ANALYSIS OF THE HOST-PATHOGEN PROTEOMICS OF ISRAELI ACUTE PARALYSIS VIRUS IN THE HONEY BEE USING MASS SPECTROMETRY

by

Sarah Natrasany

B.Sc., The University of Regina, 2009
B.A., The University of Regina, 2009

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

in

The Faculty of Graduate and Postdoctoral Studies

(Genome Science and Technology)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2013

Abstract

Recent declines in honey bee populations worldwide have spurred significant research into the impact of pathogens on colony health. The role of the Israeli Acute Paralysis Virus (IAPV) in hive mortality has become of particular concern since being correlated with colony losses, although the pathogenic mechanism used by IAPV remains largely unknown. To compound this problem, few molecular studies of the honey bee immune response exist. This lack of knowledge poses a significant barrier if we are to address the impact IAPV has on honey bee health. In this thesis, two routes of research were conducted to aid our understanding of the honey bee host-pathogen relationship. First, a cell culture system using honey bee hemocytes was established and optimized to address the lack of an available honey bee *in vitro* system, which has significantly delayed honey bee molecular research and especially host-pathogen studies. While hemocyte cultures were simple to create, cell division was not observed and attempts to immortalize the hemocytes through oncogene transfection were hindered by the fragility of these cells. Overall, hemocyte cultures are a useful tool for some experimental applications requiring medium numbers of cells and not involving substantial manipulation.

Next, to investigate changes in host protein expression during IAPV infection, mass spectrometry-based quantitative proteomics was used to compare IAPV infected and healthy pupae. This approach applied stable isotope dimethylation labeling combined with multidimensional fractionation using strong cation exhcange to identify and quantify ~800 proteins over three time points. Proteins that were significantly changing during infection were determined and clustered into four distinct expression patterns. To infer functional roles of proteins, *Drosophila* homologues were obtained for each protein in the data set and the corresponding GO Terms used for functional analysis. Proteins involved in processes including translation and the ubiquitin-proteasome pathway, among others, were identified and future investigation of these pathways will be useful in identifying host proteins required for infection. This analysis represents an important first step towards understanding the honey bee host response to IAPV infection through the systems-level analysis of protein expression and demonstrates the utility of mass spectrometry-based proteomics in honey bee research.

Preface

This dissertation is original, unpublished work by the author. The IAPV infected pupae analyzed for this thesis were produced by and obtained from Boncristiani, H. at The University of North Carolina, Greensboro, NC, USA. Ethics approval was not required for this research.

Table of contents

List of tables

List of figures

List of abbreviations

ABPV	Acute bee paralysis virus
AMP	antimicrobial peptides
AGO	Argonaute
Bp	base pair
CMV	cytomegalovirus immediate-early
DWV	Deformed wing virus
DCV	Drosophila C virus
ESI	electrospray ionization
FDR	false discovery rate
FBS	fetal bovine serum
GO	Gene Ontology
GFP	green fluorescent protein
HPLC	high-pressure liquid chromatography (HPLC)
IGR	intergenic region
iTRAQ	isotope tags for relative and absolute quantification
IAPV	Israeli acute paralysis virus
JAK	Janus Kinase
KBV	Kashmir bee virus
LC-MS/MS	liquid chromatography mass spectrometry
Mgf	Mascot generic file
MS	mass spectrometry
m/z	mass-to-charge ratio
MALDI	matrix-assisted laser desorption ionization
ORF	open reading frame
PTU	phenythiourea
PBS	phosphate buffered saline
PCR	polymerase chain reaction
Q	quadrupole
Rf	radio-frequency
RP	reverse phase
RT	reverse-transcription
RNAi	RNA interference
RdRP	RNA-dependent RNA polymerase
RISC	RNA-induced silencing complex
SCX	strong cation exchange
ToF	time-of-flight
UBC	University of British Columbia
UTR	untranslated region
VPg	virion protein (VPg)

Acknowledgements

It is hard to believe that three years have passed since I began my studies, and I am very grateful for the many people who have helped and supported me along the way. First, I would like to thank my supervisor Dr. Leonard Foster for the opportunity to work in the lab and for having confidence in me despite my total lack of experience in mass spectrometry when I started. Thank you for your support and for the encouragement to follow my interests during my project. I would also like to thank my supervisory committee members, Dr. Eric Jan and Dr. Paul Pavlidis for their advice and efforts.

I am especially grateful to all of my colleagues in the Foster lab, for helping me develop my ideas, listening to my frustrations, and making the lab a fun place to be. Thank you to Amanda for the endless entertainment watching you chase bee escapees around the lab and for the friendship over the years, even if you labeled the microwave incorrectly. I am inspired by your love and knowledge of honey bees and greatly enjoyed our trips to the farm (and for frozen yoghurt). To Joost, thank you isn't a big enough word for the time and patience spent teaching a rookie all the workings of mass spectrometry, the help with my project, and advice in writing this thesis. I apologize for getting you involved with the Q-ToF, but I enjoyed sharing the mass spec 'office' and finding the best place to go on a Sunday in Minneapolis. Thanks for the good times and for inspiring (against probability) an appreciation of mass spec in someone who arrived strictly a biologist.

Lastly and as always, I am grateful to my family. I would like to thank my parents for their encouragement throughout my career in science. Your support has always been much appreciated, from trying to proofread my undergrad biochemistry reports for me, to the many conversations and mail packages while here at UBC. I thank Francis, for your love, support, and perpetual troublemaking... and most importantly for being there for me and always being ready for the next adventure.

1 Introduction

1.1 Significance of honey bees

The honey bee *Apis mellifera* is highly valued for the production of hive products such as honey and wax, as a fascinating social insect with complex hive behaviors and, most significantly, for its ability to pollinate a wide variety of both agricultural and wild plant species. It has been estimated that the pollination services of honey bees are capable of increasing yield in 96% of crop species. This high utility in agriculture means the honey bee has become the predominant species employed to enhance the production of many crops including apple, alfalfa, almond and berry, among many others (Potts et al., 2010). In order to meet the high demand for pollination, honey bees are extensively managed and migratory beekeeping operations supply bees to regions unable to maintain the populations needed for pollination year round. Ultimately, the economic contribution of honey bees to agriculture is significant, and has been estimated at ~US$ 20 billion in North America (Gallai et al., 2009, VanEngelsdorp et al., 2010).

1.1.1 Honey bee health and the impact of pathogens

The value of honey bees has come into sharp focus recently due to significant declines in managed bee populations worldwide, including the US where losses of 23% were reported over 2006-2007 and increased to 36% from 2007-2010 (Cox-Foster et al., 2007, Johnson et al., 2009). At the same time, the proportion of agricultural crops depending on pollinators has increased rapidly in comparison to available hive numbers (Potts et al., 2010). In light of our dependence on honey bees, there is thus significant concern for the health and future of these pollinators. In effort to understand the declines in honey bee populations significant research into factors affecting colony survival has been made, including pesticide exposure, effects of highly managed migratory beekeeping, environmental pollution, and the spread of pathogens. In order to prevent further declines in honey bee health it will be necessary to understand the effect of each of these stressors as well as employ an integrative approach that looks at the system as a whole (Potts et al., 2010).

The impact of pathogen infections on colony mortality has become a major concern after the examination of failing hives revealed high rates of pathogen infection, often with multiple infections simultaneously present (Cox-Foster et al., 2007, Runckel, 2011). A wide variety of pathogens are known to infect honey bees, including fungal species, bacteria, parasitic mites,

and viruses. Among these, viruses are the most numerous with more than 18 different virus species known to infect honey bees (Bonning et al., 2009, Genersch et al., 2010). The majority of honey bee viruses are positive-sense, single stranded RNA viruses of the picorna-like virus superfamily. The most common of these include acute bee paralysis virus (ABPV), Israeli acute paralysis virus (IAPV), Kashmir bee virus (KBV), deformed wing virus (DWV), black cell queen virus, sacbrood virus, and chronic bee paralysis virus. Within this group, ABPV, KBV, and IAPV are three highly related members of the *Dicistroviridae* virus family which have been implicated as possible drivers of the honey bee population declines, and IAPV specifically has been correlated with colony losses in the United States (Cox-Foster et al., 2007). However, the mechanism of pathogenesis used by these viruses is poorly understood especially at a molecular level, and details of the host response to infection are also unknown.

In order to establish the impact of viral infections on honey bee health, we must first begin to understand the fundamental biology of the host-pathogen interaction. The sequencing of the honey bee genome in 2006 (Consortium, 2006) has provided an important platform for molecular honey bee research and many advances in our knowledge of honey bee pathogens have been made since the project's completion. This thesis discusses our current understanding of honey bee dicistroviruses with a focus on IAPV infection, as well as the increasing importance of using high-throughput techniques which have been made possible by the availability of the honey bee genome sequence. This background is critical to the experiments described in this thesis, which aim to advance the methods used study honey bee viral pathogens and specifically increase our understanding of IAPV infection in the honey bee.

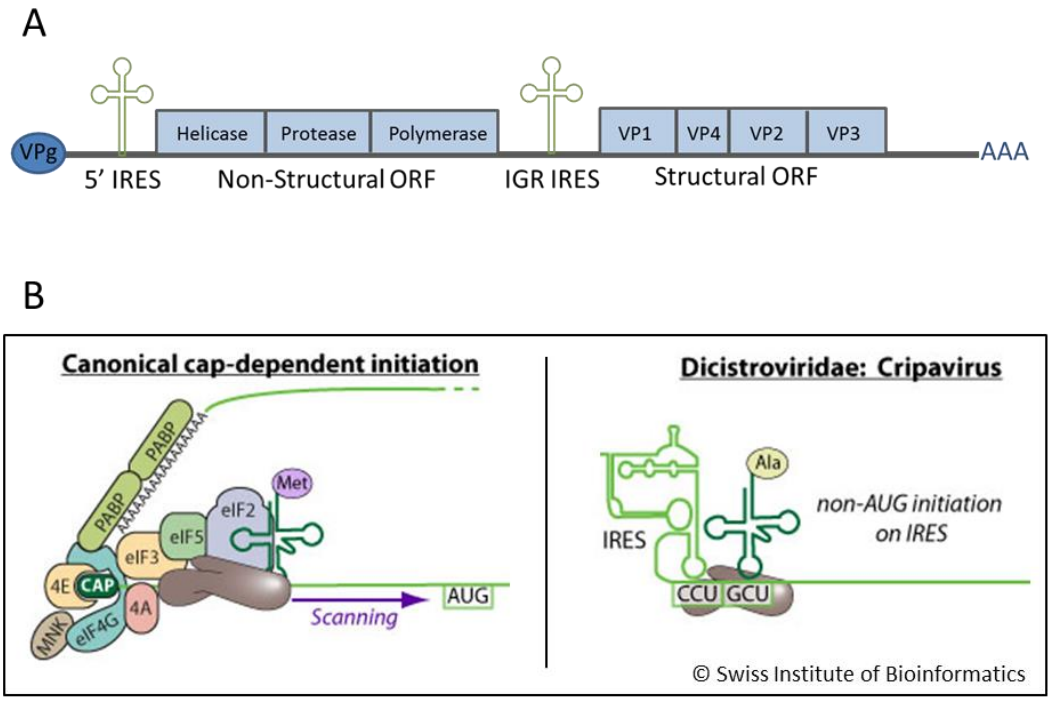
1.1.2 Biology of the honey bee *Dicistroviridae* viruses

Our current understanding of IAPV, KBV, and ABPV indicate that these highly related viruses share several similar characteristics including symptoms of infection, transmission, and genome organization, although IAPV is the least characterized of the three due to its more recent discovery following severe colony losses in Israel in 2007 (Maori et al., 2007). Like most *Dicistroviridae,* IAPV, KBV, and ABPV predominantly cause covert infections without obvious clinical symptoms, but are highly virulent when injected into adults or pupae. Lethal infections with ABPV and IAPV, although not KBV, have been characterized by a rapidly progressing paralysis of infected individuals, including symptoms of trembling, flightlessness, darkening and loss of hair from the body (de Miranda et al., 2009). Transmission of all three viruses appears to occur through both vertical and horizontal routes (de Miranda et al., 2009) and the parasitic mite

Varroa destructor is a potent vector capable of transmitting viruses with high efficiency (Bakonyi et al., 2002, Chen et al., 2007, di Prisco et al., 2011, de Miranda et al., 2009).

At the molecular level the genomes of IAPV, KBV, and ABPV are similar to those of other *Dicistroviridae*, consisting of a single-stranded RNA molecule containing two open reading frames (ORFs) separated by an intergenic region (IGR), shown in Figure 1. The 5' ORF contains non-structural genes while the structural genes are contained in the smaller 3' ORF, an organization which is characteristic of the *Dicistroviridae* family. The 5' end of the genome contains a large untranslated region (UTR) with a covalently attached 5' genome-linked virion protein (VPg), and a poly-A tail is present at the 3' end of the genome (Bonning et al., 2009, Chen et al., 2006, de Miranda et al., 2009, Nakashima et al., 2009). The two *Dicistroviridae* ORFs are initially transcribed into polyproteins, and translation proceeds from both the 5' UTR and IGR internal ribosome entry sites (IRES). The IGR IRES uniquely allows the assembly of 80S ribosomes without canonical translation initiation factors, and is thought to act as a way for the virus to avoid and disrupt traditional CAP-dependent translation in the host (de Miranda et al., 2009, Schuler et al., 2006, Wilson et al., 2000).

Figure 1 *Dicistrovirus* genome structure and translation mechanism



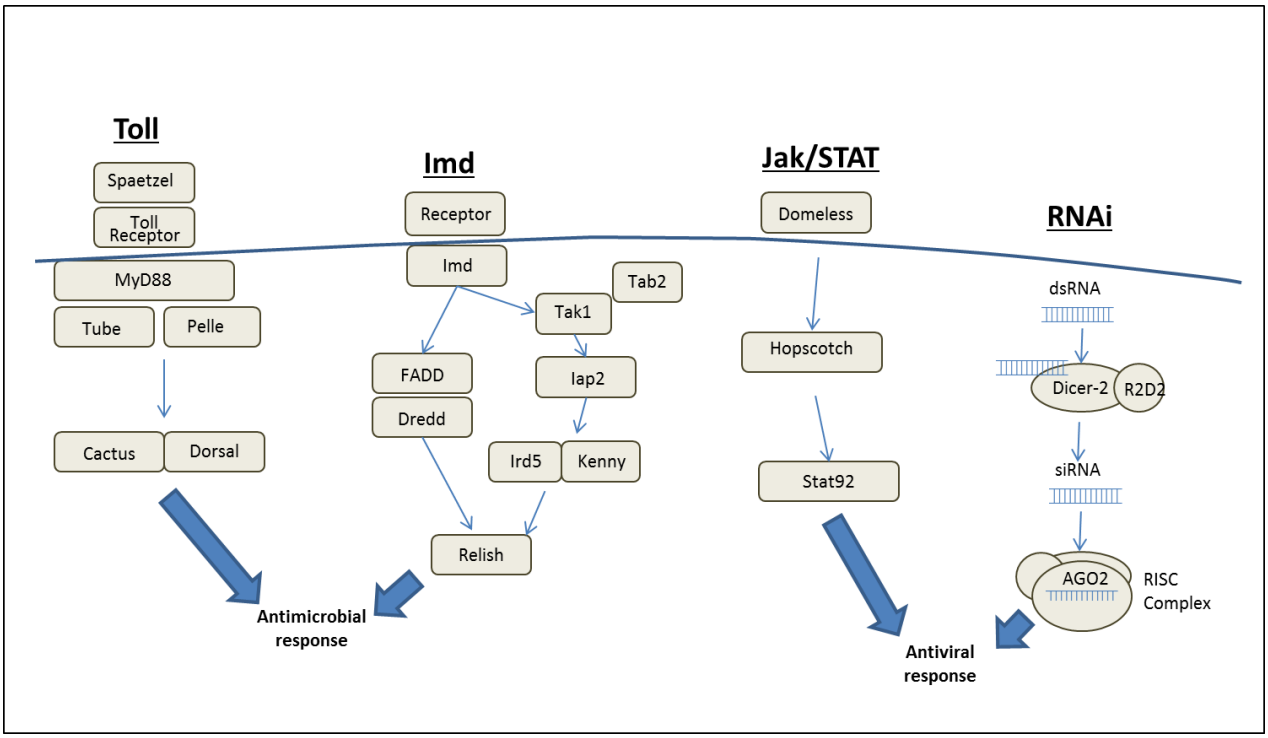
A Genomic organization of *Dicistroviridae*. *Dicistroviruses* have a ssRNA genome organized into two open reading frames (ORFs) separated by an intergenic region (IGR). Both ORFs are transcribed by internal ribosomal entry sites (IRES) present at the 5' and IGR regions of the genome. **B Cap-dependent and *Dicistrovirus* IGR IRES-mediated translation.** Canonical cap-dependent translation (left) requires the assembly of multiple factors to form an initiation complex that performs AUG scanning to begin translation. *Dicistrovirus* translation at the IGR IRES site does not require initiation factors and is initiated directly at a non-AUG start site.

1.1.3 Honey bee immune system

Like other organisms that can be infected by a multitude of pathogens, honey bees are equipped with defense mechanisms to protect against disease. However unlike higher vertebrates, insects lack adaptive immune functions and must rely solely on the innate immune response. In general, the innate defenses of insects include physical and chemical barriers, a cellular response, and a humoral response, all of which act together to defend the insect against microbial and viral pathogens (Engstrom et al., 2004, Pinheiro et al., 2006). Currently, relatively little is known about the specific genes and proteins expressed during the honey bee innate immune response. To address this lack, models for each of these immune signaling pathways have been made through the comparison of the honey bee genome with the genomes of the fruit fly *Drosophila melanogaster* and the mosquito *Anopheles gambiae* (Evans et al., 2006).

Although the number of immune-related genes in the honey bee appears to be reduced to approximately one-third of the number found in other insects, orthologues have been identified for nearly all pathway members indicating the high similarity of the antimicrobial and antiviral response of honey bees and other insects (Evans et al., 2006). An overview of the insect antiviral and antimicrobial immune response pathways is shown in Figure 2.

Figure 2 Insect immune response pathways



The insect innate immune response involves the Toll, Imd, Jak/STAT, and RNAi pathways. Bacterial and fungal pathogens primarily induce the Toll and Imd pathways, which result in the transcription of genes encoding antimicrobial peptides and other response proteins. Viral infection produces a distinct immune response through the induction of primarily Jak/STAT and RNAi pathways.

1.1.4 Antimicrobial immune response

The insect antimicrobial immune response first involves physical barriers, such as a tough outer cuticle covering the body and a hostile gut environment characterized by a low pH and the presence of reactive oxygen species. Antimicrobial peptides (AMPs), which are small peptides capable of acting directly on pathogens, are also secreted by tracheal and intestinal cells to

eradicate invading microbes. If this first line of defense is breached, the entry of microbes into the body cavity triggers an immune reaction that consists of cellular and humoral components (Vodovar et al., 2012). The cellular response involves the encapsulation and phagocytosis of pathogens by hemocyte cells. One hallmark of the humoral response is the secretion of AMPs from the fat body, which is regulated by two NFkB signaling cascades, the Toll and Imd pathways (Lemaitre et al., 2007). Both Toll and Imd are initiated through the recognition of pathogen-associated recognition patterns that are present on the invading microbe (Pinheiro et al., 2006, Vodovar et al., 2012).

