Insect facultative symbionts: immunological interactions, transgenerational effects and their role in host defence

Megan Saphana Rose Hasoon

Doctor of Philosophy

University of York

Biology

December 2020

Abstract

Facultative symbionts are found in most insects, having a range of effects on their host's biology. They often influence how these insects respond to a range of environmental stressors, including aiding in defence against natural enemies. Aphids are a model species for studying these endosymbionts and the work in this thesis uses the pea aphid to investigate some of the open questions about host-microbe interactions within this system and beyond.

Using both experimental and genomic techniques, we investigate the interactions between the host and their facultative symbionts, the role this plays in defence and the wider implications this may have for symbiont evolution. We find that facultative symbionts mediate transgenerational effects, leading to decreased susceptibility to a natural enemy across generations, likely due to an interaction between the host immune system and the bacterial symbiont. We also find that symbiont presence alters the immune response of the aphid, though this does not appear to have large impacts on the response to an invading pathogen. Using genomic techniques, we investigate the genomes of newly sequenced isolates of the two facultative symbionts used throughout this thesis. Here, we highlight the differences across and within symbiont species, outlining the variation in genome evolution and suggesting potential mechanisms for some of the phenotypes conferred by these symbionts.

This work highlights the dynamic relationship between the pea aphid host and their bacterial endosymbionts and how ecological pressures may play a role in shaping these interactions. Furthermore, it outlines how these relationships may in turn influence symbiont evolution. The work in this thesis demonstrates the multiple levels of complexity to these host-symbiont relationships and opens up new avenues for exploration within these diverse, complex communities.

Table of Contents

List of figures

List of tables

Acknowledgements

I would first like to start by thanking my supervisor Julia Ferrari for her advice, support and constant encouragement throughout. For her continued patience as I stumbled through the first years and probably made far too much noise in the office next door. I am also grateful for her allowing me to develop outside of my research, encouraging me to engage with teaching, science communication and outreach projects which often came as a welcomed break from the research. I would also like to thank my co-supervisor Alistair Darby for his, and his groups, help in the first year as I started my journey in bioinformatics. With their support, a steep learning curve that consisted mostly of angry keyboard bashing, turned into something I actually quite enjoyed and with slightly less keyboard bashing.

Thank you to my TAP panel, Ville Friman and Gavin Thomas for their guidance and valued discussion that helped me to improve my research greatly. I would also like to thank Sally James and Lesley Gilbert in the Technology Facility for being exceedingly patient with my questions and qPCR struggles.

The Department of Biology, especially D0, was a fantastic place to work and I made some incredible friends along the way. Thank you to Eleanor for keeping me going in the lab, especially on the long and difficult days. You kept me sane, albeit that your method of reading me woodlouse facts and talk of spider genitalia was a little unconventional. To my office family Leda and Jeremy, thanks for being the best biology Mum and Dad I could ask for, you guys really got me through when things got tough. To Jordan and Pierre who dragged me through the final stages; thanks for always reminding me that I'm not as stupid as I think, although many of my actions suggested otherwise. And to Pasky, thanks for all the pints.

I was lucky enough to start my PhD with an incredible cohort spread across the entire department, which of course gave me plenty of places to walk to and procrastinate. Thank you especially to Sophie, Jack, Bryden, Kate, Katie, Elena, Dom and Tom. For the best

memories, the tears over beers and always having someone to dance it out with come Friday night. I couldn't have done it without you.

I'd lastly like to thank my family. Graeme, for everything really - the reality checks when I needed them the most, for late night trips to the lab just to keep me company, becoming proficient in the art of stacking and pouring agar plates, and in the final stages, taking care of everything to ensure I could concentrate on writing my thesis. To my parents and my siblings; Mum, Dad, Thomas and Lydia for their endless support. Thank you for always being on the other side of the phone when I needed an escape, for reminding me to be proud of all my accomplishments, no matter how small, and for listening to me rattle on about my experiments, even when you didn't understand a word. And finally, for your continued love and belief in my abilities, especially when I myself wasn't sure I could do it.

Thank you.

Author's declaration

I declare that all the material in this thesis, apart from that described below, is a result of my own work. The project was funded by a NERC (Natural Environment Research Council) studentship. The thesis was written by me, with comments from Julia Ferrari. All experiments were planned by Megan Hasoon and Julia Ferrari and were carried out in the lab by Megan Hasoon at the University of York. The genome sequencing was carried out by staff at both the University of York Technology Facility and the University of Liverpool Centre for Genomic Research. This work has not been previously presented for an award at this, or any, other University.

Chapter One
General introduction

Symbiosis can be defined as a close association between two or more different organisms, with endosymbiosis being where one of the organisms lives inside the other. These relationships are extremely common and widespread and are an important part of evolutionary history. For example, it is thought that the mitochondria and the chloroplast, organelles that, to some extent, define eukaryotic life, were formed by an ancient endosymbiotic association (Dimijian 2000, López-García et al. 2017). Although endosymbiotic relationships are ubiquitous, they can have very different evolutionary histories, and varying ecological impacts (Brownlie & Johnson 2009, Feldhaar 2011). We know that harbouring endosymbionts can have a range of effects on their host's biology, however, we are still only beginning to truly understand the depth of these impacts and how these differ across the large variety of host-symbiont relationships we find in nature.

Insect endosymbiosis

One widely studied and constantly expanding field in endosymbiosis is that of insect-microbe interactions. Most insects are infected with at least one symbiont, and these relationships can have a variety of effects on the host (Feldhaar 2011, Hilgenboecker et al. 2008, Moran et al. 2008). In many cases, these symbiotic relationships can be beneficial to the host, aiding with the provision of nutrients, protection from natural enemies and resistance to climate stresses, all of which can contribute to insects' ability to survive in diverse ecological conditions. However, in some cases, these relationships can lead to negative effects including altering the host's reproductive system (Hiroki et al. 2004, Hurst et al. 1999), fecundity costs (Heyworth & Ferrari 2015, Oliver et al. 2006) and reduced longevity (Min & Benzer 1997, Vorburger & Gouskov 2011).

There are two different types of insect endosymbiont, obligate and facultative. Obligate symbionts are those which are required by the host for survival and thus form long-term, stable relationships with the host (Feldhaar 2011, Moran et al. 2008, Oliver et al. 2010). Many of these microbes are intracellular, living in specialised cells called bacteriocytes (Baumann 2005). They are often metabolic partners, ensuring the host receives the required nutrients they may not be able to get from their unbalanced diets (Douglas 1998,

Moran et al. 2008, Zientz et al. 2004). These obligate partners often have extremely small genomes, as they have undergone genome reductions due to their long-term associations, in some cases lasting millions of years, and strict vertical transmission (Moran et al. 2008, Wernegreen 2002). It is thought that around 10% of all insects require an obligate metabolic partner, and as such these relationships have been well studied (Baumann 2005; Douglas 1989, 1998).

Facultative symbionts, however, are not required for host survival and therefore have less specialised relationships than obligate symbionts (Moran et al. 2008, Oliver et al. 2010). They can transmit both vertically and horizontally, but the transmission mode is linked to their function and the dependence of the host on the symbiont (McMeniman et al. 2009). For example, those that are highly beneficial to the host are more likely to transmit vertically than horizontally, although there are some exceptions. Those that tend to transmit vertically, like obligate symbionts, often have smaller genome sizes (McMeniman et al. 2009). Facultative symbionts also often occur at much lower densities than obligates and can live in a variety of different tissues within the host (Oliver et al. 2010). Many effects of facultative symbionts on their host remain unknown, however, the current known effects largely fall into two categories; reproductive manipulators or beneficial mutualists. Reproductive manipulators encourage their vertical transmission by influencing the reproduction of their hosts. This occurs through mechanisms including cytoplasmic incompatibility, parthenogenesis, feminization, and male-killing (Hiroki et al. 2004, Hurst et al. 1999, Mariño et al. 2017, Zchori-Fein et al. 2001). *Wolbachia,* for example, is a well-known and widely studied endosymbiont and is a reproductive parasite of many insect hosts (reviewed in Werren et al. 2008). Beneficial mutualists, unlike reproductive manipulators, instead increase the fitness of their hosts (Feldhaar 2011, Oliver et al. 2010). They mostly are transmitted vertically as they increase the fitness of the host, making them more likely to survive than those without the symbiont. These mutualists can provide a range of benefits to their host such as increased resistance to natural enemies, including parasitoid wasps (Heyworth & Ferrari 2015, Oliver et al. 2003, Xie et al. 2014), fungi

(Cardoza et al. 2006, Heyworth & Ferrari 2015, Łukasik et al. 2013c, Scarborough et al. 2005) and viruses (Hedges et al. 2008), an ability to use different host plants (Tsuchida et al. 2004, 2011) and resistance to climate stresses (Brumin et al. 2011, Heyworth & Ferrari 2015, Montllor et al. 2002). The mechanisms by which many of these effects are conferred, however, are still largely unknown.

Investigating the effects of these symbionts, and the mechanisms by which they occur will give us a greater understanding of the evolution and ecology of these relationships and how they may change over time. This becomes increasingly important when the insect is of economic importance, such as those that pollinate crops or are significant pest species. As many of these facultative relationships aid in host defence, the interactions between external influence, the host and the symbiont can alter both the symbiont evolution and that of the host. For us to truly understand the complexities of these interactions, however, it requires a range of biological approaches. This includes studies into the ecological pressures on these systems and their outcomes but also understanding the genomic and mechanistic underpinnings of the effects we see, and how they may influence the evolution and interactions in these systems. Approaching these questions from a variety of angles can better inform our understanding, especially when the interactions are highly complex.

Insect defence mechanisms

Insects face a range of pathogens and parasites from which they must defend themselves. In recent years, our understanding of how they respond to these threats has significantly grown, including the beneficial outcomes of cooperation between the insect and their symbiotic bacteria (Eleftherianos et al. 2013, Feldhaar & Gross 2008).

The first line of defence for many insects is a physical barrier to infection, namely their cuticle, which protects them from the external environment (Feldhaar & Gross 2008, Vallet-Gely et al. 2008). In addition to this, in some cases, we also see behavioural evasion strategies which can be important for host-pathogen evolution but are often over-looked (De Roode & Lefèvre 2012). The cuticle of the insect provides varying levels of protection

against different environmental threats. For example, bacterial pathogens, without assistance or the presence of a wound, are unable to break through this barrier (Vallet-Gely et al. 2008). Conversely, fungal pathogens and parasitoids both have mechanisms by which they overcome the physical barrier to infection. If the pathogen or parasite can overcome this, they are then faced with the insect immune system, which employs a range of mechanisms to help protect the insect against various invaders. Insects are lacking an adaptive immune system, relying solely on their innate system, our understanding of which largely comes from investigations into the immune system of *Drosophila melanogaster* (Lemaitre & Hoffmann 2007). The innate immune system of insects can be loosely divided into two parts, although the separation between these is not clear cut. One half involves the cellular immune response, which includes hemocytes, phagocytes, and methods such as encapsulation. The other is the humoral response, which involves a range of signalling pathways resulting in the release of a range of effectors such as anti-microbial peptides (Lemaitre & Hoffmann 2007). Although there is no adaptive immune system found in insects, there is increasing evidence that the immune system of insects is more flexible and adaptive than previously thought (Milutinović & Kurtz 2016, Roth et al. 2018). For example, some insects have shown reduced susceptibility to pathogens following previous exposure and, in some cases, this reduced susceptibility can be carried through to the next generation (Milutinović & Kurtz 2016, Roth et al. 2018). This is discussed in further detail later in this chapter.

In recent years, there has been an increased focus on the interaction between the immune system of the insect and their endosymbiotic bacteria (Eleftherianos et al. 2013). Insect endosymbionts can aid in host defence in a range of ways, from employing their own defensive mechanisms which aid the host (Feldhaar 2011), to interactions between the symbionts and the host immune system, which can lead to changes in host immune response (Eleftherianos et al. 2013). Whilst symbionts are able to help the insect by protecting against these immunological stressors, they also need to ensure they are not targeted by the host immune system. This involves a careful balance between symbiont and

host immunity and these interactions can lead to a variety of outcomes in different systems, the details of which are also discussed in more detail later in this chapter. The interaction between different pathogens and parasites, the host and their endosymbionts can have large effects on the evolution of both the host and their symbionts and how they respond to different pathogens and parasites.

Aphids as a model system

Aphids are commonly used as a model system for a variety of research topics. They are ecologically important as they are a major agricultural pest, but they are also an ideal system for investigating a range of biological questions; from understanding polyphenisms and speciation, to endosymbiosis and virus transmission.

There are a large number of aphid species, around 4000 described species thus far, and they are found globally across all temperate regions (Dixon et al. 1987). Aphids are an important agricultural pest, often found in greenhouses and crop fields. It has been calculated that, without the use of insecticides, direct damage by aphids would be responsible for the loss of 8–16% of pea crops, 10–13% of wheat, 8% of barley and 5% of potato in Britain, in addition to damage of other major food crops (Dedryver et al. 2010, Tatchell 1989). This equates to an average loss of £70 million (Tatchell 1989). In addition to direct damage, crop loss can be due to plant viruses of which aphids are the most common vector, transmitting half of the insect-vectored viruses (Nault 1997, Ng & Perry 2004). Vector damage alone can lead to a 6% average loss of sugar beet crops in Britain (Tatchell 1989).

Aphids are a complex system, with a variety of traits that make them the ideal system for a range of biological questions. For example, they are often used as a model to investigate polyphenisms, as aphids express multiple phenotypic morphs of which such a wide range is not often found in other model species. This makes them highly favourable for studies into the developmental processes that allow for a genome to lead to different phenotypes, that may be limited in other systems (Brisson & Stern 2006). As aphids are vectors for many viruses, they are often used to study the transmission of viruses, including aiming to

understand the dynamics and mechanisms of these infections. Aphids, in particular, the pea aphid, have become an important model system for endosymbiosis research. They harbour several bacterial endosymbionts that can have a variety of effects on the host phenotype, which are of both evolutionary and ecological significance. Additionally, over the last two decades, we have seen the publication of the genomes of many aphid species, including the pea aphid (Gauthier et al. 2007), which has allowed for studies into the molecular basis to some of these ecologically relevant traits.

The pea aphid is used as a model system throughout this thesis and, within a laboratory setting, pea aphids have a variety of characteristics that make them a desirable model system. They reproduce through cyclic parthenogenesis and, within controlled long day light conditions, are asexual. They have rapid generation times and can parthenogenetically produce a large number of genetically identical offspring, often over a dozen per day. This leads to their fast population growth. Although different pea aphid biotypes are plant-specific, as seen in other aphid species, all are able to develop and perform well on *Vicia faba* (broad bean) (Ferrari et al. 2008), therefore, within a laboratory setting, pea aphid lines can easily be maintained as stock cultures on broad bean. In addition to a laboratory setting, aphids can also be studied in the field as natural populations, which again makes them a desirable system to work with.

Pea aphid endosymbionts

The pea aphid is a widely studied species in insect endosymbiosis, due to the fact they harbour an obligate symbiont and are often also infected with a number of facultative symbionts. Most aphids rely on their obligate symbiont *Buchnera aphidicola* to survive. *Buchnera* is a nutritional symbiont which aids in providing essential amino acids for the aphid which cannot be obtained from the aphids limited diet of plant phloem (Akman Gündüz & Douglas 2009, Douglas 1998, Moran et al. 2008). This obligate symbiont lives in specialised host cells, called bacteriocytes, and is transmitted vertically (Douglas 1997,

Wilkinson et al. 2003). It is thought that the association between *Buchnera* and aphids began over 200 million years ago (Shigenobu et al. 2000). Due to its strict vertical transmission and long-term close host association, *Buchnera* has an extremely reduced genome of around 641 kb, which is approximately one-seventh of the genome size of *E. coli* (Shigenobu & Wilson 2011, Shigenobu et al. 2000). Further to this, many of the genes retained in the reduced genome are for the biosynthesis of essential amino acids, again highlighting the close association, and co-evolution, of *Buchnera* with the aphid host (Shigenobu et al. 2000).

In addition to their obligate symbiont, pea aphids can also be infected with a number of facultative symbionts, which are not required for survival but often benefit the host (Oliver et al. 2010). There are currently eight known facultative symbionts that infect the pea aphid; *Hamiltonella defensa, Regiella insecticola, Fukatsuia symbiotica* (PAXS; X-type), *Serratia symbiotica, Ricketsiella viridis, Spiroplasma, Rickettsia* and *Wolbachia,* although Wolbachia infections are rare. *Arsenophonus* also infects some aphid species but has yet to be found in the pea aphid (Guo et al. 2017). The traits conferred by these symbionts are relatively well understood, although some of these species are better studied than others. Much of the literature investigating aphid endosymbionts focusses on investigating the effects and phenotypes associated with infection of these facultative symbionts under different environmental conditions, which will be covered in the following section.

Phenotypes associated with aphid facultative symbionts

Many studies investigating aphid facultative symbionts focus on a single ecological benefit associated with a particular symbiont, although the symbiont may confer more than one benefit to their host. Facultative symbionts can benefit their host in several ways, including providing protection against natural enemies and increasing the host tolerance to heat stress. However, the ability of symbionts to protect against these different environmental stressors ranges across symbiont species and strain.

One well-studied benefit is the parasitoid resistance conferred by *Hamiltonella defensa*. Some strains of *H. defensa* carry a phage, APSE, (*Acyrthosiphon pisum* secondary endosymbiont) which provides protection against the parasitoid wasp *Aphidius ervi* (Oliver et al. 2003, 2009; Weldon et al. 2012). Toxins produced by APSE target the parasitoid wasp eggs and prevent development (Oliver et al. 2009). There are several different variants of APSE which encode different toxins and confer different levels of protection against the parasitoid (Rouïl et al. 2020, Weldon et al. 2012). Furthermore, this effect appears to have a level of specificity, with the presence of *H. defensa* and APSE not necessarily leading to the same level of protection against other parasitoid wasp species (Hopper et al. 2018, McLean & Godfray 2015, Oliver & Higashi 2019). The presence of this defensive symbiont also affects the interaction between aphids and parasitoids, with aphids carrying *H. defensa* exhibiting less behavioural evasion tactics and alarm pheromone (Dion et al. 2011a, Oliver et al. 2012). It is thought that parasitoids have also evolved mechanisms to attempt to overcome this protection. They have been seen to exhibit behavioural adaptations, including superparasitising those infected with *H. defensa* (Oliver et al. 2012) or in some cases avoiding oviposition in those harbouring *H. defensa* (Łukasik et al. 2013a). Studies have also shown parasitoids to evolve to increase their infectivity in the presence of a defensive symbiont, after only a few generations (Dion et al. 2011b, Rouchet & Vorburger 2014), although this appears to be symbiont strain specific evolution (Rouchet & Vorburger 2014). Whilst this mechanism of parasitoid protection appears to be specific to *H. defensa*, particularly those with APSE, there are other protective phenotypes conferred by facultative symbionts that are not as specific. *Regiella insecticola*, for example, is well known for its fungal pathogen resistance (Scarborough et al. 2005), however, resistance to fungus has also been seen to be provided by *Fukatsuia* (Heyworth & Ferrari 2015)*, Spiroplasma, Rickettsia* and *Rickettsiella* (Łukasik et al. 2013c). Furthermore, some strains of *R. insecticola* have been seen to protect against the parasitoid *A. ervi*, although this is not due to the presence of APSE (Hansen et al. 2012). The mechanisms behind many of the protective phenotypes conferred by these symbionts are still unclear.

Unlike other species, where the literature has focussed on a single benefit of a symbiont, *Fukatsuia symbiotica* has been seen to provide a variety of benefits to the aphid host. These include protection against a parasitoid wasp, resistance to a fungal pathogen and heat stress (Heyworth & Ferrari 2015). The level of protection, however, varies across the different strains of this symbiont (Smee et al. 2021). Furthermore, there appears to be a large fecundity cost associated with being infected with *F. symbiotica*, which again varies across strains (Heyworth & Ferrari 2015, Smee et al. 2021). Interestingly, we also see a large spectrum of lifestyles associated with different *F. symbiotica* strains, that we do not see in other closely related aphid symbionts. For example, within the pea aphid, some strains are beneficial mutualists, conferring the benefits outlined above, however, there is also a North American strain that appears to confer no benefits to the host, but still leads to a reduction in fecundity (Doremus & Oliver 2017). This strain is often found in co-infections with *H. defensa*, which reduces the costly phenotype. In addition to the pea aphid strains, *F. symbiotica* has also been found in another genus of aphid, *Cinara* (Meseguer et al. 2017). Here, *F. symbiotica* is thought to have a co-obligate relationship with *Buchnera*. It has yet to be investigated whether this strain of *F. symbiotica* confers any of the protective phenotypes or costs we see with infection of the facultative pea aphid strains.

In addition to these defensive traits, it has also been suggested that facultative symbionts may influence how aphids perform on different host plants, although there are contradictory results in this area. There are certain symbiont species that associate with pea aphid clones from particular host plants, for example, *Regiella* is often found in aphids feeding on *Trifolium* spp., and *Hamiltonella* from *Medicago sativa* and *Lotus pedunculatus* (Ferrari et al. 2012). These associations suggest there may be a beneficial association between the symbiont species and host plant use. However, although one study found that the removal of *Regiella* decreased aphid performance on *Trifolium* (Tsuchida et al. 2004), another study found infecting different aphid genotypes with *Regiella* led to only one clone showing an increase in performance, whilst in others, this had no effect (Ferrari et al. 2007). Further, a study investigating both the introduction and removal of *Hamiltonella* provided

no evidence that symbionts can influence host plant use specialisation, although the effects differed across aphid genotype (McLean et al. 2011). This suggests that these associations likely involve an interaction between host and symbiont genotype and the host plant. Alternatively, strong associations between certain symbiont species and host plants may instead be due to tritrophic interactions, such as an aphid's likelihood to experience a particular natural enemy on a given host plant.

These facultative symbionts are usually found at moderate levels within aphids, with multiple infections of two or three different symbiont species being common (Ferrari et al. 2012, Russell et al. 2013). The frequency of multiple infections and symbiont distribution can vary with location and time (Smith et al. 2015), aphid species (Henry et al. 2015), as well as host plant biotype potentially playing a role in symbiont distribution (Ferrari et al. 2012, Leonardo & Muiru 2003, Smith et al. 2015).

Multiple infections, like single infections, are thought to be maintained either due to increased benefit to the host or through the reproductive manipulation of the hosts. Although, maintenance of multiple infections appears to vary across different symbiont groupings, and there is increasing research that suggests co-infections may be maintained due to a variety of factors including levels of protection, the cost to the host, and transmission, which in some cases favour particular combinations (Rock et al. 2018, Weldon et al. 2020). Increased benefit from multiple infections has been seen in aphids, for example, a *Hamiltonella-Serratia* co-infection showed increased protection against parasitoid wasps (Oliver et al. 2006). However, increased benefits from multiple infections are not always the case and there are variable results of co-infections. In some multiple infections, where the symbionts provided the same function, the level of protection conferred did not increase but remained the same as the most protective single infection (McLean et al. 2018). Furthermore, *Hamiltonella-Regiella* co-infections showed decreased parasitoid protection compared to singularly infected *Hamiltonella* or, in some cases maintained the parasitoid protection but led to decreased antifungal defences, depending on the combination of strains. But, no co-infections between these species showed

evidence of defensive benefits from both symbionts, i.e. both parasitoid and pathogen defence (Weldon et al. 2020). Additionally, one study investigated a range of phenotypic outcomes, including pathogen and parasitoid resistance from co-infections between five different strains of *H. defensa* and *F. symbiotica* (Smee et al. 2021*).* They found a large array of phenotypes across the combinations, with some combinations leading to decreased protection from a pathogen or parasite, some being more detrimental than others, whilst others had little to no effect on the levels of protection conferred by the protective symbiont (Smee et al. 2021).

Some multiple infections are more common than others, with double infections of *Regiella* and *Spiroplasma, Hamiltonella and Spiroplasma and Serratia and Rickettsia* being the most common (Frantz et al. 2009). Interestingly, however, there appears to be no increased benefit of these combinations (Łukasik et al. 2013b, Montllor et al. 2002) suggesting that relatedness and competition for resources may also play a role in maintaining these multiple infections. *Hamiltonella, Regiella* and *Fukatsuia* are commonly found in single infections however, co-infections of *Hamiltonella* and *Regiella* and *Regiella* and *Fukatsuia* are relatively uncommon (Ferrari et al. 2012, Russell et al. 2013). Heyworth (2015) found that *Regiella* and *Fukatsuia* were unable to co-exist although a third symbiont, *Spiroplasma*, was unaffected by the multiple infection. This is likely due to the high relatedness of these two species, thus similar niches and more competition for resources and space. Interestingly, however, *Hamiltonella* and *Fukatsuia* can co-exist, despite also being closely related. This then may also be related to the host benefit they provide. *Hamiltonella* is known for its parasitoid resistance whereas, *Regiella* and *Fukatsuia* provide similar benefits to their host, including fungal pathogen resistance and heat tolerance, however, *Fukatsuia* also induces a fecundity cost. Therefore, by harbouring both symbionts, the host may not get any increased benefit but may gain a fitness cost. There is still much to be understood about the interactions between symbionts in multiple infections, including how they communicate and compete, and how multiple infections may influence the phenotypes conferred by different symbiont species.

Invertebrate immune priming

In addition to employing defensive symbionts, insects are also able to increase their resistance to pathogens and parasites through the "priming" of their immune system. Immune priming is the enhancement of an individual's immune response following exposure to pathogens and parasites. It can occur in response to a variety of types of pathogen and parasite and can occur both within and across generations (Milutinović & Kurtz 2016, Roth et al. 2018). Immune priming is widely accepted and understood in vertebrates due to the fact they have both an innate and adaptive immune system. Their innate immune system works as the first line of defence whilst their adaptive immune system is primed through the transfer of antibodies from their mother, often transferred by milk or yolk as they develop (Boulinier & Staszewski 2008, Grindstaff et al. 2003, Hasselquist & Nilsson 2009). Within invertebrates immune priming was, for many years, assumed to not be possible. Invertebrates lack the adaptive immune system found in vertebrates and instead rely solely on protection by their non-specific innate immune system. However, in the last decade, there has been an increase in studies of the innate immune system of invertebrates, and the evidence for invertebrate immune priming has grown (reviewed in Milutinović & Kurtz 2016, Roth et al. 2018). The ability of invertebrates to exhibit immune priming was first shown in *Daphnia magna*, where studies found that the offspring of parents that were exposed to a pathogen had increased fitness when exposed to the same pathogen strain, compared to those who were exposed to a different strain (Little et al. 2003). Maternal exposure to bacterial pathogens has since been seen to increase susceptibility to pathogens in the next generation in several different invertebrate species. This includes immune priming in various insect species, including the bumblebee, *Bombus terrestris* (Sadd et al. 2005) and the mealworm beetle, *Tenebrio molitor* (Moret 2006). Furthermore, investigations into the red flour beetle, *Tribolium castaneum*, showed evidence for both maternal and paternal influence on the immune response of their offspring (Roth et al. 2010). Although it is clear that invertebrates, including insects, have the ability to exhibit immune priming, the mechanisms are still unclear. There does also appear to be a level of specificity to immune priming, with the level of protection following

maternal exposure differing with pathogen and parasite genotype, and host species (Dhinaut et al. 2018, Roth et al. 2018). There may, therefore, be more than one mechanism at work, with different hosts employing different mechanisms for protection.

There is some suggestion that bacterial endosymbionts that infect invertebrates may play a role in immune priming (Eleftherianos et al. 2013). However, much of the work in this area has focussed on priming in the sense that, the presence of the symbiont influences the immune system of the insect, which leads to increased resistance to pathogens and parasites, rather than previous exposure as described above. Endosymbionts, by living within the host, exist under the pressures of an active host immune system. It is likely then that the presence of these symbionts plays an important role in shaping the host immune system and potentially play a role in helping to prime the system, leading to increased resistance to pathogens or parasites. For example, in the mosquito, *Aedes aegypti, Wolbachia* can prime the immune system to protect the host against a range of infections, including bacteria, viruses and parasites (Kambris et al. 2009, Moreira et al. 2009). However, studies investigating *Wolbachia* in *Drosophila* found whilst *Wolbachia* is able to prime the host immune system to better resist RNA virus infections (Teixeira et al. 2008), flies infected with *Wolbachia* do not show any increased resistance to bacterial pathogens (Rottschaefer & Lazzaro 2012, Wong et al. 2011). Although, it has since been shown that *Wolbachia* can provide antibacterial protection in *Drosophila* under more natural routes of pathogen infection (Gupta et al. 2017). Studies carried out thus far suggest that, although there is evidence for endosymbionts being able to prime the immune system against future attacks, this effect appears to differ across host and symbiont genotype, as well as route of pathogen infection and whether the symbiont infection was natural or was introduced experimentally. Our understanding of this symbiont-influenced priming, including the mechanisms by which it occurs, remains incomplete and further work is needed to fully understand these complex interactions.

Aphid immune response

Investigations into the immune response of the pea aphid found that, compared to other insects, the pea aphid appears to have a reduced immune system (Gerardo et al. 2010, Laughton et al. 2011). Genomic analyses of the pea aphid genome found that whilst there are some similarities, such as the presence of heat shock proteins (HSPs) and prophenoloxidase (ProPO), many genes thought to be conserved across insect species appear to be missing from the aphid genome (Gerardo et al. 2010). Firstly, aphids appear to be lacking key components required for microbial recognition. Peptidoglycan receptor proteins (PGRPs), for example, are absent from the pea aphid genome. PGRPs can detect peptidoglycans present in the cell walls of gram-negative and gram-positive bacteria, which then leads to the activation of signalling pathways (Gerardo et al. 2010, Lemaitre & Hoffmann 2007, Steiner 2004). The receptors are highly conserved across both mammals and insects, yet are missing from the aphid immune system (Gerardo et al. 2010, Kaneko & Silverman 2005). Interestingly, gram-negative binding proteins (GNBP) appear to have paralogs in the aphid genome (Gerardo et al. 2010). GNBP are thought to be closely associated with PGRPs, forming a complex for the recognition of gram-positive bacteria and activation of signalling pathways (Gottar et al. 2006, Leulier & Lemaitre 2008, Zheng et al. 2011), so it is unclear why aphids have retained these, whilst the PGRPs are missing. There is suggestion that GNBPs may not be involved in the detection of bacteria in aphids and may instead play a role in fungal recognition (Gerardo et al. 2010).

In addition to recognition genes, aphids also show differences in their immune signalling pathways. In insects, there are four key immune pathways which play a role in the recognition and breakdown of microbial pathogens (Evans et al. 2006, Gerardo et al. 2010, Lemaitre & Hoffmann 2007, Zou et al. 2007). The first, the Toll pathway, plays a role in development as well as immunity (Leulier & Lemaitre 2008). This pathway is present in the aphid genome although it is unclear whether this pathway serves as an immune pathway or functions as a developmental pathway. The Janus kinase/Signal transducers and activators of transcription (JAK/STAT) pathway, another key immune signalling pathway, also has the

potential to play a role in development as well as its immune function (Luo & Dearolf 2001). This pathway, compared to others, is less well understood but is present in the pea aphid genome. The final two pathways, the immunodeficiency (IMD) pathway and the c-Jun N-terminal kinase (JNK) pathway, are somewhat linked. The IMD pathway, involved in fighting invasions of gram-negative bacteria, is missing important components in the aphid genome, including the PGRPs, as described above, as well as crucial signalling genes (Gerardo et al. 2010). These missing genes are conserved across many sequenced insects including flies, bees and beetles (Evans et al. 2006, Lemaitre & Hoffmann 2007, Zou et al. 2007). Although some IMD pathway genes are missing, the pea aphid genome has retained orthologs for the JNK pathway (Gerardo et al. 2010). In *Drosophila*, the IMD pathway activates the JNK pathway, so it is surprising that aphids have one pathway but not the other (Lemaitre & Hoffmann 2007). The JNK pathway has recently been shown to play an important role in the immune response of the pea aphid, including mediating and controlling phagocytosis and PPO activation (Ma et al. 2020). However, in pea aphids, the JNK pathway must rely on another mode of activation as, unlike in other insects, the JNK pathway is not activated by the IMD pathway or Eiger-Wengen (Igaki et al. 2002, Ma et al. 2020).

Antimicrobial peptides (AMPs), often activated by the signalling pathways described above, are produced by most insects in response to invading microbes (Bulet et al. 2004). AMPs are diverse with some being specialised to particular species, whilst others are found across a large variety of species. Interestingly, aphids are missing many of the common AMPs found in other insect species (Gerardo et al. 2010). For example, genes which encode defensins are found in fungi, plants and animals, including in the genome of all other sequenced insect species, but are missing from pea aphids (Gerardo et al. 2010, Zou et al. 2007). Investigations into the cellular responses of the aphid immune system also found no AMP activity within the haemolymph of aphids in response to gram-negative and gram-positive bacteria (Laughton et al. 2011).

Although the genomic investigations into the pea aphid immune system show the absence of many crucial immune pathways and genes, aphids do appear to still be able to exhibit a strong immune response. Cellular investigations found aphids have three types of haemocyte cell, which can respond to and remove invading bacteria (Laughton et al. 2011), although the number of haemocyte cells are reduced compared to other similar organisms (Laughton et al. 2011).

Aphid immune priming

Although there is much evidence for immune priming in a variety of insect species, interestingly aphids have yet to be seen to exhibit immune priming, although the research in this area has been limited.

One study found that maternal exposure to a parasitoid wasp did not increase the resistance to parasitism in the next generation (Vorburger et al. 2008). However, little is known about the immune response to parasitoid infections and, in the case of pea aphids, we know they often rely on the facultative symbiont *H. defensa* for protection against some parasitoids (Oliver et al. 2003, 2005). As a parental immune response is required for immune priming to occur, if this is not exhibited in response to a parasitoid, we would not expect immune priming in this case. A further study investigated within-generation immune priming in response to a bacterial pathogen and found no effect of pre-exposure to later infections with the same pathogen (ter Braak et al. 2013). They also found no effect on immune priming with the presence of facultative symbionts. This study, however, did not investigate these effects across generations. As detailed in the previous section, aphids have a reduced set of immune genes (Gerardo et al. 2010), in particular genes that respond to bacterial infections. Again, it is therefore not unexpected that there was no immune priming in response to these bacterial pathogens.

There is still much that is unknown about the aphid immune system generally, and immune priming in these species. Although thus far it seems immune priming in aphids may be

unlikely, further research is needed to understand the full scope of aphid immunity and priming in response to different pathogens and parasites.

Interactions between the aphid immune system and facultative symbionts

It has been suggested that the loss of many immune genes within the aphid genome is related to their close associations with their facultative symbionts. It may be that their reduced immune function has evolved as aphids have become reliant on their facultative symbionts, however, it could be that this reduced immune system has instead allowed for the invasion of these facultative symbionts. There is likely a complex association between the aphid immune system and the facultative symbionts.

In many other systems, the interactions between the host immune system and their bacterial symbionts leads to changes in host immunity. For example, mosquitoes harbouring *Wolbachia* have increased resistance to a range of infections (Kambris et al. 2009, Moreira et al. 2009), In flies carrying *Wolbachia* however, there appears to be no evidence for resistance to bacterial pathogens (Rottschaefer & Lazzaro 2012, Wong et al. 2011), although there is in response to RNA viruses (Teixeira et al. 2008). This effect, therefore, appears to depend on the host-symbiont combination, as well as the immunological threat. Studies have suggested that the presence of facultative symbionts may influence the aphid immune response, however, results are varied. Some facultative symbionts have been seen to increase the number of immune cells found in the haemolymph of the pea aphid (Laughton et al. 2016). These facultative symbionts were also found within the hemocytes, suggesting this increase in haemocyte number may be related to control of symbiont numbers. Other studies, however, found that the presence of the same facultative symbionts led to a decrease in immune cell numbers (Schmitz et al. 2012). This may have been due to the conditions of the aphid infection, i.e. new versus established infections, or could suggest the interaction between facultative symbiont and aphid immune system is variable across aphid and symbiont genotype. There are also differing effects with different

symbionts. *Serratia*, for example, appears to have no effect on the immune response or in some cases lowers the number of immune cells (Renoz et al. 2015, Schmitz et al. 2012). This symbiont is more recently acquired and thus has a more intact genome than some of the other tested facultative symbionts, such as *H. defensa* and *R. insecticola*, so may have retained some immune evasion mechanisms or an ability to suppress the immune system which may have been lost in the other symbiont species. Although there is some suggestion that the presence of symbionts may be having some effect on the aphid immune response, one study showed no evidence that harbouring symbionts led to increased resistance to bacterial pathogens, as seen in other systems (ter Braak et al. 2013). This could suggest that the facultative symbionts are either not having a large effect on the immune response of the aphid, or that any effect they are having is not necessarily benefitting the host, at least in response to bacterial pathogens. More studies are required to understand the full scope of the interaction between the aphid immune system and their facultative symbionts, including how this may differ across aphid-symbiont combinations and whether these interactions alter the response to different immunological threats.

Purpose of thesis

The remainder of this thesis aims to investigate the facultative symbionts of the pea aphid and their interactions with the host. It largely focusses on the defensive traits conferred by the symbionts, the mechanisms behind these and how harbouring a symbiont may influence the host immune system. To do this, I first investigated how these symbiotic relationships may influence defence against natural enemies across multiple generations, and how interactions between the symbiont and the host immune system may also play a role in their defence against natural enemies. I also investigated the genomes of two common facultative symbiont species, exploring within and across symbiont species differences, and highlighting potential mechanisms for some of their beneficial and costly phenotypes.

In Chapter two I investigate the role of symbionts in transgenerational immune priming in aphids in response to a fungal pathogen. Given that aphids reproduce parthenogenetically and have telescopic generations, it would be likely that trans-generational effects such as immune priming would have evolved in aphids. However, this does not appear to be the case (ter Braak et al. 2013, Vorburger et al. 2008). This is likely due to the reduced immune gene repertoire of aphids. Aphids instead often rely on their facultative symbionts for protection against pathogens and parasites, however, what is still unknown is whether maternal exposure to natural enemies can influence the resistance in the presence of a symbiont. Using aphids infected with *H. defensa* and *F. symbiotica*, I investigated the susceptibility to a fungal pathogen over two generations.

I found that when mothers were exposed to a fungal pathogen there was a reduction in susceptibility in the second generation. Although, this only occurred in those harbouring a facultative symbiont, with no change in those without facultative symbionts, suggesting that the transgenerational effect seen here was mediated by the presence of a symbiont. We found that harbouring a facultative symbiont led to the aphids being more susceptible in the first generation, but by the second generation the susceptibility decreased to a level similar to that of the symbiont free individuals. Therefore, although we saw a decrease in susceptibility in the facultative symbiont harbouring individuals in the second generation, susceptibility was not reduced compared to the symbiont-free controls and was instead due to individuals no longer having the increased susceptibility seen in the first generation. I hypothesised that this effect was due to an interaction between the symbionts and the immune system of the aphid, leading to changes in immune response to the pathogen over the two generations. In addition to the phenotypic responses to the pathogen, I also investigated the densities of the symbionts over the first two days of infection and across generations. Here, we found that there was an increase in the densities of the symbiont in generation two, with the densities of *H. defensa* being significantly increased, further suggesting an interaction between the host immune system and these facultative symbionts. Overall, the results suggest there is a transgenerational effect occurring in

response to a fungal pathogen infection which is likely due to an interaction between the facultative symbionts and the aphid immune system.

Chapter three explores the interaction between facultative symbionts and the host immune system. There is evidence in several insect species that harbouring bacterial endosymbionts can alter the host's immune response, in many cases by priming the immune system, allowing for increased resistance to invading pathogens. Aphids are a unique system in which to explore this interaction, as they harbour a number of different facultative symbionts, which protect them against a range of natural enemies, but also have a reduced immune system compared to many other insects. Here, I investigated the effect of two different symbiont species on the host immune response, and whether the presence of these facultative symbionts altered the host's response to an invading pathogen.

I found that the presence of a symbiont, without immunological stress, does not appear to alter the cellular immunity of the aphid. Although on infection with a pathogen, the cellular immune response was lessened when individuals harboured a facultative symbiont, particularly when harbouring one which has previously been seen to defend against fungal pathogens. Within the humoral response, investigated using the expression of a candidate set of immune genes, I found that harbouring a facultative symbiont did not lead to changes in the immune response to the fungal pathogen. However, I did find that aphids harbouring facultative symbionts appeared to have an overall higher expression of three out of the six immune genes tested, compared to those without. This suggests that harbouring a facultative symbiont does have some effect on the host response, although, as this was seen in treatments that were highly susceptible to the pathogen, it does not appear to be priming the host against attack, as in other systems. This further highlights the complexity of this system and the interactions between the host and its bacterial endosymbionts.

Chapter four reports the genomic analysis of novel genomes of five *Fukatsuia symbiotica* isolates. *F. symbiotica* has a wide variety of lifestyles within different aphids. Within UK pea aphids, *F. symbiotica* acts a facultative mutualist (Heyworth & Ferrari 2015) however in North American pea aphids, *F. symbiotica* confers no known benefits to its host but is costly to carry, acting almost like a parasite in this case (Doremus & Oliver 2017). Furthermore, within the genus *Cinara*, *F. symbiotica* is thought to be a co-obligate symbiont of *Buchnera*, assisting with metabolic processes. In this chapter, I built and analysed the draft genomes of five facultative mutualist isolates of *F. symbiotica*. I then compared these to the published genome of the North American and *Cinara*-associated strains, with the aim of understanding the genome evolution and transition from parasite to mutualist within this species. I also investigated the differences between the five novel isolates, for which I had phenotypic information, to explore potential mechanisms for the beneficial traits conferred by these strains. Further to this, all strains were compared to two other closely related aphid facultative symbionts, *Hamiltonella* and *Regiella*.

In this chapter, I outline the differences in genome size and metabolic abilities within and across species of symbiont, as well as describing the pathogenic factors retained or lost from the various genomes and the differences in their mobile DNA composition. Overall, certain aspects of the genomes of *F. symbiotica*, within the pea aphid, including the genome size and retention of toxin genes and metabolic pathways, suggest this symbiont is at a mid-point between its free-living relatives and other closely related pea aphid symbiont species, potentially due to a less close association with its host. Within *Cinara*, although the *F. symbiotica* genome size is reduced compared to the other strains, there are few large differences in gene content across the genomes, suggesting that the transition to co-obligate may be recent. However, differences in mobile DNA across these genomes were in line with what would be expected given their different lifestyles. Within this chapter, I also highlighted differences between isolates that provide different levels of protection against natural enemies, outlining potential mechanisms for some of these phenotypes which would benefit from further exploration.

Chapter five investigates novel genomes of five isolates of *Hamiltonella defensa* with known phenotypic effects and differing levels of parasitoid protection. I built and annotated the genomes of these five isolates using de-novo methods. The genomes were then compared to each other and previously published genomes of *H. defensa*, using a variety of comparative genomics techniques.

The five novel genomes fell into two phylogenetically distinct groups. Previous investigations into *H. defensa* genomes found that the majority of differences were across clades but overall, they showed high similarity. As in previous studies, I found large differences across clades, but also slightly more variability than expected across isolates that fell within the same clade, primarily within their mobile DNA composition. Overall, although these isolates showed high similarity, there were notable differences in their genome composition as these strains have diverged onto different evolutionary trajectories.

Chapter Two

Symbiont-mediated maternal effects on pathogen resistance

Abstract

Transgenerational immune priming, where parental exposure to an immunological threat leads to the enhancement of the offspring immune response, is well recognised in insects and often occurs in response to pathogens and parasites. Aphids' parthenogenetic reproduction and telescoping of generations (two generations of embryos, granddaughters within daughters, develop within an aphid mother) are traits that would make transgenerational effects likely to occur. However, aphids have a reduced immune system compared to that of other insects, thus transgenerational immune priming may be unlikely. Many aphids harbour facultative symbionts, which are not required for survival but can benefit the host. Aphids often rely on these symbionts to resist natural enemies, but it is unknown whether the maternal environment affects the resistance conferred by facultative symbionts.

Here, we tested the role of facultative symbionts in influencing offspring susceptibility to a fungal pathogen following maternal exposure, in the pea aphid *Acyrthosiphon pisum*. To do this, we infected aphids harbouring different facultative symbionts, and facultative symbiont free aphids, with a fungal pathogen over two generations. We found that harbouring a facultative symbiont in the first generation made individuals more susceptible to the fungal pathogen than those without. However, in the second generation, there was a decrease in susceptibility in aphids harbouring facultative symbionts but no change in susceptibility or survival in the facultative symbiont free aphids.

These results suggest there is a transgenerational effect occurring in response to this fungal pathogen, which is mediated by the presence of a facultative symbiont. The transgenerational effect seen here is likely due to an interaction between the facultative symbionts and the aphid immune system.

Introduction

An individual's phenotype is influenced not just by genotype but also by a range of environmental factors, which can lead to a variety of plastic responses (Bonduriansky 2012, Bonduriansky & Day 2009). In addition to the influence from the current environment of the individual, its phenotype can also be influenced by the parental environment, which can be a large driver of offspring phenotype (Bonduriansky 2012, Mousseau & Fox 1998a, Uller 2008). Transgenerational immune priming (TGIP) is a type of parental-influenced plasticity which involves the enhancement of the offspring immune defence, based on the parental experiences (Roth et al. 2018). This effect is seen across a wide variety of taxa, including both vertebrate and invertebrates. The study of TGIP, often referred to as immune memory, has been well studied and is widely accepted within vertebrates (reviewed in Roth et al. 2018). Vertebrates have both an innate immune response, their non-specific first line of defence, and a targeted adaptive immune system which can be enhanced by the transfer of maternal antibodies, through milk, and yolk whilst developing (Boulinier & Staszewski 2008, Grindstaff et al. 2003, Hasselquist & Nilsson 2009). Invertebrates, however, lack this adaptive immune system and rely completely on their innate immune system. For this reason, it was assumed for many years that invertebrates lacked the ability to form any immune memory (Little & Kraaijeveld 2004).

