The University of Nottingham
UNITED KINGDOM · CHINA · MALAYSIA

Snart, Charles J. P. (2011) Transcriptome analysis of honey bee larvae following neonicotinoid exposure. MRes thesis, University of Nottingham.

Access from the University of Nottingham repository:
http://eprints.nottingham.ac.uk/12333/2/Transcriptome__Analysis_of_Honey_Bee_Larvae_following_Neonicotinoid_Exposure.pdf

Copyright and reuse:

The Nottingham ePrints service makes this work by researchers of the University of Nottingham available open access under the following conditions.

This article is made available under the University of Nottingham End User licence and may be reused according to the conditions of the licence. For more details see: http://eprints.nottingham.ac.uk/end_user_agreement.pdf

A note on versions:

The version presented here may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the repository url above for details on accessing the published version and note that access may require a subscription.

For more information, please contact eprints@nottingham.ac.uk

Transcriptome Analysis of Honey Bee Larvae following Neonicotinoid Exposure

Charles Snart

Dissertation submitted to the University of Nottingham in partial fulfilment of the requirements for the degree of Masters of Research with in Advanced Genomics and Proteomics.

.

Acknowledgements:

I would like to thank my supervisor, Dr Reinhard Stoger, for the help and guidance that he has offered through this investigation, my postdoc Dr Kamila Derecka for aid in learning and carrying out the techniques described with, alongside a great deal of help in interpreting and presenting my results. I would also like to thank Jim Craigon for his help in selecting and applying statistical analysis to these results, and all the staff at Genomnia for their help in producing and interpreting the transcriptomic data.

Abstract:

The current decline of the European Honey Bee (*Apis Mellifera*) has been linked to the increasing use of neonicotinoid pesticides within agriculture. Whilst the toxicity of these pesticides to *Apis* has long been established, the possibility of low dosages inducing molecular stress has not yet been fully explored. Of particular interest is the action of these nicotine derivatives on the nicotinic acetylcholine receptor, and its association with the DNA methyltransferase family (Dnmts).

An experimental group of three hives were exposed to sugar water contaminated with a low concentration of imidacloprid (2µg/l). From these hives, 12 third instar larvae were selected. A corresponding number of larvae were also selected from three control hives, for a total of 12 samples. Using quantitative reverse transcription-PCR, known Dnmt transcripts were detected and amplified from these larval samples. Specially designed oligonucleotide primers were used containing gene specific sequences that linked to universal DNA sequences, ensuring that PCR amplification products were of predetermined sizes. Products of this amplification were resolved by capillary electrophoresis and detected by fluorescence spectrophotometry. Simultaneously, the transcriptomes of 3 larval samples each from the control and experimental groups were generated using SOLiD platform sequencing. The statistical package EdgeR was then utilised to identify differential candidates of known honey bee microRNA's.

Statistical analysis utilising a one-way Analysis of Variance (ANOVA) found no significant differences in the expression levels of known Dnmt transcripts between control and experimental groups. However, comparisons of sequenced control and experimental transcriptomes identified a number of differential microRNA candidates, most notably miR-9/14.

Table of Contents

2: Materials and Methods — Page Number

3: Results — Page Number

1: Introduction

1.1 An Introduction to the European Honey Bee (*Apis mellifera*):

1.1.1 Characterisation and Distribution:

The honey bees are a subgroup of bees characterised by a high degree of eusociality, production and storage of honey, and the construction of a protective nest from wax secretions. Classified within the genus '*Apis*', the group appears to originate within south eastern Asia (Smith et al., 2000), as all but one of the extant lineages has been confirmed as native to the area (Engel et al., 2000). Despite this, the first fossil records of Apis bees originate from the Eocene-Oligocene boundary (approx. 33 million years ago), and were obtained from European deposits. However, this implies only that the species was present in this location during the time period, and not a European origin (Arias and Sheppard, 2005). Recent discoveries have indicated that a separate species, *Apis Linnaeus*, was native to the western hemisphere during the Miocene, based from a single 14 Million year old worker specimen preserved in paper shale from Nevada (Engel et al., 2000).

Though noted for their highly sophisticated level of social behaviour, it should be observed that this level of sociality predates the origin of the genus (Engel et al., 1999). This is indicated by the conserved, though lower, level of sociality apparent in the honey bees extant relatives, the stingless bees and the bumblebees.

1.1.2 Taxonomy and Classification:

The honey bees comprise the genus *Apis*, and are the only extant members of tribe Apini. Alongside the other bee species, they form the family Apidae, one of the four families of the order Hymenoptera (alongside sawflies, wasps, and ants). There are three extant clades present within the honey bee*;* the dwarf honey bees (*Micrapis*), the giant honey bees (*Megapis*), and the common honey bees (*Apis*) (Arias and

Sheppard, 2005). Of these, *Micrapis* is the most ancient lineage, with an estimated divergence from previous extinct lineages around 40 Ma. These clades are often further differentiated into two groups, the open nesting bees (*Micrapis* and *Megapis*) and the cavity nesting bees (*Apis*) (Arias and Sheppard, 2005). This distinction is due to *Micrapis* and *Megapis* producing much less extensive hive formations. *Micrapis* produces small, exposed nests in trees and shrubs, whereas *Megapis* will often produce a very small number of exposed combs on cliffs or high tree limbs (Chapman et al., 2009).

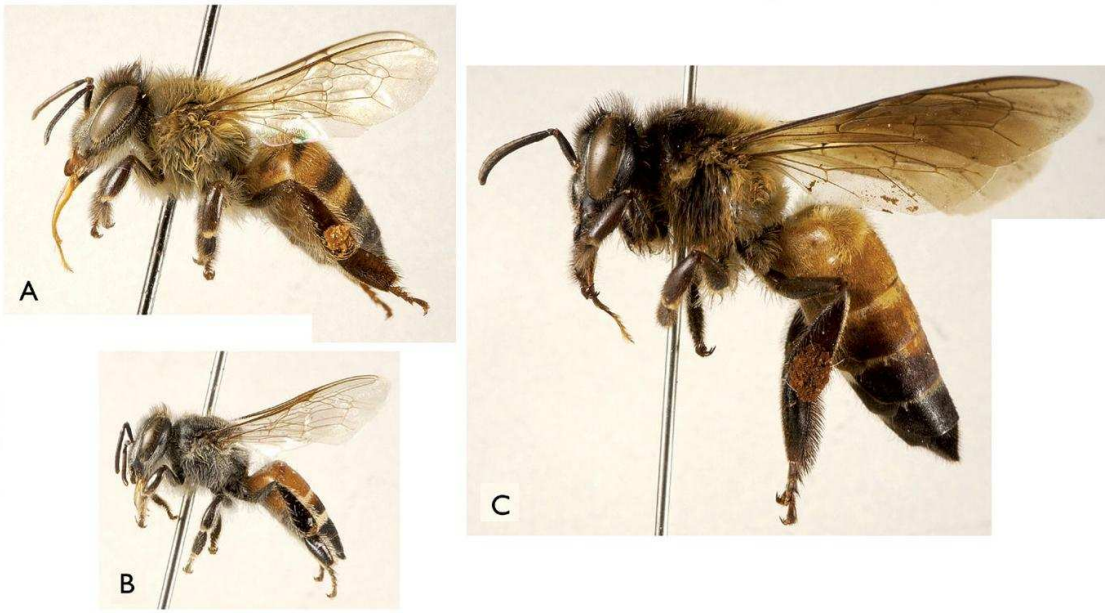

Figure 1. Comparison of workers across honey bee clades. **A** *Apis mellifera* Linnaeus **B** *Apis florea* Fabricius **C** *Apis dorsata* Fabricius (Kotthoff et al., 2011).

Micrapis is notable for its high level of evolutionary distinction from the other clades. Though occurring sympatrically alongside other wild honey bees, this distinction indicates that they likely evolved as a result of allopatric speciation, with their ranges and environments eventually overlapping. Due to its minute size, little protection is required from the species during hive examination/harvesting, though the latter activity may be unlikely due to a lack of significant honey stores. By contrast, *Megapis* has historically, received a far greater level of harvesting, however, this is tempered by the fierce nature of the species. Due to greater size and aggression, a

large degree of protection would be required for harvesting, as a provoked swarm would be quite capable of killing a human being.

The Apis bees have received the greatest amount of domestication, and are widely used throughout agriculture for both direct honey production, and pollination of flowering crops. The current consensus is that this group comprises two major species, with a disputed number of minor species:

- *Apis mellifera* – also known as the European honey bee, mellifera is the main species utilised within western apiculture. Widely considered to have diverged from an eastern ancestor, it has been managed for millennia for human benefit. The third insect to have its genome fully sequenced, *mellifera* may prove to be an important insect model for epigenetic mechanisms. It is by far the most researched of the honey bees, and is the focus of both this investigation, and many others.

- *Apis cerana* – known as the eastern honey bee, *cerana* occupies the role of *mellifera* in southern and eastern asia. Eight separate subspecies of *cerana* are known to exist, though its relationship with the Bornean honey bee (*Apis cerana nuluensis*), and the Philippine honey bee (*Apis nigrocincta*) remains difficult to resolve.

A recent hybrid strain, the Africanised honey bee (well publicised as the 'killer bee') was accidentally produced by crossing the European species with an African subspecies (Schumacher and Egen, 1995). These bees are marked by greater aggression and disease resistance, and have spread rapidly since their release (Winston, 1994). In many regions these bees are listed as a pest, but hybrids between European bees and Africanised bees have resulted in high yield, and disease resistant colonies (Mitchell, 2006).

1.1.3 Nutrition:

Honey bees feed almost exclusively on nectar and pollen gathered from flowering plants. This nectar is harvested and digested by worker bees, converting it into respective forms of brood feed. A simple, largely undigested form (honey) is utilised in order to rear future workers, whereas a specialised glandular secretion known as royal jelly is utilised to rear prospective queens. Pollen is converted into what is known as 'bee bread', which is consumed in large quantities by both larvae and adult workers. This quantity is essential, due to difficulties in its digestion (Parker, 1926).

Nectar:

Nectar is an aqueous plant secretion with varying properties, comprised of anywhere between 5 to 80% sugars, alongside small quantities of vitamins, organic acids, lips, minerals, and aromatic substances (White, 1975). Comprised of the three sugars; sucrose, glucose, and fructose, nectar can be classified into varying groups depending on their individual proportions (White, 1979). Whilst it is possible to feed nectar directly to both adults and the brood (Gary, 1975), workers usually process it into either honey or royal jelly. The nectar is first transported back to the nest in the honey stomach, before being transferred to nurse worker bees for digestion. Enzymes are then secreted from the hypopharyngeal glands into the crop, where the nectar is broken down into simpler inverted forms. This more digestively accessible compound is then gradually reduced in water content through evaporation on the workers tongue, before being deposited into a brood cell. Alternatively, if the produced honey is not needed for brood rearing at the time of production, it will be stored in a wax sealed cell for later consumption. This storing process decreases the nectar water content to less 20%, protecting the honey from any potential yeast contamination.

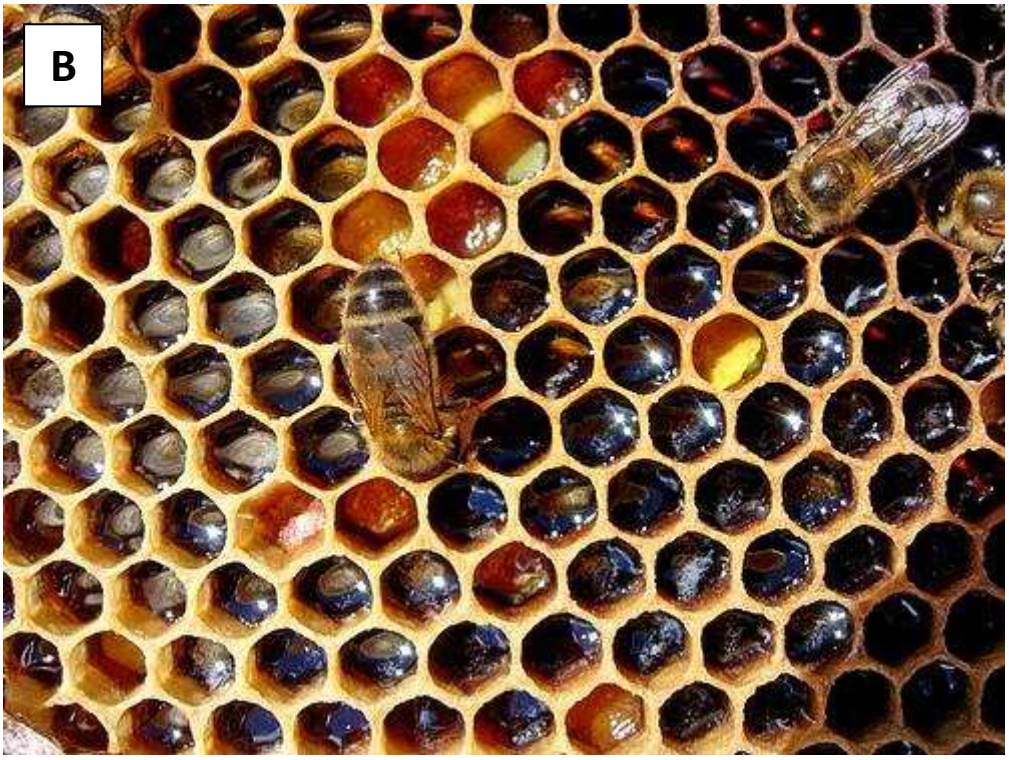

Figure 2. Nectar collection and storage. **A** Nectar extraction (European honey bee extracts nectar) **B** Honey and nectar storage in open cells (Worker/nurse bees storing nectar in honeycomb)

Pollen:

Whist nectar is the main carbohydrate source of the honey bee diet; pollen is the only source of protein that is regularly available to the honey bee. Pollen is comprised of the plant's male germplasm, designed primarily as a travelling fertiliser, whilst also appealing to potential insect vectors. Whilst various other nutritionally valuable compounds are present in harvested pollen, the variability of these

components across different plant species makes it difficult to establish their relative dietary importance.

Before initial processing, pollen is treated by workers to prevent germination and bacterial activity.