The Toll pathway plays an important role in both developmental processes and immunity. In the immune response Toll-like receptors are activated by the binding of the cytokine Spaetzle in response to the presence of Gram-positive bacteria and fungi. Receptor activation causes the recruitment of intracellular death-domain containing proteins to form a receptor complex, which can then degrade the inhibitor Cactus allowing nuclear translocation of the NFκB transcription factors Dorsal and/or Dif and subsequent expression of immune response genes including AMPs (Lemaitre et al., 2007).

The Imd pathway is triggered by Gram-negative bacteria and begins with the recruitment of the adaptor protein Imd by the activated receptor (Lemaitre et al., 2007, Vodovar et al., 2012). There are two branches of the Imd pathway, both of which lead to the activation of the transcription factor Relish. In the first branch, Imd interacts with FADD to bind the caspase Dredd, which is required for the cleavage and phosphorylation of the transcription factor Relish. The second branch involves the propagation of signal through TAK1 with the adaptor TAB2 and Ird5 to promote cleavage of Relish. Activated Relish is able to translocate to the nucleus and promote the transcription of AMPs and other immune related genes (Costa et al., 2009, Lemaitre et al., 2007).

1.1.5 The antiviral immune response

The insect antiviral response is distinct from the immune reaction induced by microbial infection and primarily involves the RNA interference (RNAi) and Jak/STAT pathways. Although there is evidence that input from the Imd and Toll pathways may be required for a complete antiviral response in *Drosophila*, the details of this requirement have not yet been determined (Costa et al., 2009, Lemaitre et al., 2007).

The JAK/STAT pathway is important for developmental processes as well as the immune response, although the precise immune function of this pathway is not completely known. Jak/STAT has been found to be unnecessary for AMP expression but required for protection of *Drosophila* to infection with Drosophila C Virus, and is suspected to be a secondary response to viral infection initiated by cytokine signals produced by damaged cells (Lemaitre et al., 2007, Vodovar et al., 2012). The Jak/STAT cascade involves three main cellular components including the receptor Domeless, the Janus Kinase (JAK) Hopscotch, and the STAT92 transcription factor (Dostert et al., 2005, Evans et al., 2006, Lemaitre et al., 2007).

RNAi is the hallmark of the insect antiviral response. RNAi pathways involve enzymes of the Argonaute (AGO) family, which function to specifically cleave target RNA molecules. Three distinct RNAi pathways exist in insects, involving three classes of small RNAs that differ in size, template origin, and processing pathway: small interfering (si)RNA, micro (mi)RNA, and Piwi-associated (pi)RNA. The siRNA pathway has been implicated as the major player in the insect antiviral response (Kemp et al., 2009, van Rij et al., 2006, Vodovar et al., 2012). In this pathway, Dicer-2 senses dsRNA produced during viral replication. Dicer-2 in complex with the binding protein R2D2 processes the dsRNA into duplex siRNAs. The guide strand of the siRNA becomes incorporated into the RNA-induced silencing complex (RISC) while the other strand is degraded. The RISC then mediates sequence specific degradation of viral RNA through the action of AGO2. Mutations in AGO2, Dicer-2, or R2D2 have been found to increase susceptibility to viral infection in *Drosophila*. The presence of RNAi suppressors in the genomes of two *Dicistroviridae*, Cricket Paralysis Virus and Drosophila C Virus, further indicates the significance of this pathway in the defense against viruses. (Kemp et al., 2009, van Rij et al., 2006, Vodovar et al., 2012, Wang et al., 2006).

1.2 Tools for the molecular study of honey bees

Advancement in our understanding of the honey bee immune system and the pathogenic effects of viral infection have been greatly inhibited by a lack of basic tools to study these processes at the molecular and biochemical level. One significant obstacle has been the absence of an *in vitro* honey bee system: bees are not only difficult to rear in a laboratory setting due to their eusocial nature and lack of a defined artificial diet, but no honey bee cell culture system is currently available.

1.2.1 Cell culture

The establishment of an *in vitro* cell line begins with the isolation of cells from a tissue of interest, which once obtained are maintained in a culture dish under conditions closely resembling the *in vivo* environment of the cells (Jones, 1962). Cultures initiated directly from tissues are called primary cells and these typically survive in culture for a limited amount of time with limited rounds of cell division occurring before senescence. Cells which overcome the barrier of senescence are capable of dividing continuously, providing a constant source of cells without needing to harvest new primary cells from tissues. Such cultures are immortalized and represent a continuous cell line (Yeager et al., 1999, Lundberg et al., 2000). Immortalization may occur spontaneously and, although this is rare, cells isolated from embryos and some tumor types are more likely to spontaneously produce continually dividing isolates (Schneider, 1972). More typically, immortalization must be forced through exposure to chemical carcinogens or radiation, DNA tumor viruses, or through the transfection of oncogenes, all of which result in genetic changes that constitutively activate growth pathways or inhibit tumor-suppressor pathways (Shay et al., 1991). However, tumor cells have altered phenotypes compared to non-cancerous cells, and thus an immortalization method which does not significantly alter normal cell properties is highly desired although often not achievable (Yeager et al., 1999). Expression of the catalytic subunit of telomerase has also been shown to induce immortalization in some cell types by avoiding the senescence signal produced by progressive telomere shortening (Gil et al., 2005, Stewart et al., 2000). Other research has indicated that cells which have bypassed senescence following inactivation of tumor-suppressor pathways often undergo a second proliferative crisis due to telomere shortening, which must be addressed for continual division to occur. Over all, the route to immortalization depends on the type of cell in question, and different methods may be required to force the immortalization of different cells (Gil et al., 2005, Lundberg et al., 2000, Yeager et al., 1999).

Cell culture using both primary and immortalized cell lines has been well established as an important research tool in both vertebrate and invertebrate systems, and over 500 insect cell lines have been created from different tissues of numerous insect orders (Lynn, 2007). Insect cell culture has been used to research diverse aspects of insect cell function such as signaling mechanisms, immune pathways, gene expression, and cell migration (Smagghe et al., 2009). Cultures have also played an important role in the evaluation of host-pathogen interactions and the effect of viruses on host cell growth and survival (Hunter, 2010, Smagghe et al., 2009).

Although to date no honey bee cell lines are available, the beneficial impact the development of such an *in vitro* system would have on the state of honey bee research is clear.

1.2.2 Viral detection

The absence of a honey bee *in vitro* system also requires that the propagation of honey bee viruses is done through injection of individual bees with virus. While effective, viral propagation using individual bees can introduce other virus species due to the presence of background infections, especially since *Dicistroviridae* typically cause asymptomatic infections and mixed virus species are often encountered within a single hive. This makes it necessary to identify multiple, often non-symptomatic infections within individual bees. Historically, this has been achieved using serology-based approaches or methods for detecting surface antigens on viral particles, including immunodiffusion, ELISA, and Western blot (de Miranda et al., 2009). These methods, however, are less accurate and sensitive than polymerase chain reaction (PCR), which has become the gold standard for detecting viral genetic material from infected hosts (Stoltz et al., 1995, Boncristiani et al., 2009, de Miranda et al., 2009). For the detection of single-stranded RNA honey bee viruses, reverse-transcription (RT) PCR is used and a multitude of RT-PCR protocols are now available including techniques to detect negative strand intermediates present only during active infection. Although PCR-based techniques are highly sensitive, specific, and quantitative (Boncristiani et al., 2009) they are unfortunately able to provide only limited and primarily diagnostic insight into viral infection in honey bees.

1.3 Mass spectrometry based proteomics

The main goal of molecular biology is to reveal the inner workings of cells and characterize the processes occurring in diverse biological states. While traditional targeted techniques have characterized gene expression in many biological processes, they provide insight into only a small part of the larger set of interacting components. In order to achieve a complete view of cellular function, a systems-level approach must instead be used. Transcriptomics using microarrays to study and quantify gene expression has been an important first step in understanding transcriptional regulation of genes through their messenger levels. However, it is proteins that directly confer the function of genes, and many events occur at the translational and post-translational levels which cannot be detected by microarray. Although mRNA quantification has been used to indirectly measure changes in protein abundance, the direct

study of proteins is ideal since mRNA levels do not necessarily correspond to protein abundance (Walther et al., 2010). Mass spectrometry (MS)-based proteomics is the most comprehensive and versatile method for the large-scale study of protein dynamics, and is capable of identifying thousands of proteins from whole cell lysates (de Godoy et al., 2008). Mass spectrometry is especially useful in the field of honey bee research since MS analysis can be done directly using tissues and does not exclusively require tools generally unavailable to honey bee researchers, such as cell lines and antibodies. In addition to identification, MS-based proteomics can be used for relative or absolute protein quantification, detection of protein modifications, and protein localization, making it a powerful tool in analysis of protein expression and biological function (Yates et al., 2009, Walther et al., 2010).

Mass spectrometers essentially measure the mass-to-charge ratio (m/z) of molecules such as peptides or proteins with high accuracy. To study complex biological samples, a method termed "bottom-up" proteomics is typically employed that analyzes peptides rather than full length proteins (Yates, 2009). Here, protein mixtures are enzymatically digested into peptides by an enzyme such as trypsin, which specifically cleaves at the C-terminal side of lysine and arginine residues, and tandem MS (or MS/MS) is used to determine the sequence identity of peptides. In a data-dependent experiment, tandem MS involves two levels of MS scans, referred to as MS1 and MS2. In the first scan, MS1, the instrument acquires a precursor scan (or full scan) in which the m/z of all peptide ions are measured, each producing a peak in the mass spectrum. In a subsequent MS2 scan, a specific m/z is selected by the instrument and the corresponding precursor ion is isolated and fragmented, typically through collisions with an inert gas which is known as collision induced dissociation. The m/z of the resulting fragments are measured to produce a MS/MS spectrum. Each generated MS/MS fragmentation spectrum is searched against a protein database using a search engine such as Mascot and Sequest (Walther et al., 2010, Deutsch et al., 2008). The search engine matches each MS/MS spectrum to predicted, in-silico-generated fragmentation patterns to identify the peptide's amino acid sequence. The identified peptides are then grouped to corresponding proteins, and a confident identification is obtained when several peptides contribute to the same protein (Yates et al., 2009).

1.3.1 MS instrumentation

Mass spectrometers involve three essential parts: the ion source, which converts molecules into gas-phase ions; the mass analyzer, which separates ions according to their m/z; and the ion detector which registers the number of ions at each m/z. To analyze peptides by MS, they must

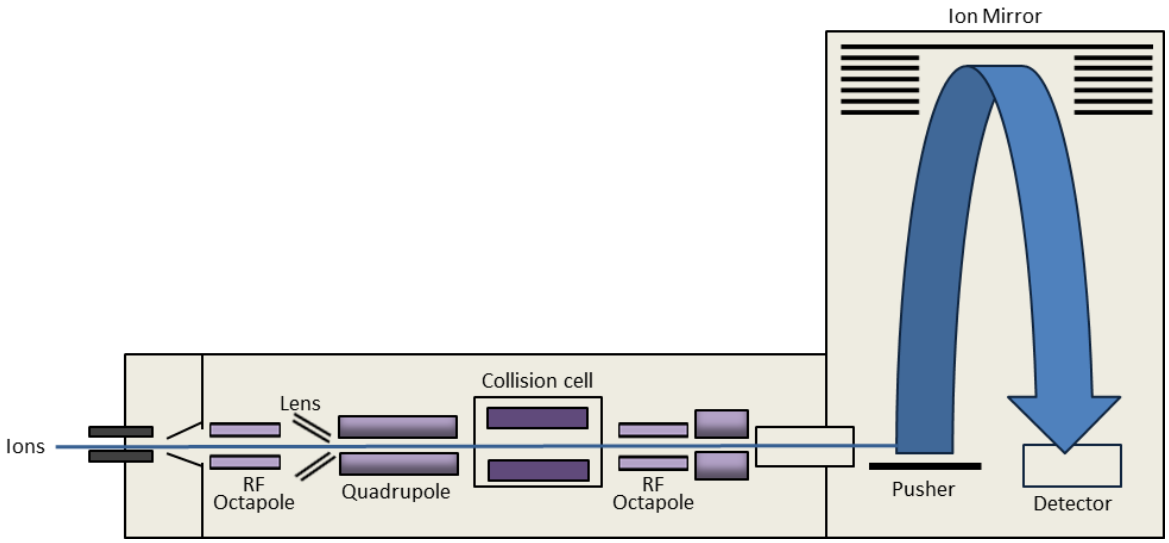
be first transferred into the gas phase as ions without degradation occurring. This is achieved mainly through two “soft” ionization techniques, matrix-assisted laser desorption ionization (MALDI), and electrospray ionization (ESI), which create ions with low internal energy that are unlikely to fragment during the ionization process. In ESI, the sample is dissolved in a solvent and infused into the mass spectrometer at very low flow rates. An electrical potential is applied to the flowing liquid to create an electrically charged spray, which generates peptide ions and in turn allows the transfer of these ions into the mass spectrometer for analysis (El-Aneed et al., 2009, Yates et al., 2009). An advantage of ESI is that it can be easily interfaced with high-pressure liquid chromatography to reduce sample complexity through chromatographic peptide separation prior to mass spectrometric analysis. The simplicity with which this can be done makes ESI the most commonly used ionization technique for the analysis of complex biological samples (Aebersold et al., 2003, Walther et al., 2010).

After ionization, the sample reaches the mass analyzer which uses physical properties such as time-of-flight (ToF) or electric or magnetic fields to separate ions based on their m/z. Each mass analyzer has unique characteristics including mass range, analysis speed, resolution, and sensitivity (Yates et al., 2009). The four main types of analyzers commonly used in proteomics are the ion trap, Fourier transform ion cyclotron, quadrupole (Q) and time of flight (ToF) analyzers. These analyzers may stand alone but are often combined in tandem to produce hybrid instruments which combine the strengths of different analyzers.

One type of hybrid mass spectrometer is the Q-ToF, which consists of a quadrupole and ToF analyzer connected in series (Figure 3). Quadrupole mass analyzers consist of four parallel electric rods with varying direct current and alternating radio-frequency (rf) potentials. The setting of the potential and frequency applied to the rods determines which ions (of a specific m/z) will pass through the quadrupole, while the remaining ions collide with the rods resulting in ejection. A quadrupole operated in rf-only mode essentially functions as an ion guide which allows all ions to pass, transferring them to other compartments of the instrument (El-Aneed et al., 2009, Guerrara et al., 2005). In the Q-ToF, the quadrupole is operated in rf-only mode for the MS1 full scan, or is operated in mass scanning mode for the MS2 scan to select ions of a specific m/z for subsequent fragmentation in the collision cell. Lastly, the intact precursor ions (MS1) or fragment ions (MS2) enter the ToF analyzer, which separates ions according to their flight time through a tube. Flight time is directly related to m/z, and larger mass ions require more time to travel the length of the flight tube than lighter mass ions of the same charge. To account for different kinetic energies among ions of the same m/z, a reflectron (or ion mirror) is

used to alter the path of ions, where ions with higher energies penetrate deeper into the reflectron before being repelled. This neutralizes differences in kinetic energies and increases resolution. This arrangement of analyzers in the Q-ToF allows ion isolation in the first analyzer while the final analyzer separates the fragment ions based on their m/z values, resulting in the identification of peptides with high sensitivity, resolution, and mass accuracy (Guerrara et al., 2005, El-Aneed et al., 2009, Yates et al., 2009).

Figure 3 Schematic view of a quadrupole time-of-flight mass spectrometer



Quadrupole time-of-flight mass (Q-ToF) mass spectrometers are hybrid instruments consisting of a quadrupole (Q) mass analyzer linked with a time-of-flight mass (ToF) analyzer. Ions enter the source and are focused by an rf-only octapole into the quadrupole, which acts either as an ion guide (during an MS1 scan) or mass filter (during an MS2 scan). During MS2 fragmentation occurs in the collision cell. Ions are focused by another rf-octapole before entering the ToF, where ions are pulsed into the flight tube. Ions are redirected by the ion mirror and finally reach the detector.

1.3.2 Fractionation of complex samples and SCX

Samples such as those produced from the digestion of whole cell lysates are often composed of hundreds of thousands of peptides. Thus, one of the main challenges in MS-based proteomics is to detect peptides within very complex mixtures. To address this, the ESI source is typically coupled to a high-performance liquid chromatography (HPLC) system so that sample separation is followed by immediate ionization and analysis as the sample elutes from the HPLC column (LC-MS/MS). Reverse phase (RP) HPLC with C18 resin, which separates compounds based on

their hydrophobicity, is the most common chromatographic material chosen for LC-MS/MS as it offers high resolution and reproducibility while being directly compatible with sample ionization by ESI.

However, often the components of a complex sample cannot be adequately resolved using only the separation performance of RP-LC due to the huge variation in peptide concentrations. This results in a dynamic range problem in which low abundant peptides are especially difficult to detect. Additional separation steps (multidimensional LC) can be used to address this by creating multiple fractions which are significantly less complex than the initial mixture prior to LC-MS/MS analysis, thus increasing the depth of proteome characterization. As RP separates compounds based on their hydrophobicity, additional separation techniques should use orthogonal properties so that the peptides are separated by two dissimilar biochemical qualities. Fractionation methods have been developed which take advantage of a variety of peptide characteristics, including peptide charge, pI, size, and polarity. All of these methods have been demonstrated similarly capable of resolving peptides and proteins; the main difference between these methods then becomes their ability to enrich for specific peptide types, such as post-translationally modified peptides, and their ease of use (Gilar et al., 2005, Mohammed et al., 2011).

Strong cation exchange (SCX) is a relatively simple and well-established technique which separates peptides based on their net positive charge. SCX is comparatively less labor-intensive and more robust than other methods, and is capable of enriching specific subclasses of modified peptides including phosphorylated and N-acetylated peptides (Gilar et al., 2005, Mohammed et al., 2011). Separation occurs such that singly charged peptides are eluted first from the SCX column, followed by doubly, triply, and higher charged peptides. The charge of a peptide is dependent on the presence of positively charged residues (arginine, lysine, and histidine) and negatively charged residues (aspartate, glutamate, and phosphorylated residues). At low pH all acidic residues are neutral, resulting in the bulk of tryptically generated peptides having a positive charge, typically of 2+ or 3+ (Mohammed et al., 2011). The power of SCX for increasing peptide identifications as well as isolating specific peptide groups such as phosphopeptides and N-acetylated peptides has been well established in several studies (Gauci et al., 2009, Mohammed et al., 2011, Peng et al., 2001).

1.3.3 Quantitative proteomics

One of the most challenging goals in molecular biology is to measure and compare the abundance of proteins between two or more physiological states. MS-based proteomics is able to achieve the comprehensive quantitation of a system through two main strategies: differential isotopic labeling, or label free quantitation. Label-free quantitation either directly compares peptide signal intensities or uses the number of acquired spectra matching a peptide to perform quantification. Quantification using this method is simple, not requiring labeling reagents or procedures, and can be applied to cell types or tissues which are difficult to label; however, the accuracy of quantification via these techniques is much lower than label-based quantification (Walther et al., 2010). Thus most quantitative proteomic methods use isotopic labeling, which enables the quantification of protein levels among multiple treatments or states within the same analysis. In this technique, peptides from two or more treatments are given isotopic labels of different masses and then combined together for analysis in a single LC-MS/MS run. Because heavy and light labeled peptides are chemically identical, the two have identical responses during MS analysis and differ only by the mass change introduced by the isotope label (Bantscheff et al., 2007, Ong et al., 2005). The ion abundances of the heavy and light peptides are measured by the mass spectrometer and compared. If the amount of one of the peptides is known absolute quantitation can be performed, otherwise ratios of the heavy/light pair are used to determine the relative quantitation of the peptides compared to each other (Boersema et al., 2008, Yates et al., 2009). Several isotopic labeling methods exist and differ mainly by the way stable isotope labels are introduced onto the peptides or proteins. These include enzymatic incorporation during digestion, metabolic incorporation directly by cells, chemical labeling, or by spiking in a labeled peptide standard (Bantscheff et al., 2007, Ong et al., 2005).