Over the last two decades, however, the evidence for immune memory has grown. It has shown that the innate immunity of invertebrates has evolved to involve immune memory (Milutinović & Kurtz 2016, Roth et al. 2018), which can transfer across generations (Roth et al. 2018). The recent surge in studies into invertebrate TGIP began with investigations into the crustacean, *Daphnia magna* (Little et al. 2003), as well as some insect species such as the bumblebee, *Bombus terrestris* (Sadd et al. 2005), and the mealworm beetle, *Tenebrio molitor* (Moret 2006). These studies showed that antibacterial responses can be enhanced following maternal exposure to bacterial pathogens, and has since been shown in a range of invertebrate taxa (Freitak et al. 2009, Kurtz & Franz 2003, Roth et al. 2009, Sadd & Schmid-Hempel 2006). Further to this, studies into the red flour beetle, *Tribolium castaneum,*

showed that there is potential for both maternal and paternal influence on immune priming (Roth et al. 2010). Although it is now widely accepted that there is TGIP within invertebrates, the mechanism behind this effect is still unclear. Additionally, the extent of the protection provided by immune priming varies with pathogen type and host species, suggesting a level of specificity to TGIP and potentially more than one mechanism at work (Dhinaut et al. 2018, Roth et al. 2018).

Aphids, when compared to other insect species, have a reduced immune system and thus immune response (Gerardo et al. 2010). There are thought to be four key immune pathways that are conserved across all insects; the Toll, immunodeficiency (IMD), c-Jun N-terminal kinase (JNK), and Janus kinase/Signal transducers and activators of transcription (JAK/STAT) pathways (Evans et al. 2006, Gerardo et al. 2010, Lemaitre & Hoffmann 2007, Zou et al. 2007). These play a major role in protecting against microbial pathogens, working to both recognise and breakdown the pathogens. These signalling pathways respond to different invading pathogens and often lead to the production of antimicrobial peptides (AMPs) (Dionne & Schneider 2008, Gerardo et al. 2010).

Investigations into the aphid genome found that aphids are lacking some of the key features of the immune system thought to be conserved across all insect species (Gerardo et al. 2010). Of the four conserved pathways, three are thought to also play a role in development in addition to immune function (Gerardo et al. 2010, Lemaitre & Hoffmann 2007, Zou et al. 2007). These three are still found in the aphid genome, however, the IMD pathway which is used to fight invasions of gram-negative bacteria is missing (Gerardo et al. 2010). In addition to the signalling pathways, aphids also have large differences in the recognition genes found in their genome, compared to other insects. Peptidoglycan receptor proteins (PGRP), for example, activate signalling pathways by recognising peptidoglycan in the cell walls of gram-positive and negative bacteria, however, the aphid genome is lacking PGRPs which are conserved across both mammals and insects (Gerardo et al. 2010, Steiner 2004). AMPs are also an important part of the innate immune response, forming the first line of defence in most insects following microbial invasion. Despite this,

aphids are missing many AMPs that are found commonly in other insects (Gerardo et al. 2010). For example, one type of AMP, defensins, have been found in all other annotated insect genomes yet has been lost in aphids (Gerardo et al. 2010, Zou et al. 2007). Although the aphid genome is lacking many important factors, aphids are still able to exhibit an immune response against invading pathogens and parasites (Laughton et al. 2011). However, unlike other insects, aphids appear to show no TGIP, though research in this area is limited (ter Braak et al. 2013, Vorburger et al. 2008). TGIP relies on the mothers environment being an accurate predictor of the environment that their offspring will experience (Roth et al. 2018). Aphids' parthenogenetic reproduction and telescoping of generations make mothers likely to share their environment with their offspring, and thus a good predictor of the pathogens and parasites their offspring may encounter. It is surprising then that these transgenerational effects do not occur in aphids.

Aphids often carry one, or more, gram-negative bacterial symbiont which live extracellularly within the aphid host, many of which protect the aphid against various pathogens and parasites (Feldhaar 2011, Moran et al. 2008, Oliver et al. 2010). It is thought that their reduced immune system may be linked to this close association, although it is unclear whether the association with these facultative symbionts has led to the loss of function in immune pathways, or whether their reduced response to gram-negative bacteria has allowed for the easy invasion of these symbionts (Gerardo et al. 2010). Aphid facultative symbionts help protect against a range of environmental stressors. For example, *Hamiltonella defensa* is well known for its ability to protect against the parasitoid wasp *Aphidius ervi* (Oliver et al. 2003, 2005), and *Regiella insecticola* against the fungal pathogen *Pandora neoaphidis* (Łukasik et al. 2013c, Scarborough et al. 2005). *Fukatsuia symbiotica*, another common symbiont of the pea aphid, can protect against both the parasitoid and the fungal pathogen when in a co-infection with *Spiroplasma* (Heyworth & Ferrari 2015), although the levels of protection vary across isolates (Smee et al. 2021). There is a good understanding of the mechanism for parasitoid protection provided by *H. defensa*, but

otherwise the mechanisms behind some of these protective phenotypes, including fungal pathogen resistance, are still unclear.

Given that aphids rely on their symbionts for protection rather than classic modes of immune defence, we might expect that that maternal exposure may influence symbiont-conferred resistance in a similar manner to TGIP. Currently, it is unknown whether the maternal environment affects resistance conferred by symbionts in insects. There is some evidence that bacteria may be able to mediate transgenerational effects, as one study in *Lepidoptera* found evidence that maternal transfer of bacteria to developing eggs led to differential expression in immune-related genes (Freitak et al. 2014). However, this involved the transfer of ingested *E. coli*, not a protective symbiont associated with an insect host as investigated here.

Here, we aimed to understand what role aphid facultative symbionts play in influencing offspring susceptibility to a fungal pathogen, following maternal exposure. We also aimed to investigate the effect the pathogen infection has on symbiont densities, across and within generations, to understand how the fungal pathogen and symbionts interact over the period of infection, and whether the infection may lead to changes across generations. To do this, we infected aphids harbouring different facultative symbionts, and facultative symbiont free aphids, with a fungal pathogen over two generations. We hypothesised that harbouring a facultative symbiont would lead to changes in susceptibility across generations, with offspring of mothers who were exposed to the pathogen, having decreased susceptibility. We did not expect to see any decrease in susceptibility in the facultative symbiont free aphids, as aphids have thus far been shown to not have TGIP.

Methods

Aphids

The experiment was carried out using five different experimental lines, which were all the same aphid genotype but harboured different facultative symbionts. The five lines included one which was carrying *Fukatsuia symbiotica* isolate F3, two *Hamiltonella defensa* lines, H1 and H3 and a co-infected line which was carrying both *H. defensa* isolate H3 and *F. symbiotica* isolate F3. We also tested a symbiont free line, which were aphids harbouring no facultative symbionts.

The experiment was run with *F. symbiotica* and *H. defensa* single and co-infected aphids to test if any transgenerational effect may be altered by harbouring different symbiont species in a single or co-infection. We also used a symbiont free aphid line which would allow us to see if any TGIP was occurring in response to this fungal pathogen and to observe how well facultative symbiont free aphids were able to resist this pathogen.

Experimental design

To produce the experimental aphids, adult aphids were first placed onto four 2-week-old *Vicia faba* (var. The Sutton) plants enclosed in a plastic, vented 2 L cage. They were left overnight to reproduce and removed the following day. Their offspring were then left to mature and, on reaching nine-to-ten days old, 20 individuals from each line were randomly selected. These were then infected with the fungal pathogen, *Lecanicillium lecanii* (identified based on morphology). The infection was carried out by exposing individuals to a pair of adult aphid cadavers on which the fungus was sporulating. To infect the individuals, the cadavers of the fungus-infected aphids were left on 1% agar overnight at 20°C in a high humidity environment to start sporulation. These were then put on wet filter paper placed over a small cylindrical glass vial (height = 35 mm, diameter = 20 mm) creating a "spore shower" over the 20 individuals. The infection process lasted 90 minutes, during which the sporulating cadavers were rotated among all replicates within a block, including across treatments, to ensure each group was exposed to an equal amount of fungal spores.

Following the infection, the individuals were separated into two groups; ten individuals were taken for qPCR at subsequent timepoints, and ten for phenotype data to investigate susceptibility. For the qPCR timepoints, all individuals were placed onto a petri dish containing a *V. faba* leaf with the stem inserted into 2% agar to keep the leaves fresh. All petri dishes contained leaves of similar size. One individual was taken at 3 hours and 48 hours post-infection (Figure 2-1) and stored at -20°C. Individuals for the phenotype data were placed onto individual petri dishes to allow us to track which offspring were produced by which mother. These were left for two days to reproduce (average number of offspring produced=13.5±0.2). The offspring were then separated onto a different petri dish, noting which aphid was their mother. These offspring would become the second-generation individuals. Seven days following the offspring separation (nine days post-infection), mothers were marked as either "alive", having survived the fungal pathogen infection, "sporulating", where the fungus had been successful in its infection, or "dead", in which the aphid had died, however there was no sporulating phenotype, suggesting this may not have necessarily been due to the fungus.

The offspring of these individuals were then separated into three generation 2 treatments, depending on the outcome of their mother's infection in the previous generation: "Mother Alive", "Mother Dead", "Mother Sporulated". On reaching nine days old these were infected with the pathogen, following the same protocol as in the previous generation. Again, following infection, these individuals were separated into two groups. Those for the qPCR data were placed on dishes and an individual taken at the appropriate timepoints. The other individuals were placed onto 2-week-old *V. faba* plants in vented 2 L cages and nine days post-infection were marked as either alive, dead, sporulating, or missing. Often when aphids are kept on plants, some individuals cannot be found and, as it is impossible to determine what has happened to these, they were marked as "missing".

There were eight biological replicates over four experimental blocks. In the first generation, the first two blocks consisted of 10 individuals per replicate, which was increased to 20 for the second two blocks to allow for more offspring to be produced for second generation

individuals. The number of generation 2 individuals and replicates varied as these were dependent on the outcome of the first-generation individuals (see Table 2-1).

Table 2-1: Number of total individuals and number of replicates for each treatment and aphid line

The number of individuals and number of replicates for the generation 2 treatments are reduced as these relied on the outcome of the fungal infection in the previous generation.

	Symbiont free		*F. symbiotica* 3		*H. defensa* 1		*H. defensa* 3		Co-infected (H3 + F3)	
	No. of individuals	No. of replicates	No. of individuals	No. of replicates	No. of individuals	No. of replicates	No. of individuals	No. of replicates	No. of individuals	No. of replicates
Gen 1	120	8	120	8	120	8	120	8	120	8
Gen 2 (Mother Alive)	73	6	47	5	19	1	10	1	41	4
Gen2 (Mother Dead)	52	5	69	6	64	6	45	5	68	6
Gen 2 (Mother Sporulated)	61	6	77	8	86	8	45	7	68	6

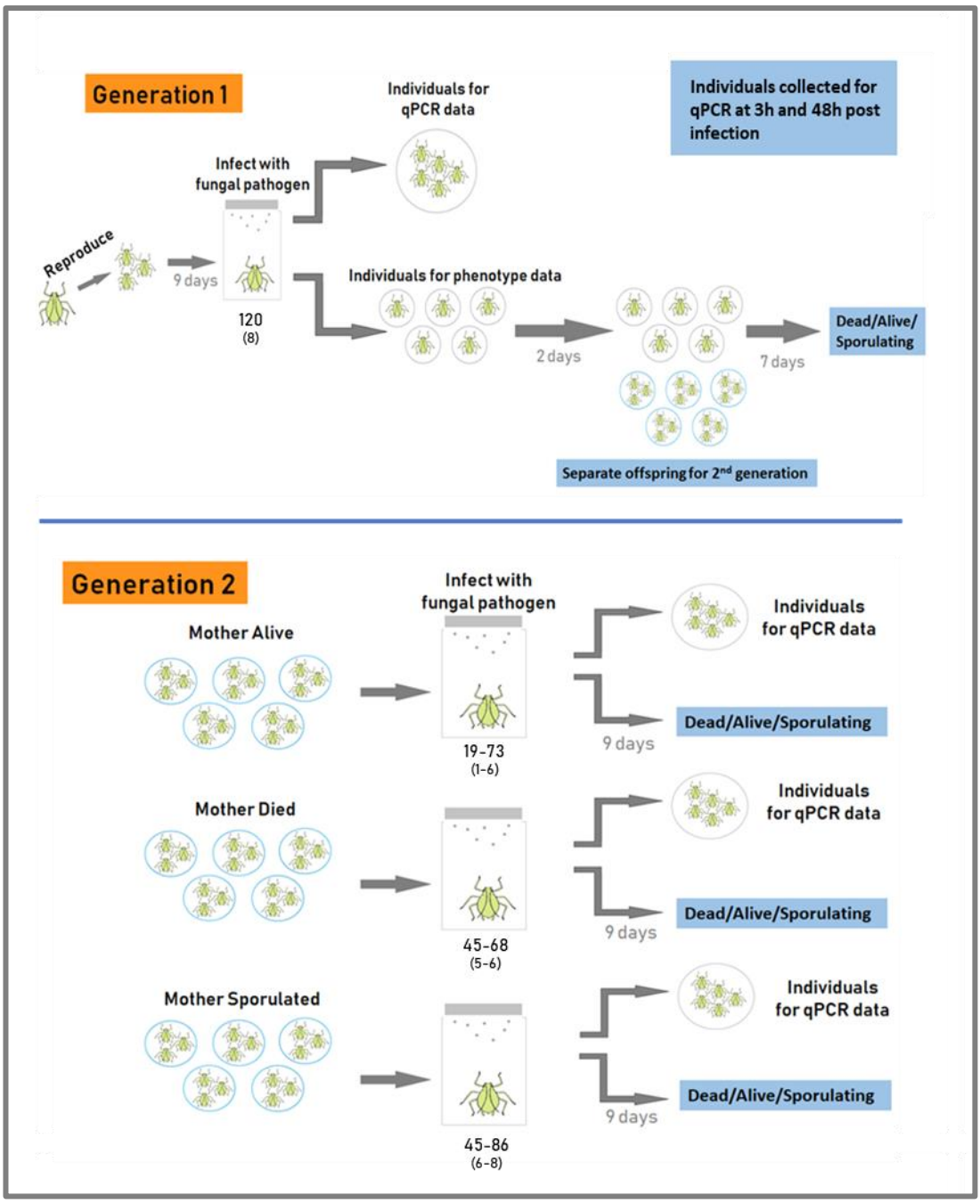


Figure 2-1: Experimental design
For each line, generation 1 individuals were exposed to a fungal pathogen. Following this, they were separated into two data collection types. For the qPCR data, an individual was taken at two timepoints, 3 hours and 48 hours post-infection. For the phenotype data, individuals were placed on separate plates, to record which offspring were produced by which mother. Two days post-infection offspring were separated from their mothers. Nine days post-infection the phenotypes of the individuals were noted. The offspring of these mothers (separated at day two) were divided into three treatments dependent on their mother's phenotype at nine days post-infection. They were infected with the pathogen as per the first generation and again individuals for qPCR data were collected and phenotype noted at nine days post-infection.
Numbers noted below the fungal infection indicate the total number of infected individuals and number of replicates across all symbiont lines. Number of individuals and replicates varied in generation 2 as these relied on the outcome of the first generation. Specific number of individuals and replicates for each symbiont line can be found in Table 2-1.

qPCR protocol

DNA was first extracted from each aphid sample. Aphids were homogenised in 200 µl of 5 % Chelex solution followed by the addition of 10 µl of proteinase K (10 mg/ml). The samples were then incubated at 56°C for four hours, followed by a further eight-minute incubation at 100°C. Following incubation, the samples were centrifuged at 16,000 xg for three minutes and the supernatant, which contained the DNA, pipetted into a 1.5 ml tube. The samples were then stored at -20°C until required.

We ran qPCR to detect the densities of the facultative symbionts in the aphids across the different timepoints and generations. The samples were run on an Applied Biosystems QuantStudio™ 3 Real-Time PCR System using SYBR® Green reagent. Within each well, there was 10 µl of Fast SYBR 2x master mix, 1 µl of forward primer (7 µM), 1 µl of reverse primer (7 µM), 5 µl of ultrapure water and 3 µl of the DNA sample. This was carried out using specific primers for both *H. defensa* and *F. symbiotica* (see Table 2-2 for sequences). The qPCR cycling conditions were 95°C for 20 seconds, then 40 cycles of 95°C for 1 second and 60°C for 20 seconds. A melt curve was also generated for each plate. The cycling conditions for the melt curve were 95°C for 1 second, 60°C for 20 seconds and then a gradual increase at 0.1°C/s to 95°C.

Each plate was analysed using the online Applied Biosystems™ Analysis Software using a standard curve of known concentrations. For the standard curve, the DNA concentrations were first calculated from control samples using an Agilent 2100 Bioanalyzer. The control samples (concentrations: *H. defensa*: 11.3 pmol/l, *F. symbiotica*: 11.5 pmol/l) were then diluted serially 1:10 with distilled water to create a 5-sample curve. There were four biological replicates for each treatment and the samples were run in technical triplicates.

The concentrations of each sample were calculated by comparing the Ct values to the equation of the standard curve. Ct values above that of the negative control were discarded. The concentrations of the sample were then standardised using the total DNA concentration for each sample, to allow for differences in aphid size and DNA extraction efficiency. From these values, the DNA copy number for each sample was then calculated.

Table 2-2: qPCR primers
Specific symbiont primer sequences for *F. symbiotica* and *H. defensa*. Primer sequences were taken from Heyworth (2015).

qPCR primers		
Symbiont	Primer Sequences	
Fukatsuia	Forward: gyrB-X397	TGGATTGGCTGGTGAAAGAAT
	Reverse: gyrb-X466	TTCATCTCTCCCAAACCTTTATAG
Hamiltonella	Forward: H1165	CGTGAAATGACAAGACGTAAAGGT
	Reverse: H1241	TCACGCTCCTGGCAATCC

Statistical analysis

Each of the four potential outcomes of the experiment (alive, dead, sporulated, missing) were analysed separately, using the proportion of individuals for each outcome. The missing individuals were analysed first using an analysis of variance (ANOVA) with maternal outcome and symbiont line as explanatory factors. These individuals were then removed from the data set and new proportions calculated for the remaining three outcomes (alive, dead, sporulated). The proportion of individuals from each outcome were arcsine square-root transformed before analysis.

The proportion of sporulating, alive and dead individuals were analysed using a linear mixed-effects model. The proportion of individuals for each outcome was the response variable. Generation, symbiont, maternal outcome, and their interactions were fitted as fixed effects, and block and strain were fitted as random effects. The data was analysed using the lme4 (Bates et al. 2015) and lmerTest (Kuznetsova et al. 2017) packages. For variables with more than two levels or significant interactions, post-hoc tests were performed using the lsmeans package (Lenth 2016). The analysis of the missing individuals

was carried out at the symbiont strain level, but as there was no difference between the outcome of individuals from the two strains of *H. defensa* (H1 and H3), we combined these and presented this data at the species level. For our second analysis strain was included into the model as a random effect, but again the two strains did not differ, and as such the data was presented at the species level.

The densities of facultative symbionts at three timepoints across the two generations were analysed using an analysis of variance (ANOVA). The four symbiont treatments (*F. symbiotica* single infection, *H. defensa* single infection, *F. symbiotica* co-infection, *H. defensa* co-infection) werse analysed together with copy number as the response variable and timepoint, symbiont, and their interaction as the explanatory variables. All analyses were carried out in R, version 3.6.1 (R Core Team 2013) and graphs were made using the ggplot2 package (Wickham 2016).

Results

Transgenerational effect of fungal pathogen infection

We first analysed the number of missing individuals to ensure these would not bias the comparison across the different treatments. We found a significant difference in the proportion of missing individuals across maternal outcomes ($F_{3,116}$=19.88, p<0.001), however, this was entirely due to the difference between generation 1 and generation 2 mothers (Figure 2-2D). This was expected as the generation 1 aphids were kept on plates, and in generation 2 they were kept on plants, only allowing for the opportunity for missing individuals in the second generation. We found no significant difference between the proportion of missing individuals across the four symbiont lines ($F_{3,116}$=0.85, p=0.469). Following this, missing individuals were removed from the data reducing the overall number of individuals per treatment (Table 2-3), and new proportions for sporulating, alive and non-sporulating dead individuals calculated (Figure 2-2A-C).

Table 2-3: The total number of individuals across all replicates before and after the removal of missing individuals. Number of replicates can be found in Table 2-1.

	Symbiont free		*F. symbiotica* 3		*H. defensa* 1		*H. defensa* 3		Co-infected (H3 + F3)	
	Total no. of individuals	Individuals minus "missing"	Total no. of individuals	Individuals minus "missing"	Total no. of individuals	Individuals minus "missing"	Total no. of individuals	Individuals minus "missing"	Total no. of individuals	Individuals minus "missing"
Gen 1	120	120	120	120	120	120	120	120	120	119
Gen 2 (Mother Alive)	73	60	47	32	19	18	10	10	41	30
Gen2 (Mother Dead)	52	43	69	61	64	49	45	33	68	52
Gen 2 (Mother Sporulated)	61	61	77	67	86	67	45	39	68	40

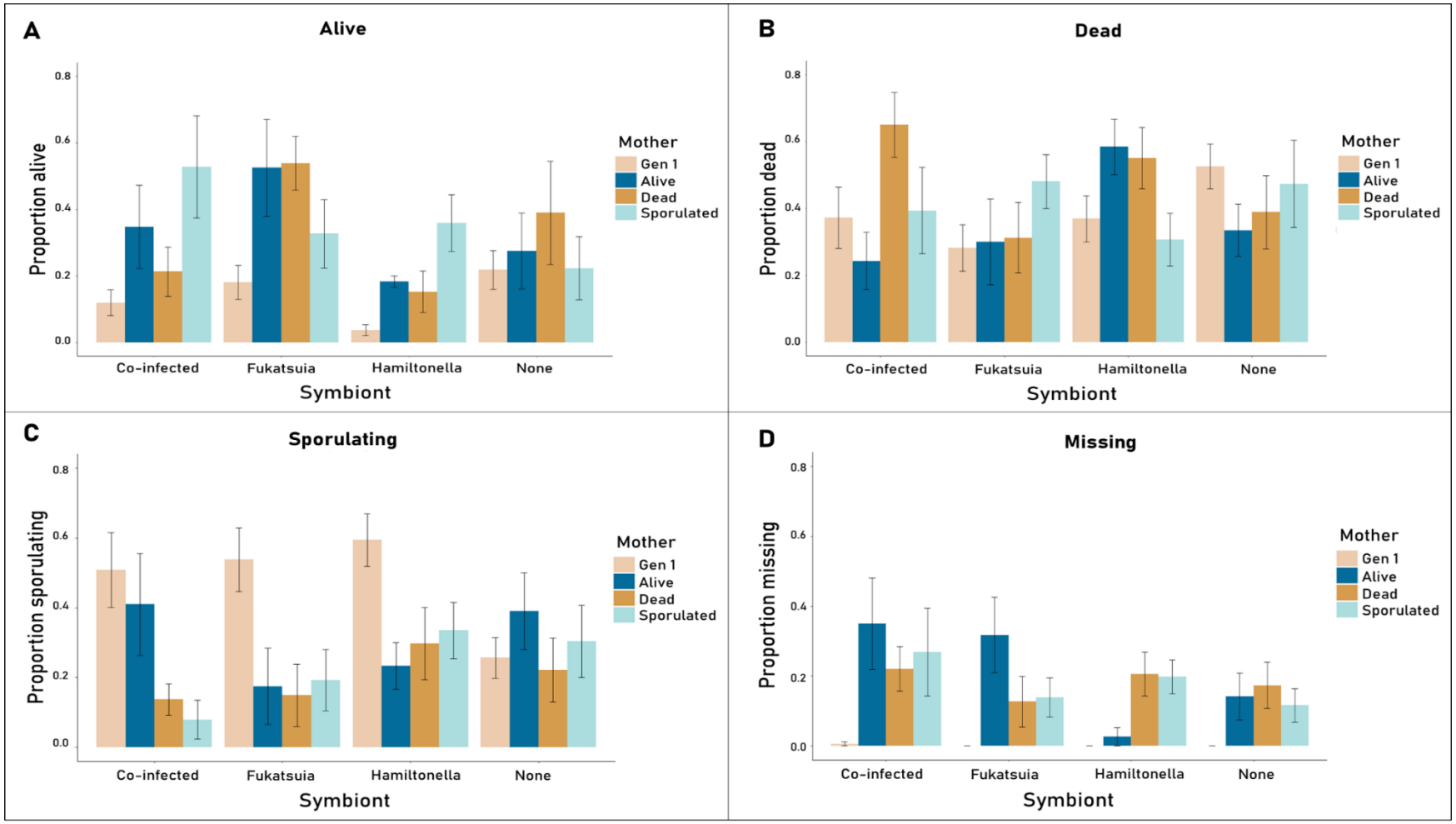


Figure 2-2: Proportion of individuals from each maternal outcome, across four aphid symbiont treatments
A) Proportion alive B) Proportion dead, C) Proportion sporulating, D) Proportion missing.
Maternal outcomes: Generation 1=peach, Alive=dark blue, Dead=gold, Sporulated=light blue. Error bars represent the standard error.
The proportion of missing individuals was calculated from the total number of individuals across all replicates. The proportions of alive, dead and sporulating individuals were calculated from the total number of individuals after the deduction of missing individuals (Table 2-3).

Firstly, we found that there was a significant difference in the overall proportion of individuals surviving the fungal pathogen infection across the different symbiont lines ($F_{3,112}$=3.28, p=0.024; Figure 2-3), with those harbouring *F. symbiotica* in a single infection having significantly higher survival than those harbouring *H. defensa* (p=0.016; Figure 2-3). However, there was no significant difference in the overall proportion of sporulating ($F_{3,112}$=1.40, p=0.247), or non-sporulating dead individuals ($F_{3,112}$=0.76, p=0.518), across the four symbiont lines.

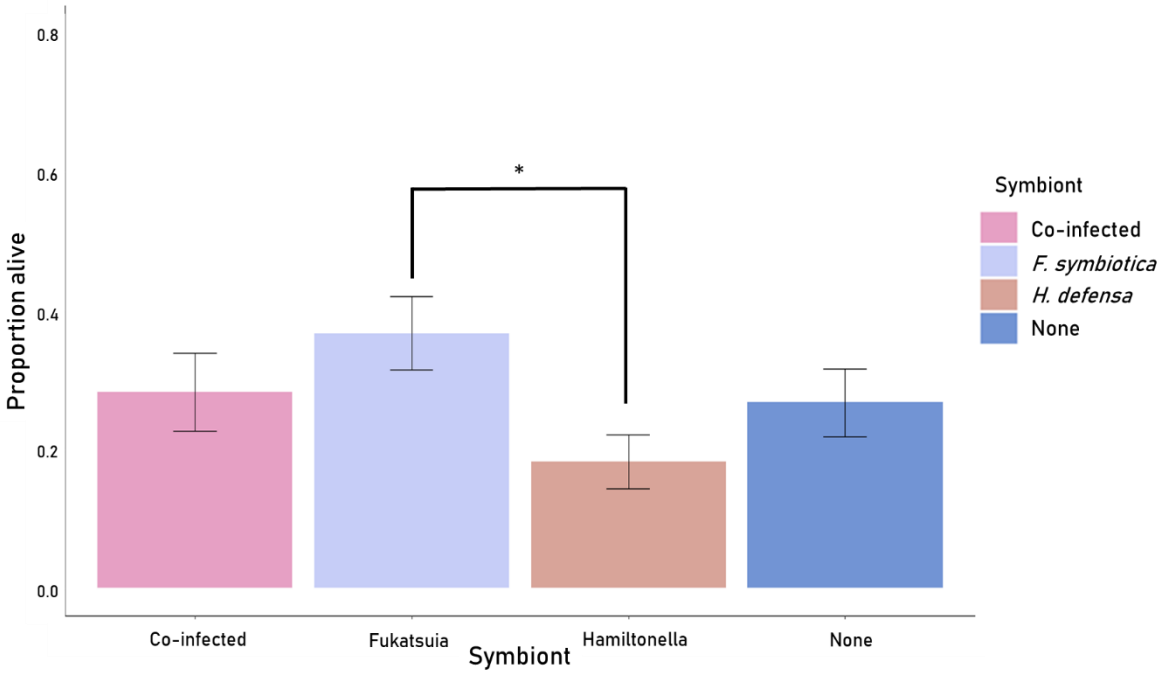


Figure 2-3: Proportion of surviving individuals over both generations from each aphid symbiont line
Symbiont lines: Co-infected=pink, *F. symbiotica*=light blue, *H. defensa*=brown, None=dark blue
Error bars represent the standard error. Asterisks show significant difference in proportion of surviving individuals across aphid symbiont treatments (*: p<0.05).

We tested whether maternal exposure to a fungal pathogen led to a decrease in susceptibility in the second generation. We found there was a significant decrease in the proportion of sporulating individuals in the second generation ($F_{1,116}$=19.3, p<0.001; Figure 2-4A) and a significant increase in survival ($F_{1,116}$=17.65, p<0.001; Figure 2-4B), showing that, on average, there was a transgenerational effect of maternal pathogen exposure. There was no significant difference in the proportion of individuals that died, but did not produce spores, between the two generations ($F_{1,116}$=0.01, p=0.91).

There was also a significant interaction between generation and symbiont line ($F_{3,106}$=2.69, p=0.049), as the decrease in sporulation differed across the different symbiont lines. Those harbouring *F. symbiotica* in a single infection showed a significant decrease in sporulation in the second generation (p=0.020; Figure 2-4A). There was a trend towards a decrease in sporulation in the *H. defensa* and the co-infected line (Figure 2-4A), although this was not statistically significant (*H. defensa*: p=0.087; Coinfected: p=0.097). In contrast, there was no change in the proportion of sporulating individuals from the facultative symbiont free lines (p=0.999; Figure 2-4A), suggesting that the presence of a facultative symbiont, in particular *F. symbiotica,* mediated this transgenerational effect. There was no significant interaction between generation and symbiont in the proportion of non-sporulating dead individuals ($F_{3,107}$=0.54, p=0.656) or the proportion of surviving individuals ($F_{3,106}$=0.69, p=0.560), although there was a trend towards a larger increase in survival in those with facultative symbionts compared to those without (Figure 2-4B).

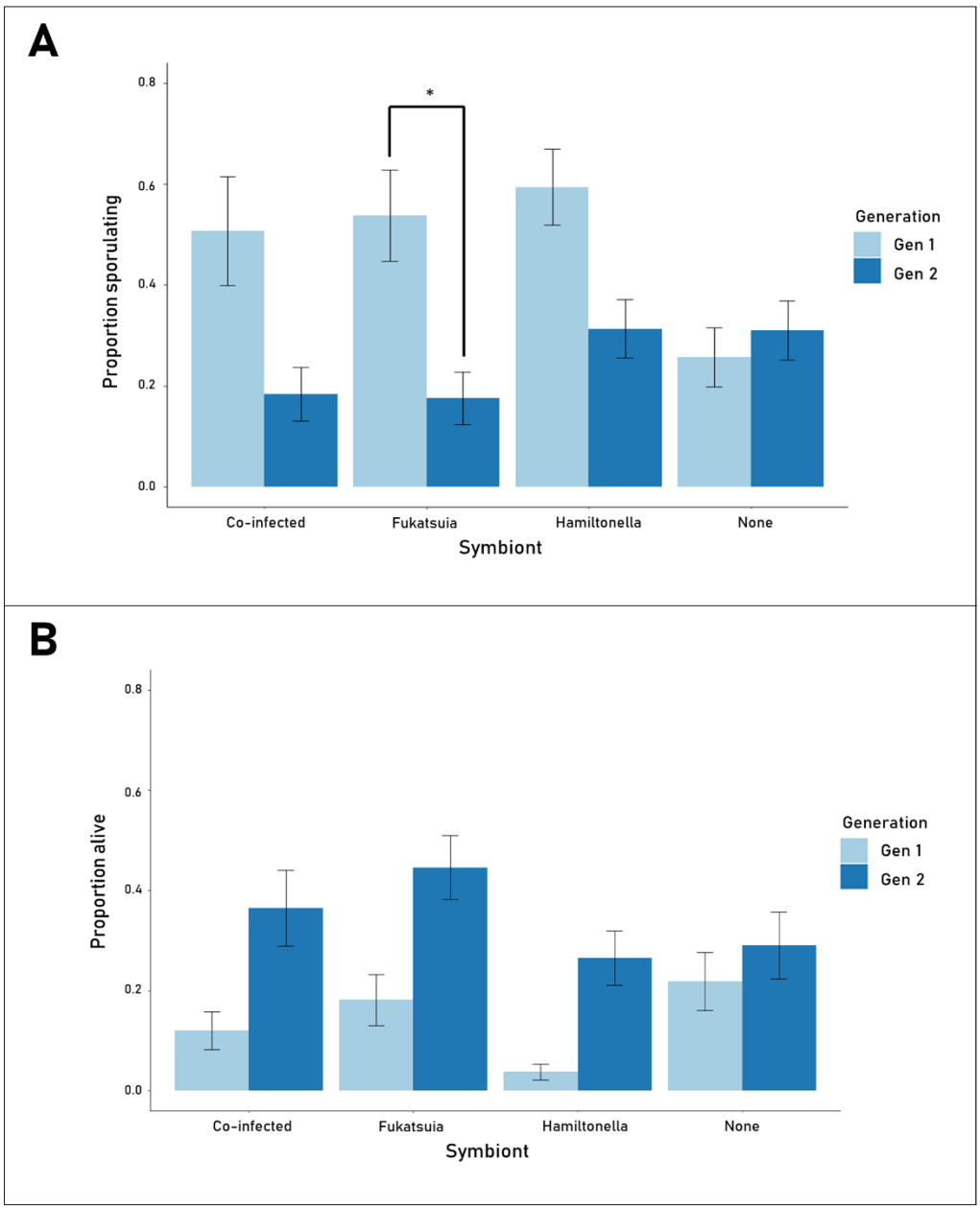


Figure 2-4: Proportion of sporulating individuals across four aphid symbiont treatments, over two generations of fungal pathogen infection

A) Proportion sporulating B) Proportion alive

Generation 1=light blue, Generation 2=dark blue.

Error bars represent the standard error. Asterisks show significant difference in proportion sporulating between generation one and two across aphid symbiont treatments (*: $p<0.05$)

The generation 2 individuals were offspring from mothers from three potential outcomes of the fungal pathogen infection; mother alive, mother sporulated and mother died. We hypothesised that mothers who had different outcomes in the first generation may lead to different levels of offspring susceptibility in the second generation. However, we found that the maternal outcome had no significant effect on the proportion sporulating ($F_{2,110}$=2.03, p=0.137; Figure 2-2C), the proportion of surviving individuals ($F_{2,111}$=0.44, p=0.646; Figure 2-2A) or the proportion of non-sporulating dead individuals ($F_{2,111}$=2.42, p=0.094; Figure 2-2B). However, when analysing the proportion alive, we did see a significant interaction between maternal outcome and symbiont line ($F_{6,101}$=2.63, p=0.021), since mothers who harboured *H. defensa* and sporulated led to significantly higher survival in the second generation (p=0.012; Figure 2-2A), an effect not seen in other symbiont lines. There was no significant interaction between maternal outcome and symbiont line in the proportion of sporulating individuals ($F_{6,100}$=0.95, p=0.461) or the proportion of non-sporulating dead individuals ($F_{6,101}$=2.04, p=0.067).

Facultative symbiont densities following fungal infection

We analysed the symbiont densities over two days of fungal pathogen infection in the first generation and in the first day of infection in the second generation following maternal sporulation ("Mother Sporulated" treatment), to investigate how the symbiont densities may have been affected by the fungal pathogen infection and may differ across the two generations.

Overall, there was a significantly higher number of copies of *H. defensa* than *F. symbiotica* in both the singular and co-infected aphids ($F_{3,33}$=17.148, p<0.001; Figure 2-5). We also found a significant difference in the symbiont densities across the timepoints ($F_{2,33}$=16.82, p<0.001; Figure 2-5), with generation 2 individuals having higher symbiont densities than generation 1 individuals at both 3- and 48-hours post-infection (p<0.001). There was no difference in the symbiont densities in generation 1 across the two timepoints (p=0.675).

There was also a significant interaction between symbiont infection and timepoint ($F_{6,33}$= 4.85, p=0.001). Post-hoc tests revealed that, although there appeared to be an increase in *F. symbiotica* densities in the second generation compared to the first, this was not significant in either the single (p=1.00; Figure 2-5A) or co-infected (p=1.00; Figure 2-5B) aphids. In those harbouring *H. defensa* however, there was a significantly higher symbiont density in generation 2 compared to generation 1 at 3 hours post-infection in the single (p=0.012; Figure 2-5C) and co-infected aphids (p<0.001; Figure 2-5D). The *H. defensa* densities in generation 2 aphids, compared those at 48 hours post-infection in generation 1, were also significantly higher in the co-infected aphids (p<0.001; Figure 2-5D) and, although there was a trend towards an increase in density in aphids harbouring *H. defensa* in a single infection, this was not significant (p=0.090; Figure 2-5C).

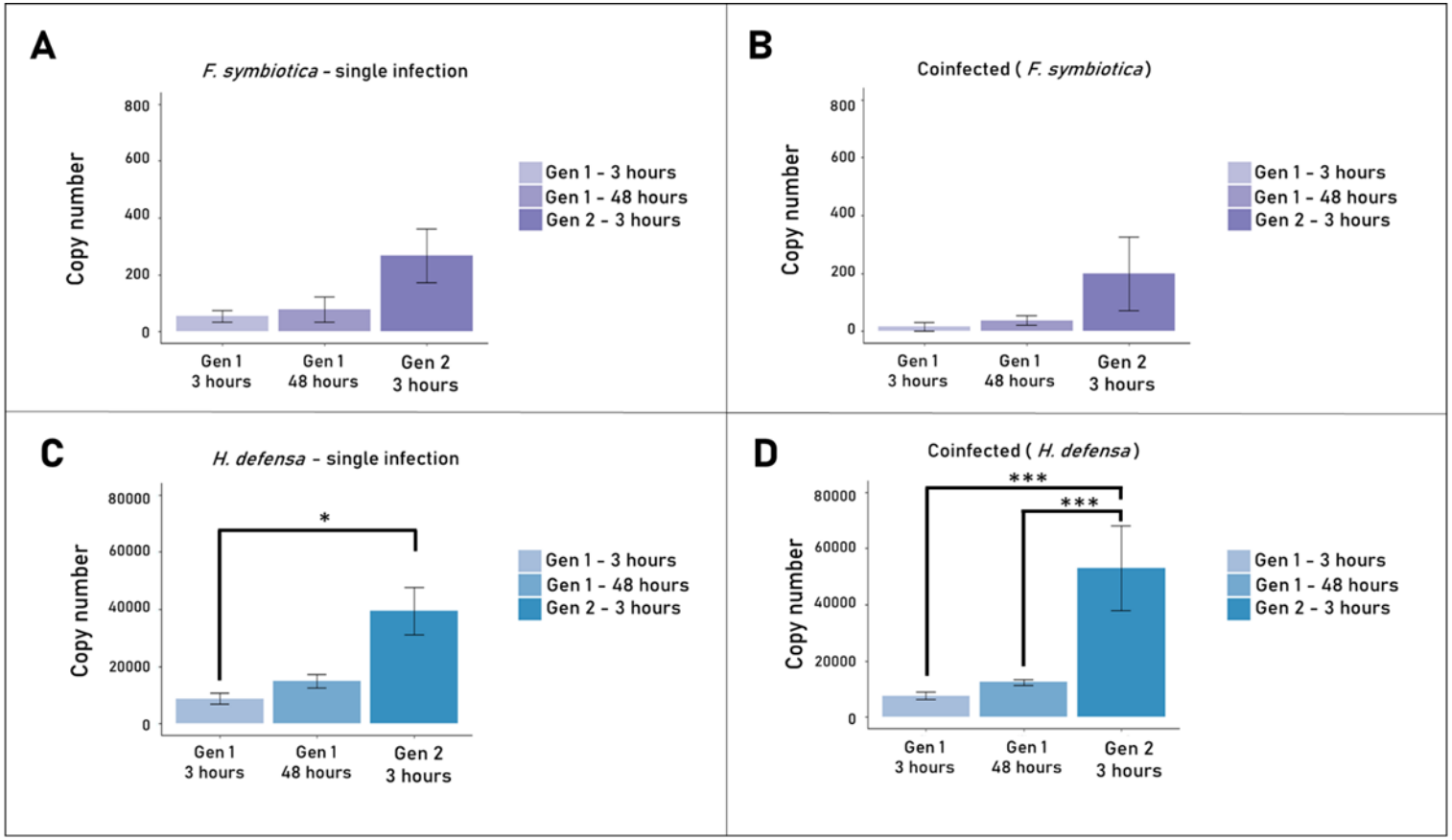


Figure 2-5: Symbiont densities following fungal pathogen infection over two generations

Mean symbiont copy number per aphid following fungal pathogen infection in aphids A) singly carrying *F. symbiotica*, B) carrying *F. symbiotica* in a co-infection C) singly carrying *H. defensa*, B) carrying *H. defensa* in a co-infection. Copy number was corrected for using the DNA concentration of each sample to allow for differences in aphid size and DNA extraction efficiency.

The graph shows two timepoints across the first two days of infection in generation 1 (3 hours and 48 hours post-infection), and one timepoint for generation 2 (3 hours post-infection).

Purple bars represent the symbiont counts for *F. symbiotica* in both singly and co-infected individuals. Blue bars represent the *H. defensa* densities. The error bars show standard errors and asterisks show significant differences in symbiont copy numbers across timepoints (*: $p<0.05$, ***: $p<0.001$).

Discussion

We found that if mothers were exposed to a fungal pathogen then their offspring were less susceptible to the same fungal pathogen in the next generation, although only in those carrying a facultative symbiont. We saw a decrease in the proportion of sporulation in the second generation in those carrying a facultative symbiont, which was statistically significant when the individuals harboured the protective symbiont *F. symbiotica*, in a single infection. We also saw a trend towards an increase in survival in the individuals with facultative symbionts, although this increase was not significant. There was no change in the proportion of sporulating or surviving individuals in the facultative symbiont free individuals, suggesting this transgenerational effect relies on the presence of a facultative symbiont. However, although we saw a decrease in susceptibility across generations in those harbouring facultative symbionts, this was not reduced relative to the symbiont-free control and was instead due to a reduction of the increased susceptibility of symbiont harbouring aphids in generation 1.

It is well known that insects are able to influence, or prime, the immune system of their offspring following exposure to pathogens and parasites (Moret 2006, Roth et al. 2018, Sadd & Schmid-Hempel 2006). The mechanisms by which this influence occurs are still unclear and may vary across host-pathogen interactions. However, the conditions necessary for this to evolve are well understood and largely relies on the mother's environment being an accurate predictor of the environment their offspring are likely to experience (Roth et al. 2018). This transgenerational effect would be expected to occur in aphids as their telescoping of generations, and fast generation time, means offspring are likely to share their environment with their parents. This makes the maternal environment a good predictor of what their offspring are likely to experience, and thus their risk of encountering the same parasite or pathogen. This is particularly evident in the case of pathogens which are immobile, such as fungal pathogens, which rely on the transmission of their spores to individuals within close proximity. Despite this, immune priming has yet to be seen in aphids. One study found no increase in resistance to a bacterial pathogen in

aphids that had been previously exposed to the pathogen, although this was not carried out across generations (ter Braak et al. 2013). Likewise, another study investigating transgenerational immune priming (TGIP) in response to a parasitoid wasp, found no change in susceptibility to parasitism following maternal exposure (Vorburger et al. 2008). Overall, however, there has been limited research investigating TGIP in aphids.

Given previous studies, it is therefore unsurprising that we found no evidence of TGIP in the facultative symbiont free treatment in response to the fungal pathogen. These aphids were already able to defend against this pathogen reasonably well, with only around 25% of the individuals sporulating in the first generation. Interestingly, aphids harbouring a facultative symbiont, in generation 1, were more susceptible to the fungal pathogen (50-60% sporulation) than those without a facultative symbiont, with the level of susceptibility varying slightly across the different lines. This was surprising for those harbouring *F. symbiotica*, as we may have expected this to be lower given that this symbiont has been seen to protect against another fungal pathogen (Heyworth & Ferrari 2015). However, in other facultative symbiont species, the fungal pathogen protection they provide has some level of specificity (Parker et al. 2013), which could be the same in this case. In the second generation, we see the susceptibility of those carrying facultative symbionts decrease, and the survival increase, suggesting there is a transgenerational effect occurring. The level to which the sporulation reduces, however, is similar to that from individuals who have no facultative symbionts. This, therefore, does not necessarily suggest a benefit, but instead a reduction of the cost of increased susceptibility that occurs when harbouring a facultative symbiont.

The mechanism by which this symbiont-mediated transgenerational effect occurs is unclear. However, one likely explanation could be that it is due to an interaction between the aphid immune system and the facultative symbionts. It is known that symbionts can play an important role in shaping host immunity (Eleftherianos et al. 2013, Gross et al. 2009), however, there is still limited understanding of the mechanisms underpinning the

interactions between the host and facultative symbionts (Laughton et al. 2016, Schmitz et al. 2012) and the interactions often differ across symbiont and host genotype combinations.

