0A, spo0F, spoIIAB, spoIIB, spoIVA* and *spsC*
Genes with less similar substitutions	*gdh, kinC, sigA, sigF, sigK, spoIIAB, spoIIB, spoIVB, spoVID, spsC* and *spsD*

Table 5.4: Details of the sporulation genes in which the substitutions were present.

The 23 partial gene sequences came from 16 sporulation genes (Table 5.4). 12 of those sporulation genes were found to contain residues substituted with similar residues. 11 of the genes were shown to contain residues substituted for less similar residues. The less varied partial gene sequences were from *kinB, sigE* and *spo0A*. The most varied partial gene sequences were from *sigK*.

5.5.6. Assessing the number of nucleotide substitutions responsible for the residue differences

Codon Position of Substitution	1st Base	2nd Base	3rd Base	Additions	Overall Sequence similarity to 6254 / BCM
6678	1	0	1	0	99.98
6264 / 6993	15	17	31	3	99.26

Table 5.5: Differences in concatenated nucleotide sequences representing the amino acid sequence described in table 5.3. The position of substitutions within the codon, number of additions, and the percentage similarity to the type strain DNA is shown.

The nucleotide sequences for the 23 partial gene sequences were extracted and combined to create a concatenated nucleotide sequence of 8958bp. The location of the substitutions within codons was calculated; 6264 and 6993 contained mainly 3rd base substitutions with an almost even number of 1st and 2nd base substitutions (Table 5.5). As expected 3 additional nucleotides were present representing the additional amino acid. 6678 contained 2 nucleotide substitutions, a 1st base substitution responsible for the 1 residue difference detected and a silent 3rd base substitution. 6254 and BCM nucleotide sequences were identical as seen in the amino acid sequences. Assuming 6254 / BCM are the type strain of the species, 6678 was found to have 0.02% sequence divergence and 6264 and 6993 had 0.74% sequence divergence.

5.6. Discussion

5.6.1. Assembly of *P. larvae* genome

Data produced from a half 454 pico titre plate produced 591,394 reads that were assembled into 199 contiguous sequences for strain 6254 (Table 5.1). Compared to the previous sequencing attempt by the Baylor College of Medicine which had

produced 646 contiguous sequences from 54,073 reads. All of the isolates sequenced were assembled in to a lesser number of contiguous sequences than the previous sequencing attempt, it was also noted that the size of contigs produced was greatly increased from an average of 11kb to an average of 23kb for strain 6254. Average contig length achieved from a $1/6^{th}$ of a plate remained comparable with the contigs previously produced. The production of these larger contigs highlights the full potential of the massively parallel sequencing method, as shorter 454 sequencing reads were able to be assembled into larger contigs than the longer sanger sequencing reads.

The size of the *P. larvae* genome remains unknown however it is estimated to be around 4.3-5 megabases. Sequencing of *Paenibacillus sp. JDR-2* revealed a genome size of 7.2 megabases (Unpublished, Acession:CP001656), whilst comparative analysis of the sequenced contigs against this complete genome revealed very little sequence similarity (data not shown), the eventual size of the *P. larvae* genome could be greater than 5 megabases. Attempts were made to produce a more complete scaffold from the contigs produced, it was established that the amount of sequence data was not sufficient from individual strains to create this scaffold.

It is proposed that all contigs are combined to produce a scaffold of the genome. Whilst this would not be an accurate representation of a complete genome, with less than 3% difference in transcribed regions analysed within this project. It would allow for better construction of future assemblies with the availability of a pre-existing scaffold. This work remains on going at time of writing, due to the complexity of the process.

5.6.2. GC content analysis

As no complete *P. larvae* genome currently exists, BLAST analysis against the NCBI database cannot accurately confirm that all sequence data is chromosomal *P. larvae* DNA. BLAST analysis revealed high sequence similarity with pre-existing BCM contigs, suggesting that sequenced DNA did belong to *P. larvae*. Plasmid mediated virulence has been reported in many closely related bacterial

species such as *Bacillus cereus*, *Bacillus anthracis* and *Bacillus thuringiensis* (Andrup et al. 2008; Pannucci et al. 2002; Van der Auwera and Mahillon 2008; Van der Auwera et al. 2005; Van der Auwera et al. 2008; Wang et al. 2008). Plasmids have been identified as existing within the *Paenibacillus* genus, in *P polmyxia* (Ma et al. 2011) and also in *P. larvae* (Alippi et al. 2007; Benada et al. 1988; Murray et al. 2007). To date plasmids recovered from *P. larvae* have only been linked to tetracycline resistance, not the virulence of the organism. The *Paenibacillus* genus like *Bacillus* has a relatively low GC content (Qin et al. 2006), analysis of the data produced can identify sequences belonging to plasmids or transposable elements within the genome. Lab based methods had failed to extract any plasmids from the strains tested within this project (data not shown), it was therefore not expected that plasmid DNA would be identified within the data produced. Whilst contiguous sequences were identified which were of a similar GC content to *Paenibacillus larvae* plasmid pMA67 (36%), rather than the expected ~44%, when the data was plotted against the existing 646 BCM contigs it was apparent that the sequence data produced was chromosomal rather than that of a plasmid (Figure 5.1). Further BLAST analysis revealed no matches in any contigs for the pMA67 sequence (data not shown). Plasmids found within this species are more likely to be community acquired resistance plasmids, rather than plasmids essential for the survival of the organism.

5.6.3. Copy number of sporulation genes

It is known that within bacterial genomes genes exist with a variety of copy numbers. BLAST analysis of the contigs revealed a range of hits from 0 – >50 (Figure 5.2). Whilst it should be acknowledged that the sequence data produced by these methods is a random process, when considering the probability of multiple BLAST hits across several contigs compared to the a single hit from each of the four strains, it is possible to suggest that those genes producing a larger number of BLAST hits exist in larger copy numbers than those with fewer hits. It is likely that those genes seemingly present in higher copy numbers are more important for the survival of the bacterium. Using this logic it is possible to

suggest that changes to these genes with higher copy numbers are more likely to affect the phenotype of the organism.

5.6.4. BLAST analysis for orthologous sporulation genes

It is believed that genetic differences rather than environmental conditions are responsible for the difference in production of viable spores. It has been noted that the control of sporulation is the result of several genes within *B. subtilis*, logically it can be assumed that amino acid substitutions within these genes will likely affect the initiation of sporulation. In *B. subtilis* at least 120 genes have been associated with sporulation many associated with the control of sporulation (Chapter 5.4.6). A BLAST search for orthologous sequences within the *P. larvae* contigs revealed that significant BLAST hits were present in 89 of those genes. 23 partial genes sequences were present in all 4 sequenced strains and the BCM contigs. These partial sequences belonged to 16 sporulation genes (Table 5.5) and were found to contain a total of 39 substitutions and 1 addition (Table 5.4). Phylogenetic analysis of the 2983 residue concatenated sequence confirmed that a clear differentiation could be made between ERIC I isolates and ERIC IV isolates (figure 5.3). The BCM sequence data was shown to cluster with the ERIC type I isolates, knowing that this isolate came from a diseased hive in the USA (pers comm. Jay Evans, USDA), it is likely to predict that this strain of *P. larvae* belongs to ERIC type I. A single substitution was responsible for strain 6678 grouping on a separate branch to BCM and MID1 strains. It was observed that genes returning more BLAST hits contained more substitutions; whilst it would be acceptable to suggest that the increase in frequency of hits was likely to return sequences with additional substitutions. Comparison of the concatenated nucleotide sequences showed that levels of sequence diversity were comparable to those detected in house keeping genes (chapter 2). Whilst the relationship between ERIC I and IV isolates has been shown here, it is unknown where ERIC types II and III would be located. These types were not found to produce large quantities of viable spores (data not shown), so it would be expected that these isolates would be closer related to ERIC type I isolates based upon sporulation genes. However this was not seen in purH, PyrE or the 16S rRNA phylogeny

where ERIC types II and III were shown to be closer related to ERIC type IV isolates.

Whilst analysis of small partial gene sequences cannot predict a change in gene function, the number of sporulation genes existing with residue substitutions, leads to the conclusion that a genetic difference must be at least partially responsible for the difference in sporulation observed. It is likely that additional genomic data would give rise to the discovery of more substitutions occurring in a greater number of sporulation linked genes. Whilst it is known that differences exist in the number of viable spores produced by isolates belonging to ERIC types I and IV, it is known that within a hive environment isolates confirmed as ERIC type I are able to produce large numbers of viable spores. This suggests that the differences observed within the sporulation genes are likely linked to the expression of other genes or the control of sporulation.

Whilst the genetic differences observed hold the potential to explain the phenotypic differences observed between ERIC I and IV isolates. It is still not possible to explain the observed difference displayed between strains 6264 and 6993. Both isolates were identified as ERIC IV isolates, whilst both isolates were capable of producing large number of viable spores and contained the same genetic differences within the sporulation genes. The level of virulence observed from isolate 6993 was not displayed by isolate 6264 despite a potentially higher initial spore dose (data not shown).

It was observed that not all isolates producing large quantities of viable spores, exhibited increased levels of *in-vitro* pathogenicity. Strain 6264 was found to produce equal or greater numbers of viable spores than strain 6993 (data not shown), however this isolate failed to deviate significantly from the control larval mortality. It is therefore not possible to suggest that sporulation alone is responsible for the elevated levels of pathogenicity; rather that it is one of several factors involved in the pathogenicity of the isolate. Further genomic analysis may reveal differences in existence between strain 6264 and strains 6254, 6678 and 6993. It would however be useful to identify a target gene family that can be linked to the pathogenicity of *P. larvae*. Without this it is likely that many

genomic differences would be observed, but the true relevance of these differences would be unknown.

5.6.5. Are the observed genetic mutations responsible for phenotypic differences?

Sporulation in *P. larvae* has not been studied, genes essential for this process are not known. Due to the close taxonomic relationship with *B. subtilis* it was assumed that orthologous genes would be present within the *P. larvae* genome. 74% of the 120 *B. subtilis* sporulation genes were shown by BLAST analysis to have orthologs within the *P. larvae* sequence data. The function of these genes within *B. subtilis* has been well documented. *spo0A* is considered to be the primary transcriptional activator and both *spo0A* and *spo0F* are response regulators of the sporulation process (Hoch 1993). *kinB* and *kinC* have been identified as essential kinases required for sporulation (Jiang et al. 2000). *spoIIA* is an essential sporulation loci containing amongst others *spoIIAB* (Burbulys and Trach, 1991). Spore coat formation has been linked to *spoVID* and *spoIVA* genes, making these genes essential for the production of viable endospores (Ozin et al. 2001). All of these genes were found to have residue substitutions in MID 2 and 4. With 51% of the amino acid substitutions considered to be non-similar residues, it is likely that these differences are at least partially responsible for the difference in viable spores produced. The most substitutions were found to exist in 3 partial sequences belonging to the *sigK* sigma factor, this has been shown to be essential for spore production in *Bacillus thuringiensis* (Bravo et al. 1996).

The genes identified to contain residue substitutions were found to be genes responsible for the control of sporulation. Many of these residue substitutions occur in genes responsible for the initiation of the sporulation process, or responsible for the forming of viable spores. It is not believed that increased sporulation is a result of these mutations, rather that these mutations alter the environmental conditions in which *P. larvae* isolates are able to sporulate. It has already been shown that ERIC type IV isolates are not thiamine deficient, whereas ERIC type I isolates required thiamine to grow on laboratory media (Chapter 1.3.2). It is known that ERIC type I infections such as MID1 are able to

produce AFB scales containing >100,000 spores, so it is likely that laboratory based techniques are not activating the sporulation pathway due to environmental conditions. Environmental factors has previously been shown to affect sporulation in *Bacillus mycoides* (Knaysi 1945) and *Bacillus subtilis* (Msadek 1999). Further more sporulation is ultimately known to be a survival mechanism for bacteria, utilised for survival when conditions become unfavourable. Further expressional studies of sporulation in *P. larvae* would likely reveal the basis behind the difference in sporulation exhibited between 6993 and 6254/6678.

5.6.6. Can strain types be identified?

Based upon phenotypic and genotypic analysis of isolates belonging to ERIC types I and IV. It is proposed that 3 strains types are in existence (figure 5.4), Highly Pathogenic strains (6254 and 6678?), Strains producing many viable spores (6264) and Highly Pathogenic strains capable of producing many viable spores (6993). Additional work is required to confirm these findings, not least the addition of strains belonging to ERIC groups II and III. This has the potential of intra-species discrimination based upon observed phenotypic traits.

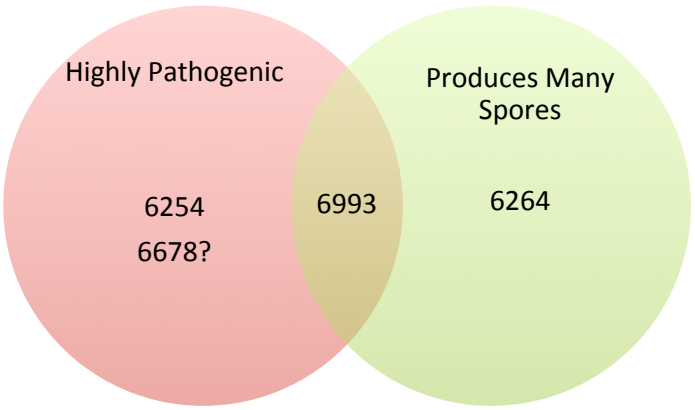


Figure 5.4: Possible strain types present within ERIC types I and IV. Showing the observed phenotypic differences.

5.7. Conclusions

The primary assembly of the *P. larvae* genome for strain 6254 (MID 1) has produced the lowest number of contiguous sequences, whilst maintaining coverage of approximately 4.3Mb. Upon a conservative estimate of a 5Mb genome this represents assembled coverage of x0.93, the average contig length is also >10kb longer than previous assemblies of the genome. It remains unlikely that ½ a sequencing plate has provided enough genomic sequence to completely assemble the *P. larvae* genome. Currently the genomic data from all isolates is being assembled to produce a more complete scaffold of the genome, whilst not a true representation of a single *P. larvae* genome, the scaffold can be used as a reference for future assemblies.

Laboratory based methodologies had shown that plasmids could not be isolated from any of the strains of *P. larvae* tested (data not shown). Shotgun sequencing methods often highlight the presence of plasmids or transposable elements as they form contigs differing in GC content to that of the genome. Whilst contigs were identified with GC content expected from plasmid sequences, upon mapping the distribution of contigs based on GC content no difference was identified compared to the previous sequencing attempt by the Baylor College of Medicine. BLAST analysis revealed no pMA67 (*P. larvae* plasmid) homologous sequences existed within the contigs. The combination of laboratory based methodologies and the analysis of the genomic data produced suggests that virulence is not plasmid mediated, and any plasmids are likely community acquired antibiotic resistance plasmids.