Chemical methods introduce the stable isotope label by tagging specific amino acids at the protein or peptide level. The main advantage of chemical labeling is that the methods can easily be applied to tissue samples where metabolic incorporation is difficult, although some variability is introduced due to the numerous processing steps that occur prior to labeling and sample mixing. Many chemical methods have been developed, including isotope-coded affinity tags (ICAT), isotope tags for relative and absolute quantification (iTRAQ), and dimethyl labeling. Each of these methods involves various chemistries and have different advantages and requirements (Bantscheff et al., 2007). Dimethyl labeling is an appealing method since it is very affordable, quick and simple to perform, and can be applied to any proteomics experiment including those using tissues or body fluids. The affordability of the dimethyl reagents means

that application is not limited by sample type or amount. The dimethyl labeling process uses isotopologues of formaldehyde and sodium cyanoborohydride to incorporate dimethyl labels at the α- and ε-amino groups of all proteolytic peptides. This occurs first through a reductive amination reaction of the peptide primary amines (N-terminus or ε-amino group of lysine residues) with the formaldehyde to generate a Schiff base that is then reduced by sodium cyanoborohydride to the final label form. Light, intermediate, and heavy labels are available, and the majority of labeled peptides differ in mass by 4 – 8 Da which is the minimum mass difference required to ensure adequate separation of each labeled species such that interference does not occur during LC-MS/MS analysis (Bantscheff et al., 2007, Boersema et al., 2008).

1.3.4 Analysis of quantitative data

LC-MS/MS experiments typically generate thousands of spectra, all of which must be analyzed to identify and quantify the peptides or proteins in a sample. This requires several steps, including assigning peptides to acquired spectra via a search engine, validating putative peptide matches and removing incorrect identifications, assigning quantitation ratios, and using the assigned peptides to infer protein identifications (Deutsch et al., 2008, Keller et al., 2005).

For the identification process, each generated MS/MS spectrum is searched and matched against a peptide database containing theoretical spectra for each possible peptide and assigned a score based on the quality of the match. Validation of this process is typically done through decoy searching, where reversed protein sequences are included in the database and the number of matches to the decoy peptides is used to estimate a false discovery rate (FDR). By applying cut off values, the data set can be limited to a number of false and true identifications that produces a desired FDR, typically 1%. Validated peptides can then finally be mapped to their corresponding proteins. Several programs are available to do the search and validation process, each of which use different algorithms, settings, and produce slightly different results (Keller et al., 2005, Mueller et al.,2007).

Following identification, quantitative data must be extracted. This may be done at the MS1 or MS2 level depending on the quantitation method employed, and in this respect isotopic labeling techniques fall into the categories of isobaric or isotopic. Isobaric labeling reagents such as iTRAQ produce specific fragment ions during MS/MS that are used for quantification. Thus two differently labeled peptides will have the same MS1 profile but are distinguished and quantified in MS2 upon fragmentation. Alternatively, isotopic labeling methods such as dimethyl labeling

produce peptide pairs with characteristic mass differences due to the label. The two peptides display identical elution profiles with a mass shift between them, and quantitation is performed at the MS1 level by computing the ratios between the ion chromatograms of the isotopic pairs, and is independent of MS/MS fragmentation (Mueller et al., 2007).

Most mass spectrometer vendors have developed proprietary analysis programs designed to extract, search, and quantify data generated from vendor specific instruments. Substantial heterogeneity exists in the data formats used by different mass spectrometers to output and process raw results and has made the comparison of results from different experiments or different instruments challenging (Keller et al., 2005, Pedrioli et al., 2004). To address this issue vendor independent software has been developed based on generic data formats including mzXML, mzData, and mzML. These open formats can be produced through instrument specific converters and enable the uniform analysis and exchange of MS/MS data (Keller et al., 2005, Pedrioli et al., 2004, Deutsch et al., 2008). Both open source and commercial programs are now available with different features and requirements. Some common examples include SEQUEST and Mascot for peptide searching and identification, and MaxQuant and ProteoIQ for quantification.

1.4 Project outline

In this thesis two aspects of honey bee research are examined. First, the development of a honey bee *in vitro* cell culture system is investigated in chapter two using honey bee primary hemocyte cultures. Considerations for the setup, optimization, and manipulation of hemocyte cultures are described, and through the transfection of oncogenes the immortalization of hemocyte cells is attempted with the goal of creating a continuous hemocyte cell line.

Chapter three describes the application of mass spectrometry-based proteomics to evaluate the honey bee host response to infection with the virus IAPV. A quantitative strategy to compare protein expression between healthy and IAPV infected pupae is described, including sample fractionation by SCX and the application of a triplex dimethylation labeling protocol. Factors influencing the analysis of the quantitative data are discussed and the expression profiles of proteins significantly changing during infection are generated and used to investigate the functional roles of proteins involved in the infection process.

2 Investigation into a honey bee hemocyte primary cell line and attempted immortalization

2.1 Introduction

The availability of a honey bee cell culture system would provide several benefits to honey bee research and specifically the study of host-virus interactions. Recognizing the utility of a cell culture system, many groups have attempted to create a honey bee cell line. Primary cultures have been established from several different tissues, including eggs, neural cells, and whole pupal preparations (Bergem et al., 2006, Gisder et al., 2011, Hunter, 2010, Kitagashi et al., 2011, Kreissl et al., 1992), although none of these have been successfully developed into an available, continuous cell line. One honey bee tissue which has been relatively unexplored for the purpose of cell culture is honey bee hemocytes, or blood cells. Despite the ease of obtaining large numbers of these cells, no published studies attempting culture of honey bee hemocytes currently exist.

In this study the utility of honey bee larval hemocyte primary cultures as an in vitro system is investigated. A growth medium designed to reflect the composition of larval hemolymph was optimized and hemocyte cultures established, and standard cell culture techniques such as trypsinization, cryopreservation, and transfection were attempted using primary hemocytes. Transfection was performed by two methods, lipofection and electroporation, with the aim of inducing hemocyte immortalization and continuous proliferation, as a continuous cell line is the ideal standard for cell culture work. Of the several methods available to force cells to overcome senescence and become immortalized, this was attempted through the transfection of oncogenes. Transfection of human c-myc was previously reported to immortalize embryonic bee cells (Kitagashi et al., 2011) and *Drosophila Ras85D*, the ortholog of mammalian *H-ras, K-ras,* and *N-ras* (Neuman-Spilberg et al., 1984), has been reported to induce sustained hemocyte proliferation via the Raf-MAPK pathway in flies (Asha et al., 2003). Thus honey bee *Ras85D* and human c-myc were hypothesized to be two good candidate genes to aid the proliferation and immortalization of hemocytes. The difficulty met in effectively transfecting hemocyte cells indicates that although these cultures are simple to establish and maintain, creation of a continuous cell line from primary hemocytes represents a significant challenge.

2.2 Experimental methods

2.2.1 Materials

All cell culture media and supplements were obtained from Invitrogen (Burlington, ON) and salts or chemicals of analytical grade or better were obtained from Sigma-Aldrich. The following materials were obtained as described: forceps and scissors, BioQuip (Rancho Dominguez, CA); disposable scalpels, Fisher Scientific; protease inhibitor tablets, Roche Applied Science.

2.2.2 Extraction and culture of hemocytes and hemolymph from honey bees

Honey bee larvae and adult workers were collected from hives at the University of British Columbia (UBC) farm apiary. Larvae were briefly washed in 70% ethanol and air-dried. An incision was made into the side of larvae using a sterile scalpel and forceps and hemolymph was allowed to pool from the incision for 1 to 5 s. Hemolymph was collected by pipette and transferred to an eppendorf tube containing Anticoagulant Drosophila's Ringer Solution (20 mM EDTA, 5 mM KCl, 130 mM NaCl, and 10 mM Tris base) (Scherfer et al., 2004). To harvest hemolymph from adult bees, workers were anesthetized on ice and an incision was made at the neck using scissors. Hemolymph was collected by pipette as above and transferred to a volume of Ringer's solution. Aliquots of adult or larval hemolymph in Ringer's solution were seeded into untreated six-well cell culture plates with a volume of 0.5 – 1.0 mL culture media per 9.4cm^2 well and cultures were incubated at 28^{0}C. Protease inhibitor tablets (1/40 tablet per culture) were added to adult hemolymph cultures to prevent phenoloxidase activity and melanization.

2.2.3 Optimization of culture media composition

Leibovitz's L15 Media, Grace's Insect Media, and Schneider's Drosophila Media were supplemented with 10% fetal bovine serum (FBS) and 0.1% Gentamycin. Hemolymph cultures were initiated in each media type at a range of adjusted pH values including 6.8, 7.0, and 7.2. Medias were additionally modified by adding supplements singly and in combination at the following concentrations: 35 mM proline, 20 mM glutamine, 55 mM trehalose, 3% PF-68, and 1X yeastolate. Larval hemocyte cultures were monitored visually for cell health, adherence to culture dish, duration of survival, and cell division in each adjusted media type. Schneider's Drosophila Media supplemented with 10% FBS, 0.1% Gentamycin, 35 mM proline, and an adjusted pH of 6.8 was used for all subsequent experiments and is referred to as modified Schneider's Media.

2.2.4 Cryopreservation of larval hemocytes

Approximately 100 ul of hemolymph was centrifuged at 200 rcf for 15 min immediately following extraction from larvae. Pelleted cells were resuspended in modified Schneider's Media, and a 20% DMSO solution was added dropwise to the cell suspension for a final DMSO concentration of 10%. The cell solution was transferred to a cryovial and frozen at -80°C overnight and subsequently transferred to liquid nitrogen. As a control 100 ul of hemolymph was centrifuged, resuspended with and without 10% DMSO, and directly cultured. Cryopreserved cultures were thawed at 24 h and 1 week after freezing by adding 1mL warmed media and gentle mixing prior to transfer to culture dish. Media of all cultures was refreshed once cells began to adhere to the culture dish surface to remove DMSO.

2.2.5 Trypsinization of larval hemocytes

To trypsinize adherent hemocytes, media was removed and cells were washed twice with 1X PBS. 0.5 mL of 0.25% Trypsin solution was added, and cells were incubated at 5 to 15 minutes at room temperature. Following incubation 500 ul of modified Schneider's media was added and culture dishes were sharply knocked on a hard surface to lift up cells. The cell solution was transferred to a new 9.4 cm^2 well of a six-well plate containing 500 ul fresh media and incubated.

2.2.6 Oncogene constructs and transfection of larval hemocytes

The oncogenes human c-myc (NCBI reference sequence NM_002467, gene length 1377 bp) and honey bee *Ras85D* (ras at 85 orthologue LOC410812; NCBI reference sequence XM_394288, gene length 981 bp) were inserted into the pcDNA3.1+ vector at the Hind III and Xho I cloning sites by GenScript USA Inc. Oncogene constructs and a control construct of pcDNA3.1 with GFP were transfected into larval hemocyte cells by lipofection and by electroporation. For lipofection, the Lipofectamine LTX kit was used to transfect 500ng DNA using 0.5 ul Plus Reagent and 2 ul lipofectamine in a 100ul volume. Transfection was performed on both adhered cells and cultures immediately after seeding with cells still in suspension. Media was refreshed within 12 h of transfection to remove lipofection reagents. Electroporation was done with a BioRad GenePulser Xcell electroporation system. 100 ul of trypsinized cells were resuspended in serum free media and transferred to a 0.1mm cuvette. Electroporation was tested using exponential decay settings of 500 uF, 1000 Ω, and 250-450v and square wave settings of 20 ms at 150-400v.

2.3 Results

2.3.1 Establishment of larval and adult hemocyte cultures

Hemolymph is distributed throughout the insect body and is a source of proteins, nutrients, hormones, waste products produced by various organs, and most importantly blood cells or hemocytes (Crailsheim et al., 1985). To harvest hemocytes for culturing, hemolymph was collected from honey bee larvae at approximately days 5 to 6 of development, immediately before cell capping and the prepupal period. Older larvae were preferred due to their large size and increased amounts of hemolymph (5 – 10 ul per larvae) compared to younger, smaller larvae. Cultured cells prepared from prepupal hemolymph, however, did not adhere and showed poor survival. Hemolymph from approximately 3 larvae produced hemocyte cultures of 60% confluency in a 9.4cm^2 dish and were fibroblastic in shape, with predominantly bipolar but also multipolar cells frequently observed (Figure 4A). Cells appeared to contain numerous granules, and attached strongly to the culture plate surface with adherence occurring within 2-3 hours after culture initiation. This morphology is consistent with reports of other insect hemocyte cell cultures (Mitsuhashi et al., 1985, Nakahara et al., 2010). Cultured cells did not appear to divide, although healthy cells could be maintained for up to 1 month. Although melanization has been reported to occur in cultures prepared directly using hemolymph due to phenoloxidase activity (Lynn, 2001), this was not found to affect larval hemocyte cultures. Minimal exposure of hemolymph to ambient air, combined with immediate dilution in anticoagulant Drosophila's Ringer Solution was likely sufficient to prevent melanization of cultures.

A main drawback to primary cultures produced from honey bee larvae is that fresh larvae are available only during the summer months. Adult bees may be obtained year round, however, and the possibility of deriving cells from adult bee hemolymph was next investigated. However adult hemolymph was difficult to extract and low volumes (~1 to 3 ul) were typically obtained per adult. Cultures also melanized quickly within 1 to 2 h of establishment and washing cells in Anticoagulant Ringer's solution did not prevent melanization. Addition of inhibitors did prevent melanization, but cultures with noticeable cells (Figure 4B) were obtained only when protease inhibitor tablets were used. These cells were very small, bipolar, and did not attach to the dish surface. As the larval hemocytes were the most promising of the two culture types, further investigation to determine optimal growth conditions and other cell culture applications was performed.

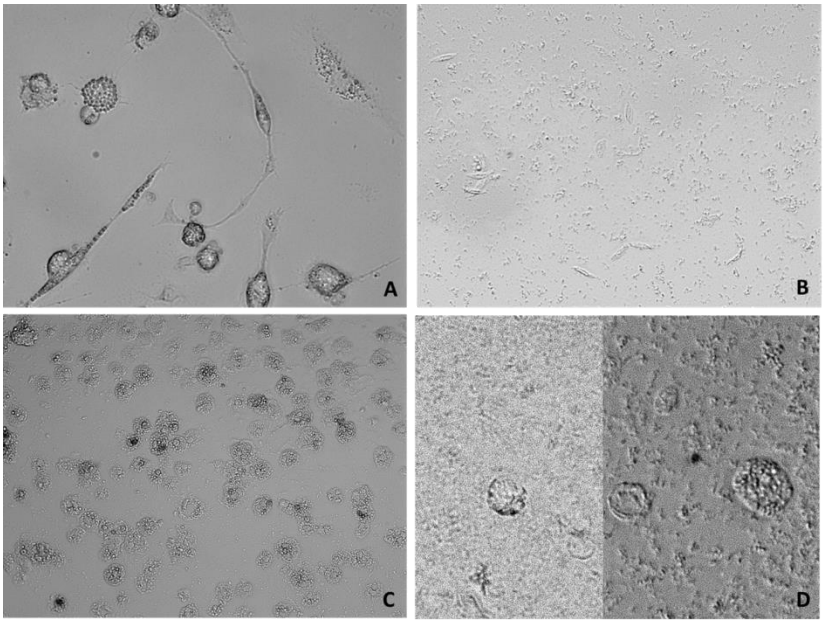
Many honey bee primary cell and tissue cultures have been created and maintained using a medium closely reflecting the composition of hemolymph (Kaatz et al., 1985, Rachinsky et al., 1998), and in general primary cell cultures thrive under conditions closely mimicking that of their in vivo conditions (Jones, 1962). These considerations make the creation of a culture medium with a similar composition to the honey bee hemolymph a logical first step in culturing hemocytes. To optimize larval hemocyte growth conditions, three types of insect culture media were explored: Leibovitz's L15 Media, Schneider's Drosophila Media, and Grace's Insect Media. Optimal cell survival was observed at pH 6.8 for all media types, which is also the reported pH of larval hemolymph (Bishop, 1923). All media were further optimized by observing the effect of supplementation with proline (final concentration of 35 mM), glutamine (25 mM), and trehalose (35 mM) to more closely reflect the composition of honey bee hemolymph (Arslan et al., 1986, Brandt et al., 1979, Wang, 1970). Pluronic F-68 and yeastolate, two common culture additives, were also tested for any benefit to the hemocyte cultures. Of the three media types, cultures appeared healthiest when grown in Schneider's Drosophila media. Cultures grown without additional supplements had a typical survival period of 2 to 3 wk, frequently with significant debris and higher numbers of visibly unhealthy cells. Supplementation with proline produced very apparently healthy cultures, with viable cells still observed after as long as 1 month. Yeastolate addition may also have shown a small positive effect on culture health although this was not consistently observed. No positive effect was observed with the addition of the other supplements, including glutamine and trehalose. Therefore, all subsequent hemocyte cultures were grown in Schneider's Drosophila Media with a pH of 6.8 and proline supplemented at a concentration of 35 mM.

Manipulation of hemocytes by basic cell culture techniques was found to be difficult. Trypsinization of cultured hemocytes resulted in high levels of cell death regardless of conditions and cells did not completely loosen from the dish surface even after a 15 min incubation period in trypsin. While cells could be knocked loose by mechanical force after approximately 5 minutes of incubation, cells often did not re-adhere when transferred to a new culture dish, and centrifugation and/or resuspension by pipetting appeared to further increase cell death. Cryopreserved cultures similarly showed poor viability, further indicating the fragility of these cells. Thawed cultures survived approximately one week, had few adherent cells and many unhealthy-looking cells in suspension with large amounts of debris.

2.3.2 Transfection of larval hemocyte cultures

The difficulties caused by the lack of cell division in hemocyte cultures and seasonal availability of larval hemolymph could potentially be overcome by forcing the immortalization of hemocytes to produce a continuously dividing cell line. One way in which forced immortalization can be accomplished is through the transfection of oncogenes into cells. As a first step towards achieving this, transfection efficiency and promoter activity in hemocytes was tested using two constructs consisting of GFP paired with the human cytomegalovirus immediate-early (CMV) promoter or the Drosophila Actin 5C distal promoter. Transfection of the constructs was attempted using both lipofection and electroporation. Following lipofection, cells remained adhered to the culture dish but appeared visibly distressed (Figure 4C) although viable cells were observed up to 1 month following transfection. Electroporation of cells resulted in significant debris and few remaining viable cells (Figure 4D). Square wave electroporation at 150v, 20ms resulted in the highest cell survival. GFP-induced fluorescent cells were not observed following lipofection regardless of whether adherent or newly seeded (suspended) cells were used. Occasional weak fluorescence was observed following electroporation, although this was difficult to distinguish from background fluorescence among cellular debris. Overall, promoter activity could not be established and successful transfection of cells was not conclusively achieved.

Figure 4 Primary hemocyte cultures and transfections



A Primary larval hemocytes. Cells adhere strongly and are characterized by the presence of granules and pseudopodia. **B Primary adult hemocytes.** Cells are small and oval shaped, and remain in suspension. **C Larval hemocytes following lipofection.** Cells remain adhered but appear unhealthy. **D Larval hemocytes following electroporation.** Cells do not adhere and are unhealthy in appearance, with significant debris.