Endosymbionts live under the active immune system of their host and, whilst some have mechanisms that enable them to evade their hosts immune system (Burke & Moran 2011a, Hurst et al. 2003), when detected this can lead to a variety of outcomes. In some cases, the presence of a symbiont leads to increased immune response or investment, priming the immune system against future attacks. This effect is evident in mosquitos that carry *Wolbachia,* which can lead to increased resistance against a range of pathogens and parasites (Kambris et al. 2009, Rancès et al. 2012). Likewise, it has been shown recently that the beneficial symbiont *Snodgrassella alvi*, can prime the immune system of honeybees, increasing their resistance to pathogens (Horak et al. 2020). However, in some systems, harbouring a symbiont can lead to the host being more vulnerable to attack. For example, increased immune investment due to symbiont presence could lead to the immune capacity of the host becoming depleted (Vigneron et al. 2012), alternatively, the host may down-regulate their immune response, to ensure they are not fighting their beneficial microbes, again, leaving them vulnerable to invading pathogens (Vigneron et al. 2012).

In aphids, the results in this area are mixed. One study found that the presence of *H. defensa* and *R. insecticola* altered the immune response of the host, although this differed across symbiont species and aphid genotype (Laughton et al. 2016). Laughton et al. (2016) found that carrying *H. defensa* led to increased immune cell counts, whilst *R. insecticola* led to an increased encapsulation response. These symbionts, however, have also been found within phagocytic immune cells, suggesting that immune cells within the aphid may be playing a role in symbiont control (Laughton et al. 2011, Schmitz et al. 2012). In this case, given that *H. defensa* is known to have a fast replication rate and has maintained virulence factors within its genome (Degnan et al. 2009, Laughton et al. 2014), it is thought that the increased immune cells seen in those harbouring *H. defensa* were due to the host using some of its immune capacity to control the populations of this symbiont. There is also evidence, however, that aphids harbouring *H. defensa* and *R. insecticola* have reduced immune cell

numbers compared to uninfected aphids (Schmitz et al. 2012). Furthermore, a study investigating whether there was any immune priming in aphids in response to a bacterial pathogen found no evidence of increased resistance when harbouring a facultative symbiont (ter Braak et al. 2013).

Here, we find that in the first generation, aphids harbouring facultative symbionts are more susceptible to the pathogen than those without. This could be due to the aphid concentrating some of their immune capacity to control of symbiont populations, rather than investing in pathogen defence, which may leave these individuals more vulnerable to attack. Following the invasion of this pathogen, it may be that mothers influence the immune response of their offspring, to increase investment in defence over symbiont control, thus leading to a decrease in the susceptibility of these individuals.

The question remains whether this is a general effect that would be typical of a response to a range of invading pathogens, or whether this only occurs when harbouring a symbiont leads to individuals being more susceptible to a particular pathogen. If, for example, individuals with symbionts did not have increased susceptibility to a pathogen, as seen in bacterial pathogens for example (ter Braak et al. 2013), we may expect this transgenerational effect not to occur. Likewise, if the host was harbouring a symbiont that was able to protect against the pathogen, we may not see a transgenerational effect, as there would be no immunological cost, in the first generation, in which to mitigate against.

We hypothesised that the fungal pathogen infection may influence changes in symbiont densities, within and across generations. There was no change in the symbiont densities across the first two days of infection in generation 1, suggesting that the fungal pathogen did not directly affect symbiont densities in the aphid. However, there was an increase in symbiont densities in the second generation, which was significantly increased in those carrying *H. defensa* in both a single and co-infection. If it is the case that, in the second generation, individuals are investing more immune capacity in defence rather than

symbiont control, this would explain why we see an increase in the symbiont densities. Additionally, *H. defensa* has been seen to reproduce quickly in other aphid genotypes (Laughton et al. 2014), which would explain the larger increase seen in *H. defensa* compared to *F. symbiotica*.

As discussed above, aphids without facultative symbionts did not show any increased resistance to the fungal pathogen but showed relatively low susceptibility (25-30% sporulation) across the two generations. In the first generation, those harbouring symbionts had increased susceptibility to the pathogen, which decreased in generation 2. We hypothesise this is due to a switch in immune resource prioritisation from control of symbiont populations in generation 1, to pathogen defence in generation 2. If this was the case, providing the symbiont is not protective against the pathogen, we might expect these aphids, now investing only in pathogen defence, to have similar levels of susceptibility to those without symbionts, who too only invest in pathogen defence. We saw the levels of sporulation in aphids carrying *H. defensa* in generation 2 decrease to around 30%, similar to that in the symbiont free aphids, and given that *H. defensa* has not been seen to provide any fungal pathogen protection, this is as expected given the above hypothesis.

However, in aphids harbouring *F. symbiotica*, in both a single and co-infection we saw the individuals in generation 2 having a more reduced susceptibility, with around 17% sporulation. This is lower than that of the facultative symbiont free individuals suggesting that the protection in this line may stretch beyond just an immune response, with potentially more than one mechanism at work. *F. symbiotica* is able to protect against the fungal pathogen, *Pandora neoaphidis* (Heyworth & Ferrari 2015, Smee et al. 2021). Therefore, in addition to the change in prioritisation of their immune resources towards pathogen defence, that we hypothesise may be occurring, there could also be a resistance mechanism provided by *F. symbiotica*, which is working to further protect against this pathogen. Thus far, *F. symbiotica* has only been seen to provide resistance against the aphid specific fungal pathogen, *P. neoaphidis*, however, this could suggest that

F. symbiotica may be able to also provide resistance against other fungal pathogens. Although, further research would be required to confirm whether *F. symbiotica* can exhibit generalist fungal pathogen resistance, rather than being limited to aphid-specific pathogens as seen in other symbionts, such as *R. insecticola* (Parker et al. 2013). Additionally, if it is the case that this strain of *F. symbiotica* can protect against this pathogen, it is unclear why this was not exhibited in the first generation.

Overall, we found a transgenerational effect occurring in response to this fungal pathogen. We saw a reduction in susceptibility in the second generation, which only occurred when aphids harboured a facultative symbiont and was only statistically significant in those harbouring the protective symbiont, *F. symbiotica*, in a single infection. This suggests that this transgenerational effect is mediated by the presence of a symbiont, rather than the classic form of transgenerational immune priming found in other insects. It appears that the aphids may be influencing their offspring to invest their immune capacity towards pathogen defence, where they otherwise may have focussed their resources on controlling symbiont populations. There was an increase in symbiont densities in the second generation, especially in those harbouring *H. defensa*, which further supports the idea that the aphid is less invested in symbiont control as they were in generation 1, allowing the symbionts to replicate and reach higher densities.

This is the first evidence for symbiont-mediated transgenerational effects in insects. Much of the work investigating the interaction between facultative symbionts and their hosts has focussed on how the presence of a symbiont may prime the immune system of the host (reviewed in Eleftherianos et al. 2013). However, this work has failed to explore how this may transmit across generations. Aphids are an interesting model in which to explore this question, as they rely heavily on their facultative symbionts for protection against natural enemies, and have a reduced immune system compared to other insect species (Gerardo et al. 2010). In this study, we find that harbouring a facultative symbiont makes the individual more susceptible to this pathogen in the first generation, and it may be that this

increased susceptibility is what allows for this transgenerational effect to occur. It appears the aphid responds to this invading organism by plastically altering the prioritisation of its offspring's immune response, in this case potentially switching from symbiont control to pathogen defence. As the susceptibility in generation 2 does not decrease below that of symbiont free individuals, this could be considered not necessarily beneficial to the host, but instead mitigation of a cost. Therefore, it may be the case that this effect would not be seen in other systems unless they harboured a symbiont which led to increased susceptibility to a pathogen.

The transgenerational effect seen here is likely due to an interaction between the facultative symbionts and the aphid immune system. This highlights that we still have much to understand about the interaction between facultative symbionts and their hosts, how they may influence the immune system, and how exposure to pathogens may alter this.

Chapter Three

Effect of facultative symbionts on host immunity and response to a fungal pathogen

Abstract

Insects face a range of pathogens and parasites and often rely heavily on their innate immune system to protect against these. Many also carry bacterial endosymbionts, which often aid in protecting their host. In addition to the direct protection provided by symbionts, this defence is thought to also be due to the interaction between the symbiont and the immune system where, in some cases, the presence of the symbiont can influence host immunity, "priming" them against future attacks. Aphids carry several bacterial symbionts, but unlike many other insects, aphids have a reduced repertoire of immune genes, likely due to a complex interaction between their symbionts and the host immune system. Thus far, there have been mixed results on the effect of symbionts on insect immunity, including within aphids. In this chapter, we investigated the effect of two facultative symbiont species (*Fukatsuia symbiotica* and *Hamiltonella defensa*) on host immunity, and whether the presence of these symbionts alters how the aphid immune system responds to a fungal pathogen infection. We investigated the cellular immune response and the expression of six candidate immune genes in the days following infection with a fungal pathogen.

We found a difference in the cellular immune response between those with and without facultative symbionts. There was a large increase in immune cell counts following fungal pathogen infection in aphids without facultative symbionts, which was lessened in those carrying the non-protective *H. defensa*, and not seen in those harbouring *F. symbiotica*. In the expression of the immune genes, generally, those with and without symbionts showed similar patterns of gene expression changes following fungal pathogen infection, suggesting the presence of a symbiont did not influence how these genes responded to the pathogen. However, before the fungus had penetrated the aphid cuticle, there was a significantly higher level of expression of many of the genes in aphids with facultative symbionts, particularly *F. symbiotica*. So, although the symbionts may not have had a large influence on the response to the pathogen, the symbionts did appear to be altering host immune response in the absence of a pathogen. This, however, was found in aphids

harbouring both symbiont species, and as those carrying *H. defensa* are highly susceptible to the pathogen, this suggests that this alteration of the immune response did not prime the host against attack. There is a very complex interaction occurring between the aphid immune system and their facultative symbionts, which is still largely not understood.

Introduction

Insects face a range of pathogens and parasites within their environment. To protect against these, they often rely on their non-specific innate immune response. Insects also often carry bacterial symbionts which can help them in a variety of ways, including aiding with protection against natural enemies (Feldhaar 2011). These endosymbionts live under the pressure of their host's active immune system, and so there are often interactions occurring between these symbionts and the host immune system (Eleftherianos et al. 2013). Outside of the protection provided directly by the symbionts, these interactions may also play a role in helping protect against invading pathogens, through the enhancement of the host immune response (Eleftherianos et al. 2013). For example, this effect can be observed in the host's response to *Wolbachia*, a common bacterial symbiont that infects multiple host species. *Wolbachia* can prime the immune system of mosquitoes which helps protect against a range of threats, including bacterial and viral infections (Kambris et al. 2009, Moreira et al. 2009). However, in flies, although there are studies suggesting *Wolbachia* helps protect against RNA viruses (Teixeira et al. 2008), there is also evidence that carrying *Wolbachia* can have neutral or negative effects on susceptibility to pathogens and parasites (Rottschaefer & Lazzaro 2012, Wong et al. 2011). Although, more recently it has been shown that antibacterial protection can be provided by *Wolbachia* in *Drosophila* through different, more natural, routes of pathogen infection (Gupta et al. 2017). This shows that the interaction between these symbionts and their host, and the effect it may have on their immune response, can vary greatly across host and symbiont genotype, as well as the infectious organism they are facing.

Aphids carry a number of bacterial symbionts, which help protect the aphid against a range of natural enemies (Oliver et al. 2010). For example, *Hamiltonella defensa* is a well-studied symbiont of the pea aphid which is known for its ability to protect the host against the parasitoid *Aphidius ervi*. This protection is provided by a phage (APSE), which encodes toxins that prevent the wasp from developing (Degnan & Moran 2007; Oliver et al. 2003, 2005). There are also several symbiont species, including *Regiella* (Scarborough et al.

2005), *Fukatsuia* (Heyworth & Ferrari 2015), *Spiroplasma*, *Rickettsia* and *Rickettsiella* (Łukasik et al. 2013c), that can protect against a fungal pathogen *Pandora neoaphidis*, although, the mechanisms by which this protection is conferred is still unknown. Whilst it is thought that these protective phenotypes are directly due to the symbiont, as we have little understanding of these mechanisms, it is unclear whether these are direct symbiont effects or due to an interaction between the symbiont and the host.

Our understanding of the cellular and humoral response of the innate system of insects has been largely reliant on descriptions from the model species *Drosophila melanogaster* (Ferrandon et al. 2007, Lemaitre & Hoffmann 2007), although these are similar across a number of insects including mosquitoes (Christophides et al. 2004), beetles (Altincicek et al. 2008, Zou et al. 2007) and bees (Evans et al. 2006). The cellular response involves the use of cells that are able to encapsulate and phagocytise invading pathogens and parasites, whereas the humoral response relies on a range of signalling pathways which in turn leads to the release of effectors, commonly phenoloxidases and anti-microbial peptides (AMPs) (Lemaitre & Hoffmann 2007).

Aphids, however, have a reduced repertoire of immune genes, compared to other insects (Gerardo et al. 2010). They have maintained the cellular responses seen in other insect species, although the number of cells is reduced (Laughton et al. 2011). However, their humoral responses, in particular those related to bacterial pathogen defence, differ, and they are lacking many of the signalling pathways and AMPs, conserved across many insects (Gerardo et al. 2010, Zou et al. 2007). It is thought that this reduced immune system may be linked to the strong association these hosts form with their facultative symbionts, and because of this, other phenomena that occur in other insect species, such as the priming of the immune system, within and across generations, are thought to be unlikely.

Thus far, studies investigating the effects of aphid facultative symbionts on the immune system of the aphid have had mixed results. One study found that harbouring facultative symbionts led to an increase in the number of circulating immune cells in the haemolymph of the pea aphid (Laughton et al. 2016). Symbionts have also been found within the

hemocytes, suggesting that hemocytes are potentially playing a role in the control of symbiont populations (Laughton et al. 2016). However, another study found contradictory results, suggesting that harbouring the same symbionts led to a decrease in immune cell numbers (Schmitz et al. 2012). Additionally, an investigation into symbiont influence on within-generation priming against bacterial pathogens found symbiont presence did not affect survival, suggesting they were not able to prime against future attacks (ter Braak et al. 2013).

Work thus far has concentrated on the symbionts influence on the aphid's ability to respond to bacterial pathogens. Yet, we know that the aphid genome is missing many of the important components to be able to elicit an immune response to bacterial pathogens (Gerardo et al. 2010, ter Braak et al. 2013). In response to a fungal pathogen, however, pea aphids do appear to exhibit an immune response (Barribeau et al. 2014, Parker et al. 2017, Xu et al. 2019). For example, there is evidence that infecting facultative symbiont free aphids with fungal pathogens leads to the fitness cost of exhibiting an immune response, which is not seen in response to bacterial pathogens (Barribeau et al. 2014). However, unlike in response to bacterial pathogens, there have been no investigations into how harbouring facultative symbionts may influence the immune response to fungal pathogens and/or whether this differs with symbiont species.

Here, we investigated whether harbouring a facultative symbiont alters the host immune response, and how this may influence their response to infection with a fungal pathogen. Using two different facultative symbiont species, we also aimed to examine whether this response changed when the aphid harboured either a protective or non-protective symbiont. We firstly examined the cellular immunity of the aphids, with and without facultative symbionts, looking at the number of circulating immune cells at four different timepoints following infection. We hypothesised that harbouring a facultative symbiont may alter the cellular immunity of the aphid and thus how the aphid responds to a fungal pathogen infection over time. We expected that harbouring facultative symbionts may lead to increased cellular immunity, which may, in turn, mean the aphid does not need to further

increase this immune response following pathogen infection. Although in those harbouring *F. symbiotica*, if this symbiont protects against this pathogen through its own mechanisms, we might expect to see less of a cellular response in this treatment.

We also investigated the expression of six immune genes found in the pea aphid, the difference in expression in those harbouring different facultative symbionts and how these changed over the first two days following pathogen infection. We hypothesised that there would be a difference in the expression of the candidate immune genes due to the presence of a symbiont altering the immune response of the aphid.

The genes used in this study were chosen from a previous study which investigated differential expression of immune genes in winged and non-winged pea aphids in response to the fungal pathogen *P. neoaphidis*, using transcriptome sequencing (Parker et al. 2017). Parker et al. (2017) identified nine immune genes which were differentially expressed in the transcriptome and analysed these using qPCR. From those genes, we chose the five genes which, following qPCR, showed increased expression in response to the fungal pathogen in the unwinged aphids and one which did not show any changes in expression in response to *P. neoaphidis* infection. The included phenoloxidase, known to be an important part of the insect innate immune response (Cerenius et al. 2008), playing a role in protecting against pathogens in a range of insects, including the pea aphid (Laughton et al. 2011, Schmitz et al. 2012, Xu et al. 2019), legumain, which forms part of the lysosomal pathway and has been found to upregulate in response to viruses in aphid vectors (Li et al. 2020), in addition to the response to a fungal pathogen, and four cathepsin genes. Cathepsins are proteases, expressed in hemocytes, which are known to have immunological properties. They have been shown to have lysosomal activity against a range of immunological threats including bacteria, viruses, as well as previously being found to be involved in fungal pathogen responses in grain aphids (Grell et al. 2011). Three of the cathepsin genes were cathepsin B variants. Within aphids, cathepsin B genes have diverged into a large number of copies, with 28 copies found in the pea aphid genome, compared to just two copies in *D. melanogaster* (Rispe et al. 2008). This large divergence also occurs in other insect

species which feed on unbalanced diets, and so is thought to be related to the aphid diet of plant sap (Rispe et al. 2008). These genes appear to have a range of functions across different aphid species. For example, in the green peach aphid, *Myzus persicae*, they play a role in the ability of the aphid to colonise different host plants (Mathers et al. 2017), affect the growth of the aphids (Pinheiro et al. 2017) and alter the transmission efficiency of a plant virus (Pinheiro et al. 2017). Additionally, in the social aphid *Tuberaphis styraci*, a cathepsin B gene forms an important component of the venom of soldier aphids (Kutsukake et al. 2004). The fourth cathepsin gene was cathepsin L, which was not differentially expressed in response to *P. neoaphidis* in Parker et al. (2017), and instead used as a control. However, this gene may potentially play a role in regulations of *Buchnera* populations (Nishikori et al. 2009), as well as symbiont populations in other systems (Byeon et al. 2015). Therefore, here it would be interesting to investigate whether this gene is differentially expressed in those with and without facultative symbionts and whether the pathogen infection would alter this, potentially due to three-way interactions between the host, symbiont and pathogen.

Methods

Experimental design

The experiment was carried out using four different treatments to investigate how harbouring facultative symbionts alters the immune response of the aphid, and whether this differed when harbouring a protective (*F. symbiotica*) or non-protective (*H. defensa*) symbiont. The four treatments included one which carried no facultative symbionts and was infected with the fungal pathogen (Symbiont free F+) and one which had no facultative symbionts but did not undergo fungal pathogen infection, acting as a symbiont free control (Symbiont free F-). All treatments would be compared back to the symbiont free control (Symbiont free F-). There were two facultative symbiont treatments, one which consisted of aphids carrying *F. symbiotica* isolate F1 and one with aphids carrying *H. defensa* isolate H1, both of which were infected with the fungus. All the aphids used were the same aphid genotype.

There were no symbiont harbouring fungus free treatments included in this analysis due to issues that occurred during the experiment, which led to the symbiont harbouring fungus free treatments being infected with the fungus. However, these treatments were not required to answer the questions presented in this chapter; firstly, whether harbouring a symbiont alters host immunity compared to those without symbionts. This could be answered by comparing across treatments at 0 h as, at this timepoint, the fungus would have not yet penetrated the aphid cuticle. However, it is important to note that although the fungus had not passed across the cuticle at 0 h, it is not to say there is not another process by which the aphid can sense the imminent infection. Secondly, we aimed to understand how the immune gene expression changed in the two days following fungal pathogen infection. Here, we compared the gene expression across different timepoints, within a treatment, all of which were compared back to the same reference (symbiont free F- at 0 h), again allowing for the lack of symbiont harbouring fungus free treatments. This was also true for understanding the differences in post-infection gene expression in aphids with and without facultative symbionts. The expression at each timepoint could be compared across treatments as they were all calculated using the same reference, meaning all the fold changes were relative to the symbiont free, fungus free control.

Experimental procedure

To produce the experimental aphids, 10 adult aphids were placed onto 2-week-old *Vicia faba* (var. The Sutton) plants enclosed in a vented 2 L cage. These were left overnight to reproduce and removed the following day. The offspring were left to mature until they were nine-to-ten days old. At this stage, 22 individuals for each treatment were randomly selected. The individuals were then infected with the fungal pathogen, *Lecanicillium lecanii* (identified based on morphology), as described below. The fungal pathogen used in the experiment was a natural isolate extracted from our aphid cultures. Following the fungal pathogen infection, individuals were placed onto a petri dish containing a *V. faba* leaf with the stem inserted into 2% agar to keep the leaf fresh. All petri dishes contained leaves of

similar sizes. Individuals were taken at four timepoints for the cellular immunity assay and at three timepoints for analysis of the expression of the immune genes (methods detailed below). The remaining aphids were kept for a measure of susceptibility to the fungus, and at nine days post-infection were marked as either sporulating, alive or dead (where the aphid had died but not necessarily due to the pathogen as there were no spores formed). Individuals from the symbiont free F- treatment were treated identically throughout but were not infected with the fungal pathogen. The experiment was carried out using five biological replicates.

Fungal pathogen infection

Fungus infected adult aphid cadavers were left on 1% agar overnight at 20°C in a high humidity environment to start sporulation. The fungal pathogen infection was carried out by exposing aphids to a pair of these aphid cadavers. For the infection, the cadavers were placed onto wet filter paper placed over a small cylindrical glass vial (height=35 mm, diameter=20 mm) which created a "spore shower". The spore shower lasted 90 minutes, during which the cadavers were rotated among the replicates to ensure that each replicate was exposed to an equal amount of spores.

Cellular immunity assay

To measure the cellular immunity of the aphids in the different treatments, we counted the number of circulating immune cells in the haemolymph of the aphids. Haemolymph was collected from leg wounds of two adult aphids using a hand-held micro-injector and glass capillary. The haemolymph was smeared onto a slide, then fixed and stained using a Clin-Tech Speedy Diff Kit according to the manufacturer's instructions.

The number of cells were then counted using a Nikon Eclipse 50i microscope at 200x magnification. As the volume of haemolymph able to be obtained from each replicate differed, we counted the number of cells within a 0.25 cm^2 area. We did three counts per slide and calculated the mean count. The immune cell counts were carried out at 0 h (immediately following 90-minute spore shower), 24 h, 48 h and 72 h post-infection. Due to the lack of symbiont harbouring fungus free treatments, we used the cell counts at 0 h as

a baseline for the number of cells without any pathogen influenced response. At 0 h the fungal pathogen would not yet have affected the aphid's cellular immunity as it would not have penetrated the cuticle of the aphid.

Gene expression analysis

We investigated the patterns of gene expression across the different symbiont lines and timepoints using quantitative PCR. The aphids were infected with the fungal pathogen, as described above (except for the fungus free treatment) and were flash-frozen in liquid nitrogen at 0 h, 24 h and 48 h post-infection. We collected five replicates of each treatment at each timepoint. The aphids were stored at -80°C until required. The RNA extractions were carried out using a Qiagen RNeasy Midi Kit with on-column DNase digestion using a Qiagen RNase-Free DNase set for the removal of genomic DNA contamination. RNA was then reverse transcribed to cDNA using a Thermo Scientific RevertAid First Strand cDNA Synthesis kit using Oligo(dT) 18 primers, according to the manufacturer's instructions. 1 µg of RNA was used for the cDNA synthesis and the sample was diluted by adding 80 µl of dH_2O before being used in the qPCR.

The cDNA samples were then used in a qPCR reaction, run on an Applied Biosystems QuantStudio™ 3 Real-Time PCR System, using SYBR® Green reagent. Within each well, there was 10 µl of Fast SYBR 2x master mix, 1 µl of forward primer (7 µM), 1 µl of reverse primer (7 µM), 6 µl of ultrapure water and 2 µl of cDNA. The qPCR cycling conditions were 95°C for 20 seconds, then 40 cycles of 95°C for 1 second and 60°C for 20 seconds. A melt curve was also generated for each plate. The cycling conditions for the melt curve were 95°C for 1 second, 60°C for 20 seconds and then a gradual increase at 0.1°C/s to 95°C. There were five biological replicates for each treatment and all samples were run in technical triplicates.

The qPCR was carried out using six different immune gene primers taken from Parker et al. (2017) (see Table 3-1 for sequences and efficiencies) and Ef1α was used as the endogenous control. To calculate the efficiency of the primers, first, a 5-sample curve was created by serially diluting a cDNA sample 1:10 with distilled water. The efficiency was then calculated

from the slope of the curve using the equation $Efficiency\ (\%) = (10^{\frac{-1}{slope}} - 1) \times 100$. The relative gene expression for each of the immune genes was calculated using the $2^{-\Delta\Delta Ct}$ method (Livak & Schmittgen 2001) using the control treatment (symbiont free, fungus free) at 0 h as the reference. These were then $\log_2$ transformed.

Statistical analysis

To investigate the susceptibility to the pathogen, the effect of treatment on the outcome of infection was analysed using a multiple logistic regression. The response variable was the outcome of the infection, at three levels (alive, dead, sporulating), and the explanatory variable was treatment. The treatments were; symbiont free fungus free (F-), symbiont free fungus infected (F+), *H. defensa* (F+) and *F. symbiotica* (F+). Post-hoc tests were carried out using the lsmeans package (Lenth 2016).

For the cellular immunity analysis, we first analysed the differences in cell counts across the treatments at timepoint 0 h, using an ANOVA with cell count as the response variable and treatment as the explanatory variable. Following this, all the data was analysed using a repeated measures two-way ANOVA with cell count as the response variable, timepoint, treatment and their interaction fitted as fixed effects and replicate as a random effect. This was carried out using the nlme package in R (Pinheiro J et al. 2020). For significant factors with more than two levels or significant interactions, post-hoc tests were performed using the lsmeans package (Lenth 2016).

The analysis of the expression of the immune genes was carried out as the cellular immunity analysis, but here the response variable was the ΔΔCt value. All six genes were analysed separately. All analyses were carried out in R, version 3.6.1 (R Core Team 2013) and graphs were made using the ggplot2 package (Wickham 2016).

Table 3-1: qPCR primer sequences used in gene expression analysis.
The table includes the gene name, ACYPI number and the primer efficiency for all six immune genes used and the endogenous control gene. Primer sequences were taken from Parker et al. (2017).

Gene name	Gene ID	Efficiency	Primer sequence
Immune genes			
Cathepsin B	ACYPI000014	89.6%	F: AACCGGGAACGCATCGT
			R: TCTACGCCCCATCCAATCAT
Cathepsin B	ACYPI89528	89.6%	F: CCCAGTTACAAGTCCGGTGTTT
			R: CACTGCGTGTCCTCCCAAT
Cathepsin B-n	ACYPI089577	86.3%	F: CTTCTACAGTTACAAGTCCGGTGTTT
			R: TGCGTGTCCGCCCAAT
Cathepsin L	ACYPI006974	83.3%	F: GCGGCAACGGAAAACG
			R: GCAAATTGACGAAACGCTTTATG
Legumain	ACYPI001325	93.1%	F: CCAACGCCAGGAAAAATCA
			R: CGCCTTTGTAGTCGATTTGAACT
Phenoloxidase	ACYPI004484	98.4%	F: CCACTGGGACTTAAGAGAGATATGC
			R: TGAGCTTTGCGGCCATTT
Endogenous control			
Ef1α	ACYPI006711	101.1%	F: CTGATTGTGCCGTGCTTATTG
			R: TATGGTGGTTCAGTAGAGTCC

Results

Susceptibility to the fungal pathogen

First, we analysed the outcome of infection for individuals across the different treatments following fungal pathogen infection. Overall, we found that treatment had a significant effect on the outcome of the infection ($F_{9,44}$=16.08, p<0.001, Figure3-1). As the symbiont free fungus free control (symbiont free F-) was not infected with the fungal pathogen, there were no sporulating aphids in this treatment (Figure 3-1). There was a significant difference in the sporulating individuals across the different treatments, with significantly higher sporulation in both the symbiont free fungus infected (symbiont free F+) treatment (p<0.001) and the *H. defensa* treatment (p<0.001) compared to symbiont free F-. Although those carrying *F. symbiotica* showed some susceptibility to the pathogen, this was not significantly higher than the symbiont free F- treatment (p=0.704). There was no significant difference in the sporulation rate across the fungus infected treatments, however, as the aphids carrying *H. defensa* were the most susceptible there was a trend towards higher sporulation in this treatment compared to aphids with *F. symbiotica* (p=0.098).

Although there was less sporulation in aphids harbouring *F. symbiotica*, there was no more survival in this treatment, with only around 10% of the individuals surviving the fungal pathogen infection. This was significantly lower than those not infected with the pathogen (p<0.001; Figure 3-1). As expected, there was also a significantly lower survival rate in those infected with the fungal pathogen, in the symbiont free F+ treatment (p<0.001; Figure 3-1) and the *H. defensa* treatment (p<0.001; Figure 3-1). All of the fungus infected treatments showed significantly lower survival than the fungus free treatment, but there was no difference in the survival rates across those that did undergo fungal infection.

Many of the individuals died, but did not sporulate, following the fungal pathogen infection, with those carrying *F. symbiotica* showing the highest rate of death. It is known that high spore dosages can lead to high mortality in aphids (Parker et al. 2014) and, although we did not measure spore dose in this experiment, lower dosages may have led to lower mortality across the infected treatments. However, across all treatments, there was no

significant difference in death rate, including the symbiont free F- treatment, which had a similar death rate to those infected with the pathogen (Figure 3-1) suggesting other factors, such as stress, may have caused the high death rate.

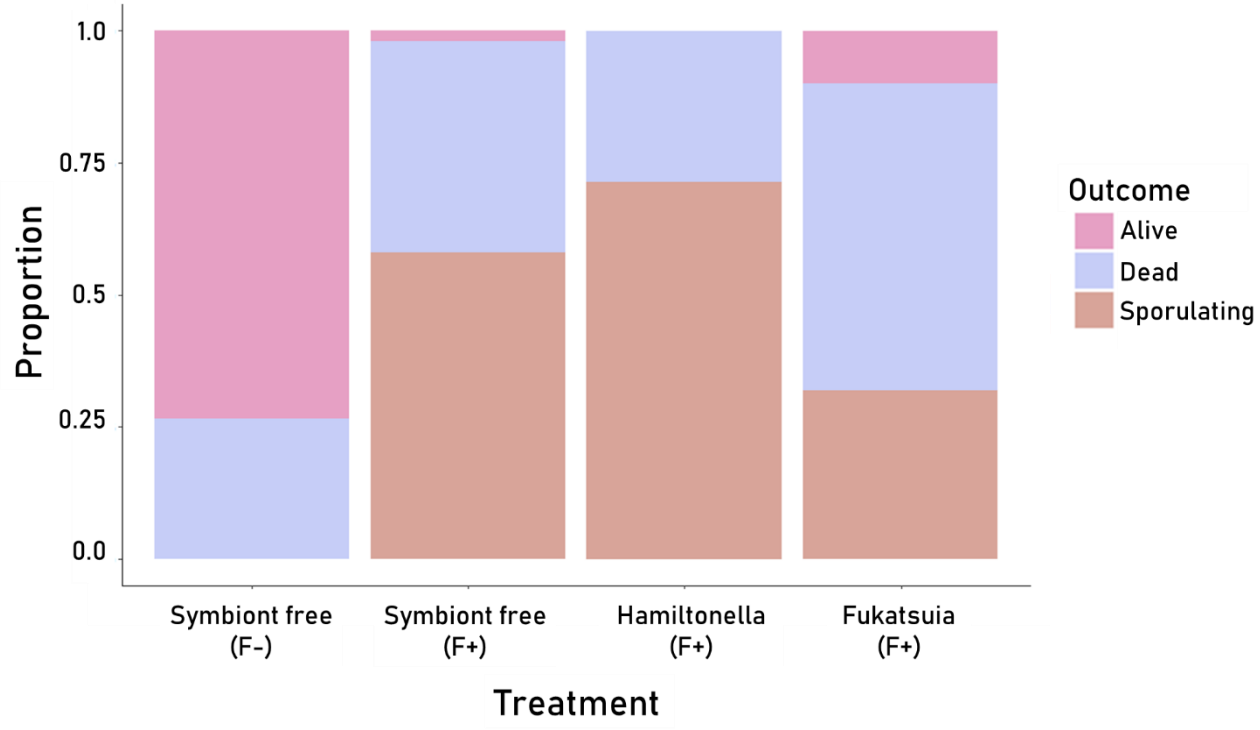


Figure 3-1: Proportion of individuals from each phenotype following fungal pathogen infection across the four treatments.
Alive = pink, Dead = light blue, Sporulating = brown
Experimental treatments: Symbiont free (F-), Symbiont free (F+), *Fukatsuia* (F+), *Hamiltonella* (F+)
F+ =infected with fungal pathogen, F- =no fungal pathogen infection

Cellular immunity

We investigated the number of circulating immune cells in the haemolymph of the different aphid treatments, across four timepoints, as a measure of cellular immunity towards the fungal pathogen infection. We first compared the number of cells across treatments at 0 h, at which the fungus would not yet have penetrated the aphid cuticle leading to a cellular immune response, to see if harbouring a facultative symbiont led to any differences in the

number of circulating cells. We also looked within treatments to investigate how the number of circulating cells differed following fungal pathogen infection, and whether this varied when harbouring a facultative symbiont.

For all the treatments, at timepoint 0 h, there was no difference in the number of cells ($F_{3,16}$=1.50, p=0.253; Figure 3-2). This suggests that harbouring a facultative symbiont does not lead to any increased cell numbers in these particular aphid-symbiont genotype combinations. We then explored the differences across timepoints and treatments. We found no overall effect of treatment on the number of cells ($F_{3,16}$=2.17, p=0.131). However, there was a significant effect of timepoint ($F_{3,28}$=5.70, p=0.003), with there being a significantly higher number of cells at 48 h post-infection compared to 0 h (p=0.002), 24 h (p=0.003) and 72 h (p=0.04). There was also a significant interaction between treatment and timepoint ($F_{9,28}$=4.22, p=0.001). Post-hoc tests showed that, as expected, there was no change in the number of circulating cells across all timepoints in the symbiont free F- treatment (p=1.00; Figure 3-2). There was, however, a significant increase in cell numbers at 48 h in the symbiont free F+ treatment (Figure 3-2). This was significantly increased compared to its own cell numbers at 0 h (p<0.001) and 24 h (p<0.001) but also compared to the number of cells in the symbiont free F- treatment (p<0.001) and *F. symbiotica* harbouring aphids (p=0.002) at the same timepoint. There was a trend towards there being more cells at 48 h in the symbiont free F+ than the *H. defensa* harbouring aphids, but this was not significant (p=0.056). This is likely due to the rise in the number of cells in the *H. defensa* treatment that also occurred at 48 h post-infection (Figure 3-2). However, unlike the symbiont free F+ treatment, this was not a significant increase in cell numbers. This may be because, although there was an increase in numbers overall, the number of cells counted at this timepoint was quite variable across the five replicates in this treatment. Interestingly, unlike in the symbiont free F+ and the *H. defensa* treatments, we saw no change in the number of cells across all the timepoints following pathogen infection in aphids carrying the protective symbiont *F. symbiotica* (p=0.999; Figure 3-2).

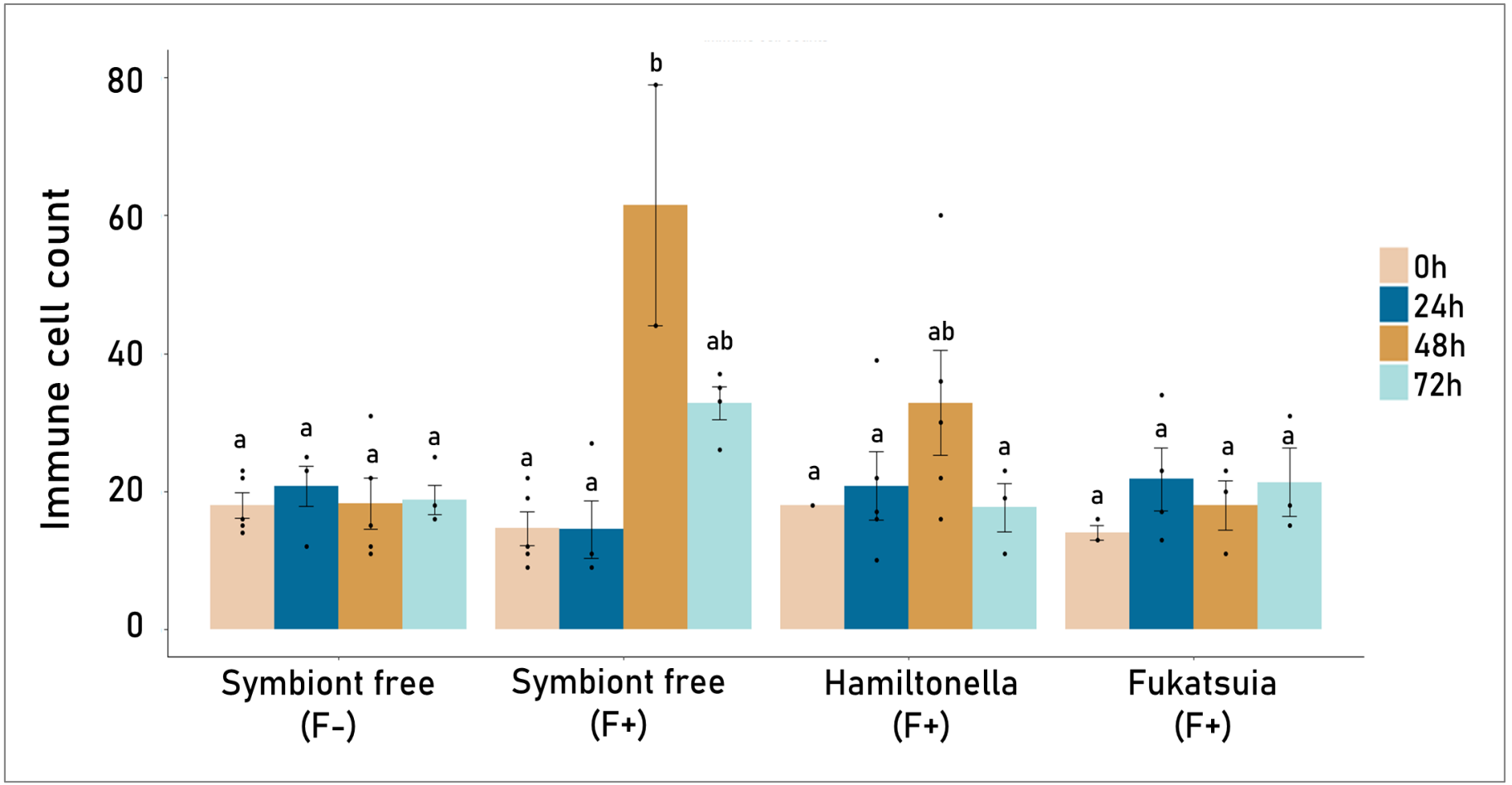


Figure 3-2: Number of circulating immune cells in aphids from four different treatments over four timepoints.

Cells were counted using a light microscope at 200x magnification within a 0.25cm^2 area. Three counts were taken per slide and an average taken.
Timepoints: 0 h post-infection=peach, 24 h post-infection=dark blue, 48 h post-infection=gold, 72 h post-infection=light blue.
F+ =infected with fungal pathogen, F- =no fungal pathogen infection.
Bars with the same letter are not significantly different. Error bars represent the standard error.

Gene expression analysis

Effect of fungal pathogen infection on immune gene expression

We investigated how the fungal pathogen affected the expression of the immune genes, by firstly looking at the differences in gene expression in the symbiont free aphids with and without fungal pathogen infection. We found that in the symbiont free F- treatment, although most genes showed a slight decrease in expression at 24 h, there was no significant change in gene expression across the three timepoints in any of the six genes (Figure 3-3A-F; full summary of stats found in Table 3-2). In those infected with fungus (symbiont free F+), there was also no significant change in gene expression across the timepoints, with the exception of Cathepsin B (00014) in which there was a strong trend towards higher expression than symbiont free F- at 48 h ($p=0.068$; Figure 3-3A). This suggests that Cathepsin B (00014) may be playing a role in the immune response to this pathogen, although overall there was not a large immune response within this set of genes exhibited in response to this pathogen. Comparing across treatments, in some genes (Cathepsin B (89528), Legumain and Phenoloxidase) there was no decrease in expression at 24 h, seen in the symbiont free F- treatment, occurring in the symbiont free F+ treatment (Figure 3-3B-D), suggesting the fungus may be having some effect on the expression of these genes. Although these differences in expression were not large enough to be statistically significant.

Looking then at the changes in expression across all treatments, including aphids carrying facultative symbionts, in two of the immune genes, there was a significant decrease in expression at 24 h compared to 0 h (Phenoloxidase ($p=0.009$; Figure 3-3D) and Cathepsin B-n ($p=0.007$; Figure 3-3E)). In Phenoloxidase, there was also a trend towards significantly lower expression at 48 h compared to 0 h (Figure 3-3D), but in Cathepsin B-n, the expression had increased again by 48 h (Figure 3-3E). There was also an overall higher expression of Cathepsin B (00014) at 48 h compared to 24 h ($p=0.03$; Figure 3-3A) and a trend towards this being higher than the expression at 0 h ($p=0.07$; Figure 3-3A).

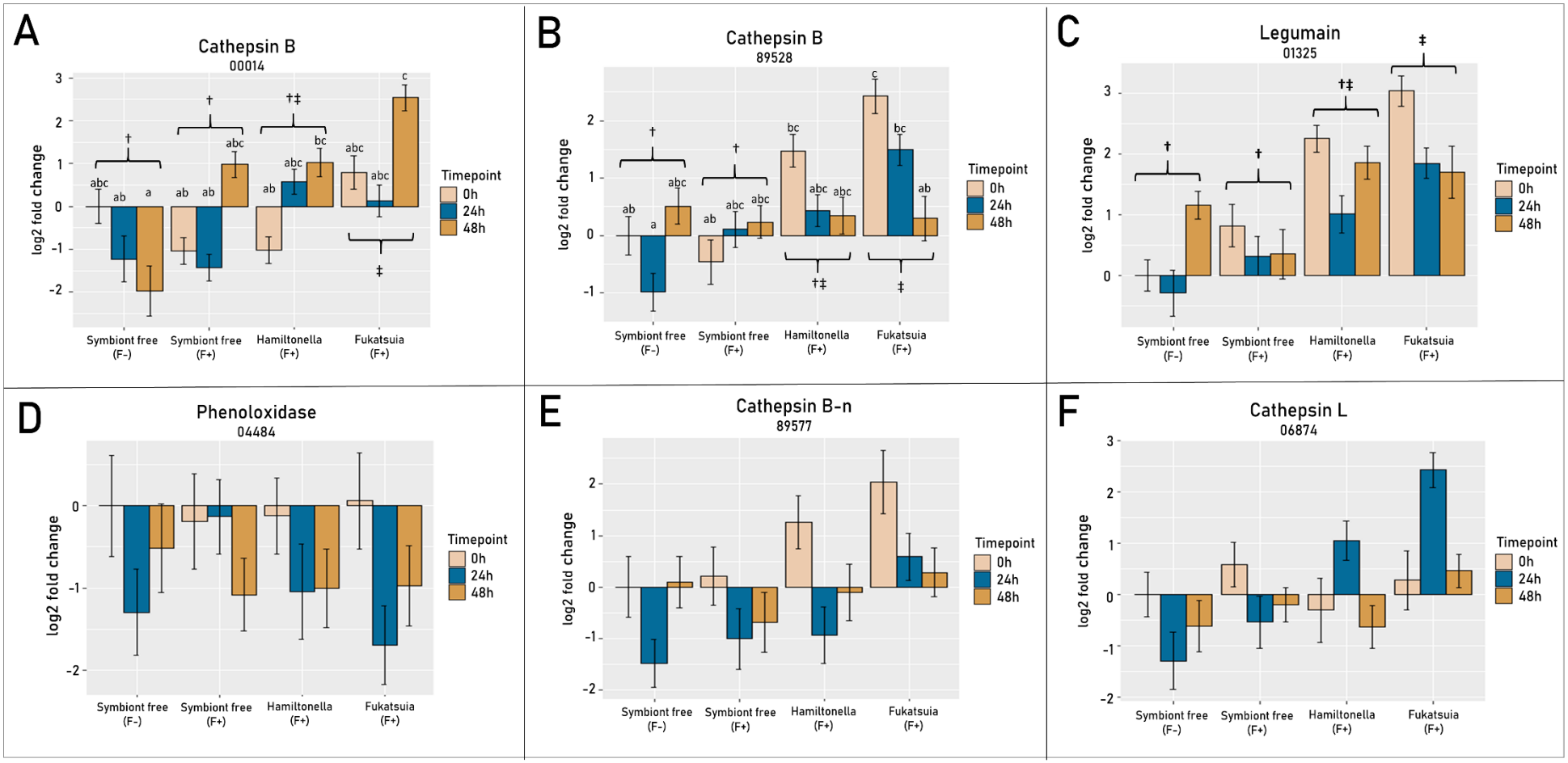


Figure 3-3: Log$_2$ fold change in the gene expression of six different immune genes, over three timepoints following infection with a fungal pathogen across four different treatments.