Following the identification of 3 phenotypes of *P. larvae* belonging to ERIC types I and IV, it has been shown that genetic differences exist within sporulation genes. Phylogenetic analysis based upon 23 partial sporulation gene sequences, revealed that the observed difference in sporulation between ERIC type IV and ERIC type I isolates could be identified based upon amino acid substitutions within the sporulation genes. Three different sequences were identified, strain

6254/BCM, 6678 and 6993/6264. 39 substitutions and 1 addition were identified between strains 6254/BCM and 6993/6264. Strain 6678 differed from strain 6254/BCM by 1 amino acid. It is likely that a larger sample size would reveal single amino acid substitutions existed amongst ERIC type I isolates. It is unknown where ERIC type II and III isolates would sit within this phylogeny, based upon MLST sequencing or 16S rRNA phylogenies (Chapter 2) these isolates were shown to be closer related to ERIC type IV isolates. Sporulation testing revealed that these isolates did not form large quantities of viable spores under laboratory conditions, so it is predicted that a phylogeny based upon sporulation genes from these isolates would identify them as closer relatives to ERIC type I isolates. Limited funding available for this project dictated the number of isolates that could be sequenced, whilst many hypotheses remain un-answered, strains were carefully chosen to best answer the most pertinent questions.

Examination of the nucleotide sequences from which these partial gene sequences are coded, revealed high levels of sequence similarity. Less than 1% sequence divergence existed between ERIC type I and ERIC type IV isolates, similar to the levels of divergence seen in housekeeping genes whilst developing an MLST scheme (Chapter 2). These genes again have been shown to have less sequence diversity than the 16S rRNA subunit, which was unexpected. However with only 5 strains examined, it is likely that more diverse strains do exist and increasing the sample size would highlight this. Unlike the MLST and 16S rRNA phylogenies the grouping of isolates here, whilst based upon genetic data, does represent an observed phenotypic difference in production of viable spores.

Whilst the observed differences in pathogenicity between strain 6678/6254 and 6993 have been partially explained due to the vast difference in the number of spores produced (chapter 4). It remains unclear why strain 6264 displays low levels of pathogenicity, despite producing more viable spores than strain 6993 (data not shown). It is believed that 3 strain types have been identified within ERIC types I and IV (Figure 5.4), these strain types consist of isolates capable of causing disease but lacking the ability to produce sufficient numbers of spores within a lab, strains capable of producing sufficient numbers of spores within a

lab and strains combining both of the phenotypic traits. However, further comparative analysis combined with phenotypic studies of the isolates may reveal new phenotypic and genotypic differences that exist between *P. larvae* isolates.

Completion of a *P. larvae* genome will highlight complete gene sequences; this study suggests that genes relating to sporulation will show intra-species genetic differences based upon the observed phenotypic differences. Based upon the large number of BLAST hits and the significant proportion of residue substitutions found, it is proposed that *sigK spo0A* and *kinB* would be a reasonable starting points for a more detailed study into the sporulation differences between isolates. These genes have been shown to be essential for the initiation of sporulation or production of viable spores in *Bacillus sp.* (Burbulys and Trach, 1991; Jiang et al. 2000; Ozin et al. 2001).

Further data mining should be performed on these existing *P. larvae* contigs, potentially revealing the difference in pathogenicity observed between strains 6993 and 6264. However this experiment aimed to show that observed sporulation phenotypes, had a genetic basis. The complete understanding of *P. larvae* pathogenicity remains unknown; virulence factors responsible for this infection have yet to be discovered. So differences in existence between strains 6993 and 6264 could only be suggested as a potential reason for the observed difference in virulence. Further analysis of a complete *P. larvae* genome should highlight potential virulence factors, combining this information with phenotypic analysis of the strains would reveal the relevance in relation to the virulence of the organism.

5.8. Acknowledgements

I would like to thank Celina Whalley for preparing the DNA library for sequencing, and Naveed Aziz for additional help troubleshooting the concentration of DNA required. Thank you to Ummey Hany for loading the pico titre plate and running the sequencer. Finally a huge thank you to Rachel Glover

for being my ‘Perl and Bio-perl Guru’ and for assistance on performing the bioinformatical analysis.

Chapter 6
General Discussion

6.1. Understanding intra-species differentiation of *P. larvae*

At the point of original isolation *P. larvae* was believed to exist as two separate species. *P.* (Formerly *Bacillus*) *larvae* was shown to be the causative agent responsible for a lethal brood disease, American Foulbrood (AFB) (White 1906). *P.* (Formerly *Bacillus*) *pulvifaciens* was believed to be the causative agent responsible for a relatively benign brood disease, Powdery Scale (Katznelson 1950; Nakamura 1984). AFB was commonly identified in honeybee colonies, and *P. larvae* was successfully isolated, only a handful of occurrences of powdery scale are reported and as a result <10 isolates are known to exist.

The two species were closer related than originally anticipated, and as such they were taxonomically grouped as *P. larvae* subspecies *larvae* and subspecies *pulvifaciens* (Heyndrickx et al. 1996). At this time it was still believed that all instances of AFB were caused by isolates identified as subspecies *larvae*, subspecies *pulvifaciens* listed as the causative agent responsible for powdery scale disease despite the low incident rate.

In-vitro pathogenicity studies of both subspecies revealed that subspecies larvae was capable of causing AFB as previously identified, but also that many strains of subspecies *pulvifaciens* were capable of causing AFB (Ashiralieva and Genersch 2006; Genersch et al. 2006). In some instances the resultant infection was deemed to be more virulent than a subspecies *larvae* infection. As a result the subspecies differentiation was dropped at this time (Genersch et al. 2006).

6.1.1. Differences within the species

The true relationship between isolates of *P. larvae* is still not understood. Whilst the reclassification based upon virulence data was widely accepted by the research community in 2007, differences still existed between isolates. It is known that ERIC fingerprinting produces 4 distinct profiles, a range of phenotypes exist within these ERIC groupings (Table 6.1).

	ERIC type I	ERIC type II	ERIC type III	ERIC type IV
Subspecies[1]	*larvae*	Both	*pulvifaciens*	*pulvifaciens*
Thiamine deficient[2]	yes	some	no	no
Average time taken to kill 100% of larvae[1]	12 Days	7 Days	7 Days	7 Days
Capable of systemic colony infection[1]	yes	yes	no	no

Table 6.1: Summary of the phenotypic differences known to exist between isolates and how these relate to the subspecies differentiation and ERIC types. [1] (Genersch et al. 2006) [2] (Heyndrickx et al. 1996).

So whilst it had been shown that all tested isolates of *P. larvae* held the potential to cause an AFB infection on a larval level (Ashiralieva and Genersch 2006; Genersch et al. 2005; Genersch et al. 2006; Yue et al. 2008), the ability for isolates belonging to ERIC type III and IV to cause a systemic colony infection has never been documented. Some discrepancy also exists around ERIC type II isolates, as examples exist here from both of the former subspecies. This suggested that whilst the dropping of the subspecies differentiation based upon virulence data was likely the best course of action, the correlation between observed intra-species differences and virulence was still not understood.

6.2. New findings and suggestions in this thesis

The results in this thesis agree with the dropping of the subspecies differentiation. The data also highlights significant differences that exist between isolates of *P. larvae*, which have not yet been explored. The findings of this thesis are documented below in relation to increasing the understanding of the intra-species differences that exist within *P. larvae*.

6.2.1. Genetic differences exist amongst *P. larvae* isolates

Identification of isolates as *P. larvae* was best performed using 16S rRNA sequencing. Commonly used methods such as biochemical analysis and colony morphology, were not suitable for the identification of *P. larvae* due to the often-ambiguous results achieved (chapter 1). Fatty Acid Profiling was commonly used at FERA for the identification of unknown plant pathogenic bacteria; this method was only able to identify *P. larvae* isolates as belonging to the *Paenibacillus* genus. The sequencing of the 16S rRNA subunit has been shown to be a suitable identification method for use when conventional methods fail to correctly discriminate (Petti et al. 2005). With a minimum 16S rRNA sequence identity of 97.5%, sequencing of this subunit was accurately able to identify isolates to the species level. This was a higher level of sequence divergence than previous reported (Genersch et al. 2006), the true relevance of this higher than expected sequence divergence is unknown as this was only displayed on a pairwise basis between 4 isolates.

The 16S rRNA phylogenetic relationship between *P. larvae* isolates was explored, as it had previously been shown that whilst both subspecies grouped closely, they existed on separate branches (Alippi et al. 2002; Shida et al. 1997a; Shida et al. 1997b). Whilst later research highlighted that isolates identified, as subspecies *pulvifaciens* were more phenotypically diverse than isolates identified as subspecies *larvae*, no phylogeny was provided to assess the phylogenetic relationship (Genersch et al. 2006). The results in this thesis (chapter 2) showed

that within a 16S rRNA phylogeny isolates grouped according to their original subspecies differentiation.

As it has been suggested that 16S rRNA similarity of over 97% can equate to overall genome similarity of just 60% (Stackebrandt and Goebel 1994), it was decided to explore the possibility of using housekeeping genes to more accurately understand the intra-species differences that existed. Multi Locus Sequence Typing had been successfully used to distinguish isolates of *Bacillus cereus, anthracis and thuringiensis* (Helgason et al. 2004). Whilst they are differentiated as different species, the level of sequence similarity existing between them is very high (Helgason et al. 2000). Two housekeeping genes were initially sequenced, *purH* and *PyrE*. Unexpected was the high levels of average sequence identity that existed between isolates, in both instances being greater than that of 16S rRNA. Once again isolates grouped upon phylogenetic analysis based upon previous subspecies differentiation.

In all phylogenies strains 6260 and 6261 were shown to share higher sequence similarity with subspecies *pulvifaciens*, despite being previously identified as subspecies *pulvifaciens*. Within the 16S rRNA phylogeny strain 6255 was shown to group incorrectly, strain 6264 was grouped incorrectly on the *purH* phylogeny and strain 6993 was incorrectly grouped on both *purH* and *PyrE* phylogenies. It is suggested that strains 6260 and 6261 were likely incorrectly identified to subspecies level at time of isolation, this has previously been documented as a challenging process (Forsgren et al. 2008). Due to the inconsistency amongst the collective phylogenies, it is suggested that strains 6255, 6264, 6993 all represent anomalies within the data set and were likely correctly identified at time of isolation.

In summary the phylogenetic analysis of 16S rRNA and 2 housekeeping genes agreed with the dropped subspecies differentiation. With the high levels of sequence similarity, it is not suggested that the taxonomy of *P. larvae* is incorrect. It is believed that significant differences exist between isolates and that these agree with the previous subspecies differentiation. It is therefore suggested that sequencing of 16S rRNA or housekeeping genes cannot be used to

determine the virulence of isolates, as the differences that exist within the species are not yet understood.

6.2.2. ERIC types display congruence with the previously observed genetic differences

ERIC repeats were originally discovered within members of the *Enterobacteriacae* (Hulton et al. 1991). Despite many suggestions that these repeats do not exist in all species of bacteria (Gillings and Holley 1997), it has been shown that primers designed to this region can be used for fingerprinting studies in most bacteria (De Bruijn 1992). Lower annealing temperatures are thought to be responsible, allowing priming of less homologous regions (Gillings and Holley 1997). Adjusting the annealing temperature used within this thesis, agreed with this. At annealing temperatures close to the melting point of the primers, fewer bands were visible (data not shown). Annealing temperatures were intentionally kept low to produce similar banding patterns to those previously documented.

All 4 ERIC types previously identified within *P. larvae* (Ashiralieva and Genersch 2006; Genersch et al. 2006), were shown to be present within our range of isolates. All UK environmental isolates were found to belong to ERIC type I (Chapter 3). The link between ERIC types and 16S rRNA phylogenies had not been previously explored. It was shown here that ERIC type I isolates were monophyletic within the phylogeny. Isolates representing ERIC types II-IV grouped together within the phylogeny. It is not known if this is an accurate representation, or a result of the low sample size of these isolates.

At first glance it appears that this observation is equal to that of the link between subspecies differentiation and 16S rRNA phylogeny. It should be noted that the inconsistencies relating to the congruence between 16S rRNA and subspecies differentiation do not exist within the link between ERIC type and 16S rRNA. Within purH and PyrE phylogenies, only isolates 6264 (*purH*) and 6993 (*purH*

and *PyrE*) were incorrectly grouped. This added further conviction to the suggestion that 6260 and 6261 were previously incorrectly identified.

As previously mentioned, the relationship between these gene phylogenies and virulence is unknown. Whilst ERIC types were believed to represent the virulent sub-type at one time, due to their reflection of the subspecies differentiation, this has been shown not to be the case (Genersch et al. 2006). Even upon the addition of ERIC types to the gene sequence differences, it is not possible to accurately determine the virulence of isolates.

6.2.3. Lateral Flow Devices not capable of detecting isolates belonging to all 4 ERIC types

Lateral Flow Devices (LFDs) are rapid field based diagnostic kits originally developed for the detection of plant pathogens (Danks and Barker 2000). This method was adopted for the detection of AFB infections (Vita, Europe). It was suggested that the antibodies used within this test kit were raised against an ERIC type I specific target (pers comm. Dr Giles Budge, FERA), however it was unknown if the commercially available test kits would be capable of detecting *P. larvae* from all 4 ERIC types.

As expected LFD kits were capable of detecting ERIC type I isolates, representing the majority of worldwide AFB infections (Antunez et al. 2009; Genersch et al. 2006). It was also shown that LFD kits were not capable of detecting isolates representing ERIC types II-IV. The significance of this finding was considered, whilst it could not be said that the assay was capable of detecting all AFB infections. Only a handful of ERIC type II infections have been reported causing systemic colony infections displaying typical AFB symptoms. Due to the low incidence rate, it is possible to identify these outbreaks using other methods, and crucially they can be identified by experienced beekeepers due to the AFB symptoms produced. ERIC types III and IV have never been shown to be responsible for systemic colony infections

displaying AFB symptoms, instead it is suggested that they are more likely capable of acute larval level infections (Genersch 2009).

It is not recommended at this time that the LFD kit be redesigned, as it appears that ERIC type I isolates are responsible for nearly all reported AFB infections. It should however be made clear that the LFD device should not be solely used for the identification of AFB, if unexpected results are achieved further diagnostic assistance should be sought for the appropriate identification. To fully understand AFB epidemiology, more studies would be required to explore the incidence of all 4 ERIC types (chapter 6.3.2)

6.2.4. No correlation observed between proteolytic activity and *in-vitro* virulence

Proteolytic activity is the only virulence factor consistently associated with the virulence of *P. larvae* (Dancer and Chantawannakul 1997; Holst and Sturtevant 1940). Considering proteases as virulence factors has often been debated, as often they are considered to be for metabolic processes rather than specific to infection (Lantz 1997). As in *Bacillus thuringiensis* (Andrews et al. 1985; Oppert 1999) it is likely that proteolytic activity is essential for the pathogenicity of the organism. The symptoms produced by AFB infections are indicative of proteolytic activity, however it remains unknown if the proteolytic activity is crucial in the infection pathway or merely involved in the degradation of larval remains.