2.4 Discussion

Hemocytes have been studied in several insect species including *Drosophila* and *Lepidoptera* (Smagghe, 2009, Nakahara, 2010, Mitsuhashi, 1985). Types of insect hemocytes that have been defined include plasmatocytes, oenocytes, granular haemocytes, and spherule cells (Ribeiro, 2006). Insect species vary in the types of hemocytes present and numbers of each cell type. One of the most predominant classes of hemocyte is the plasmatocyte. These cells are primarily involved in capsule / nodule formation and also limited phagocytosis. Characteristically, Lepidopteran plasmatocytes spread rapidly upon surface contact, develop numerous pseudopodia, and have a characteristic fibroblast-like morphology (Mitsuhashi et al., 1985, Nakahara et al., 2010). In *Drosophila* the equivalent cells, named lamellocytes, often lack pseudopodia but are flat, thin cells in monolayer (Ribeiro et al., 2006). The predominant cell type cultured from honey bee hemolymph was morphologically similar to the plasmatocytes of other insect species, indicating that these cells may be of the same type.

These cultures were simple to establish and were capable of surviving approximately one month in optimized culture media, although without visible cell division. As honey bee larvae are only available during the summer months, alternatives to increase availability of cells were investigated including cryopreservation of larval hemocytes and hemocyte cultures prepared from adults, which are present in the hive during all seasons. Cryopreserved cells had poor survival, likely due to sensitivity of cells to mechanical manipulation also observed during cell trypsinization. Adult hemocyte cultures were also significantly more difficult to culture and maintain, making adult hemocytes an inefficient route for primary culture establishment.

One or more immortalized cell lines capable of continuous division is the ideal standard for cell culture work, and a continuous honey bee line would be highly advantageous. Immortalization may occur spontaneously or be forced through application of carcinogens, UV radiation, infection with DNA tumor viruses, or introduction of oncogenes (Shay et al., 1991). Induced division and immortalization of larval honey bee hemocytes was attempted through the transfection of oncogenes, including the human c-myc gene and the honey bee *Ras85D* oncogene. Transfection of human c-myc was previously reported by Kitagashi et al., 2011 to immortalize embryonic bee cells. *Drosophila Ras85D*, the ortholog of mammalian *H-ras, Ki-ras,* and *N-ras* (Neuman-Spilberg et al., 1984), has been reported to induce sustained hemocyte proliferation via the Raf-MAPK pathway in flies (Asha et al., 2003). Transfection of each oncogene was attempted using honey bee hemocyte cells, but was unsuccessful. Inefficient transfection and possibly poor promoter activity was demonstrated through the attempted

transfection of GFP. Promoter choice is likely a critical element in successfully inducing immortalization through oncogene transfection. Mammalian promoters including CMV have been reported to function very poorly or not at all in several insect cell lines (Bourouis et al., 1983, Pfeifer et al., 1997). A promoter more likely to express well in honey bees is the strong Drosophila Actin 5C promoter, which has been successfully used for gene expression in cultured *Drosophila* cells (Chung et al., 1990). However GFP expression was still not seen following transfection of honey bee hemocytes and may have been the result of the inherent difficulty in transfecting primary cells (Gresch et al., 2004, Jordan et al., 2008). This difficulty may be compounded by the apparent fragility of the hemocytes and relative difficulty in obtaining high cell counts ideal for lipofection or electroporation.

2.5 Conclusions

Primary hemocytes are moderately simple and fast to culture, requiring minimal culture preparation. Although division does not occur, cells can survive up to one month. As such, hemocytes represent a potentially useful *in vitro* system for applications requiring medium amounts of cells and which are not dependent on mechanical manipulation of cells. Further optimization of techniques for handling hemocytes, including trypsinization and cryopreservation would be beneficial, especially the establishment of effective transfection protocols to allow transfection of oncogenes that may encourage cell division.

3 Analysis of IAPV infected honey bee pupae using mass spectrometry based proteomics

3.1 Introduction

Most experiments investigating honey bee host–pathogen relationships have focused on organism level effects, and of the few molecular studies available, those using a systems level approach are limited (Cox-Foster et al., 2007, Johnson et al., 2009, Runckel et al., 2011). While targeted approaches examining single or a select few genes or proteins have historically been highly successful in characterizing many cellular processes, these methods are time consuming and require prior knowledge regarding the target of interest (Walther et al., 2010, Yates et al., 2009). Instead, high-throughput approaches have become recognized as capable of producing unbiased and detailed insight into complex biological processes such as the host immune response to infection. While the sequencing of the honey bee genome in 2006 (Consortium, 2006) has provided a great foundation for high-throughput genomic and proteomic strategies to investigate the honey bee molecular and biochemical pathways, considerable effort is still required to optimize these methods and ensure they become routinely used at their full potential.

Viral infection can induce significant changes in host protein expression that characteristically reflect the pathogenic mechanism used by the virus to infect host cells and promote viral replication (Kellam et al., 2001). The identification of proteins which are altered during infection can provide both fundamental biological insight as well as direct the development of treatments for infection. However, it is currently unknown how host protein expression changes over the course of IAPV infection in the honey bee. To address this lack of understanding, high-throughput, mass-spectrometry based proteomics was used to investigate the honey bee immune response and pathogenic effects of IAPV infection. This approach allowed infection induced changes in honey bee protein expression to be identified and quantified at the systems-level, representing the first quantitative proteomic evaluation of IAPV infection in honey bees. Several proteins differentially regulated during infection were identified, and this information will be useful in directing further investigation of the biochemical pathways and processes important to IAPV infection. Overall, this analysis is an important first step towards increasing our fundamental understanding of the molecular function of the honey bee immune system and the pathogenic strategy used by IAPV.

3.2 Experimental methods

3.2.1 Materials

All salts and chemicals were of analytical grade or better and were obtained from Sigma-Aldrich unless otherwise indicated. HPLC grade solvents were obtained from ThermoFisher Scientific. The following materials were obtained as follows: endopeptidase Lys-C, Wako Chemicals (Osaka, Japan); porcine modified trypsin, Promega (Nepean, Ontario, Canada); Taq polymerase and SuperScript III enzymes, Invitrogen (Burlington, Ontario, Canada); PolySULFOETHYL A column 50x1.0 mm (PolyLC inc.); Large Capacity Chip (II), Agilent Technologies.

3.2.2 IAPV propagation and diagnostic RT-PCR

Tissues of IAPV infected bees were homogenized for 1 min in 100 ul PBS by mechanical disruption with ceramic beads using a MP FastPrep-24 homogenizer (MP Biomedicals). Homogenate was purified by centrifugation at 14,100 rcf for 10 min and the supernatant containing the viral particles was harvested. For viral propagation, 3 to 4 ul of lysate was injected into the abdominal cavity of healthy pupae using a Hamilton syringe, resulting in significant IAPV infection. Pupae were harvested approximately 3 d after injection to produce new viral stocks.

Presence of IAPV infection and absence of background viruses was confirmed through RT-PCR. Details about the primers used for the specific detection of each virus and the honey bee actin gene are shown in Table 1. Briefly, total RNA was extracted from infected bees using Trizol reagent as described by the manufacturer and quantified using a NanoDrop 1000 Spectrophotometer. Approximately 500 ng RNA was amplified with random hexamer primers and the reverse transcription enzyme SuperScript III as indicated by the manufacturer. Following reverse transcription, 1 ul of the resulting template was mixed with primer pairs corresponding to IAPV, KBV, ABPV, DWV, or honey bee actin and amplified by Taq polymerase enzyme using a MyCycler Thermal Cycler (BioRad) and the following program: 94^0C, 2 min; 30 cycles of 94^0 for 1 min, 54^0 for 1min, and 72^0 for 1min; 72^0C for 7 min. PCR products were visualized on a 1% agarose gel exposed under UV light.

Table 1 Primers and gene product size used for diagnostic multiplex RT-PCR.

Virus	Primer Pair	Product length (bp)	Source
DWV	5'-ATCAGCGCTTAGTGGAGGAA-3' 5'-TCGACAATTTTCGACATCA-3'	702	Chen et al., 2006
IAPV	5'-AGACACCAATCACGGACCTCAC-3' 5'-AGATTTGTCTGTCTCCCAGTGCACT-3'	475	Maori et al., 2007
KBV	5'-GTTTCTATGCAAATCGCA-3' 5'-CCATCCAGGCACATTCTG-3'	282	Todd et al., 2007
ABPV	5'-GGAACATGGAAGCATTATTG-3' 5'-AATGTCTTCTGAACCATAG-3'	687	Bakonyi et al., 2002
Actin	5'-TGCCAACACTGTCCTTTCTG-3' 5'-AGAATTGACCCAATCCA-3'	156	Zufelato et al., 2004

3.2.3 Protein solubilization and triplex dimethylation labeling

IAPV infected and control honey bee pupae were prepared at the University of North Carolina, Greensboro, as follows: white-eyed pupae (approximately day 12 to 13 of development) were collected from two healthy hives of different genetic backgrounds. Pupae were injected with a high dose of IAPV (500,000 viral genomes / bee), a low dose of IAPV (10,000 genomes / bee) or PBS as control. Six pupae from each treatment group were harvested at 0, 12, and 36 h post injection and divided into three replicates for each time point of each hive, with the exception that only two replicates were made for the 0 h time point of hive two. Heads from between one and three pupae from each treatment were used per replicate and combined for subsequent labeling and analysis.

For proteomic analysis using triplex dimethylation labeling, pupae were dissected in PBS containing Roche protease inhibitor tablets (1 tablet / 50 mL PBS) to obtain head tissues. Tissues were homogenized in lysis buffer (50 mM Tris-Cl, 150 mM NaCl, 1% NP-40) by mechanical disruption with ceramic beads using a MP FastPrep-24 (MP Biomedicals). Clarified supernatant was separated from debris following centrifugation at 10,000 rcf for 5 min and proteins were precipitated using the ethanol/acetate method (Foster, 2003). Protein samples were finally resuspended and solubilized in 6 M urea, 2 M thiourea, 100 mM Tris-Cl, 20 mM DTT. Protein concentrations of the samples were measured by Bradford assay. Proteins were alkylated using iodoacetamide at a 1:10 iodoacetamide:protein ratio and the in solution digestion of proteins was performed using Lys-C (1:50 ratio of enzyme:protein) followed by Trypsin (1:50 ratio of enzyme:protein) as described in Foster, 2003. Digested peptides were

purified and concentrated on C18 STAGE-tips (Rappsilber et al., 2003) and eluted in 80% acetonitrile, 0.5% acetic acid, and dried in a vacuum concentrator (Eppendorf). Dried peptides were resuspended in 100 mM triethylammonium bicarbonate and chemical dimethylation labeling was performed using light (CH_2O), medium (CD_2O), and heavy ($^{13}CD_2O$) isotopologues of formaldehyde as described in Boersema, 2009. Labelled samples were finally combined together in equal amounts for STAGE-tip purification and subsequent elution in 80% acetonitrile, 0.5 % acetic acid. Eluted samples were dried and resuspended in 20% acetonitrile and 0.1% formic acid for subsequent fractionation by strong cation exchange.

3.2.4 Strong cation exchange fractionation

HPLC-SCX fractionation using an Agilent 1100 HPLC system was performed by loading 300 to 600 ug of combined labeled sample onto a PolySULFOETHYL A column 50x1.0 mm (PolyLC inc.). Samples were eluted from the column over a 60 minute gradient: 0 to 7 min 100% Buffer A (20% acetonitrile and 0.1% formic acid); 7-15 min up to 13% Buffer B (20% acetonitrile, 0.1% formic acid, 0.5 M KCl); and 15-25 min up to 25 % B. At 27 min, the concentration of B was 100% which was held for 5 min. The flow rate applied during the gradient was 50 ul / min and fractions were collected at 1 min intervals for 40 min. Fractions were pooled according to their intensity so that each final fraction had a similar concentration of peptides. Fractions were dried and resuspended in 0.5% acetic acid for LC-MS/MS analysis.

3.2.5 Liquid chromatography tandem mass spectrometry

The analysis of each fraction was performed using an Agilent 6550 Q-ToF mass spectrometer equipped with a HPLC-Chip Cube system (Agilent Technologies). Analysis was done using the Large Capacity Chip (II) which has a 150 mm x 75 mm separation column and 160 nL enrichment column (G4240-62010). The trapping of peptides was done at 2 ul/min using Buffer A (3% acetonitrile, 0.1% formic acid) and analysis was performed at a flow rate of 300 nL/min using a gradient of 3% to 40% Buffer B (90% acetonitrile, 0.1% formic acid) over 120 minutes. The Q-ToF mass spectrometer was operated in positive ion mode to acquire spectra in the range of 300 – 1700 m/z to an accumulation target value of 25000. Up to 20 precursors with intensities above a threshold of 1000 counts were selected for collision-induced fragmentation.

3.2.6 Data analysis

Raw acquisition files were converted to Mascot generic format using the Agilent MassHunter Qualitative Analysis software (Version 5.0) and searched using the Mascot search engine (Matrix Science, London, UK) against a honey bee and virus database containing reversed decoy sequences. A peptide mass tolerance of 10 ppm and fragment mass tolerance of 0.1 Da were chosen and Trypsin selected as the proteolytic enzyme, allowing one missed cleavage. mzXML files were generated from raw data files using the conversion program Trapper. mzXML and Mascot search results were uploaded to the program ProteoIQ (version 2.3.08, Premier Biosoft) to perform relative quantitation. For each replicate, quantitated data was filtered to a 1% FDR using a Mascot ion score cut off. Proteins identified with <1 peptide or quantified with <2 peptides were removed and proteins with quantitation standard deviations above 1 were manually verified and corrected. Proteins quantified at all three time points were analyzed using a one-way ANOVA test to identify proteins which significantly changed in expression during infection. The ANOVA test was calculated by the program Perseus using a threshold of 0.05 and permutation-based FDR for multiple-testing correction. Of the 323 proteins quantified at all three time points, the ANOVA test identified 71 proteins that were significantly changing. These proteins were clustered according to their expression profiles by a fuzzy K-means method using the language R for statistical computing (version 2.15.2).

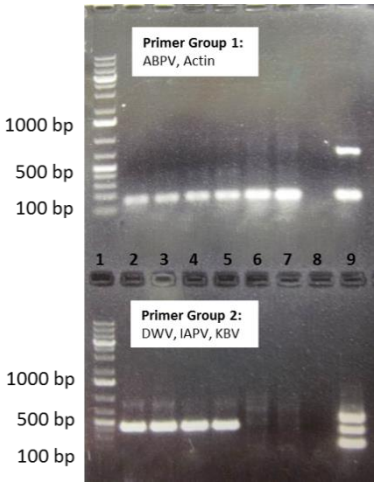
3.3 Results

3.3.1 Detection and propagation of IAPV monoinfection in honey bee pupae

Honey bees may be infected with multiple viruses without showing any observable symptoms, making the specific detection of viral infections essential to the study of any honey bee virus. To confirm the presence of IAPV and absence of other potentially contaminating viruses in the pupae experimentally infected in this study, reverse transcription polymerase chain reaction (RT-PCR) was used to detect genes unique to each of four different viruses: IAPV, KBV, ABPV, and DWV. To significantly reduce the time required to perform these diagnostic PCR reactions, primers were combined into two groups such that all fragments could be sufficiently separated by agarose gel electrophoresis: 1) Actin and ABPV; 2) IAPV, KBV, and DWV. The specific detection of viral RNA from singly and multiply infected samples is shown in Figure 5. Pupae experimentally injected with IAPV produced only two bands (lanes 2 to 5) corresponding to the

positive control actin and IAPV. PBS injected pupae (lanes 6 to 7) only showed signal consistent with actin and the negative control (lane 8) did not show any signal at all. Bands corresponding to all 5 genes were observed for pupae that were infected with all 4 viruses (lane 9). These results indicate that IAPV injected pupae used in this study were free from coinfection with DWV, KBV, or ABPV.

Figure 5 Multiplexed RT-PCR of honey bee pupae.



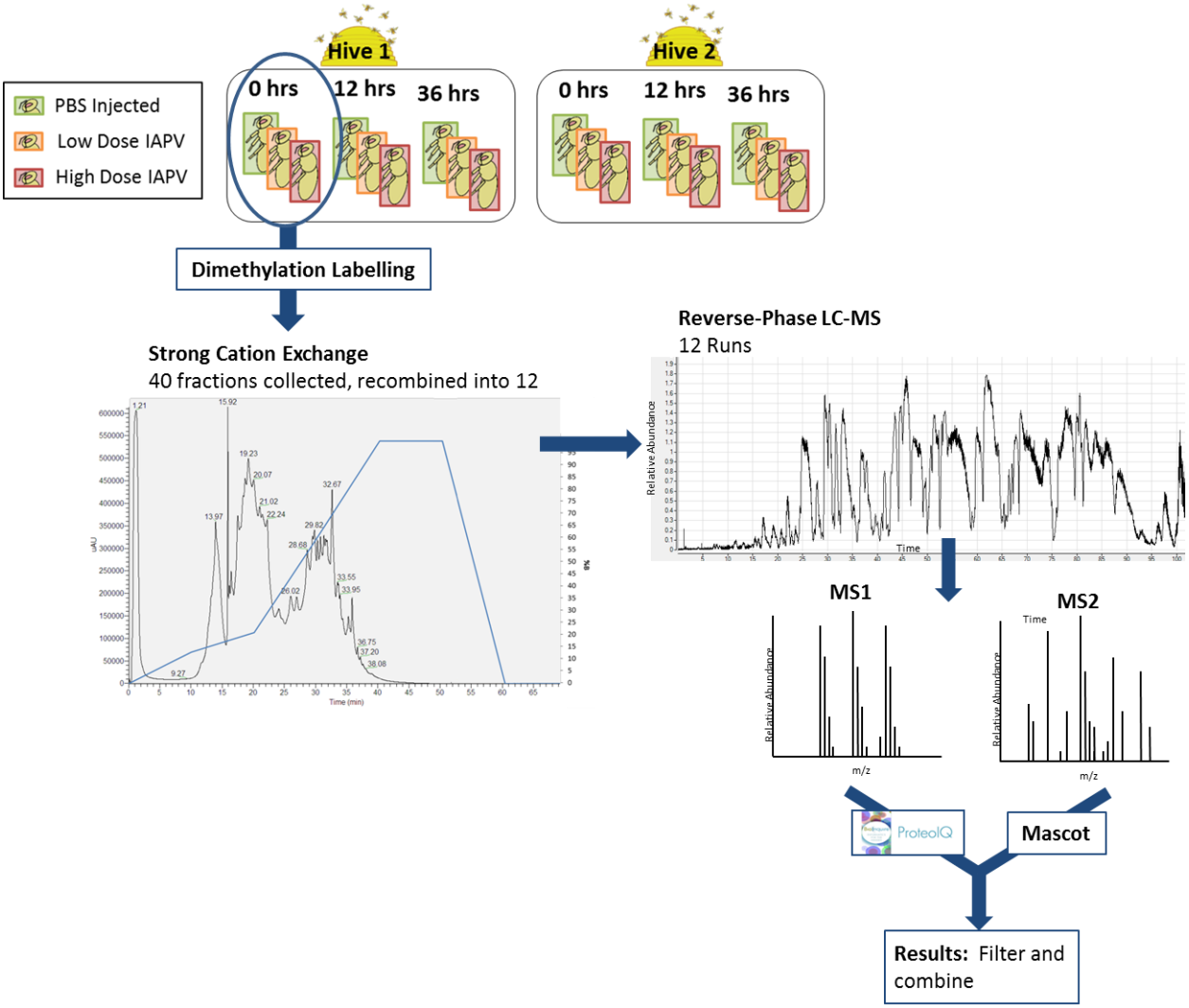
Ladder (lane 1), IAPV injected pupae (lanes 2 to 5), PBS injected pupae (lanes 6 to 7), negative control using water as template (lane 8) and pupae of mixed infection as positive control (lane 9). Note that a ~156bp band corresponding to actin is present for all honey bee samples but not for the negative control.