A) Cathepsin B (00014) B) Cathepsin B (89528) C) Legumain D) Phenoloxidase E) Cathepsin B-n F) Cathepsin L

The relative gene expression for each of the immune genes was calculated using the $2^{-\Delta\Delta Ct}$ method and were log$_2$ transformed.
Timepoints: 0 h post-infection=peach, 24 h post-infection=dark blue, 48 h post-infection=gold.
F+ =infected with fungal pathogen, F- =no fungal pathogen infection. Error bars represent the standard error.
Letters are presented above bars of those with significant interactions and bars with the same letter are not significantly different.
Treatments are grouped by symbol where there was a significant effect of treatment. Treatment groups with the same symbol are not significantly different.
The significant effect of timepoint seen in Cathepsin B (00014), Phenoloxidase and Cathepsin B-n was unable to be denoted on the graph.
See Table 3-2 for details of statistical output.

Symbiont influence on the expression of immune genes

Following the investigation of the overall effects of the fungal pathogen infection, we further explored the expression of the immune genes within the facultative symbiont harbouring treatments, to see how the presence of a symbiont altered the host immune response and whether this changed following the fungal pathogen infection. We first analysed the expression across the treatments at 0 h post-infection, where the fungus would not yet have penetrated the cuticle of the aphid, to investigate whether simply harbouring a facultative symbiont alters the expression of these immune genes.

We found that in some of the immune genes there was an effect of harbouring a facultative symbiont at 0 h post-infection. In Cathepsin B (89528), there was a significant difference in expression ($F_{3,16}$=6.849, p=0.004; Figure 3-3B), with aphids harbouring *F. symbiotica* having higher expression than both symbiont free treatments, symbiont free F- (p=0.018) and symbiont free F+ (p=0.005). There was also a trend towards an increased expression of this gene in the *H. defensa* treatment, however, this was not significantly higher than either symbiont free F+ (p=0.069) or symbiont free F- (p=0.211) (Figure 3-3B). We saw a similar effect in Legumain ($F_{3,16}$=8.97, p=0.001; Figure 3-3C). Aphids harbouring *F. symbiotica* had higher expression of Legumain at 0 h than both symbiont free F- (p=0.001) and symbiont free F+ (p=0.016). In this gene, we found that those harbouring *H. defensa* also had significantly higher expression than the symbiont free F- treatment (p=0.015), although this was not significantly higher than symbiont free F+ (p=0.161), due to there being a slightly higher expression in symbiont free F+ than symbiont free F- at 0 h. There was no significant difference in the expression of Phenoloxidase ($F_{3,16}$=0.02, p=0.996; Figure 3-3D), Cathepsin L ($F_{3,16}$=0.16, p=0.922; Figure 3-3F), Cathepsin B (00014) ($F_{3,16}$=3.33, p=0.05; Figure 3-3A) or Cathepsin B-n ($F_{3,16}$=1.18, p=0.348; Figure 3-3E) at time 0 h across the different treatments, although the expression of Cathepsin B (00014) in the *F. symbiotica* treatment and Cathepsin B-n in both symbiont treatments was increased (Figure 3-3A,E).

Looking collectively across all timepoints, we found that, overall, aphids carrying facultative symbionts had higher expression of the immune genes than the symbiont free treatments, particularly those harbouring *F. symbiotica*. In the *F. symbiotica* treatment, there was significantly higher expression of Cathepsin B (00014), Cathepsin B (89528) and Legumain (See Table 3-2 for p values; Figure 3-3A-C). There was also a trend towards those carrying *H. defensa* having a higher overall expression of Cathepsin B (00014) and Legumain compared to the symbiont free F- treatment (Table 3-2, Figure 3-3A, C).

Within Cathepsin L, there was no change in expression across the three timepoints in the symbiont free treatments. There did appear to be an increase in expression at 24 h in the *H. defensa* and *F. symbiotica* treatments (Figure 3-3F), although these changes were not statistically significant (*F. symbiotica*; p=0.614, *H. defensa*; p=0.968).

Symbiont influence on immune gene expression trajectories following fungal infection

We examined the variation in the course of gene expression following fungal pathogen infection across treatments, to see if the immune gene trajectories were altered by the presence of a facultative symbiont. There was a significant interaction between timepoint and treatment in two of the six immune genes, Cathepsin B (00014) and Cathepsin B (89528). In the expression of Cathepsin B (00014), there was a similar pattern across all treatments infected with the fungus. There was a significant increase in expression at 48 h in the *H. defensa* (p=0.038) and *F. symbiotica* (p<0.001) harbouring aphids, compared to the expression of these genes in the symbiont free F- aphids, but not relative to their own expression at the previous timepoints (Figure 3-3A). There was also a trend towards an increase in the expression at 48 h in the symbiont free F+ treatment, when compared to symbiont free F- (p=0.068; Figure 3-3A), suggesting that the fungal pathogen is leading to a change in expression in this gene. In Cathepsin B (89528), unlike in the other treatments, the gene expression in the *F. symbiotica* carrying aphids was significantly reduced at 48 h, compared to 0 h (p=0.039), although the expression did not fall below the levels of

expression seen in the other treatments (p=1.00; Figure 3-3B). There was also a slight reduction in expression in the *H. defensa* treatment, though this was not significant. This shows that aphids harbouring symbionts, specifically *F. symbiotica*, have a different gene expression trajectory of Cathepsin B (89528) following fungal pathogen infection, compared to aphids without symbionts.

Table 3-2: Full table of statistics for each gene.

The table includes results from the repeated measures two-way ANOVA, including F statistic, degrees of freedom and p-value. The table also shows results from post-hoc tests carried out on significant results for factors where there were more than two levels and on significant interactions. For the results from post hoc tests following significant interactions, only a subset of the results are included. Across treatment-timepoint interactions were not relevant here and therefore are excluded.

Gene		*F* statistic	df	p-value
Phenoloxidase	Timepoint	5.51	2	**0.009**
	Treatment	0.16	3	0.923
	Treatment x timepoint	0.96	6	0.467
	Post-hoc comparison for timepoint			
	0 h vs 48 h			0.055
	0 h vs 24 h			**0.009**
	24 h vs 48 h			0.768
Legumain	Timepoint	2.75	2	0.0799
	Treatment	5.62	3	**0.008**
	Treatment x timepoint	1.19	6	0.339
	Post-hoc comparisons for treatment			
	Symb free (F-) vs Symb free (F+)			0.973
	Symb free (F-) vs *F. symbiotica*			**0.014**
	Symb free (F-) vs *H. defensa*			0.08
	Symb free (F+) vs *F. symbiotica*			**0.036**
	Symb free (F+) vs *H. defensa*			0.181
	F. symbiotica vs *H. defensa*			0.806
Cathepsin B-n	Timepoint	5.94	2	**0.007**
	Treatment	2.03	3	0.151
	Treatment x timepoint	0.56	6	0.758
	Post-hoc comparison for timepoint			
	0 h vs 48 h			0.115
	0 h vs 24 h			**0.005**
	24 h vs 48 h			0.402
Cathepsin L	Timepoint	0.77	2	0.47
	Treatment	1.92	3	0.167
	Treatment x timepoint	1.65	6	0.168

Gene		*F* statistic	df	p-value
Cathepsin B (00014)	Timepoint	4.28	2	**0.023**
	Treatment	8.42	3	**0.001**
	Treatment x timepoint **(¥)**	3.81	6	**0.006**
	Post hoc-comparison for timepoint			
	0 h vs 48 h			0.07
	0 h vs 24 h			0.908
	24 h vs 48 h			**0.03**
	Post-hoc comparison for treatment			
	Symb free (F-) vs Symb free (F+)			0.641
	Symb free (F-) vs *F. symbiotica*			**0.001**
	Symb free (F-) vs *H. defensa*			0.069
	Symb free (F+) vs *F. symbiotica*			**0.018**
	Symb free (F+) vs *H. defensa*			0.489
	F. symbiotica vs *H. defensa*			0.231

Subset of post-hoc comparisons for treatment x timepoint (¥)

within treatments	
Symbiont free (F-)	***p-value***
0h vs 24h	0.927
0h vs 48h	0.422
24h vs 48h	0.998
Symbiont free (F+)	
0h vs 24h	1.000
0h vs 48h	0.476
24h vs 48h	1.000
F. symbiotica	
0h vs 24h	1.000
0h vs 48h	0.606
24h vs 48h	0.237
H. defensa	
0h vs 24h	0.237
0h vs 48h	0.375
24h vs 48h	1.000

across treatments at same timepoint			
0h			***p-value***
Symb free (F-)	vs	Symb free (F+)	0.976
Symb free (F-)	vs	*F. symbiotica*	0.997
Symb free (F-)	vs	*H. defensa*	0.979
Symb free (F+)	vs	*F. symbiotica*	0.525
Symb free (F+)	vs	*H. defensa*	1.000
F. symbiotica	vs	*H. defensa*	0.541
24h			
Symb free (F-)	vs	Symb free (F+)	1.000
Symb free (F-)	vs	*F. symbiotica*	0.908
Symb free (F-)	vs	*H. defensa*	0.548
Symb free (F+)	vs	*F. symbiotica*	0.802
Symb free (F+)	vs	*H. defensa*	0.392
F. symbiotica	vs	*H. defensa*	1.000
48h			
Symb free (F-)	vs	Symb free (F+)	0.068
Symb free (F-)	vs	*F. symbiotica*	**<0.001**
Symb free (F-)	vs	*H. defensa*	**0.038**
Symb free (F+)	vs	*F. symbiotica*	0.803
Symb free (F+)	vs	*H. defensa*	1.000
F. symbiotica	vs	*H. defensa*	0.774

Gene		*F* statistic	df	p-value
Cathepsin B (89528)	Timepoint	2.39	2	0.108
	Treatment	4.67	3	**0.016**
	Treatment x timepoint **(¥)**	3.71	6	**0.007**
	Post-hoc comparison for treatment			
	Symb free (F-) vs Symb free (F+)			0.983
	Symb free (F-) vs *F. symbiotica*			**0.022**
	Symb free (F-) vs *H. defensa*			0.262
	Symb free (F+) vs *F. symbiotica*			**0.048**
	Symb free (F+) vs *H. defensa*			0.445
	F. symbiotica vs *H. defensa*			0.531

Subset of post-hoc comparisons for treatment x timepoint (¥)

within treatments	
Symbiont free (F-)	***p-value***
0h vs 24h	0.854
0h vs 48h	0.999
24h vs 48h	0.335
Symbiont free (F+)	
0h vs 24h	0.958
0h vs 48h	0.997
24h vs 48h	1.000
F. symbiotica	
0h vs 24h	0.891
0h vs 48h	**0.039**
24h vs 48h	0.653
H. defensa	
0h vs 24h	0.807
0h vs 48h	0.727
24h vs 48h	1.000

across treatments at same timepoint			
0h			***p-value***
Symb free (F-)	vs	Symb free (F+)	0.999
Symb free (F-)	vs	*F. symbiotica*	**0.044**
Symb free (F-)	vs	*H. defensa*	0.565
Symb free (F+)	vs	*F. symbiotica*	**0.008**
Symb free (F+)	vs	*H. defensa*	0.201
F. symbiotica	vs	*H. defensa*	0.950
24h			
Symb free (F-)	vs	Symb free (F+)	0.883
Symb free (F-)	vs	*F. symbiotica*	**0.037**
Symb free (F-)	vs	*H. defensa*	0.619
Symb free (F+)	vs	*F. symbiotica*	0.652
Symb free (F+)	vs	*H. defensa*	1.000
F. symbiotica	vs	*H. defensa*	0.903
48h			
Symb free (F-)	vs	Symb free (F+)	1.000
Symb free (F-)	vs	*F. symbiotica*	1.000
Symb free (F-)	vs	*H. defensa*	1.000
Symb free (F+)	vs	*F. symbiotica*	1.000
Symb free (F+)	vs	*H. defensa*	1.000
F. symbiotica	vs	*H. defensa*	1.000

Discussion

In this chapter, we aimed to explore the influence of facultative symbionts on the immune response of aphids, and how this affects their response to a fungal pathogen infection. Firstly, we found that those without a facultative symbiont and those harbouring the non-protective symbiont *H. defensa* were highly susceptible to the pathogen, whereas those harbouring *F. symbiotica* only showed a moderate level of susceptibility. This suggests there is some level of protection provided by *F. symbiotica* to this pathogen. However, although *F. symbiotica* was able to prevent sporulation, this did not lead to increased survival, and there was a high amount of non-sporulating dead aphids in this treatment. But there was also a high proportion of aphids not infected with the pathogen that died, suggesting these may have died due to other explanations, such as stress. This may in turn explain why the survival rate in *F. symbiotica* harbouring aphids was lower than expected given the sporulation proportion.

In addition to understanding how the symbionts influence the immune response of the aphids, we also aimed to understand whether these responses differed when the symbiont was protective, in the case of *F. symbiotica* or not, as in *H. defensa*. To this end, we investigated both the cellular immune response and the expression of candidate immune genes of aphids harbouring the different facultative symbionts, and those without, following fungal pathogen infection.

Firstly, we looked at the levels of cellular immunity before the fungal pathogen had invaded the aphid, to see whether simply harbouring a facultative symbiont altered the cellular immunity of the aphid. Previous studies have shown links between harbouring symbionts and the number of immune cells in aphids, however, there have been mixed results in this area. Some show that harbouring facultative symbionts leads to increased numbers of circulating immune cells (Laughton et al. 2016), potentially used by the host to control the population of the symbionts within the haemolymph (Laughton et al. 2016). Alternatively, others have shown that there are decreased numbers of immune cells circulating in those

with facultative symbionts (Schmitz et al. 2012), whilst some had no effect (Schmitz et al. 2012). This could suggest that there are symbiont and aphid genotype interactions occurring, which may be leading to different host-symbiont combinations having different levels of effect on the host immune response. In this study, we found no difference in the number of cells across all treatments, suggesting that in these particular aphid-symbiont combinations, harbouring a facultative symbiont does not affect the cellular immunity of the host.

We then investigated the cellular immune response to the fungal pathogen over the first 72 hours of infection. We found that in the facultative symbiont free aphids there was a cellular immune response to the pathogen, with a significant increase in cell counts at 48 hours post-infection, which did not occur in those not infected with the fungus. This, however, did not occur in all treatments. In *H. defensa* there was also an increase in the number of immune cells at 48 hours, though this was not significant as the number of cells counted in each replicate was quite variable in this treatment. However, it seems as though, in those harbouring *H. defensa*, there was some cellular response to the pathogen, as seen in the symbiont free F+ treatment. Interestingly, there was no change in the cell counts across all the timepoints in the *F. symbiotica* treatment. This is particularly interesting as this symbiont has previously been seen to protect against a fungal pathogen (Heyworth & Ferrari 2015) and appears to be providing some form of protection against sporulation from this pathogen too. This could suggest that here, the aphid is relying on its symbiont to protect against this pathogen, not needing to invest in the cellular immunity it would have to without the presence of this symbiont. Alternatively, as the increase in cell counts we see in *H. defensa* is not as large as that in the symbiont free treatment, it could be that both symbionts are leading to a depression of the immune response here, *F. symbiotica* more so than *H. defensa*.

Previous studies have interpreted changes in immune cell numbers differently. In some cases, an increased immune response has led to higher numbers of immune cells (Kacsoh & Schlenke 2012, McGonigle et al. 2017), however, there have also been studies which

suggest that in response to some invading pathogens, such as fungal pathogens, immune cell numbers decrease (Gillespie et al. 2000, Hung & Boucias 1992). Studies on the pea aphid response to a fungal pathogen showed no change in immune cell numbers following infection with the fungal pathogen, *P. neoaphidis* (Parker et al. 2017), yet in response to the pathogen used in this study (*L. lecanii*), there was an increase in cell numbers following infection. This could suggest that, potentially, this fungal pathogen is more virulent than *P. neoaphidis* and thus needs a greater immune investment. Or it may be that this pathogen does not infect the aphid in the same way as *P. neoaphidis,* leading to differences in the aphid immune response across these two pathogens.

Given there is often difficulty interpreting the cellular immune responses, and to better understand the full effect of the interaction between the symbiont, the host and the fungal pathogen, we also examined the changes in expression of six immune genes. These candidate genes were chosen as all, but one had previously been shown to be upregulated in the pea aphid in response to infection with *P. neoaphidis* (Parker et al. 2017).

Interestingly, overall, we saw less of a response in this set of genes than expected. Across all the genes, we saw no significant change in expression in the symbiont free aphids across the first 48 hours of infection. This is particularly intriguing given that the rise in immune cells would suggest there is some form of immune response being exhibited here. However, in the expression of Cathepsin B (00014), we saw a strong trend towards an increase in expression at 48 hours post-infection, suggesting that at least this gene is involved in the immune response to the pathogen. Similarly, the expression of Phenoloxidase, Legumain and Cathepsin B (89528) did not decrease at 24 hours post-infection in the symbiont free fungus infected treatment as much as they did in the fungus free treatment. So, although there was no significant difference in the expression of these genes compared to the fungus free treatment, and they did not upregulate as we might expect, the presence of the fungal pathogen did appear to have some effect on their expression.

We hypothesised that the presence of facultative symbionts may alter the course of immune gene expression in response to the pathogen. We found no significant change in expression of Cathepsin B (89528) in the symbiont free F+ treatment across the three timepoints, however, this differed in those with facultative symbionts. In the *H. defensa* treatment, there was a non-significant reduction in the levels of gene expression at 24 hours, relative to 0 hours, leading to similar levels of gene expression to the symbiont free treatments at this timepoint. The expression was then unchanged at 48 hours. In the *F. symbiotica* treatment, there was also a reduction in gene expression relative to the expression in this treatment at 0 hours. By 24 hours, the expression in this treatment had reduced slightly but was still higher than the expression within the other treatments at the same timepoint. It then significantly reduced by 48 hours, bringing the expression of this gene down to a similar level to that seen in the other treatments. It may be that this gene is not involved in the immune response to the fungus, given we saw no change in expression in the symbiont free fungus infected aphids. But, retaining increased expression of the gene seen at 0 hours in the symbiont treatments may be detrimental following the pathogen infection, as the host cannot concentrate resources to the most dominant immune threat. In weevils, for example, their primary endosymbiont leads to a decreased immune response to pathogens, compared to those without symbionts, which was hypothesised to be due to the presence of the symbiont focussing the host immune resources elsewhere (Vigneron et al. 2012). This may explain why we see a decrease in expression in these treatments, as aphids reallocate immune resources to pathogen defence rather than those for symbiont related functions. On the other hand, the changes in expression we see could instead be due to an interaction occurring between the symbiont and the fungus, which in turn is leading to the reduction of the expression of these genes.

Beyond this, however, in most of the genes explored here, the symbionts did not appear to alter the course of gene expression following fungal pathogen infection, following a similar pattern to the symbiont free F+ treatment. For example, in Cathepsin B (00014) there is also an increase in expression at 48 hours post-infection in the facultative symbiont

treatment, again suggesting this gene is playing a role in the immune response to this pathogen. In others, where there is downregulation at 24 hours, for example in Legumain, this also occurs in the symbiont treatments, although the expression often remains higher than the symbiont free treatments.

We also saw differences across treatments in immune gene expression following infection in Cathepsin L, though these were not significant. Previous studies investigating aphid gene expression following *P. neoaphidis* infection saw no change in the expression of this gene (Parker et al. 2017), suggesting it was not playing a role in immune defence. Here, we similarly saw no change in expression in the symbiont free treatments. However, we did see an increase in expression in the symbiont harbouring treatments at 24 hours (although not significant). Cathepsin L is thought to respond to bacteria and has previously been seen to play a role in the control of symbiont populations (Futahashi et al. 2013, Renoz et al. 2015). As such, it is likely that the increase in expression of this gene at 24 hours in the facultative symbiont harbouring treatments is related to the control of these symbiont populations, rather than an immune response to the pathogen infection.

One of the most interesting findings of this chapter was the increased expression of the immune genes at 0 hours post-infection in those harbouring facultative symbionts, particularly *F. symbiotica*. We hypothesised that facultative symbionts, outside of pathogen influence, may independently be impacting the host immune response and we found this to be the case in some of the genes tested. It should be noted, however, that although the fungus has not yet penetrated the cuticle of the aphid at 0 hours, it is not to say the aphid is not able to sense this infection another way. Although, if this were the case, we would expect to see this effect occurring in all aphid treatments infected with the fungus, which would not explain the higher expression in only those harbouring facultative symbionts. In two of the six genes investigated (Cathepsin B (89528) and Legumain), there was significantly higher expression at 0 h post-infection, in the *F. symbiotica* treatment compared to both symbiont free treatments. There was also a trend towards increased

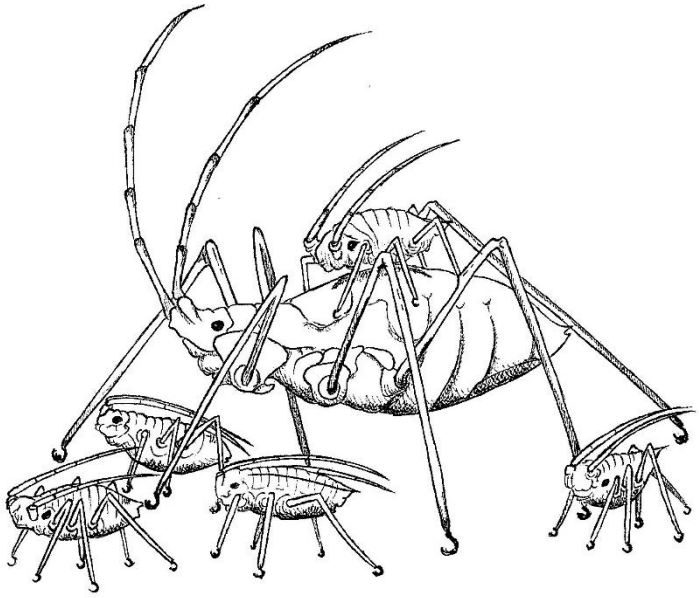
expression at 0 hours in the *H. defensa* treatments in these genes. Cathepsin B-n also had increased levels of expression at 0 hours in the symbiont harbouring treatments, although this was not significant. Given the increased expression at 0 hours is only seen in the facultative symbiont treatments, this suggests that the presence of a symbiont is leading to the increase in expression in these genes. There was no increased expression of Cathepsin L at 0 hours and, as highlighted above, Cathepsin L may play a role in facultative symbiont control. Therefore, unless the populations were required to be controlled at this timepoint, we may not expect to see expression levels of this gene being any higher than in the symbiont free treatment. Likewise, the lack of increased expression at 0 hours in Phenoloxidase (PO) is not unexpected as PO is cytotoxic and needs to be carefully regulated (González-Santoyo & Córdoba-Aguilar 2012). If PO was constantly activated by the presence of a symbiont, this may lead to unnecessary host death. It is unclear, however, why there was no upregulation of Cathepsin B (00014) at 0 hours in the facultative symbiont treatments, especially given the other Cathepsin B genes appeared to show this to some extent. Interestingly though, many of the other genes showed down-regulation following the fungal pathogen infection, whereas Cathepsin B (00014) showed upregulation. This could indicate that these genes are responding in different ways to the fungal pathogen and/or the presence of the symbionts, potentially playing different immunological roles within the host. Given the diversification of Cathepsin B genes across aphids (Rispe et al. 2008) and their different roles they have been seen to play (Kutsukake et al. 2004, Mathers et al. 2017, Pinheiro et al. 2017), it is perhaps not unexpected to see these genes responding if different ways, as they may provide different functions in this host. Investigating where these genes are differentially expressed within the aphid, may give more insight into the roles these genes are playing both before and after pathogen infection.

Although there is increased expression of many of these immune genes at 0 hours in the facultative symbiont lines, this does not necessarily benefit the immunity of the aphid to, at least this, particular invading pathogen. Some studies suggest harbouring symbionts can

lead to an increased immune response to pathogens (Eleftherianos et al. 2013, Horak et al. 2020, Moreira et al. 2009), essentially "priming" the immune system against attack. However, although those with *F. symbiotica* had some protection against the fungal pathogen, the increased expression was also seen in *H. defensa* harbouring individuals, which showed high susceptibility to the fungal pathogen. So, whilst there was increased immune gene expression, this did not lead to increased resistance. It may be that this priming of the immune system only occurs in response to pathogens of the similar "type" i.e. bacterial symbionts prime against bacterial pathogens, rather than simply causing a heightened immune response, able to deal with any threat. Although, studies investigating the effect of facultative symbionts and aphids protection against bacterial pathogens suggest that this is not necessarily the case either (ter Braak et al. 2013), however, this may be more related to the aphids lacking many of the components thought to be involved in bacterial pathogen defence (Gerardo et al. 2010). It is not uncommon to see mixed results from symbiont influence on host immunity against different invading organisms. For example, whilst *Wolbachia* in *Aedes aegypti* can provide resistance against a range of organisms, including pathogens, viruses and parasites (Kambris et al. 2009, Moreira et al. 2009), in *Drosophila, Wolbachia* is only able to protect against RNA viruses, not bacteria or parasitoids (Rottschaefer & Lazzaro 2012, Teixeira et al. 2008, Wong et al. 2011). It is still unclear what the molecular basis to differences in symbiont ability to prime against various invaders is, although these likely differ across host-symbiont relationships.

Overall, we found that following fungal pathogen infection there was a cellular immune response, although this was only in the symbiont free treatment, was lessened in the *H. defensa* treatment, and did not occur in the *F. symbiotica* treatment. This suggests that harbouring a facultative symbiont affects the cellular immune response to this fungal pathogen, and, in particular, harbouring a protective symbiont leads to no cellular immune response. In the six candidate immune genes explored in this study, we found some evidence of an immune response, although not as strong as previously reported in response to a different fungal pathogen in the same set of genes. Furthermore, outside of

two genes, the expression patterns of these genes following the infection were largely unaffected by the presence of a facultative symbiont. Interestingly, however, we did find that before the influence of the fungal pathogen, the presence of a facultative symbiont, particularly, *F. symbiotica* led to increased gene expression. This was also seen, to a lesser extent, in aphids harbouring *H. defensa* which are highly susceptible to the pathogen. This, therefore, suggests that the presence of these symbionts is not "priming" the immune system of the host against attack from this pathogen, as is seen in other systems. There is clearly a highly complex interaction between the facultative symbionts and the host immune system which, in this system at least, does not appear to be altering their defence against a pathogen. Further work potentially exploring the entire transcriptome of aphids with and without symbionts, following infection, would be beneficial in order to fully understand the extent to which these symbionts are influencing host immunity, and how this host-microbe interaction may be altered by the introduction of a pathogen or parasite.



Chapter Four

From parasite to mutualist: the genome evolution of an aphid symbiont within and across host species

Abstract

Close associations between eukaryotic hosts and bacteria are ubiquitous and these can be both beneficial and disadvantageous to the host. Aphids harbour a number of bacterial endosymbionts, many of which benefit the aphid, helping tackle various environmental stressors. Here we focus on the symbiont *Fukatsuia symbiotica*. *F. symbiotica* is known to protect against natural enemies but also carries a cost, lowering the fecundity of the host. The level of benefit and cost conferred by *F. symbiotica* varies across isolates. Within this symbiont species there is also a range of lifestyles. Five isolates identified within pea aphids in the UK live largely as a facultative mutualist, providing some defensive traits, however a North American strain appears parasitic, conferring no known benefits to its host. Furthermore, *F. symbiotica* has been found in the aphid genus *Cinara*, where it lives as a co-obligate partner with *Buchnera*, the obligate endosymbiont of aphids, helping with the synthesis of amino acids, a function usually provided by *Buchnera*.

We have built and analysed the genomes of five facultative mutualist isolates of *F. symbiotica* which convey different host phenotypes. We compared these genomes to previously published *F. symbiotica* strains and other closely related symbiont species. Investigating genomes across the large variation in lifestyle of *F. symbiotica* may help inform how these symbionts evolve along the continuum of parasite to mutualist. In addition, we compared the genomic differences within the five facultative mutualists from UK pea aphids, to attempt to uncover potential mechanisms for defensive traits conferred by these isolates.

F. symbiotica has a larger genome size than other closely related aphid facultative symbionts, who have undergone genome reductions due to close associations with their hosts. *F. symbiotica* has also retained more metabolic abilities and pathogenic factors such as toxin genes. Within *F. symbiotica*, the isolate found in *Cinara* has a smaller genome size and carries less mobile DNA, although beyond this there are few differences across the genomes. This may suggest that the switch to co-obligate lifestyle in *Cinara* may be quite recent. We also investigated the variation in the facultative isolates that confer different

levels of protection from natural enemies, outlining potential mechanisms for this protection.

Although they share the same host species, the *F. symbiotica* genome carries factors which suggest they are more pathogenic or less host-restricted than other aphid facultative symbionts. The genome suggests it lies at a mid-point between its free-living relatives and the other facultative symbionts, being more mutualistic than *Yersinia* spp. but having not evolved the same close host associations as *Hamiltonella defensa* and *Regiella insecticola*.

Introduction

Most insects form symbiotic relationships with bacterial species, many of which live inside their host as endosymbionts. These can greatly contribute to an insects' ability to survive in diverse conditions (Feldhaar 2011, Hilgenboecker et al. 2008, Moran et al. 2008, Oliver et al. 2010). The relationship between these microbes and their hosts can vary greatly, and lead to diverse ecological and evolutionary effects on their hosts. Some relationships can negatively affect the hosts, such as those who manipulate the reproduction of their host species (Stouthamer et al. 1999). Others, however, can be beneficial, protecting their hosts from natural enemies (Łukasik et al. 2013c, Oliver et al. 2003), heat stress (Heyworth & Ferrari 2015) or in the case of some obligate mutualists, assisting with nutrient acquisition (Douglas 1998).

There are some examples of bacterial clades which contain a mix of lifestyles, from obligate to facultative, suggesting these obligate associations evolve from facultative mutualists which undergo gene losses. One well-studied example is *Wolbachia*. *Wolbachia* is, within some hosts, a reproductive manipulator (Stouthamer et al. 1999), however, it has also been found to be beneficial in other insect hosts (Dedeine et al. 2001). In addition to this, in a nematode, *Wolbachia* has an obligate association with its host and is required for development and fertility (Taylor et al. 2005). Genomic insights into these symbionts can help us understand their evolution and how they transition from free-living species to obligate mutualists (Comandatore et al. 2013, Fenn & Blaxter 2006).

The pea aphid, *Acyrthosiphon pisum*, has become a model system for studying insect-microbe interactions. Almost all aphids carry the obligate, or primary, symbiont *Buchnera aphidicola*, which is required by the aphid to synthesise the essential amino acids absent from the aphids basic diet of plant sap (Akman Gündüz & Douglas 2009, Douglas 1998, Hansen & Moran 2011). Aphids can also be infected with a number of facultative, or secondary, endosymbionts (Oliver et al. 2010) and whilst these are not required for survival, they can often benefit the host. There are currently nine known species of aphid facultative symbionts, many of which provide benefits to their host (Guo et al. 2017). For example,

Hamiltonella defensa is known to defend aphids against the parasitoid wasp *Aphidius ervi* (Oliver et al. 2003, 2005) and a variety of symbiont species, including *Fukatsuia, Regiella*, *Rickettsia, Rickettsiella* and *Spiroplasma*, can protect against a fungal pathogen (Łukasik et al. 2013c, Scarborough et al. 2005). The parasitoid resistance conferred by *H. defensa* is attributed to a bacteriophage, APSE, with the level of protection varying with different APSE variants (Degnan & Moran 2008, Oliver et al. 2009, Weldon et al. 2012). There is also a strain of *Regiella insecticola* (strain 5.15) that provides protection against *A. ervi* and comparative genomics of this species highlighted potential mechanisms for this (Hansen et al. 2012). Beyond this, however, we have little understanding of the mechanisms behind these protective phenotypes.

Here we concentrate on one of the nine aphid symbionts, *Fukatsuia symbiotica*. This symbiont is found across a number of aphid species, although not as commonly as other aphid symbionts such as *H. defensa* and *R. insecticola* (Henry et al. 2015). Within the pea aphid, however, this symbiont appears to be fairly widespread and frequent, within certain populations (Ferrari et al. 2012, Henry et al. 2015, Russell et al. 2013, Smith et al. 2015). The isolates of this symbiont found in the UK can provide multiple benefits to their host. These include protection against the parasitoid *A. ervi* and the fungal pathogen *Pandora neoaphidis,* as well as heat stress when co-infected with *Spiroplasma* (Heyworth & Ferrari 2015). There are differences, however, between the isolates of *F. symbiotica* as to which, and how much of a, benefit they confer (Smee et al. 2021). There is another strain that is found commonly in North American pea aphids (Ap5D). However, unlike the isolates found in the UK pea aphids, this North American strain of *F. symbiotica* provides no known benefits to its host (Doremus & Oliver 2017). There are still high costs associated with being infected with this isolate of *F. symbiotica*, however this appears to be lessened when a co-infection occurs with *H. defensa* (Doremus & Oliver 2017). Further to the isolates found within pea aphids, *F. symbiotica* is also found in some species of the aphid genus, *Cinara*. Within *Cinara F. symbiotica* is prevalent, although appears to be quite species-specific, with 90% of the bacterial haplotypes found to be associated with a single *Cinara* species

(Meseguer et al. 2017). Many *Cinara* possess a di-symbiotic system, which usually occurs between *Buchnera* and *Serratia,* where *Serratia* complements the biosynthetic abilities of *Buchnera* (Lamelas et al. 2008). Interestingly, here it is thought that *F. symbiotica* replaces *Serratia* and lives as a co-obligate partner with *Buchnera* as part of the di-symbiotic system found in this genus (Meseguer et al. 2017). Compared to other closely related symbiont species, there is a high amount of variation within species in *F. symbiotica*. The lifestyles of the different isolates of this species range from parasitic in the case of Ap5D, being costly to the host, but conferring no benefits, to the strain found in *Cinara* which lives as co-obligate partner. The UK pea aphid isolates appear to sit somewhere between these two extremes, as facultative mutualists. They carry some cost, but also benefit their host. Further to the lifestyle differences, there is also a range of phenotypes associated with the facultative mutualists, providing different levels of defence against different natural enemies and some carrying more of cost than others (Smee et al. 2021).

Genomic investigations into the evolution of aphid facultative symbionts, including different strains of *H. defensa* and *R. insecticola*, have provided useful insights into the evolution of these species and the role they play in aphid ecology (Degnan et al. 2009, 2010; Hansen et al. 2012). The genomes of these facultative symbionts are often reduced compared to their free-living relatives, suggesting genomic reductions have occurred due to long-term, stable, associations with their hosts (Burke & Moran 2011b; Degnan et al. 2009, 2010). They still retain some pathogenic factors, such as toxin genes, and have a higher percentage of mobile DNA than obligate species, however, this is common of many insect facultative mutualists (Moran et al. 2008). Comparing across isolates of the same species has also allowed for the identification of potential mechanisms for defensive traits conferred by these symbionts, for example, the parasitoid resistance provided by one strain of *R. insecticola* (Hansen et al. 2012).

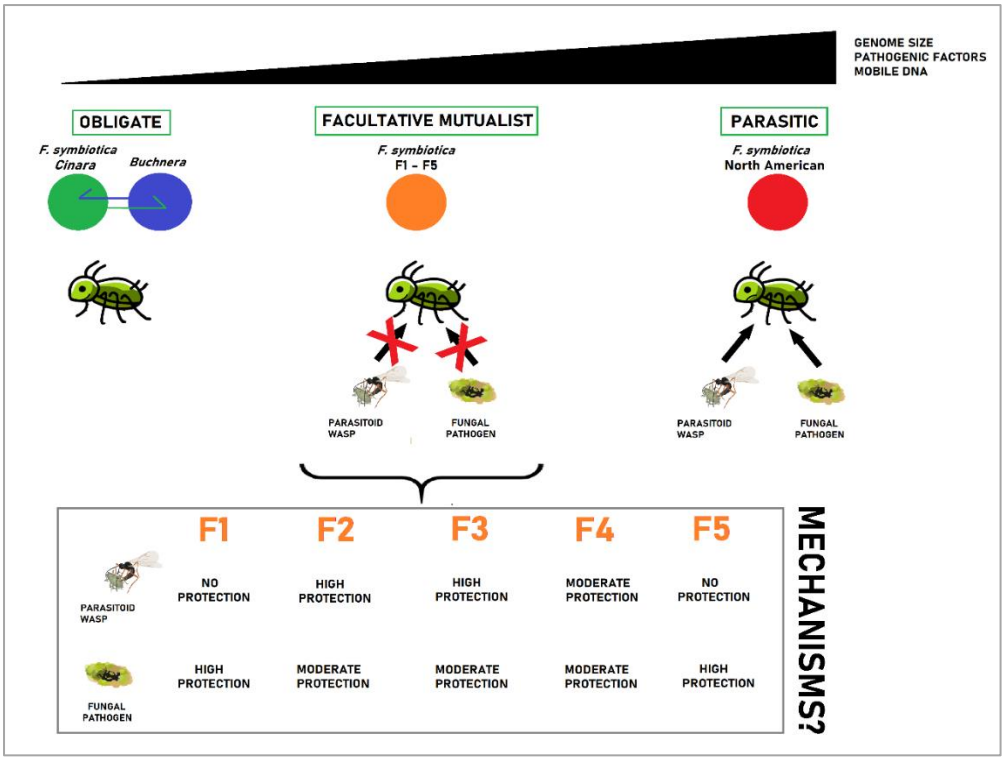


Figure 4-1: Graphical representation of chapter including background information, aims and hypotheses.

Here, we present draft genomes of five facultative mutualist isolates of *F. symbiotica* which convey different host phenotypes, comparing them to previously published *F. symbiotica* genomes and other closely related symbiont species. There is a large amount of phenotypic variation across hosts harbouring the different isolates of *F. symbiotica*, from the parasitic North American strain (Ap5D) to the *Cinara* obligate mutualist. By comparing across these isolates, we aim to further our understanding of how these genomes evolve; which elements are lost and retained in the transition into a more mutualistic lifestyle. In addition to the overall variation within *F. symbiotica*, there are differences in the defensive properties and costs associated with the five facultative mutualist isolates that were found in UK pea aphids (Smee et al. 2021). By investigating the genomic differences between these isolates, we aim to uncover any potential mechanisms for the defensive traits conferred by these isolates of *F. symbiotica*. In addition to comparing within species, we also compare *F. symbiotica* to the closely related aphid facultative symbionts, *H. defensa*

and *R. insecticola*, to highlight how the *F. symbiotica* genome differs from closely related species and how this may reflect the aphid phenotypes it confers.

Methods

Isolate origins

The five isolates, F1-5 were from pea aphids collected from two different UK locations, all of which were collected from *Medicago sativa*. Table 4-1 outlines the aphid clone, symbionts present in each aphid clone, and the location and date of collection. Most of the *F. symbiotica* isolates were in co-infections with *H. defensa* except for F4 which was a single infection, and F5 which was in a co-infection with *Spiroplasma*. We were unable to cure the *Spiroplasma* infection, and so all analyses of isolate F5 also contain *Spiroplasma*.

Table 4-1: Details of the origin of the five facultative pea aphid *F. symbiotica* isolates F1-5. Details include the aphid clone, symbionts present in each infection and the collection site and date. All lines were collected from *Medicago sativa*.

Isolate	Aphid clone	Symbionts	Collection site	Collection date
F1	218	*Hamiltonella & Fukatsuia*	Eling	May 2010
F2	238	*Hamiltonella & Fukatsuia*	Beaconsfield	May 2010
F3	236	*Hamiltonella & Fukatsuia*	Beaconsfield	May 2010
F4	211	*Fukatsuia*	Eling	May 2010
F5	217	*Fukatsuia & Spiroplasma*	Eling	May 2010

Genome sequencing, assembly and annotation

To reduce the amount of eukaryotic DNA being sequenced, haemolymph was extracted from the symbiont harbouring aphids and used for DNA extractions. The paired-end DNA sequencing was carried out using an Illumina MiSeq, and the library preparation and sequencing was carried out at the University of York Technology Facility (F2-5) and the University of Liverpool Centre for Genomic Research (F1). The read lengths for isolates F1-5 were 4521071, 6386539, 7271257, 7681809 and 6548524 respectively.

Raw sequences were trimmed using Trimmomatic (Bolger et al. 2014) using default parameters. As the sequences contained aphid and *Buchnera* DNA, the aphid and *Buchnera* sequences needed to first be removed before the symbiont genomes could be assembled. To do this the sequences were mapped to the published pea aphid (GCA_000142985.2) and *Buchnera aphidicola* str. APS (BA000003.2) genomes and the mapped sequences removed. The remaining sequences were used to assemble the genomes using A5-miseq (Coil et al. 2014), using the default parameters. Genes were predicted using PROKKA v1.12 (Seemann 2014) and RAST-Classic (Aziz et al. 2008), again using default parameters.

Phylogenetic reconstructions

Two core gene phylogenetic reconstructions were created. Firstly, a phylogeny of the seven *F. symbiotica* isolates was created using 1854 single copy orthologous genes (SICO) shared across the isolates. Secondly, a larger phylogenetic reconstruction was created using 550 SICO shared across several strains of nine bacterial species. The species included were chosen using previously published phylogenies of aphid facultative symbionts (Degnan et al. 2009). Here, we also included five newly sequenced genomes of *H. defensa* (H1-5), the genomic details of which are discussed in Chapter Five. Previously published genomes (NCBI references can be found in Appendix A) were downloaded from NCBI and annotated using the same pipeline to avoid bias due to different gene annotation methods. RAST-Classic annotations of all genomes were used for this analysis. For both phylogenetic reconstructions, the one-to-one orthologs were identified using OrthoMCL (Chen et al. 2006). The gene sequences of the orthologous genes were then concatenated and aligned using MAFFT (Katoh et al. 2002). The phylogenetic reconstructions were created using MAFFT with 1000 bootstrap replicates. All analyses were carried out using the default parameters unless specified.

Phylogenetic reconstructions of the SPI-1- and SPI-2-like type III secretion systems and of the Toxin B genes found within the seven isolates were also created. The type III secretion system phylogeny was created using concatenated sequences of homologs of four of the

export apparatus genes (SsaR/SpaP, SsaS/SpaQ, SsaT/SpaR, SsaU/SpaS), across eight different bacterial species. The Toxin B phylogeny was created using the annotated Toxin B genes found within the seven *F. symbiotica* genomes. For both reconstructions, the sequences were aligned, and the reconstructions created using MAFFT with 1000 bootstrap replicates.

Comparative genome analyses

Genomes were first compared using the dnadiff comparison tool from MUMmer (Kurtz et al. 2004). dnadiff outlines the base to base similarity of the genomes, which is particularly useful when comparing different isolates of the same species, allowing for the confirmation that the isolates are genetically distinct and then to assess how similar those isolates are at the nucleotide level. Orthologous gene group were found using OrthoMCL (Chen et al. 2006), which uses an all-against-all protein blast to identify, and then cluster, orthologous gene groups across different isolates of the same and different species. Previously published genomes of the closely related aphid facultative symbiont *R. insecticola* (strain LSR1 and 5.15; GCA_000143625.1, GCA_000284655.1) were included in these analyses and were also annotated using PROKKA and RAST to avoid gene annotation bias. Functional annotation was carried out using KAAS (KEGG Automatic Annotation Server) (Moriya et al. 2007) using the BLAST search program and a template data set of 35 different bacterial species (Appendix B). We also searched for secondary metabolite gene clusters within the genome using antiSMASH v5.0 (Blin et al. 2019). Secondary metabolites are non-essential molecules which are not required for survival but are produced by a range of organisms (O'Connor 2015). They are often specialised, and different metabolites carry out a wide range of biological functions (Jenke-Kodama et al. 2008). We identified three different types of mobile DNA; insertion sequences, prophages and plasmids. The online tool ISsaga (Varani et al. 2011) was used to predict the insertion sequences within each genome, the prophage regions were identified using PHASTER (Arndt et al. 2016) and PLSDB was used to search for plasmids (Galata et al. 2019). All analyses were carried out using default parameters unless specified.

Results and discussion

We assembled and annotated draft genomes of five isolates of *F. symbiotica*. The genome size of *F. symbiotica* is around 3.2 Mb, with some variation between isolates (F1 = 3.6 Mb, F2 = 3.2 Mb, F3 = 3.3 Mb, F4 = 3.2 Mb, F5 = 4.8 Mb). The much larger genome size of F5 is due to the inclusion of the *Spiroplasma* genome. The average %GC content is around 42 %, although this too varies with isolate (F1 = 41 %, F2 = 42.7 %, F3 = 42.4 %, F4 = 42.6 %, F5 = 36.2 %). Again, the %GC content of F5 varied likely due to *Spiroplasma*. The genome of the published Ap5D strain of *F. symbiotica* is a similar size, although slightly smaller (3.1 Mb) (Patel et al. 2019). The genome of the *F. symbiotica* strain found in *Cinara* was originally assembled to be around 2 Mb, much smaller than the other isolates (Meseguer et al. 2017). However, a recent study reassembled and annotated this isolate and found it to be 2.79 Mb (Patel et al. 2019). This is much larger than previously assumed, but still smaller than the pea aphid *F. symbiotica* isolates. What is still unclear, however, is whether this difference is due to genome reductions that have occurred due to this strains switch in lifestyle to co-obligate (Meseguer et al. 2017), or whether this is purely due to the quality of the sequencing and thus assembly of the genome. Throughout the remainder of this analysis, we have used the originally assembled genome of this isolate (2 Mb), as at the time of carrying out these analyses we did not have access to the sequence data or the re-assembled genome of this strain. Comparing the phenotypes of isolates F1-5, although they all differ slightly, isolate F1 appears to produce phenotypes that differ most from the other isolates (Smee et al. 2021). For example, isolate F1 confers a very high level of protection against the fungal pathogen, *Pandora neoaphidis,* whereas others only provide moderate protection. F1 also provides much less protection against the parasitoid *A. ervi* than the other isolates (Smee et al. 2021). Additionally, F1 is the mostly costly to carry with aphids infected with this isolate having the lowest fecundity (Smee et al. 2021). Interestingly, isolate F1 has a slightly larger genome than the other isolates, which may suggest more phenotypic differences compared to the other isolates. Usually, more parasitic or free-living bacteria will have larger genome sizes as they have undergone less genome reduction, which occurs as bacteria become host-restricted and vertically

transmitted. As the genome of F1 is slightly less reduced than the other three facultative symbionts (unclear for F5 due to *Spiroplasma*), this suggests it may not have evolved an as close relationship with the host, which may in turn be related to the more costly phenotype associated with isolate F1 (Smee et al. 2021).