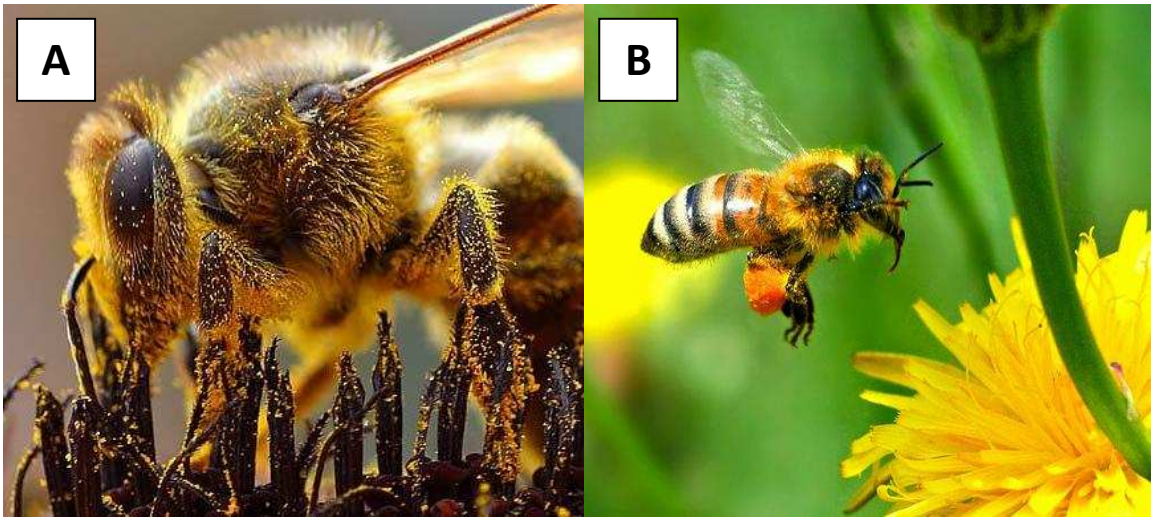

Figure 3. Worker pollen collection. **A** Worker coated in pollen (What is bee pollen used for?) **B** Worker displaying full pollen sac. Utilisation of pollen as a food source results in the honey bee acting as a natural vector for pollination. During collection, workers store pollen in a small sac-like organ on their hind legs, known as the pollen sac (Save the bees – a global crisis).

This is accomplished through the addition of phytocidal acid, which is also produced from the hypopharyngeal glands (Pain and Maugenet, 1966). Long term storage is dependent on this treatment, as the contained enzymes prevent anaerobic breakdown and fermentation. As a result, pollen is stored in its pre-processed form, unlike nectar. This is most likely due to inherent challenges involved in pollen digestion itself as its wall components are indigestible by bees. Digestion takes place within the midgut, where protective enzymes are secreted to cope with the abrasive nature of the pollen wall (Barker and Lehner, 1972). Osmotic shock is then utilised to rupture the pollen wall, at which point the internal components are digested (Klungness and Peng, 1984). Each worker larva has been shown to require approximately 125mg of pollen, comprising around 30mg of protein (Rosov, 1944).

Royal Jelly:

Royal jelly is a separate, milky white component that is secreted from the mandibular glands of nurse workers. Fed in significant quantities (20-40% of brood feed) to worker larvae for the first two days of development, and continued throughout the development of queen larvae, royal jelly is an integral component in female honey bee differentiation. The effects of royal jelly will be reviewed in greater detail in the section on Reproduction and Development.

Figure 4. Developing queen larvae immersed in royal jelly.

1.1.4 Caste System:

An individual honey bee colony acts as a super organism, undergoing a level of selection as a single unit that is distinct from that of gene and individual (Wilson and Sober, 1989). Throughout this superorganism, a division of labour occurs across three phenotypically distinct castes: Workers, Drones, and Queens. This division is an integral part of the eusocial superorganism, as individual altruism provides for maximisation of production at a colony level.

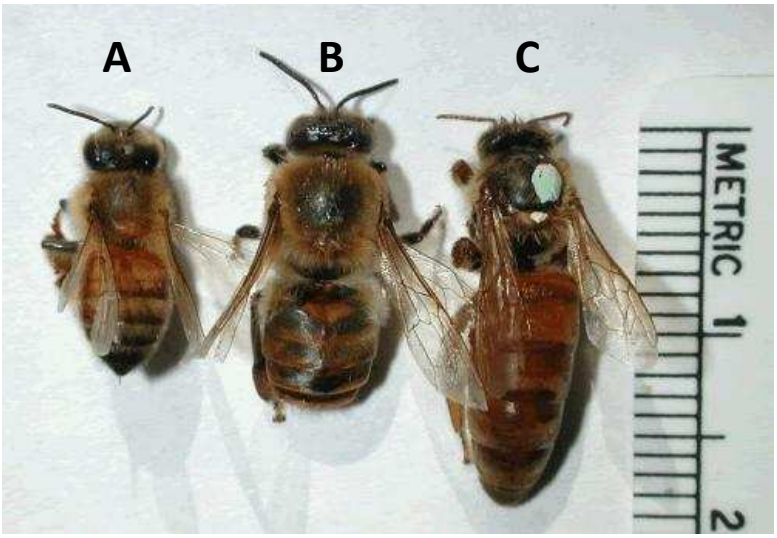

Figure 5. Honey Bee caste comparisons. **A** Worker **B** Drone **C** Queen. The haploid male drone can be easily distinguished from the other female castes by their bulbous head and enlarged body. Within the diploid castes, workers can be distinguished from queens by the queen's enlarged thorax/ovipositor (Honey bee castes).

Workers perform the vast majority of manual labour, being responsible for foraging, hive maintenance, brood rearing, care of the queen, and food production. This caste consists of sterile females, with differing theories being produced for task allocation within the group. The degree of physiological specialisation being required for different tasks, led to the Caste Concept, that under stable colony conditions, workers are organised into specialist groups based on age and size (Seeley, 1982) However subsequent studies indicated a large amount of task switching occurring even in stable colonies. This cast some doubt on previous conclusions of a caste system, with Gordon (1989) suggesting that task group is a more appropriate term than caste, as it does not imply a stable specialisation. Despite this, it remains applicable for certain physiologically specified groups. A good example of this is the distinction between nurse bees and foragers, as foragers do not possess the nurse bee's active hypopharyngeal glands, and are thus incapable of producing brood food (Johnson, 1995). Any re-specification from forager to nurse would as a result require extensive glandular regeneration. Alongside physiology, other possible caste

differences may be founded on morphological differences, or differential information supply for task learning (Johnson, 1995).

The queen caste performs the paramount role within the colony, that of reproduction. Most hives will have only one fertile queen, who is constantly attended to by a specialised group of workers known as nurse bees. In a fully functioning hive, queen rearing is rare, but the rate of replacement upon the death or infertility of the existing queen is remarkable. The reliable rearing of a new, fully fertile queen can be accomplished in as little as two weeks since the disappearance of the previous queen, or in anticipation of the colony splitting.

Drones consist of the only caste of male bees, and have little function within the hive other than to mate with new queens. Possessing no pollen baskets, wax glands, or the ability to produce royal jelly, drones are entirely dependent on worker food provision for survival, lacking a stinger; they provide little in terms of colony defence. Though originally thought to only mate once, it is now known that queens will mate with five to six drones. These encounters usually take place on the wing, and although drones die shortly after, their highly motile sperm soon arrives in the spermatheca, a small organ in the queen's abdomen. Here it remains viable for the rest of the queen's reproductive lifespan. Though highly prevalent throughout the hive during spring months, workers usually drive out any remaining drones in the autumn months, preceding the annual winter cull.

1.1.5 Development and Reproduction:

Insects are unique in the degree of their utilisation of polyphenisms, the production of two or more distinct adult phenotypes, with distinct differences in both behaviour and morphology. In hymenoptera, this distinction is usually tied to the concept of haplodiploidy, with the number of chromosomes possessed by an individual determining its adult phenotype (Tyvand and Thorvaldsen, 2010).

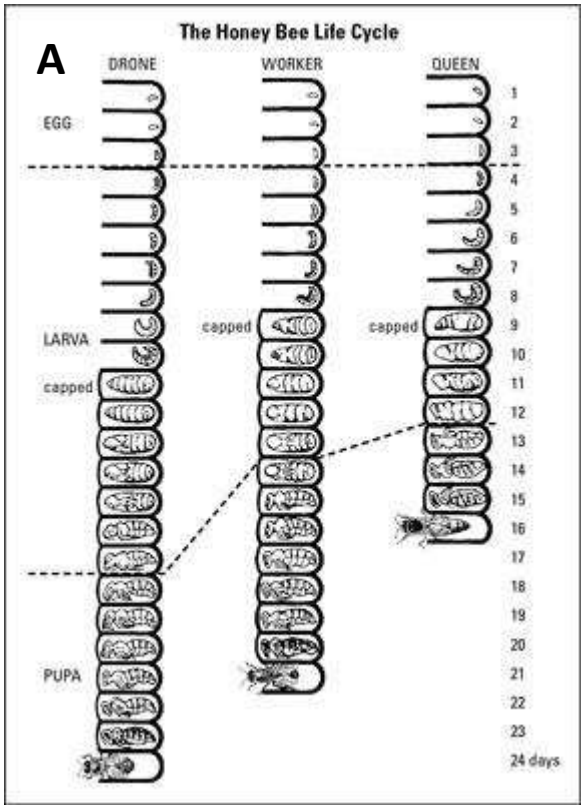

Figure 6. Larval development. **A** Individual caste larval development (The honey bee life cycle) **B** Progression cycle from egg to larvae to pupa. Honey bees undergo five major larval instars (L) (distinct periods of development), within which they undergo significant metamorphosis (The bee as a pupa). At the end of each instar, the developing larva sheds its skin, with the exception of the final instar. Whilst drone fate is set by their haploid genome, the final caste of diploid larvae is not set until the L2/L3 transition period. From this point onwards, final caste fate is set, though it is possible to produce queen-like individuals through interference. These individuals display significantly lower levels of ovariole maturation.

In the honey bee, as in most eusocial hymenoptera, an environmental component also exists, specifically in relation to queen rearing. The distinction between worker

and drone, or male and female, is largely genetic, with haploid males and diploid females being reared from unfertilised and fertilised eggs respectively (Cho et al., 2006). Whilst the male development into the drone caste is from this point set, the eventual adult phenotype of the female relies on larval nutrition. Queen destined larvae receive a highly specialised secretion from the mandibular gland of attendant workers, known as royal jelly (Colhoun and Smith, 1960). Though largely uncharacterised, this substance has been shown to stimulate wide scale epigenetic remodelling in exposed larvae. By comparison, the eventual worker larvae only receive royal jelly exposure for two days, before being switched to standard brood feed. From this point, developing queen larvae display both an increased weight gain when compared to developing worker larvae, due to a much accelerated rate of feeding. The expression of Juvenile Hormone (JHE) is also noticeably elevated in these larvae. In the honey bee, ovariole development and maturation is highly linked to the level of JHE produced by the larvae. Juvenile hormone is also utilised to a degree by workers, to control division of labour across different age groups (Hagenguth and Rembold, 1978).

A recent study by Kamakura (2011) has indicated that Royalactin, a 57-kDa protein within royal jelly, is responsible for inducing queen development. However, it also surprisingly indicated that *Drosophila melanogaster* experiences similar changes in development, suggesting the possibility of a conserved differentiation trigger. Treatment with Royalactin resulted in an increase in body size, due to the activation of p70 S6 kinase and mitogen-activated protein kinase within the larval fat body. This process was modulated by Egfr iRNA, in the fat body, which was shown to be induced upon Royalactin application.

1.1.6 The Worker Caste and the Division of Labour:

The honey bee worker caste operates alongside an age related behavioural development system in adult workers (Winston, 1987). Younger bees work within the hive within the first 2-3 weeks of life, performing hive maintenance and working behaviour (Sullivan et al., 2003). Throughout this period, younger bees also take part in orientation flights, in order to learn hive location and prominent local landmarks (Capaldi et al., 2000). After this early role, older bees transfer to active nectar foraging, which typically continues for the rest of their 4-6 week lives (Sullivan et al., 2003). Much of this regulation is due to the action of Juvenile hormone (JHE). Alongside its previously detailed metamorphic role in development and reproduction, JHE also has a significant regulatory role in behavioural development within insects (Wyatt and Davey, 1996).

Figure 7. Nurse bees attending queen. Younger honey bees act as nurses for the queen, providing sustenance and aiding in egg placement, before transferring to foraging behaviour as they age. The change in behaviour is highly correlated with an overall change in the concentration of juvenile hormone that is produced.

Of the known JHE's, Juvenile Hormone III is the only one currently known to be present within the honey bee (Hagenguth and Rembold, 1978). It is synthesised by the corpora allata (CA), a small, glandular pair of bodies located alongside the

foregut (Gäde et al., 1997). Disruption or incorrect function of the CA has been shown to contribute to significant behavioural impairments in adult workers, particularly affecting orientation during early foraging flights (Sullivan et al., 2000). After receiving an allatectomy, young workers would often vanish during onset of flight activity, apparently unable to relocate the hive. Two plausible suggestions were put forward as a result of this, that either workers were unable to learn or remember the hive location, or that there was a significant deficit in physical aspects related to flight. Due to the correlation of the application of JHE and an increase in cytochrome oxidase activity and protein synthesis in muscle wing mitochondria, it could be concluded that an allatectomy would result in a decrease in the workers capacity for flight. This would leave workers physically unable to make the return journey, rather than incapable of locating the hive.

1.1.7 Communication:

It has long been noted that bees possess a highly developed communication system. The celebrated ethologist Karl von Frisch first observed non-primitive communication between honey bee workers in 1923, describing what has now become known as the ‘waggle dance’. Workers returning from new food sources were observed to report their findings through an intricate series of tail waggling movements. It was hypothesised that this, combined with the scent of pollen both on the worker and its collected nectar, was utilised to communicate the presence, quality and location of a food source. This dancing action was shown to only occur if the discovered food source was in abundance, with dancing rapidly decreasing when the new source becomes depleted. Several patterns were also established within the dance itself, with relatively close sources (50-100 metres) being communicated through the use of circular dances, with actual tail-waggling only being utilised in the case of sources further afield.

Later observers have since deduced a more sophisticated level of communication, noting subtle differences between dances that allow the generation of much more specific quantitative information. This quantification allows the valuing of resources, based on factors, such as the current nutritional needs of the colony, and the potential cost-benefit ratio of the foraging trip (Waddington, 1982). The role of pheromones in this recruitment has also been established (Thom et al., 2007). Through the use of gas chromatography paired with mass spectrometry, it was indicated that two specific alkanes, tricosane and pentacosane, are released from the abdomens of returning workers. This release has a significant effect on worker behaviour, increasing the number of foragers that leave the hive. By comparison, non-dancing workers returning from the same food source only produce very limited amounts of these alkanes, indicating a chemical element in recruitment behaviour (Thom et al., 2007).

1.1.8 Human Cultivation:

Despite the presence of over 20,000 wild bee species, *Apis mellifera* is almost universally utilised for western apiculture. This distinction is due to the highly social nature of *mellifera*, which dwells in colonies large enough for the necessary levels of honey production. By comparison, many wild species are solitary by nature (e.g. Mason Bees) (Cardinal and Danforth, 2011), or dwell in colonies too small for honey exploitation (e.g. Bumblebees). This preference for sociality also exists within tropical honey production, which often utilises the Asian honey bee, *Apis cerana*. On top of these established species, several hybrids, such as the Buckfast bee, have been produced in order to enhance favourable production traits (e.g. increased honey yield, disease and parasite resistance, and more prolific breeding) (Danka and Villa, 2000).