One of the earliest documented AFB detection methods was the use of milk agar to detect proteolytic activity (Alippi 2001); this was considered confirmation of an AFB infection. This method was adapted for the comparison of proteolytic activity of *P. larvae* strains in the laboratory. It was noted that very few isolates expressed proteolytic activity, however 4 isolates were shown to exhibit levels of proteolytic activity and crucially strain 6993, previously deemed to be non-pathogenic (Watkins et al. 2003), was shown to express no proteolytic activity.

The use of an *in-vitro* honeybee rearing method showed that isolates expressing proteolytic activity on the lab assay, were not responsible for increased larval mortality. Conversely, and unexpected, was the greatly increased levels of larval mortality produced by strain 6993. This experiment was run on 2 separate occasions to confirm the observed result. At this time it was concluded that laboratory based proteolytic assays could not be used as a determinant of *P. larvae* virulence. It could not be determined if a correlation existed between proteolytic activity and *in-vitro* virulence, it remained unknown if the laboratory assay was an accurate representation of *P. larvae* proteolytic activity. It was believed that whilst laboratory conditions may not result in the expression of all proteolytic enzymes, the prior use of a milk agar assay suggested that this was an accurate representation.

This experiment did show, that unlike previously stated, strain 6993 was capable of being highly virulent within an *in-vitro* pathogenicity assay. It was concluded that this strain should not be considered for use as a biological control strain against EFB. It seemed likely that larvae fed with spores belonging to this strain, would likely perish in less than 4 days. This experiment did not reveal any honeybee larva that survived infection with this strain, so even if this bacterial isolate is capable of inhibit EFB infections. The resultant AFB infection negates the benefits of avoiding EFB.

It is suggested that symptomatic AFB infections, caused predominantly by ERIC type I isolates, involve proteolytic enzymes. The infection caused by ERIC type IV strain 6993 did not produce comparable symptoms and lacked the presence of proteolytic activity (chapter 4). It seems likely that proteolytic enzymes are involved in ERIC type I infections, however the increased levels of larval mortality caused by strain 6993, combined with no detectable proteolytic activity, imply that proteolytic enzymes are likely responsible for the symptoms produced during AFB infections, rather than being essential for host invasion. It seems logical that futures studies on the pathogenicity of *P. larvae* should not exclusively focus on proteolytic enzymes.

6.2.5. Sporulation likely to be responsible for the *in-vitro* virulence differences observed between strain 6993 and ERIC type I

It is widely reported that ERIC type I infections are responsible for the majority of AFB outbreaks worldwide (Antunez et al. 2009; Genersch et al. 2006). However *in-vitro* pathogenicity testing within this thesis showed that strain 6993 (an ERIC type IV isolate) was capable of causing greater larval mortality (chapter 4). A more accurate comparison of the differences in *in-vitro* virulence between ERIC types I and IV was planned, it was realised that ERIC type I isolates produced a maximum of 200 viable spores ml^{-1} compared to ERIC type IV isolates capable of producing >20,000 viable spores ml^{-1}. The approximate spore doses fed to larvae in previous testing were calculated, whilst all doses were capable of causing infection (Woodrow and Holst 1942), ERIC type IV isolates were fed to larvae in a higher concentration.

It was assumed that this difference in sporulation might be responsible for the observed difference in *in-vitro* virulence. Whilst not possible to accumulate large quantities of ERIC type I spores, it was possible to reduce the initial spore dose of strain 6993 (ERIC type IV). This reduction in spore dose caused a reduction in larval mortality. A replicate plate of this experiment still showed increased levels of larval mortality at low spore doses. The larvae for this test were selected from a limited section of brood comb, with each plate using larvae from a single side of the comb. It was suggested that the differences in larval mortality could be explained by the age of the larvae. It was calculated that an approximate difference of 12 hours was likely to exist between the 2 sides of the comb. This age difference has been shown to affect the spore dose required to cause infection. It is therefore proposed that larvae from plate 2 were 12 hours younger and as a result the lower spore doses were capable of causing infection. Further testing of this hypothesis is required, however this was not possible during the time frame of this thesis.

It is important to consider the implications of this finding on the previous proteolytic study. It has been suggested that *P. larvae* produces an array of proteolytic enzymes at the point of sporulation. As many isolates failed to produce large quantities of viable spores in laboratory media, it is likely that the proteolytic study was not a true reflect of proteolytic activity displayed by all isolates. The result from isolate 6993 still remains however, as it was shown this isolate produced large numbers of viable spores, and no proteolytic activity was detected. It still seems unlikely that a laboratory based proteolytic assay could be used a virulence determinant, however it is important to consider that different results may have been obtained should adequate numbers of spores been produced.

6.2.6. Genetic basis for the observed differences in sporulation

With the cost associated with genome sequencing and computational requirements for analysis reducing. It was possible to amplify large amounts of genomic data from 4 isolates of *P. larvae* representing ERIC types I and IV. It had been observed that ERIC type I isolates belonged to a phenotype not capable of producing viable spores under laboratory conditions (chapter 4). ERIC type IV was shown to represent a phenotype capable of producing large numbers of spores under laboratory conditions. It was hypothesised that these observed phenotypes would be represented by genetic differences within sporulation genes.

Sporulation was not well studied in *P. larvae* but had been adequately documented in *Bacillus subtilis*, a closely related species. It was proposed that sporulation genes from *B. subtilis* would have orthologs within the *P. larvae* genome. BLAST analysis revealed 89 of the 120 sporulation genes had orthologs within *P. larvae*. Further analysis of these sequences showed that 23 partial sequences representing 16 genes contained residue substitutions. These substitutions represented 2 genotypes, ERIC type 1 and ERIC type IV. The genes in which substitutions were shown to exist were linked to the initiation of

sporulation and the formation of viable spores, this suggested that these mutations were likely responsible for the differences in sporulation observed between ERIC type I and ERIC type IV isolates.

It was concluded that these substitutions were likely responsible for the phenotypic difference observed, ERIC type I isolates are known to produce >200,000 spores in a larval infections, as these have been recovered from AFB scales. Therefore it has to be assumed that all ERIC type I isolates possess the ability to produce viable spores and the difference was likely due to a deficiency found in ERIC type I isolates (as with thiamine deficiency (Heyndrickx et al. 1996)). This would explain why ERIC type I isolates do not sporulate in laboratory conditions. Further analysis of the ERIC type I genomes would likely reveal potential deficiencies to be addressed in the production of ERIC type I spores. At this time a true study of the differences observed in *in-vitro* pathogenicity could be performed.

6.3. To further understand the intra-species differences within *P. larvae*

This thesis aimed to highlight the differences in existence between strains of pathogenic *P. larvae* and a non-pathogenic biological control strain of *P. larvae*. However upon the implementation of an *in-vitro* pathogenicity assay, the biological control strain was shown to be highly pathogenic. As such there are no known non-pathogenic isolates of *P. larvae* in existence, so it was not possible to differentiate between these isolates. However the data collected in this thesis does further the knowledge of the differences in existence between isolates of *P. larvae*. How this data can be applied to further studies of *P. larvae* is discussed in the following sections.

6.3.1. Do non-pathogenic isolates of *P. larvae* exist?

This project started in the belief that we held a non-pathogenic isolate within our collection (Watkins et al. 2003). This was shown not to be the case, as observed by Genersch (2006) this isolate was shown to be highly virulent *in-vitro* (chapter 4). It remains unknown if non-pathogenic isolates of *P. larvae* exist, it might be suggested that the notion of non-virulent isolates simply refers to ERIC types III and IV. Without further evidence it is suggested that these isolates are considered virulent, and are capable of causing high levels of *in-vitro* virulence or rapid acute cell infections within a colony. Not all isolates displayed these characteristics, and it would be recommended that in the search for non-virulent isolates ERIC types III and IV remain the primary focus.

6.3.2. Exploring the epidemiology of ERIC type II-IV isolates

Whilst unlikely that ERIC type II-IV infections are responsible for symptomatic AFB infections. The prevalence of the organism amongst colonies is unknown. ERIC type II infections have been shown to cause symptomatic AFB infections that are not detected by LFD; no symptomatic AFB infections have currently been linked to ERIC type III or IV. Due to the production of symptoms and a potential colony level infection, ERIC II isolates are more likely to be detected. However not all ERIC type II isolates were previously identified as belonging to subspecies *larvae* some were identified as subspecies *pulvifaciens* it is unknown if the same disease progression is exhibited by ERIC type II isolates from both subspecies. ERIC types III and IV are believed to produce cell localised acute infections within a colony, it is plausible to suggest that many larval deaths per annum could be due to ERIC type III and IV infections.

To determine the epidemiology of all 4 ERIC types of *P. larvae* a large scale LFD independent random sampling method would need to be adopted, sampling a significant proportion of the total number of colonies. Only after such a study

would the true epidemiology of these isolates be understood. The significance of isolates belonging to ERIC types II-IV in relation to larval loss could be better determined at this point. Only after such a study would it be possible to accurately comment on which isolates should be detected by any future detection methods.

6.3.3. Completion of a *P. larvae* genome

Despite the genome of *P. larvae* being published in 2006, this remains unfinished (Qin et al. 2006). A subsequent sequencing project resulted in a better primary assembly of the genome (Chan et al. 2011), however this assembly was used to study the genetics of *P. larvae* rather than the creation of a complete genome. The primary assembly of strain 6254 produced in this thesis (chapter 5), is currently the best primary assembly of the *P. larvae* genome known to exist. Whilst more genetic data would be required to complete the genome of this single isolate, it is suggested that with minimal sequence divergence detected (chapter 2 and 5) that sequence data is combined to create a scaffold of the *P. larvae* genome. This scaffold will assist in the assembly of future *P. larvae* sequencing projects, acting as a reference. Completion of a *P. larvae* genome will ultimately lead to a better understanding of the organism and any potential virulence factors that are yet unknown.

6.3.4. Exploring the differences in pathogenicity existing between ERIC type IV isolates 6993 and 6264.

It was shown that the difference observed in *in-vitro* virulence was likely due to differing initial spore doses fed to larvae (chapter 4). This explains the difference observed between strains 6254/6678 and 6993, whereby it was shown that upon reduction to similar spore doses of 6254, an equal level of larval mortality was achieved from strain 6993. Sporulation data showed that strain 6264 was able to produce an often greater amount of viable spores than strain 6993, and possessed the same residue substitutions as were also found in strain 6993 (chapter 5). Comparisons of the *in-vitro* virulence of these isolates showed that strain 6264,

whilst of a similar spore dose to strain 6993, produced similar levels of mortality to that of strains 6254 and 6678 with far lower initial spore doses.

It is unknown what is responsible for this difference in *in-vitro* virulence, but it is likely that further data mining of the genome data would reveal potential factors responsible for this difference. Whilst it is possible that this strain represents a non-pathogenic subset of *P. larvae*, it would seem more likely at this time that the requirements for that strain to express it's virulent phenotype were not met.

6.3.5. Creation of a virulence specific *P. larvae* PCR assay

A potential outcome of this project was the creation and validation of a *P. larvae* assay capable of differentiating isolates based upon virulence. It was quickly determined that whilst intra-species groupings could be made, it was unknown what correlation existed between these groupings and virulence. At present time it is recommended that any new assays designed to detect AFB are designed to detect all isolates of *P. larvae*, this may result in the detection of false positives (based upon virulence). It seems preferential to detect all isolates of *P. larvae*, as at present time it is still assumed that all isolates possess the ability to cause infection.

With increased knowledge of the differences that exist within the species, and *in-vitro* testing on a wider range of *P. larvae* isolates. It may be possible to determine the differences that exist between virulent and avirulent isolates, at this time an assay could be developed based on these virulence determinants. If a further epidemiological studied was carried out on ERIC II-IV isolates and the true impact of these isolates was calculated, then it may be possible to conclude that previously determined intra-species groupings can be used to identify the *P. larvae* isolates of interest.

6.4. Conclusion

It was not possible to identify the differentiation between pathogenic and non-pathogenic isolates of *P. larvae* within this project. This was due to the previously identified biological control strain (6993), being shown to cause high levels of *in-vitro* larval mortality. Without the availability of a non-pathogenic strain it was possible to explore the differences that exist within the species as a whole, rather than specific differences existing between pathogenic and non-pathogenic isolates.

	ERIC type I	**ERIC type IV**
Previous subspecies label	*larvae*	*pulvifaciens*
Position on 16S Phylogeny	Monophyletic	Groups with ERIC types II and III
Position on MLST Phylogeny	Monophyletic	Groups with ERIC types II and III
Detected by LFD	No	Yes
Virulent *in-vitro*	No*	Yes / No
Sporulates in laboratory media	No	Yes

Table 6.2: An overview of the differences observed between ERIC type I and IV isolates within this thesis. *Within this study, previously shown to cause infection (Ashiralieva and Genersch 2006; Genersch et al. 2005; Genersch et al. 2006; Yue et al. 2008).

Phenotypic differences were shown to exist between ERIC type I and IV isolates (table 6.2), and comparative genomics revealed that this difference in sporulation was likely the result of multiple amino acid substitutions. The data presented within this thesis aids the study of variability amongst *P. larvae* isolates and has suggested many potential areas of interest to focus further studies.

Appendices

Appendix 1. Growth of *Paenibacillus larvae*

All handling of isolates of *P. larvae* beyond the point of initial isolation from hive material was performed in a laminar flow hood, ensuring a more sterile work environment. Disposable culturing tools were used, and a fresh pack opened prior to any experimental work. Incubation of *P. larvae* was at 34°C and liquid cultures were incubated on an orbital incubator operating at 170rpm.

Appendix 1.1. *P. larvae* growth Media

No specific media exists for *P. larvae*, as a result media was used which was developed for closely related species. Within this project many different solid media were trialed, most were unable to support the growth of *P. larvae* (data not shown). Two types of solid media were identified as useful for the growth of *P. larvae* isolates (table 7.1).

Media Name	Ingredients	Additives
BHI-T	Oxoid Brain Heart Infusion media (37g/L) and 5% Oxoid Agar No.1	Thiamine (final conc. 1mg/L)
Columbia Horse Blood Agar	Oxoid Columbia Blood Base agar (cat# CM0331) (39g/L) and 2% NaCl	5% Laked Horse Blood

Table 7.1: Solid media used for culturing *P. larvae* isolates.

No observable difference was detected between BHI-T or Columbia Horse blood agar, under the same growth conditions equal numbers of colonies were produced with near identical morphology. It was decided to continue using BHI-T agar due to the reduced cost and reduced production time.

Several experiments required a larger number of bacterial cells, for these experiments it was necessary to utilise liquid media. Two liquid media were found to support the growth of *P. larvae* isolates (Table 7.2).

Media Name	Ingredients	Additives
BHI-T	Oxoid Brain Heart Infusion media (37g/L)	Thiamine (final conc. 1mg/L)
TMYGP	1.5% Yeast Extract and 0.1% Sodium Pyruvate	0.4% Glucose and 0.03M Tris-maleate buffer

Table 7.2: Liquid media used for growth of *P. larvae*

It was noted that BHI-T broth cultures were capable of producing larger numbers of vegetative cells. TMYGP broth cultures produced lesser numbers of vegetative cells, but these cells produced more viable endospores after heat shocking. BHI-T media was routinely used for the growth of *P. larvae* in liquid media where spores were not expressively required. In instances where spores were required, TMYGP was used.