Overall, with the exception of the *Cinara* isolate, the genome of *F. symbiotica* is larger than that of its most closely related symbionts, with *H. defensa* having a genome size of 2.1 - 2.6 Mb and *R. insecticola* around 2 Mb (Degnan et al. 2010), both having undergone large genome reductions.

F. symbiotica's genome size is reflected in the number of coding sequences (CDS) with isolates F1-5 having 3114, 2981, 3055, 3007 and 5268 CDS respectively. Of these, around 30 % were annotated as hypothetical proteins with around 60 % of those being <100 amino acids in length. The *F. symbiotica* genomes also contained 204 of the 205 single copy genes found in all *Gammaproteobacteria* (Lerat et al. 2003) suggesting the genomes are well represented and there are no large gaps in sequencing.

We re-annotated the published sequences of Ap5D and the *Cinara* strain to ensure during comparison there was no bias due to different analysis techniques. Within the genome of the Ap5D we found 3419 CDS, which is slightly higher than that found previously (2929 CDS; Patel et al. 2019), with a similar percentage of hypothetical proteins to the other pea aphid *F. symbiotica* isolates. The *Cinara F. symbiotica* genome contained 1827 CDS, due to the smaller size, with only 18% of those being hypothetical proteins. The newly assembled genome of the *Cinara* strain found 2800 annotated CDS (Patel et al. 2019).

Table 4-2: Comparison of genome features of the seven *F. symbiotica* isolates and other relevant bacterial species.
F. symbiotica has a larger genome than other closely related facultative symbionts. Ap5D and isolates F1-4 are similar in size but the genome of the co-obligate *Cinara* strain is slightly smaller. All isolates have a similar %GC content. F1-5 and Ap5D carry the same three plasmids, one of which is absent from the *Cinara* strain.

	B. aphidicola APS	H. defensa 5AT	R. insecticola 5.15	F. symbiotica (1)	F. symbiotica (2)	F. symbiotica (3)	F. symbiotica (4)	F. symbiotica (5)	F. symbiotica (*Cinara*)	F. symbiotica (Ap5D)	S. glossindius	Y. pestisCO92
Chromosome, bp	640,681	2,110,331	2,035,106	3,607,515	3,173,350	3,271,889	3,206,892	4,760,724	2,794,581	3,131,594	4,171,146	4,653,728
Extrachromosomal elements	2	1	1	3	3	3	3	3	2	3	3	3
Total G + C (%)	26.2	40.1	42.4	41	42.7	42.4	42.6	36.2	43.4	43.5	54.7	47.6
Total predicted CDS	571	2,100	1,761	3,114	2,981	3,055	3,007	5,268	1827 / 2800	3,419	2,432	4,012
Average CDS size (bp)	984	812	856	803	837	821	832	604	917	795	873	998
rRNA operons	2	3	4	3	3	2	3	5	7	5	7	6
tRNAs	32	42	36	30	29	26	30	55	42	48	69	70
Lifestyle	Obligate	Facultative	Facultative	Facultative	Facultative	Facultative	Facultative	Facultative	Co-obligate	Parasitic	Facultative	Pathogen

Genome comparison: within and between species

The base to base comparison of the sequences of the *F. symbiotica* genomes show that, although all are similar, none are identical. Isolates F2-4 appear to be equally similar, showing a 96 % similarity to one another, as expected from isolates of the same species. Isolate F1 however, has a lower base similarity of 93 %. Due to the F5 genome also containing *Spiroplasma*, this isolate also has a reduced similarity of around 80 %. As expected, given they are from the same host species and the same geographical area, F1-5 are more dissimilar to Ap5D and *Cinara* strains than they are to each other. Isolates F2-4 showed around 91 % similarity to Ap5D, with F1 again differing slightly more (87 %). F2-4 share around 94 % base similarity with the *Cinara* strain, and F1 shares around 89 % base similarity. F5 has reduced similarities to Ap5D and the *Cinara* strain, with base similarities of 78 % and 80 % respectively, however this is due to the presence of *Spiroplasma*. Interestingly, isolates F1-5 share more base similarity with the *Cinara* strain than Ap5D, which is surprising given they come from a different host species. This could suggest that isolates F1-5 and *Cinara* have become more dissimilar to the North American strain as they have both evolved towards a more mutualist lifestyle. Alternatively, it may be that the North American strain has diverged from the other isolates, whilst they have remained similar, due to different selection pressures. It is surprising that isolate F1-5 are more similar to the *Cinara* strain than the North American strain, given they have a different host species, and this highlights that factors outside of the host may be playing a role in this symbiont evolution.

Investigations into orthologous genes found across the different isolates showed there were 1854 single copy one-to-one orthologues (SICO) across all seven *F. symbiotica* genomes. Across just the five UK pea aphid isolates (F1-5) there were a total of 2325 SICO. We used the 1854 SICO found across the seven genomes, to build an unrooted phylogeny of the seven *F. symbiotica* isolates (Figure 4-2). The phylogeny highlights the strong similarity across all these isolates, as the divergences that do occur are very small-scale. However, it shows that Ap5D, the North American strain, is the most diverged from the other isolates (Figure 4-2) which is also indicated by the percentage differences above.

Within isolates F1-5 from the same host species and similar location, F1 has diverged slightly from the other four isolates (Figure 4-2). Interestingly, F1 provides differing benefits and results in a higher fecundity cost to the host compared to that conferred by the other isolates (Smee et al. 2021).

Looking across species, we found that there were 1108 SICO between the isolates F1-5 of *F. symbiotica* and *H. defensa* isolate (H1), and 1240 between the *F. symbiotica* isolates and *R. insecticola*. This suggests *F. symbiotica* is more similar to *R. insecticola* than to *H. defensa*.

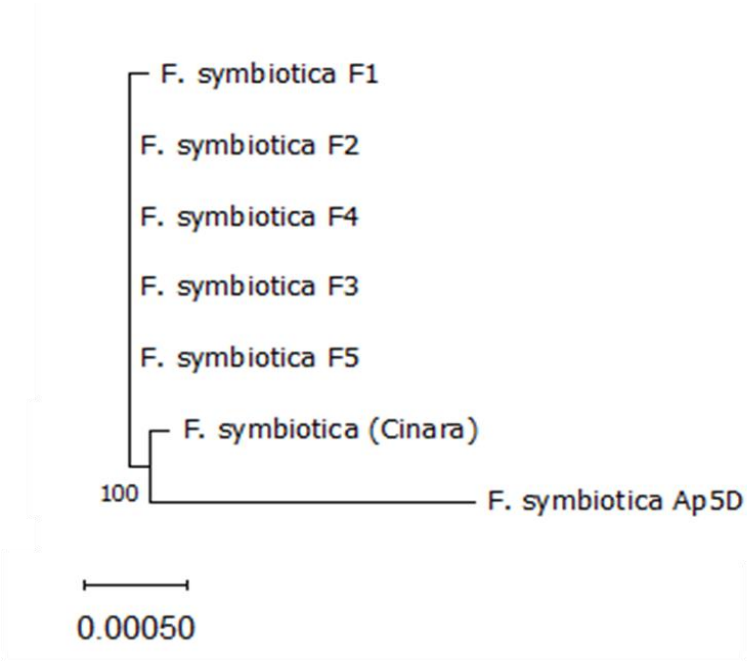


Figure 4-2: Phylogenetic reconstruction of the five newly sequenced isolates of *F. symbiotica* F1-5 and the two previously published *F. symbiotica* genomes, Ap5D (North American strain) and *Cinara* (strain found within *Cinara*).
The North American strain is the most diverged from the other six isolates. Within isolates F1-5 which are from UK pea aphids, F1 has slightly diverged from the other four isolates. F1 confers slightly different phenotypes to the host than the other isolates and leads to a higher fecundity cost. The differences across the isolates however are extremely small, with all isolates being highly similar.
The reconstruction was produced using concatenated amino acid sequences of single copy orthologous genes across the seven isolates (1854 genes) using MAFFT with 1000 bootstrap replicates.
The scale bar corresponds to 0.0005 amino acid substitutions per site.

Phylogenetic reconstruction

We created a phylogeny of the different isolates of *F. symbiotica* with other closely related aphid symbiont species, to see how much these isolates differed in terms of their core gene structure, and how diverged they were from other aphid facultative symbionts. The phylogenetic reconstruction (Figure 4-3) shows the three aphid facultative symbionts form a linage that is distinct from the other species included in the phylogeny, including *Yersinia* spp. which were previously reported to fall into a clade with *H. defensa* and *R. insecticola* (Degnan et al. 2009). Likewise, *F. symbiotica* appears to be most closely related to *R. insecticola*, with these two species forming a sister group, separate from *H. defensa* (Figure 4-3). Of the three symbionts, *H. defensa* has undergone the most divergence from the other closely related species (Figure 4-3). Looking across the *F. symbiotica* isolates, the North American strain, Ap5D, has diverged slightly from the other isolates (Figure 4-3), as seen in the phylogeny built from the 1854 SICO across the seven *F. symbiotica* isolates (Figure 4-2), but here too the difference is extremely small. This shows that the core SICO (550 genes) are highly conserved within the *F. symbiotica* isolates, including those within the genome of the *Cinara* strain. Therefore, although the *Cinara* may have undergone some genome reduction, the core genes within the genome remain conserved across *F. symbiotica,* regardless of their host species and lifestyle.

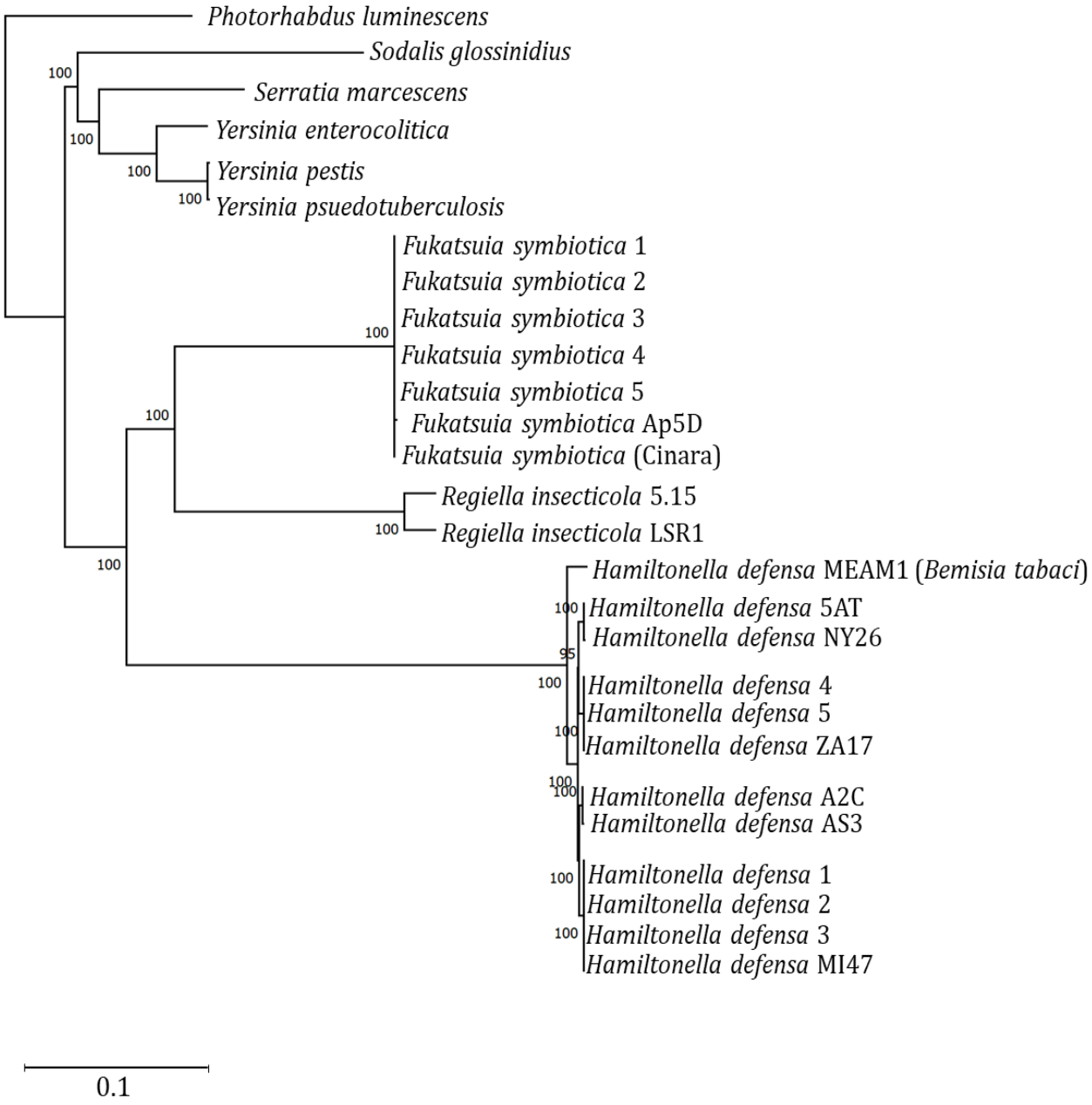


Figure 4-3: Phylogenetic reconstruction of the five newly sequenced isolates of *F. symbiotica* (F1-5), the two previously published *F. symbiotica* genomes and other closely related *Enterobacteriaceae*.

Here, we also included five newly sequenced genomes of *H. defensa* (H1-5).

All three aphid facultative symbiont species have diverged from their closely related free-living relatives. *R. insecticola* and *F. symbiotica* are sister species and form a sperate clade from *H. defensa* which is the most diverged of the three aphid facultative symbiont species. The reconstruction was produced using concatenated sequences of the single copy orthologous genes across the nine species (550 genes) using MAFFT with 1000 bootstrap replicates. The scale bar corresponds to 0.1 amino acid substitutions per site.

Metabolic reconstruction

Like other closely related symbionts, all the isolates of *F. symbiotica* can complete glycolysis, the tricarboxylic acid (TCA) cycle and the pentose phosphate pathway. They are also able to produce pyrimidines and purines. However, *F. symbiotica* differs from the other aphid facultative symbionts in essential amino acid biosynthetic pathways. *H. defensa* and *R. insecticola* have lost all the essential amino acid biosynthetic pathways except for threonine and lysine (Table 4-3). This is likely due to these symbionts having a long and stable relationship with the aphid host, leading to them also becoming reliant on *Buchnera* to synthesise these. It is unclear however why the symbionts have retained these two pathways. One reason for the retention of the lysine biosynthetic pathway may be due to *Buchnera*'s usage of this amino acid. Lysine usage in *Buchnera* is twice that required for the growth of *E. coli* (Shigenobu et al. 2000), and so *H. defensa* and *R. insecticola* may have retained this pathway to ensure they receive enough lysine for their own growth. Further to this, both lysine and threonine are energetically inexpensive to produce (Douglas 2016), so producing these would not come at much of a cost to the symbiont. In the whitefly system, *H. defensa* is known to produce threonine and lysine for its own growth, releasing none to the host (George & Toole 2017). This is linked to its ability to synthesise its own aspartate, required for these pathways. Studies have shown that when the aspartate needed for the biosynthesis was provided by the host, *H. defensa* would release some lysine, however, it would not release any threonine. Due to this, within whitefly, *H. defensa* is considered a 'nutritional parasite' (George & Toole 2017). These symbionts, therefore, may be synthesising these two amino acids for their own growth, whilst they can obtain the others from *Buchnera*'s synthesis.

In addition to threonine and lysine, *F. symbiotica* has retained the biosynthetic pathways for three other amino acids; valine, leucine and isoleucine (Table 4-3). These are also very energetically inexpensive to produce (Douglas 2016) and are three of the six amino acids that *Buchnera* lacks part of the pathway for, usually compensated for by the aphid (Hansen & Moran 2011). *F. symbiotica*, therefore, could also help generate these amino acids for the

aphid similar to *Buchnera*, or act as a co-obligate symbiont as expected in the isolate of *F. symbiotica* found in *Cinara*.

Table 4-3: Ability of five bacterial species, including three aphid facultative symbionts, to synthesise the ten essential amino acids.
Blue indicates the pathway is present.
Aphid facultative symbionts have lost the ability to synthesise many of the essential amino acids, although *F. symbiotica* has retained more of these than the other two aphid symbiont species (*H. defensa* and *R. insecticola*).

	F. symbiotica	*H. defensa*	*R. insecticola*	*S. marcescens*	*Y. pestis*
Arginine					
Histidine					
Isoleucine					
Leucine					
Lysine					
Methionine					
Phenylalanine					
Threonine					
Tryptophan					
Valine					

The loss of ability to synthesise many of the essential amino acids in *H. defensa* and *R. insecticola* has likely occurred due to their close associations with insect species which have an obligate symbiont, which they too can rely on for the synthesis of the essential amino acids. It is unclear, however, why *F. symbiotica* has not also lost the ability to synthesise these essential amino acids. This may be due to the relationship between *F. symbiotica* and the host, which may be less stable than the other closely related symbionts, which is also suggested by some of the other genome features of *F. symbiotica*. Alternatively, the retention of this function could be due to a benefit it confers to either the host, potentially supplementing *Buchnera*'s synthesis, or to the symbiont itself. For example, *F. symbiotica* is often found in co-infections with other facultative symbionts, and

at lower densities (Doremus & Oliver 2017; Chapter 2), and as such this symbiont may benefit from being able to synthesise more of its own amino acids, reducing resources to compete for.

Similar to *H. defensa* and *R. insecticola*, *F. symbiotica* can synthesise seven nonessential amino acids and can synthesise most essential vitamins. *H. defensa* is unable to synthesise thiamine and pantothenate, two essential vitamins, but *F. symbiotica*, whilst also lacking the ability to synthesise pantothenate, can synthesise thiamine. *R. insecticola* strain 5.15, like *H. defensa*, lacks the ability to synthesise either, however, strain LSR1, like *F. symbiotica*, can synthesis thiamine.

Mobile DNA

Mobile genetic elements (MGE) within the *F. symbiotica* genomes consisted of transposable elements (TEs) as insertion sequences (ISs), plasmids and phage islands. Insertion sequences are widespread, small transposable elements, that are an important component of many bacterial genomes and how they evolve (Chandler & Siguier 2013, Siguier et al. 2014). They are widespread, and due to there being so many types, these are often split into smaller families and groups (Chandler & Siguier 2013). Within isolates F1-5 and the *Cinara* strain, there were 16 IS families found. Previous explorations into Ap5D found 20 different IS families (Patel et al. 2019). There were slightly more found within F5, however this was likely due to the presence of *Spiroplasma*, as these did not occur in any of the other *F. symbiotica* isolates (Table 4-4). The number of open reading frames (ORFs) associated with ISs also differed across the isolates, and somewhat reflects what we might expect given the different lifestyles. Ap5D had the highest number (446), isolates F1-4 had between 295 and 324 and the *Cinara* strain had the least, with only 152 (Table 4-4). The number of ISs found within genomes can often be a predictor of how host dependent the bacteria are, as more mobile DNA is often found in more free-living bacteria, which do not go through strict bottlenecks as they are vertically transmitted (Siguier et al. 2014). All the isolates (except for F5) showed similarities in the abundance of ORFs associated with the

different families, with IS630 being the most abundance IS (Table 4-4). IS5 was the second most abundant IS. The genomes all carried a few different group types within IS5 but much of this was made up of the ISL2 group. A large number of ORFs associated with IS256 were found in F5, although this IS family was missing from F1-4 and *Cinara*. We initially assumed this was due to the presence of *Spiroplasma* in this genome, but Ap5D also has one ORF which is associated with IS256 (Table 4-4; Patel et al. 2019).

Table 4-4: Number of insertion sequence (IS) families contained in the seven *F. symbiotica* isolates. Light grey rows symbolise ISs that are likely in *Spiroplasma* as they only occur in F5. Ap5D had the highest number of ISs and the *Cinara* strain had the least, but all isolates shared similarities in their IS composition and the abundance of each family.

IS families	F1	F2	F3	F4	F5	Cinara	Ap5D
IS630	121	122	125	128	143	44	151
IS5	51	58	61	58	294	34	47
IS4	30	24	29	28	24	14	27
IS481	17	18	19	16	18	9	19
ISAs1	16	19	20	16	15	5	11
ISNCY	15	12	16	11	17	10	13
IS200/IS605	9	21	18	19	20	4	14
IS701	8	8	9	9	9	9	12
IS3	7	7	7	5	24	7	8
IS110	6	4	6	10	10	3	14
IS982	4	3	3	2	5	4	9
IS21	3	3	3	3	2	3	10
ISKra4	3	3	3	3	3	2	2
Tn3	2	2	2	2	2	1	44
ISL3	2	2	2	2	2	2	0
IS1	1	1	1	1	1	1	7
IS256	0	0	0	0	181	0	1
Retron-type RNA-directed DNA polymerase	0	0	0	0	0	0	21
Transposase	0	0	0	0	0	0	18
IS91	0	0	0	0	0	0	7
Mobile element protein	0	0	0	0	0	0	11
IS1595	0	0	0	0	2	0	0
IS982	0	0	0	0	5	0	0
IS30	0	0	0	0	50	0	0
TOTAL	**295**	**307**	**324**	**313**	**827**	**152**	**446**

The isolates of *F. symbiotica* differed in their prophage composition. Isolate F1 contained 16 potential prophage regions, although only one of these was annotated as a complete prophage. Within F1 however, many of the prophages found in other isolates were present,

but these were annotated as questionable in their completeness (Table 4-5). The number of intact prophages in the other *F. symbiotica* genomes were higher. In isolates F2, F3 and F4 there were 19 (three intact), 21 (five intact) and 20 (five intact) identified regions respectively (Table 4-5). There were 20 regions (nine intact) found in Ap5D and 25 regions found in the *Cinara* strain but only five of these were intact. Isolate F5 contained 33 (six intact) prophage regions (Table 4-5). Of those six intact regions, however, two were not found in any other *F. symbiotica* isolate, one was unclassified and there were two copies of the *Erwinia*-like phage, one of which had a much lower %GC content than the other copy and those found in the other isolates. This suggests that four of these prophages may be related to the *Spiroplasma* genome rather than F5 itself (Table 4-5).

In addition to the number of complete prophages varying, the composition of these prophages also varied across isolates. The one intact prophage in isolate F1 was not present in any other isolate, and there were a number of phages found in Ap5D which were not present in any of the other isolates (Table 4-5). Some prophages were commonly found across almost all isolates, but, overall, there were no isolates that shared more prophages than others, all being equally dissimilar to each other. None of the *F. symbiotica* isolates carried the phage APSE, which is often found in *H. defensa* and is responsible for the parasitoid wasp resistance (Degnan & Moran 2008, Oliver et al. 2009, Weldon et al. 2012).

There are three plasmids contained within the genome of strain Ap5D; pFS5D.1, pFS5D.2 and pFS5D.3. By searching for plasmids within the remaining six genomes we found, as previously reported, the *Cinara* strain carried two of these plasmids, pFS5D.1 and pFS5D.3 (Patel et al. 2019) and we found that the pea aphid isolates F1-5 carried all three of the plasmids found in Ap5D.

Table 4-5: Table of prophage regions identified in the seven *F. symbiotica* isolates.
The table includes the total number of prophage regions, the number of regions that are intact, each intact regions most similar published phage and the number of gene hits for that phage, and any questionably complete phage that occurred as intact phage in other isolates. Light blue text signifies phage with low %GC content and those specific to F5, suggesting they may be part of the *Spiroplasma* genome.

Ap5D had the highest number of intact regions and many of these were not found in the other isolates. F1 contained the lowest number and this phage was not present in any other isolates. There were some phages that were commonly found across almost all isolates, and overall, there were no isolates that shared more phages than others

Isolate	Number of prophage regions	Number intact	Most common phage of intact regions (gene hit counts)	Questionable phage found as intact in other isolates
F1	16	1	1. Salmonella virus P22 - NC 002371 (2)	Enterobacteria phage Fels-2 (intact in F2,3,4,5, Ap5D, Cin) Salmonella phage RE-2010 (intact in F3, Cin) Erwinia phage ENT90 (intact in F3,4,5,Ap5D,Cin)
F2	19	3	1. Sodalis phage phiSG1 - NC007902 (4) 2. Escherichia phage HK639 - NC016158 (11) 3. Enterobacteria phage Fels-2 - NC010463 (9)	Salmonella phage RE-2010 (intact in F3, Cin) Erwinia phage ENT90 (intact in F3,4,5, Ap5D,Cin) Escherichia virus N15 (intact in F4) Salmonella phage 118970_sal3 (intact in F3,4,Ap5D)
F3	21	5	1. Salmonella phage RE-2010 - NC019488 (6) 2. Escherichia phage HK639 - NC016158 (10) 3. Enterobacteria phage Fels-2 - NC010463 (6) 4. Erwinia phage ENT90 - NC019932 (6) 5. Salmonella phage 118970_sal3 - NC031940 (2)	Enterobacteria phage mEp235 (intact in F4)
F4	20	5	1. Escherichia virus N15 - NC001901 (3) 2. Enterobacteria phage Fels-2 - NC010463 (15) 3. Erwinia phage ENT90 - NC019932 (6) 4. Enterobacteria phage mEp235 - NC019708 (3) 5. Salmonella phage 118970_sal3 - NC031940 (2)	Escherichia phage HK639 (intact in F2,3,Ap5D) Escherichia virus HK97 (intact in Ap5D)
F5	33	6	1. Gordonia phage Nyceirae - NC031004 (1) 2. Erwinia phage ENT90 - NC019932 (5) 3. Escherichia virus phiX174 - NC001422 (6) 4. No similar phage 5. Erwinia phage ENT90 - NC019932 (6) 6. Enterobacteria phage Fels-2 - NC010463 (5)	
Ap5D	20	9	1. Erwinia phage ENT90 - NC019932 (11) 2. Sodalis phage phiSG1 - NC007902 (2) 3. Pseudomonas phage PS-1 - NC029066 (5) 4. Salmonella phage 118970_sal3 - NC031940 (3) 5. Escherichia virus HK97 - NC002167 (2) 6. Enterobacteria phage HK620 - NC002730 (4) 7. Escherichia phage HK639 - NC016158 (8) 8. Salmonella phage SJ46 - NC031129 (3) 9. Enterobacteria phage Fels-2 - NC010463 (15)	
Cinara	25	5	1. Erwinia phage ENT90 - NC019932 (3) 2. Sodalis phage phiSG1 - NC007902 (3) 3. Salmonella phage RE-2010 - NC019488 4. Enterobacteria phage phi80 - NC021190 5. Erwinia phage ENT90 - NC019932 (7)	Enterobacteria phage mEp235 (intact in F4)

Virulence mechanisms

Flagellar assembly

F. symbiotica has retained a full flagellar assembly pathway, unlike *H. defensa* in which all the genes have been lost. Whether *F. symbiotica* expresses this flagellum and is motile within the aphid is still unknown. Previous work on *R. insecticola* has shown that it has also retained some flagellar assembly genes, however, has not been seen to be motile (Degnan et al. 2010). Additionally, *R. insecticola* encodes the gene lafA, which is required for a lateral flagellar, rather than the typical flagellum, fliC (Degnan et al. 2010). Like *R. insecticola* *F. symbiotica* also carries lafA. Many endosymbiont species engaging in long term or obligate associations with insects have lost many, or all, of the flagellar genes, as they are no longer necessary for their intracellular lifestyle. Flagella are also expensive to express (Macnab 1996), so the retention of a flagellum suggests that it plays an important role in *F.* symbioticas lifecycle, potentially playing a role in its transmission (Rio et al. 2012) or giving *F. symbiotica* a competitive advantage, as seen in *Photorhabdus*, where a functioning flagellum benefits this bacteria during insect infection (Easom & Clarke 2008).

Bacterial secretion systems

Secretion systems are an important component of bacteria, required for the transport of a range of substrates (Costa et al. 2015). All strains of *F. symbiotica* carry homologous genes which encode for type I-IV secretion systems, with the exception of *Cinara* which lacks one of the type IV systems which lies on the plasmid which is absent from this genome (Patel et al. 2019). All strains also carry a twin-arginine translation (Tat) pathway. The Tat pathway transports folded proteins, playing a role in a range of key processes (Palmer & Berks 2012), and is also important for virulence in almost all animal and plant pathogens (De Buck et al. 2008). This pathway is also found in the *R. insecticola* genome but is absent in *H. defensa*.

Type III secretion systems (T3SS) play an important role in the interaction between gram-negative bacteria and their host (Hueck 1998). They are required for the invasion of host cells, and can also interfere with the host innate immune system, allowing for pathogens to colonise the host cells (Egan et al. 2014, Hueck 1998, Matsumoto & Young 2009, Waterman

& Holden 2003). *F. symbiotica* carries two T3SS. These are similar to the SPI-1 and SPI-2 T3SS found in *Salmonella* species and are both found in the genomes of *H. defensa* (Degnan et al. 2009) and the *R. insecticola* strain 5.15 (Degnan et al. 2010). The *R. insecticola* strain LSR1, which lacks the ability to protect against the parasitoid wasp, *A. ervi*, only encodes an SPI-2-like T3SS (Hansen et al. 2012). In *H. defensa*, both T3SS are split in two (Degnan et al. 2009, 2010) and in *R. insecticola* the SPI-1 is also broken into fragments (Degnan et al. 2010). Similarly, SPI-1 in *F. symbiotica* is fragmented. The SPI-1 T3SS was originally associated with virulence in pathogens, but since has been found in a variety of bacterial species, including insect symbionts. It is thought that this T3SS may have more than one role, associating more with invasion and proliferation in these insect symbionts than virulence (Egan et al. 2014). The SPI-1-like T3SS that resides in the *Cinara* strain has diverged slightly from the other six isolates (Figure 4-4). In the genome of the *Cinara* strain, there is also no needle (PrgI) gene present for SPI-1. In the other isolates, this gene is present but lies on a different contig to other SPI-1 genes within the genome. This could suggest that there are differences between the isolates as to how they invade the host cells. Alternatively, it could suggest slightly different roles for these T3SS.

The SPI-2 T3SS in *F. symbiotica* appears to be mostly complete, with no fragmented sections like that seen in SPI-1. This T3SS however, is lacking many of the regulators and effectors usually found in this T3SS. The core export apparatus genes in this T3SS are conserved across all *F. symbiotica* isolates and are most closely related to the SPI-2-like T3SS in *H. defensa* and *R. insecticola* (Figure 4-4).

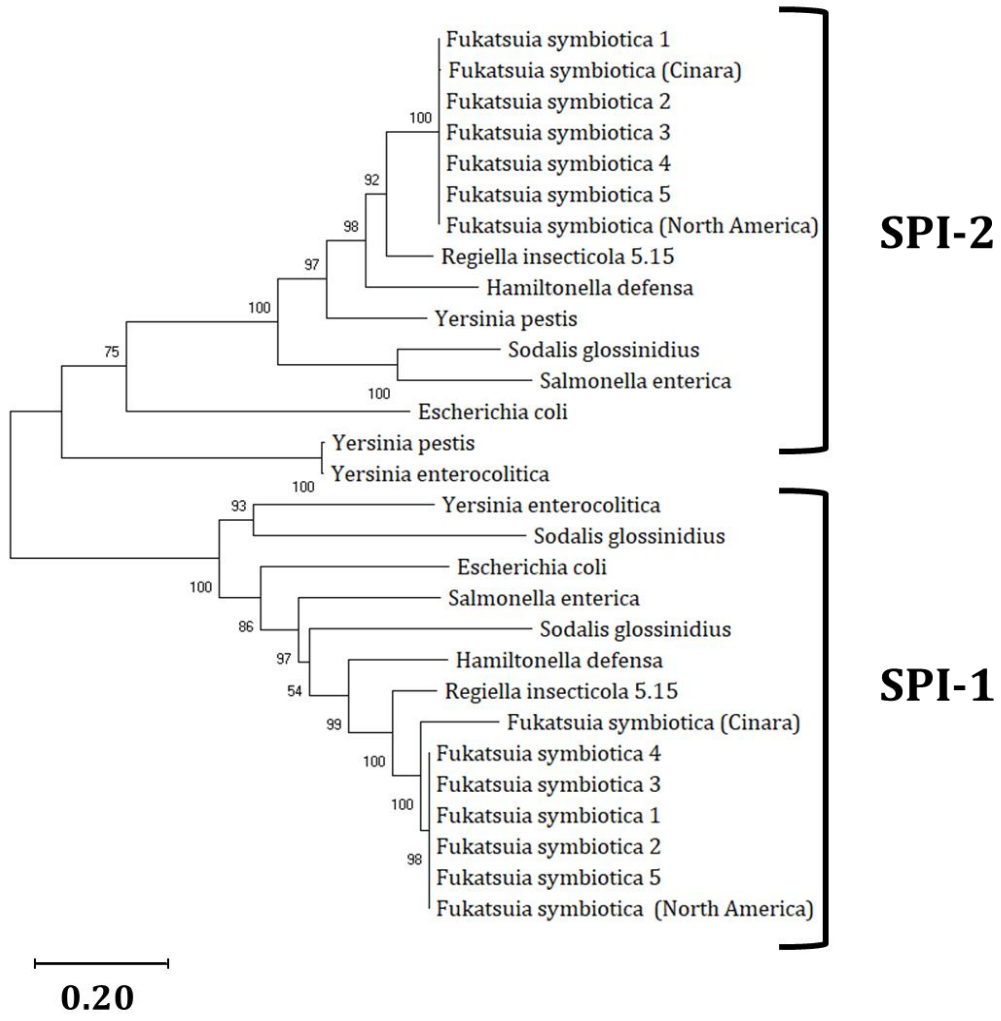


Figure 4-4: Phylogenetic reconstruction of the type III secretion systems of *F. symbiotica*, *H. defensa*, *R. insecticola* and other *Enterobacteriaceae* known to encode SPI-1- and SPI-2-like type III secretion systems.

The reconstruction was produced using concatenated sequences of the homologs of four of the export apparatus genes (SsaR/SpaP, SsaS/SpaQ, SsaT/SpaR, SsaU/SpaS) using MAFFT with 1000 bootstrap replicates.

The scale bar corresponds to 0.2 amino acid substitutions per site.

PhoPQ

F. symbiotica has retained PhoPQ, a two-component system which is often associated with virulence and anti-microbial peptide (AMP) resistance. The PhoPQ two-component system has been well studied in *Salmonella* (Dalebroux & Miller 2014, Dalebroux et al. 2014, Groisman 2001), where it plays several roles including sensing changes to the extracellular environment, specifically monitoring changes in Mg^{2+} levels, and regulating the expression of many genes, some of which are virulence genes (Groisman 2001, Groisman et al. 1997, Miller et al. 1989). PhoPQ is able to change lipopolysaccharides (LPS) which can lead to a reduction in the susceptibility to AMPs (Dalebroux & Miller 2014, Groisman 2001, Groisman et al. 1997, Gunn & Miller 1996). The PhoPQ system is crucial for pathogens, and some symbionts, to be able the evade the immune system of the animal host. Aphids, however, have a reduced innate immune system compared to that of other insect species, including a lack of common insect AMPs (Gerardo et al. 2010) therefore, it is surprising that *F. symbiotica* has retained these AMP resistance pathways, and PhoPQ may instead be playing a different role here.

PhoPQ, whilst being involved in AMP resistance, is also associated with virulence in many other bacterial species, including *Shigella* spp. (Lin et al. 2018), *Yersinia pestis* (Oyston et al. 2000), *Yersinia pseudotuberculosis* (Pisano et al. 2014), and *Psudeomonas aeruginosa* (Gooderham et al. 2009). In *Photorhabdus*, the PhoPQ pathway is essential for virulence against its insect host, with PhoP mutants losing their virulence (Derzelle et al. 2004). This two-component system has been lost in both *H. defensa* and *R. insecticola* strain LSR1, but remains in strain 5.15 (Hansen et al. 2012). These genes have previously been highlighted as potentially being involved in parasitoid resistance in *R. insecticola* (Hansen et al. 2012), and it may be that this two-component system also plays a role in virulence in *F. symbiotica*.

Toxin genes

The number of toxin-related genes within *F. symbiotica* differs across isolates, and also across species when compared to closely related aphid symbionts. There are large numbers of RTX (repeats in toxins) toxins, which are known to be toxic against eukaryotes (Coote 1992), found across the facultative symbionts of the pea aphid, including within the genome of *H. defensa* and *R. insecticola*. Previous investigations of RTX toxins within the Ap5D strain found a large number of RTX toxins in the *F. symbiotica* genome, which were also found in *H. defensa* and *R. insecticola* (Patel et al. 2019). We searched our draft genomes for the RTX toxins found in Ap5D and they were all present, including those that were absent in the *Cinara* isolate. However, whilst some of these were identical to those found in Ap5D, some showed reduced similarity (~93 %), so there was some slight difference in the RTX toxins across the isolates. This could suggest different roles for these RTX toxins and potentially differences in toxicity, which may be important if these genes are exhibited towards parasitoids. It may also suggest that these toxins have diverged within the different isolates due to different external or internal pressures, with their environment potentially having some influence on this.

One copy of a cytolethal distending toxin subunit B (cdtB) was also found within the genomes of all the isolates. cdtB is found in variants 2 and 8 of the phage APSE, which infects *H. defensa*, and is responsible for the parasitoid resistance conferred by *H. defensa* (Rouïl et al. 2020). APSE is absent from *F. symbiotica*, but they still carry a copy of this toxin gene. However, this particular cdtB is more similar to that from certain enteric pathogens, such as *Yersinia* spp., *Helicobacter* spp. and *Campylobacter* spp., suggesting that this likely did not originate from a transfer across the two facultative symbionts.

F. symbiotica also carries several other toxin genes annotated as Toxin A and Toxin B. These toxins are completely absent from the *H. defensa* genome. In the genome of *R. insecticola* strain 5.15, there are five Toxin A copies, although Toxin B is absent. One of the genes annotated as Toxin A in the genome of *R. insecticola* 5.15 was found to be the well-known

insecticidal gene, *makes caterpillars floppy (mcf)*, found in this strain. This strain of *R. insecticola* confers protection against the parasitoid wasp *A. ervi* and it is thought that *mcf* may play a role in providing this protection (Hansen et al. 2012). All but one of the Toxin B genes, and some of the Toxin A genes, found in the *F. symbiotica* genome, are orthologous to *mcf*. Furthermore, within a large orthologous cluster of genes containing this toxin, we identified a number of hypothetical proteins within the *F. symbiotica* genome (96 genes total; F1 = 13, F2 = 12, F3-5 = 13, F(*Cinara*) = 14, Ap5D = 19). Some of these genes are likely unannotated toxins that may have a similar function to *mcf*. Previous investigations of the Ap5D genome found nine *mcf* homologs (Patel et al. 2019), and it is likely some of those are included here.

Interestingly, there was a higher number of Toxin B genes present in Ap5D. This isolate is not known to convey any benefits to the aphid host (Doremus & Oliver 2017), however it is costly to the aphid, reducing its fecundity. The two extra Toxin B genes found in Ap5D have also diverged, not being phylogenetically similar to any of the Toxin B genes found in the other isolates (Figure 4-5). It could be that this toxin is an insecticidal toxin and is negatively impacting the host, in particular the development of their offspring, resulting in lower fecundity. Additionally, F2 has the fewest Toxin B genes and, of the five facultative mutualists, is the isolate that causes the smallest decrease in fecundity of the aphid (Smee et al. 2021). Furthermore, two of the Toxin B genes from isolate F1, which leads to the greatest decrease in fecundity, fall into a clade within which there are no Toxin B genes from isolate F2 (Figure 4-5).

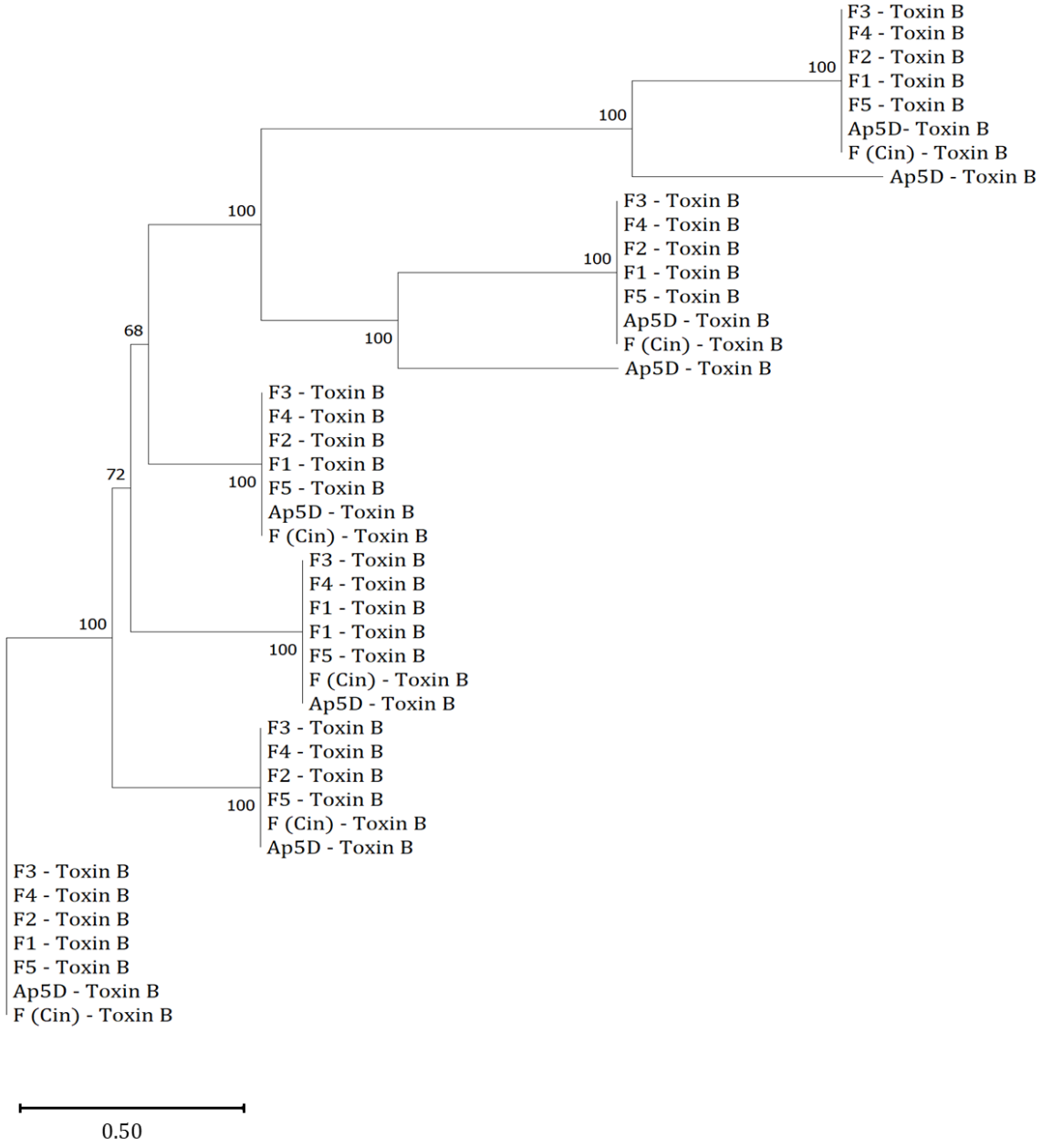


Figure 4-5: Phylogenetic reconstruction of Toxin B genes found in the seven isolates of *F. symbiotica*. F1-5 = facultative mutualists from UK pea aphids, Ap5D = North American strain, F (Cin) = strain found within *Cinara*.

The reconstruction was produced using MAFFT with 1000 bootstrap replicates.
The scale bar corresponds to 0.5 amino acid substitutions per site.

Another set of toxin genes found within the *F. symbiotica* genome include putative toxin complex genes. There were two copies of a gene which shared some similarity (~37 % identity) to toxins from *Photorhabdus* and *Xenorhabdus*, in particular, component A of the insecticidal toxin complex (Tc) – TcdA. The toxin complex is made up of three different proteins; A, B and C and the toxicity expressed by the Tc is dependent on which components are present. For example, it is known that singly, component A confers some level of toxicity, as well as the combination of components B and C, but for full toxicity to occur, all three parts are required (Waterfield et al. 2007). Tc genes have thus far been shown to primarily show oral toxicity to insects, although some strains of bacteria containing Tcs are still toxic on injection into the host (Rangel et al. 2016, Waterfield et al. 2007). The toxicity of this complex also has a level of specificity, as different strains that contain Tcs provide different levels of toxicity in different insects, and under different infection conditions (Rangel et al. 2016). So, the toxicity of this particular toxin complex relies, in part, on an interaction between the bacteria and the insect host.

Further to the similarity to the insecticidal toxin, these genes both returned a non-specific hit against the conserved domain within the A component of the Tc genes, VRP1. Additionally, a B component-like gene (47 % similarity to TcaB) and C component-like (55 % similarity TccC) are located adjacent to the putative TcdA gene. These two genes both returned hits against the conserved domains within the B (SpvB, MidN and MidC) and C (RhsA) components of the toxin complex mentioned above, suggesting a full toxin complex within the *F. symbiotica* genome. It may be that this toxin complex is having some effect on the aphid host, but if so, it does not appear to be expressing full toxicity and causing death, like that seen in other species. It could be that this toxin complex is having an effect on the development of the aphid offspring and could be resulting in the reduction in fecundity we see in aphids infected with *F. symbiotica*. Further exploration into these genes and the role they may be playing in this species would be beneficial.