3.3.2 Strategy for quantitative proteomics of IAPV infected honey bee pupae

To investigate the effects of IAPV infection in honey bees at the protein level, wild-type and visibly healthy pupae from two different hives were infected with IAPV (low or high dose) or mock infected with PBS. Pupae were harvested at 0, 12, and 36 h post injection and divided into three replicates per time point for each hive. Each replicate was processed and analyzed using the strategy for quantitative proteomic analysis described in Figure 6. As a paralytic virus, IAPV is expected to significantly affect the brain, and in fact preliminary tests indicated IAPV proteins accumulate predominantly in the honey bee head. Therefore in this strategy heads were harvested from pupae for subsequent proteomic analysis. Proteins were extracted from pooled head tissues from low dose IAPV, high dose IAPV, or PBS injected pupae. Proteins were digested into peptides using trypsin and labeled as “heavy” (high dose IAPV), “medium” (low dose IAPV) or “light” (PBS) by dimethylation labeling. Labeled samples were combined in equal amounts, leading to 17 differentially labeled samples, and each sample was subjected to SCX separation with 40 fractions collected at one minute per fraction. Several fractions were

combined due to low amounts of peptide resulting in a final 12 fractions per sample. Each fraction was subsequently analyzed by LC-MS/MS over a two-hour gradient. Figure 6 shows typical chromatograms for both SCX and RP LC-MS separation.

Figure 6 Strategy for quantitative proteomic analysis of IAPV infected pupae.



Proteins from IAPV (red and orange boxes) and PBS (green box) injected pupae were extracted, digested, and labeled in parallel. Labeled samples were combined and fractionated by SCX into 12 fractions that were analyzed by LC-MS/MS. The resulting raw data was processed for peptide identifications (MS2) and quantification (MS1) to establish relative ratios for peptides.

3.3.3 Protein identification and quantification

Raw data from each LC-MS analysis was searched against a combined honey bee / virus protein database using the Mascot search engine. The database contained “Forward” and “Reverse” sequences to provide an estimate of the false discovery rate (FDR) of identification. Using this, results were filtered to a 1% FDR across all fractions of each replicate. In addition, proteins identified by less than two peptides were removed from each replicate. These criteria ensured a high quality dataset.

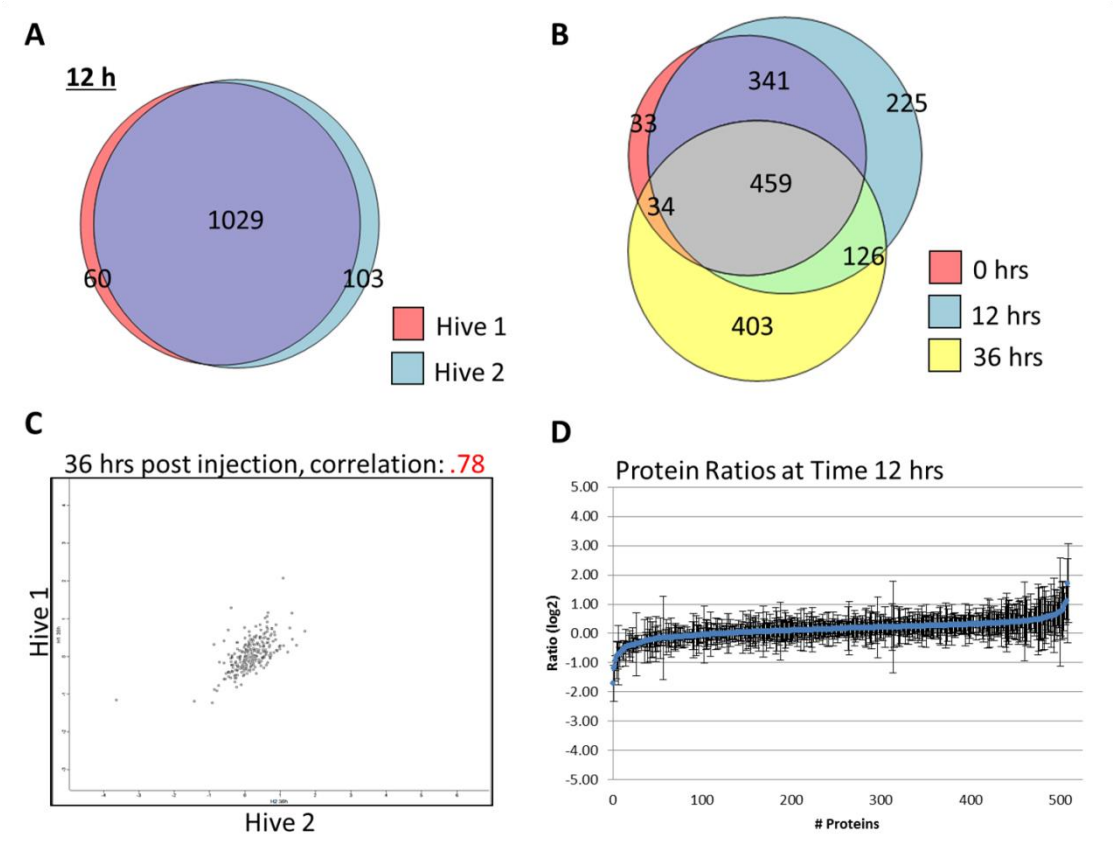
Next, the relative quantification of proteins was performed for each replicate using the program ProteoIQ. Relative expression levels were obtained by calculating the area under the extracted ion chromatogram of each species. Ratios of all quantified peptides belonging to a protein were then averaged to produce a final ratio reflecting relative protein abundance in IAPV infected versus uninfected bees. Within each replicate, only proteins with two or more quantifiable peptides were included and proteins with standard deviations greater than 1.0 were manually inspected for erroneous quantitation. These proteins often included peptides with extreme ratios incorrectly assigned by the program algorithm. Examples of problematic quantitation typically include low abundant peptides with intensities near the noise level or quantitation channels that were contaminated with other co-eluting species. In both cases it is difficult, if not impossible, for the software algorithm to correctly estimate the peptide abundance leading to high variation between peptide ratios that belong to the same protein. Manual curation of these suspicious peptides ensured good quality of quantitative data.

A preliminary analysis of the quantitative data was done by using scatter plots to compare protein expression between low dose and high dose IAPV injected pupae. Correlation between the two doses was determined for each replicate of both hives by Spearman rank. While lower correlation was seen at the 0 h time points of both hives compared to the later time points, on average the spearman rank value was >0.6 indicating that the high dose and low dose treatments produced highly similar host responses and could be considered replicates. This increased the total number of replicates to six per time point of each hive.

Next, differences in both protein identification and quantitation were compared between the two hives. Figures 7A shows the total number of protein identifications overlapping between hive 1 and hive 2 for the 12h time point (replicates combined). Interestingly, it can be seen that the majority of identified proteins are identical between the hives. Although both hives were expected to have similar responses to infection, such high overlap in identifications is unusual

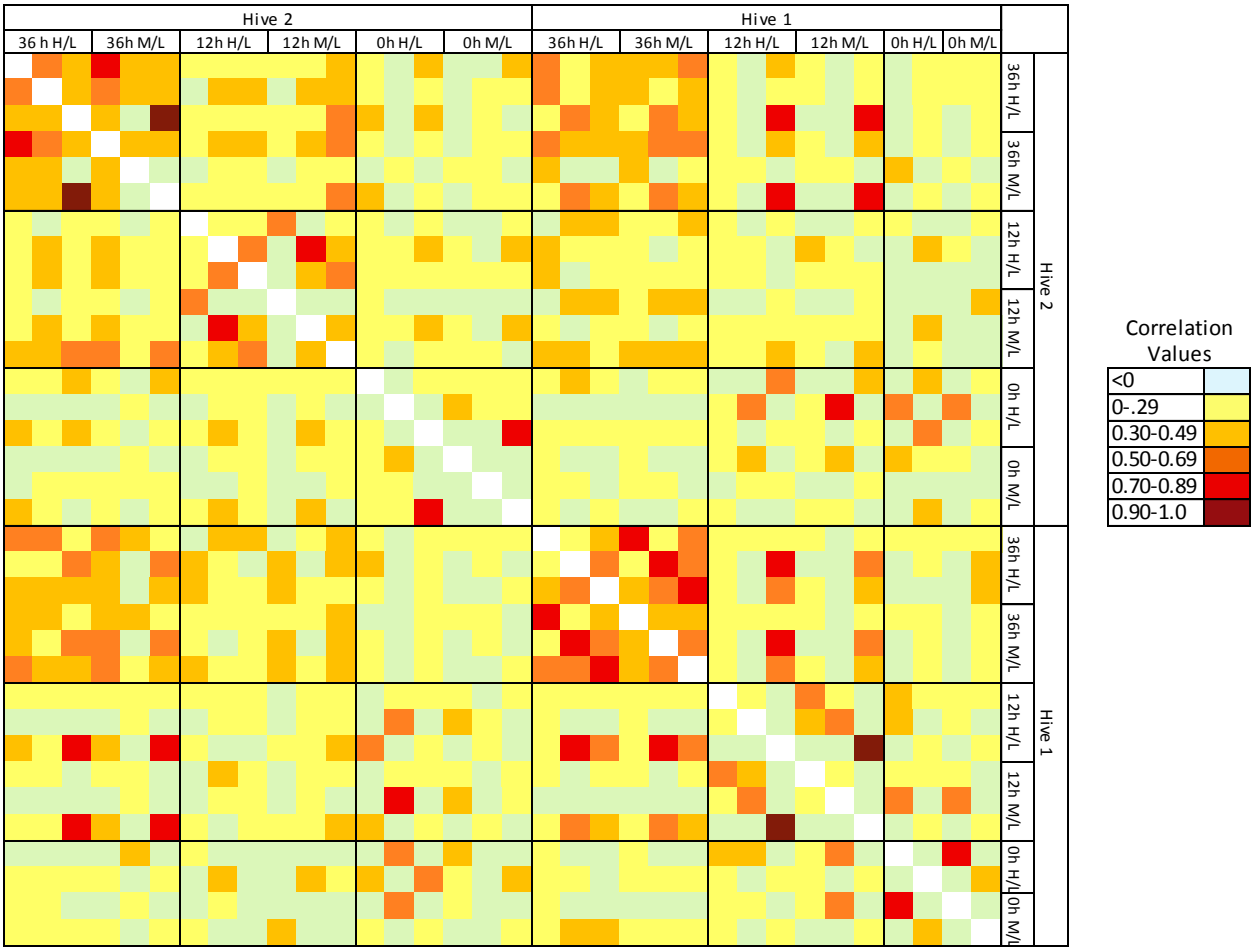
because large variation typically exists even between replicates. Protein ratios were also compared between hives by scatter plot for each time point and correlation was calculated by spearman rank. Figure 7C shows the scatter plot for the 36 hour time point and a spearman rank value of 0.78 indicates good correlation for this time point. Reasonable correlation was found for all time points between the hives, allowing data from both hives to be combined for a final total of twelve replicates (six from each hive) per time point. All correlation coefficients are displayed as a heatmap in Figure 8.

Figure 7 Protein identification and quantification.



A Overlap of identified proteins between Hive 1 and Hive 2 at 12 h. Identified proteins were highly similar between Hives 1 and 2 at all time points. **B Identification of peptides across all time points.** Overlap in protein identification across time points is low. **C Correlation between protein ratios of Hive 1 and Hive 2.** High correlation (spearman rank coefficient 0.78) indicates the protein ratios between hives are in good agreement and that the two data sets can be reasonably combined. **D Log2 quantitative ratios for host proteins identified 12 hours post infection with IAPV.** Expression patterns of protein log2 ratios 12 hours post injection. Error bars represent the standard deviation of the log2 ratio. Similar patterns were seen with 0 and 36 hour time points.

Figure 8 Heatmap of correlation values.



Spearman rank correlation values comparing all samples including replicates, low dose IAPV, high dose IAPV, hive 1 and hive 2 are displayed as a heatmap.

Lastly, proteins from the twelve replicates of each time point were combined to create a final set of proteins identified at each stage of infection. Proteins identified in only one replicate were discarded. This resulted in a very comparable number of identified proteins per time point, ranging from ~900 to ~1200 proteins (Table 2). An example of the overlap in identified proteins between the three time points is shown in Figure 7B. Around 500 proteins were identified in all three time points which is typical of the analysis of complex samples and is partly due to the large dynamic range of different protein concentrations within cells. Next the quantitative data from all replicates was averaged together, and relative expression levels were obtained for close to 80% of all identified proteins per time point (Table 2). The expression pattern of proteins at 12 h post infection is shown in Figure 7D and is representative of protein expression

at the other time points. Proteins with ratios above a log2 value of 0 indicate proteins that were found at higher levels in IAPV infected pupae compared to controls, while negative ratios indicate proteins which had decreased abundance in infected pupae compared to controls.

3.3.4 Identification of IAPV viral proteins

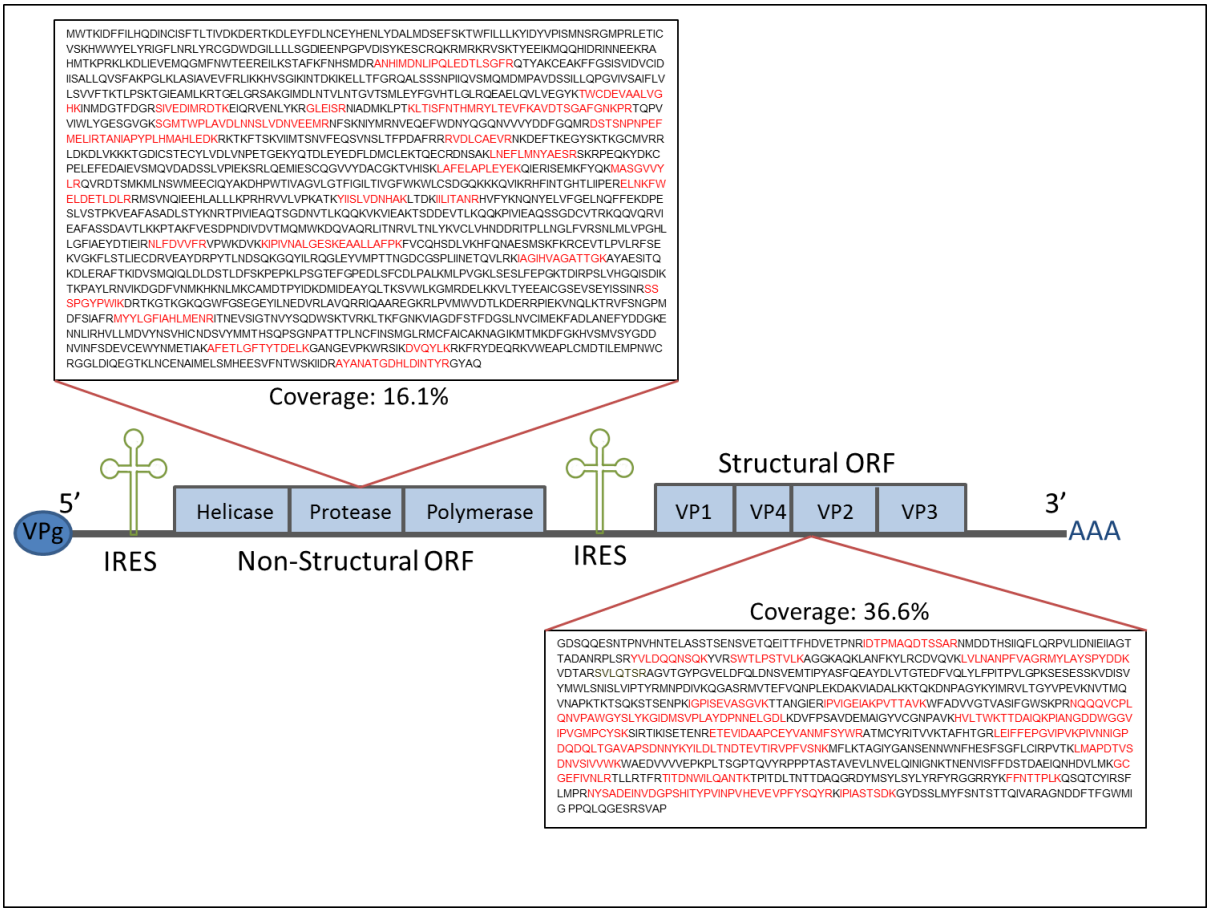
Viral peptides corresponding to IAPV ORF 1 and 2 polyproteins were identified at 12 and 36 h post injection, but not at 0 h, indicating that in addition to RT-PCR a proteomics approach can also be used as a diagnostic tool. These peptides corresponded specifically to a USA strain of IAPV designated by Genbank accession EU218534. This IAPV strain was documented as found in Western honey bees first collected in the USA in 2003 (Chen, Y., Unpublished). Viral peptides were identified from all regions of both the structural and non-structural polyproteins. Figure 9 shows the viral genome structure and complete polyprotein sequences, with peptides identified by LC-MS/MS shown in red. Percentages of total sequence identified corresponded to a coverage of 16.1% of the non-structural polyprotein and 36.6% of the structural polyprotein.

Table 2 Total number of identified and quantified proteins at each IAPV infection time point.

	0hrs	12hrs	36hrs
Identified	892	1180	1022
Quantified	700	924	805

Identified proteins were detected in two or more replicates. Quantified proteins were defined as having two or more unique quantified peptides in two or more replicates. Totals include proteins identified and/or quantified at one or more time point.

Figure 9 Detection of polyproteins from USA IAPV strain EU218534.



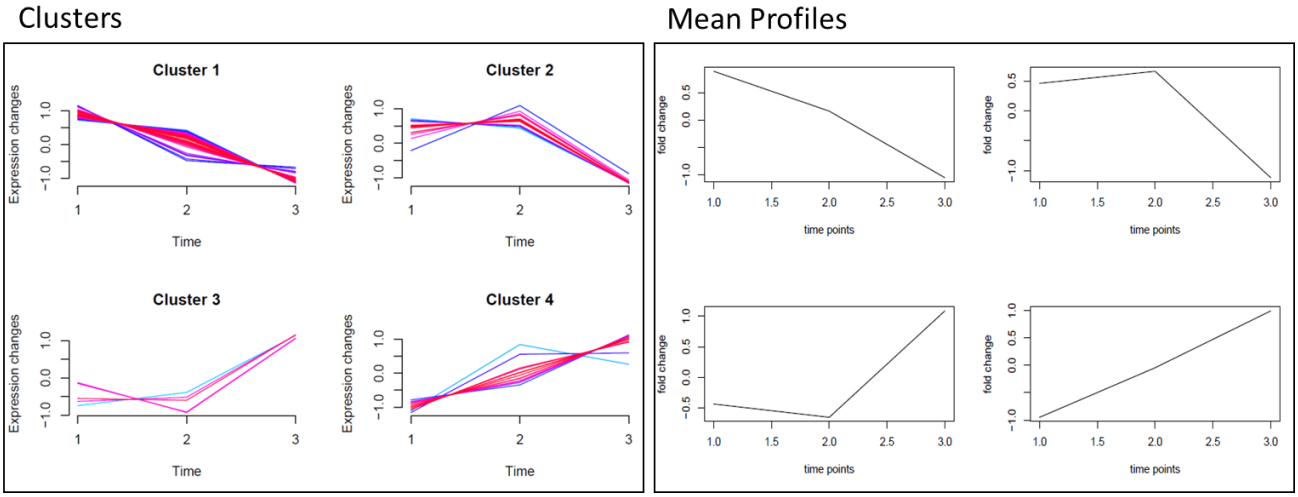
Genomic structure of IAPV and corresponding structural and non-structural polyproteins. Red sequences indicated peptides detected by LC-MS/MS analysis.