Appendix 2. Storage of *Paenibacillus larvae* isolates

Bacterial isolates were stored on BHI-T plates at 4°C for a maximum period of 7 days. Storage beyond this point utilised PROTECT cryopreservation beads (LabM, Product Code: D530). This allowed for long term storage of isolates at -80°C. All LMG reference isolates were cultured and stored using the PROTECT system, all subsequent experiments cultured LMG isolates from the frozen beads.

Appendix 3. DNA Extraction

All DNA extraction was performed using a Wizard® Genomic DNA Purification kit (Promega cat.#A1125). The protocol for purification of DNA from gram-positive bacteria was followed; only 60µl of 10mg/ml lysozyme was used to lyse the bacterial cell walls. Steps to remove RNA from the sample were not required so were not undertaken. Once the DNA pellet had been re-suspended in 100µl of DNA Rehydration Solution a subsequent 400µl of ddH_2O were added due to the high concentration of recovered genomic DNA.

Appendix 4. DNA Storage

All DNA was stored at -20°C to ensure minimal degradation over time. When required the DNA extracts were allowed to thaw at room temperature for the minimum possible time, to reduce exposure to the increased temperatures. All DNA isolates were handled in a semi-clean room after initial extraction. Un-extracted samples and amplified PCR reactions were excluded from this area to maintain low levels of contamination. If DNA extracts were repeatedly needed in a short time period, they were stored at 4°C, to avoid degradation due to repeated freeze-thaw cycles.

Appendix 5. PCR

All PCR was performed using GoTaq® Hot Start Polymerase (Promega, Cat. #M5006). 0.3U of Taq was added to PCR reactions. $MgCl_2$ was added at a final concentration of 1.5mM unless otherwise stated. A dNTP mix (Qiagen, Cat. #201900) was used in all reactions, at a final concentration of 200µM. To avoid contamination issues DEPC treated H_2O (Invitrogen, Cat. #AM9915G) was used. All PCR reactions were performed on the 96 well GeneAmp® PCR system (AppliedBiosystems, Cat. #N8050200) or the Veriti® 96 well thermal cycler (AppliedBiosystems, Cat. #4375786). The thermal cyclers were regular serviced and monitored under a UKAS accreditation scheme; differences in the thermal properties of the machines were within acceptable guidelines, allowing the use of both systems.

Appendix 6. PCR purification

PCR reaction clean-up prior to further analysis was performed using the QIAquick PCR purification kit (Qiagen, Cat. #28106), Following the PCR purification spin protocol. No modifications were made to this protocol as acceptable quantities of purified PCR product were recovered on all attempts.

Appendix 7. *In-vitro* rearing of honeybee larvae

The original method, from which the exposure bioassay was developed, was developed from a pesticides testing method in development by a consortium of European laboratories, led by INRA. Selwyn Wilkins at FERA made this method available for this project. It is believed that this method remains unpublished.

Unité Expérimentale d'Entomologie
LE MAGNERAUD

RÉPUBLIQUE FRANÇAISE
MINISTÈRE DE LA RECHERCHE
MINISTÈRE DE L'AGRICULTURE

METHOD FOR TESTING PESTICIDE TOXICITY TO HONEYBEE BROOD IN LABORATORY CONDITIONS

INTRODUCTION

The test provides quantitative oral toxicity data on honeybee brood. It is designed for *in vitro* treatments of active substances or formulated pesticides. These data should be used in an appropriate brood risk assessment scheme.

1. EXPERIMENTAL CONDITIONS

1.1 PRINCIPAL OF THE TRIAL

Larvae at the L1 stage are fed standardized amounts of artificial diet. Test products are incorporated into the food at the different concentrations within an appropriate range in order to compute the end points: LD50, LC50, NOAEL and NOAEC. The reference product is dimethoate.

1.2 TRIAL CONDITIONS

1.2.1 ORIGIN OF THE LARVAE

For one replicate, larvae are collected preferably from a unique colony. If two colonies are necessary, larvae originated from each colony must be distributed in two samples of equal size (24 larvae) in each plate. The colonies have to be healthy and must not show any visible clinical signs. Tests are realised with summer larvae during a period from the middle spring to the middle autumn, variable according to the country. In case of sanitary treatment, the date of application and the kind of product has to be supplied. No treatment has to be applied within the 4 weeks preceding the beginning of experiments.
At D-3 (Fig. 4) the queen of the chosen colony is confined in its own colony in an excluder cage containing a comb with emerging worker brood and empty cells (Fig. 1). The excluder cage is placed close to combs containing larvae. At D-2, about 30 hours after encaging, the queen is removed from the cage, after checking the presence of fresh laid eggs. The comb containing the eggs is left in the cage, near the brood, during the incubation stage and until hatching (D1).

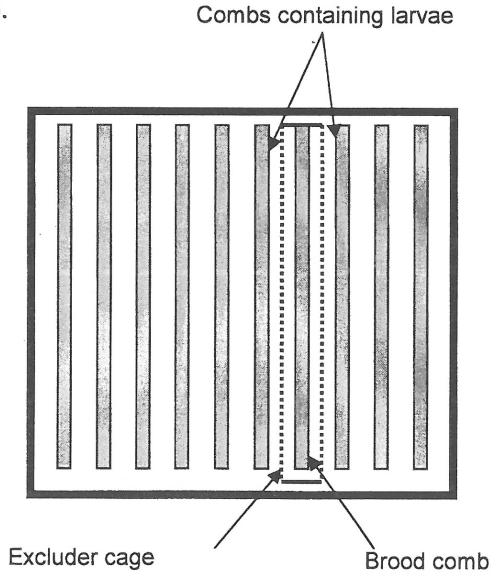


Fig. 1: Cross section of a hive with the excluder cage

1.2.2 COLLECTING LARVAE

At D1 (Fig. 3), the comb containing fresh laid eggs is carried from the hive to the laboratory (regulated at a constant temperature of 25°C if possible) in a special wooden container in order to avoid temperature variation. It is then introduced into a laminar-flow hood for grafting.

1.2.3 PREPARATION OF REARING MATERIAL

1.2.3.1 LARVAL FOOD

Larvae are reared in crystal polystyrene grafting cells (ref CNE/3, NICOPLAST Society) having an internal diameter of 9 mm. The cells are firstly immerged for 30min in 0.4% MBC (methyl benzethonium chloride) in water, and then dried in a laminar-flow hood. Each cell is placed into a well of a 48-well tissue culture plate, which was previously half filled with a piece of dental roll wetted with 15.5% glycerol in 0.4% MBC (Fig. 2).

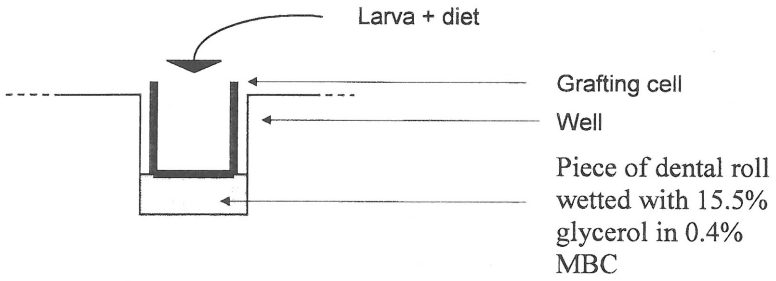


Fig. 2: Larval cell in a tissue culture well.

These plates are placed into a hermetic Plexiglas desiccator (NALGENE 5314-0120 or 5317-0180 according to the volume required), with a dish filled with K2SO4 saturated solution in order to keep a water saturated atmosphere. The dessicator is placed into an incubator at 34°C.

The food is composed of the three following diets:

Diet A (D1):
50% of fresh royal jelly + 50% of an aqueous solution containing 2% of yeast extract, 12% of glucose and 12% of fructose.

Diet B (D3):
50% of fresh royal jelly + 50% of an aqueous solution containing 3% of yeast extract, 15% of glucose and 15% of fructose.

Diet C (from D4 to D6):
50% of fresh royal jelly + 50% of an aqueous solution containing 4% of yeast extract, 18% of glucose and 18% of fructose.

For each diet, the sugar solution is filtered at 0.22µm with a syringe filter before being mixed to the royal jelly.
We consider that a "fresh royal jelly" is a royal jelly collected during the year, devided into 5g aliquots, and stored in a freezer at -20°C. The diets prepared for a test are stored in a fridge at +5°C during the test. We use pure osmosed water, free from all organic contaminant.
The micropipettes used to provide the diet into the cells are equipped with disposable tips.
The larvae are collected with a thin paintbrush (N° 3/0).

1.2.3.2 PUPATION AND EMERGENCE

At D7 (pre pupa stage), the plates are transferred into a hermetic container containing a dish filled with a saturated NaCl solution in order to keep 80% of relative humidity. The container is then placed into an incubator at 34°C.
At D15, each plate is transferred into a crystal polypropylene box (11 x 15 x 12cm) with a cover aerated with a wire gauze, and containing a piece of comb with a small plastic royal pheromone diffuser in its centre (bee boost®), fixed with a wire (Fig. 3). Emerging bees are fed with syrup and pollen powder provided by bird feeders. The boxes are kept in the hermetic container.

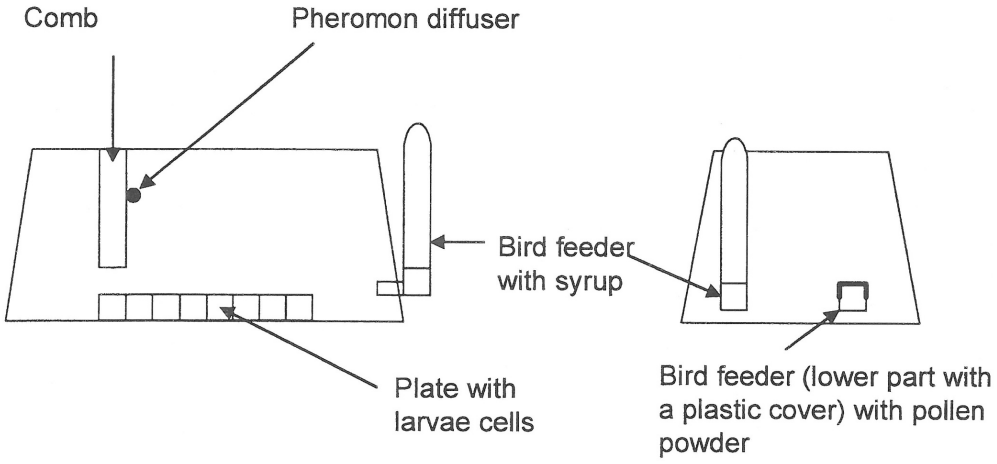


Fig. 3: Emergence box

1.2.4 GRAFTING AND FEEDING OF LARVAE

20µl of diet A is dropped into each cell, and one larva is delicately transferred from the comb to each cell, on the surface of the diet, using the wetted paintbrush. When a plate is completed with 48 larvae, it is placed into the hermetic container, and then into the incubator at 34°C.
Larvae are fed once a day (except at D2) at the same time with a sterilised pipette, according to the schedule of Fig. 4. The diet is warmed at 34°C prior to each using.

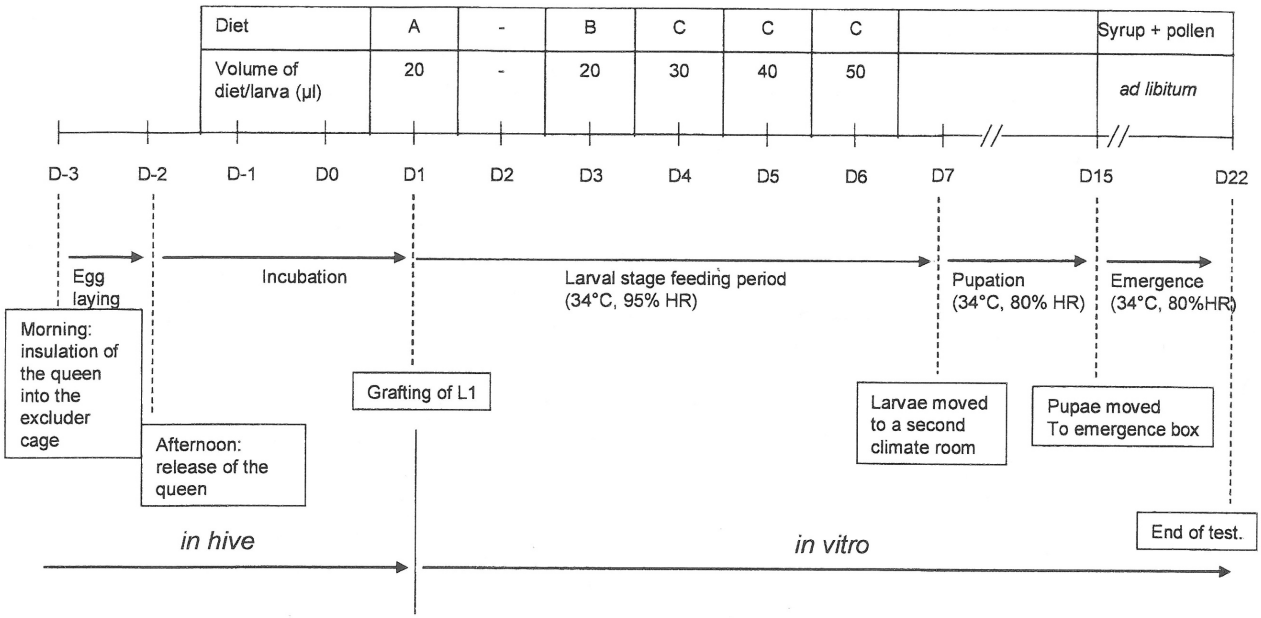


Fig. 4: Steps of an *in vitro* test.

2. EXPERIMENTAL LAY OUT

The experimental unit is the 48 larvae plate. For each test, the following treatments should be used:

- control without solvent (1 plate)
- control with solvent (1 plate) if necessary
- 5 treatments i.e. the 5 doses or concentrations to test (1 plate per treatment)
- reference treatment with dimethoate (1 plate)
- 1 additive plate (totally or partially filled, according to the number of larvae available) can be used at D4, in the acute toxicity test to replace the larvae which died before D4..

One test has a minimum of three replicates with one different larvae origin and new tested solutions for each replicate.

3. TOXICITY TESTING

3.1 PREPARATION OF THE PESTICIDE SOLUTIONS

3.1.1 SOLUTIONS OF THE TEST PESTICIDE

The test pesticide is dissolved in water. If it is not soluble in water at the experimental concentrations, we can use an other solvent as acetone. In that case, it is required to prepare a second control fed with diet containing the solvent at the same concentration as the treated samples.
Dilutions of the stock solutions are made preferably with osmosed water, using disposable pipette tips equipped with a filter.

The rate of the tested solution must not exceed 10% of the final volume. A different rate has to be justified. In all cases, we must use a constant volume for the different dilutions in order to have a constant rate between the diet and the test pesticide solution.