The collection of honey from wild bee colonies is one of the most ancient activities still performed by humans, and its practise still exists at a varying level amongst

aboriginal tribes across the world. Its usage is depicted in numerous cave paintings, the earliest dating to around 13,000BC. Later settled communities began the first domestications through the use of artificial hives, including wooden boxes, hollowed logs, and porcelain containers. The presence of apiculture within ancient Egypt is well documented, with pots of honey being contained within the tomb of Tutankhamun. A form of high class culture was also prevalent within ancient

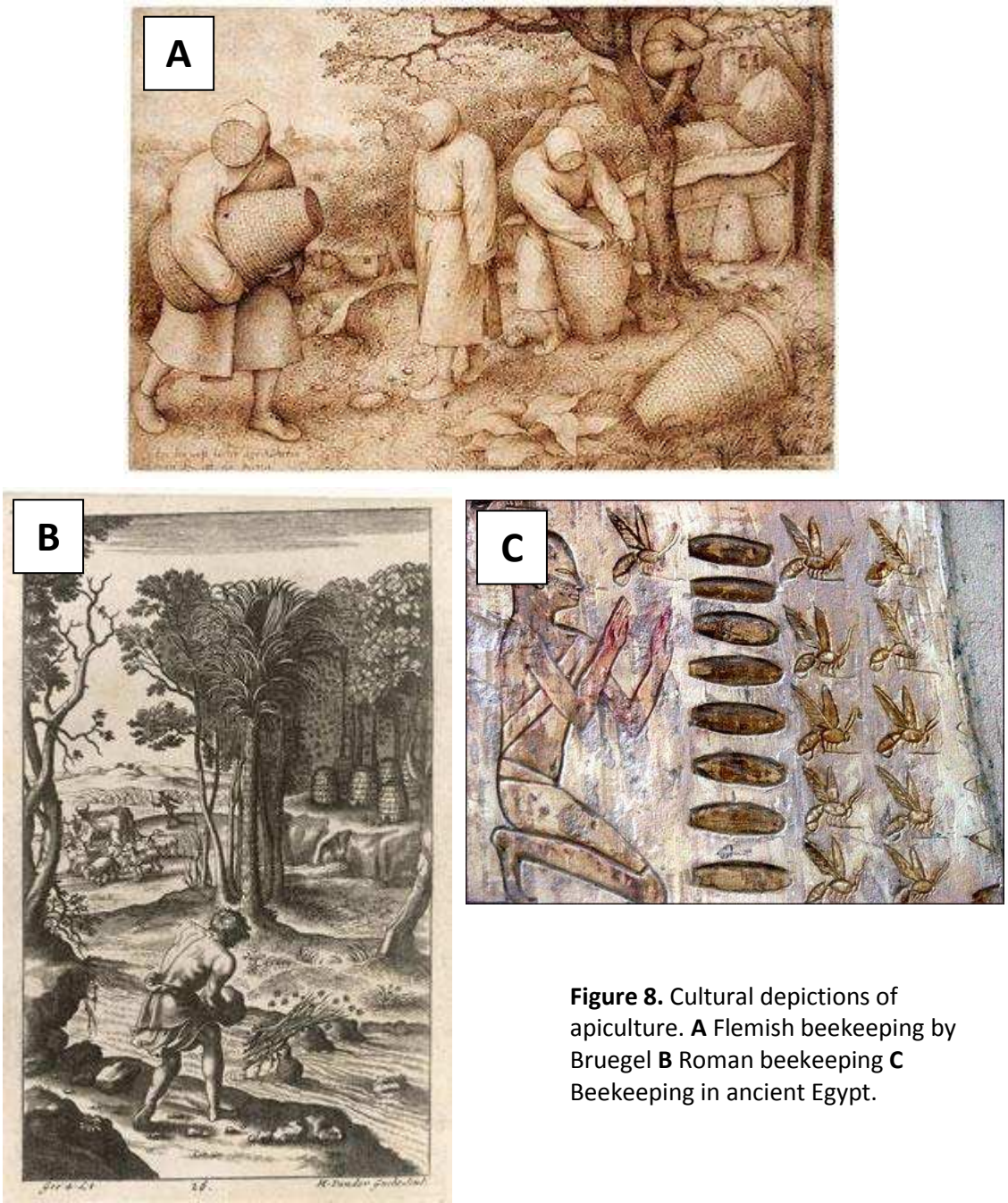

Figure 8. Cultural depictions of apiculture. **A** Flemish beekeeping by Bruegel **B** Roman beekeeping **C** Beekeeping in ancient Egypt.

Greece; the existence of gold rings engraved with depictions of bee keeping indicates its high cultural status.

Apiculture was widely utilised throughout the Roman and medieval periods, but it was not until the 18th century that Europeans began to undertake scientific study of beekeeping. The most significant of these early researchers was Francois Huber. Huber was the first to examine the internal workings of colonies, and to observe physical insemination of young queens on the wing (Huber, 1814). The most groundbreaking of these discoveries, was the determination of 'bee space', the required distance between frames to prevent gaps from being filled with a combination of honey and wax (Huber, 1814). This, combined with the removal frames designed by Lorenzo Lorraine Langstroth, allowed honey production on a much larger, sustainable level (Langstroth, 1853). As a result, it was now possible to harvest honey and wax without previous extermination of the bee population, allowing long term colony use.

1.1.9 Agricultural Role and Importance:

Agricultural economy heavily depends on efficient pollination of plants and crops, with approximately 35% of farmed crops requiring some form of pollination. Within western agriculture, *Apis mellifera* is the most widely utilised pollinator, supplying roughly 90% of flowering crops (Aizen and Harder, 2006). Whilst it is likely that in the event of the species extinction, another pollinator would eventually rise to fill its ecological niche, the cost of such a sudden extinction to worldwide agriculture would be catastrophic. Whilst such a disappearance would not necessarily bring about mankind's own extinction, it is clear that such an event would leave the global diet severely impoverished.

Figure 9. Modern Beekeeping. Due to utilisation of modern techniques, honey production/pollination has grown into a mass industry. The widespread use of the removable frame, alongside the application of smoke treatment has allowed individual hives to remain sustainable from year to year. Smoke treatment is mainly utilised as a calming agent, as it causes workers to react as if a real fire is present. In this situation individual workers will proceed to the colony stores of nectar/brood feed, where they will gorge in an attempt to retrieve as much as possible in the event of hive destruction.

1.2 Colony Collapse Disorder:

1.2.1 An Introduction:

Ever since the late early 1960's, the western honey bee population has been subject to a significant decline and general stagnation. Whilst the global hive stock is still displaying significant growth worldwide (~45% in the last half century), this increase is not keeping pace with the current growth in agricultural demand (Aizen and Harder, 2009). More alarmingly, whilst the worldwide population has increased, the number of hives in the USA continues to display a steady decline of around 1.8% (Oldroyd, 2007). Concurrently, there has been a widespread reduction in the abundance and diversity of wild European bees (Biesmeijer et al., 2006), and many studies have detailed a decline in various local pollinators (Winfree et al., 2009). Whilst this long term decline has been blamed on a wide range of factors, public interest has been increased by perceived intensity of this slump beginning from 1991. This interest is correlated with the appearance of the mysterious 'disappearing disease', where viable hives are suddenly abandoned despite the lack of an obvious cause. Widespread within western populations, a number of symptoms of this newly termed Colony Collapse Disorder have been regularly recognised (vanEngelsdorp et al., 2007), and include the following:

- Rapid loss of adult worker bees,
- Few or no dead bees found in the hive
- Presence of immature bees (brood)
- Small cluster of bees with live queen present, and pollen and honey stores in hive.

This rapid vanishing of the main worker population has greatly complicated attempts to study the disorder, as the few remaining dead bees are able to give little information about whole colony health. However, the modified behaviour of the remaining cluster has been noted, displaying a much reduced amount of foraging behaviour (vanEngelsdorp et al., 2007). As these clusters are often comprised of

young adult workers, this trait is not unexpected (Sullivan et al., 2003), but the cause of their reluctance to consume newly presented feed is unknown.

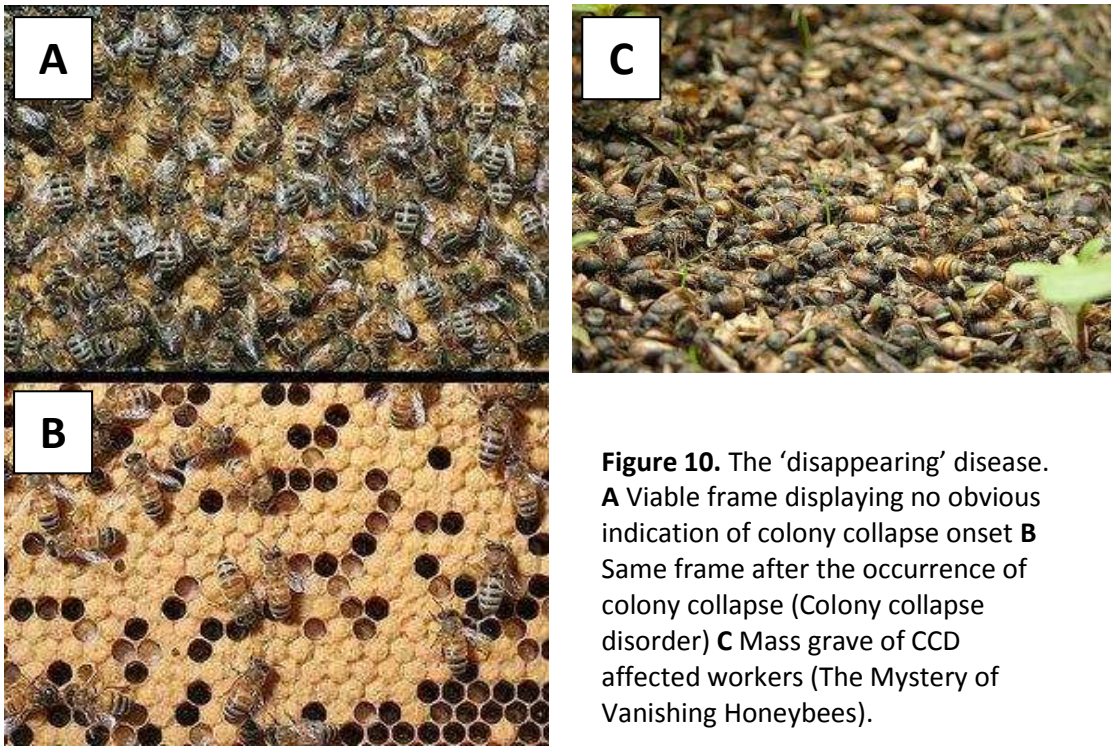

Figure 10. The 'disappearing' disease. **A** Viable frame displaying no obvious indication of colony collapse onset **B** Same frame after the occurrence of colony collapse (Colony collapse disorder) **C** Mass grave of CCD affected workers (The Mystery of Vanishing Honeybees).

The devastating consequences of current colony collapse episodes have had a significant effect on hive levels with Europe and USA. The extreme nature of these losses, combined with the scale of the outbreak, has resulted in a far greater amount of interest than any previous episodes. However it shares many common factors with these previous outbreaks, and their causes may shed some light on current developments.

1.2.2 The 'Disappearing Disease' – A history:

The occurrence of large scale colony collapse may have previously been considered rare, but it was hardly unknown. A very similar disease was noted as far back as the summer of 1869, and was widely referred to as the "May disease". Losses were also described by Aikin (1897), with large Colorado clusters being rapidly reduced to

small, non-viable populations. Several causes were originally suggested, such as the unusually hot summer or a lack of pollen. Later investigations by Burnside (1930) drew a correlation between this May disease, and the presence of a specific strain of the *Aspergillus* fungi. The occurrence of Stonebrood throughout colonies was also shown to be caused by *Aspergillus flavus*, with infected larvae suffering a complete mummification of soft tissues (Burnside, 1930). However, adult worker observations indicate a superficially normal phenotype, with the exception of mass hive desertions seen in other colony collapse incidents (Burnside, 1930). Whilst the onset of the disease was traced to *Aspergillus* comb contamination (Betts, 1919), Giauffret et al. (1967) supported the hypothesis that the rapid spread of the fungi was due to disruption of the bee's intestinal fauna, most likely due to previous attempts to introduce pest controlling antibiotics into their diet.

Three similar outbreaks occurred in Britain between 1905 and 1919. As seen in the aforementioned Stonebrood outbreak, workers afflicted with the disorder crawled a consider distance from the hive entrance before succumbing. However, these workers were incapable of flight (Anderson, 1930), leading to some workers claiming that starvation was the cause of the outbreak (Bailey and Gibbs, 1964). Notably, this instance saw implication of the Nosema disease, with the symptoms being dubbed 'Isle of Wight Disease' (Fantham and Porter, 1912). Work by Bullamore (1922) also indicated a genetic component to the outbreak, with certain colony strains indicating a greater amount of resistance.

Whilst Australia has seen little of the effects of the most recent series of colony collapses, it has also historically suffered from outbreaks. The first of these occurred in 1910, with much more widespread occurrence of the 'disappearing disease' (Oertel, 1965) in the 1970's (Olley, 1976). Both outbreaks were attributed to somewhat similar causes, namely stress, poor food supply, and dampness. The first series in particular were attributed to bees harvesting from *Eucalyptus leucoxylon*, which produced honey unsuitable for consumption (Beuhne, 1910). However, the

second outbreak occurred alongside a similar set of colony losses in Mexico, and the United States. Repeated investigations into their cause were largely inconclusive, with genetics (Kulinčević et al., 1984), viral infection, microorganisms, pesticide poisoning, and pollen contamination all being ruled out as possible factors (Mraz, 1977). Like the current disappearing disorder, no single factor has been conclusively linked to these colony losses, indicating that a combination of the above factors may have been responsible (Kulinčević et al., 1984).

The most recent string of heavy losses is noted to have begun in France during the winters of 1999-2000. The vast majority of infected hives (>70%) were found to contain some form of honey bee disease, though no particular combination of diseases resulted in a significant change in mortality (Faucon et al., 2002). As these losses largely coincided with the implementation of the systemic neonicotinoid pesticide imidacloprid, many attributed its widespread use in agriculture to the decline in honey bee stocks. However, whilst acute treatment with these imidacloprid has shown to result in impairments to communication and orientation flights in gatherers (Bortolotti, 2003), the sub lethal affects of chronic imidacloprid contamination remain largely unexplored (Smirle et al., 1984).

Further complicating the matter is the recent transmission of parasites and pathogens from the asian honey bee, *Apis cerana*. Both the tracheal mite, *Acarapis woodi*, and the varroa mite, *Varroa destructor*, are present within asian honey bee populations, which has developed a natural resistance to them (Peng et al., 1987). However, their recent introduction into North American colonies has had devastating effects (Munoz et al., 2008). Though the North American honey bee has developed a resistance to the tracheal mite, they have been unable to resist the life shortening effects of varroa mite infestation. The emergence of a new strain of *Nosema* has also resulted in non-specific symptoms within hives, such as low honey yield, and greater winter colony losses (Fries et al., 2006).