3.3.5 Analysis of protein expression in honey bees during IAPV infection

For the expression analysis of host proteins, only proteins that were quantified over all 3 time points were included to create a final dataset of 323 host proteins. Many high-throughput studies apply a threshold cut-off to distinguish protein expression changes. Often the choice of threshold is based on criteria such as fold change, which may exclude significantly regulated proteins and include non-significantly changing proteins (Albaum et al., 2011). To attempt to separate significant from non-significant changes, a statistical one-way ANOVA test with a threshold value of 0.05 was applied to infer significance. Using this method 71 host proteins were found to be regulated significantly out of the total 323 proteins quantified at all three time points. Next, the expression patterns of these 71 differentially regulated proteins were

investigated for similarity in over time. To this end fuzzy K-means clustering was performed using the software environment R (R Core Team, 2013). Analysis was performed using between 3 and 8 clusters, and four clusters were ultimately chosen because this number of clusters best separated the expression patterns into distinct groups with multiple protein members per group (Figure 10). Cluster 1 and 2 generally show a decrease in protein expression over time. In cluster 1 protein expression is decreased at each successive time point, and may indicate proteins which are quickly and continually suppressed during viral infection. Proteins in cluster 2 appear to maintain their current expression from 0 to 12 hours but are decreased at the third time point. These proteins do not appear to be suppressed during initial infection but decrease quickly with sustained viral replication. Clusters 3 and 4 include proteins with increased expression over time and expression trends opposite to clusters 1 and 2. Proteins in cluster 4 are upregulated immediately upon infection while the small number proteins in cluster 3 become upregulated after 12 hours and have increased expression at 36 hours post infection. A complete list of significant proteins including cluster membership is shown in Table 3.

Figure 10 Fuzzy K-means clustering of significantly changing proteins.



Clusters (left) and cluster mean profiles (right) created from 71 proteins determined by ANOVA test to be significantly regulated over time.

Table 3 Proteins significantly changing during IAPV infection

Sequence Id	0h	12h	36h	Cluster	Dmel_Gene_ID	Dmel_Name
gnl|Amel_4.5|GB41443-PA	0.45	-0.48	-2.23	1	#N/A	#N/A
gnl|Amel_4.5|GB50000-PA	-0.03	-0.22	-1.23	1	#N/A	#N/A
gnl|Amel_4.5|GB47839-PA	0.30	0.06	-0.98	1	FBgn0025682	scf-PC
gnl|Amel_4.5|GB40503-PA	0.07	-0.18	-0.74	1	FBgn0032350	CG6287-PA
gnl|Amel_4.5|GB52999-PA	0.28	-0.04	-0.61	1	#N/A	#N/A
gnl|Amel_4.5|GB41818-PA	0.39	0.23	-0.61	1	FBgn0031992	CG8498-PA
gnl|Amel_4.5|GB52520-PA	0.61	0.36	-0.56	1	FBgn0034345	CG5174-PO
gnl|Amel_4.5|GB49410-PA	0.29	0.08	-0.55	1	FBgn0000253	Cam-PA
gnl|Amel_4.5|GB44869-PA	0.10	-0.08	-0.51	1	FBgn0250791	Snap-PA
gnl|Amel_4.5|GB43750-PA	0.39	0.03	-0.44	1	FBgn0038976	CG7048-PA
gnl|Amel_4.5|GB54861-PA	0.42	0.17	-0.37	1	#N/A	#N/A
gnl|Amel_4.5|GB49117-PA	0.11	0.00	-0.37	1	FBgn0001218	Hsc70-3-PB
gnl|Amel_4.5|GB54211-PA	0.14	0.02	-0.36	1	FBgn0039802	dj-1beta-PA
gnl|Amel_4.5|GB45947-PA	0.21	0.05	-0.34	1	FBgn0024841	Pcd-PA
gnl|Amel_4.5|GB48604-PA	0.41	-0.22	-0.34	1	FBgn0260010	rump-PA
gnl|Amel_4.5|GB47103-PA	0.18	-0.03	-0.33	1	FBgn0028737	Ef1beta-PA
gnl|Amel_4.5|GB44398-PA	0.14	0.04	-0.32	1	FBgn0016697	Prosalpha5-PA
gnl|Amel_4.5|GB46214-PA	0.24	0.08	-0.30	1	FBgn0014869	Pglym78-PA
gnl|Amel_4.5|GB44422-PA	0.20	-0.14	-0.29	1	#N/A	#N/A
gnl|Amel_4.5|GB51065-PA	0.24	0.07	-0.28	1	FBgn0261593	RpS10b-PB
gnl|Amel_4.5|GB53550-PA	0.40	0.19	-0.28	1	FBgn0031037	CG14207-PB
gnl|Amel_4.5|GB49757-PA	0.68	0.25	-0.27	1	FBgn0037913	fabp-PC
gnl|Amel_4.5|GB45569-PA	0.31	0.19	-0.25	1	FBgn0000181	bic-PA
gnl|Amel_4.5|GB50356-PA	0.13	-0.08	-0.24	1	FBgn0003274	RpLP2-PA
gnl|Amel_4.5|GB41633-PA	0.25	0.02	-0.23	1	FBgn0016120	ATPsyn-d-PB
gnl|Amel_4.5|GB42297-PA	0.12	-0.06	-0.22	1	FBgn0001220	Hsc70-5-PA
gnl|Amel_4.5|GB50917-PA	0.16	0.05	-0.22	1	FBgn0002593	RpLP1-PA
gnl|Amel_4.5|GB44693-PA	0.19	0.09	-0.21	1	FBgn0034259	CG6459-PA
gnl|Amel_4.5|GB40461-PA	0.27	0.06	-0.20	1	FBgn0005585	Crc-PA
gnl|Amel_4.5|GB52736-PA	0.17	0.08	-0.18	1	FBgn0010217	ATPsyn-beta-PA
gnl|Amel_4.5|GB49047-PA	0.28	0.13	-0.16	1	FBgn0028684	Tbp-1-PA
gnl|Amel_4.5|GB50942-PA	0.25	0.08	-0.14	1	FBgn0004907	14-3-3zeta-PD
gnl|Amel_4.5|GB41207-PA	0.28	0.13	-0.13	1	FBgn0028694	Rpn11-PA
gnl|Amel_4.5|GB54446-PA	0.18	-0.05	-0.13	1	FBgn0000116	Argk-PA
gnl|Amel_4.5|GB45099-PA	0.28	0.19	-0.11	1	FBgn0010213	Sod2-PA
gnl|Amel_4.5|GB42526-PA	0.21	0.08	-0.10	1	FBgn0262559	Mdh2-PA
gnl|Amel_4.5|GB53138-PA	0.35	0.12	-0.10	1	FBgn0016687	Nurf-38-PA
gnl|Amel_4.5|GB43300-PA	0.23	0.01	-0.07	1	FBgn0263396	sqd-PB
gnl|Amel_4.5|GB49097-PA	0.40	-0.01	-0.06	1	FBgn0001280	janA-PB
gnl|Amel_4.5|GB54343-PA	0.58	0.42	-0.06	1	FBgn0036334	CG11267-PA

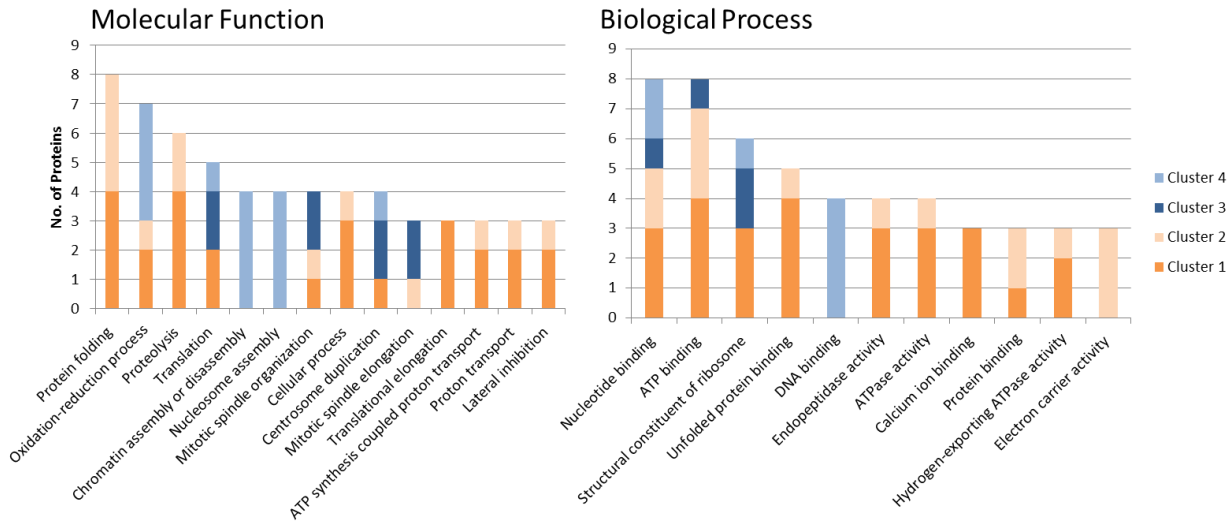
Sequence Id	0h	12h	36h	Cluster	Dmel_Gene_ID	Dmel_Name
gnl|Amel_4.5|GB54343-PA	0.58	0.42	-0.06	1	FBgn0036334	CG11267-PA
gnl|Amel_4.5|GB43161-PA	0.26	0.66	-0.88	2	FBgn0026084	cib-PA
gnl|Amel_4.5|GB48820-PA	-0.09	-0.12	-0.65	2	FBgn0028987	Spn2-PA
gnl|Amel_4.5|GB44431-PA	0.07	0.14	-0.60	2	FBgn0015282	Pros26.4-PA
gnl|Amel_4.5|GB55232-PA	0.17	0.12	-0.60	2	FBgn0021765	scu-PA
gnl|Amel_4.5|GB54862-PA	0.23	0.13	-0.49	2	FBgn0250848	26-29-p-PA
gnl|Amel_4.5|GB41867-PA	0.19	0.23	-0.36	2	FBgn0039562	Gp93-PA
gnl|Amel_4.5|GB42809-PA	0.25	0.17	-0.33	2	FBgn0037874	Tctp-PA
gnl|Amel_4.5|GB47336-PA	0.11	0.14	-0.32	2	FBgn0014002	Pdi-PA
gnl|Amel_4.5|GB54222-PA	0.08	0.12	-0.30	2	FBgn0023529	CG2918-PA
gnl|Amel_4.5|GB55643-PA	0.00	0.50	-0.26	2	FBgn0016691	Oscp-PA
gnl|Amel_4.5|GB45752-PA	0.15	0.24	-0.12	2	FBgn0000173	ben-PA
gnl|Amel_4.5|GB47553-PA	0.24	0.20	-0.12	2	FBgn0010516	wal-PA
gnl|Amel_4.5|GB55318-PA	0.14	0.18	-0.10	2	FBgn0086904	Nacalpha-PB
gnl|Amel_4.5|GB47960-PA	0.16	0.20	-0.10	2	FBgn0011661	Moe-PB
gnl|Amel_4.5|GB47462-PA	0.17	0.28	-0.10	2	FBgn0033663	ERp60-PA
gnl|Amel_4.5|GB44039-PA	0.26	0.38	0.08	2	FBgn0262782	Mdh1-PA
gnl|Amel_4.5|GB40280-PA	0.12	0.19	0.49	3	FBgn0027580	CG1516-PI
gnl|Amel_4.5|GB45334-PA	-0.03	-0.05	0.49	3	FBgn0014026	RpL7A-PA
gnl|Amel_4.5|GB54419-PA	0.09	0.14	0.76	3	#N/A	#N/A
gnl|Amel_4.5|GB45394-PA	0.47	0.15	0.99	3	FBgn0038771	CG4390-PB
gnl|Amel_4.5|GB43559-PA	0.26	-0.27	1.04	3	FBgn0020910	RpL3-PA
gnl|Amel_4.5|GB43731-PA	-0.11	0.41	0.26	4	FBgn0015222	Fer1HCH-PF
gnl|Amel_4.5|GB50902-PA	0.11	0.24	0.45	4	FBgn0001091	Gapdh1-PA
gnl|Amel_4.5|GB44339-PA	-0.19	0.18	0.51	4	#N/A	#N/A
gnl|Amel_4.5|GB50272-PA	0.03	0.29	0.54	4	FBgn0031418	CG3609-PA
gnl|Amel_4.5|GB50158-PA	-0.18	0.06	0.55	4	FBgn0003279	RpL4-PA
gnl|Amel_4.5|GB44760-PA	-0.04	0.12	0.64	4	#N/A	#N/A
gnl|Amel_4.5|GB51071-PA	0.01	0.33	0.71	4	FBgn0033735	CG8525-PA
gnl|Amel_4.5|GB46366-PA	0.02	0.24	0.72	4	#N/A	#N/A
gnl|Amel_4.5|GB44318-PA	-0.31	0.30	0.72	4	#N/A	#N/A
gnl|Amel_4.5|GB47380-PA	-0.77	-0.26	0.80	4	#N/A	#N/A
gnl|Amel_4.5|GB53755-PA	0.06	0.38	1.18	4	#N/A	#N/A

Significant honey bee proteins (n = 71) determined by ANOVA from all proteins quantified at all time points (n = 323). Column A: Protein sequence ID. Column B-D: Protein expression values at 0, 12, and 36 h time points. Column E: Cluster membership. Column F: Gene ID of homologous *Drosophila* proteins if available. Column G: Protein name corresponding to *Drosophila* gene ID.

The expression of host proteins may be modulated during infection for several reasons. Immune related proteins may be upregulated by the host to defend against the virus. Upregulation can also be seen in host proteins hijacked by the virus to aid viral replication, or as part of the stress response to prolonged infection. Downregulated proteins may be the result of viral suppression of normal host cell activities or the diversion of normal cell function by the host to increase the immune response. Furthermore, the specific proteins regulated during these processes depend on the type of pathogen. Pathogens, whether fungal, bacterial, or viral, typically employ unique mechanisms of pathogenesis which will involve different host proteins and trigger the host immune pathways to varying degrees (Cherry et al., 2005). The functional analysis of proteins with different expression patterns can be used as a first step to gain insight into the host-pathogen relationship and the mechanisms employed by the pathogen studied. However, a well-annotated genome is required to perform functional analysis, making this difficult in the honey bee. Although the complete sequence of the honey bee genome was initially published in 2006 and updates including some genome annotation have since been released (Consortium, 2006), a complete annotation of the most recent release of the honey bee genome is still unavailable. *Drosophila* homologues have been identified for the majority of bee genes, including several 1:1 homologues. The similarity between honey bee and *Drosophila* pathways and the utility of using *Drosophila* to model honey bee molecular function including the immune system has been previously discussed (Evans et al., 2006). Therefore, proteins in each cluster were functionally classified based on homologous *Drosophila* proteins which were assigned one or more Gene Ontology (GO) terms. However, some honey bee proteins did not have *Drosophila* homologues, and further investigation of these proteins may reveal unique honey bee proteins which are involved in infection. GO Term enrichment can be determined by statistical analysis and is often performed using online bioinformatic programs such as DAVID (Database for Annotation, Visualization, and Integrated Discovery). This is ideal, as the program calculates a measure of the significance of enrichment of each GO Term by comparing the GO Terms associated with the differentially expressed proteins against a “background” containing the GO Terms of all proteins identified. However, for this study, the number of proteins in each cluster was small (between 5 and 37 proteins per cluster) and attempts to assign statistical significance for the 71 differentially expressed proteins against the background of 323 total quantified proteins using the program DAVID was negative. Instead, to provide a visual representation of the most prominent GO Terms in the data set, the total number of proteins associated with each GO term was plotted in Figure 11. The general expression trends of proteins associated with these terms including individual proteins assigned to each GO Term

are listed in Table 4 and can be used to hypothesize the involvement of different protein members.

Figure 11 GO Term assignment.



Total number of proteins assigned to each molecular function (left) and biological process (right) GO term. Terms with less than 3 assigned proteins are not shown. Cluster membership of the proteins associated with each GO term is indicated by bar colour: dark orange (cluster 1), light orange (cluster 2), dark blue (cluster 3) or light blue (cluster 4).

Several general trends of protein expression were identified from the GO Term assignments. For example, proteins associated with proteolysis (GO:0006508) were members of clusters 1 and 2 which had downregulated expression profiles. Proteins in this group included 26-29-p and proteosomal subunits Tbp-1, Rpn11, Prosα5, and Pros26.4. The decreased expression of these proteins over time may indicate that proteolytic pathways are inhibited during IAPV infection. Proteins associated with translation or translational elongation (GO:0006412 and GO:0006414), including Ef1beta, RpLP2, and RpLP1 were also downregulated over the course of IAPV infection. Other ribosomal proteins, however, were associated with clusters which had strongly upregulated expression profiles, including RpL4, RpL3, and RpL7A. Histone proteins including His4, His2B, His2A, and His3 were exclusively found in cluster 4, indicating strong upregulation during infection. Other processes represented in the dataset include calcium ion binding (GO005509), proton transport (GO:0015986), and oxidation-reduction (GO:0055114).

Table 4 Cluster membership and GO Term assignment of individual proteins regulated during IAPV infection.

Cluster	GO Term BP	Description	Proteins
2,3	GO:0000022	Mitotic spindle elongation	RpL3, RpL7A, Pros26.4
4	GO:0006333	Chromatin assembly or disassembly	His4:CG33909, His2B:CG33898, His2A:CG33856, His3:CG33851
4	GO:0006334	Nucleosome assembly	His4:CG33909, His2B:CG33898, His2A:CG33856, His3:CG33851
1,3,4	GO:0006412	Translation	RpL4, RpL3, RpL7A, RpLP2, RpLP1
1	GO:0006414	Translational elongation	Ef1beta, RpLP2, RpLP1
1,2	GO:0006457	Protein folding	Gp93, Pdi, ERp60, CG2918, Crc, Hsc70-5, CG7048, 14-3-3zeta
1,2	GO:0006508	Proteolysis	Pros26.4, 26-29-p, Rpn11, Prosalpha5, Tbp-1, 26-29-p
1,2,3	GO:0007052	Mitotic spindle organization	RpL3, RpL7A, Pros26.4, Cam
1,2	GO:0009987	Cellular process	Pros26.4, Rpn11, Prosalpha5, Tbp-1
1,2	GO:0015986	ATP synthesis coupled proton transport	Oscp, ATPsyn-d, ATPsyn-beta
1,2	GO:0015992	Proton transport	Oscp, ATPsyn-d, ATPsyn-beta
1,2	GO:0046331	Lateral inhibition	Mdh1, Cam, ATPsyn-beta
1,3,4	GO:0051298	Centrosome duplication	RpL4, RpL3, RpL7A, Hsc70-3
1,2,4	GO:0055114	Oxidation-reduction process	Fer1HCH-PF, Pdh, CG3609, Gapdh1, Mdh1, CG6287, Sod2

Cluster	GO Term MF	Description	Proteins
1,2,3,4	GO:0000166	Nucleotide binding	Pdh, CG3609, antdh, Mdh1, scu, Mdh2, sqd, rump
4	GO:0003677	DNA binding	His4:CG33909, His2B:CG33898, His2A:CG33856, His3:CG33851
1,3,4	GO:0003735	Structural constituent of ribosome	RpL4, RpL3, RpL7A, RpLP2, RpLP1, RpS10b
1,2	GO:0004175	Endopeptidase activity	Pros26.4, Rpn11, Prosalpha5, Tbp-1
1	GO:0005509	Calcium ion binding	Crc, scf, Cam
1,2	GO:0005515	Protein binding	Moe, Nacalpha, 14-3-3zeta
1,2,3	GO:0005524	ATP binding	CG1516-PI, Gp93, Pros26.4, CG2918, Hsc70-5, Tbp-1, Hsc70-3, ATPsyn-beta
1,2	GO:0008553	Hydrogen-exporting ATPase activity	Oscp, ATPsyn-d, ATPsyn-beta
2	GO:0009055	Electron carrier activity	Pdi, ERp60, wal
1,2	GO:0016887	ATPase activity	Pros26.4, Hsc70-5, Tbp-1, Hsc70-3
1,2	GO:0051082	Unfolded protein binding	Gp93, Crc, Hsc70-5, CG7048, CG11267

Individual GO Terms for biological process (GO Term BP; upper table) and molecular function (GO Term MF; lower table) contain proteins from single and multiple clusters.