3.1.2 SOLUTIONS OF THE TOXIC REFERENCE

The toxic reference is dimethoate:

- in acute toxicity test: 3 µg/larva mixed with diet C and provided at D4,
- in chronic toxicity test: mixed with the three diets at the constant concentration of 20,000 µg / kg diet..

3.2 TREATMENTS

In acute toxicity test, larvae are treated at D4 with the diet C containing the preparation to test at the suitable concentration.
For chronic toxicity test, larvae are treated every day (except D2) with the diets containing the preparation to test at a constant concentration.

3.3 PRELIMINARY EXPERIMENT FOR DOSE RANGE FINDING

In order to assess the adequate LD 50 range, it is recommended to run a preliminary experiment where doses of the test preparation may vary according to a geometrical ratio from 5 to 10.

4. RESULTS

4.1 DEFINITION OF MORTALITY

LARVA: An immobile larva or a larva which does not react to the contact of the paintbrush is noted as dead.

PUPA: A non emerged individual at D22 is noted as dead during pupal stage.

ADULT: An immobile adult which does not react to a tactile stimulation is noted as dead.

4.2 TIME TABLE OF MORTALITY CHECKS

LARVA: At the feeding moment, dead larvae are systematically removed for sanitary reasons. Specific mortality checks are made according to the type of test.
In the test where exposure is at D4 (acute toxicity), a first mortality check is made at D4 in order to replace the dead larvae before they have started consuming the diet containing the insecticide. Then we note the mortality at D5, D6 and D7.
In the test with chronic exposure mortality is noted at D7.

PUPA: Non emerged bees are counted at D22.

ADULT: Alive adult bees and dead adults which have left their cell and show a nomal development are both counted at D22.

5. STATISTICAL ANALYSIS AND PRESENTATION OF RESULTS

5.1 VALIDITY RANGE OF DATA

-In control samples larval mortality (number of dead larvae/48), pupal mortality (number of dead pupae at D22/number of alive pre pupae at D7) and adult mortality (number of dead emerged bees at D22/total number of emerged bees) must be lower or equal to 15% for the assessment of a LD50 or a LC50, 20% for the assessment of a NOAEL or a NOAEL. In case of higher mortality in control sample the replicate is invalidated..
- The rate mortalities with dimethoate must be:
    - higher than or equal to 50% at D6 for larvae exposed to 3 µg / larva at D4
    - higher than or equal to 50% at D7chronic exposure of larvae to the concentration: 20,000 µg / kg diet.
- The calculated LD 50 and LC50 must be in each case between the two extreme tested doses. They must not be extrapolated out of the tested limits.

Any deviation from the above conditions will invalidate the test

5.2 DATA ANALYSIS

5.2.1 LD50 AND LC50 CALCULATION

Mortalities are expressed in percentage of the reference populations after an adjustment according to the Abbott formula (1925):

$$M = \frac{(P - T)}{S} \times 100$$ raw mortalities

$$M = \frac{(\%P - \%T)}{100 - \%S} \times 100$$ percent mortalities

M: adjusted mortality expressed in percent of the initial population, initial number of larvae (48) for a larval mortality, number of alive pre pupae at D7 for a pupal mortality, number of emerged bees at D22 for an adult mortality

P: mortality due to the treatment

T: control mortality

S: surviving number in control

%P: mortality percentage due to the treatment

%T: control mortality percentage

The results will be analysed using a regression model with a high adjustment level, which can be checked with the determination coefficient value. All raw and adjusted data must appear in

the study report. The lethality graphs and their equations must be reported. The results must be notified for a LD50 at 24 and 48h in µg per individual, and for a LC50 in µg per litre of solution (ppb). These calculated variables must be given with a confidence interval at 5%.

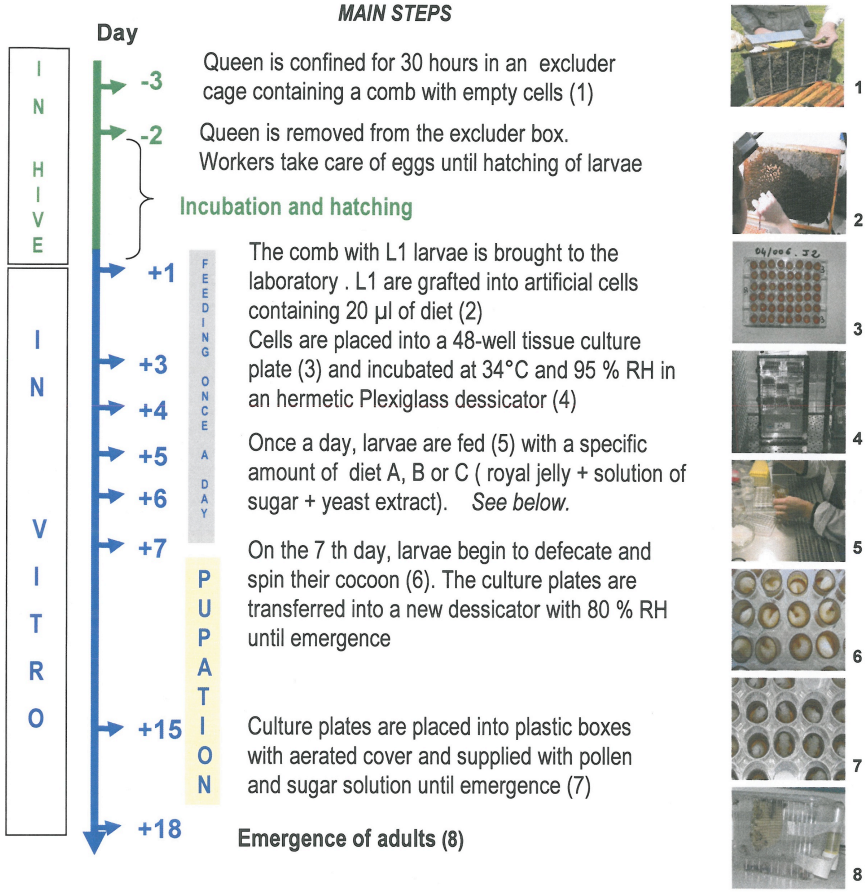
5.2.2 Determination of the NOAEL and NOAEC

The NOAEL an NOAEC are the highest dose and concentration respectively, which do not induce a mortality significantly higher than that observed in controls. This analysis will be done by using a Chi2 test.

In vitro method for rearing *Apis mellifera* larvae

D. Fortini, P. Aupinel, B. Michaud, J-F. Odoux, J-N Taseï
Unité expérimentale d'Entomologie

MAIN STEPS



Queen is confined for 30 hours in an excluder cage containing a comb with empty cells (1)

Queen is removed from the excluder box. Workers take care of eggs until hatching of larvae

Incubation and hatching

The comb with L1 larvae is brought to the laboratory . L1 are grafted into artificial cells containing 20 µl of diet (2)
Cells are placed into a 48-well tissue culture plate (3) and incubated at 34°C and 95 % RH in an hermetic Plexiglass dessicator (4)

Once a day, larvae are fed (5) with a specific amount of diet A, B or C (royal jelly + solution of sugar + yeast extract). *See below.*

On the 7 th day, larvae begin to defecate and spin their cocoon (6). The culture plates are transferred into a new dessicator with 80 % RH until emergence

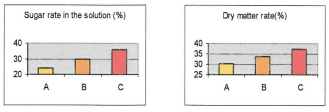
Culture plates are placed into plastic boxes with aerated cover and supplied with pollen and sugar solution until emergence (7)

Emergence of adults (8)

Mode of feeding
(larvae are fed once a day, volume total 160 µl)

DAY	D1	D3	D4	D5	D6
Diet	A	B	C	C	C
Volume/larvae	20 µl	20 µl	30 µl	40 µl	50 µl

Diets compositions
(50% royal jelly+ 50% sugar solution + yeast extract)



INSTITUT NATIONAL DE LA RECHERCHE AGRONOMIQUE
CENTRE Poitou-Charentes
Unité expérimentale d'entomologie
Le Magneraud BP 52 • 17700 Surgères
www.poitou-charentes.inra.fr

ALIMENTATION
AGRICULTURE
ENVIRONNEMENT

INRA

Appendix 8. Development of an exposure bioassay for the study of *P. larvae* infections.

The following sections document the changes made to the method listed above

Appendix 8.1. Selecting larvae of the correct age

The method aims to uniformly control the age of brood on the desired comb; this is achieved by confining the queen to a chosen new comb within her colony. Combs containing young brood are placed on either side of the chosen comb, to encourage the laying of eggs. The queen is left caged on the comb for 30 hours, at which point she is released and the comb is examined for freshly laid eggs.

Queens from several colonies were confined to combs in this way, however none of them laid eggs. Queens will only lay eggs if all factors meet her standards. Lower outside temperature reduces the foraging activity of a hive. A lack of foraging reduces the amount of food available to the hive. In addition workers prior to the laying of eggs by the queen must clean cells. A lack of stored food or available cleaned cells reduces the number of eggs laid by a queen (Free and Williams 1974). Several attempts were made at this method, but none of the confined queens laid eggs as expected. An alternative approach was needed to obtain young brood for grafting. Beekeepers from the National Bee Unit (NBU) are able to quickly age the brood on a comb with a simple visual inspection in the field. The NBU beekeepers selected an appropriate comb from within a colony, unfortunately adding a degree of variability to the experiment. Larvae of 24-48 hours old are very small and often harder to see than the eggs that they emerge from. As a result of this inspection in the field by eye proves very difficult and could potentially result in the use of larvae of ±12 hours. This was deemed as an acceptable age difference, as further selection of appropriate aged larvae can then take place in the laboratory with magnifying equipment.

Appendix 8.2. Collection of larvae

The NBU beekeepers chose the appropriate combs from the colonies, which were removed from the colony and shaken to return the adult bees to the donor colony. The brood comb was placed in a paper potato sack before transporting to the laboratory. This maintained a temperature close to the ~34°C maintained within a hive (Dunham 1931), thus keeping the brood in a healthy condition ready for grafting. It also shelters the vulnerable larvae from wind and rain that could be detrimental to their health. Enclosing the comb also helps to maintain the high level of humidity required for brood development.

Appendix 8.3. Artificial cell set up

The larvae are reared in sterilised polystyrene grafting cells (Thorne, http://www.thorne.co.uk), these are contained within a 48 well tissue culture plate (Griener bio-one, Cat.#677102). To maintain a hygienic environment Methylbenzethonium Chloride (MBC) (Sigma-Aldrich, Cat.#M7379-10G) is used at a 0.4% concentration to soak the grafting cells for at least 30 minutes before leaving to dry in a sterile flow hood. In to each well on the plate goes a small piece of cotton dental roll; this is then saturated with a 15.5% glycerol (Sigma-Aldrich, Cat.#G5516-1L) and 0.4% MBC solution. This set-up ensures that each well can be treated as an individual replicate, as there is no cross contamination between grafting cells. It is worth noting that the dental roll must be kept small to ensure the grafting cells do not come into contact with the lid of the plate (Figure 7.1).

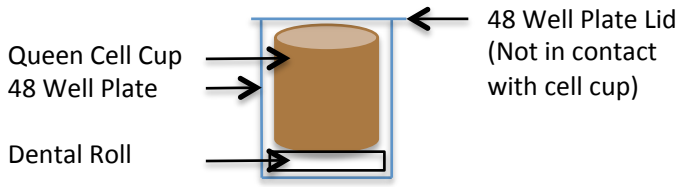


Figure 7.1: Diagram of the artificial cell set up used in this experiment

Appendix 8.4. Environmental conditions

The optimum temperature for this process is 34°C, the average temperature recorded at the centre of brood chamber in a colony (Dunham 1931). Controlled humidity is also required, to achieve this a hermetic acrylic desiccator (Thermo Scientific, Cat.#5317-0120) is used. The required humidity is 95%, this can be achieved by placing several petri dishes in the bottom drawer of the dessicator each containing a saturated solution of potassium sulphate (K_2SO_4) (Sigma-Aldrich, Cat.#PO772-250G), which has been experimentally shown to maintain a relative humidity of >95% (Rockland 1960).

Appendix 8.5. Longevity of the experiment

The method calls for the use of a special emergence box; this provides different conditions that are required towards the later stages of pupation and ultimately for the successful emergence of adult bees. To reduce the complexity of this experiment, it was decided that this was not required. Experimental data showed that the slowest AFB infections take place within ~10 days (Genersch et al. 2006); emergence normally does not occur until day 22. The end point of this experiment was set at 10 days, allowing the usage of just one hermetic chamber.

Plate Number	Number of deceased larvae
1	4
2	5
3	7
4	10

Table 7.3: Control larval mortality achieved on 4 48 well plates.

With in-house testing of this method ranging from 25-80% control mortality (pers com), it was necessary to assess the average mortality rates achieved with this experimental set up. The average mortality across the 4 plates were observed (Table 7.3) The average larval mortality observed was 13.54%, with the lowest mortality rate of 8.33% and highest mortality rate of 20.83%. The rate was below rates previously achieved in house, and provided the basal level for the design of the experiment.

Appendix 8.6. Number of larvae per test group

It was determined that a test group of 24 larvae per bacterial isolate would be used. Provided that control mortality rates did not increase, this would result in the loss of approximately 3 larvae from each group not due to infection. This was deemed an acceptable sample size to show the overall mortality rates induced by the strains of bacteria. It was likely that control mortality rates would decrease with increased experience at performing the experiment. Larval mortality in groups infected with a bacterial isolate, would be compared with the associated control allowing for calculation of number of deaths attributed to infection.

Appendix 8.7. Grafting

Pre-warmed plates are filled with 20µl of pre-warmed Diet A food (Table 7.4), the plates are returned to the desiccator to allow the food to settle to the bottom of the cell cup prior to grafting.

Various methods of grafting exist, ranging from the use of paintbrushes to the usage of specialist tools. This method favors the use of a special stainless steel grafting tool with the occasional use of a Chinese grafting tool (Thorne, http://www.thorne.co.uk). Working within a laminar flow hood the comb is examined to find an area that contains similar sized brood, these are assumed to be of approximately the same age due to the typical laying pattern followed by a queen (Camazine 1991). Using the steel-grafting tool, larvae are removed from the cells on the comb paying careful attention to the orientation of the larvae. The larvae must be placed on the pre-warmed food within the grafting plate in the same orientation in which it was removed from the cell on the comb. This avoids accidental drowning of the larva in the food. This process is repeated until the plate has been filled. The lid is added and the plate is returned to the desiccator within the incubator.

In order to treat each cell as an individual it is important to sterilise the grafting tool each time it is used, wiping away excess food before dipping in 100%

ethanol and flaming seemed the best method of sterilization that would be the least toxic to the next larva. The tool is cooled in double distilled sterile distilled water before grafting the next larva.

Appendix 8.8. Larval diet

During the larval stage of life the larvae have differing nutritional requirements, with concentrations increasing with the age of the larvae (Table 7.4).