It is still unknown at this stage whether any of the previously seen disappearing diseases are linked to the current epidemic, despite a passing similarity in both symptoms and possible causes. However, previous theories for colony collapse incidents have often indicated that multiple factors may be combining to cause an overall decline, rather than there being a singular cause.

1.2.3 Colony Collapse Disorder – The Culprit:

So far, three major factors have been linked to the onset of CCD. Whilst none of these have yet been shown to be directly responsible for colony collapse, it is generally accepted that a combination of these factors are contributing the current honey bee hive decline. Specifically, these contributors are (vanEngelsdorp et al., 2010):

- Environmental Stressors, such as Pesticides
- Infectious diseases and Parasites (Varroa, Nosema, IAPV)
- Poor colony nutrition and heritable weaknesses

In previous colony collapse episodes, the blame has been largely placed on the second category. Despite the current focus on damage caused by *Varroa destructor*, historically it is *Nosema* that has received the greatest amount of attention (Fantham and Porter, 1912), due its recurring outbreaks coinciding with major colony collapse events. However, non-varroa exposed colonies in Australia have so far shown little to no sign of colony collapse disorder onset.

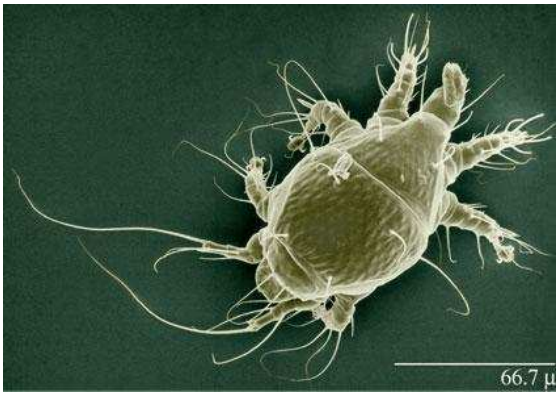

Figure 11. Possible viral agents of CCD **A** *Varroa destructor* **B** Worker larvae infested by varroa (Varroa on a pupae) **C** *Nosema apis*.

Furthermore, its introduction to North American colonies is correlated with an immediate downturn in hive lifespan. Varroa has also recently been shown to act as a vector for Israeli Acute Paralysis Viris (IAPV) (de Miranda et al., 2010; Di Prisco et al., 2011), further implicating it as a major factor in the decline in North American stocks. However, colony collapse has also occurred in non-infested hives, indicating that the presence of Varroa is not the only cause of its onset.

Despite the devastating effects of Varroa introduction, newly developed commercial pesticides have received the main share of the blame for the current wave of colony deaths. The introduction of neonicotinoid pesticides on a widespread basis within agriculture in the early 1990's coincided roughly with the most recent wave of

colony losses, leading to claims of a detrimental effect on honey bee mortality. Whilst the toxicity of the most widely used neonicotinoid, Imidacloprid, to honey bees has been established, there are concerns that low levels of exposure may be having sub lethal effects on colony health. Alongside this direct action, it has been hypothesised that chronic Imidacloprid exposure is resulting in long term heritable weaknesses (such as long term changes in DNA methylation pattern), leaving the colony vulnerable to disease and parasite infestation. However, the direct molecular action of long term exposure to low levels of imidacloprid has yet to be investigated.

Though poor colony nutrition has been the main factor in previous waves of colony collapse, it is considered to be more likely as a secondary contributor to the current outbreak. As malnutrition is a known stressor, it is plausible that incorrect nutrition is weakening honey bee immune systems, contributing to the spread of previously documented diseases. However, little data exists to support these theories, due to the greater focus on pesticide and parasite effects within contemporary research.

Due to recent studies indicating a possible link between low level neonicotinoid exposure and differential gene expression in *Apis mellifera*, this investigation will aim to explore its effects in terms of chronic treatment. This will focus on both its possible effects on the DNA-Methyltransferase family, and on a more global level on genome wide regulation of microRNAs.

1.2.4 The Neonicotinoids:

The systemic nitroguanidine pesticide family known as the neonicotinoids have seen widespread use within commercial agriculture. This utilisation is due to their competitive action as an inhibitor for nicotinic acetylcholine receptors (nAChRs) (Bai et al., 1991). Indeed, their relative toxicities have been correlated with the abundance of these receptors (Buckingham et al., 1995). nAChRs are widely abundant within targeted insect nervous systems, and as such, these compounds are highly potent insecticides even at low concentrations (Breer and Sattelle, 1987). Their appeal as a pest control is enhanced by low toxicity to mammals and relatively low toxicity for non-target insects during acute treatment. However, the affects of both their long-term and chronic use remain controversial, and they have often been linked to the recent decline in *Apis mellifera*. Of the neonicotinoids, the most widely utilised is Imidacloprid.

1.2.5 Imidacloprid:

The neonicotinoid Imidacloprid has recently seen a high level of scrutiny within the agricultural community, due to a perceived link between its use and the onset of colony collapse disorder. First synthesised in 1985 by Nihon Bayer, it was approved for agricultural use on a wide scale in 1994. Concurrently, the 1995-1996 season saw a large scale decline in honey bee stocks, with French farmers being the first to notice the onset of what was known as ‘Mad Bee Disease’. When exposed to large amounts of imidacloprid, honey bees indicated a sharp decrease in learning and memory processes.

Imidacloprid is a nitromethylene, and is therefore derived from the 2-(nitromethyl) pyridine structure, which possesses a low amount of insecticidal activity (Matsuda et al., 2001). Soloway et al. (1979) conducted an early structure-activity investigation based on this prototype, that lead to the discovery of 2-(nitromethylene)-tetrahydro-1,3-thiazine, also known as nithiazin. Whilst nithiazin displayed a greater level of

insecticidal activity, its high photo-instability (Kagabu and Medej, 1995) prevented its commercial adaptation. Substitutions at the 1-position lead to the development of nitromethylenes with greater insecticidal activity, in the form of PMNI, followed by Imidacloprid, which possessed enhanced photosensitivity (Kagabu, 1997).

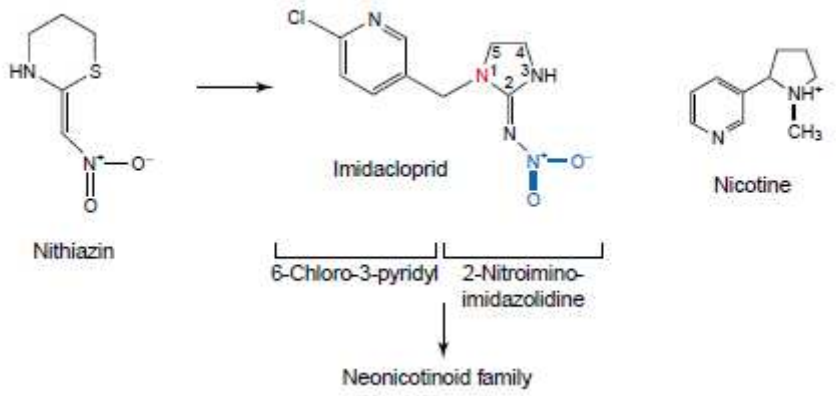

Figure 11. Imidacloprid synthesis pathway. Nicotine is included to allow a comparison of the aromatic ring structure. Imidacloprid was originally created via substitution at the N^1 position of Nithiazin, itself a derivative of Nicotine (from Matsuda et al., 2001).

The high level of toxicity for these commercial neonicotinoids (Imidacloprid, acetamiprid, and nitenpyram) in insects has been shown to largely be due to them all containing a 6-chloro-3-pyridyl moiety, thus resembling both nicotine and epibatidine. Buckingham et al. (1995) showed that the affinity of several neonicotinoids for insect nicotinic acetylcholine receptors (nAChRs) correlated strongly with its relative toxicity. This suggests that nAChRs may present the primary sites of action for these compounds (Liu et al., 1993). It is also notable that imidacloprid displays a higher affinity for insect nAChRs than vertebrate nAChRs, contributing to its effectiveness as a pesticide (Liu and Casida, 1993).

1.2.6 Previous Investigations into Imidacloprid:

Imidacloprid has long been implicated as a contributor in the 1990's/2000's colony collapse. As such, several experiments have been performed in order to determine the effects of its toxicity to the honey bee. Some investigations have focused on the effects of acute treatment, which result in a significant increase in colony mortality, alongside significant physical and behavioural impairments among workers. However, the few investigations into sublethal/chronic exposure that do exist have not indicated any connection between this exposure and an increase in colony mortality.

Beliën et al. (2009) conducted an extensive investigation on global colony vitality, comparing imidacloprid contaminated colonies (orally ingested through sugar water) with a non-contaminated control. This investigation focused on several parameters, such as the total amount of active and dead bees, capped brood surface area, and overall colony weight. Results indicated that this sublethal exposure had no significant difference in terms of colony weight or direct mortality; these findings were in line with those obtained by Nguyen et al. (2009). However, the number of active bees dropped at a significantly higher rate than the control bees (drops were expected due to end of the summer season), along with the number of capped brood cells. However, these recovered towards the control rates at late assessment rates, though it must be remembered that contamination treatment was conducted only once. Actual contamination may be a more gradual development due to long term harvesting of contaminated tissues.

Prior studies have also indicated a direct relationship between the effects of low level imidacloprid contamination and the Proboscis Extension Reflex. Cano Lozano et al. (2001) demonstrated that direct injections of nicotinic antagonists into the honey bee brain disturbed memory retrieval of olfactory conditioning of the proboscis extension reflex (PER). This reflex remains unaffected at lower doses (around

1.25ng/bee), and instead results in a habituation of the PER, alongside an increase in motor function (Lambin et al., 2001).

1.2.7 Imidacloprid and the nAChR receptor:

Nicotinic acetylcholine receptors are largely confined to the nervous system in insects (Matsuda et al., 2001), with insect nervous tissue being the richest of all nervous tissues in nAChR densities (Littleton and Ganetzky, 2000). Some evidence indicates the presence of nAChR subtypes present in insect tissue (Gundelfinger, 1992). This would explain the discrepancy between chronic and acute toxicities that imidacloprid has been shown to cause in honey bee. Suchail et al. (2001) were able to indicate an unusual mortality profile, where mortality rose significantly under a low dosage, fell with an intermediate dosage, only to rise again with higher doses.

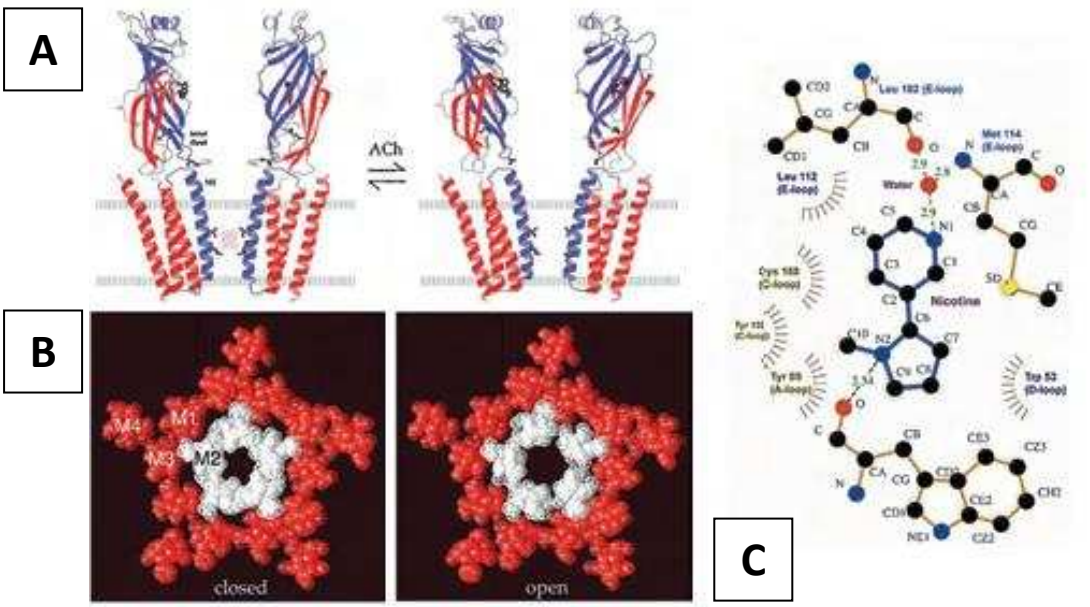

Figure 12. Nicotine action on the nAChR. **A** Two of the nAChR subunits in both open and closed states. Nicotine is capable of acting at the ligand binding site shown **B** Cross section of the nAChr in open and closed form (from Unwin, 2003) **C** Molecular binding of nicotine to the nAChR. Nicotine and its derivates are capable of this binding due to its ring structure.

Investigations have implicated two different theories for this varying level of mortality, either that there are multiple metabolic pathways for imidacloprid processing, or that more than one subtype of nAChR exists, with a varying level of nicotinic binding affinity (Suchail et al., 2001).

1.3 The DNA Methyltransferases and DNA Methylation:

1.3.1 DNA Methylation:

Cytosine methylation is highly prevalent within mammalian cells, and occurs predominantly in CpG di-nucleotides. This process results in the addition of a methyl group (CH_3) at the C5 position of the cytosine pyrimidine ring. Long known to be an essential component in correct mouse development, this repressive marker is generally associated with gene silencing (Bestor, 2000). As a result, changes in DNA methylation patterns are closely linked to the progression of cancers and other diseases (Momparler and Bovenzi, 2000). Whilst the mammalian model receives the most focus within DNA methylation research, Cytosine-5 methylation is a highly evolutionarily conserved, modification (Bestor, 1990). Present within bacterial, fungal, animal and plant genomes, it is likely that bacterial enzymes coding against restriction enzyme activity represent the evolutionary origins of the DNA methyltransferases (Bestor, 1990). This family of enzymes is responsible for C5 cytosine methyl group addition, functioning as a mechanism for both gene silencing and cell memory (Bestor, 2000). This system is present in most model organisms, with the main exception, the nematode *Caenorhabditis elegans* having possibly lost the system very recently in evolutionary terms (Bestor, 1990).

By comparison to the established Dnmt system within mammals, the functional existence of DNA methylation in insects has remained a controversial issue (Field et al., 2004). The recent sequencing of the honey bee genome revealed widespread CpG island methylation, alongside a fully catalytically active complement of Dnmt's (Wang et al., 2006). Due to absence of a previous insect Dnmt model, the discovery of this system may have provide new evolutionary insights into Dnmt's in invertebrates.