3.4 Discussion

To increase our fundamental knowledge of the host-pathogen relationship between IAPV and the honey bee, a high-throughput quantitative proteomics strategy was used to examine protein expression changes between IAPV infected and healthy honey bee pupae. Exclusive infection of pupae with IAPV and the absence of background infections was confirmed prior to proteomic analysis by RT-PCR. The speed, sensitivity, and accuracy of this method make it an excellent diagnostic tool for the detection of honey bee viruses, especially in light of the lack of cell culture based methods which are traditionally used for the purification and propagation of viruses (Boncristiani et al., 2009, de Miranda et al., 2009, Stoltz et al., 1995). The experimental approach chosen for LC-MS/MS analysis included the stable isotope labeling of pupal tissues by chemical dimethylation followed by multidimensional separation of peptides using SCX fractionation. This resulted in the identification of between ~900 to ~1200 proteins at each time point of infection with a false discovery rate of <1%. Following manual validation of the quantitation process, quantitative data was obtained for nearly 80% of proteins identified at each

time point. The validation of quantitative data is an important step, as there are many potential sources of error that can affect the quantitation process which should be identified and addressed. Spectral background noise can affect quantitation by causing interference, although this can typically be filtered by data processing packages. Interference from co-eluting peptides of similar m/z can also result in the incorrect assignment of intensities, and low-intensity or noisy data may also distort the mean value of computed ratios (Bantscheff et al., 2007). The identification and removal of such spectra ensured a high quality data set.

Initial comparison of quantitation ratios revealed good correlation between high and low dose IAPV injected pupae, indicating that both dosages were sufficient to cause severe disease. This is a testament to the virulent nature of *Dicistroviridae* viruses when directly injected, and likely only a few virus particles are required for infection through this route (Bakonyi et al., 2002, Chen et al., 2007, de Miranda et al., 2009). Protein expression ratios were also compared between the two hives as well as between replicates within each hive. Ratio correlation was found to be similar for both groups, although interestingly correlation was much less strong at the 0 h time point than at 12 or 36 hours post infection. This might reflect the diverse physiological states of healthy pupae prior to infection and the diversion of host processes to defense during later time points resulting in less variation in protein expression. Systemic biological noise is also likely prominent in the dataset due to the use of small numbers of control pupae at each replicate. Pronounced individual variability could have been reduced by combining all control PBS-injected pupae into a uniform sample that is spiked into each replicate. Homogenous material such as cultured cells would be another route to eliminate variability, and use of cell cultures should be attempted if a honey bee line becomes available.

A comparison of protein identifications between the two hives indicated that nearly the same proteins were identified in each hive. This is unusual since typically only a subset of peptides are identified in a single analysis due to the complexity of peptides at varying concentrations present in any one sample. Thus, even the repeat analysis of a single sample will result in the identification of different proteins (Bantscheff et al., 2007). It is possible that the redundancy of identification seen between hives in this study reflects a problem in the protein database used to identify proteins from mass spectra, where erroneous or missing database entries would prevent the identification of good spectra acquired during LC-MS/MS analysis. MS-based proteomics is made possible by the availability of gene and genome sequence databases, from which protein sequences can be predicted and annotated. As the honey bee genome has been sequenced relatively recently and protein annotations are currently being updated this may be a

likely scenario, one which will hopefully soon be remedied through ongoing annotation efforts. Identifications were less similar between time points, with approximately 500 proteins identified across 0, 12, and 36 hours post infection. Although this degree of overlap is more typical of a complex proteomic analysis, given the high similarity between hive replicates it is also possible that the decreased overlap is due to the expression of different proteins in response to progressing infection. This may indicate that proteins detected at 12 and 36 hours but not 0 hours, or conversely proteins detected at 0 and 12 hours but not 36 hours are in fact regulated during infection to such an extent that they are below the level of detection at the very early or very late time points. Further examination of these proteins could provide additional information on biological processes that are suppressed or initiated during infection, however the approach of clustering protein expression profiles used in this study would be difficult to apply to proteins that are not quantified at all three time points. It is therefore important for future work to first identify the presence and number of biologically relevant proteins lacking quantification at one of the three time points, and accordingly apply a method that allows the analysis of these proteins together with other quantified proteins as one complete dataset.

The viral life cycle involves several steps, including entry into the host cell cytoplasm, uncoating, translation, replication, and virion assembly. While the molecular functionality of the honey bee *Dicistroviridae* has not yet been described, it may be similar to related viral species such as the Drosophila C Virus (DCV), a dicistrovirus that infects the fruit fly *Drosophila melanogaster*. Studies of DCV indicate that the virus enters the host cell through clathrin-mediated endocytosis and once inside, uncoats and releases its genomic RNA into the cytoplasm to begin replication. The viral RNA genome is copied by the ORF 1-encoded RdRP into a negative-sense complementary ssRNA intermediate which serves as the template for the synthesis of new viral genomic RNA (Cherry et al., 2006). It is well known that close association with the host cell machinery is needed for viral replication. However, in order to gain access to the host factors required for replication, the virus must compete directly with the host cell. One feature of IRES-containing dicistroviruses is that they are capable of bypassing canonical 5'-cap dependent translation by directly binding specific ribosomal proteins at the IRES site to initiate translation (Cherry et al., 2005, Gilbert, 2011). The modulation of host ribosomal proteins was observed during IAPV infection. Proteins associated with translation or translational elongation, including Ef1beta, RpLP2, and RpLP1 were downregulated over the course of IAPV infection, while other ribosomal proteins including RpL4, RpL3, and RpL7A had strictly upregulated expression profiles. The upregulation of some host ribosomal proteins may thus indicate that these are

involved in IRES-mediated translation, while other translational proteins specific to host cell translation are downregulated. IRES-containing RNA viruses such as DCV have also been found to be unusually sensitive to host ribosome levels. The identification of host factors required for IRES-dependent translation will not only aid the elucidation of the pathogenic mechanism of IAPV infection, but may also potentially provide direction for screens of small molecules that inhibit IRES function (Cherry et al., 2005). RNA-binding proteins are known to represent another important factor in the control of transcriptional initiation. Many RNA-binding proteins have been implicated in viral infection, and proteins previously described as affecting translational processes were identified in this study as being both upregulated and downregulated to infection. Some of these factors may be crucial to IAPV translation and replication, and the role of these proteins including Gapdh1, sqd, and Hsc70 should be further investigated (Pacheco et al., 2009).

The products of viral translation and replication were clearly detected in this analysis, with proteins corresponding to both structural and non-structural IAPV ORFs identified with high coverage at 12 and 36 h post infection. The accumulation of viral proteins by 12 h indicates the ability of IAPV to replicate quickly when injected into its host. During replication, the viral polyproteins translated from each ORF are proteolytically digested into their functional proteins, including the capsid proteins encoded by ORF2 and the helicase, protease, and RNA-dependent RNA polymerase (RdRP) proteins encoded by ORF1. The protease activity is primarily provided by the ORF1 encoded protease, but also host cellular proteases (de Miranda et al., 2009). The identification of peptides from all polyproteins regions reflects the presence of both precursor and final viral proteins in the host cell.

Viral infection induces significant proteomic changes within the host, although differences in the antiviral response to different virus species highlights the variation in pathogenic strategies used by different virus families (Cherry et al., 2005, Kellam et al., 2001). In response to DCV infection, the Toll and Imd pathways are not activated, indicating that canonical NF-κB signaling pathways are not required for the antiviral response. Instead infection strongly upregulates the RNAi and Jak/STAT pathways, which if inhibited result in greatly increased severity of DCV infection (Cherry et al., 2006). In this study, the honey bee AMPs including Abaecin, Apidaecin, Defensin, Hymenoptaecin, and Jellein were not detected during any time point of infection. While this does not necessarily indicate these proteins were completely absent, it is also likely that infection with IAPV does not result in the substantial production of AMPs. This would be

consistent with the activation of primarily non-AMP producing RNAi and Jak/STAT pathways in flies infected with Drosophila C Virus (Cherry et al., 2003).

Several other functional categories were represented which can reflect the disease process, including calcium ion binding, histone activity, and protein degradation. Expression of histone proteins His4, His2B, His2A, and His3 was altered during infection, although histone involvement in the immune response has been characterized as typical of DNA virus infection (Kobiyama et al., 2010). Protein degradation genes such as those of the ubiquitination-proteasome pathway may be induced during viral infection to aid immune evasion or release of viral progeny (Gao et al., 2006). Proteins associated with proteolysis were found to be significantly changing in this analysis, such as 26-29-p and proteosomal subunits Tbp-1, Rpn11, Prosα5, and Pros26.4 (also known as Rpt2). However, these proteins had downregulated expression profiles over the course of infection, possibly indicating that proteolytic pathways are inhibited during IAPV infection.

3.5 Conclusion

The quantitative proteomic approach used in this study allowed the investigation of many proteins important to progression of IAPV infection in honey bees. Experimental design involving stable isotope labeling of tissues and extensive LC-MS/MS analysis resulted in the identification and quantification of hundreds of honey bee proteins, representing one of few large-scale investigations into honey bee protein biology. Expression profiling of proteins involved in infection identified several functional categories which provided insight into the IAPV infection process. This study demonstrates that quantitative proteomics is a valuable tool for honey bee research that can be applied to other biological processes to provide accurate and reproducible information that cannot be gained through other methods.

4 Conclusion

Recent declines in honey bee populations worldwide have spurred significant research into the impact of pathogens on colony health. The Israeli Acute Paralysis Virus has become of particular concern since being correlated with colony losses (Cox-Foster et al., 2007, Potts et al., 2010). However, little is known regarding the mechanism of pathogenesis used by IAPV, and currently only a few studies have addressed the response of the honey bee immune system to pathogen infection at a molecular level (Evans et al., 2006, Maori et al., 2009). While this lack of knowledge poses a considerable barrier if we are to address the impact IAPV has on honey bee health, significant advances have been made recently in the area of honey bee research. The sequencing of the honey bee genome in 2006 (Consortium, 2006) has resulted in available gene and protein sequence libraries, providing a basis for high-throughput analyses such as mass spectrometry-based proteomics. Mass spectrometry is a powerful tool in analysis of protein expression and biological function and is capable of identifying thousands of proteins from whole cell lysates (de Godoy et al., 2008, Yates et al., 2009). Mass spectrometry is especially useful in the field of honey bee research since MS analysis can be done directly using tissues and does not exclusively require tools often unavailable to honey bee researchers such as cell lines and antibodies. Thus, while the development of a honey bee in vitro system such as cell culture will be highly advantageous and should be pursued, current mass spectrometry-based proteomic methods can provide significant insight into the molecular processes and pathways of the honey bee immune response.

4.1 Addressing the project aims and hypotheses

In this thesis, two routes of research were conducted to aid our understanding of the honey bee host-pathogen relationship. First, a cell culture system using honey bee hemocytes was established and optimized in chapter 2 with the aim of addressing the lack of an available honey bee *in vitro* system, which has significantly hindered honey bee molecular research and especially host-pathogen studies. The use of larval hemocytes is an obvious route for the exploration of honey bee cell culture, as these cells are easy to obtain in relatively large numbers. However, while hemocyte cultures were simple to initiate and maintain, cell division was not observed. Although the honey bee *Ras85D* and human c-myc oncogenes were hypothesized to be two good candidate genes to induce cell proliferation and immortalization through transfection into the cultured hemocytes, attempts to transfect these oncogenes were hindered by the fragility of hemocyte cells. This indicates that further optimization of current

transfection protocols or use of alternate methods will be necessary for the successful immortalization of hemocytes. It is ideal to have access to continuously proliferating cells for experimentation using cell culture. However, isolating or inducing such proliferating cultures of honey bee cells appears to be highly difficult, as shown both in this work as well as in other documented attempts to culture a variety of honey bee cell types. Currently, larval honey bee hemocyte cultures represent a useful tool for some experimental applications requiring medium amounts of cells and not involving substantial manipulation.

Next, to investigate changes in host protein expression during IAPV infection, mass spectrometry-based quantitative proteomics was used to compare IAPV infected and healthy pupae in chapter 3. This approach applied a stable isotope dimethylation labeling strategy combined with multidimensional fractionation using SCX to identify and quantify ~800 proteins over three time points of infection. Because host protein expression changes significantly during viral infection, it was hypothesized that the expression profile of several proteins would be differentially regulated in response to IAPV infection. As a result, a predominantly discovery-based approach was taken in this this study. Of the over 300 proteins identified across all three time points nearly 80 were determined to be significantly changing over the course of the 36 h infection period, indicating substantial changes in host proteins during IAPV infection. To further investigate the honey bee host response to IAPV the expression profiles of significantly changing proteins were clustered into four distinct patterns, providing a sense of which proteins were up and downregulated during infection. To next infer specific functional roles of these proteins, *Drosophila* homologues were obtained for each protein in the data set and corresponding GO Terms were identified. Proteins involved in processes including translation, the ubiquitin-proteasome pathway, as well as several histone proteins, were identified indicating that these processes are critically involved during infection. The involvement of these pathways corresponds well with our knowledge of other closely related viruses, and future investigation of these pathways will be important for identifying the role of specific host proteins involved during infection. This analysis represents an important first step towards understanding the honey bee host response to IAPV infection through the systems-level analysis of protein expression. The identified expression changes in host honey bee proteins provides important direction for future research into the effect of IAPV and potential treatment options, and this study demonstrates the utility of mass spectrometry-based proteomics in honey bee research.

4.2 Future directions

4.2.1 Transfection of primary honey bee hemocyte cells

Primary cells are notoriously difficult to transfect, and in this thesis both electroporation and lipofection proved to be too damaging to hemocytes to allow cell survival while successfully transfecting the gene of interest. Various methods of transfection exist, and use of alternate techniques such as nucleofection, a modified method of electroporation, or through baculovirus-mediated transfection may more successfully allow transfection and encourage hemocyte division and subsequent immortalization. Investigation into which of these techniques most effectively transfect honey bee primary hemocytes while maintaining cell viability will be necessary for further efforts to create an immortalized, continuous honey bee cell line. Once the method of transfection is adequately optimized and efficient promoter function can be determined through the expression of a marker protein such as fluorescent GFP, then the transfection of oncogenes may be attempted singly, in combination, or with other genes such as telomerase.

4.2.2 Coinfection studies with IAPV

An important question is to determine the effect of IAPV both individually as well as in combination with other pathogens (Potts et al., 2010). Coinfection studies where honey bees are infected with IAPV as well as other viruses including DWV, KBV, or ABPV, fungal pathogens such as Nosema species, or the parasitic Varroa mite will provide insight into potential synergetic or antagonistic interactions between pathogens. In the highly social setting of the honey bee hive, single infections are rarely found (Cornman et al., 2012, Runckel et al., 2011). Instead, the entire pathogen community present in the hive should be considered. Thus, the experimental design described in this thesis should be applied to coinfection studies with multiple pathogens to guide our understanding of the effect of compound infections as well as aid the development of treatments for bees infected with several pathogens in the field.

4.2.3 Role of histone proteins in IAPV infection

One interesting result obtained during the quantitative proteomic comparison of IAPV infected and uninfected honey bee pupae in this thesis was the increased expression of host histone proteins over the course of IAPV infection. Although DNA viruses are most characteristic of manipulating the host histone proteins, many different virus types are capable of altering the chromatin environment of the host to enhance viral replication or survival. Positive single strand

RNA viruses have been found to be reverse transcribed in *Drosophila* cells (Goic et al., 2013), and it is further known that segments of the IAPV genome are integrated into the honey bee host genome (Maori et al., 2007). It may be possible that during replication or integration, IAPV interacts with or recruits host histone proteins. Further investigation into the role of histones during IAPV infection may provide interesting insight into the mechanism of pathogenesis used by *Dicistroviridae* such as IAPV to infect their hosts.

4.2.4 Analysis of the host and viral phosphoproteome during IAPV infection

Post-translational modifications are known to play an important role in the host-pathogen interaction for several virus species. Phosphorylation of viral proteins may affect viral activities such as entry into the host cell, localization, and replication (Hutchinson et al., 2012). Similarly, binding of the virus to the host cell can induce intracellular signaling cascades through changes in the phosphorylation status of proteins, affecting viral infectivity (Wojcechowskyj et al., 2013). Determination of protein phosphorylation is easily done by MS-based proteomics using SCX-based separation as described in this study, and is capable of identifying hundreds or thousands of phosphorylation sites. Analysis of both the phosphorylation status of IAPV and changes in the host phosphoproteome during different time points of infection would provide insight into the mechanisms and pathways affected by IAPV infection.

4.3 Closing

Our molecular knowledge of the honey bee immune system has been greatly limited due to a lack of *in vitro* tools and the absence of systems-level studies of honey bee biochemical pathways. While continuous cell culture remains a difficult to overcome obstacle for honey bee research, this thesis also demonstrates the value of MS-based proteomics in furthering our understanding of the honey bee immune response at a systems-level. The importance of achieving a robust understanding of the interaction between the honey bee and its many pathogens has become very clear in recent years due to significant declines in managed honey bee populations. IAPV is one pathogen which has been correlated with poor colony health, and it is hoped that this thesis will provide future direction for identification of the impact of IAPV infection singly as well as in combination with other stressors. Ultimately this knowledge will contribute to treatment of infected hives and protection of this economically important and biologically fascinating organism.

References

Aebersold, R., & Mann, M. (2003). Mass spectrometry-based proteomics. *Nature, 422,* 198-207.

Albaum, S. P., Hahne, H., Otto, A., Haußmann, U., Becher, D., Poetsch, A., Goesmann, A., et al. (2011). A guide through the computational analysis of isotope-labeled mass spectrometry-based quantitative proteomics data: an application study. *Proteome science*, *9*(1), 30.

Asha, H., Nagy, I., Kovacs, G., Stetson, D., Ando, I., & Dearolf, C. R. (2003). Analysis of Ras-induced overproliferation in Drosophila hemocytes. *Genetics*, *163*(1), 203–15.

Bakonyi, T., & Grabensteiner, E. (2002). Phylogenetic analysis of acute bee paralysis virus strains. *Applied and environmental microbiology, 68*(12), 6446–6450.

Bantscheff, M., Schirle, M., Sweetman, G., Rick, J., & Kuster, B. (2007). Quantitative mass spectrometry in proteomics: a critical review. *Analytical and bioanalytical chemistry*, *389*(4), 1017–31.

Bergem, M., Norberg, K., & Aamodt, R. M. (2006). Long-term maintenance of in vitro cultured honeybee (Apis mellifera) embryonic cells. *BMC developmental biology*, *6*, 17.