Non-ribosomal peptides

Non-ribosomal peptide synthases (NRPSs) produce non-ribosomal peptide products, some of which have anti-fungal, -bacterial, and -viral properties, as well as other functions (Bloudoff & Schmeing 2017). The genomes of the seven *F. symbiotica* isolates encode a non-ribosomal peptide synthase (NRPS) cluster but these genes are absent from both the *H. defensa* and *R. insecticola* genomes.

The five UK pea aphid isolates of *F. symbiotica* are known to protect against a fungal pathogen, in the presence of *Spiroplasma* (Heyworth & Ferrari 2015), although the mechanism for this is currently unknown. Additionally, the isolates vary in their ability to protect against the fungal pathogen. Of these five isolates, F1 and F5 confer almost complete protection to the fungal pathogen whereas isolates F2, F3 and F4 confer only a moderate level of protection (Smee et al. 2021). Although all five mutualist pea aphid isolates carry the NRPS cluster, isolates F1 and F5 share the same NRPS genes and thus likely core chemical structure of the product (Figure 4-6). Isolates F2, F3 and F4 all have slightly different gene sequences and predicted chemical structure. The *Cinara* strain contains the NRPS cluster, however, it is unknown whether this strain provides any protection against the fungal pathogen, and the genetic structure of the NRPS in this isolate also differs from isolates F1 and F5. Ap5D also appears to carry this cluster, though it is known not to provide fungal pathogen resistance (Doremus & Oliver 2017). The structure of the NRPS also differs in this genome compared to the other isolates, and in particular from those that confer fungus resistance (F1, F5).

The NRPS cluster shows some similarity to Hassallidin E (18 % similarity), which is known to have anti-fungal properties (Pancrace et al. 2017) and Ralsolamycin (40 % similarity). Ralsolamycin is produced by an NRPS/PKS cluster within *Ralstonia solanacearum* which has close associations with plant-pathogenic fungus in the soil. This molecule appears to contribute to the bacteria's ability to invade the hyphae of the fungus and to also co-exist with plant-pathogenic fungus (Spraker et al. 2016).

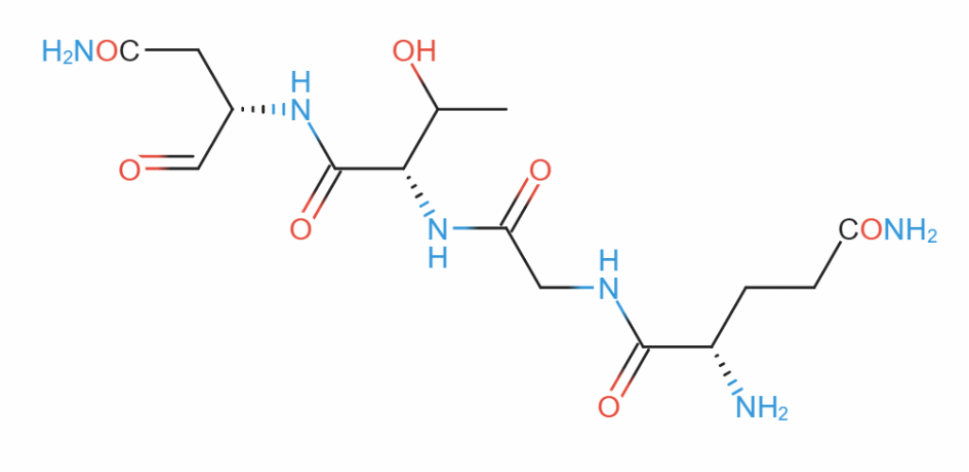


Figure 4-6: Predicted core chemical structure of the NRPS product from *F. symbiotica* isolates F1 and F5. Taken from antiSMASH v5.0 (Blin et al., 2017).

Given this cluster's similarity to that from bacteria which interact with fungus, and the genetic variation that appears to somewhat mirror the levels of resistance in isolates F1-5, there is a strong possibility this NRPS cluster may be playing a mechanistic role in the fungal pathogen resistance conferred by *F. symbiotica*. Further investigation would be needed to confirm this, however. This idea is, to some extent, conflicted by the fact it is found within the genome of Ap5D which provides no fungal pathogen resistance. However, the predicted chemical structure of the product differs, which may significantly alter the role of this particular NRPS, or how effective the products are against different pathogens. Interestingly, the cluster is absent from *R. insecticola,* which is also known to provide fungal pathogen resistance, although there may be more than one mechanism for fungal resistance, differing across symbiont species, as we see in parasitoid resistance.

Conclusions

The genomes of the five new *F. symbiotica* isolates presented here and the North American strain, Ap5D, are around 3.2 Mb in size, approximately 1 Mb larger than the genomes of the two closely related species *H. defensa and R. insecticola*. The more reduced genome size of *H. defensa* and *R. insecticola* potentially suggests a more stable or long-term association with the aphid hosts than *F. symbiotica*. The *Cinara* isolate, however, like *R. insecticola* and *H. defensa,* has a reduced genome size of around 2.8 Mb, although this is not as reduced as *H. defensa* and *R. insecticola*, and larger than we might expect given its co-obligate lifestyle.

F. symbiotica is most closely related to the facultative symbiont *R. insecticola* and *H. defensa* has undergone the most divergence from their free-living relatives. The genomes of the five new isolates of *F. symbiotica* all have high sequence similarity, and although the North American strain Ap5D showed a slight divergence, the core genes within all seven *F. symbiotica* isolates are highly conserved, although they appear to have different lifestyles.

Looking at the single copy orthologs that occur across just the seven *F. symbiotica* isolates shows F2-5 do not differ. However, F1 has diverged slightly from the other isolates from the same host species and location (F2-5). F1 is more costly to the aphid host and confers different protective phenotypes compared to isolates F2-4. The *Cinara* strain also forms a different clade to the five facultative mutualists and Ap5D has diverged the most from the other six isolates. This is interesting as, given it shares a host species with isolates F1-5, and these share some similarity in their phenotype (all confer a fecundity cost), we might expect that these isolates would be more similar to each other than the *Cinara* strain. This could suggest that isolates F1-5 have evolved to be more mutualistic, like the strain found within *Cinara,* whereas Ap5D has remained more pathogenic. However, instead, it may be that the North American strain has undergone more divergence than the other isolates. This could be due, in part, to the internal interactions between other symbionts within the North American pea aphids. For example, in North American pea aphids, *F. symbiotica* is

frequently found in a co-infection with *H. defensa*. It is thought that *F. symbiotica* exploits this mutualism by "hitchhiking" alongside *H. defensa* (Doremus & Oliver 2017), not providing any benefits to the host, which could, in turn, affect the evolution of this symbiont. It is unclear, however, whether there is an untested benefit which is conferred by this symbiont that we are yet to uncover. In addition, the UK isolates of this symbiont also often occur alongside *H. defensa*, and yet still provide some benefits to the host. In the genome evolution of different strains of *H. defensa,* strains that provide different levels of protection against a parasitoid have diverged into different clades (Chevignon et al. 2018; Chapter Five). A similar pattern could be occurring here due to different benefits conferred by the isolates of this symbiont and the external selection pressures that may drive these differences. It is unclear whether the isolate found within *Cinara* confers any defensive benefits or any costs to its host. Insight into this may further aid in understanding how strongly these genomic similarities are linked to the defensive host phenotypes that these different isolates confer.

The metabolic capabilities of *F. symbiotica* show some similarities to *R. insecticola* and *H. defensa*, however, *F. symbiotica* has retained some pathways that have been lost in others. *F. symbiotica* having retained the ability to synthesise more of the essential amino acids than *H. defensa* and *R. insecticola*, makes it is less reliant on the aphid host and *Buchnera*. *F. symbiotica* has also retained a full flagellar pathway, and whilst *R. insecticola* has retained some genes required for the flagellar apparatus (Degnan et al. 2010), these genes are absent in *H. defensa*. It is likely the flagellum plays an important role in *F. symbiotica*s, and potentially *R. insecticola*s, lifecycle.

F. symbiotica has retained virulence genes including toxin genes which are absent from other symbionts. The toxin genes found in *F. symbiotica* are similar to the toxin complex genes found in many insect pathogenic bacteria. Some of these genes found in the *F. symbiotica* genome may give some indication as to why carrying *F. symbiotica* is more costly to the aphid host. Furthermore, *F. symbiotica* has retained genes which have been seen to play an important role in sensing environmental changes, virulence and evasion of

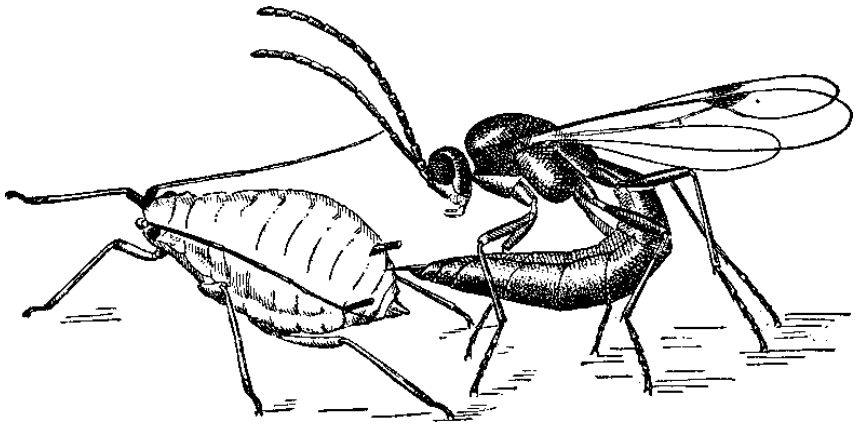
the immune system, such as the PhoPQ two-component system. Aphids have reduced immune responses and are missing AMPs found in other insect species (Gerardo et al. 2010), therefore it is more likely that these genes are playing a role in virulence than host evasion, in this species. *H. defensa* is lacking this system and is present only in *R. insecticola* strain 5.15, the strain which can protect against a parasitoid wasp (Hansen et al. 2012).

In addition to the genomic differences between *F. symbiotica* and the two closely related aphid facultative symbionts, there are also differences within the species. We found differences in their mobile DNA, specifically in the insertion sequences across the seven isolates. We found Ap5D had the highest number of ISs whereas the *Cinara* strain had the fewest, and isolates F1-4 sat in the middle of these two. This difference in mobile DNA was the largest indicator of lifestyle change across these genomes.

F. symbiotica carries an NRPS cluster which is absent from the genomes of *R. insecticola* and *H. defensa*. The isolates that confer almost complete protection against a fungal pathogen have the same genes within the cluster, and the same predicted chemical structure of the NRPS product. These also show similarities to clusters in other bacterial species that are known to interact with fungus. It may be that these NRPS genes may be playing a role in the fungal resistance conferred by *F. symbiotica*. However, this NRPS cluster is also found in Ap5D, which does not provide any resistance to a fungal pathogen, so this would benefit from further exploration.

As symbiont species transition from free-living species to mutualists and obligate species, there are genome reductions which often include a reduction in mobile DNA and the loss of pathogenic factors which are no longer necessary (Moran et al. 2008). Much of this genome reduction occurs due to genetic drift, due to the small population sizes and extreme bottlenecks that occur as they are vertically transmitted (Moran et al. 2008). *R. insecticola* and *H. defensa* appear to have evolved a close association with the aphid host, and as a result, there are similarities in their genome sizes and features. *F. symbiotica*s genome, however, suggests a more pathogenic or free-living lifestyle than the other aphid facultative symbionts, although they are all from the same host species. The genome size,

mobile DNA and pathogenic factors found in the genome of *F. symbiotica* suggests it has evolved to be more mutualistic than its free-living relatives, such as *Yersinia* spp., however, it has not evolved the same close association we see in *H. defensa* and *R. insecticola*. Although some pathogenic factors retained by this species could be related to the benefits it provides to its hosts, it may also be that these factors have been retained due to a less stable relationship with the aphid host. Within *Cinara*, we see *F. symbiotica* has a slightly smaller genome size, and differences in the amount mobile DNA, but otherwise this isolate shows no large difference in genome structure compared to the pea aphid isolates. This may suggest that the transition to the obligate lifestyle within *Cinara* may have been relatively recent. Within the pea aphid in some cases, *F. symbiotica* appears more like a pathogen than a symbiont, and this is somewhat reflected in the costly phenotype we see occurring with infection with this symbiont species.



Chapter Five

Comparative genomics of the facultative symbiont, *Hamiltonella defensa*

Abstract

Many insects carry bacterial facultative symbionts which often confer benefits to their host. *Hamiltonella defensa*, a common facultative symbiont of aphids, is well known for protecting its host against a parasitoid wasp. This protection is conferred by a bacteriophage, named APSE. The level of protection conferred by *H. defensa* varies across strains, due to different variants of APSE. Previous comparative studies of *H. defensa* genomes found different *H. defensa* strains fall into separate clades, aligning with the APSE variant they carry. They found no differences in the core processes across the strains, though there was variation across clades, especially in their mobile DNA composition.

Here we present the draft genomes of five isolates of *H. defensa* from pea aphids, which confer different host phenotypes, most notably different levels of protection against a parasitoid. We compare these genomes to each other and the previously published strains of *H. defensa*. The isolates appeared to split into two distinct groups, differing phylogenetically and falling into separate clades. In addition, these isolates carry different APSE variants, again falling into two distinct groups, with one group providing full protection from the parasitoid and the other providing moderate protection. Our results support previous findings with regards to similarities across strains. All the isolates shared the same genes in term of core processes, including metabolism and housekeeping functions. There were however large across clade differences, including their toxin gene composition and mobile genetic elements. We also found some within clade variation with one group varying more than the other.

Introduction

Many insects harbour bacterial facultative symbionts, which often benefit their host (Feldhaar 2011, Moran et al. 2008, Oliver et al. 2010). This involves protecting their host from a variety of natural enemies (Degnan et al. 2010, Heyworth & Ferrari 2015, Łukasik et al. 2013c, Oliver et al. 2003, Scarborough et al. 2005). The pea aphid is a model system for studying these facultative symbionts, and there are currently nine known species of facultative symbionts that infect aphids (Guo et al. 2017).

Hamiltonella defensa is a common aphid facultative symbiont, found in several aphid species (Henry et al. 2015, Zytynska & Weisser 2016), where it primarily resides extracellularly in the haemolymph of the insect (Moran et al. 2005, Tsuchida et al. 2005). *H. defensa* is well known for its ability to protect its aphid host against the parasitoid wasp *Aphidius ervi* (Oliver et al. 2003, 2005). Studies have shown that this protective phenotype is due to *H. defensa* being infected with a bacteriophage, APSE, with those lacking APSE conferring no parasitoid resistance (Oliver et al. 2009, Weldon et al. 2012). Strains of *H. defensa* vary in their ability to protect against the parasitoid wasp, due to different APSE variants providing different levels of protection (Oliver et al. 2009). The protection conferred by APSE also appears to be specific to *A. ervi*, as the same protective phenotypes are not conferred in response to other parasitoid species (McLean & Godfray 2015). *H. defensa* has also previously been found to protect against heat stress (Russell & Moran 2006). However, this effect is not universal, as others found no protection against heat stress, suggesting different strains confer different levels of protection (Heyworth et al. 2020). The mechanism for heat tolerance is currently unknown, but these effects differ across strains and it may be that host genotype and host-symbiont interactions also have an effect on this.

H. defensa also occurs in other insects, including psyllids and whiteflies (Sandstrom et al. 2001). It is well studied in two biotypes of the whitefly species *Bemisia tabaci,* however, there has been no evidence of *H. defensa* providing any parasitoid resistance in this species, like that seen in aphids (Rollat-Farnier et al. 2015). Within this species *H. defensa* is

limited to bacteriocytes, occurring with *B. tabaci*'s obligate nutritional symbiont, *Portiera aleyrodidarum*. It is thought that these *H. defensa* strains are acting as 'nutritional parasites' in this species, rather than acting as a beneficial microbe as they are in aphids (George & Toole 2017).

Previous genomic studies of *H. defensa* strains have shown it has a reduced genome compared to that of its free-living relatives, suggesting a long association with its insect hosts (Chevignon et al. 2018, Degnan et al. 2009). Those found within aphids, however, have also retained a number of pathogenic factors (Chevignon et al. 2018, Degnan et al. 2009). Although strains show differences in their ability to protect the host against a parasitoid, comparative studies of previously published *H. defensa* genomes have shown little genomic differences between strains (Chevignon et al. 2018) and the differences in protective phenotype are largely down to the APSE variant they are infected with (Chevignon et al. 2018, Oliver et al. 2009). The genomes of the strains studied thus far show no changes in core processes, such as metabolic processes and essential housekeeping processes, including those found in *B. tabaci* (Chevignon et al. 2018).

Here we present the draft genomes of five different *H. defensa* isolates from the pea aphid, *Acyrthosiphon pisum*. These five isolates confer different levels of parasitoid protection, with three providing complete protection against *A. ervi* and two providing moderate levels of protection (Smee et al. 2021). Some of the isolates also provide more tolerance to heat shock than others. In some cases, fecundity is only slightly reduced following heat shock compared to benign conditions, whereas others suffer a large fecundity cost (Smee et al. 2021). The survival of the aphid post-heat shock is largely unaffected by the presence of these five *H. defensa* isolates, although there was some across isolate variation (Smee et al. 2021). We aimed to investigate what similarities and differences these five isolates share with those already studied and how this may further inform our understanding of *H. defensa*'s evolution.

Methods

Isolate origins

The five isolates, H1-5, were collected from pea aphids from three different locations in the UK. Table 5-1 outlines the aphid clone, the symbionts present in each aphid and the location and date of collection. Isolates H1, H2 and H3 were in co-infections with isolates F1, F2, and F3 of *F. symbiotica*, analysed in Chapter Four. Isolates H4 and H5 were from single infections. All lines were collected from the same host plant, *Medicago sativa*.

Table 5-1: Details of the origin of the five *H. defensa* isolates H1-5.
Details include the aphid clone, symbionts present in each infection and the collection site and date. All lines were collected from *Medicago sativa*.

Isolate	Aphid clone	Symbionts	Collection site	Collection date
H1	218	*Hamiltonella & Fukatsuia*	Eling	May 2010
H2	238	*Hamiltonella & Fukatsuia*	Beaconsfield	May 2010
H3	236	*Hamiltonella & Fukatsuia*	Beaconsfield	May 2010
H4	207	*Hamiltonella*	Lincoln	August 2012
H5	216	*Hamiltonella*	Lincoln	August 2012

Genome sequencing, assembly and annotation

For the genome sequencing, DNA was first extracted from the haemolymph of aphids carrying the different isolates. These were then paired-end sequenced using an Illumina MiSeq. The library preparation and sequencing were carried out at the Technology Facility at the University of York for all five isolates. The number of reads for isolates H1-5 were 4030336, 5174154, 5579329, 5900612 and 5540336 respectively.

Sequences of the new *H. defensa* isolates were first trimmed using Trimmomatic (Bolger et al. 2014), using default parameters. Sequences were then mapped to the pea aphid (GCA_000142985.2) and *Buchnera aphidicola* str. APS (BA000003.2) genomes and any mapped sequences were removed. Genomes were assembled from the remaining sequence data, using the A5-miseq assembly tool (Coil et al. 2014) using default

parameters. PROKKA v1.12 (Seemann 2014) and RAST-Classic (Aziz et al. 2008) were used to annotate the genomes, again using default parameters.

Phylogenetic reconstructions

A phylogenetic reconstruction of the *H. defensa* strains and eight other closely related *Enterobacteriaceae* was created using 550 single copy orthologues across the nine species (NCBI references for species can be found in Appendix A). Genomes that had previously been published were downloaded and re-annotated using the same annotation pipeline to ensure consistency in annotation. The orthologs were identified using OrthoMCL (Chen et al. 2006) and sequences concatenated. The alignment and phylogenetic reconstruction were created using MAFFT (Katoh et al. 2002) with 1000 bootstrap replicates, and default parameters.

We also created phylogenetic reconstructions of the type III secretion systems, and the RTXA toxins found within the *H. defensa* genomes. The type III secretion system phylogeny was created using the concatenated sequences of homologs of four of the export apparatus genes (SsaR/SpaP, SsaS/SpaQ, SsaT/SpaR, SsaU/SpaS), across eight different bacterial species. The RTXA phylogeny was created using the annotated RTXA genes found in the genomes of the five isolates. For both reconstructions, the sequences were aligned, and the reconstruction created using MAFFT with 1000 bootstrap replicates, and default parameters.

Comparative genome analyses

The dnadiff comparison tool from MUMmer (Kurtz et al. 2004) was used to align the sequences of the five isolates and get an overall base similarity across the isolates, to see how much they differed at the nucleotide level. Orthologous gene groups were identified using OrthoMCL (Chen et al. 2006). Previously published genomes of seven *H. defensa* strains were included (5AT, A2C, AS3, ZA17, NY26, MI47, MEAM1) and were annotated

using the same pipeline to avoid gene annotation bias. KAAS (KEGG Automatic Annotation Server) (Moriya et al. 2007) was used to get a functional annotation of the genomes. This was carried out using the BLAST search program with a template data set of 35 bacterial species (Appendix B). The insertion sequences were predicted using ISSaga (Varani et al. 2011) and phage regions using PHASTER (Arndt et al. 2016). All analyses were carried out using default parameters unless specified.

Results and discussion

Draft genomes of the five isolates of *H. defensa* were assembled and annotated. The sizes of the genomes were similar to those of other *H. defensa* strains, with some slight variation (H1 = 2.3 Mb, H2 = 2.3 Mb, H3 = 2.3 Mb, H4 = 2.5 Mb, H5 = 2.4 Mb). The average %GC content was around 40 % and did not vary much across isolates (H1 = 40.5 %, H2 = 40.2 %, H3 = 40.2 %, H4 = 39.9 %, H5 = 40.1 %). The number of coding sequences (CDS) in the *H. defensa* isolates were 2181, 2162, 2154, 2300, 2292 respectively with around 35 % being annotated as hypothetical proteins. Similar to other closely related facultative symbionts, all five genomes contained 204 of the 205 single copy genes found in all *Gammaproteobacteria* (Lerat et al. 2003), suggesting no large gaps in sequencing.

The base to base comparison showed all five isolates differed slightly but appeared to fall into two distinct groups. Isolates H1, H2 and H3 were very similar to each other showing around 97 % similarity, which we would expect to see in isolates of the same species. However, these three isolates only shared around 89 % similarity to isolates H4 and H5. Isolates H4 and H5, although different from isolates H1-3, showed a high similarity to each other (96 %). There were 1695 single copy one-to-one orthologues (SICO) across all five of the isolates, and within the two groups, there were a further 373 SICO across isolates H1-3 and 469 between isolates H4 and H5 (Figure 5-1).

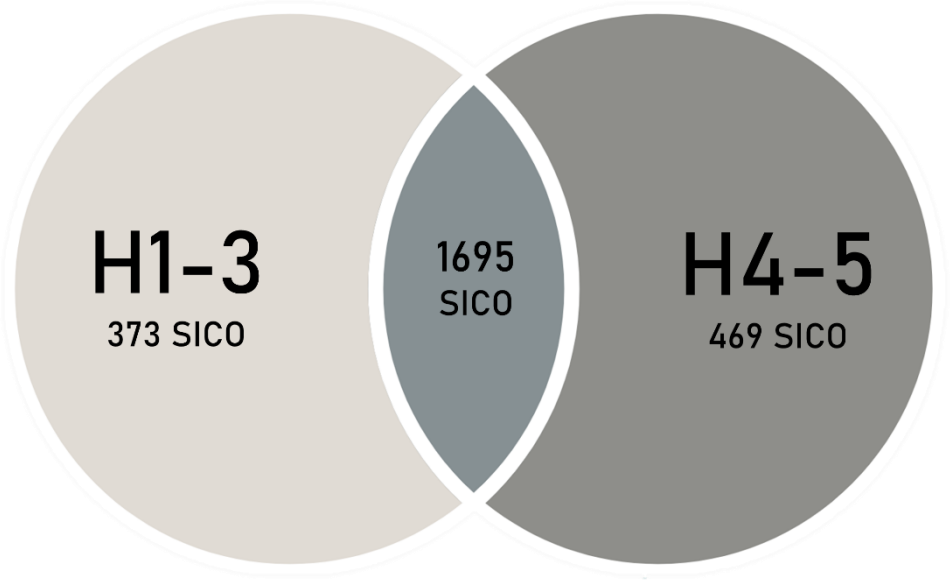


Figure 5-1: Venn diagram of the single copy orthologues (SICO) found across the five *H. defensa* isolates.
The five isolates were highly similar so had a large number of SICO across all the isolates. There was also a number of SICO shared within the two groups within which these isolates appeared to fall.

In terms of their metabolic capabilities, the isolates did not differ. They all shared the genes required for glycolysis as well as intact tricarboxylic acid (TCA) and pentose phosphate pathways. These isolates are only able to synthesise two essential amino acids, lysine and threonine, as are the other *H. defensa* strains, including those in whitefly (Chevignon et al. 2018, George & Toole 2017). This suggests that the *H. defensa* isolates that associate with an aphid host, have likely become reliant on *Buchnera*, like the aphid, to produce the other eight essential amino acids.

APSE variant

We searched the five isolates for the phage APSE, which is responsible for the parasitoid resistance against *A. ervi* conferred by this symbiont (Oliver et al. 2003, 2009; Weldon et al. 2012). We found that isolates H1-3 carry the same APSE variant, APSE-1, which confers high levels of protection against *A. ervi*. Isolates H4 and H5 carry APSE-2, which confers moderate levels of protection to the host (Smee et al. 2021).

Phylogenetic reconstruction

A phylogeny was created to understand how related these isolates were to the previously published *H. defensa* strains, and to other closely related symbiont species (Figure 5-2). The phylogeny showed the sister species of *H. defensa* are *Regiella insecticola* and *Fukatsuia symbiotica*. Within *H. defensa*, the whitefly and aphid strains form two distinct clades. Previous studies found the aphid strains of *H. defensa* divided into three clades (Chevignon et al. 2018), but here we found they split into four clades with isolates H1, H2 and H3, presented here, forming a fourth clade with the previously published strain MI47 which did not appear to be included in the previous study. MI47 has also previously been seen to share a clade with strain 5D (Rouïl et al. 2020) which was not included in this analysis. The branch lengths of the clades suggest a recent divergence. As seen in the base comparison of these isolates, the phylogeny also shows the five isolates separating into two distinct groups. Whilst isolates H1, H2 and H3 form the fourth clade with MI47, isolates H4 and H5 fall into a clade with the previously published strain, ZA17 (Figure 5-2). The previously known clades were thought to represent the APSE variant carried by the strains (Chevignon et al. 2018). Whilst isolate H1-3 follow this pattern, forming a clade with another isolate that also carries the APSE-1 variant, isolates H4 and H5 share the same clade as ZA17 which carries APSE-8, rather than with 5AT and NY26 which also carry APSE-2 (Figure 5-2). However, a recent analysis of the APSE variants in previously published strains of *H. defensa* instead considered ZA17 to carry an APSE-2 subtype, as the only difference was a pseudogenisation of one of the toxin genes (Rouïl et al. 2020). This study also highlighted that the APSE-2 variant carried by 5AT and NY26, namely APSE2.2a, differs from that carried

by ZA17, APSE2.2b. The cdtB toxin carried by these different APSE has diverged, as seen in the phylogeny of the *H. defensa* strains (Rouïl et al. 2020). The fact that H4 and H5 share a clade with ZA17 rather than 5AT and NY26 suggests that H4 and H5 may carry an APSE-2 variant which is more similar to APSE2.2b found in ZA17. Interestingly however this co-diversification between APSE and *H. defensa* is not always the case. Strain 5D, which has been seen to share a clade with MI47, and thus likely our isolates H1-3, carries an APSE2.2b variant almost identical to that found in strain ZA17 (Rouïl et al. 2020). This shows that although these divergences are usually in line with the APSE variant they carry, there may also be horizontal transmission of APSE across different *H. defensa* clades.

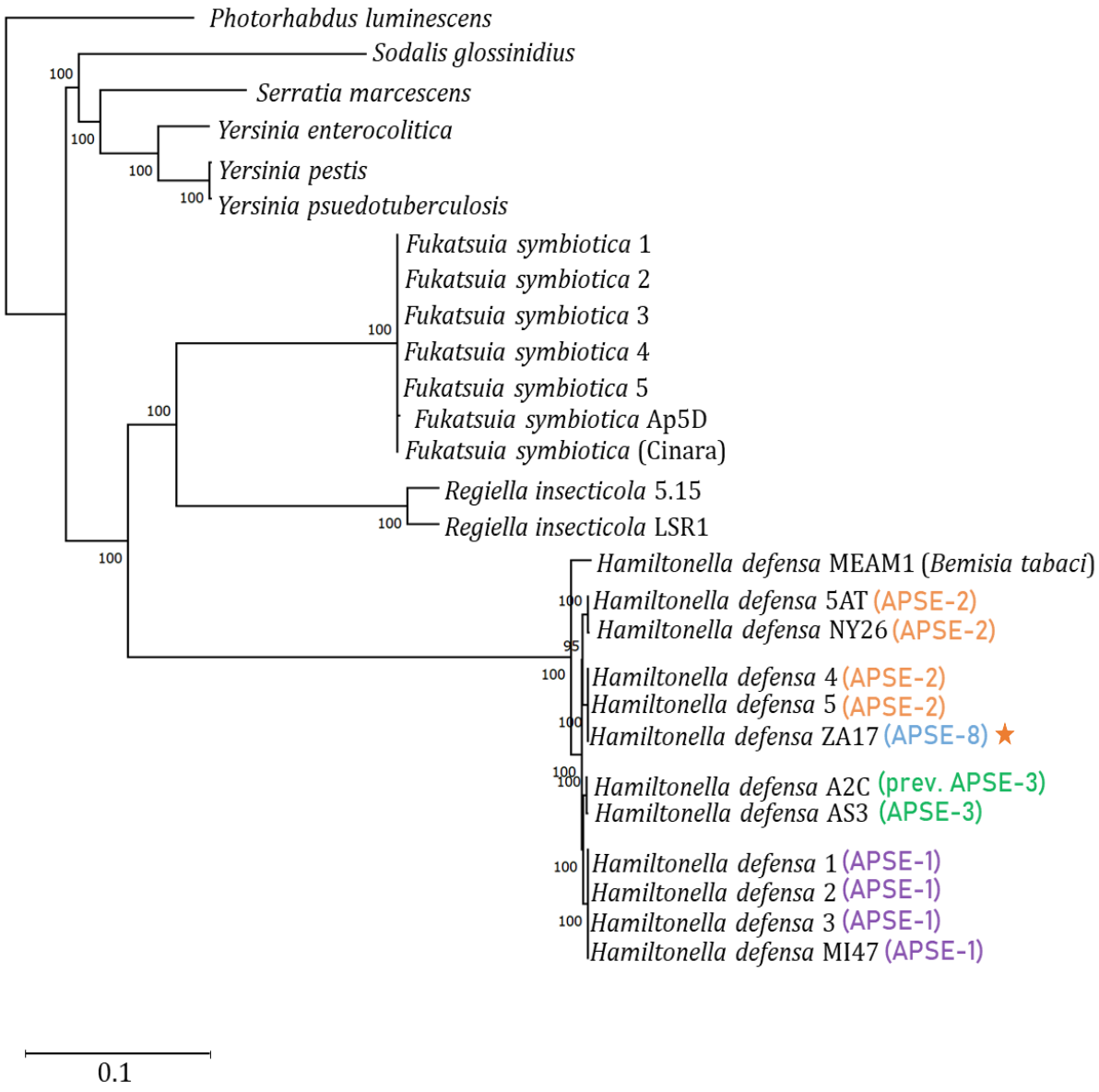


Figure 5-2: Phylogenetic reconstruction of 12 *H. defensa* strains and other closely related *Enterobacteriaceae*, including the five new isolates (H1-5).
The reconstruction was produced using concatenated sequences of the single copy orthologous genes across the nine species (550 genes) using MAFFT with 1000 bootstrap replicates.
Scale bar represents 0.1 amino acid substitutions per site.
Isolates H1-5 fall into two distinct clades, with H1-3 forming a clade with the previously published strain MI47 and H4 and H5 forming a clade with strain ZA17.
Also shown is the APSE variant associated with each H. defensa. ZA17 is marked as it has been previously reported to be carrying both APSE-8 and APSE-2 (Rouïl et al. 2020).

Virulence mechanisms

Whilst many features of *H. defensa*'s genome suggest a long-term stable relationship with its host, it also retains a number of pathogenic factors, and likely evolved from pathogenic origins (Degnan et al. 2009). These include a variety of virulence mechanisms, including bacterial secretion systems and toxin genes, some of which may be effective either against the host or against other invading organisms.

Bacterial secretion systems

Secretion systems are required by bacteria for the transportation of various substrates, many of which are important for a range of key processes (Costa et al. 2015). In *H. defensa* these secretion systems are thought to play an important role in host cell invasion (Degnan et al. 2009). All isolates carry genes that encode type I and type II secretion systems. They also carry two type III secretion systems (T3SS), similar to SPI-1 and SPI-2 found in *Salmonella*. However, within the genomes, these secretion systems are fragmented. The T3SS also appear to be highly conserved across all four of the aphid *H. defensa* clades (Figure 5-3). There is a slight divergence in isolate H4, H5 and strain ZA17 in the SPI-2-like T3SS, however, the difference is negligible, suggesting a very recent divergence. All isolates also encode an I-type type IV secretion system, which is not found in *H. defensa* strains that infect whitefly (Chevignon et al. 2018). There were also strain-specific systems which differed across the isolates. H4 and H5 encoded an F-type type IV secretion system, which was absent from isolates H1-3. This secretion system is also found in strains A2C, AS3 and ZA17 (Chevignon et al. 2018). It is unclear why this secretion system is only present in these strains, but this may impact how these particular strains invade the host cells, or affect their response to other invading organisms, such as parasitoids.

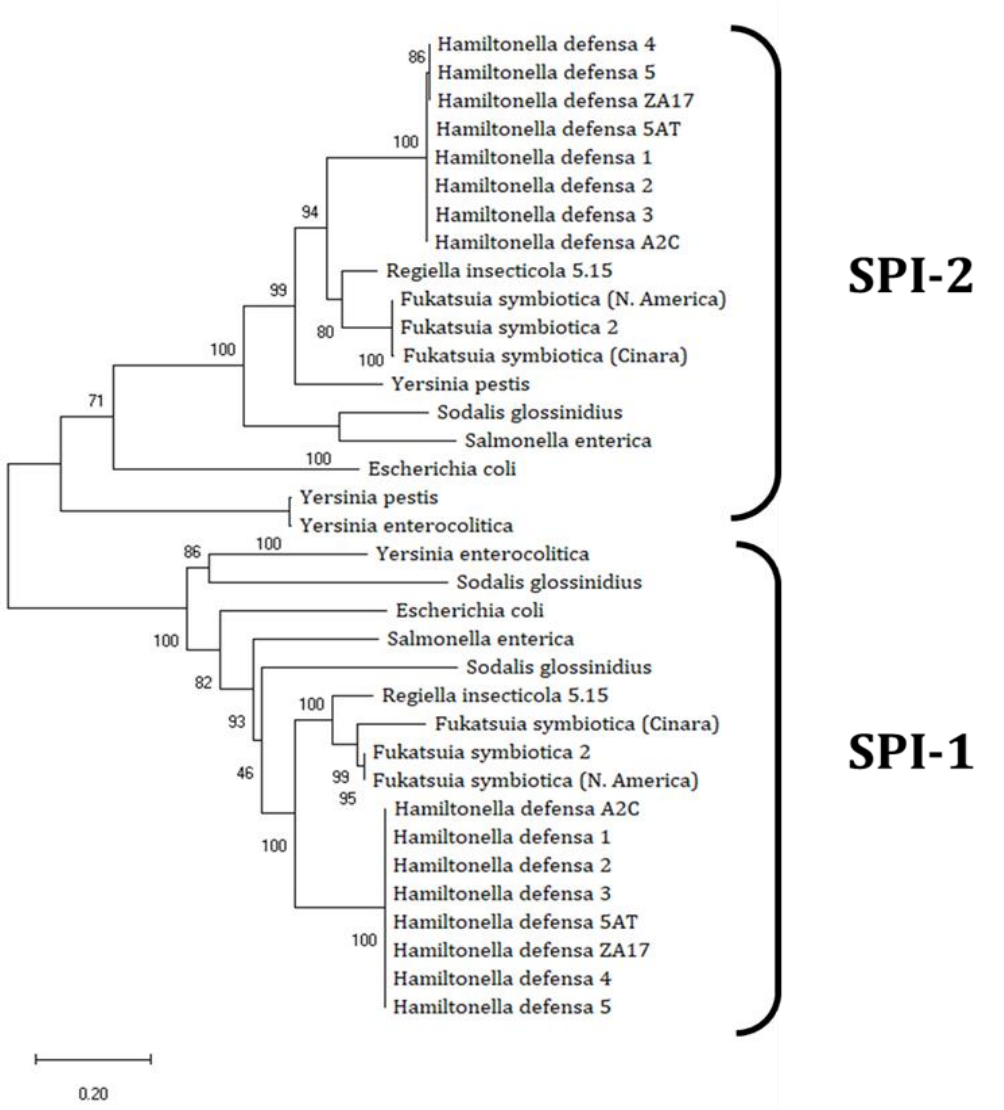


Figure 5-3: Phylogenetic reconstruction of the SPI-1- and SPI-2-like type III secretion systems of *H. defensa* and closely related *Enterobacteriaceae*.
Isolates H1-5 and one strain from each other aphid related *H. defensa* clade was included. The reconstruction was produced using concatenated sequences of four of the T3SS export apparatus genes (SsaR/SpaP, SsaS/SpaQ, SsaT/SpaR, SsaU/SpaS) using MAFFT with 1000 bootstrap replicates.
Scale bar represents 0.2 amino acid substitutions per site.

Toxins

We investigated the number of toxin genes found within the genome of the five isolates of *H. defensa*, specifically looking at the number of RTX (repeats in toxin) toxins present. RTX toxins are a superfamily of toxins known to be toxic against eukaryotes (Coote 1992) and there are large numbers of these found in other *H. defensa* strains (Chevignon et al. 2018, Degnan et al. 2009). We also found a large number of RTX toxins present in the genome of isolates H1-5, which included a range of adenylate cyclase (AC) toxins and leukotoxins. The number of RTX toxins was the same across all isolates, with all the genomes containing 17 different RTX toxins, although the numbers of each type varied (Table 5-2). Across all five isolates, RTXA was the most common type of RTX toxin, although they did not all have the same number of RTXA genes.

The numbers of each type of toxin carried by the different isolates again fell into two groups. In H1-3 there were nine RTXA toxins, whereas in H4-5 there were only seven (Table 5-2). However, in H4 and H5 there was a larger number of the AC toxin, cyaA. We further explored the similarity of the most common type of RTX toxin, RTXA, across the five isolates, to see how much these differed across the isolates. Here we found that, like the core genes in these five isolates, the RTXA toxins had too diverged into two groups (Figure 5-4).

Table 5-2: Number and type of each RTX toxin found in isolates H1-5.
All isolates contain the same number of RTX toxins but have different numbers of each type. In all isolates, the majority of the RTX toxins were RTXA toxins.

RTX toxins	H1	H2	H3	H4	H5
RTXA	9	9	9	7	7
AC toxin (cyaA)	3	3	3	5	5
Leukotoxin (lktA)	3	3	3	3	4
Leukotoxin (ltxA)	2	2	2	2	1
TOTAL	**17**	**17**	**17**	**17**	**17**

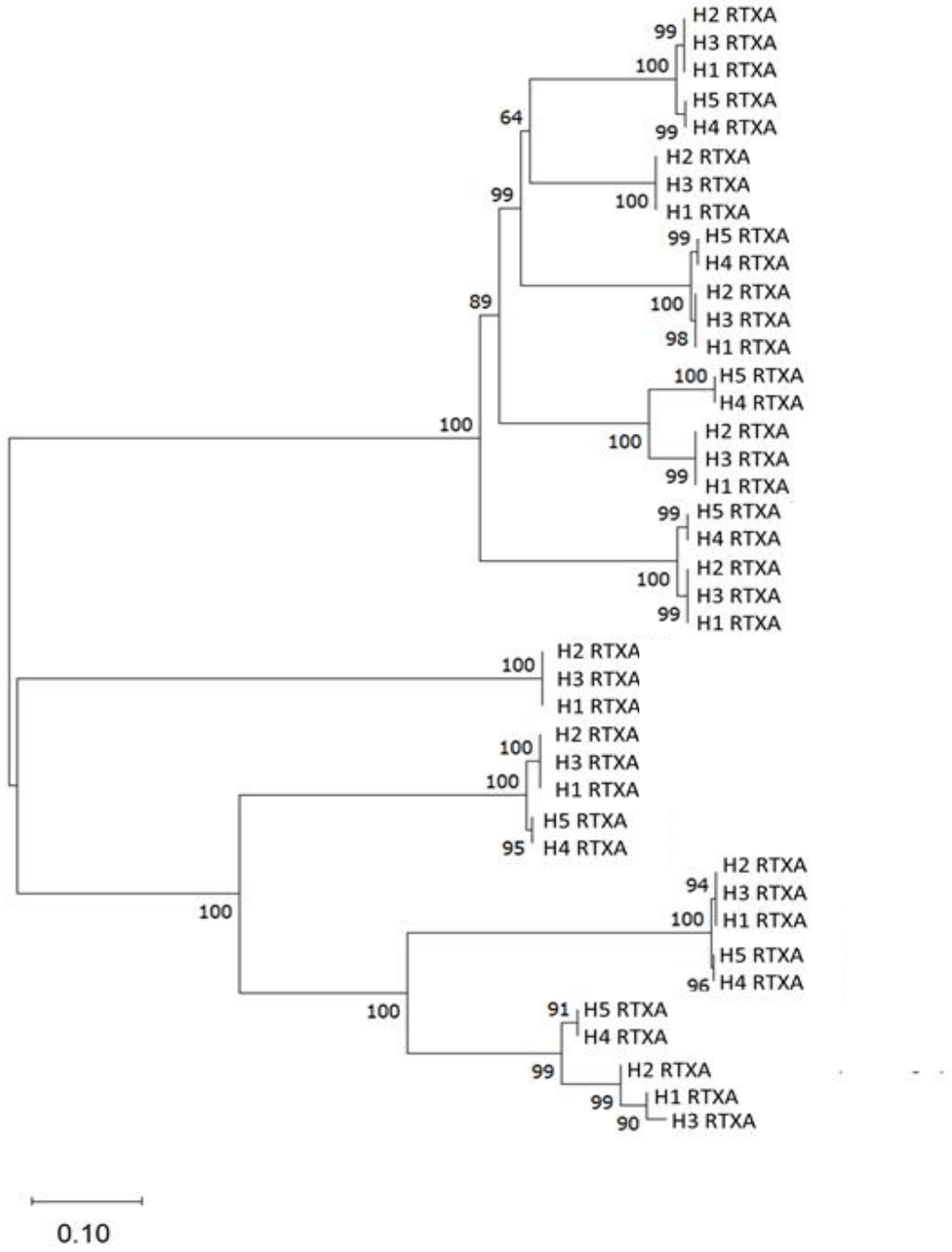


Figure 5-4: Phylogenetic reconstruction the RTXA toxins of *H. defensa* isolates H1-5.
The sequences were aligned, and the reconstruction was produced using MAFFT with 1000 bootstrap replicates.
Scale bar represents 0.1 amino acid substitutions per site.
The RTXA toxins from the different isolates fall into two groups with H1-3 forming a distinct group separate from H4-5. This is similar to the grouping and divergence seen in the core genes of these five isolates.

Mobile DNA

Investigations of previously published strains of *H. defensa* found that the greatest amount of variation within the strains came from their mobile genetic elements (MGEs) rather than those involved with the core processes (Chevignon et al. 2018). We investigated the MGE in the genomes of isolates H1-5, including insertion sequences (ISs), plasmids and phages.

We found a total of 12 different IS families across all five *H. defensa* genomes, although these differed slightly between isolates in both the families that were present and the abundance of each family (Table 5-3). The IS composition of the isolates again fell into two distinct groups. H1-3 shared the same 11 families and there was little difference in the number of open reading frames (ORFS) associated with these families. The most abundant type was IS630, which similarly is the most common IS family found in the genome of the closely related *F. symbiotica* (Chapter Four). Overall, in H1-3 there were 100-102 ISs. Within H4 and H5 however, there were fewer ISs, 79 in H4 and 77 in H5, but again, these were similar to each other in number and composition (Table 5-3). These were also made up of 11 IS families, but these isolates were missing ISKra4, found in H1-3, and instead carried a small number of IS200/IS605. H4 and H5 also differed to H1-3 in their abundance of each family and in H4 and H5, the most abundant IS family was IS3, rather than IS630. H4 and H5 also only carried one ORF associated with IS6, but there were much more of this family found within isolates H1-3 (Table 5-3). ZA17, which shares a clade with H4 and H5, has previously been reported have 148 ISs (Chevignon et al. 2018), which is much higher than the number found in H4 and 5, but these too are mostly from the most abundant family, IS3 (28%) (Chevignon et al. 2018). This suggests that the IS composition of isolates is similar between isolates from the same clade, but overall, the number may vary more than previously reported.

Table 5-3: Number of each insertion sequence (IS) family contained in the five *H. defensa* isolates.
Darker blue shows a higher abundance of the IS family in that isolate.
Isolate H1-3 and H4-5 fell into two groups with regards to there IS similarities. H1-3 had a higher number of ISs than H4-5, with IS630 being the most abundant. H4-5 carried less ISs and had a higher abundance of IS3 than IS630.