1.3.2 The DNA Methytransferase Family:

Though long detected in the form of crude cellular extracts, the first DNA methyltransferase was purified as a protein by Bestor and Ingram (1983). This first 200- kDa protein was named DNA methyltransferase 1 (Dnmt1), and was found to be specific to CpG dinucleotides (Pfeifer et al., 1990), with significant activity occurring on non-methylated DNA. Despite this, its preferred DNA substrate was shown to be hemi-methylated DNA, DNA methylated at CpG on only one strand (Pfeifer et al., 1990). As such, its role as the maintenance Dnmt was proposed, with subsequent evidence in the mouse showing that inactivation of Dnmt1 in embryonic stem cells results in genome wide loss of CpG methylation (Sen et al., 2010). The current view is that Dnmt1 maintains DNA methylation by completing hemi-methylated sites (Sen et al., 2010).

However, this did not account for direct *de novo* DNA methylation. Searches of expressed tag sequences using Dnmt1 revealed three potential transcripts that could encode other members of the Dnmt family (Robertson et al., 1999). Two of these candidates, later termed Dnmt3a and Dnmt3b were shown to have no preference for non-methylated or hemi-methylated DNA, and further disruption in mice indicated that these were the two missing de novo Dnmts (Robertson et al., 1999). This was shown by ES cells with respective knockouts being unable to methylate viral genomes and other repetitive elements (Robertson et al., 1999). Similarly, the lack of an intermediate factor Dnmt3L (like), results in a failure to establish distinct methylation patterns.

Considered the third subfamily (Jurkowski et al., 2008) of the DNA methyltransferases, the role of Dnmt2 proteins in DNA methylation remain enigmatic. Originally assigned to the family due to its extensive homology to both eukaryotic and prokaryotic DNA-C5-Methyltransferases (Yoder and Bestor, 1998), it displays the highest level of evolutionary conservation (Okano et al., 1998; Tang et

al., 2003). However, it displays low enzymatic activity on DNA substrates, likely due to the lack of nucleic acid binding cleft (Goll and Bestor, 2005), complicating any interactions with duplex DNA (Goll et al., 2006).

1.3.3 Dnmts and Caste Differentiation:

A recent investigation by Kucharski et al. (2008) indicated that widespread epigenetic reprogramming is responsible for controlling female reproductive status in the honeybee. As both fertile queens and sterile workers are genetically identical, it had long been considered that a nutritional trigger was resulting in phenotypic plasticity. Recent findings support the idea that this plasticity is due to somatic developmental imprinting, induced through selective feeding of female larvae (Barchuk et al., 2007). Whilst normal larvae receive much less complicated nutrition, those marked to become queens consume royal jelly, a poorly characterised biological substance. This results in behavioural & reproductive differences between the worker and queen castes, despite a clonal level of DNA similarity. These phenotypic alterations are likely due to changes in epigenetic state, which is accomplished through highly sophisticated enzyme machinery (Goll and Bestor, 2005). Dnmts form a crucial part of this regulation, and are notable for being highly conserved throughout developmental biology. Despite their absence in other known insect sequences, the honey bee genome has been shown to contain a full complement of CpG island methylating Dnmts (Wang et al., 2006).

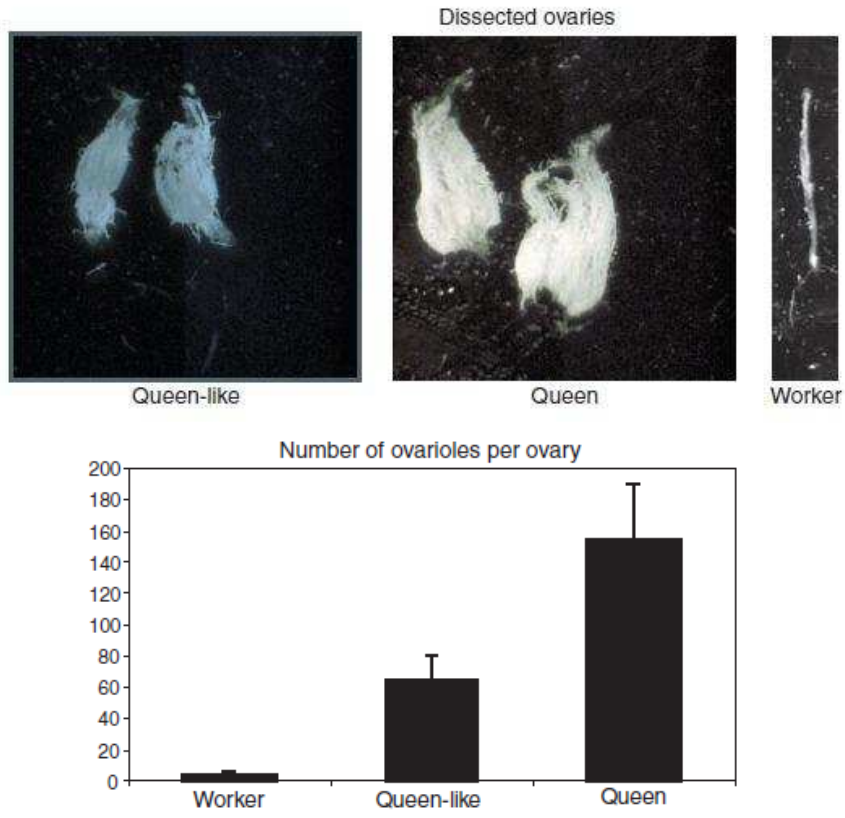

Figure 14. Comparison of ovary development under Dnmt3 silencing treatment. **A** Comparison of dissected ovaries between queen-like individuals (Dnmt3 impaired), full queens, and workers **B** Comparison of number of functional ovarioles between queen, queen-like, and worker phenotypes (Kucharski et al., 2008). Analysis of the dissected ovaries from the queen-like individuals reveals a level of ovarian development similar to that of a full queen. However, direct comparisons of the number of functional ovarioles show a lower level of functionality in the queen-like phenotype. This indicates that whilst the visible structure of the ovaries across the phenotypes has much in common, the actual reproductive capacity of the queen like individuals is greatly diminished. If imidacloprid treatment were to result in Dnmt3 silencing, it may result in the regular emergence of the queen-like phenotype. These individuals would be incapable of sustaining the required reproductive rate within the hive. A decreased reproductive rate between each generation could result in a sustained heritable weakness in the hive.

Impairment of Dnmt3, especially at the L2/L3 developmental transition, is known to distort the distinctions between caste developments, particularly in regards to ovarian development (i.e. Kucharski et al., 2008, see figure 14). Injections of small interfering RNAs directed against Dnmt3 (Kucharski et al., 2008) during the L2/L3 transition result modifies the cytosine methylation pattern of an illustrator gene. This shifts development from the worker phenotype towards a Queen-like

phenotype. The change in methylation pattern is thought to play a role in the differential expression of Dnmt3 between these 'queen-like' individuals, and non-silenced 'workers'. This was confirmed from analysis of *dynactin p62*, which was conducted in order to test differential methylation under dietary change (*p62* is regulated by diet in *Drosophila*)(Kucharski et al., 2008). This study indicated significant differences in the level of gene methylation in conserved gene sequences between Dnmt3 siRNA treated 'queens' and control 'workers' (48% for queens versus 58% in workers), particularly in DNA extracted from the larval head (63% for queens versus 79% in workers)(Kucharski et al., 2008). As this silencing results in remarkably similar characteristics to royal jelly treatment, there is some indication that it may utilise a similar form of silencing as a gene expression differentiator. It should also be noted that treatment using a non-larval control siRNA, *uth*, produced a much lower proportion of queen-like characteristics than direct Dnmt3 silencing (72% of larvae in direct siRNA treated, compared to 27% in non-larval treated), with a lesser degree of ovary development.

1.3.4 Neonicotinoids and RNA Expression (mRNA, micro-RNA, non-coding):

A recent study involving the nAChR in mice indicated a potential link between nicotine treatment and the DNA methyltransferases (Satta et al., 2008); however, the inspiration for this work is much earlier in origin. It has long been noted that human schizophrenia patients are frequent abusers of tobacco (Martin et al., 2004). Due to nicotine being a potent cholinergic receptor agonist, and the general down regulation of nAChR expression in schizophrenia patient brains, it is possible to conclude that this addictive behaviour is an attempt to self medicate (Breese et al., 2000). This is likely due to a decrease in GABAergic function with the disease state brain. However, these same receptors are known to be regulated by DNA methyltransferase 1 (Dnmt1) expression (Breese et al., 2000). Satta et al. (2008) found that repeated stimulation of nAChR's in the mouse through nicotine injections resulted in an increase in cortical and hippocampal expression of glutamic acid decarboxylase (GAD) 67. GAD 67 plays a key role in final GABA synthesis, hence

causing an increase in the amount of GABA synthesised. This is correlated with an overall decrease in Dnmt1 plasma levels, and a lower level of GAD67 promoter methylation. This decrease in methylation is the most likely cause of the increased GAD67 synthesis.

It has been previously noted that the neonicotinoids are a high affinity agonist for the nAChR (Liu et al., 1993), due to their structural similarity to nicotine. As such, it is entirely plausible that imidacloprid nAChR binding would result in a similar down regulation of Dnmt1. Despite this correlation having only being drawn so far from a single study in mice, the Dnmts are a highly evolutionarily conserved enzyme group, and it is possible that this same down regulation in expression is present within the honey bee. Furthermore, due to significant role of the Dnmt family within caste differentiation, it is possible that residual levels of imidacloprid contamination could be having an effect on worker-queen ratio/development.

1.4 Investigative Background

1.4.1 Background:

This possible association between Imidacloprid and molecular stress has been of significant interest for CCD investigators within the recent years. However, none have focused on the effect of low dose imidacloprid exposure in terms of differential gene expression. The Dnmts play an important regulatory role in honey bee development and differentiation, and any disruption of this regulation could have dire effects on overall colony fitness. Through the route of action established in section 3.4, it is plausible that small levels of imidacloprid contamination are causing differential gene expression within honey bee larvae. As shown in section 3.4, this differential gene expression may be interfering with larval differentiation. Due to direct conversion of contaminated nectar into brood feed, larvae, along with their attendant workers, receive the greatest amount of exposure within the hive. This exposure during such a critical period of development may result in caste abnormalities, lowering the overall health of the colony, and possibly resulting in a hereditary weakness.

This investigation will attempt to explore this through transcript analysis of larvae of Imidacloprid contaminated colonies. We will operate under the working hypothesis that this contamination results in subtle, yet statistically significant alteration in gene expression.

1.4.2 Aims:

1. To determine whether low level Imidacloprid exposure results in differential transcript expression in honey bee larvae. This will involve the examination of specific genes that have the most significant potential for differential expression, along with the type and level of alteration.

2. To conduct a bioinformatic analysis of transcriptomic data generated by SOLiD platform sequencing in order to draw conclusions about the effects of larval imidacloprid contamination. These results will be experimentally confirmed using qRT-PCR.

1.4.3 Plan of Work:

Larvae from hives that have been exposed for two weeks to the imidacloprid contaminated sugar water will be analysed. The length of exposure was chosen to simulate a longer term chronic contamination of each colony. These L3 larvae were then stored using RNAlater in order to inhibit RNase activity. The transcriptome, including mRNA, microRNAs and long non-coding RNA of these larvae will be determined by high-throughput sequencing (SOLiD platform) at both the University of Nottingham and Genomnia, Italy.

Whilst the transcription data is being produced, the first phase of the investigation will focus on a list of gene candidates that may show changes in gene expression upon exposure to imidacloprid. Housekeeping genes will also be analysed, in order to normalise mRNA levels between different samples (Silver et al., 2006). This will consist of larval DNA extraction using a RiboPure Kit, followed by multiplex reverse transcription PCR (rt-PCR) to measure gene expression levels. The addition of marker tagged primers allows this mechanism to universally amplify cDNA and mRNA gene copies with no loss of specificity. Quantitative real-time PCR (qPCR), will also be used to quantify mRNA levels. Bioinformatic modelling will then be used to deduce any differential expression amongst transcripts.

2: Materials and Methods

2.1 Honey Bee Larvae and Isolation of RNA

2.1.1 Larval RNA Isolation:

Individual larvae corresponding to the third developmental instar were selected from each hive and stored using Ambion RNAlater, before freezing at -70°C. RNA isolation was then performed on these larval samples. An initial method of phenol based extraction was used to test the concentration of RNA produced from whole tissue homogenisation. Phenol was selected due to the high fat content of larval tissues. Whole larvae were thawed then homogenised directly in 10-20 volumes TRI Reagent (1mL TRI Reagent per 50mg of tissue), followed by incubation at room temperature for 5 minutes. Due to a range of different individual larval weights, the actual volume of reagent differed between samples. At this point 1mL of the homogenate was transferred to a 1.5mL microcentrifuge tube, to which was added 100µL of bromochloropropane (BCP). Due to a large amount of homogenate generated, sample five was split into two separate tubes. These tubes were then vortexed for 15 seconds and incubated at room temperature for 5 minutes. The mixture was then separated through centrifugation at 12,000 x g for 10 minutes. From this, 400µL of the aqueous phase (top layer) was removed and transferred to a new, labelled microcentrifuge tube.

Samples from this homogenisation method were able to confirm that a high RNA concentration can be obtained for whole tissue samples. Further homogenisation switched from this phenol method to a column based method. This was due to the need to isolate RNA which could be used for both specific testing of gene candidates and the generation of the larval transcriptome.

For RNA purification, a RiboPure kit was utilised. 200µL of 100% ethanol was added to the recovered aqueous phase, the mixture was then vortexed immediately for 5 seconds to prevent RNA precipitation. The sample was transferred to a Filter Cartridge-Collection Tube assembly and centrifuged for 30 seconds, after which the flow-through was discarded. 500µL of a wash solution was applied to the assembly, which was then centrifuged for 30 seconds. The flow through was once again discarded, and the wash procedure was repeated to prevent ethanol carryover. The Filter Cartridge was then moved to a new Collection Tube, and 100µL of Elution Buffer was added. The filter column was incubated for 2 minutes at room temperature, followed by 30 seconds of centrifugation to elute the RNA. This elution process was repeated with a second aliquot of buffer. The recovered RNA underwent long-term storage at -70°C.

2.1.2 DNase Treatment and Removal:

Before RNA quantification, a Turbo DNA-free kit was utilised in order to remove any contaminating DNA fragments. 0.1 volume 10X TURBO DNase Buffer and 1µL TURBO DNase (2 units/µL) was added to the RNA, and mixed gently. This was followed by incubation for 30 minutes at 37°C. A 0.1 volume of resuspended DNase Inactivation Reagent was then added and thoroughly mixed, before being incubated for 5 minutes at room temperature. The samples were centrifuged at 10,000 x g for 1.5 minutes, after which the supernatant (Containing the RNA) was removed and transferred to a fresh tube.

2.1.3 RNA Quantification:

Upon completion of DNase treatment, the RNA concentration was determined by measuring its absorbance at 260 nm (Absorbance wavelength 280/260). This was accomplished by the use of a NanoDrop 1000A Spectrophotometer, which was used to directly measure 1µL of the sample.