Bishop, G.H. (1923). Body fluid of the honey bee larva. I. Osmotic pressure, specific gravity, pH, O2 capacity, CO2 capacity, and buffer value, and their changes with larval activity and metamorphosis. *Journal of biological chemistry,* 543-565.

Bishop, G. H. (1925). Chemistry of Bee Larval Blood. II. Chemical constituents of the blood, and their osmotic effects. *Journal of biological chemistry,* 77–88.

Boersema, P. J., Aye, T. T., Van Veen, T. B., Heck, A. J. R., & Mohammed, S. (2008). Triplex protein quantification based on stable isotope labeling by peptide dimethylation applied to cell and tissue lysates. *Proteomics*, *8*(22), 4624–32.

Boersema, P. J., Raijmakers, R., Lemeer, S., Mohammed, S., & Heck, A. J. R. (2009). Multiplex peptide stable isotope dimethyl labeling for quantitative proteomics. *Nature protocols*, *4*(4), 484–94.

Boncristiani, H. F., Di Prisco, G., Pettis, J. S., Hamilton, M., & Chen, Y. P. (2009). Molecular approaches to the analysis of deformed wing virus replication and pathogenesis in the honey bee, Apis mellifera. *Virology journal*, *6*, 221.

Bonning, B. C. (2009). The Dicistroviridae: An emerging family of invertebrate viruses. *Virologica sinica*, *24*(5), 415–427.

Bonning, B. C., & Miller, W. A. (2010). Dicistroviruses. *Annual review of entomology*, *55*, 129–50.

Bourouis, M., & Jarry, B. (1983). Vectors containing a prokaryotic dihydrofolate reductase gene transform Drosophila cells to methotrexate-resistance. *The EMBO journal*, *2*(7), 1099–1104.

Brandt, N. R., & Huber, R. E. (1979). Carbohydrate utilization in the thoraces of honey bees (Apis mellifera) during early times of flight. *Journal of insect physiology*, *25*(6), 483–486.

Chen, Y., & Pettis, J. (2006). Prevalence and transmission of honeybee viruses. *Applied and environmental microbiology, 72*(1), 606–611.

Cherry, S., Doukas, T., Armknecht, S., Whelan, S., Wang, H., Sarnow, P., & Perrimon, N. (2005). Genome-wide RNAi screen reveals a specific sensitivity of IRES-containing RNA viruses to host translation inhibition. *Genes & development*, *19*(4), 445–52.

Cherry, S., & Perrimon, N. (2004). Entry is a rate-limiting step for viral infection in a Drosophila melanogaster model of pathogenesis. *Nature immunology*, *5*(1), 81–7.

Chung, Y., & Keller, E. (1990). Positive and negative regulatory elements mediating transcription from the Drosophila melanogaster actin 5C distal promoter. *Molecular and cellular biology*, *10*(12), 6172–6180.

Cornman, R. S., Tarpy, D. R., Chen, Y., Jeffreys, L., Lopez, D., Pettis, J. S., vanEngelsdorp, D., et al. (2012). Pathogen webs in collapsing honey bee colonies. *PloS one*, *7*(8), e43562.

Costa, A., Jan, E., Sarnow, P., & Schneider, D. (2009). The Imd pathway is involved in antiviral immune responses in Drosophila. *PloS one*, *4*(10), e7436.

Cox-Foster, D. L., Conlan, S., Holmes, E. C., Palacios, G., Evans, J. D., Moran, N., Quan, P.-L., et al. (2007). A metagenomic survey of microbes in honey bee colony collapse disorder. *Science*, *318*(5848), 283–7.

Crailsheim, K. (1985). Distribution of haemolymph in the honeybee (Apis mellifica) in relation to season, age and temperature. *Journal of insect physiology*, *31*(9), 707–713.

De Godoy, L. M. F., Olsen, J. V, Cox, J., Nielsen, M. L., Hubner, N. C., Fröhlich, F., Walther, T. C., et al. (2008). Comprehensive mass-spectrometry-based proteome quantification of haploid versus diploid yeast. *Nature*, *455*(7217), 1251–4.

De Miranda, J. R., Cordoni, G., & Budge, G. (2010). The Acute bee paralysis virus-Kashmir bee virus-Israeli acute paralysis virus complex. *Journal of invertebrate pathology*, *103*, S30–47.

Deutsch, E. W., Lam, H., & Aebersold, R. (2008). Data analysis and bioinformatics tools for tandem mass spectrometry in proteomics. *Physiological genomics*, *33*(1), 18–25.

Di Prisco, G., Pennacchio, F., Caprio, E., Boncristiani, H. F., Evans, J. D., & Chen, Y. (2011). Varroa destructor is an effective vector of Israeli acute paralysis virus in the honeybee, Apis mellifera. *The journal of general virology*, *92*, 151–5.

Dostert, C., Jouanguy, E., Irving, P., Troxler, L., Galiana-Arnoux, D., Hetru, C., Hoffmann, J. a, et al. (2005). The Jak-STAT signaling pathway is required but not sufficient for the antiviral response of drosophila. *Nature immunology*, *6*(9), 946–53.

El-Aneed, A., Cohen, A., & Banoub, J. (2009). Mass spectrometry, review of the basics: Electrospray, MALDI, and commonly used mass analyzers. *Applied spectroscopy reviews*, *44*(3), 210–230.

Engström, Y., Loseva, O., & Theopold, U. (2004). Proteomics of the Drosophila immune response. *Trends in biotechnology*, *22*(11), 600–5.

Evans, J. D., Aronstein, K., Chen, Y. P., Hetru, C., Imler, J., Jiang, H., Kanost, M., et al. (2006). Immune pathways and defence mechanisms in honey bees Apis mellifera. *Insect molecular biology*, *15*(5), 645–56.

Foster, L. J., De Hoog, C. L., & Mann, M. (2003). Unbiased quantitative proteomics of lipid rafts reveals high specificity for signaling factors. *Proceedings of the national academy of sciences of the United States of America*, *100*(10), 5813–8.

Gallai, N., Salles, J.-M., Settele, J., & Vaissière, B. E. (2009). Economic valuation of the vulnerability of world agriculture confronted with pollinator decline. *Ecological economics*, *68*(3), 810–821.

Gao, G., & Luo, H. (2006). The ubiquitin-proteasome pathway in viral infections. *Canadian journal of physiology and pharmacology*, *84*(1), 5–14.

Gauci, S., Helbig, A. O., Slijper, M., Krijgsveld, J., Heck, A. J. R., & Mohammed, S. (2009). Lys-N and trypsin cover complementary parts of the phosphoproteome in a refined SCX-based approach. *Analytical chemistry*, *81*(11), 4493–501.

Genersch, E. (2010). Honey bee pathology: current threats to honey bees and beekeeping. *Applied microbiology and biotechnology*, *87*(1), 87–97.

Gil, J., Kerai, P., Lleonart, M., & Bernard, D. (2005). Immortalization of primary human prostate epithelial cells by c-Myc. *Cancer research*, *65*(6), 2179–2185.

Gilar, M., Olivova, P., Daly, A. E., & Gebler, J. C. (2005). Orthogonality of separation in two-dimensional liquid chromatography. *Analytical chemistry*, *77*(19), 6426–34.

Gilbert, W. V. (2011). Functional specialization of ribosomes? *Trends in biochemical sciences*, *36*(3), 127–32.

Gisder, S., Möckel, N., Linde, A., & Genersch, E. (2011). A cell culture model for Nosema ceranae and Nosema apis allows new insights into the life cycle of these important honey bee-pathogenic microsporidia. *Environmental microbiology*, *13*(2), 404–13.

Goic, B., Vodovar, N., Mondotte, J., Monot, C., Frangeul, L., Blanc, H., Gausson, V., et al. (2013). RNA-mediated interference and reverse transcription control the persistence of RNA viruses in the insect model Drosophila. *Nature immunology*, *14*(4), 396–403.

Gresch, O., Engel, F. B., Nesic, D., Tran, T. T., England, H. M., Hickman, E. S., Körner, I., et al. (2004). New non-viral method for gene transfer into primary cells. *Methods*, *33*(2), 151–63.

Guerrera, I. C., & Kleiner, O. (2005). Application of mass spectrometry in proteomics. *Bioscience reports*, *25*(1-2), 71–93.

Hunter, W. B. (2010). Medium for development of bee cell cultures (Apis mellifera: Hymenoptera: Apidae). *In vitro cellular & developmental biology. Animal*, *46*(2), 83–6.

Hutchinson, E. C., Denham, E. M., Thomas, B., Trudgian, D. C., Hester, S. S., Ridlova, G., York, A., et al. (2012). Mapping the phosphoproteome of influenza A and B viruses by mass spectrometry. *PLoS pathogens*, *8*(11), e1002993.

Johnson, R. M., Evans, J. D., Robinson, G. E., & Berenbaum, M. R. (2009). Changes in transcript abundance relating to colony collapse disorder in honey bees (Apis mellifera). *Proceedings of the national academy of sciences of the United States of America*, *106*(35), 14790–5.

Jones, B. (1962). The cultivation of insect cells and tissues. *Biological reviews*, (37), 512–536.

Jordan, E., Collins, M., & Terefe, J. (2008). Optimizing electroporation conditions in primary and other difficult-to-transfect cells. *Journal of biomolecular techniques*, (19), 328–334.

Kaatz, H., Hagedorn, H., & Engels, W. (1985). Culture of honeybee organs: development of a new medium and the importance of tracheation. *In vitro cellular & developmental …*, *21*(6).

Kellam, P. (2001). Post-genomic virology: the impact of bioinformatics, microarrays and proteomics on investigating host and pathogen interactions. *Reviews in medical virology*, *11*, 313–329.

Keller, A., Eng, J., Zhang, N., Li, X., & Aebersold, R. (2005). A uniform proteomics MS/MS analysis platform utilizing open XML file formats. *Molecular systems biology*, *1*, 2005.0017.

Kemp, C., & Imler, J.-L. (2009). Antiviral immunity in drosophila. *Current opinion in immunology*, *21*(1), 3–9.

Kitagishi, Y., Okumura, N., Yoshida, H., Nishimura, Y., Takahashi, J., & Matsuda, S. (2011). Long-term cultivation of in vitro Apis mellifera cells by gene transfer of human c-myc proto-oncogene. *In vitro cellular & developmental biology. Animal*, *47*(7), 451–3.

Kreissl, S., & Bicker, G. (1992). Dissociated neurons of the pupal honeybee brain in cell culture. *Journal of neurocytology*, *21*(8), 545–56.

Lemaitre, B., & Hoffmann, J. (2007). The host defense of Drosophila melanogaster. *Annual review of immunology*, *25*, 697–743.

Lundberg, A., & Hahn, W. (2000). Genes involved in senescence and immortalization. *Current opinion in cell biology, 12*, 705–709.

Lynn, D. (2007). Available lepidopteran insect cell lines. *Baculovirus and insect cell expression protocols*, *338*, 117–137.

Lynn, D. E. (2001). Novel techniques to establish new insect cell lines. *In vitro cellular & developmental biology. In vitro cellular & developmental biology. Animal, 37*(6), 319–21.

Maori, E., Lavi, S., Mozes-Koch, R., Gantman, Y., Peretz, Y., Edelbaum, O., Tanne, E., et al. (2007). Isolation and characterization of Israeli acute paralysis virus, a dicistrovirus affecting honeybees in Israel: evidence for diversity due to intra- and inter-species recombination. *The journal of general virology*, *88*(Pt 12), 3428–38.

Maori, E., Tanne, E., & Sela, I. (2007). Reciprocal sequence exchange between non-retro viruses and hosts leading to the appearance of new host phenotypes. *Virology*, *362*(2), 342–9.

Mitsuhashi, J., & Shozawa, A. (1985). Continuous Cell Lines from Larval Hemocytes of the Cabbage Armyworm, Mamestra brassiae. *Development, growth & differentiation, 27*(5), 599–606.

Mohammed, S., & Heck, A. (2011). Strong cation exchange (SCX) based analytical methods for the targeted analysis of protein post-translational modifications. *Current opinion in biotechnology*, *22*(1), 9–16.

Mueller, L., & Brusniak, M. (2008). An assessment of software solutions for the analysis of mass spectrometry based quantitative proteomics data. *Journal of proteome research, 7*, 51–61.

Nakahara, Y., Kanamori, Y., Kiuchi, M., & Kamimura, M. (2010). Two hemocyte lineages exist in silkworm larval hematopoietic organ. *PloS one*, *5*(7), e11816.

Nakashima, N., & Uchiumi, T. (2009). Functional analysis of structural motifs in dicistroviruses. *Virus research*, *139*(2), 137–47.

Neuman-Silberberg, F. S., Schejter, E., Hoffmann, F. M., & Shilo, B. Z. (1984). The Drosophila ras oncogenes: structure and nucleotide sequence. *Cell*, *37*(3), 1027–33.

Ong, S. E., & Mann, M. (2005). Mass spectrometry-based proteomics turns quantitative. *Nature chemical biology*, *1*(5), 252–62.

Pacheco, A., & Martinez-Salas, E. (2010). Insights into the biology of IRES elements through riboproteomic approaches. *Journal of biomedicine & biotechnology*, *2010*, 458927.

Pedrioli, P. G., Eng, J. K., Hubley, R., Vogelzang, M., Deutsch, E. W., Raught, B., Pratt, B., et al. (2004). A common open representation of mass spectrometry data and its application to proteomics research. *Nature biotechnology*, *22*(11), 1459–66.

Peng, J., & Elias, J. (2003). Evaluation of multidimensional chromatography coupled with tandem mass spectrometry (LC/LC-MS/MS) for large-scale protein analysis: the yeast proteome. *Journal of proteome research*, (2), 43–50.

Pfeifer, T., Hegedus, D. D., Grigliatti, T., & Theilmann, D. a. (1997). Baculovirus immediate-early promoter-mediated expression of the Zeocin resistance gene for use as a dominant selectable marker in dipteran and lepidopteran insect cell lines. *Gene*, *188*(2), 183–90.

Pinheiro, V. B., & Ellar, D. J. (2006). How to kill a mocking bug? *Cellular microbiology*, *8*(4), 545–57.

Potts, S. G., Biesmeijer, J. C., Kremen, C., Neumann, P., Schweiger, O., & Kunin, W. E. (2010). Global pollinator declines: trends, impacts and drivers. *Trends in ecology & evolution*, *25*(6), 345–53.

Rachinsky, A., & Hartfelder, K. (1998). In vitro biosynthesis of juvenile hormone in larval honey bees: comparison of six media. *In vitro cellular & developmental biology. Animal, 34*(8), 646–648.

Rappsilber, J., Ishihama, Y., & Mann, M. (2003). Stop and go extraction tips for matrix-assisted laser desorption/ionization, nanoelectrospray, and LC/MS sample pretreatment in proteomics. *Analytical chemistry*, *75*(3), 663–70.

Ribeiro, C., & Brehélin, M. (2006). Insect haemocytes: what type of cell is that? *Journal of insect physiology*, *52*(5), 417–29.

Runckel, C., Flenniken, M. L., Engel, J. C., Ruby, J. G., Ganem, D., Andino, R., & DeRisi, J. L. (2011). Temporal analysis of the honey bee microbiome reveals four novel viruses and seasonal prevalence of known viruses, Nosema, and Crithidia. *PloS one*, *6*(6), e20656.

Scherfer, C., Karlsson, C., Loseva, O., & Bidla, G. (2004). Isolation and Characterization of Hemolymph Clotting Factors in Drosophila melanogaster by a Pullout Method. *Current biology*, *14*, 625–629.

Schneider, I. (1972). Cell lines derived from late embryonic stages of Drosophila melanogaster. *Journal of embryology and experimental morphology, 27*(2), 353–365.

Schüler, M., Connell, S. R., Lescoute, A., Giesebrecht, J., Dabrowski, M., Schroeer, B., Mielke, T., et al. (2006). Structure of the ribosome-bound cricket paralysis virus IRES RNA. *Nature structural & molecular biology*, *13*(12), 1092–6.

Shay, J. W., Wright, W. E., & Werbin, H. (1991). Defining the molecular mechanisms of human cell immortalization. *Biochimica et biophysica acta*, *1072*(1), 1–7.

Smagghe, G., Goodman, C. L., & Stanley, D. (2009). Insect cell culture and applications to research and pest management. *In vitro cellular & developmental biology. Animal*, *45*(3-4), 93–105.

Stewart, S., & Weinberg, R. a. (2000). Telomerase and human tumorigenesis. *Seminars in cancer biology*, *10*(6), 399–406.

Stoltz, D., Shen, X., Boggis, C., & Sisson, G. (1994). Molecular diagnosis of Kashmir bee virus infection. *Journal of apicultural research, 34*(3), 153-160.

Todd, J., Miranda, J. De, & Ball, B. (2007). Incidence and molecular characterization of viruses found in dying New Zealand honey bee (Apis mellifera) colonies infested with Varroa destructor. *Apidologie*, *38*, 354–367.

Van Rij, R. P., Saleh, M.-C., Berry, B., Foo, C., Houk, A., Antoniewski, C., & Andino, R. (2006). The RNA silencing endonuclease Argonaute 2 mediates specific antiviral immunity in Drosophila melanogaster. *Genes & development*, *20*(21), 2985–95.

VanEngelsdorp, D., & Meixner, M. D. (2010). A historical review of managed honey bee populations in Europe and the United States and the factors that may affect them. *Journal of invertebrate pathology*, *103*, S80–95.

Vodovar, N., & Saleh, M. (2012). Of Insects and Viruses: The Role of Small RNAs in Insect Defence. *Advances in insect physiology*, *42*.

Walther, T. C., & Mann, M. (2010). Mass spectrometry-based proteomics in cell biology. *The Journal of cell biology*, *190*(4), 491–500.

Wang, D., & Moeller, F. (1970). Comparison of the free amino acid composition in the hemolymph of healthy and Nosema-infected female honey bees. *Journal of invertebrate pathology*, *15*, 202–206.

Wang, X., Aliyari, R., Li, W., & Li, H. (2006). RNA interference directs innate immunity against viruses in adult Drosophila. *Science*, *312*(April), 452–454.

Weinstock, G. M., Robinson, G. E., Gibbs, R. a., Worley, K. C., Evans, J. D., Maleszka, R., Robertson, H. M., et al. (2006). Insights into social insects from the genome of the honeybee Apis mellifera. *Nature*, *443*(7114), 931–949.

Wilson, J. E., Pestova, T. V, Hellen, C. U., & Sarnow, P. (2000). Initiation of protein synthesis from the A site of the ribosome. *Cell*, *102*(4), 511–20.

Wojcechowskyj, J., Didigu, C., Lee, J. Y., Parrish, N. F., Sinha, R., Hahn, B. H., Bushman, F. D., et al. (2013). Quantitative Phosphoproteomics Reveals Extensive Cellular Reprogramming during HIV-1 Entry. *Cell host & microbe*, *13*(5), 613–23.

Yates, J. R., Ruse, C. I., & Nakorchevsky, A. (2009). Proteomics by mass spectrometry: approaches, advances, and applications. *Annual review of biomedical engineering*, *11*(c), 49–79.

Yeager, T. R., & Reddel, R. R. (1999). Constructing immortalized human cell lines. *Current opinion in biotechnology*, *10*(5), 465–9.

Zufelato, M. S., Lourenço, A. P., Simões, Z. L. P., Jorge, J., & Bitondi, M. M. G. (2004). Phenoloxidase activity in Apis mellifera honey bee pupae, and ecdysteroid-dependent expression of the prophenoloxidase mRNA. *Insect biochemistry and molecular biology*, *34*(12), 1257–68.