IS families	H1	H2	H3	H4	H5
IS630	28	30	29	14	12
IS3	19	16	19	23	25
ISNCY	13	12	12	14	14
IS5	12	15	14	12	10
IS6	9	9	9	1	1
IS110	8	9	7	4	4
IS701	4	4	4	1	1
IS982	3	3	3	3	3
IS4	2	2	1	4	4
IS481	1	1	1	1	1
ISKra4	1	1	1	0	0
IS200/IS605	0	0	0	2	2
TOTAL	**100**	**102**	**100**	**79**	**77**

In terms of their prophage composition, whilst having some differences, the five isolates were all fairly similar, with each isolate containing between three and five intact prophages (H1: 9 regions, 4 intact; H2: 12 regions, 5 intact; H3: 7 regions, 3 intact; H4: 8 regions, 3 intact; H5: 10 regions, 5 intact). All isolates contained ASPE, as discussed earlier. In addition to APSE, there was another prophage that was shared across all isolates, and one that occurred in all but isolate H3. Overall, the prophages that occurred in the isolates were similar, with only two types that were not shared across at least two of the five isolates. There also did not appear to be any distinct split between clades in their prophage composition as in the other genome features. The phage found in strain ZA17 and MI17, which the five new isolates share a clade with, were also similar to H1-5 in number and composition.

We investigated the plasmid content of the isolates and how similar they were to the plasmids found in the other *H. defensa* strains. The five isolates differed in their plasmid content and fell into the two groups described earlier, although there were also more differences within these groups than expected. H1 contained four plasmids, which were identical to four of the five plasmids found in the previously published strain 5D (p5D.1, p5D.2, p5D.4, p5D.5; Table 5-4). In H2 we found two of these same plasmids (p5D.4 and p5D.5) and, although p5D.1 and p5D.2 also appeared to be present in the genome, these were only partially complete in this strain (85%). Finally, within H3, p5D.2 was found although, again, this was not complete (85%), and no other plasmids were found within the genome. There were no other plasmids that showed similarity to those from other previously published *H. defensa* strains found in H1-3.

Isolates H4 and H5 differed greatly to H1-3 in their plasmid composition but there was less variation in this group (Table 5-4). As expected, given they share a clade with strain ZA17, two of the plasmids found in this strain were also present in isolates H4 and H5 (pZA17.1 and pZA17.3). pZA17.2 also appeared to be present in the genomes, but this was not complete in either H4 (78%) or H5 (72%). Again, no other plasmids that showed similarity to those in previously published *H. defensa* strains were found here. Although not included in our core gene phylogeny, strain 5D has previously been shown to share a clade with MI47, which here shares a clade with isolates H1-3, so it is not unexpected that these plasmids are found in isolates H1-3 and those from ZA17 found in H4 and H5. This again highlights that many of the *H. defensa* genome similarities lie within the different clades of this species.

Table 5-4: Summary of plasmids found in H1-5.
The table shows the name of the plasmids present in each isolate and where a complete plasmid was not found, the percentage found is shown in brackets.

H1-3 carry a similar set of plasmids, with H1 and 2 having the highest number overall. H4 and H5 carry plasmids that are different from those found in H1-3 but the same as strain ZA17, with which they share a clade (Figure 5-2). There is also a higher similarity between the plasmids found in the genomes of H4 and H5 than there is within the H1-3 group.

H1	H2	H3	H4	H5
p5D.1	p5D.1 (85%)		pZA17.1	pZA17.1
p5D.2	p5D.2 (85%)	p5D.2 (85%)	pZA17.2 (78%)	pZA17.2 (72%)
p5D.4	p5D.4		pZA17.3	pZA17.3
p5D.5	p5D.5			

Conclusions

We built and analysed five genomes of the aphid facultative symbiont *H. defensa*, comparing them to previously published strains of this species. The genomes were between 2.3 – 2.5 Mb in size, similar to that of other strains found in aphid hosts (Chevignon et al. 2018, Degnan et al. 2009). Comparing across the five isolates, we found that they fell into two distinct groups. Isolates H1, H2 and H3 were more similar to each other than to H4 and H5, and H4 and H5 also shared higher similarity to each other than to isolates H1-3. This was seen in the phylogeny of the core genes within these genomes, with the two groups falling into different clades. Previous strains had fallen into three clades (Chevignon et al. 2018) and isolates H4 and H5 fell into one of these, with the previously published strain ZA17. Isolates H1-3, however, formed a separate fourth clade, with the previously published strain MI47. The isolates also carried different APSE variants, therefore providing different levels of parasitoid resistance. Again, these were split into two groups, isolates H1-3 carrying APSE-1 and isolates H4 and H5 carrying APSE-2.

As seen in previous studies, the isolates did not differ from each other, or previously published strains, with regards to their core processes. The metabolic capabilities of the five isolates were the same, retaining the ability to synthesise only two of the ten essential amino acids. In terms of their virulence factors, there were some similarities but also differences between the isolates. The type I-III secretion systems of these five isolates were the same but isolates H4 and H5 also carried an F-type type IV secretion system, found in strains AS3, A2C and ZA17, which was missing from isolates H1-3. The isolates also differed in their toxin composition. These differences were again split by the two clades these isolates fell into. Overall, all isolates carried the same number of RTX toxins but isolates H1-3 contained slightly more RTXA toxins than isolates H4 and H5, with H4 and H5 having slightly more AC toxins. Furthermore, the RTXA toxins found across the five isolates appear to have diverged into two clades, as seen in the core gene phylogeny. This suggests these toxins have changed following the divergence of these isolates into different clades.

Previous comparative studies of *H. defensa* found large differences in the strains mobile genetic elements (MGEs) (Chevignon et al. 2018). These five isolates have some differences in their MGEs, however, again these appear to be mainly across clade differences. Isolates H1-3 had a higher number of ISs found in their genomes compared to H4-5. These also differed in the abundance of each IS family within the genome, with IS630 being the most abundant in H1-3 and IS3 in H4 and H5. They also differed in their plasmid content, with H1 carrying four complete plasmids, similar to those found in strain 5D. There was more difference than expected across isolates H1-3 as, although they all showed some similarity in plasmid content, H2 and H3 appeared to be lacking some of the complete plasmids found in H1. Isolates H4 and H5 were again very different to H1-3, carrying plasmids found in the genome of strain ZA17, which is not unexpected given the shared clade. H4 and H5 had more similarity to one another in their plasmid content, than within the H1-3 group.

H. defensa is known for its parasitoid resistance, provided by the phage APSE. There is variation between the level of this resistance to parasitoids conferred by these five isolates, which is due to the strains carrying different APSE variants. There are currently seven known

APSE variants, six from aphid species, and one found in *H. defensa* from whitefly (Rouïl et al. 2020). APSE-7, found in whitefly-associated *H. defensa*, has diverged from those found in aphid species, as seen in the *H. defensa* genome. The difference in APSE variant in these five isolates falls into clades, with strains carrying the same APSE variant sharing a clade. However, there are exceptions to this, so there is likely horizontal transmission occurring across these strains (Rouïl et al. 2020). Outside of this benefit, in some cases, *H. defensa* has also been seen to decrease the lifespan and lifetime reproduction of aphids (Cayetano et al. 2015, Vorburger & Gouskov 2011), and as such symbiont presence could be selected against in the absence of the parasitoids. Interestingly, however, the most protective strains also appear to be less costly (Cayetano et al. 2015). Given that these strains have diverged into clades with particular APSE variants, it is likely ecological influences are playing a role in this symbiont's evolution, potentially having different effects on strains that confer different levels of protection, as well as their relationship with the host.

In this study, we found genomic variation across both the newly sequenced genomes of pea aphid *H. defensa*, and across the previously published strains. The five isolates split into two clades associated with the APSE variant they carry, and thus parasitoid protection conferred. Although the difference in APSE is one of the largest and most ecologically relevant differences across these isolates, there is also a large amount of variation between these two groups outside of this. The variation across isolates laid outside of the core processes, and was found in the virulence factors, such as toxin genes, and mobile DNA composition. As in previous studies, we found that there is a high amount of similarity between those in the same clade (Chevignon et al. 2018), however, there were some within clade differences, and isolates H1-3 appeared to differ more than H4-5 did.

H. defensa occurs across a wide range of host species, including different aphid species, as well as whitefly. There is evidence for the horizontal transfer of these strains across species, with near identical *H. defensa* strains being found in different host species (Łukasik et al. 2015). However, studies have shown that the genetic distance of symbiont strains can affect the ability of these strains to form stable relationships with novel hosts. For example, strains

that are genetically distant from those that previously infected a host, are not able to form stable long-lasting relationships with that host (Łukasik et al. 2015). This shows there are limitations to the horizontal transfer of these strains across host species, which may be due to host-symbiont relationships and coevolution, limiting novel strains from establishing. The divergence we see occurring in these isolates may suggest less horizontal transmission as strains become more genetically distinct, evolving towards a strictly vertical transmission.

This chapter further highlights the range of strains within this symbiont species and their different evolutionary trajectories. It will be interesting to see how the evolutionary and ecological pressures on these various strains of *H. defensa* will lead their genomes to continue to evolve, potentially diverging further from one another, and the future implications of this.

Chapter Six
General discussion

Overview

This thesis aims to further our current understanding of the complex interactions between facultative symbionts and their hosts, particularly the role this plays in defence against natural enemies. The pea aphid is an interesting system within which to explore these interactions, given they harbour several different facultative symbionts and have a reduced immune system compared to other insect species.

I first explored whether facultative symbionts can influence changes in susceptibility to a pathogen across two generations. I found that harbouring symbionts leads to transgenerational effects, which do not occur in aphids without symbionts. This resulted in offspring having reduced susceptibility to a pathogen compared to their mothers, likely due to an interaction between the symbiont and the host immune system. I then went on to explore the interaction between facultative symbionts and the host immune system further, investigating whether the presence of a symbiont alters the host immune response and if this influences how the host responds to an invading pathogen. Here I found further evidence of the complex interaction between the immune system of the pea aphid and its facultative symbionts, showing facultative symbionts do alter the immune response of the host, although this does not necessarily affect their ability to protect against a pathogen. Lastly, I investigated the genomes of the two facultative symbionts used throughout this thesis, *Fukatsuia symbiotica* and *Hamiltonella defensa*, where I detailed the genomic differences across a variety of isolates of these two species, and highlighted potential mechanisms for fungal pathogen resistance and the fecundity cost exhibited by *F. symbiotica*. In this final chapter, I bring together the findings from the previous chapters outlining general conclusions, applications and future directions.

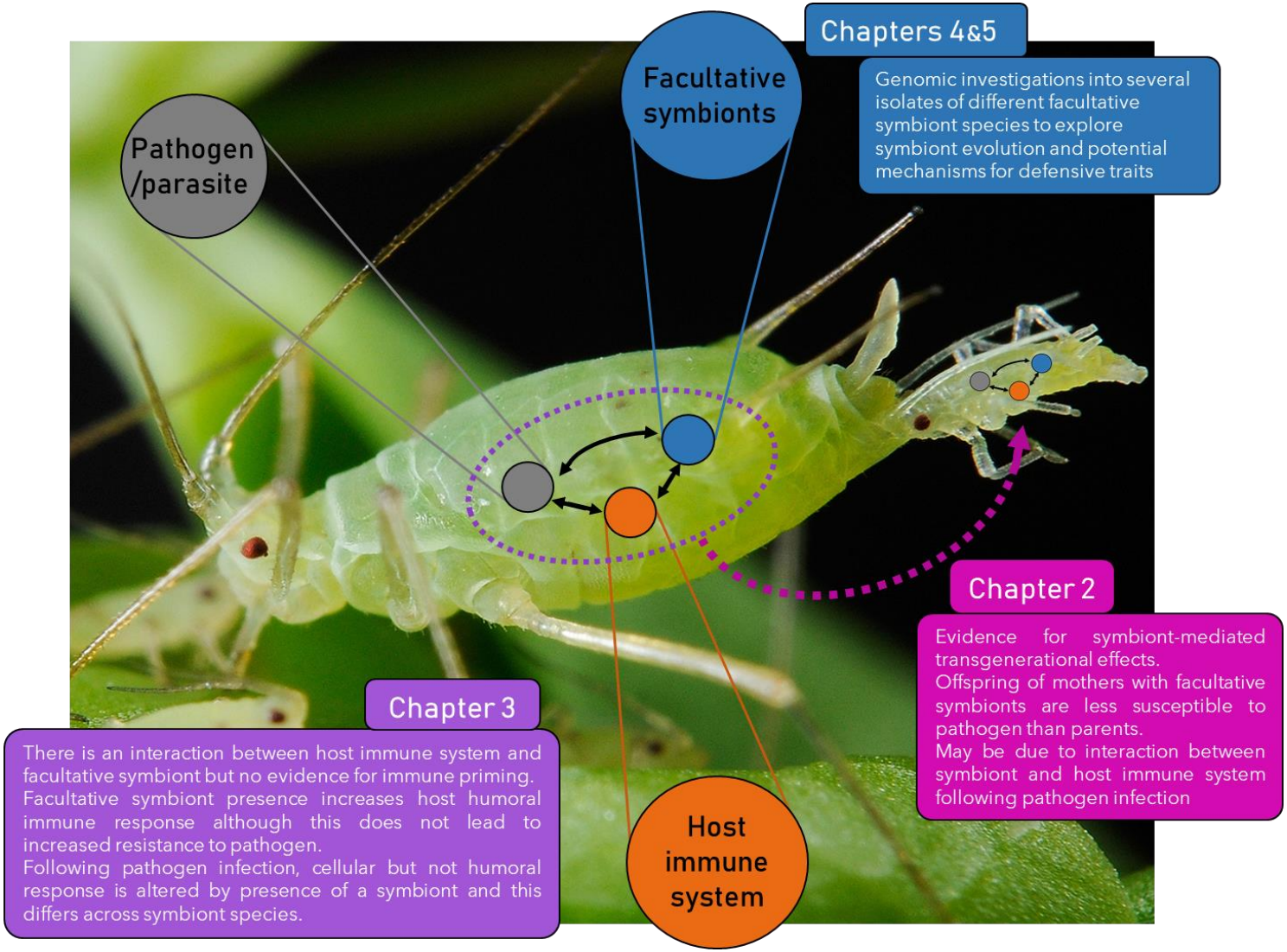


Figure 6-1: Graphical summary of focus and findings of each chapter within the thesis. The thesis broadly investigates the interactions between aphids and their facultative symbionts and the role this plays in host defence. Chapter 2 shows, for the first time, symbiont-mediated transgenerational effects, with a reduction of susceptibility to a pathogen across two generations in those harbouring facultative symbionts. Chapter 3 investigates the interaction between the facultative symbionts and the host immune system showing symbionts do influence the host immune system, though this interaction is complex and is altered in the presence of an invading pathogen. Chapters 4 and 5 outline the genome evolution of two aphid facultative symbionts, outlining the differences across and within species and highlighting potential mechanisms for the phenotypes conferred by these species.

Implications of thesis findings

The four chapters presented in this thesis provide new understanding of host-symbiont interactions, their role in defence and symbiont evolution, whilst opening up new areas of exploration within these host-symbiont relationships. They reinforce the knowledge that the interactions between the host and symbiont are highly complex and that the environment can have large impacts on these interactions and thus host and symbiont evolution.

Scope for symbiont influence on transgenerational effects

Transgenerational effects, where mothers influence their offspring's phenotype, are seen throughout nature (Bonduriansky et al. 2012, Mousseau & Fox 1998b, Uller 2008). In recent years we have seen increasing evidence that these transgenerational effects can decrease insect susceptibility to a variety of infectious organisms, by priming the host immune system against pathogens that previous generations have already experienced (reviewed in Roth et al. 2018). This can in some sense be likened to the adaptive immune system of vertebrates, which carries "memories" of previously encountered threats, and the ability to protect against them. Beyond the immune system, symbionts, many of which are transmitted vertically from parent to offspring, also play a role in defending insect hosts against some of these immunological threats (Feldhaar 2011, Oliver et al. 2010). There is a plethora of evidence that symbionts can influence host biology and of the complex interactions between the host and the symbiont. Yet, the presence of symbionts, their interactions and the role this may play in transgenerational effects has been largely overlooked.

Aphids harbour several facultative symbionts, many of which aid in their defence against natural enemies (Feldhaar 2011, Moran et al. 2008, Oliver et al. 2010). Given their parthenogenetic reproduction, aphids are a system where transgenerational effects would be likely to occur. However, thus far there has been limited evidence for transgenerational effects in aphids (Vorburger et al. 2008). In Chapter 2, we show that transgenerational effects do occur in response to a pathogen infection in aphids harbouring facultative

symbionts, an effect which is not found in those without symbionts. We found offspring from mothers harbouring symbionts had decreased susceptibility to a pathogen but there was no change in susceptibility in those without. This is the first time symbiont-meditated transgenerational effects have been shown in insects. We also found that the symbiont densities increased in the second generation which we hypothesised was related to the interaction between symbiont and host immune system, which may have driven these transgenerational effects. It may be that offspring of mothers that have been infected with the pathogen concentrate their immune resources towards protecting against future attacks from the pathogen, where they may previously have been investing more heavily in the control of symbiont populations. This would explain why there were increased densities of these symbionts in the second generation. Further, *H. defensa* has previously been found to be a fast replicator (Laughton et al. 2014) and this is where there was the greatest increase in symbiont numbers.

In Chapter 2, we found that harbouring these symbionts, in the first generation, made aphids more susceptible to the pathogen than those without and it could be this increased susceptibility that allowed for these transgenerational effects to occur. Aphids may have been able to counteract the negative effects of harbouring facultative symbionts, in this case, increased susceptibility, although they may not be able to "prime" their immune response against future attacks, explaining why we only see the transgenerational effects occurring in symbiont harbouring individuals. It would be interesting to explore whether this effect is found in response to other infectious organisms, such as parasitoids, or whether this effect is specific and requires a negative effect of symbiont presence to mitigate against. Symbiont involvement in shaping host immunity across generations could lead to a range of consequences for host, symbiont and pathogen evolution.

Beyond this, it would be beneficial to explore whether these transgenerational effects only occur in response to threats to the immune system of the host or whether there could be other ecological influences on these interactions, such host plants and climate stresses, that could too see symbiont-mediated transgenerational effects. For example, these

transgenerational effects may allow for aphids to become increasingly tolerant of heat stress or be able to more easily adapt to host plant shifts, allowing populations to spread more easily.

There is much evidence for transgenerational effects in a range of insect species, especially in increasing resistance to pathogens and parasites (Freitak et al. 2009, Moret 2006, Roth et al. 2009, Sadd et al. 2005) and it has been shown that both maternal and paternal environment can influence the phenotype of their offspring (Roth et al. 2010). Likewise, there has been a large body of research on how symbionts may be able to increase the host's protection to pathogens and parasites (Feldhaar 2011). Yet these two factors have been kept separate, with no investigations into whether the symbiotic bacteria, that often exists in insects, can influence transgenerational effects. There was one study which investigated whether bacteria, though not symbionts, could influence immunity across generations in *Lepidoptera*, where they found that eggs from mothers that had ingested the bacteria had increased immune gene expression (Freitak et al. 2014). Freitak et al. (2014) also discovered that this transgenerational immune priming was due to the maternal transfer of the bacteria to the developing eggs, highlighting the transfer of bacteria across generations as a potential mechanism for transgenerational immune priming. This suggests there is scope for the transfer of bacterial species, including symbionts, leading to transgenerational effects in other insect species.

Overall, the potential for these transgenerational effects highlights another level of complexity to these already complicated interactions, which we are yet to even begin to understand. However, further research is required to determine the full scope of these effects, their generality and the mechanisms by which they occur.

Increasing complexities in host-symbiont interactions

Whilst symbionts might aid in host defence through their own defensive mechanisms, it is also important to consider the role host-symbiont interactions may play in defence. There is increasing evidence that the interaction between symbionts and their host can lead to

increased resistance to pathogens and parasites (Eleftherianos et al. 2013), yet these are not clear-cut effects seen across all host-symbiont relationships. In some cases, symbiont presence can increase host protection (Kambris et al. 2009, Moreira et al. 2009, Teixeira et al. 2008), but there are also examples where this provides no change in susceptibility in the host (Rottschaefer & Lazzaro 2012, Wong et al. 2011). In Chapter 2 we showed, within the pea aphid, susceptibility can be decreased across generations, likely due to an interaction between the symbionts and the host immune system. In the subsequent chapter, therefore, we investigated symbiont influence on the host immune response and whether this altered the host response to a pathogen. In Chapter 3 we found that harbouring a facultative symbiont led to increased expression of some immune genes, although there were no differences in cellular immunity. However, following pathogen infection, symbiont presence did alter the host's cellular immunity, with those harbouring symbionts having a reduced cellular immune response. Increases in immune gene expression due to symbiont presence does occur in other insect species. For example, some strains of *Wolbachia* increase immune gene expression in mosquitoes, which is thought to protect the host against invading organisms (Kambris et al. 2009, 2010). However, other strains have not induced increased expression suggesting these effects may be strain-specific (Hughes et al. 2011). Increased immune gene expression has also recently been found in bees in response to their beneficial symbiont *Snodgrassella alvi*, which too increased their resistance to a pathogen (Horak et al. 2020). Interestingly, the increased immune gene expression due to the presence of symbionts in Chapter 3 did not lead to the aphids being more resistant to a pathogen as we see in other systems, and most of the expression was downregulated following pathogen infection. One caveat of the experiment carried out in this chapter was that we only investigated a subset of the genes within the immune system of the aphid, which were chosen due to their likely involvement in the immune response to fungal pathogens. However, across all treatments, there were no large changes in the expression of these genes in response to the pathogen, and, given that we saw symbiont influence on the cellular immunity following the infection, it is likely that the humoral response was also affected. As these immune responses are often hard to pin-point,

looking across the entire transcriptome of the host may give more insight to this. Nevertheless, we did see that the presence of symbionts altered the expression of these genes, be it that they were involved in pathogen resistance or not. This highlights an interaction between the host immune system and their symbiotic partners, and this may be due to the host employing these immune genes to control symbiont populations. It would be beneficial to further explore these host-symbiont interactions and the role they may play in aiding the host in defence against a range of pathogens and parasites. It would be particularly interesting to explore the transcriptome of the host harbouring different isolates of the same species, which confer different levels of benefit and cost to the host, to investigate how these may be influencing the host gene expression differently. It would also be valuable to examine the symbiont transcriptome to understand how the symbiont too may be responding to changes in the host, and how this may differ after the introduction of a pathogen.

In Chapter 3 the changes in gene expression that occurred in those harbouring symbionts, following pathogen infection, suggest that the introduction of another organism may lead to changes in host-symbiont interactions. This is also what we hypothesised was leading to the changes across generations in Chapter 2. One interesting avenue for exploration may be how these interactions may lead to changes in the ability for new symbiont relationships to form. Often aphids carry more than one symbiont species, but there are differences in which symbionts are able to coexist, some more easily than others (Heyworth 2015, Weldon et al. 2020). This is thought to be largely due to resource allocation and how closely related the two symbionts are, however as the immune response of aphids is altered by the presence of a symbiont this may then have some effect on the ability of new symbionts to invade. For example, if this immune response is, in fact, working to control symbiont populations, as hypothesised, newly invading symbiont species may find it difficult to populate within a host that is already working to control symbiont populations, as it may too be targeted. Alternatively, this could work in favour of these symbionts, which are able

to compete for resources more effectively against a symbiont population that is being controlled by the host.

Uncovering whether interactions between the host and symbiont alters the aphid's ability to transfer viruses would be of high importance. Aphids are a significant agricultural pest as they transmit half of the insect-vectored plant viruses, many of which cause a huge amount of damage to crops (Ng & Perry 2004, Tatchell 1989). In other systems, symbionts can alter the hosts ability to transmit viruses (Kliot et al. 2014) and so this may also be effective here. If there are symbiont influences on the hosts ability to transmit viruses, this could be harnessed to better control the transmission of these viruses, having applications for agriculture and food security.

Our understanding of the interactions between the host and symbionts, especially in economically important species, can have a variety of applications from food security to biocontrol (Douglas 2007, Vorburger 2018). It is also important for understanding how ecological impacts, such as climate change, may alter these relationships and lead to different outcomes. Overall, there is still much to be explored before we truly understand the complexities of these host-symbiont interactions and how they can influence the evolution of these species.

Interactions and their role in symbiont evolution

The interactions between host and symbiont can have large effects on both host and symbiont evolution. In addition to experimental studies investigating the interactions between these organisms, to truly understand the relationships we too must explore these effects at the genomic level. Analysing genomes can play an important part in understanding the cause and consequences of different aspects of these interactions, giving insight into what determines a successful host-symbiont relationship. We investigated the genomes of two different species of aphid facultative symbiont, *F. symbiotica* (Chapter 4) and *H. defensa* (Chapter 5), to complement our understanding of the evolution of these two species, that have different relationships with their host. In

addition to comparing across species, we looked at different isolates within species to highlight the genomic differences that can occur within a species and whether these may influence the host relationship.

The genomic investigations into various *H. defensa* strains showed that this species has diversified more than other closely related symbionts. The strains found in whitefly are very different from those within aphids, but there is also divergence across strains from the same host species. These fall into different clades and those that share a clade have more similarities to one another than to the other isolates. These clades also outline the different APSE variant, which provides the protection against a parasitoid wasp, carried by each isolate.

Within *F. symbiotica* we see large phenotypic differences across isolates, but these are not partnered with large genomic variation. Unlike *H. defensa* the core genes of *F. symbiotica* appear to fall within one clade. However, although the *F. symbiotica* isolates had more similarities within their core genes than *H. defensa*, overall, there was more variation outside of these, for example in their mobile DNA and toxin genes. The isolates also did not appear to fall into any distinct groups like the various isolates of *H. defensa*. In *F. symbiotica*, some of the metabolic pathways that are lost in other closely related facultative symbionts were retained, suggesting that this species has not undergone the same level of genome reduction as these other species. This could indicate that the relationship between *F. symbiotica* and aphids is a more recent one, or it may be that this symbiont has a more turbulent relationship with the host than others. This, however, may be beneficial for this symbiont species, which often occurs in co-infections, as it has less resource to compete for.

As well as within-host effects, external environment may also influence symbiont evolution. Because of this, it may be expected to see a range of evolutionary trajectories for the same species, within the same host from different environments. For example, aphid populations that often come into contact with parasitoids may be more likely to form close relationships with their *H. defensa* strains, and thus these particular isolates may undergo a different

evolutionary path than those that are carried by hosts less dependent on *H. defensa*s defensive mechanisms.

Given that insects also carry an obligate symbiont on which they rely for the production of amino acids, it is also important to consider how this too may be influencing the evolution of the facultative symbionts. *H. defensa* in whitefly, for example, lives within bacteriocytes with their obligate symbiont, and acts as a nutritional parasite (George & Toole 2017). Within the pea aphid, *H. defensa* has lost the ability to synthesise many of the ten essential amino acids, suggesting it relies on the aphid's obligate symbiont to provide these. This suggests that *H. defensas* relationship with the aphid host is a stable, long-term relationship, which has allowed for the loss of these vital pathways. As well as existing beside the obligate symbiont, many of these facultative symbionts co-exist within the aphid, *F. symbiotica* and *H. defensa* being a common partnership (Doremus & Oliver 2017). Other symbiont partnerships, however, such as *F. symbiotica* and *R. insecticola* do not appear to be able to co-exist as easily (Heyworth 2015), although it is not fully understood why this is. It may be that this is due to symbionts competing for resources, or due to host control of redundant symbiont populations, given they provide similar benefits. Exploring the genomes of species that are more, or less, often found together may give insight into whether this is directly due to the symbionts or whether it is related to host-symbiont interactions.

As our understanding of these symbionts and their interactions with the host grows, we should be further exploring the concept of the hologenome; "the union of all the genes within the host and the microbiome" (Bordenstein & Theis 2015, Rosenberg & Zilber-Rosenberg 2018). This requires a new understanding of these genomes, their interactions, and how the evolution of the host and its microbes come together as a collective. There has been some controversy regarding this concept, which in turn has led to it being the subject of a number of reviews (Douglas & Werren 2016, Moran & Sloan 2015, Morris 2018, Rosenberg & Zilber-Rosenberg 2018, Theis et al. 2016). It appears those who are sceptical of the concept are so due to the lack of clarity over the definition, limitations and

importance of the idea of the hologenome (Douglas & Werren 2016, Moran & Sloan 2015, Theis et al. 2016). Some argue that the relevance of the holobiont as a unit of selection has been exaggerated and is likely the exception rather than the rule, especially given that microbes are not always transmitted vertically and do not always share a "unified purpose" with the host (Douglas & Werren 2016, Moran & Sloan 2015). However, others feel this scepticism is unfounded (Morris 2018, Theis et al. 2016) and highlight that the concept does not consider the hologenome as the only, or most important, level at which selection can work (Bordenstein & Theis 2015, Rosenberg & Zilber-Rosenberg 2013). The pea aphid, in particular, is an ideal model in which to explore this concept given there is a large amount of variation across biotypes and their microbial composition, as well as harbouring microbes that have different modes of transmission. Analysing these organisms at the level of the hologenome may further our understanding of these interactions, the ecological impacts and their evolutionary outcomes.

Mechanisms behind phenotypes

Although there has been a large body of work investigating the benefits conferred by facultative symbionts, the mechanisms behind many of these are still largely unknown. We do have a good understanding of the mechanism of parasitoid protection conferred by *H. defensa*, which is due to the bacteriophage APSE, however, this is one of the few well-understood mechanisms (Degnan & Moran 2007; Oliver et al. 2003, 2005). Likewise, investigations into the genomes of different strains of *R. insecticola* have aimed to uncover the mechanisms behind one strain's parasitoid protection, although this only provided speculative results, highlighting a range of pathogenic factors which may be involved (Hansen et al. 2012). In Chapter 4 we investigated the genomes of different isolates of *F. symbiotica* which lead to different host phenotypes (Smee et al. 2021). We highlighted some potential mechanisms for both the beneficial and costly phenotypes we see with infection of this species. We outlined a potential set of toxins found within the *F. symbiotica* genomes which were not found in other closely related symbiont species and have insecticidal effects in other bacterial species such as *Photorhabdus* and *Xenorhabdus*.

These could potentially be having an effect on the ability of the aphids to reproduce or directly affecting the embryos themselves. However, this would not explain why there are different levels of cost to harbouring different isolates. It may be interesting to explore the expression of this gene across different aphid ages to see if these genes become more effective as aphids reach adulthood and begin to reproduce. Likewise, exploring the expression of these within the embryos may also give insight into whether these genes are playing a role in the fecundity cost. However, costs with host-symbiont relationships are not uncommon (Zytynska et al. 2019), and these may instead be due to a more subtle side-effect of these interactions rather than a particular symbiont mechanism that is directly affecting the reproduction of the host.

We also highlighted a possible mechanism for the fungal pathogen resistance conferred by *F. symbiotica*. We identified a non-ribosomal peptide synthase (NRPS) cluster within the genome of all isolates, which was the same in the two isolates that confer high levels of protection against a pathogen. Furthermore, this NRPS shared similarity to those found in other bacterial species that closely interact with fungi (Chapter 4, Spraker et al. 2016). There is some contradictory evidence for this, as these genes were also found in a previously published strain of *F. symbiotica*, Ap5D, which does not provide any resistance to a pathogen (Chapter 4, Doremus & Oliver 2017). However, the predicted product of the NRPS cluster was different to that in the protective isolates, and that may be what distinguishes between these isolates ability to defend against the pathogen. This is the first time the genomes of different isolates of *F. symbiotica* within the same host species have been investigated and any mechanisms highlighted. Future work could focus on exploring this NRPS cluster further, and its possible involvement in fungal pathogen resistance, potentially using gene expression studies in response to a fungal pathogen. The protective phenotype in *F. symbiotica* was previously found to be in response to *Pandora neoaphidis* (Heyworth & Ferrari 2015), although the results in this thesis are from infection with a different fungal pathogen, *Lecanicillium lecanii*, and there is some suggestion in Chapter 2 and 3 that *F. symbiotica* may be able to provide some level of protection against this

pathogen too. In the case that the NRPS is having an impact on fungal pathogen resistance, it may also be beneficial to explore the specificity of the bioactive product, and whether different NRPSs may be more, or less, effective against different fungal species.

Given insect symbionts are largely unculturable, it can be difficult to carry out confirmatory studies on these potential mechanisms. However recent advances in culturing *H. defensa* have allowed for the transfer of APSE into phage-free strains to confirm the role of APSE in the protection against parasitoids (Brandt et al. 2017, Chevignon et al. 2018). *F. symbiotica* has also recently been successfully cultured, though at lower growth rates than *H. defensa* (Patel et al. 2019). Therefore, future in-vitro studies of this symbiont may be able to better investigate the potential mechanisms behind these symbiont-mediated phenotypes.

Applications for insect symbiont research

As our research in the field of symbiosis continues to expand, there has been a notable increase in evidence for the potential applications of many of these symbiotic microbes, including those found in insects. Applications range from those that are ecologically relevant, including pest management and bioremediation to those for human health and the prevention of disease spread.

Some of the best-studied insect-microbe relationships are within pest species, such as aphids, whitefly and mealybugs, as well as others. Understanding the interactions between these species and their symbionts could potentially help control these pest species, having important agricultural implications. Aphids, for example, transmit half of all insect-vectored plant viruses, leading to a large amount of crop damage (Ng & Perry 2004, Tatchell 1989). Within whitefly, symbionts have been seen to impact this species ability to transmit viruses (Kliot et al. 2014). Understanding the interactions between plant viruses, the insect vectors and their symbiotic bacteria, may allow us to harness these effects to better control the spread of viruses within important crop plants. Likewise, as symbionts can also influence the susceptibility to natural enemies, biocontrol agents can quickly become less effective.

Symbiont-mediated protection against parasitoids, a common biocontrol agent, provided by *H. defensa* in aphids is a common example of this (Oliver et al. 2003). As infected aphids are less susceptible to these, overuse of parasitoids as biocontrol could lead to an increase in *H. defensa* infected aphids, as they have a fitness advantage. This could then go on to have wider implications for other interactions that may also be affected by the presence of these defensive symbionts. Additionally, this could influence the genome evolution of these symbiotic partners (Chapter 5), which in turn could lead to changes in how these symbionts interact with their host and environmental influences. Further to this, in Chapter 2, we showed facultative symbionts led to decreased susceptibility to a pathogen across generations. If these symbiont-mediated transgenerational effects occur in response to other organisms, such as parasitoids or plant viruses, this too could influence the effectiveness of repeated exposure to the same biocontrol agent or the transmission of economically important plant viruses.

Whilst it is necessary to be able to control pest species, it is also crucial to consider the health of important pollinators, such as bees. Here too, symbiotic bacteria can play a vital role. There is evidence that a supplement of bacteria from the honeybee gut to a larval diet can help reduce infection with American foulbrood disease (AFB), a disease that prevents the development of bee larvae (Olofsson & Vásquez 2008). Furthermore, new molecular approaches have found other ways to improve honeybee health. Genetic alteration of one of their symbiotic species led to changes in the expression of immune genes which provided protection against viral infection as well as killing their parasitic mites (Leonard et al. 2020).

Outside of these agriculturally relevant applications, insect symbionts have been found to have applications in human health. In Chapter 3 we found that aphid symbionts altered the expression of immune genes, although in this case, we did not find it influenced the hosts resistance to a pathogen. However, in other systems, these interactions between the host immune system and their symbionts do affect insects ability to be infected with different

organisms, and this has been harnessed for human disease control. One well-studied example of this is the research into *Wolbachia* and the effects it can have on the transmission of vector-borne diseases such as malaria, dengue and yellow fever, to name a few (Ant et al. 2018, Kambris et al. 2009, Moreira et al. 2009, Walker et al. 2011). Research has shown that the presence of *Wolbachia* can reduce the ability of invading pathogens to reproduce, leading to a reduction in pathogens carried by *Wolbachia*-infected vectors (Caragata et al. 2019, Hussain et al. 2013, Kambris et al. 2010, Pacidônio et al. 2017). This can also be enhanced by harnessing some of the "costs" that are usually associated with *Wolbachia* infection, as a reproductive manipulator. This is done via two methods. Firstly, by releasing *Wolbachia* infected males that are incompatible with the females, there is a reduction in vector numbers, due to the production of nonviable offspring (Flores & O'Neill 2018). Secondly, rather than reducing the vector numbers, releasing *Wolbachia* infected individuals to instead replace the uninfected individuals, will, in turn, reduce the pathogen load as described above (Hoffmann et al. 2011). These methods are being heavily researched and in some parts of the world have begun to be implicated.

Beyond this biocontrol of human diseases, there is also the scope for the use of symbiont produced novel compounds and enzymes to be used in human medicine. For example, polyketide synthases (PKS) and non-ribosomal peptide synthetases (NRPS) produce bioactive compounds and are found widely in insect defensive symbionts (Flores & O'Neill 2018). These compounds, and the like, may be able to produce new antibiotic and anti-fungal medicines which could be used to treat human infections. Genomic investigations into various symbiotic bacteria allows us to continue to identify novel compounds which may have uses beyond the effects they confer in their insect host. In the genome of *F. symbiotica* for example (Chapter 4), we found an NRPS gene cluster which we highlighted as a potential mechanism for the fungal pathogen resistance conferred by this symbiont. It is not to say, however, that the compound produced by this NRPS could not have other applications outside of aphid pathogen defence. This highlights the importance

of genomic investigations into these symbionts, in addition to other experimental approaches.

Many of the applications described above rely on a better understanding of the interaction between bacterial symbionts and their host. For applications for disease and vector control, we also need to further our knowledge on how these interactions can alter the evolution of infectious organisms, and how infectious organisms may too alter the interaction between symbionts and their host. Across many systems, including the pea aphid, we still have little knowledge of the mechanisms behind many of the defensive traits conferred by symbionts and whether these are direct symbiont effects or are due to host interactions. Understanding these effects and the mechanisms by which they occur may allow us to harness some of these phenotypes more effectively for our own benefit. Overall, however, there is still much to discover about these complex communities, their capabilities and interactions, and their importance for plant and animal health.

Summary and conclusions

Insects face a range of environmental stressors, from natural enemies to changes in climate and their endosymbionts often help them overcome many of these threats. The interactions between the host and their endosymbionts, however, can lead to a range of phenotypic and genotypic outcomes for both organisms.

My work begins to explore these interactions and the impact this may have on defence against natural enemies. I highlight the complexities of these systems, showing that whilst symbionts may influence the host immune response, this may not always alter how a host responds to a pathogen, yet pathogens may alter the host-symbiont interactions. I also show that these interactions may impact, not just the host, but may also be able to transmit across generations, impacting the host's offspring. I also explore potential mechanisms for defence and the genome evolution of endosymbionts which have different relationships

with the host, showing that there are still gaps in our understanding of these systems, their effects and the underlying mechanisms.

Host-symbiont interactions, and the external pressures that may alter them, can lead to the evolution of both the symbionts and their hosts, and to fully understand the impacts we must continue to explore both the genotypic and phenotypic aspects of these relationships. The pea aphid as a model system has taught us much and continues to teach us about the interactions between endosymbionts and their hosts, however, there is still much to be learnt within this model and beyond. As with many explorations, it appears the more we learn, the more questions we have.

Appendix A

NCBI reference numbers for the bacterial species included in the core gene phylogenies in Chapter 4 (Figure 4-3) and Chapter 5 (Figure 5-2).

Fukatsuia symbiotica (*Cinara*) – FQSP02000375.1

Fukatsuia symbiotica Ap5D – CP021659.1

Hamiltonella defensa 5AT – GCA_000021705.1

Hamiltonella defensa A2C – GCA_002777195.1

Hamiltonella defensa AS3 – GCA_002777215.1

Hamiltonella defensa MEAM1 - GCA_002285855.1

Hamiltonella defensa MI47 - GCA_002269405.1

Hamiltonella defensa NY26 – GCA_002777295.1

Hamiltonella defensa ZA17 – GCA_002777235.1

Photorhabdus luminescens – BX470251.1

Regiella insecticola 5.15 – GCA_000284655.1

Regiella insecticola LSR1 – GCA_000143625.1

Serratia marcescens – NZ_HG326223.1

Sodalis glossindius – NC_007712.1

Yersinia enterocolitica – NC_008800.1

Yersinia pestis – NC_003143.1

Yersinia pseudotuberculosis – NZ_CP008943.1

Appendix B

35 bacterial species included in the KEGG KAAS annotation in Chapters 4 and 5.