2.2 GeXP Multiplex PCR:

2.2.1 GeXP Overview:

Multiplex gene expression analysis relies on the combination of a reverse transcriptase reaction with a two-primer PCR process, in order to amplify multiple genes within a single reaction. Through this method, it is possible to efficiently analyse up to 30 genes per reaction whilst retaining sensitivity and speed. Whilst this use of a universal priming strategy maintains the gene ratio of the RNA sample, it also allows variations in individual amplification efficiency of individual genes to be overcome. Most importantly, this is achieved without compromising overall detection sensitivity.

Candidate mRNA's are initially converted into cDNA utilising the enzyme reverse transcriptase. The reverse primers utilised for this conversion are chimeric, containing a 19-nucleotide universal tag sequence at the 5' end, and a gene specific sequence at the 3' end. This results in each newly synthesised single stranded cDNA containing the full universal tag sequence at the 5' end. After this initial conversion, GeXP Multiplex PCR can then be performed. This process is comprised of a ten minute incubation of each reaction at 95°C, followed by 35 seperate PCR cycles. Each of these consists of the following conditions: 94°C for 30 seconds, 55°C for 30 seconds, and finally 68°C for 1 minute. The first three PCR cycles utilise chimeric forward primers, which contain a second 18-nucleotide universal tag-sequence at the 5' end, and a gene specific sequence at the 3' end. These initial reactions result in a two-stranded cDNA containing both forward and reverse universal tags at the 5' ends, as well as establishing the ratios of specific genes before universal amplification. This forward primer also contains the D4 dye label, which is integral for later analysis via capillary electrophoresis. It should be noted that these primers were also designed in order to amplify a kanamycin RNA transcript that is included within each reaction to act as an external control.

Cycles 3-35 see the deliberate introduction of an excessive concentration of universal forward and reverse primers. This results in simultaneous and uniform amplification of multiple target fragments, helping to prevent primer bias whilst providing adequate amplification. Each reaction is then subjected to capillary electrophoresis, which is used to separate the dye-labelled fragments.

2.2.2 Primer Design and Singlets:

Whilst GeXP Multiplex PCR operates using a universal primer during later amplification, specific Forward and Reverse Primers are required for initial reverse transcription. These specific primers were designed using the Express Designer tool of the GenomeLab Genetic Analysis System. The full structure of these F/R primers, alongside the Genes of Interest, can be found in Table 1.

2.2.3 Testing Individual Primer Pairs (Singlets):

In order to test individual primers, singlet reactions were run for each constituent gene within the multiplex. Simultaneously, a sample multiplex rt/PCR was run for each of the test RNA samples. This utilised a Reverse-plex (comprising of 900µL Tris buffer + 5µL of suspended primer) and a Forward-plex (comprising of 480µL Tris buffer +1µL of suspended primer). Preliminary results indicated primer specificity, along with no evidence of DNA contamination.

Primer Name	Product Size	Left Sequence	Right Sequence	Universal Left Sequence	Universal Right Sequence
Dnmt3	100	ACAGGAACCGGACGTGATAG	GGACGATAACAGACGTGGCT	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
MBD_b	110	TTTCGTCTCAACCAGTCACG	AGCTATCGAGACGGCCTGTA	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
Dnmt1a	117	CGAGTAGTAAGCGTGCGTGA	AAGTGGTGGAGGAACTGCAT	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
Dnmt2	126	CCCTTGTCAATTTGGCATTC	TGTGATCCTTCAACATGCTCT	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_GAPDH	133	ACTTGGCAAAGGTGCAGACT	GGCATGGTCATCACCAATAA	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
Dnmt1b	142	TGGAATGAATTGGGAAATGC	TTCTTGGGCATTCAGGATCT	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_SAM_TM6	149	GAAAAAGACGCGAAAAAGCA	CAGCCGACTTCGAAAAGAAC	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_actin	156	TGCCAACACTGTCCTTTCTG	AGAATTGACCCACCAATCCA	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_cyt_c	163	AGGAAAACTGGTCAAGCACC	CACGTTCTTGCGGTTTCTTC	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_JHE	170	CTTTTCTCGCTTCCACAACC	TCCTGGTCCAGCAATGTGTA	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_gst1	177	CCCGTTGAAATTAAAATGCC	TATCGTGTGTTGCGGGTTTA	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_ubq	184	TAGTATTGCGTTTGCGAGGA	CGACCATCCTCCAGTTGTTT	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_TIF_2_mt	198	ATAAAGCAGCTGCCGATCAA	TTCCTGCCACATCACCTTTT	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_rbp18	206	GGGCACGAATATTGACGGTA	TCAAAAACCAATCGGGAATC	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_sdha	213	GGCAAAGCTGCAAAAATCTC	ACCCAAGGCAGCATTGATAC	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_aTub	234	GATATTTGCCGACGCAATCT	GACCGAAAGTTGCTCGTGAT	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_DNA	244	GCACAGACCCGAGTGAATAGA	CATATGAACCGGTGAAAAATGT	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_Ef-1a-f1	252	CTTGGGGTCAAGCAGTTGAT	CTTTCCGTCAGCGTTACCAT	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
a_cox-1	272	TCTTCACCATCTGTTGATATTGC	ATTGGATCCCCACCTCCTAT	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA
Kan(r)	288	ATCATCAGCATTGCATTCGATTCCTGTTTG	ATTCCGACTCGTCCAACATC	AGGTGACACTATAGAAT	GTACGACTCACTATAGGGA

Table 1. Properties of individual multiplex primers. Included is the Left (Forward) and Right (Reverse) primer sequences for both the genes of interest (in bold), and reference gene candidates. The Universal tag sequences for non-specific amplification during rt-qPCR are also included

Though specificity was confirmed for the major genes of interest and housekeeping genes, a number of gene candidates were highly over amplified within this early multiplex. As a result, attenuation of specific primers was required to bring their associated genes within a dynamic range of quantification (Mesarch et al., 2000).

2.2.4 Primer Attenuation:

As peaks produced by preliminary multiplex use were not within a quantifiable range, a series of primer attenuations were conducted. Genes exceeding quantification levels were organised into a preliminary group of high expressers (HE1), constituting of JHE, rbp 18, actin, GAPDH, and Ubq. After initial dilution (see figure 16), a new wave of highly-amplified genes emerged, consisting of MDB, Dnmt 1a &1b, SAM, Cyt-c, and TIF2. This emergence was due to the lower amplification level of HE1 allowing a greater amount of free nucleotides within the PCR reaction. These were categorised as a second group of high expressors (HE2), which received individual dilution. The remaining primers were organised into a final group of expressers (HE3). Within this group, Dnmt2, sdha, and Cox-1 did not present visible peaks within the original multiplex reactions. As it was possible that the high expression within the groups HE1/HE2 was preventing any amplification of these genes, their primer concentration was tripled.

2.2.5 RNA Dilution:

In order to further lower total RNA amplification to within a quantifiable range, a series of dilutions were performed on the pure RNA. This was achieved through the addition of DNase-free H_2O, resulting in a final RNA concentration of 30ng/ml.

2.2.6 Capillary Electrophoresis:

The post-PCR samples were analysed utilising Beckman-CQ capillary electrophoresis, at a dilution range of 1 in 20 (2µL of qPCR product diluted in 38µL Sample Loading Solution (SLS)), in order to enhance the resolution of the results. Each sample included a marker dye in order to indicate size and to confirm that the sample has been correctly run. This analysis then separates individual dye labelled fragments based on their respective sizes. Whilst this method produces a signal that can be used to judge quantity, a certain level of clean-up is required before data analysis can be performed. It must be noted that this quantification relies more on the ratio of the level of signal between samples, rather than absolute quantification.

2.3 Data Analysis:

2.3.1 Sample Profiling:

Preliminary data analysis was conducted utilising the GeXP Profiling tool of the Genome Lab System software. This module allowed analysis of the specific fragments generated from the GeXP. Each sample was screened for quantification level and possible DNA contamination. The raw expression levels of these fragments were then exported to Microsoft Excel for Gene Normalisation.

2.3.2 Gene Normalisation:

Normalisation of real-time PCR data is critical in order to produce reliable mRNA quantification. Compounding problems in the PCR can easily result in changes in expression profile, due to such factors as the integrity of the RNA, RT efficiency, and variation between loadings of cDNA samples. It is also possible that edge effects during machine loading may have resulted in sample variation. The most common form of normalisation is by relating the mRNA level of the gene of interest to the mRNA level of a reference gene. These reference genes ideally act as an internal control, displaying a highly stable level of in vivo expression regardless of cell type. As such, these genes are often referred to as housekeeping genes, and are typically cytoskeletal (e.g. actin) or ribosomal in origin (e.g. rbp18). The inclusion of these confirmed endogenous controls increases the reliability of the rt-PCR, allowing correction for any errors due to sample variation and quantification.

A large number of transcript analyses currently only utilise a single housekeeping gene as a normalisation factor, however, recent studies have shown that a normalisation factor generated for a larger amount of housekeeping genes may reduce variation errors (Nicot et al., 2005). In order to directly compare the two methods, separate normalisation factors were generated from the average expression of both two housekeeping genes, and from five housekeeping genes.

However, the reliability of this normalisation is heavily tied to quality of the normalisation factor employed. Any variation within this factor can mask real changes, as well as produce artificial ones. Identification of the most suitable candidate genes to act as a normalisation factor was conducted from analysis performed on the bee RNA data using the statistical package NormFinder. This identified Actin and Juvenile Hormone as the combination presenting the highest stability value. The top five candidate genes identified were Actin, JHE, GAPDH, aTUB, and EF-1a-f1.

2.3.3 Statistical Analyses:

The data was analysed using a general linear modelling with the statistical package Genstat. A one-way Analysis of Variance (ANOVA) was utilised in order to directly compare between the two treatment groups. Individual hives were utilised as the main experimental units, with larval weight being incorporated as a co-variate.

2.4 Transcriptome Sequencing

Whilst a majority of larval samples were utilised for experimental validation, 4 extracted mRNA samples were subjected to next generation sequencing (NGS) in order to determine their transcriptome. This high-throughput sequencing was conducted at both the University of Nottingham and Genomnia, Italy. Applied Biosystem's SOLiD sequencing platform (AB SOLiD) was utilised for this purpose. AB SOLiD uses a di-base sequencing chemistry in order to generate one of four possible calls (colours) combinations for each nucleotide pair, with SNPs resulting in a change of two adjacent colour positions. Any sequencing errors are indicated by single call changes, allowing easy differentiation between the two. A series of algorithms were used in order to map both short and long reads. Any identified transcripts will then be analysed for potential differential expression between control and experimental groups. This analysis will utilise the R based software package Bioconductor, specifically the Empirical analysis of digital gene expression data in R (edgeR) package.

3: Results

3.1 Multiplex PCR:

3.1.1 Test Samples:

The RNA Isolation method was originally performed on five test samples from the larval control group. This was to confirm whether or not it was possible to isolate a satisfactory concentration of larval RNA from a whole tissue sample. As shown in figure 8.1, a significant total concentration of RNA was produced. However, total RNA quality remained unknown at this stage and it was not possible to rule out any DNA contamination.

Sample ID	RNA Concentration (ng/µL)	A260	260/230	260/280	Constant	Weight (mg)
C1	1592.6	39.817	1.23	2.04	40	150
C2	214.37	5.359	1.94	2.16	40	100
C3	787.94	19.699	1.69	2.14	40	100
C4	169.4	4.235	2.10	2.16	40	150
C5	254.95	6.374	2.19	2.16	40	150

Table 2. Larval Weights and RNA Concentrations of Control Samples. 230/260/280 indicates the wavelength of the spectrum that is currently being sampled.

Due to a low number of test samples, it was not possible to match up individual larval weights. Nanodrop analysis indicated a large variation in the concentration of RNA isolated, even between the individual larvae of the same weight (C2 and 3, C1, 4, 5). A possible explanation for this individual variation is the overall metabolic state of the larvae. Whilst all selected larvae were visually analogous to the third developmental instar, it was difficult to identify how long each specimen had been at this stage. Due to the occurrence of widespread epigenetic activity triggered at this stage (associated with the trigger between queen/worker development), it is possible that some of the control larvae had reached this transitional period, accounting for a greater concentration of isolated RNA. Unsurprising the two samples generated from the fourth larvae displayed the lowest concentration of

RNA. A wide range of concentrations was generated from these samples; however this not immediately indicative of differences in size or metabolic weight. It must be noted that these concentrations depend on the amount of the aqueous phase that is recovered during the interphase stage of phenol based homogenisation. The fact that this was the first time utilising this method may also have limited the amount of aqueous phase that was recovered.

With the presence of a large quantity of larval RNA confirmed, singlet testing was performed to check for individual gene expression.

3.1.2 Singlets and Multiplex Testing:

Individual singlet PCR's were carried out in order to ascertain whether or not specific genes present in amplifiable amounts in the test RNA isolations. Alongside this, a multiplex PCR reaction involving all the primers of interest was also conducted. These results are summarised in Table 3.

Singlet runs resulted in the confirmation of the presence of the vast majority of genes of interest. Despite this, GAPDH, Dnmt1b, SAM, and Actin were either not confirmed as amplified within their respective singlet reactions, or were reactions that failed. However, these genes were present with the multiplex run, indicating that these negative singlet reactions most likely resulted from pipetting error. The possible presence of a DNA peak in its singlet was somewhat alarming, as it may indicate that the RNA sample was contaminated by genomic DNA. A further repetition of multiplex testing confirmed the presence of this peak (284), but it was not possible to identify it as DNA.

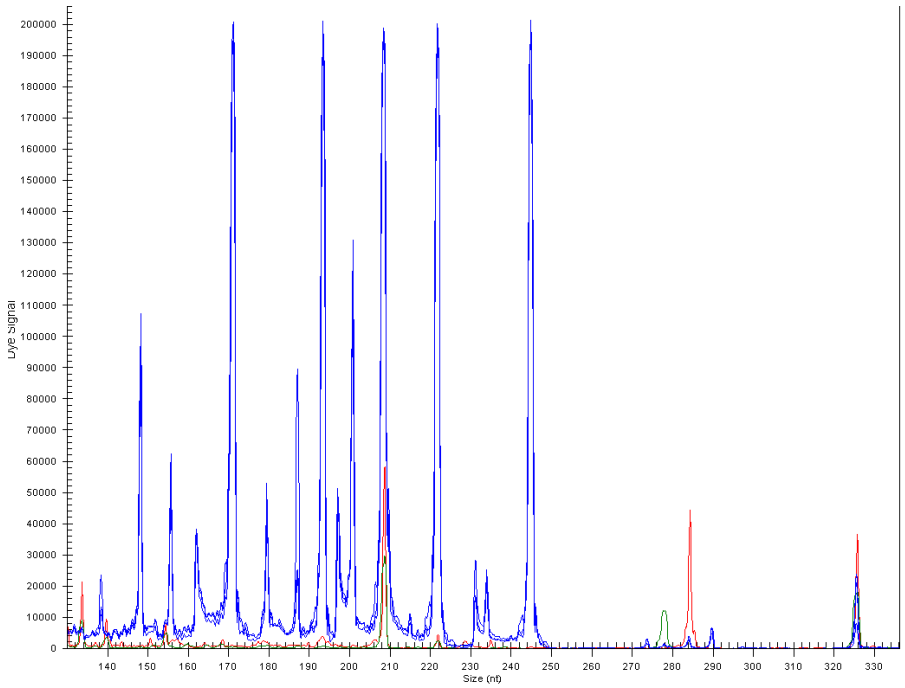

Figure 15. Test overlay multiplex spectra. Whilst a large number of reactions did fail, enough succeeded to generate some useful first impressions of the data. Alongside the five over amplified genes, several other genes were amplified beyond a dynamic range for accurate quantification. The only notable peak presented beyond the fragment size of 250 was the internal control Kanamycin; this is most likely due to free nucleotide exhaustion by the over-amplified peaks. Most importantly, this trial multiplex indicated that it is possible to produce specific RNA peaks from whole larval tissue isolations. Red spectra peaks indicate the dye marker added to confirm the success of the capillary electrophoresis.