	Diet A	Diet B	Diet C
Royal Jelly	50%	50%	50%
Yeast Extract	1%	1.5%	2%
Glucose	6%	7.5%	9%
Fructose	6%	7.5%	9%

Table 7.4: Shows the final concentrations (% Volume) added to the three required larval diets.

The larval diet is prepared from a stock solution, this was filter sterilised using a 0.2μM pore sized filter. Royal jelly was collected from hives known to be free of disease, this was separated into 5 and 10g aliquots and then frozen at -20°C. The food was prepared for use each day by defrosting the royal jelly and adding the corresponding amount of filtered stock solution to the measured weight of royal jelly. The food was pre-warmed to 34°C prior to use. Ensuring the larvae were not subjected to rapid temperature changes during the course of the experiment.

Appendix 8.9. Feeding

Day	**1**	**2**	**3**	**4**	**5**	**6**
Diet	A	-	B	C	C	C
Volume (μl)	20	-	20	30	40	50

Table 7.5: Feeding regime followed during the rearing of larvae

The larvae are fed a varied diet during the first 6 days (table 7.5). Previously determined volumes of larval food formed the basic understanding for the experiment, however in practice the amount of food pipetted into each cell had to be considered on an individual basis. Sterile pipette tips were used to ensure no contamination was added to the feed. Larvae were fed until the point at which

they would be capped over by worker bees in preparation for the pupation stage. Whilst the feeding of different quantities of food adds a degree of variability between replicates, no bacterial spores were added to food after the initial feed. Adjusting the quantity of food pipetted ensured that were not underfed leading to starvation or overfed leading to the drowning of the submerged larva.

Appendix 8.10. Bacterial inoculants

In order to regulate the dose of bacteria given to the infant larvae, the strains were all grown up in sporulation media for 5 days. From previous experimental work, it had been observed that sporulation occurred readily under these conditions. Broths were grown until they reached a uniform optical density (OD 0.6); at which time the bacterial cells were harvested by centrifugation of a 15ml aliquot. The broth supernatant was removed and the cells were re-suspended in ddH_2O. The resultant cell suspension was then heated for 20 minutes at 80°C to remove any vegetative cells. This was then added to larval diet A to be used for the initial feed and vortexed before pipetting in to the brood cups.

Appendix 8.11. Examining the larvae

Starting on day three larvae are examined individually using a plate microscope (Figure 7.2), this low magnification allows for the visualisation of respiration and subtle body movements. If no signs of movement or respiration are detected in 15 seconds, the larvae are touched with a sterile steel-grafting tool to see if a reflex reaction is present. If no movement was detected, then the larvae were considered to be deceased. As with feeding this process starts on day three, but unlike feeding this process continues on till the end of the experiment. Deceased larvae are given the time of checking as the designated time of death, unfortunately without more regular checking or a video imaging system it is not possible to give more accurate times of death.

Larvae were not checked more frequently as it is highly likely that increased observations would have dramatically increased the number of control

mortalities due to increased exposure to lower levels of humidity and lower temperatures combined with an increased level of handling using the grafting tool to see the desired movement.

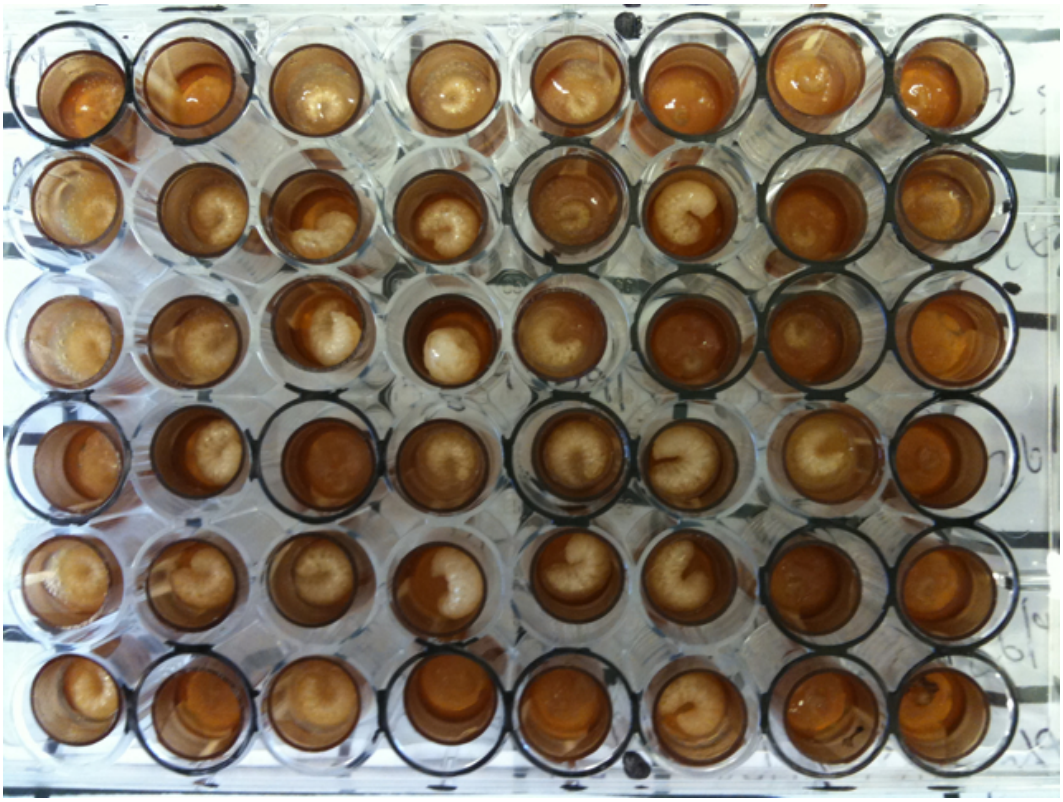


Figure 7.2: Image showing grafted larvae at approximately day 7. Wells circled with black pen contain deceased larvae.

Appendix 8.12. Grafting capacity

Combs were provided by NBU beekeepers at approximately 10:30am, after morning inspections had revealed an appropriately aged comb. All equipment was prepared in advance ready for the grafting process. The process of grafting larvae was very labor intensive, with the completion of a single 48 well plate containing 2 experimental groups taking >1.5 hours. The daily quota of 4 plates of grafted larvae required 6 hours of delicate work. Beyond this point it was not possible to graft subsequent larva without a highly increased degree of human error. Due to this increased error and the resulting larval deaths, 4 plates was set as the maximum to be grafted in a day.

Appendix 8.13. Incubation of grafted larvae

Incubator space did become a limiting factor after several days, the desiccator used contained 3 shelves each shelf could hold 4 plates un-stacked or 8 plates stacked. It was decided that un-stacked plates were preferable as these were unvented 48 well plates, and with the high levels of humidity it was decided that the risk of condensation creating a seal around the plates was too great. So to avoid the possible suffocation of the larvae it was decided to place 4 plates on each shelf. The increased space required by the desiccator within the incubator, and the unavailability of another desiccator set the maximum number of plates to be incubated as 12 at any time.

Appendix 8.14. Time required for feeding and observing larvae

Larvae were checked each 24 hours after an initial 48-hour period, these checks on the very young larvae were incredibly time consuming. Larvae was examined under a low powered microscope and due to the lack of movement at this age, they had to be observed for up to 30 seconds to witness them breathing. If after this time no signs of life were visible, they had to be gently touched with a clean grafting tool and observed for a reaction to this stimulus. This process was very time consuming, and especially whilst relatively inexperienced at this process it could take over 20 minutes to check each plate of larvae. When the time taken to check for larval mortality was combined with the time required for the preparation of food and the feeding of larvae, a significant proportion of a day had been consumed.

References

Alippi A.M. . Bacterial diseases, 1999 In Colin M.E. (ed.), Ball B.V. (ed.), Kilani M. (ed.) . Bee disease diagnosis . Zaragoza : CIHEAM-IAMZ,.

Alippi, A. M., and Aguilar, O. M. (1998). "Characterization of isolates of *Paenibacillus larvae* subsp. *larvae* from diverse geographical origin by the polymerase chain reaction and BOX primers." *Journal of Invertebrate Pathology*, 72(1), 21-7.

Alippi, A. M., Albo, G. N., Reynaldi, F. J., and De Giusti, M. R. (2005). "In vitro and in vivo susceptibility of the honeybee bacterial pathogen *Paenibacillus larvae* subsp. *larvae* to the antibiotic tylosin." *Veterinary Microbiology*, 109(1-2), 47-55.

Alippi, A. M., López, A. C., and Aguilar, O. M. (2002). "Differentiation of *Paenibacillus larvae* subsp. *larvae*, the cause of American foulbrood of honeybees, by using PCR and restriction fragment analysis of genes encoding 16S rRNA." *Applied And Environmental Microbiology*, 68(7), 3655-60.

Alippi, A. M., López, A. C., and Aguilar, O. M. (2004). "A PCR-based method that permits specific detection of *Paenibacillus larvae* subsp. *larvae*, the cause of American Foulbrood of honey bees, at the subspecies level." *Letters in Applied Microbiology*, 39(1), 25-33.

Alippi, A. M., López, A. C., Reynaldi, F. J., Grasso, D. H., and Aguilar, O. M. (2007). "Evidence for plasmid-mediated tetracycline resistance in *Paenibacillus larvae*, the causal agent of American Foulbrood (AFB) disease in honeybees." *Veterinary Microbiology*, 125(3-4), 290-303.

Alippi, A. M., and Reynaldi, F. J. (2006). "Inhibition of the growth of *Paenibacillus larvae*, the causal agent of American foulbrood of honeybees, by selected strains of aerobic spore-forming bacteria

isolated from apiarian sources." *Journal of Invertebrate Pathology*, 91(3), 141-6.

Altschul, S., Gish, W., Miller, W., Myers, E. W., and Lipman, D. (1990). "Basic local alignment search tool (BLAST)." *Journal of Molecular Biology*, 215, 403-410.

Andrews, R. E., Bibilos, M. M., and Bulla, L. A. (1985). "Protease activation of the entomocidal protoxin of *Bacillus thuringiensis* subsp. *kurstaki*." *Applied And Environmental Microbiology*, 50(4), 737-42.

Andrup, L., Barfod, K. K., Jensen, G. B., and Smidt, L. (2008). "Detection of large plasmids from the *Bacillus cereus* group." *Plasmid*, 59(2), 139-43.

Antúnez, K., Anido, M., Schlapp, G., Evans, J. D., and Zunino, P. (2009). "Characterization of secreted proteases of *Paenibacillus larvae*, potential virulence factors involved in honeybee larval infection." *Journal of Invertebrate Pathology*, 102(2), 129-132.

Antúnez, K., Arredondo, D., Anido, M., and Zunino, P. (2011). "Metalloprotease production by *Paenibacillus larvae* during the infection of honeybee larvae." *Microbiology*, 157(5), 1474-1480.

Antúnez, K., D'Alessandro, B., Piccini, C., Corbella, E., and Zunino, P. (2004). "*Paenibacillus larvae larvae* spores in honey samples from Uruguay: a nationwide survey." *Journal of Invertebrate Pathology*, 86(1-2), 56-8.

Antúnez, K., Piccini, C., Castro-Sowinski, S., Rosado, A. S., Seldin, L., and Zunino, P. (2007). "Phenotypic and genotypic characterization of *Paenibacillus larvae* isolates." *Veterinary Microbiology*, 124(1-2), 178-83.

Ash, C., Priest, F. G., and Collins, M. D. (1993). "Molecular identification of rRNA group 3 bacilli (Ash, Farrow, Wallbanks and Collins) using a PCR probe test." *Antonie van Leeuwenhoek*, 64(3), 253-260.

Ashiralieva, A., and Genersch, E. (2006). "Reclassification, genotypes and virulence of Paenibacillus larvae, the etiological agent of American Foulbrood" *Apidologie,* 37, 411-420.

Bailey, L. (1963). "The pathogenicity for honey-bee larvae of microorganisms associated with European foulbrood." *Journal of Insect Pathology*, 5, 198–205.

Bakonyi, T., Derakhshifar, I., Grabensteiner, E., and Nowotny, N. (2003). "Development and evaluation of PCR assays for the detection of *Paenibacillus larvae* in honey samples: comparison with isolation and biochemical characterization." *Applied And Environmental Microbiology*, 69(3), 1504-10.

Benada, O., Drobnikova, V., Kalachova, L., and Ludvik, J. (1988). "Plasmid DNA in *Bacillus larvae*." *Journal of Apicultural Research,* 27, 35-39.

Bohlin, J., Skjerve, E., and Ussery, D. W. (2008). "Reliability and applications of statistical methods based on oligonucleotide frequencies in bacterial and archaeal genomes." *BMC Genomics*, 9(1), 104-122.

Brar, S. K., Verma, M., Tyagi, R., Surampalli, R., Barnabe, S., and Valero, J. (2007). "*Bacillus thuringiensis* proteases: Production and role in growth, sporulation and synergism." *Process Biochemistry*, 42(5), 773-790.

Bravo, A., Agaisse, H., Salamitou, S., and Lereclus, D. (1996). "Analysis of cryIAa expression in sigE and sigK mutants of *Bacillus thuringiensis*." *Molecular & general genetics: MGG*, 250(6), 734-741.

Brehm, S. P., Staal, S. P., and Hoch, J. A. (1973). "Phenotypes of pleiotropic-negative sporulation mutants of *Bacillus subtilis*." *Journal of bacteriology*, 115(3), 1063-1070.

Brødsgaard, C. J., Ritter, W., and Hansen, H. (1998). "Response of in vitro reared honey bee larvae to various doses of *Paenibacillus larvae larvae* spores." *Apidologie*, 29, 569-578.

Brown, M., Ball, R., Morton, J., and Wilkins, S. (2007). "Foul brood disease of honey bees: recognition and control." *FERA Foulbrood Publication*.

Burbulys, D., and Trach, K. (1991). "Initiation of sporulation in *B. subtilis* is controlled by a multicomponent phosphorelay." *Cell,* 64, 545-552.

Camazine, S. (1991). "Self-organizing pattern formation on the combs of honey bee colonies." *Behavioral Ecology and Sociobiology*, 28(1), 61-76.

Catalano, F. A., Meador-Parton, J., Popham, D. L., and Driks, A. (2001). "Amino acids in the *Bacillus subtilis* morphogenetic protein SpoIVA with roles in spore coat and cortex formation." *Journal of bacteriology*, 183(5), 1645-1654.

Chan, Q. W., Cornman, R. S., Birol, I., Liao, N. Y., Chan, S. K., Docking, T. R., Jackman, S. D., Taylor, G. A., Jones, S. J., de Graaf, D. C., Evans, J. D., and Foster, L. J. (2011). "Updated genome assembly and annotation of *Paenibacillus larvae*, the agent of American foulbrood disease of honey bees." *BMC Genomics*, 12(1), 450-455.

Chung, M. C., Popova, T. G., Millis, B. A., Mukherjee, D. V., Zhou, W., Liotta, L. A., Petricoin, E. F., Chandhoke, V., Bailey, C., and Popov, S. G. (2006). "Secreted neutral metalloproteases of *Bacillus anthracis* as candidate pathogenic factors." *Journal of Biological Chemistry*, 281(42), 31408-31418.