Aeropyrum pernix

Aquifex aeolicus

Arabidopsis thaliana (thale cress)

Archaeoglobus fulgidus DSM 4304

Bacillus subtilis subsp. subtilis 168

Borreliella burgdorferi B31

Buchnera aphidicola APS (Acyrthosiphon pisum)

Candidatus Hamiltonella defensa (Acyrthosiphon pisum)

Chlamydia trachomatis D/UW-3/CX

Clostridium acetobutylicum ATCC 824

Drosophila melanogaster (fruit fly)

Escherichia coli K-12 MG1655

Haemophilus influenzae Rd KW20 (serotype d)

Halanaerobium hydrogeniformans

Helicobacter pylori 26695

Lactococcus lactis subsp. lactis Il1403

Mesorhizobium japonicum MAFF 303099

Methanocaldococcus jannaschii

Mycobacterium tuberculosis H37Rv

Mycoplasma genitalium G37

Neisseria meningitidis MC58 (serogroup B)

Photorhabdus laumondii subsp. laumondii TTO1

Plasmodium falciparum 3D7

Pseudomonas aeruginosa PAO1

Pyrococcus horikoshii

Rickettsia prowazekii Madrid E

Saccharomyces cerevisiae (budding yeast)

Salmonella enterica subsp. enterica serovar Typhi CT18

Serratia proteamaculans

Staphylococcus aureus subsp. aureus N315 (MRSA/VSSA)

Streptococcus pneumoniae TIGR4 (virulent serotype 4)

Synechocystis sp. PCC 6803

Yersinia enterocolitica subsp. enterocolitica 8081 (serotype 0:8)

Yersinia pestis CO92 (biovar Orientalis)

Vibrio cholerae O1 El Tor N16961

References

Akman Gündüz E, Douglas AE. 2009. Symbiotic bacteria enable insect to use a nutritionally inadequate diet. *Proc. R. Soc. B*. 276:987–91

Altincicek B, Knorr E, Vilcinskas A. 2008. Beetle immunity: Identification of immune-inducible genes from the model insect Tribolium castaneum. *Dev. Comp. Immunol.* 32(5):585–95

Ant TH, Herd CS, Geoghegan V, Hoffmann AA, Sinkins SP. 2018. The Wolbachia strain wAu provides highly efficient virus transmission blocking in Aedes aegypti. *PLOS Pathog.* 14(1):e1006815

Arndt D, Grant JR, Marcu A, Sajed T, Pon A, et al. 2016. PHASTER: a better, faster version of the PHAST phage search tool. *Nucleic Acids Res.* 44(W1):W16–21

Aziz RK, Bartels D, Best AA, DeJongh M, Disz T, et al. 2008. The RAST Server: Rapid Annotations using Subsystems Technology. *BMC Genomics*. 9(1):75

Barribeau SM, Parker BJ, Gerardo NM. 2014. Exposure to natural pathogens reveals costly aphid response to fungi but not bacteria. *Ecol. Evol.* 4(4):488–93

Bates D, Mächler M, Bolker BM, Walker SC. 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67(1):1–48

Baumann P. 2005. Biology Of Bacteriocyte-Associated Endosymbionts Of Plant Sap-Sucking Insects. *Annu. Rev. Microbiol.* 59(1):155–89

Blin K, Shaw S, Steinke K, Villebro R, Ziemert N, et al. 2019. antiSMASH 5.0: updates to the secondary metabolite genome mining pipeline. *Nucleic Acids Res.* 47(W1):W81–87

Bloudoff K, Schmeing TM. 2017. Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: discovery, dissection and diversity. *Biochim. Biophys. Acta - Proteins Proteomics*. 1865(11):1587–1604

Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics*. 30:2114–20

Bonduriansky R. 2012. Rethinking heredity, again. *Trends Ecol. Evol.* 27(6):330–36

Bonduriansky R, Crean AJ, Day T. 2012. The implications of nongenetic inheritance for evolution in changing environments. *Evol. Appl.* 5(2):192–201

Bonduriansky R, Day T. 2009. Nongenetic Inheritance and Its Evolutionary Implications. *Annu. Rev. Ecol. Evol. Syst.* 40(1):103–25

Bordenstein SR, Theis KR. 2015. Host Biology in Light of the Microbiome: Ten Principles of Holobionts and Hologenomes. *PLOS Biol.* 13(8):e1002226

Boulinier T, Staszewski V. 2008. Maternal transfer of antibodies: raising immuno-ecology issues. *Trends Ecol. Evol.* 23(5):282–88

Brandt JW, Chevignon G, Oliver KM, Strand MR. 2017. Culture of an aphid heritable symbiont demonstrates its direct role in defence against parasitoids. *Proc. R. Soc. B Biol. Sci.* 284(1866):20171925

Brisson JA, Stern DL. 2006. The pea aphid,Acyrthosiphon pisum: an emerging genomic model system for ecological, developmental and evolutionary studies. *BioEssays*. 28(7):747–55

Brownlie JC, Johnson KN. 2009. Symbiont-mediated protection in insect hosts. *Trends Microbiol.* 17:348–54

Brumin M, Kontsedalov S, Ghanim M. 2011. Rickettsia influences thermotolerance in the whitefly Bemisia tabaci B biotype. *Insect Sci.* 18(1):57–66

Bulet P, Stocklin R, Menin L. 2004. Anti-microbial peptides: from invertebrates to vertebrates. *Immunol. Rev.* 198(1):169–84

Burke GR, Moran NA. 2011a. Responses of the pea aphid transcriptome to infection by facultative symbionts. *Insect Mol. Biol.* 20(3):357–65

Burke GR, Moran NA. 2011b. Massive Genomic Decay in Serratia symbiotica, a Recently Evolved Symbiont of Aphids. *Genome Biol. Evol.* 3:195–208

Byeon JH, Seo ES, Lee JB, Lee MJ, Kim JK, et al. 2015. A specific cathepsin-L-like protease purified from an insect midgut shows antibacterial activity against gut symbiotic bacteria. *Dev. Comp. Immunol.* 53(1):79–84

Caragata EP, Rocha MN, Pereira TN, Mansur SB, Dutra HLC, Moreira LA. 2019. Pathogen blocking in wolbachia-infected aedes aegypti is not affected by zika and dengue virus co-infection. *PLoS Negl. Trop. Dis.* 13(5):

Cardoza YJ, Klepzig KD, Raffa KF. 2006. Bacteria in oral secretions of an endophytic insect inhibit antagonistic fungi. *Ecol. Entomol.* 31(6):636–45

Cayetano L, Rothacher L, Simon J-C, Vorburger C. 2015. Cheaper is not always worse: strongly protective isolates of a defensive symbiont are less costly to the aphid host. *Proc. R. Soc. B Biol. Sci.* 282(1799):20142333

Cerenius L, Lee BL, Söderhäll K. 2008. The proPO-system: pros and cons for its role in invertebrate immunity. *Trends Immunol.* 29(6):263–71

Chandler M, Siguier P. 2013. Insertion Sequences. *Brenner's Encycl. Genet. Second Ed.* 62(3):86–94

Chen F, Mackey AJ, Stoeckert CJ, Roos DS. 2006. OrthoMCL-DB: querying a comprehensive multi-species collection of ortholog groups. *Nucleic Acids Res.* 34(90001):D363–68

Chevignon G, Boyd BM, Brandt JW, Oliver KM, Strand MR. 2018. Culture-Facilitated Comparative Genomics of the Facultative Symbiont Hamiltonella defensa. *Genome Biol. Evol.* 10(3):786–802

Christophides GK, Vlachou D, Kafatos FC. 2004. Comparative and functional genomics of the innate immune system in the malaria vector Anopheles gambiae. *Immunol. Rev.* 198:127–48

Coil D, Jospin G, Darling AE. 2014. A5-miseq: an updated pipeline to assemble microbial genomes from Illumina MiSeq data

Comandatore F, Sassera D, Montagna M, Kumar S, Koutsovoulos G, et al. 2013. Phylogenomics and analysis of shared genes suggest a single transition to mutualism in Wolbachia of nematodes. *Genome Biol. Evol.* 5(9):1668–74

Coote JG. 1992. Structural and functional relationships among the RTX toxin determinants of gram-negative bacteria. *FEMS Microbiol. Rev.* 8(2):137–61

Costa TRD, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, et al. 2015. Secretion systems in Gram-negative bacteria: Structural and mechanistic insights

Dalebroux ZD, Matamouros S, Whittington D, Bishop RE, Miller SI. 2014. PhoPQ regulates acidic glycerophospholipid content of the Salmonella Typhimurium outer membrane. *Proc. Natl. Acad. Sci. U. S. A.* 111(5):1963–68

Dalebroux ZD, Miller SI. 2014. Salmonellae PhoPQ regulation of the outer membrane to

resist innate immunity. *Curr. Opin. Microbiol.* 17:106–13

De Buck E, Lammertyn E, Anné J. 2008. The importance of the twin-arginine translocation pathway for bacterial virulence

De Roode JC, Lefèvre T. 2012. Behavioral immunity in insects. *Insects*. 3(3):789–820

Dedeine F, Vavre F, Fleury F, Loppin B, Hochberg ME, Bouletreau M. 2001. Removing symbiotic Wolbachia bacteria specifically inhibits oogenesis in a parasitic wasp. *Proc. Natl. Acad. Sci.* 98(11):6247–52

Dedryver CA, Le Ralec A, Fabre F. 2010. The conflicting relationships between aphids and men: A review of aphid damage and control strategies. *Comptes Rendus - Biol.* 333(6–7):539–53

Degnan PH, Leonardo TE, Cass BN, Hurwitz B, Stern D, et al. 2010. Dynamics of genome evolution in facultative symbionts of aphids. *Environ. Microbiol.* 12(8):2060–69

Degnan PH, Moran NA. 2007. Evolutionary genetics of a defensive facultative symbiont of insects: exchange of toxin-encoding bacteriophage. *Mol. Ecol.* 17:916–29

Degnan PH, Moran NA. 2008. Diverse phage-encoded toxins in a protective insect endosymbiont. *Appl Env. Microbiol.* 74(21):6782–91

Degnan PH, Yu Y, Sisneros N, Wing RA, Moran NA. 2009. Hamiltonella defensa, genome evolution of protective bacterial endosymbiont from pathogenic ancestors. *Proc. Natl. Acad. Sci.* 106(22):9063–68

Derzelle S, Turlin E, Duchaud E, Pages S, Kunst F, et al. 2004. The PhoP-PhoQ Two-Component Regulatory System of Photorhabdus luminescens Is Essential for Virulence in Insects. *J. Bacteriol.* 186(5):1270–79

Dhinaut J, Chogne M, Moret Y. 2018. Immune priming specificity within and across generations reveals the range of pathogens affecting evolution of immunity in an insect. *J. Anim. Ecol.* 87(2):448–63

Dimijian GG. 2000. Evolving Together: The Biology of Symbiosis, Part 2. *Baylor Univ. Med. Cent. Proc.* 13(4):381–90

Dion E, Polin SE, Simon Jean-Christophe, Outreman Y. 2011a. Symbiont infection affects aphid defensive behaviours. *Biol. Lett.* 7(5):743–46

Dion E, Zélé F, Simon J-C, Outreman Y. 2011b. Rapid evolution of parasitoids when faced with the symbiont-mediated resistance of their hosts. *J. Evol. Biol.* 24(4):741–50

Dionne MS, Schneider DS. 2008. Models of infectious diseases in the fruit fly Drosophila melanogaster. *DMM Dis. Model. Mech.* 1(1):43–49

Dixon AFG, Kindlmann P, Lepš J, Holman J. 1987. Why There are So Few Species of Aphids, Especially in the Tropics. *Source Am. Nat.* 129(4):580–92

Doremus MR, Oliver KM. 2017. Aphid heritable symbiont exploits defensive mutualism. *Appl. Environ. Microbiol.*

Douglas AE. 1989. Mycetocyte Symbiosis in Insects. *Biol. Rev.* 64(4):409–34

Douglas AE. 1997. Parallels and contrasts between symbiotic bacteria and bacterial-derived organelles: Evidence from Buchnera, the bacterial symbiont of aphids. *FEMS Microbiol. Ecol.* 24(1):1–9

Douglas AE. 1998. Nutritional interactions in insect-microbial symbioses: aphids and their symbiotic bacteria Buchnera. *Annu. Rev. Entomol.* 43:17–37

Douglas AE. 2007. Symbiotic microorganisms: untapped resources for insect pest control.

Trends Biotechnol. 25(8):338–42

Douglas AE. 2016. How multi-partner endosymbioses function. *Nat. Rev. Microbiol.* 14(12):731–43

Douglas AE, Werren JH. 2016. Holes in the hologenome: Why host-microbe symbioses are not holobionts

Easom CA, Clarke DJ. 2008. Motility is required for the competitive fitness of entomopathogenic Photorhabdus luminescens during insect infection. *BMC Microbiol.* 8(1):168

Egan F, Barret M, Beuzón CR, Narberhaus F, Schikora A. 2014. The SPI-1-like Type III secretion system: more roles than you think. *Front. Plant Sci.* 5:

Eleftherianos L, Atri J, Accetta J, Castillo JC. 2013. Endosymbiotic bacteria in insects: Guardians of the immune system? *Front. Physiol.* 4 MAR:46

Evans JD, Aronstein K, Chen YP, Hetru C, Imler JL, et al. 2006. Immune pathways and defence mechanisms in honey bees Apis mellifera. *Insect Mol. Biol.* 15(5):645–56

Feldhaar H. 2011. Bacterial symbionts as mediators of ecologically important traits of insect hosts. *Ecol. Entomol.* 36:533–43

Feldhaar H, Gross R. 2008. Immune reactions of insects on bacterial pathogens and mutualists. *Microbes Infect.* 10(9):1082–88

Fenn K, Blaxter M. 2006. Wolbachia genomes: Revealing the biology of parasitism and mutualism. *Trends Parasitol.* 22(2):60–65

Ferrandon D, Imler JL, Hetru C, Hoffmann JA. 2007. The Drosophila systemic immune response: Sensing and signalling during bacterial and fungal infections. *Nat. Rev. Immunol.* 7(11):862–74

Ferrari J, Scarborough CL, Godfray HCJ. 2007. Genetic variation in the effect of a facultative symbiont on host-plant use by pea aphids. *Oecologia*. 153(2):323–29

Ferrari J, Via S, Godfray HCJ. 2008. Population differentiation and genetic variation in performance on eight hosts in the pea aphid complex. *Evolution (N. Y).* 62(10):2508–24

Ferrari J, West JA, Via S, Godfray HCJ. 2012. Population genetic structure and secondary symbionts in host-associated populations of the pea aphid complex. *Evolution (N. Y).* 66:375–90

Flores HA, O'Neill SL. 2018. Controlling vector-borne diseases by releasing modified mosquitoes. *Nat. Rev. Microbiol.* 16(8):508–18

Frantz A, Calcagno V, Mieuzet L, Plantegenest M, Simon J-C. 2009. Complex trait differentiation between host-populations of the pea aphid Acyrthosiphon pisum (Harris): implications for the evolution of ecological specialisation. *Biol. J. Linn. Soc.* 97(4):718–27

Freitak D, Heckel DG, Vogel H. 2009. Dietary-dependent trans-generational immune priming in an insect herbivore. *Proc. R. Soc. B Biol. Sci.* 276(1667):2617–24

Freitak D, Schmidtberg H, Dickel F, Lochnit G, Vogel H, Vilcinskas A. 2014. The maternal transfer of bacteria can mediate trans-generational immune priming in insects. *Virulence*. 5(4):547–54

Futahashi R, Tanaka K, Tanahashi M, Nikoh N, Kikuchi Y, et al. 2013. Gene Expression in Gut Symbiotic Organ of Stinkbug Affected by Extracellular Bacterial Symbiont. *PLoS One*. 8(5):64557

Galata V, Fehlmann T, Backes C, Keller A. 2019. PLSDB: A resource of complete bacterial plasmids. *Nucleic Acids Res.* 47(D1):D195–202

Gauthier JP, Legeai F, Zasadzinski A, Rispe C, Tagu D. 2007. AphidBase: A database for aphid genomic resources. *Bioinformatics*. 23(6):783–84

George E, Toole O. 2017. Cooperative Metabolism in a Three-Partner Insect-Bacterial Symbiosis Revealed by Metabolic Modeling. *jb.asm.org 1 J. Bacteriol.* 199:872–88

Gerardo NM, Altincicek B, Anselme C, Atamian H, Barribeau SM, et al. 2010. Immunity and other defenses in pea aphids, Acyrthosiphon pisum. *Genome Biol.* 11(2):R21

Gillespie JP, Burnett C, Charnley AK. 2000. The immune response of the desert locust Schistocerca gregaria during mycosis of the entomopathogenic fungus, Metarhizium anisopliae var acridum. *J. Insect Physiol.* 46(4):429–37

González-Santoyo I, Córdoba-Aguilar A. 2012. Phenoloxidase: a key component of the insect immune system. *Entomol. Exp. Appl.* 142(1):1–16

Gooderham WJ, Gellatly SL, Sanschagrin F, McPhee JB, Bains M, et al. 2009. The sensor kinase PhoQ mediates virulence in Pseudomonas aeruginosa. *Microbiology*. 155(3):699–711

Gottar M, Gobert V, Matskevich AA, Reichhart JM, Wang C, et al. 2006. Dual Detection of Fungal Infections in Drosophila via Recognition of Glucans and Sensing of Virulence Factors. *Cell*. 127(7):1425–37

Grell MN, Jensen AB, Olsen PB, Eilenberg J, Lange L. 2011. Secretome of fungus-infected aphids documents high pathogen activity and weak host response. *Fungal Genet. Biol.* 48(4):343–52

Grindstaff JL, Brodie ED, Ketterson ED. 2003. Immune function across generations: Integrating mechanism and evolutionary process in maternal antibody transmission. *Proc. R. Soc. B Biol. Sci.* 270(1531):2309–19

Groisman EA. 2001. The pleiotropic two-component regulatory system PhoP-PhoQ. *J. Bacteriol.* 183(6):1835–42

Groisman EA, Kayser J, Soncini FC. 1997. Regulation of polymyxin resistance and adaptation to low-Mg2+ environments. *J. Bacteriol.* 179(22):7040–45

Gross R, Vavre F, Heddi A, Hurst GDD, Zchori-Fein E, Bourtzis K. 2009. Immunity and symbiosis. *Mol. Microbiol.* 73(5):751–59

Gunn JS, Miller SI. 1996. PhoP-PhoQ activates transcription of pmrAB, encoding a two-component regulatory system involved in Salmonella typhimurium antimicrobial peptide resistance. *J. Bacteriol.* 178(23):6857–64

Guo J, Hatt S, He K, Chen J, Francis F, Wang Z. 2017. Nine facultative endosymbionts in aphids. A review. *J. Asia. Pac. Entomol.* 20(3):794–801

Gupta V, Vasanthakrishnan RB, Siva-Jothy J, Monteith KM, Brown SP, Vale PF. 2017. The route of infection determines Wolbachia antibacterial protection in Drosophila

Hansen AK, Moran NA. 2011. Aphid genome expression reveals host-symbiont cooperation in the production of amino acids. *Proc. Natl. Acad. Sci. U. S. A.* 108:2849–54

Hansen AK, Vorburger C, Moran NA. 2012. Genomic basis of endosymbiont-conferred protection against an insect parasitoid. *Genome Res.* 22(1):106–14

Hasselquist D, Nilsson JÅ. 2009. Maternal transfer of antibodies in vertebrates: Trans-generational effects on offspring immunity. *Philos. Trans. R. Soc. B Biol. Sci.*

364(1513):51–60

Hedges LM, Brownlie JC, O'Neill SL, Johnson KN. 2008. Wolbachia and virus protection in insects. *Science (80-.).* 322(5902):702

Henry LM, Maiden MCJ, Ferrari J, Godfray HCJ. 2015. Insect life history and the evolution of bacterial mutualism. *Ecol. Lett.* 18(6):516–25

Heyworth ER. 2015. Insect facultative symbionts: phenotypic effects and competitive interactions. *PhD thesis*. University:

Heyworth ER, Ferrari J. 2015. A facultative endosymbiont in aphids can provide diverse ecological benefits. *J. Evol. Biol.* 28:1753–60

Heyworth ER, Smee MR, Ferrari J. 2020. Aphid Facultative Symbionts Aid Recovery of Their Obligate Symbiont and Their Host After Heat Stress. *Front. Ecol. Evol.* 8:56

Hilgenboecker K, Hammerstein P, Schlattmann P, Telschow A, Werren JH. 2008. How many species are infected with Wolbachia? - a statistical analysis of current data. *FEMS Microbiol Lett*. 281:215–20

Hiroki M, Tagami Y, Miura K, Kato Y. 2004. Multiple infection with Wolbachia inducing different reproductive manipulations in the butterfly Eurama hecabe. *Proc. R. Soc. B Biol. Sci.* 271(1549):1751–55

Hoffmann AA, Montgomery BL, Popovici J, Iturbe-Ormaetxe I, Johnson PH, et al. 2011. Successful establishment of Wolbachia in Aedes populations to suppress dengue transmission. *Nature*. 476(7361):454–59

Hopper KR, Kuhn KL, Lanier K, Rhoades JH, Oliver KM, et al. 2018. The defensive aphid symbiont Hamiltonella defensa affects host quality differently for Aphelinus glycinis versus Aphelinus atriplicis. *Biol. Control*. 116:3–9

Horak RD, Leonard SP, Moran NA. 2020. Symbionts shape host innate immunity in honeybees. *Proc. R. Soc. B Biol. Sci.* 287(1933):20201184

Hueck CJ. 1998. Type III Protein Secretion Systems in Bacterial Pathogens of Animals and Plants. *Microbiol. Mol. Biol. Rev.* 62(2):379–433

Hughes GL, Ren X, Ramirez JL, Sakamoto JM, Bailey JA, et al. 2011. Wolbachia infections in Anopheles gambiae cells: Transcriptomic characterization of a novel host-symbiont interaction. *PLoS Pathog.* 7(2):

Hung SY, Boucias DG. 1992. Influence of Beauveria bassiana on the cellular defense response of the beet armyworm, Spodoptera exigua. *J. Invertebr. Pathol.* 60(2):152–58

Hurst GDD, Anbutsu H, Kutsukake M, Fukatsu T. 2003. Hidden from the host: Spiroplasma bacteria infecting Drosophila do not cause an immune response, but are suppressed by ectopic immune activation. *Insect Mol. Biol.* 12(1):93–97

Hurst GDD, Jigging FM, Von Der Schulenburg JHG, Bertrand D, West SA, et al. 1999. Male-killing Wolbachia in two species of insect. *Proc. R. Soc. B Biol. Sci.* 266(1420):735–40

Hussain M, Lu G, Torres S, Edmonds JH, Kay BH, et al. 2013. Effect of Wolbachia on Replication of West Nile Virus in a Mosquito Cell Line and Adult Mosquitoes. *J. Virol.* 87(2):851–58

Igaki T, Kanda H, Yamamoto-Goto Y, Kanuka H, Kuranaga E, et al. 2002. Eiger, a TNF superfamily ligand that triggers the Drosophila JNK pathway. *EMBO J.* 21(12):3009–18

Jenke-Kodama H, Müller R, Dittmann E. 2008. Evolutionary mechanisms underlying

secondary metabolite diversity. *Nat. Compd. as Drugs Vol. I*. 119–40

Kacsoh BZ, Schlenke TA. 2012. High hemocyte load is associated with increased resistance against parasitoids in drosophila suzukii, a relative of D. melanogaster. *PLoS One*. 7(4):

Kambris Z, Blagborough AM, Pinto SB, Blagrove MSC, Godfray HCJ, et al. 2010. Wolbachia Stimulates Immune Gene Expression and Inhibits Plasmodium Development in Anopheles gambiae. *PLoS Pathog.* 6(10):e1001143

Kambris Z, Cook PE, Phuc HK, Sinkins SP. 2009. Immune activation by life-shortening wolbachia and reduced filarial competence in mosquitoes. *Science (80-.).* 326(5949):134–36

Kaneko T, Silverman N. 2005. Bacterial recognition and signalling by the Drosophila IMD pathway. *Cell. Microbiol.* 7(4):461–69

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. *Nucleic Acids Res.* 30(14):3059–66

Kliot A, Cilia M, Czosnek H, Ghanim M. 2014. Implication of the Bacterial Endosymbiont Rickettsia spp. in Interactions of the Whitefly Bemisia tabaci with Tomato yellow leaf curl virus. *J. Virol.* 88(10):5652–60

Kurtz J, Franz K. 2003. Evidence for memory in invertebrate immunity. *Nature*. 425(6953):37–38

Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, et al. 2004. Versatile and open software for comparing large genomes. . 5(2):

Kutsukake M, Shibao H, Nikoh N, Morioka M, Tamura T, et al. 2004. Venomous protease of aphid soldier for colony defense. *Proc. Natl. Acad. Sci. U. S. A.* 101(31):11338–43

Kuznetsova A, Brockhoff PB, Christensen RHB. 2017. lmerTest Package: Tests in Linear Mixed Effects Models . *J. Stat. Softw.* 82(13):1–26

Lamelas A, Pérez-Brocal V, Gómez-Valero L, Gosalbes MJ, Moya A, Latorre A. 2008. Evolution of the secondary symbiont "Candidatus serratia symbiotica" in aphid species of the subfamily lachninae. *Appl. Environ. Microbiol.* 74(13):4236–40

Laughton AM, Fan MH, Gerardo NM. 2014. The Combined Effects of Bacterial Symbionts and Aging on Life History Traits in the Pea Aphid, Acyrthosiphon pisum. *Appl. Environ. Microbiol.* 80(2):470–77

Laughton AM, Garcia JR, Altincicek B, Strand MR, Gerardo NM. 2011. Characterisation of immune responses in the pea aphid, Acyrthosiphon pisum. *J. Insect Physiol.* 57(6):830–39

Laughton AM, Garcia JR, Gerardo NM. 2016. Condition-dependent alteration of cellular immunity by secondary symbionts in the pea aphid, Acyrthosiphon pisum. *J. Insect Physiol.* 86:17–24

Lemaitre B, Hoffmann J. 2007. The Host Defense of Drosophila melanogaster . *Annu. Rev. Immunol.* 25(1):697–743

Lenth R V. 2016. Least-squares means: The R package lsmeans. *J. Stat. Softw.* 69(1):1–33

Leonard SP, Powell JE, Perutka J, Geng P, Heckmann LC, et al. 2020. Engineered symbionts activate honey bee immunity and limit pathogens. *Science (80-.).* 367(6477):573–76

Leonardo TE, Muiru GT. 2003. Facultative symbionts are associated with host plant specialization in pea aphid populations. *Proc. R. Soc. B Biol. Sci.* 270(SUPPL. 2):

Lerat E, Daubin V, Moran NA. 2003. From Gene Trees to Organismal Phylogeny in

Prokaryotes:The Case of the γ-Proteobacteria. *PLoS Biol.* 1(1):e19

Leulier F, Lemaitre B. 2008. Toll-like receptors - Taking an evolutionary approach. *Nat. Rev. Genet.* 9(3):165–78

Li D, Zhang C, Tong Z, Su D, Zhang G, et al. 2020. Transcriptome response comparison between vector and non-vector aphids after feeding on virus-infected wheat plants. *BMC Genomics*. 21(1):638

Lin Z, Cai X, Chen M, Ye L, Wu Y, et al. 2018. Virulence and Stress Responses of Shigella flexneri Regulated by PhoP/PhoQ. *Front. Microbiol.* 8:2689

Little TJ, Kraaijeveld AR. 2004. Ecological and evolutionary implications of immunological priming in invertebrates. *Trends Ecol. Evol.* 19(2):58–60

Little TJ, O'Connor B, Colegrave N, Watt K, Read AF. 2003. Maternal transfer of strain-specific immunity in an invertebrate. *Curr. Biol.* 13(6):489–92

Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2-ΔΔCT method. *Methods*. 25(4):402–8

López-García P, Eme L, Moreira D. 2017. Symbiosis in eukaryotic evolution. *J. Theor. Biol.* 434:20–33

Łukasik P, Dawid MA, Ferrari J, Godfray HCJ. 2013a. The diversity and fitness effects of infection with facultative endosymbionts in the grain aphid, Sitobion avenae. *Oecologia*. 173(3):985–96

Łukasik P, Guo H, van Asch M, Ferrari J, Godfray HCJ. 2013b. Protection against a fungal pathogen conferred by the aphid facultative endosymbionts Rickettsia and Spiroplasma is expressed in multiple host genotypes and species and is not influenced by co-infection with another symbiont. *J. Evol. Biol.* 26(12):2654–61

Łukasik P, Guo H, van Asch M, Henry LM, Godfray HCJ, Ferrari J. 2015. Horizontal transfer of facultative endosymbionts is limited by host relatedness. *Evolution (N. Y).* 69(10):2757–66

Łukasik P, van Asch M, Guo H, Ferrari J, Charles J. Godfray H. 2013c. Unrelated facultative endosymbionts protect aphids against a fungal pathogen. *Ecol. Lett.* 16:214–18

Luo H, Dearolf CR. 2001. The JAK/STAT pathway and Drosophila development. *BioEssays*. 23(12):1138–47

Ma L, Liu L, Zhao Y, Yang L, Chen C, et al. 2020. JNK pathway plays a key role in the immune system of the pea aphid and is regulated by microRNA-184. *PLOS Pathog.* 16(6):e1008627

Macnab R. 1996. Flagella and motility. *Escherichia coli Salmonella*

Mariño YA, Rodrigues JCV, Bayman P. 2017. Wolbachia affects reproduction and population dynamics of the coffee berry borer (Hypothenemus hampei): Implications for biological control. *Insects*. 8(1):

Mathers TC, Chen Y, Kaithakottil G, Legeai F, Mugford ST, et al. 2017. Rapid transcriptional plasticity of duplicated gene clusters enables a clonally reproducing aphid to colonise diverse plant species. *Genome Biol.* 18(1):27

Matsumoto H, Young GM. 2009. Translocated effectors of Yersinia

McGonigle JE, Leitão AB, Ommeslag S, Smith S, Day JP, Jiggins FM. 2017. Parallel and costly changes to cellular immunity underlie the evolution of parasitoid resistance in three Drosophila species. *PLOS Pathog.* 13(10):e1006683

McLean AHC, Godfray HCJ. 2015. Evidence for specificity in symbiont conferred protection against parasitoids. *Proc. R. Soc. B Biol. Sci.* 282(1811):1–8

McLean AHC, Parker BJ, Hrček J, Kavanagh JC, Wellham PAD, Godfray HCJ. 2018. Consequences of symbiont co-infections for insect host phenotypes. *J. Anim. Ecol.* 87(2):478–88

McLean AHC, van Asch M, Ferrari J, Godfray HCJ. 2011. Effects of bacterial secondary symbionts on host plant use in pea aphids. *Proc. R. Soc. B Biol. Sci.* 278(1706):760–66

McMeniman CJ, Lane R V., Cass BN, Fong AWC, Sidhu M, et al. 2009. Stable introduction of a life-shortening Wolbachia infection into the mosquito Aedes aegypti. *Science (80-.).* 323(5910):141–44

Meseguer AS, Manzano-Marín A, Coeur d'Acier A, Clamens A-L, Godefroid M, Jousselin E. 2017. *Buchnera* has changed flatmate but the repeated replacement of co-obligate symbionts is not associated with the ecological expansions of their aphid hosts. *Mol. Ecol.* 26(8):2363–78

Miller SI, Kukral AM, Mekalanos JJ. 1989. A two-component regulatory system (phoP phoQ) controls Salmonella typhimurium virulence. *Proc. Natl. Acad. Sci. U. S. A.* 86(13):5054–58

Milutinović B, Kurtz J. 2016. Immune memory in invertebrates. *Semin. Immunol.* 28(4):328–42

Min KT, Benzer S. 1997. Wolbachia, normally a symbiont of Drosophila, can be virulent, causing degeneration and early death. *Proc. Natl. Acad. Sci. U. S. A.* 94(20):10792–96

Montllor CB, Maxmen A, Purcell AH. 2002. Facultative bacterial endosymbionts benefit pea aphids Acyrthosiphon pisum under heat stress. *Ecol. Entomol.* 27(2):189–95

Moran NA, McCutcheon JP, Nakabachi A. 2008. Genomics and evolution of heritable bacterial symbionts. *Annu. Rev. Genet.* 42:165–90

Moran NA, Russell JA, Koga R, Fukatsu T. 2005. Evolutionary relationships of three new species of Enterobacteriaceae living as symbionts of aphids and other insects. *Appl. Environ. Microbiol.* 71(6):3302–10

Moran NA, Sloan DB. 2015. The Hologenome Concept: Helpful or Hollow? *PLOS Biol.* 13(12):e1002311

Moreira LA, Iturbe-Ormaetxe I, Jeffery JA, Lu G, Pyke AT, et al. 2009. A Wolbachia Symbiont in Aedes aegypti Limits Infection with Dengue, Chikungunya, and Plasmodium. *Cell.* 139(7):1268–78

Moret Y. 2006. "Trans-generational immune priming": Specific enhancement of the antimicrobial immune response in the mealworm beetle, Tenebrio molitor. *Proc. R. Soc. B Biol. Sci.* 273(1592):1399–1405

Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Res.* 35:W182-5

Morris JJ. 2018. What is the hologenome concept of evolution? [version 1; peer review: 2 approved]

Mousseau T, Fox C. 1998a. The adaptive significance of maternal effects. *Trends Ecol. Evol.* 13:403–7

Mousseau T, Fox C. 1998b. Maternal effects as adaptations

Nault LR. 1997. Arthropod Transmission of Plant Viruses: a New Synthesis. *Ann. Entomol. Soc. Am.* 90(5):521–41

Ng JCK, Perry KL. 2004. Transmission of plant viruses by aphid vectors. *Mol. Plant Pathol.* 5(5):505–11

Nishikori K, Morioka K, Kubo T, Morioka M. 2009. Age- and morph-dependent activation of the lysosomal system and Buchnera degradation in aphid endosymbiosis. *J. Insect Physiol.* 55(4):351–57

O'Connor SE. 2015. Engineering of Secondary Metabolism. *Annu. Rev. Genet.* 49(1):71–94

Oliver KM, Degnan PH, Burke GR, Moran NA. 2010. Facultative Symbionts in Aphids and the Horizontal Transfer of Ecologically Important Traits. *Annu. Rev. Entomol*. 55:247–66

Oliver KM, Degnan PH, Hunter MS, Moran NA. 2009. Bacteriophages Encode Factors Required for Protection in a Symbiotic Mutualism. *Science (80-.).* 325(5943):992–94

Oliver KM, Higashi CH. 2019. Variations on a protective theme: Hamiltonella defensa infections in aphids variably impact parasitoid success. *Curr. Opin. Insect Sci.* 32:1–7

Oliver KM, Moran NA, Hunter MS. 2005. Variation in resistance to parasitism in aphids is due to symbionts not host genotype. *Proc. Natl. Acad. Sci. U. S. A.* 102:12795–800

Oliver KM, Moran NA, Hunter MS. 2006. Costs and benefits of a superinfection of facultative symbionts in aphids. *Proc. R. Soc. B Biol. Sci.* 273(1591):1273–80

Oliver KM, Noge K, Huang EM, Campos JM, Becerra JX, Hunter MS. 2012. Parasitic wasp responses to symbiont-based defense in aphids. *BMC Biol.* 10(1):1–10

Oliver KM, Russell JA, Moran NA, Hunter MS. 2003. Facultative bacterial symbionts in aphids confer resistance to parasitic wasps. *Proc. Natl. Acad. Sci. U. S. A.* 100:1803–7

Olofsson TC, Vásquez A. 2008. Detection and identification of a novel lactic acid bacterial flora within the honey stomach of the honeybee Apis mellifera. *Curr. Microbiol.* 57(4):356–63

Oyston PCF, Dorrell N, Williams K, Li S-R, Green M, et al. 2000. The Response Regulator PhoP Is Important for Survival under Conditions of Macrophage-Induced Stress and Virulence in Yersinia pestis. *Infect. Immun.* 68(6):3419–25

Pacidônio EC, Caragata EP, Alves DM, Marques JT, Moreira LA. 2017. The impact of Wolbachia infection on the rate of vertical transmission of dengue virus in Brazilian Aedes aegypti. *Parasites and Vectors*. 10(1):296

Palmer T, Berks BC. 2012. The twin-arginine translocation (Tat) protein export pathway

Pancrace C, Jokela J, Sassoon N, Ganneau C, Desnos-Ollivier M, et al. 2017. Rearranged Biosynthetic Gene Cluster and Synthesis of Hassallidin E in *Planktothrix serta* PCC 8927. *ACS Chem. Biol.* 12(7):1796–1804

Parker BJ, Barribeau SM, Laughton AM, Griffin LH, Gerardo NM. 2017. Life-history strategy determines constraints on immune function. *J. Anim. Ecol.* 86(3):473–83

Parker BJ, Garcia JR, Gerardo NM. 2014. Genetic variation in resistance and fecundity tolerance in a natural host-pathogen interaction. *Evolution (N. Y).* 68(8):n/a-n/a

Parker BJ, Spragg CJ, Altincicek B, Gerardo NM. 2013. Symbiont-mediated protection against fungal pathogens in pea aphids: A role for pathogen specificity. *Appl. Environ. Microbiol.* 79(7):2455–58

Patel V, Chevignon G, Manzano-Marın A, Brandt JW, Strand MR, et al. 2019. Cultivation-assisted genome of candidatus fukatsuia symbiotica; the enigmatic "X-type" symbiont of aphids. *Genome Biol. Evol.* 11(12):3510–22

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. 2020. nlme: Linear and Nonlinear

Mixed Effects Models. *R Packag. version 3.1-149*

Pinheiro P V., Ghanim M, Alexander M, Rebelo AR, Santos RS, et al. 2017. Host plants indirectly influence plant virus transmission by altering gut cysteine protease activity of aphid vectors. *Mol. Cell. Proteomics*. 16(4):S230–43

Pisano F, Heine W, Rosenheinrich M, Schweer J, Nuss AM, Dersch P. 2014. Influence of PhoP and Intra-Species Variations on Virulence of Yersinia pseudotuberculosis during the Natural Oral Infection Route. *PLoS One*. 9(7):e103541

R Core Team. 2013. R: A Language and Environment for Statistical Computing.

Rancès E, Ye YH, Woolfit M, McGraw EA, O'Neill SL. 2012. The Relative Importance of Innate Immune Priming in Wolbachia-Mediated Dengue Interference. *PLoS Pathog*. 8(2):e1002548

Rangel LI, Henkels MD, Shaffer BT, Walker FL, Davis EW, et al. 2016. Characterization of Toxin Complex Gene Clusters and Insect Toxicity of Bacteria Representing Four Subgroups of Pseudomonas fluorescens. *PLoS One*. 11(8):e0161120

Renoz F, Noël C, Errachid A, Foray V, Hance T. 2015. Infection dynamic of symbiotic bacteria in the pea aphid Acyrthosiphon pisum gut and host immune response at the early steps in the infection process. *PLoS One*. 10(3):

Rio RVM, Symula RE, Wang J, Lohs C, Wu Y, et al. 2012. Insight into the transmission biology and species-specific functional capabilities of tsetse (Diptera: glossinidae) obligate symbiont Wigglesworthia. *MBio*. 3(1):e00240-11

Rispe C, Kutsukake M, Doublet V, Hudaverdian S, Legeai F, et al. 2008. Large gene family expansion and variable selective pressures for cathepsin B in aphids. *Mol. Biol. Evol.* 25(1):5–17

Rock DI, Smith AH, Joffe J, Albertus A, Wong N, et al. 2018. Context-dependent vertical transmission shapes strong endosymbiont community structure in the pea aphid, Acyrthosiphon pisum. *Mol. Ecol.* 27(8):2039–56

Rollat-Farnier P-A, Santos-Garcia D, Rao Q, Sagot M-F, Silva FJ, et al. 2015. Two Host Clades, Two Bacterial Arsenals: Evolution through Gene Losses in Facultative Endosymbionts. *Genome Biol. Evol.* 7(3):839–55

Rosenberg E, Zilber-Rosenberg I. 2013. *The Hologenome Concept: Human, Animal and Plant Microbiota*. Springer International Publishing

Rosenberg E, Zilber-Rosenberg I. 2018. The hologenome concept of evolution after 10 years. *Microbiome*. 6(1):78

Roth O, Beemelmanns A, Barribeau SM, Sadd BM. 2018. Recent advances in vertebrate and invertebrate transgenerational immunity in the light of ecology and evolution. *Heredity (Edinb).* 121(3):225–38

Roth O, Joop G, Eggert H, Hilbert J, Daniel J, et al. 2010. Paternally derived immune priming for offspring in the red flour beetle, Tribolium castaneum. *J. Anim. Ecol.* 79(2):403–13

Roth O, Sadd BM, Schmid-Hempel P, Kurtz J. 2009. Strain-specific priming of resistance in the red flour beetle, Tribolium castaneum. *Proc. R. Soc. B Biol. Sci.* 276(1654):145–51

Rottschaefer SM, Lazzaro BP. 2012. No Effect of Wolbachia on Resistance to Intracellular Infection by Pathogenic Bacteria in Drosophila melanogaster. *PLoS One*. 7(7):e40500

Rouchet R, Vorburger C. 2014. Experimental evolution of parasitoid infectivity on symbiont-protected hosts leads to the emergence of genotype specificity. *Evolution (N. Y).* 68(6):1607–16

Rouïl J, Jousselin E, Coeur d'acier A, Cruaud C, Manzano-Marın A. 2020. The protector within: Comparative genomics of apse phages across aphids reveals rampant recombination and diverse toxin arsenals. *Genome Biol. Evol.* 12(6):878–89

Russell JA, Moran NA. 2006. Costs and benefits of symbiont infection in aphids: variation among symbionts and across temperatures. *Proc. R. Soc. B Biol. Sci.* 273(1586):603–10

Russell JA, Weldon S, Smith AH, Kim KL, Hu Y, et al. 2013. Uncovering symbiont-driven genetic diversity across North American pea aphids. *Mol. Ecol.* 22(7):2045–59

Sadd BM, Kleinlogel Y, Schmid-Hempel R, Schmid-Hempel P. 2005. Trans-generational immune priming in a social insect. *Biol. Lett.* 1(4):386–88

Sadd BM, Schmid-Hempel P. 2006. Insect Immunity Shows Specificity in Protection upon Secondary Pathogen Exposure. *Curr. Biol.* 16(12):1206–10

Sandstrom JP, Russell JA, White JP, Moran NA. 2001. Independent origins and horizontal transfer of bacterial symbionts of aphids. *Mol. Ecol.* 10(1):217–28

Scarborough CL, Ferrari J, Godfray HCJ. 2005. Aphid protected from pathogen by endosymbiont. *Science (80-.).* 310:1781–1781

Schmitz A, Anselme C, Ravallec M, Rebuf C, Simon J-C, et al. 2012. The Cellular Immune Response of the Pea Aphid to Foreign Intrusion and Symbiotic Challenge. *PLoS One*. 7(7):e42114

Seemann T. 2014. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*. 30(14):2068–69

Shigenobu S, Watanabe H, Hattori M, Sakaki Y, Ishikawa H. 2000. Genome sequence of the endocellular bacterial symbiont of aphids Buchnera sp. APS. *Lett. to Nat.* 407(6800):81–86

Shigenobu S, Wilson ACC. 2011. Genomic revelations of a mutualism: The pea aphid and its obligate bacterial symbiont. *Cell. Mol. Life Sci.* 68(8):1297–1309

Siguier P, Gourbeyre E, Chandler M. 2014. Bacterial insertion sequences: Their genomic impact and diversity. *FEMS Microbiol. Rev.* 38(5):865–91

Smee MR, Raines SA, Ferrari J. 2021. Genetic identity and genotype × genotype interactions between symbionts outweigh species level effects in an insect microbiome. *ISME J.* 1–10

Smith AH, Łukasik P, O'Connor MP, Lee A, Mayo G, et al. 2015. Patterns, causes and consequences of defensive microbiome dynamics across multiple scales. *Mol. Ecol.* 24(5):1135–49

Spraker JE, Sanchez LM, Lowe TM, Dorrestein PC, Keller NP. 2016. Ralstonia solanacearum lipopeptide induces chlamydospore development in fungi and facilitates bacterial entry into fungal tissues. *ISME J.* 10(9):2317–30

Steiner H. 2004. Peptidoglycan recognition proteins: on and off switches for innate immunity. *Immunol. Rev.* 198(1):83–96

Stouthamer R, Breeuwer JAJ, Hurst GDD. 1999. Wolbachia Pipientis: Microbial Manipulator of Arthropod Reproduction. *Annu. Rev. Microbiol.* 53(1):71–102

Tatchell GM. 1989. An estimate of the potential economic losses to some crops due to aphids in Britain. *Crop Prot.* 8(1):25–29

Taylor MJ, Bandi C, Hoerauf A. 2005. Wolbachia.Bacterial Endosymbionts of Filarial Nematodes. *Adv. Parasitol.* 60:245–84

Teixeira L, Ferreira Á, Ashburner M. 2008. The Bacterial Symbiont Wolbachia Induces Resistance to RNA Viral Infections in Drosophila melanogaster. *PLoS Biol.* 6(12):e1000002

ter Braak B, Laughton AM, Altincicek B, Parker BJ, Gerardo NM. 2013. Exposure to Bacterial Signals Does Not Alter Pea Aphids' Survival upon a Second Challenge or Investment in Production of Winged Offspring. *PLoS One*. 8(8):e73600

Theis KR, Dheilly NM, Klassen JL, Brucker RM, Baines JF, et al. 2016. Getting the Hologenome Concept Right: an Eco-Evolutionary Framework for Hosts and Their Microbiomes. *mSystems*. 1(2):

Tsuchida T, Koga R, Fukatsu T. 2004. Host Plant Specialization Governed by Facultative Symbiont. *Science (80-.).* 303:1989

Tsuchida T, Koga R, Matsumoto S, Fukatsu T. 2011. Interspecific symbiont transfection confers a novel ecological trait to the recipient insect. *Biol. Lett.* 7(2):245–48

Tsuchida T, Koga R, Meng XY, Matsumoto T, Fukatsu T. 2005. Characterization of a Facultative Endosymbiotic Bacterium of the Pea Aphid Acyrthosiphon pisum. *Microb. Ecol.* 49(1):126–33

Uller T. 2008. Developmental plasticity and the evolution of parental effects. *Trends Ecol. Evol.* 23(8):432–38

Vallet-Gely I, Lemaitre B, Boccard F. 2008. Bacterial strategies to overcome insect defences. *Nat. Rev. Microbiol.* 6(4):302–13

Varani AM, Siguier P, Gourbeyre E, Charneau V, Chandler M. 2011. ISsaga is an ensemble of web-based methods for high throughput identification and semi-automatic annotation of insertion sequences in prokaryotic genomes. *Genome Biol.* 12(3):R30

Vigneron A, Charif D, Vincent-Monégat C, Vallier A, Gavory F, et al. 2012. Host gene response to endosymbiont and pathogen in the cereal weevil Sitophilus oryzae. *BMC Microbiol.* 12(SUPPL. 1):1–15

Vorburger C. 2018. Symbiont-conferred resistance to parasitoids in aphids – Challenges for biological control. *Biol. Control*. 116:17–26

Vorburger C, Gegenschatz SE, Ranieri G, Rodriguez P. 2008. Limited scope for maternal effects in aphid defence against parasitoids. *Ecol. Entomol.* 33(2):189–96

Vorburger C, Gouskov A. 2011. Only helpful when required: A longevity cost of harbouring defensive symbionts. *J. Evol. Biol.* 24(7):1611–17

Walker T, Johnson PH, Moreira LA, Iturbe-Ormaetxe I, Frentiu FD, et al. 2011. The wMel Wolbachia strain blocks dengue and invades caged Aedes aegypti populations. *Nature*. 476(7361):450–55

Waterfield N, Hares M, Hinchliffe S, Wren B, ffrench-Constant R. 2007. The insect toxin complex of Yersinia. *Adv. Exp. Med. Biol.* 603:247–57

Waterman SR, Holden DW. 2003. Functions and effectors of the Salmonella pathogenicity island 2 type III secretion system

Weldon SR, Russell JA, Oliver KM. 2020. More is not always better: Coinfections with defensive symbionts generate highly variable outcomes. *Appl. Environ. Microbiol.* 86(5):1–14

Weldon SR, Strand MR, Oliver KM. 2012. Phage loss and the breakdown of a defensive symbiosis in aphids. *Proc. R. Soc. B.* 280:20122103

Wernegreen JJ. 2002. Genome evolution in bacterial endosymbionts of insects. *Nat. Rev.*

Genet. 3(11):850–61

Werren JH, Baldo L, Clark ME. 2008. Wolbachia: Master manipulators of invertebrate biology. *Nat. Rev. Microbiol.* 6(10):741–51

Wickham H. 2016. ggplot2: Elegant Graphics for Data Analysis

Wilkinson TL, Fukatsu T, Ishikawa H. 2003. Transmission of symbiotic bacteria Buchnera to parthenogenetic embryos in the aphid Acyrthosiphon pisum (Hemiptera: Aphidoidea). *Arthropod Struct. Dev.* 32(2–3):241–45

Wong ZS, Hedges LM, Brownlie JC, Johnson KN. 2011. Wolbachia-Mediated Antibacterial Protection and Immune Gene Regulation in Drosophila. *PLoS One*. 6(9):e25430

Xie J, Butler S, Sanchez G, Mateos M. 2014. Male killing Spiroplasma protects Drosophila melanogaster against two parasitoid wasps. *Heredity (Edinb).* 112(4):399–408

Xu L, Ma L, Wang W, Li L, Lu Z. 2019. Phenoloxidases are required for the pea aphid's defence against bacterial and fungal infection. *Insect Mol. Biol.* 28(2):176–86

Zchori-Fein E, Gottlieb Y, Kelly SE, Brown JK, Wilson JM, et al. 2001. A newly discovered bacterium associated with parthenogenesis and a change in host selection behavior in parasitoid wasps. *Proc. Natl. Acad. Sci. U. S. A.* 98(22):12555–60

Zheng LP, Hou L, Chang AK, Yu M, Ma J, et al. 2011. Expression pattern of a Gram-negative bacteria-binding protein in early embryonic development of Artemia sinica and after bacterial challenge. *Dev. Comp. Immunol.* 35(1):35–43

Zientz E, Dandekar T, Gross R. 2004. Metabolic Interdependence of Obligate Intracellular Bacteria and Their Insect Hosts. *Microbiol. Mol. Biol. Rev.* 68(4):745–70

Zou Z, Evans JD, Lu Z, Zhao P, Williams M, et al. 2007. Comparative genomic analysis of the Tribolium immune system. *Genome Biol.* 8(8):

Zytynska S, Thighiouart K, Frago E. 2019. A meta-analysis on the benefits and costs of hosting secondary endosymbionts in sap-sucking insects. *bioRxiv*. 563031

Zytynska SE, Weisser WW. 2016. The natural occurrence of secondary bacterial symbionts in aphids. *Ecol. Entomol.* 41(1):13–26