Whilst a large amount of genes did give a quantifiable level of amplification, GAPDH, Actin, Cyt-c, JHE, ubq, and rbp18 were amplified beyond the limit of capillary electrophoresis quantification (See figure 15). Alongside this, cox-1 did not present a confirmed peak during either its singlet run, or the multiplex reaction. Two possible explanations for this lack of amplification are that either cox-1 does not have a large enough quantity of message produced within the fat body for polymerase to amplify it, or that the primers themselves did not match the sequence used for their generation. Alternatively the sequence itself could be incorrectly annotated, though this is unlikely. Due to these genes being beyond the dynamic range of quantification, a series of primer attenuations were conducted. Futhermore, it must be taken into account that singlet runs are not always perfect, due to a differing ratio of magnesium from multiplex reactions.

Gene ID	Singlet Result	Multiplex Result
Dnmt3 (137)	Positive (KAN)	Positive
MBD (147)	Positive (KAN)	Positive
Dnmt1a (154)	Positive (KAN)	Positive
Dnmt2 (163)	Low Positive (KAN)	Negative
GAPDH (170)	Negative (No KAN)	Positive
Dnmt1b (179)	Negative (No KAN)	Positive
SAM (186)	Negative (No KAN)	Positive
Actin (193)	Negative (No KAN)	Positive
Cyt-c (200)	Positive (KAN)	Positive
JHE (207)	Positive (KAN)	Positive
gst1 (214)	Positive (KAN)	Low Positive
ubq (221)	Positive (KAN)	Positive
TIF (235)	Positive (KAN)	Low Positive
rbp (243)	Positive (KAN)	Positive
sdha (250	Possible (KAN)	Negative
aTUB (271)	Positive (KAN)	Negative
DNA (281)	Possible (KAN)	Positive
Ef-1a-f1 (289)	Positive (KAN)	Low Positive
Cox-1 (309)	Negative (KAN)	Negative

Table 3. Singlet results for Gene Primers. Whilst some primers did not yield identifiable peaks during singlet runs, enough were present within the multiplex itself to all the investigation to proceed. Dnmt2 notably failed to present a specifc peak during the multiplex, but its presence within the singlet produced such a small signal that it is likely masked by more abundant gene amplification. Though Dnmt2 is technically a gene of interest, it has relatively insignificant catabolic action in vivo (see section 1.3.2) As such, it is unlikely to have much effect on DNA methylation patterns. As a result, Dnmt2 was largely excluded from the investigation, as an insufficient amount of quantification was produced to allow statistical analysis.

3.1.3 Primer Attenuation:

A preliminary set of attenuations were conducted in order to bring GAPDH, Actin, Cyt-c, JHE, ubq, and rbp18 into a dynamic range of quantification. These genes were categorised as High Expresser Group 1 (HE1). These attenuations comprised of producing a Reverse Multiplex primer with HE1 primers excluded. Simultaneously, a reverse multiplex was produced containing only the HE1 primers. This multiplex was then utilised to generate an attenuation series, utilising DNase-free H_2O. The series consisted of 1/2, 1/4. 1/16, 1/32, 1/64, 1/128, and 1/256. The two multiplexes were

then combined during the PCR reaction. This initial series resulted in very little in terms of improvement, as the results remained well beyond a quantifiable range. However, as 1/256 began to show a slight lowering of peak levels, two further dilutions of 1/512 and 1/1024 were created.

Alongside this attenuation, initial primer concentrations were increased for genes indicating a lower level of amplification. These genes included Dnmt3, Dnmt2, sdha, aTub, Cox-1, and DNA, and were collectively termed High Expresser Group 3 (HE3). For the 1/512 and 1/1025 multiplexes, an additional 10µL of HE3 primers was added to the Multiplex for each of these genes, for a total of 15µL. This primer boosting resulted in unquantifiable levels of Dnmt3 and aTub amplification.

Despite the over amplification of some of the HE3 primers, a significant reduction in HE1 group expression did occur. Concurrently, a second group of high expressers emerged, comprised of MDB, Dnmt1a, Dnmt1b, SAM, Cyt-c, and TIF2. Due to the confusion surrounding the changes in amplification for HE1 (it was unknown whether this change was due to primer boosting or the further attenuation), the experiment was repeated utilising the 1/256 sample of group HE1. Alongside this, HE2 primers were attenuated 1/32, with the exception of Dnmt1a, this was attenuated to 1/16 due to a lower level of expression. The results of this attenuation were much closer to the dynamic range for HE2, but HE1 samples were still far beyond quantifiable.

Alongside this dilution, RNA samples were generated using a mixture of 5 homogenised larvae. This was performed in order to reduce variation between the levels of RNA isolated from individual larvae. This was combined with the 1/512 sample for HE1, with the exceptions of JHE and rbp18, which were both attenuated to 1/1024. HE2 was kept at the same dilution as the previous experiment, due to the stable, quantifiable range produced. From group HE3, Dnmt2, sdha, and cox-1

received primer boosting, due to a continued lack of amplification. Whilst this experiment did not result in a quantifiable range for the HE1 genes, it did result in an even amount of amplification both across and between individual samples. In order to retain this even range, this multiplex was kept for RNA elution.

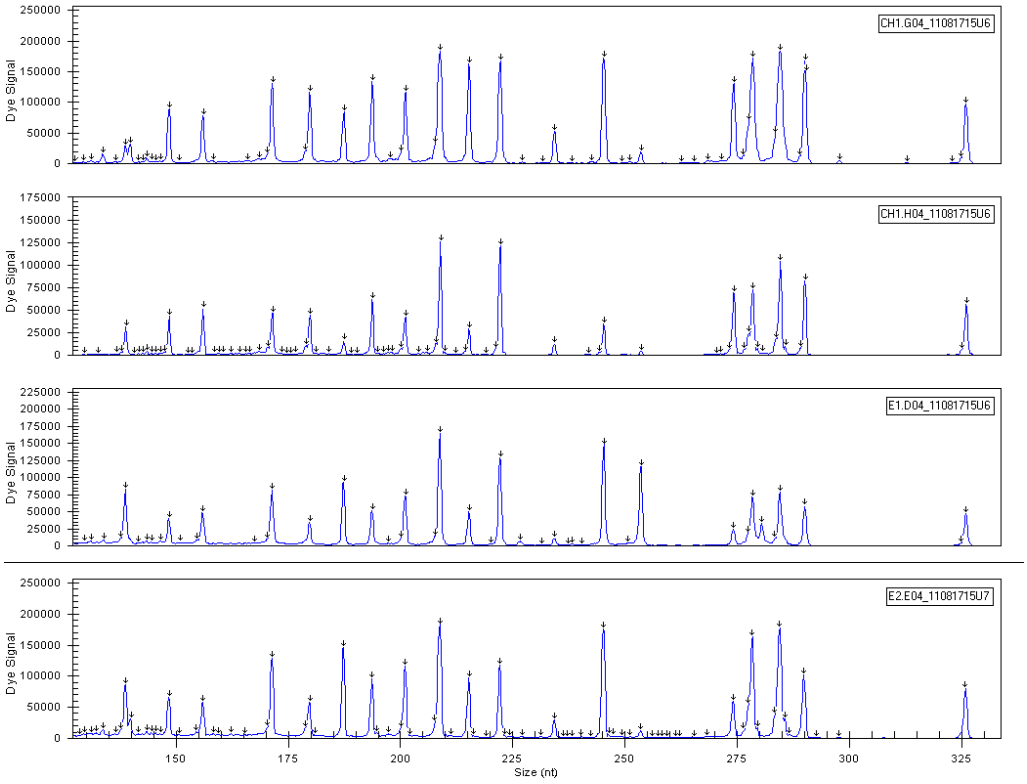

Figure 16. Multiplex spectra after HE1 primer attenuation. Several specific peaks beyond the size of 250 did emerge after this attenuation, most notably a series of peaks between 274 and 284.

Whilst there was still no identifiable fragment corresponding to DNA, two unknown peaks of similar size (278 and 284 respectively) appeared during the first wave of HE1 primer attenuation. In the initial multiplex, these peaks were present but largely unnoticed due to displaying a very small signal. This gave the first indication that non-specific primer binding may be taking place in these multiplexes. Also implicated was the presence of DNA contamination, if these peaks were in fact directly related to DNA.

3.1.4 DNase Treatment:

Whilst the close proximity of the 278 and 284 products to the size of a DNA product, could not yet be linked with DNA contamination, their unexplained appearance left its possibility open. Though the RiboPure kit utilised during RNA isolation is designed to produce highly purified RNA, no test is 100% reliable during actual usage. As a result, it was decided to DNase the RNA samples to remove any outside chance of contamination. In order to confirm successful DNA digestion, a PCR multiplex reaction was performed (see figure 17). Two samples were also prepared with reverse transcriptase excluded during the initial rt-PCR. As a result of this exclusion, only initial non-RNA contamination could be amplified. Alongside these two -RT reactions, two samples were also prepared with no template controls, in order to test for the possibility of outside contamination (see figure 18). As these reactions did not present either of these two mystery peaks, or the specific DNA peak itself, indicating that the chance of DNA contamination was low.

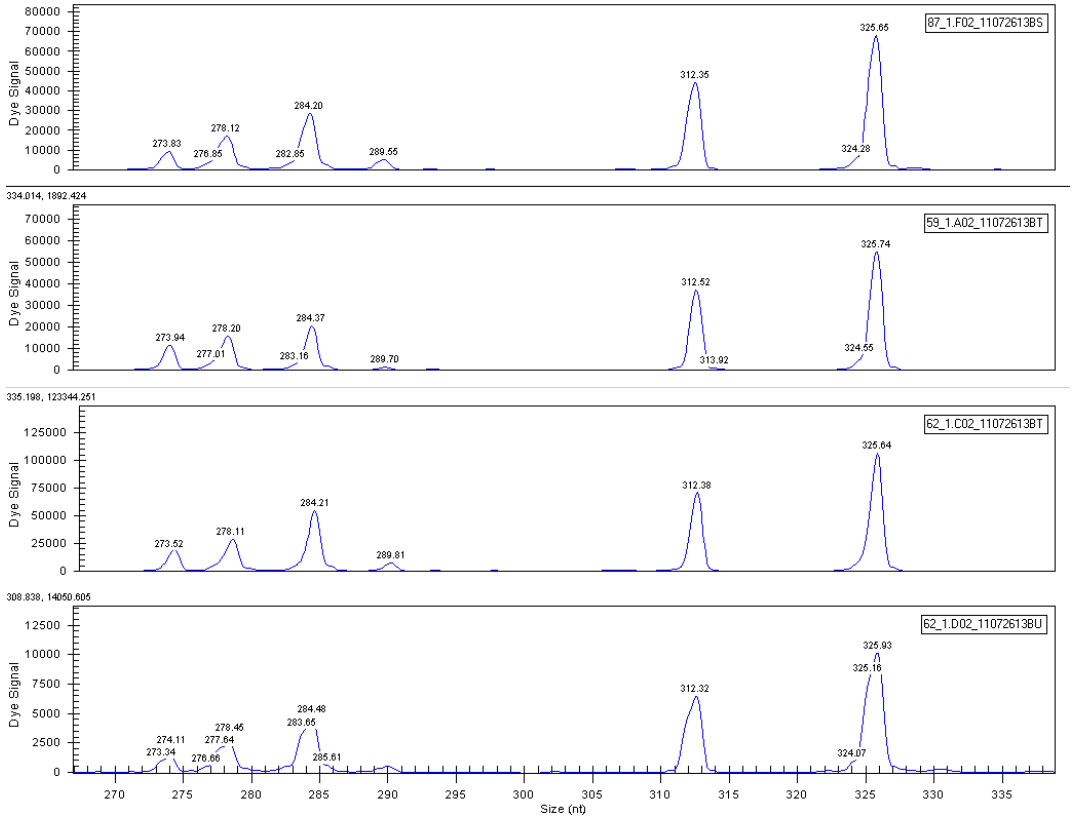

Figure 17: Multiplex spectra after DNase treatment. Whilst these peaks were notably reduced after DNase treatment, their continued presence indicates that they are unlikely to indicate actually DNA contamination. It is also possible that this reduction was a result of heat degradation during the DNase process.

3.1.5 Primer Comparisons:

Due to the continuing presence of peaks 278/284, different primer combinations were used to ascertain their group of origin. The first three samples contained a multiplex made up of only one primer group respectively (HE1, HE2, HE3), the second two contained a combination of two groups (HE1/HE2, HE2/HE3). These combinations were included to evaluate whether or not primers from each expresser groups were interacting in a non-specific manner. Alongside these reactions, isolation of genomic DNA was attempted from larval samples, and a PCR reaction was performed using this in the place of RNA. A final sample containing the full Reverse multiplex was prepared utilising 1µL of this larval DNA.