Clarridge III, J. E. (2004). "Impact of 16S rRNA gene sequence analysis for identification of bacteria on clinical microbiology and infectious diseases." *Clinical microbiology reviews*, 17(4), 840-862.

Cole, J. R., Wang, Q., Cardenas, E., Fish, J., Chai, B., Farris, R. J., Kulam-Syed-Mohideen, A. S., Mcgarrell, D. M., Marsh, T., Garrity, G. M., and Tiedje, J. M. (2009). "The Ribosomal Database Project: improved alignments and new tools for rRNA analysis." *Nucleic Acids Research*, 37(Database), D141-D145.

Cox-Foster, D. L., Conlan, S., Holmes, E. C., Palacios, G., Evans, J. D., Moran, N. A., Quan, P.-L., Briese, T., Hornig, M., Geiser, D. M., Martinson, V., Vanengelsdorp, D., Kalkstein, A. L., Drysdale, A., Hui, J., Zhai, J., Cui, L., Hutchison, S. K., Simons, J. F., Egholm, M., Pettis, J. S., and Lipkin, W. I. (2007). "A Metagenomic Survey of Microbes in Honey Bee Colony Collapse Disorder." *Science*, 318(5848), 283-287.

Dancer, B., and Chantawannakul, P. (1997). "The Proteases of American Foulbrood Scales." *Journal of Invertebrate Pathology*, 70(2), 79-87.

Danks, C., and Barker, I. (2000). "On-site detection of plant pathogens using lateral-flow devices." *EPPO Bulletin*, 30(3-4), 421-426.

De Arruda, M., Miller, R., Ferreira, M., and Felipe, M. (2003). "Comparison of *Crinipellis perniciosa* isolates from Brazil by ERIC repetitive element sequence‐based PCR genomic fingerprinting." *Plant Pathology*, 52(2), 236-244.

De Bruijn, F. (1992). "Use of repetitive (repetitive extragenic palindromic and enterobacterial repetitive intergeneric consensus) sequences and the polymerase chain reaction to fingerprint the genomes of *Rhizobium meliloti* isolates and other soil bacteria." *Applied and environmental microbiology*, 58(7), 2180-2187.

de Graaf, D. C., Alippi, A. M., Brown, M., Evans, J. D., Feldlaufer, M., Gregorc, A., Hornitzky, M., Pernal, S. F., Schuch, D. M. T., Titera, D., Tomkies, V., and Ritter, W. (2006a). "Diagnosis of American foulbrood in honey bees: a synthesis and proposed analytical protocols." *Letters in Applied Microbiology*, 43(6), 583-90.

de Graaf, D. C., De Vos, P., Heyndrickx, M., Van Trappen, S., Peiren, N., and Jacobs, F. J. (2006b). "Identification of *Paenibacillus larvae* to the subspecies level: an obstacle for AFB diagnosis." *Journal of Invertebrate Pathology*, 91(2), 115-23.

Disayathanoowat, T., Young, J., Helgason, T., and Chantawannakul, P. (2011). "T-RFLP analysis of bacterial communities in the midguts of *Apis mellifera* and *Apis cerana* honey bees in Thailand." *FEMS Microbiology Ecology*, (In Press).

Djordjevic, S. P., Forbes, W. A., Smith, L. A., and Hornitzky, M. A. (2000). "Genetic and biochemical diversity among isolates of *Paenibacillus alvei* cultured from Australian honeybee (Apis mellifera) colonies." *Applied And Environmental Microbiology*, 66(3), 1098-106.

Dobbelaere, W., de Graaf, D. C., Reybroeck, W., Desmedt, E., Peeters, J. E., and Jacobs, F. J. (2001). "Disinfection of wooden structures contaminated with *Paenibacillus larvae* subsp. *larvae* spores." *Journal of Applied Microbiology*, 91(2), 212-6.

Dubnau, D., Smith, I., Morell, P., and Marmur, J. (1965). "Gene conservation in *Bacillus* species. I. Conserved genetic and nucleic acid base sequence homologies." *Proceedings of the National Academy of Sciences of the United States of America*, 54(2), 491.

Dunham, W. E. (1931). "Hive temperatures for each hour of a day." *Ohio Journal of Science*, 31, 181–188.

Dustmann, J. H., and von der Ohe, W. (1997). "Efficient prophylactic measures against American foulbrood by bacteriological analysis of honey for spore contamination." *American Bee Journal*, 137(8), 603-606.

Eckburg, P. B. (2005). "Diversity of the Human Intestinal Microbial Flora." *Science*, 308(5728), 1635-1638.

Errington, J. (1993). "*Bacillus subtilis* sporulation: regulation of gene expression and control of morphogenesis." *Microbiology and Molecular Biology Reviews*, 57(1), 1-33.

Evans, J. D., and Pettis, J. S. (2005). "Colony-level impacts of immune responsiveness in honey bees, *Apis mellifera*." *Evolution*, 59(10), 2270-4.

Forsgren, E., Olofsson, T. C., Vasquez, A., and Fries, I. (2010). "Novel lactic acid bacteria inhibiting *Paenibacillus larvae* in honey bee larvae." *Apidologie*, 41(1), 99-108.

Forsgren, E., Stevanovic, J., and Fries, I. (2008). "Variability in germination and in temperature and storage resistance among *Paenibacillus larvae* genotypes." *Veterinary Microbiology*, 129(3-4), 342-9.

Fortnagel, P., and Freese, E. (1968). "Analysis of sporulation mutants II. Mutants blocked in the citric acid cycle." *Journal of bacteriology*, 95(4), 1431-1438.

Free, J., and Williams, I. H. (1974). "Factors determining food storage and brood rearing in honeybee (*Apis mellifera* L.) comb." *Journal of Entomology Series A, General Entomology*, 49(1), 47-63.

Fricke, B., Drössler, K., Willhardt, I., Schierhorn, A., Menge, S., and Rücknagel, P. (2001). "The cell envelope-bound metalloprotease (camelysin) from *Bacillus cereus* is a possible pathogenic factor." *Biochim Biophys Acta*, 1537(2), 132-46.

Fries, I., Lindström, A., and Korpela, S. (2006). "Vertical transmission of American foulbrood (*Paenibacillus larvae*) in honey bees (*Apis mellifera*)." *Veterinary Microbiology*, 114(3-4), 269-74.

Genersch, E. (2010a). "American Foulbrood in honeybees and its causative agent, *Paenibacillus larvae*." *Journal of Invertebrate Pathology*, 103, S10-S19.

Genersch, E. (2010b). "Honey bee pathology: current threats to honey bees and beekeeping." *Applied Microbiology and Biotechnology*, 87(1), 87-97.

Genersch, E., Ashiralieva, A., and Fries, I. (2005). "Strain- and genotype-specific differences in virulence of *Paenibacillus larvae* subsp. *larvae*, a bacterial pathogen causing American foulbrood disease in honeybees." *Applied And Environmental Microbiology*, 71(11), 7551-5.

Genersch, E., Forsgren, E., and Pentikainen, J. (2006). "Reclassification of *Paenibacillus larvae* subsp. *pulvifaciens* and *Paenibacillus larvae* subsp. *larvae* as *Paenibacillus larvae* without subspecies differentiation." *International Journal of Systematic and Evolutionary*, 56, 501-511.

Genersch, E., and Otten, C. (2003). "The use of repetitive element PCR fingerprinting (rep-PCR) for genetic subtyping of German field isolates of *Paenibacillus larvae* subsp. *larvae*." *Apidologie*, 34, 195-206.

Gillings, M., and Holley, M. (1997). "Repetitive element PCR fingerprinting (rep‐PCR) using enterobacterial repetitive intergenic consensus (ERIC) primers is not necessarily directed at ERIC elements." *Lett Appl Microbiol*, 25(1), 17-21.

Govan, V. A., Allsopp, M. H., and Davison, S. (1999). "A PCR detection method for rapid identification of *Paenibacillus larvae*." *Applied And Environmental Microbiology*, 65(5), 2243-5.

Hall, N. (2007). "Advanced sequencing technologies and their wider impact in microbiology." *Journal of Experimental Biology*, 210(9), 1518-1525.

Han, C. S., Xie, G., Challacombe, J. F., Altherr, M. R., Bhotika, S. S., Bruce, D., Campbell, C. S., Campbell, M. L., Chen, J., and Chertkov, O. (2006). "Pathogenomic sequence analysis of *Bacillus cereus* and *Bacillus*

thuringiensis isolates closely related to *Bacillus anthracis*." *Journal of bacteriology*, 188(9), 3382.

Harmsen, D., and Karch, H. (2004). "16S rRNA for Diagnosing Pathogens: a Living Tree." *ASM News-American Society for Microbiology*, 70(1), 19-24.

Haseman, L. (1961). "How long can spores of American foulbrood live?" *American Bee World*, 101, 298-299.

Haydak, M. H. (1970). "Honey Bee Nutrition." *Annual Review of Entomology*, 15(1), 143-156.

Helgason, E., Okstad, O. A., Caugant, D. A., Johansen, H. A., Fouet, A., Mock, M., Hegna, I., and Kolstø, A. B. (2000). "*Bacillus anthracis*, *Bacillus cereus*, and *Bacillus thuringiensis*--one species on the basis of genetic evidence." *Applied And Environmental Microbiology*, 66(6), 2627-30.

Helgason, E., Tourasse, N. J., Meisal, R., Caugant, D. A., and Kolstø, A.-B. (2004). "Multilocus sequence typing scheme for bacteria of the *Bacillus cereus* group." *Applied And Environmental Microbiology*, 70(1), 191-201.

Hermans, P., Sluijter, M., Hoogenboezem, T., Heersma, H., van Belkum, A., and de Groot, R. (1995). "Comparative study of five different DNA fingerprint techniques for molecular typing of *Streptococcus pneumoniae* strains." *Journal of clinical microbiology*, 33(6), 1606-1612.

Heyndrickx, M., Vandemeulebroecke, K., Hoste, B., Janssen, P., Kersters, K., De Vos, P., Logan, N. A., Ali, N., and Berkeley, R. C. (1996). "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*." *International Journal of Systematic Bacteriology*, 46(1), 270-9.

Hierro, N., Gonzalez, A., Mas, A., and Guillamon, J. M. (2004). "New PCR-based methods for yeast identification." *Journal Applied Microbiology*, 97(4), 792-801.

Higes, M., Martín-Hernández, R., Garrido-Bailón, E., González-Porto, A. V., García-Palencia, P., Meana, A., Del Nozal, M. J., Mayo, R., and Bernal, J. L. (2009). "Honeybee colony collapse due to *Nosema ceranaein* professional apiaries." *Environmental Microbiology Reports*, 1(2), 110-113.

Hoch, J. (1993). "Regulation of the phosphorelay and the initiation of sporulation in *Bacillus subtilis*." *Annual Reviews in Microbiology*, 47, 441-465.

Hoch, J. A. (1970). "Genetic Analysis of Pleiotropic Negative Sporulation Mutants in *Bacillus subtilis*." *Journal of bacteriology*, 105(3), 896-901.

Holst, E. C., and Sturtevant, A. P. (1940). "Relation Of Proteolytic Enzymes To Phase Of Life Cycle Of *Bacillus Larvae*, And Two New Culture Media For This Organism." *Journal of Bacteriology*, 40(5), 723-731.

Hornitzky, M., and Anderson, D. (2003). "Honeybee Diseases." *Australia and New Zealand Standard Diagnostic Procedures*, 1-13.

Hrabak, J., and Martinek, K. (2007). "Screening of secreted proteases of *Paenibacillus larvae* by using substrate-SDS-polyacrylamide gel electrophoresis." *Journal of Apicultural Research*, 46, 160-164.

Huang, X. (1999). "CAP3: A DNA Sequence Assembly Program." *Genome Research*, 9(9), 868-877.

Hulton, C., Higgins, C., and Sharp, P. (1991). "ERIC sequences: a novel family of repetitive elements in the genomes of *Escherichia coli*, *Salmonella*

typhimurium and other enterobacteria." *Molecular Microbiology*, 5(4), 825-834.

Iurlina, M. O., and Fritz, R. (2005). "Characterization of microorganisms in Argentinean honeys from different sources." *International Journal of Food Microbiology*, 105(3), 297-304.

Janda, J. M., and Abbott, S. L. (2007). "16S rRNA Gene Sequencing for Bacterial Identification in the Diagnostic Laboratory: Pluses, Perils, and Pitfalls." *Journal of Clinical Microbiology*, 45(9), 2761-2764.

Jarosz, J., and Glinski, Z. (1990). "Selective Inhibition of Cecropin-like Activity of Insect Immune Blood by Protease from American Foulbrood Scales." *Journal of Invertebrate Pathology*, 56, 143-149.

Jiang, M., Shao, W., and Perego, M. (2000). "Multiple histidine kinases regulate entry into stationary phase and sporulation in *Bacillus subtilis*." *Molecular microbiology*, 38(3), 535-542.

Kaltenpoth, M. (2011). "Honeybees and bumblebees share similar bacterial symbionts." *Molecular Ecology*, 20(3), 439-440.

Katznelson, H. (1950). "*Bacillus pulvifaciens* (N. SP.), an organism associated with powdery scale of honeybee larvae." *Journal of Bacteriology*, 59(2), 153-5.

Khan, A. A., McCarthy, S., Wang, R. F., and Cerniglia, C. E. (2002). "Characterization of United States outbreak isolates of *Vibrio parahaemolyticus* using enterobacterial repetitive intergenic consensus (ERIC) PCR and development of a rapid PCR method for detection of O3: K6 isolates." *FEMS microbiology letters*, 206(2), 209-214.

Kim, M., Nishiyama, Y., Mura, K., Tokue, C., and Arai, S. (2004). "Gene cloning and characterization of a *Bacillus vietnamensis* metalloprotease." *Bioscience Biotechnology Biochemistry*, 68(7), 1533-40.

Klofat, W., Picciolo, G., Chappelle, E. W., and Freese, E. (1969). "Production of adenosine triphosphate in normal cells and sporulation mutants of *Bacillus subtilis*." *Journal of Biological Chemistry*, 244(12), 3270-3276.

Knaysi, G. (1945). "A study of some environmental factors which control endospore formation by a strain of *Bacillus mycoides*." *Journal of bacteriology*, 49(5), 473-493.

Lantz, M. (1997). "Are bacterial proteases important virulence factors?" *Journal of periodontal research*, 32(1), 126-132.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., Mcgettigan, P. A., Mcwilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J., and Higgins, D. G. (2007). "Clustal W and Clustal X version 2.0." *Bioinformatics*, 23(21), 2947-2948.

Lauro, F. M., Favaretto, M., Covolo, L., Rassu, M., and Bertoloni, G. (2003). "Rapid detection of *Paenibacillus larvae* from honey and hive samples with a novel nested PCR protocol." *International Journal of Food Microbiology*, 81(3), 195-201.

Li, E., and Yousten, A. A. (1975). "Metalloprotease from *Bacillus thuringiensis*." *Applied and Environmental Microbiology*, 30(3), 354-361.

Lindström, A. (2008). "Distribution of *Paenibacillus larvae* spores among adult honey bees (*Apis mellifera*) and the relationship with clinical symptoms of American foulbrood." *Microbial Ecology*, 56(2), 253-9.

Lindström, A., Korpela, S., and Fries, I. (2008). "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." *Journal of Invertebrate Pathology*, 99(1), 82-6.

Loncaric, I., Derakhshifar, I., Oberlerchner, J., Köglberger, H., and Moosbeckhofer, R. (2008). "Genetic diversity among isolates of *Paenibacillus larvae* from Austria." *Journal of Invertebrate Pathology*, 100 (1), 44-46.

Ma, M., Wang, C., Ding, Y., Li, L., Shen, D., Jiang, X., Guan, D., Cao, F., Chen, H., Feng, R., Wang, X., Ge, Y., Yao, L., Bing, X., Yang, X., Li, J., and Du, B. (2011). "Complete Genome Sequence of *Paenibacillus polymyxa* SC2, a Strain of Plant Growth-Promoting *Rhizobacterium* with Broad-Spectrum Antimicrobial Activity." *Journal of Bacteriology*, 193(1), 311-312.

Maori, E., Paldi, N., Shafir, S., Kalev, H., Tsur, E., Glick, E., and Sela, I. (2009). "IAPV, a bee‐affecting virus associated with Colony Collapse Disorder can be silenced by dsRNA ingestion." *Insect Molecular Biology*, 18(1), 55-60.

Marston, C. K., Gee, J. E., Popovic, T., and Hoffmaster, A. R. (2006). "Molecular approaches to identify and differentiate *Bacillus anthracis* from phenotypically similar *Bacillus* species isolates." *BMC Microbiology*, 6, 22-29.

Martin, B., Humbert, O., Camara, M., Guenzi, E., Walker, J., Mitchell, T., Andrew, P., Prudhomme, M., Alloing, G., and Hakenbeck, R. (1992). "A highly conserved repeated DNA element located in the chromosome of *Streptococcus pneumoniae*." *Nucleic Acids Research*, 20(13), 3479-3483.

Martinson, V. G., Danforth, B. N., Minckley, R. L., Rueppell, O., Tingek, S., and Moran, N. A. (2011). "A simple and distinctive microbiota associated with honey bees and bumble bees." *Molecular Ecology*, 20(3), 619-628.

Matheson, A. (1993). "World Bee Health." *Bee World*, 74(4), 176-212.

Msadek, T. (1999). "When the going gets tough: survival strategies and environmental signaling networks in *Bacillus subtilis*." *Trends in microbiology*, 7(5), 201-207.

Murray, K. D., Aronstein, K. A., and de León, J. H. (2007). "Analysis of pMA67, a predicted rolling-circle replicating, mobilizable, tetracycline-resistance plasmid from the honey bee pathogen, *Paenibacillus larvae*." *Plasmid*, 58(2), 89-100.

Nakamura, L. (1984). "Bacillus pulvifaciens sp. nov., nom. rev." *International Journal of Systematic Bacteriology*, 34 (4), 410-413.

Neuendorf, S., Hedtke, K., Tangen, G., and Genersch, E. (2004). "Biochemical characterization of different genotypes of *Paenibacillus larvae* subsp. *larvae*, a honey bee bacterial pathogen." *Microbiology (Reading, Engl)*, 150(Pt 7), 2381-90.

Olsen, G. J., and Woese, C. R. (1993). "Ribosomal RNA: a key to phylogeny." *The FASEB Journal*, 7(1), 113-123.

Oppert, B. (1999). "Protease interactions with *Bacillus thuringiensis* insecticidal toxins." *Archives of Insect Biochemistry and Physiology*, 42, 1-12.

Ozin, A. J., Samford, C. S., Henriques, A. O., and Moran Jr, C. P. (2001). "SpoVID guides SafA to the spore coat in *Bacillus subtilis*." *Journal of bacteriology*, 183(10), 3041-3049.

Pannucci, J., Okinaka, R. T., Sabin, R., and Kuske, C. R. (2002). "*Bacillus anthracis* pXO1 plasmid sequence conservation among closely related bacterial species." *Journal of Bacteriology*, 184(1), 134-41.

Patel, J. (2001). "16S rRNA gene sequencing for bacterial pathogen identification in the clinical laboratory." *Molecular Diagnosis*, 6 (4), 313-321.

Peng, C. Y., Mussen, E., Fong, A., Cheng, P., Wong, G., and Montague, M. A. (1996). "Laboratory and field studies on the effects of the antibiotic tylosin on honey bee *Apis mellifera* L. (Hymenoptera: *Apidae*) development and prevention of American foulbrood disease." *Journal of Invertebrate Pathology*, 67(1), 65-71.

Peng, Y. S. C., Mussen, E., Fong, A., Montague, M. A., and Tyler, T. (1992). "Effects of chlortetracycline of honey bee worker larvae reared in vitro." *Journal of Invertebrate Pathology*, 60(2), 127-133.

Petti, C., Polage, C., and Schreckenberger, P. (2005). "The role of 16S rRNA gene sequencing in identification of microorganisms misidentified by conventional methods." *Journal of clinical microbiology*, 43(12), 6123-6125.

Qin, X., Evans, J. D., Aronstein, K. A., Murray, K. D., and Weinstock, G. M. (2006). "Genome sequences of the honey bee pathogens *Paenibacillus larvae* and *Ascosphaera apis*." *Insect Molecular Biology*, 15(5), 715-718.

R Okinaka, K Cloud, O Hampton, A Hoffmaster, K Hill, P. K., T Koehler, G Lamke, S Kumano, and Manter, D. (1999). "Sequence, assembly and analysis of pX01 and pX02." *Journal of Applied Microbiology*, 87(2), 261-262.

Ratnieks, F. L. W. (1992). "American foulbrood: The spread and control of an important disease of the honey bee." *Bee World*, 73(4), 177-191.

Rauch, S., Ashiralieva, A., Hedtke, K., and Genersch, E. (2009). "Negative correlation between individual-insect-level virulence and colony-level virulence of *Paenibacillus larvae*, the etiological agent of American foulbrood of honeybees." *Applied And Environmental Microbiology*, 75(10), 3344-7.

Reboli, A. C., Bryan, C. S., and Farrar, W. E. (1989). "Bacteremia and infection of a hip prosthesis caused by *Bacillus alvei*." *Journal of Clinical Microbiology*, 27(6), 1395-6.

Rieg, S., Martin Bauer, T., Peyerl-Hoffmann, G., Held, J., Ritter, W., Wagner, D., Kern, W. V., and Serr, A. (2010). "*Paenibacillus larvae* Bacteremia in injection drug users." *Emerging Infectious Diseases*, 16(3), 487-9.

Ritter, W. (2003). "Early detection of American foulbrood by honey and wax analysis." *Apiacta*, 38, 125-30.

Rockland, L. B. (1960). "Saturated Salt Solutions for Static Control of Relative Humidity between 5 and 40 C." *Analytical Chemistry*, 32(10), 1375-1376.

Sacchi, C., Whitney, A., and Mayer…, L. (2002). "Sequencing of 16S rRNA gene: a rapid tool for identification of *Bacillus anthracis*." *Emerging infectious diseases*, 8 (10), 1117-1123.

Setlow, P. (2006). "Spores of *Bacillus subtilis*: their resistance to and killing by radiation, heat and chemicals." *Journal of Applied Microbiology*, 101(3), 514-525.

Shida, O., Takagi, H., Kadowaki, K., and Nakamura, L. (1997a). "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*." *International Journal of Systematic and Evolutionary Biology*, 47 (2), 289-298.

Shida, O., Takagi, H., Kadowaki, K., Nakamura, L. K., and Komagata, K. (1997b). "Emended Description of *Paenibacillus amylolyticus* and Description of *Paenibacillus illinoisensis* sp. nov. and *Paenibacillus chibensis* sp. nov." *International journal of systematic bacteriology*, 47(2), 299-306.

Shin, J. H., Sung, S. K., Shim, J. D., Kim, H. R., Park, S. W., and Lee, a. J. N. (2005). "A Case of Paenibacillus alvei Cellulitis in Immunocompetent Patient." *Korean Journal of Laboratory Medicine*, 25(1), 53-55.

Sonenshein, A. (2000). "Control of sporulation initiation in Bacillus subtilis." *Current opinion in microbiology*, 3, 561-566.

Sorokin, A., Candelon, B., Guilloux, K., Galleron, N., Wackerow-Kouzova, N., Ehrlich, S. D., Bourguet, D., and Sanchis, V. (2006). "Multiple-locus sequence typing analysis of *Bacillus cereus* and *Bacillus thuringiensis* reveals separate clustering and a distinct population structure of psychrotrophic strains." *Applied And Environmental Microbiology*, 72(2), 1569-78.

Stabler, R. A., He, M., Dawson, L., Martin, M., Valiente, E., Corton, C., Lawley, T. D., Sebaihia, M., Quail, M. A., Rose, G., Gerding, D. N., Gibert, M., Popoff, M. R., Parkhill, J., Dougan, G., and Wren, B. W. (2009). "Comparative genome and phenotypic analysis of *Clostridium difficile* 027 strains provides insight into the evolution of a hypervirulent bacterium." *Genome Biology*, 10(9), R102-R107.

Stackebrandt, E., and Goebel, B. (1994). "Taxonomic note: a place for DNA-DNA reassociation and 16S rRNA sequence analysis in the present species definition in bacteriology." *International Journal of Systematic Bacteriology*, 44(4), 846-849.

Stead, D. E. (1988). "Identification Of Bacteria By Computer-Assisted Fatty Acid Profiling." *ISHS Acta Horticulturae*, 225, 39-46.

Stern, M. J., Ames, G. F. L., Smith, N. H., Clare Robinson, E., and Higgins, C. F. (1984). "Repetitive extragenic palindromic sequences: a major component of the bacterial genome." *Cell*, 37(3), 1015-1026.

Takahashi, T., Kataoka, A., Kamikawa, M., and Nishimoto, S. (2001). "Sporicidal activity of radiation on *Paenibacillus larvae* in bee hives." *Honeybee Science*, 22 (4), 181-187.

Takamatsu, H., Kodama, T., Imamura, A., Asai, K., Kobayashi, K., Nakayama, T., Ogasawara, N., and Watabe, K. (2000). "The *Bacillus subtilis* yabG gene is transcribed by SigK RNA polymerase during sporulation, and yabG mutant spores have altered coat protein composition." *Journal of bacteriology*, 182(7), 1883-1888.

Tarr, P. I. (1995). "*Escherichia coli* O157: H7: clinical, diagnostic, and epidemiological aspects of human infection." *Clinical Infectious Diseases*, 20(1), 1-8.

Tjalsma, H., Antelmann, H., Jongbloed, J. D. H., Braun, P. G., Darmon, E., Dorenbos, R., Dubois, J. Y. F., Westers, H., Zanen, G., and Quax, W. J. (2004). "Proteomics of protein secretion by *Bacillus subtilis*: separating the" secrets" of the secretome." *Microbiology and Molecular Biology Reviews*, 68(2), 207-233.

Tomkies, V., Flint, J., Johnson, G., Waite, R., Wilkins, S., Danks, C., Watkins, M., Cuthbertson, A. G. S., Carpana, E., Marris, G., Budge, G., and Brown, M. A. (2009). "Development and validation of a novel field test kit for European foulbrood." *Apidologie*, 40(1), 63-72.

Van der Auwera, G., and Mahillon, J. (2008). "Transcriptional analysis of the conjugative plasmid pAW63 from *Bacillus thuringiensis*." *Plasmid*, 60(3), 190-9.

Van der Auwera, G. A., Andrup, L., and Mahillon, J. (2005). "Conjugative plasmid pAW63 brings new insights into the genesis of the *Bacillus anthracis* virulence plasmid pXO2 and of the *Bacillus thuringiensis* plasmid pBT9727." *BMC Genomics*, 6, 103-127.

Van der Auwera, G. A., Timmery, S., and Mahillon, J. (2008). "Self-transfer and mobilisation capabilities of the pXO2-like plasmid pBT9727 from *Bacillus thuringiensis* subsp. *konkukian* 97-27." *Plasmid*, 59(2), 134-8.

van Passel, M., Bart, A., Luyf, A., van Kampen, A., and Van Der Ende, A. (2006). "Compositional discordance between prokaryotic plasmids and host chromosomes." *BMC Genomics*, 7(1), 26-34.

Versalovic, J., Koeuth, T., and Lupski, J. R. (1991). "Distribution of repetitive DNA sequences in eubacteria and application to fingerprinting of bacterial genomes." *Nucleic Acids Res*, 19(24), 6823-6831.

Vinatzer, B. A., and Yan, S. (2007). "Mining the genomes of plant pathogenic bacteria: how not to drown in gigabases of sequence." *Mol Plant Pathol*, 9(1), 105-118.

Wang, J., Steggles, J. R., and Ellar, D. J. (2008). "Molecular characterization of virulence defects in *Bacillus thuringiensis* mutants." *FEMS Microbiology Letters*, 280(1), 127-34.

Watkins, M., Thompson, H., and Dancer, B. (2003). "Development of a method of biological control of European foulbrood in honeybees to supersede antibiotic treatments and sanitary colony destruction by burning." *Horticulture LINK Project*, 1-75.

White, G. F. (1906). "The bacteria of the apiary, with special reference to bee diseases." *Bureau of Entomology Technical Series*, 14.

Williams, D. (2000). "A Veterinary Approach to the European Honey Bee (Apis mellifera)." *The Veterinary Journal*, 160, 61-73.

Wilson, W. T. (1971). "Resistance to American foulbrood in honey bees. XI* 1:: Fate of *Bacillus larvae* spores ingested by adults." *Journal of Invertebrate Pathology*, 17(2), 247-255.

Woodrow, A., and Holst, E. (1942). "The mechanism of colony resistance to American foulbrood." *Journal Economical Entomology*, 35(3), 327-330.

Yoshiyama, M., and Kimura, K. (2009). "Bacteria in the gut of Japanese honeybee, *Apis cerana japonica*, and their antagonistic effect against *Paenibacillus larvae*, the causal agent of American foulbrood." *Journal of Invertebrate Pathology*, 102(2), 91-6.

Yu, Y. T., and Kroos, L. (2000). "Evidence that SpoIVFB is a novel type of membrane metalloprotease governing intercompartmental communication during *Bacillus subtilis* sporulation." *Journal of Bacteriology*, 182(11), 3305-9.

Yue, D., Nordhoff, M., Wieler, L. H., and Genersch, E. (2008). "Fluorescence in situ hybridization (FISH) analysis of the interactions between honeybee larvae and *Paenibacillus larvae*, the causative agent of American foulbrood of honeybees (*Apis mellifera*)." *Environmental Microbiology*, 10(6), 1612-20.

Zuckerkandl, E., and Pauling, L. (1965). "Molecules as documents of evolutionary history." *Journal of Theoretical Biology*, 8(2), 357-366.