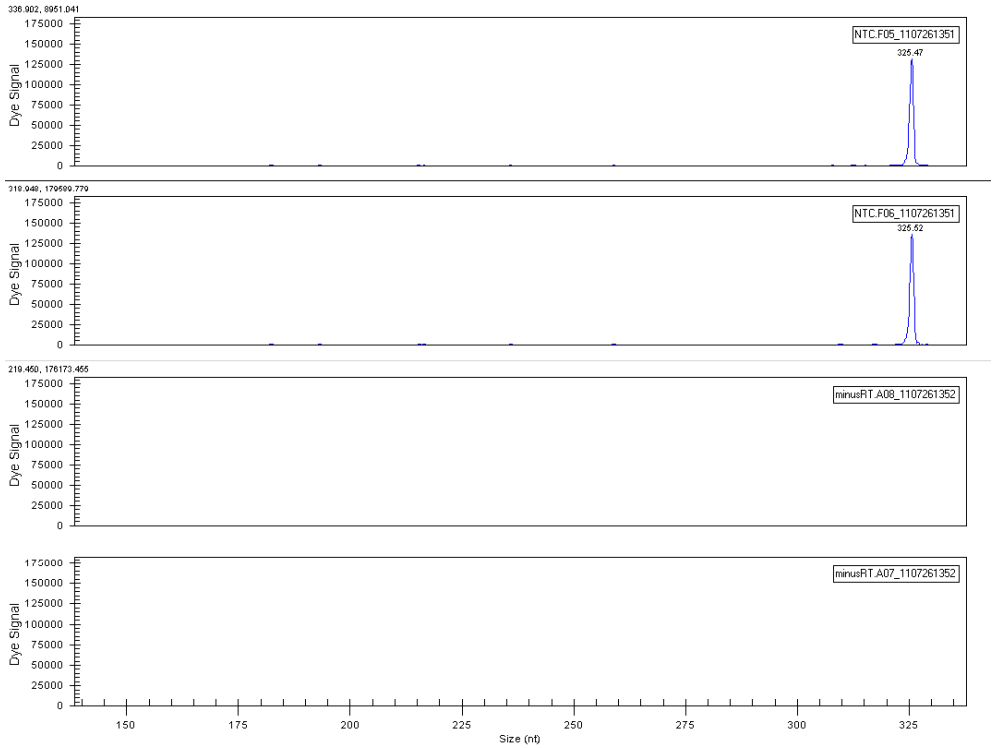

Figure 18: -RT and NTC multiplex reactions. The bottom two -RT reactions did not result in any identifiable peaks, failing to give any signs of non-RNA contamination. The top two NTC reactions only displayed an identifiable Kanamycin peak, again giving no sign of non-RNA or chemical contamination, and also indicating that the reaction itself was successful.

The results for the samples containing only individual groups immediately indicated that the two peaks were only present with the inclusion of primers from HE3. This result was reinforced by the combination samples, as only the HE2/HE3 reaction contained the 278/284 peaks. Rather unusually, peaks corresponding to JHE (207), and Actin (193) were also present within samples that did not contain HE1 primers. Though unlikely, it is possible that their presence is due to primer contamination. As such, a new mix for HE3 was formulated.

The sample containing genomic DNA fortunately did not present any notable peaks. Of the few peaks that did present any signal distinguishable from the baseline, 278/284 were the most prominent. However, their level was far below any previously registered. There was still no sign of a peak for 281, which was previously

thought to be the size of larval genomic DNA. This may indicate that a primer is not binding in a specific manner or that one of the two unknown peaks was the true peak for DNA. Alternatively the amount of extracted message was too low for quantification. Due to their lack of appearance in the (-)RT reaction, and the DNase treatment, the former of these two possibilities seems more probable.

3.1.6 Primer-Peak Identification:

In order to identify the specific primers responsible for the two unknown peaks, a series of singlet reactions were performed using different forward and reverse primer combinations. This process indicated that the two peaks only presented during the presence of DNA and Dnmt2 primers. The results of different combinations of their forward and reverse primers are listed in table 4, and displayed in figure.

Reverse Primer	Forward Primer	Peak 278	Peak 280	Peak 284
DNA	DNA	Negative	Negative	Positive
Dnmt2	DNA	Negative	Negative	Negative
DNA	Dnmt2	Negative	Negative	Negative
Dnmt2	Dnmt2	Positive	Negative	Negative
Dnmt2 + DNA	DNA	Negative	Negative	Positive
None	Dnmt2	Positive	Negative	Negative
None	DNA	Negative	Positive	Positive
R-plex + DNA	Dnmt2	Negative	Negative	Negative
R-plex + Dnmt2	DNA	Negative	Negative	Negative
None	Dnmt2	Positive	Negative	Negative
R-plex + DNA	DNA	Negative	Positive	Positive

Table 4. Singlet results for DNA-Dnmt2 primer combinations.

These results display that the peak at 278 is only present when Dnmt2 primers are present within the reaction, whereas 284 is only with present with the inclusion of DNA primers (see

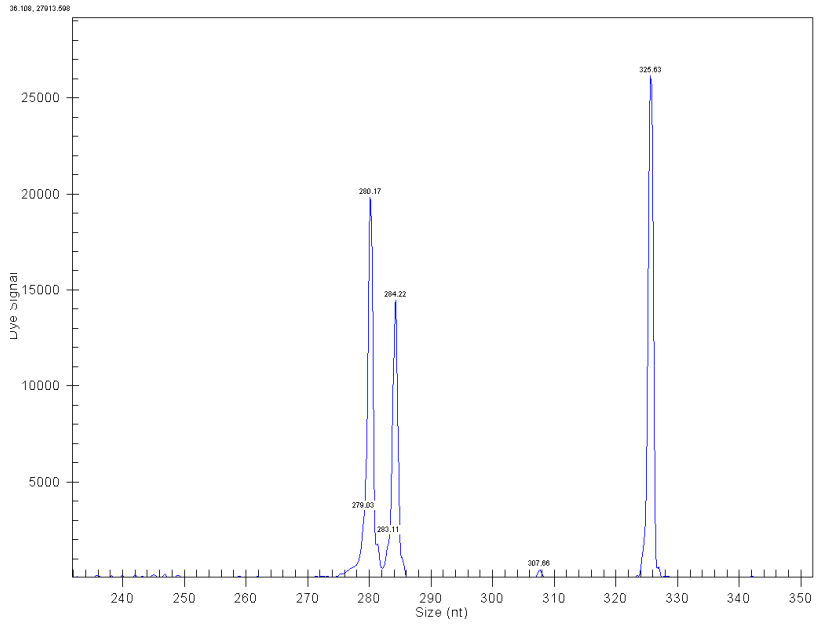

Figure 19: DNA 281 peak. This spectrum indicates that the peak corresponding to genomic DNA does appear when both DNA contamination and DNA primers are present. Alongside this, the unknown fragment 284 is amplified.

figure 20). The peak for DNA contamination, 280 was also present within certain larval samples, again only alongside the inclusion of DNA primers (see figure 19). This indicates that certain larval samples do contain the peak for genomic DNA, but more importantly, that the two unknown peaks are not directly related to DNA contamination. Furthermore, no cross contamination between DNA and Dnmt2 primers occurred.

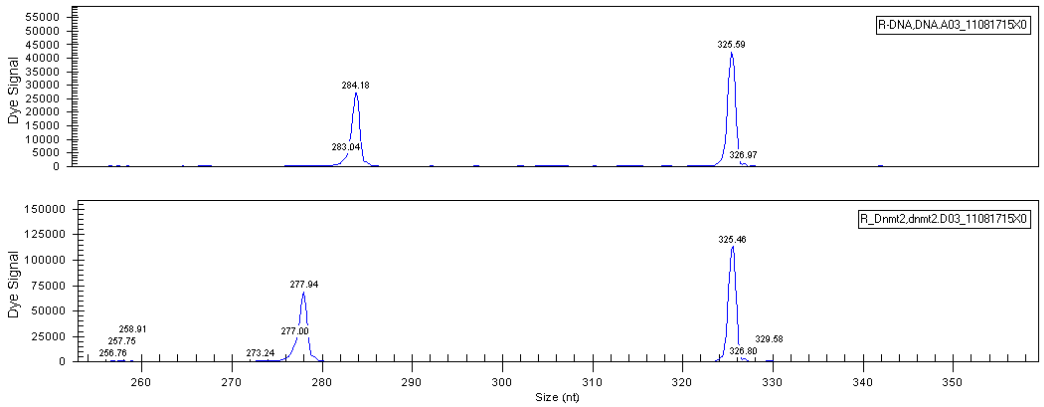

Figure 20: DNA-Dnmt2 primer combinations. No overlap occurred between the DNA primers and Dnmt2 primers.

3.1.7 RNA Dilution:

In order to reduce quantification to a measurable range, the RNA was diluted to a series of concentrations. These concentrations were 70ng/µL, 35ng/µL, 18ng/µL, and 4ng/µL. Preliminary attempts at using these dilutions to generate an RNA gradient were not successful, due to a large amount of failed samples. Whilst quantification varied widely even within samples of the same RNA concentration, this was not yet correlated with total quantification, as no normalisation had been performed on these results.

As it is most likely that these unknown peaks are due to non-specific binding, rather than any genuine threat of DNA contamination, the original primer combinations were kept, and dilution of the RNA recommenced. Rather than a mixture of RNA from multiple larvae, RNA was extracted from specific larvae of known weight and treatment type. 86 samples of larval RNA were collected, representing a weight range of 50 to 190mg. Of these, 24 samples were utilised for the final rt-qPCR (12 control and experimental respectively). Each RNA sample was diluted to a concentration of 50ng/ml, in an attempt to produce a signal within a quantifiable range. Multiplex PCR was then performed on 12 control samples and 12 experiment samples. Alongside this, a (-)RT reaction was performed to check for preliminary DNA contamination.

This experiment was hampered by the failure of a full replicate of samples; however, those examples that did not fail produced peaks that consistently displayed a quantifiable level of amplification. The experiment was then repeated in order to ensure that the results were indeed replicable. Notably, several results did produce a peak at 280, indicating some level of DNA contamination. As such, the next experiment was conducted utilising the same RNA samples, in order to assess possible contamination of the samples themselves.

The procedure for the replication was not changed from the previous experiment, with the exception of the individual primer concentrations for JHE (HE1) and cyt-c (HE2). This alteration was due to over expression of these primers in certain samples during the previous experiment. This dilution resulted in the final reverse multiplex for this investigation, and was utilised throughout the remaining procedures (see fig.4). Whilst this set of reactions did result in a quantifiable level of amplification, the DNA peak of 280 was still present throughout several samples. As a result these RNA samples were discarded, in order to eliminate the possibility of DNA contamination. A large amount of variance also existed within each sample, with large fluctuations occurring in peak size. Whilst these peaks were still within quantifiable means, it was possible that further RNA dilution would result in a more uniform level of amplification.

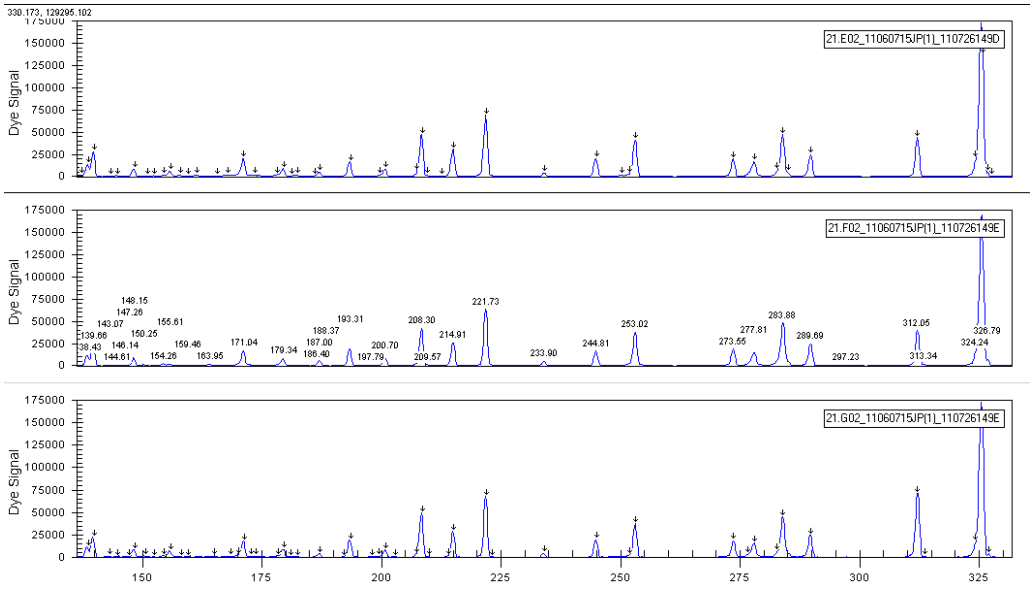

Figure 21: Final RNA spectra. These final spectra gave peaks that were well within the dynamic range of quantification, allowing further data analysis to occur.

A new stock of RNA samples were subjected to DNase treatment, followed by dilution to a concentration of 30ng/µL. The reverse-transcription reaction was conducted as previously mentioned, after which the cDNA was then separated into two separate samples. From these samples two individual PCR reactions were then

run, in order to increase the number of replicates. Previous experiments had resulted in a large amount of failed reactions, and this method was utilised in order to ensure that an adequate amount of replicates could be produced for each larval sample. Despite this, the experiment did not result in enough successful reactions to determine a valid average for quantification. Due to this, the experiment was again repeated, with a slight change in methodology. During the reverse transcription PCR, the reaction was run at a total of 15µL, from which 3 individual PCR runs were produced. This resulted in an adequate number of replicates for further analysis to be conducted (see figure 21). The investigation then continued to the data analysis stage.

3.2 Data Analysis:

3.2.1 Preliminary Analysis:

After Capillary Electrophoresis was utilised in order to generate appropriate spectra for each sample, quantification analysis was performed. Individual spectra were analysed using the GeXP Profiling tool of the Genome Lab System software package. Each sample was assessed for peak quality, alongside possible DNA contamination. The raw expression level of each sample was then exported into Microsoft Excel, in order to conduct data normalisation.

3.2.2 Normalisation of Gene Expression Data:

With a raw level of expression now established, gene normalisation was conducted in order to reduce the overall variation between individual samples. This normalisation process involves using the most stable reference (housekeeping) genes to build a normalisation factor. The quantification level for each gene of interest is then divided by the factor, in order to produce a final level of quantification. This reduces the overall level of variation between individual samples, as the final quantifications are analysed in regard to the level of reference gene expression. As this investigation aims to find any change in the level of expression of the genes of interest, any change in the ratio of amplification is more relevant than any changes in the amplification signals within samples. This is due to variance in the dye signal between individual samples, which may suggest a real difference in expression unless their relation to stable reference genes is also taken into account.

However, there is a growing debate within the literature as to how many genes should be utilised in order to generate a stable and accurate normalisation factor. The vast majority of real-time qPCR investigations currently use a minimum of two housekeeping genes in order generate a normalisation, though some investigations occasionally still use a single gene as a reference, usually actin (Thellin et al., 1999).

Recently a growing trend has emerged for investigations using a larger amount of reference genes, generating normalisation factors using up to five reference genes. The main argument for this approach is that this decreases the residual variance between individual samples, and eliminates any fluctuations within the reference genes themselves (Thellin et al., 1999). The main problem with this approach is that these reference genes have to be carefully screened in order to access their overall stability. Several software packages have been created in order to test reference gene candidates for their suitability in bees. The most commonly used of these packages are GeNorm, Bestkeeper, and Normfinder (Scharlaken et al., 2008) . This investigation used Normfinder in order to select our reference genes. The results of this analysis can be found in figure 22.

Normfinder was originally chosen for candidate gene analysis due to its calculation of not only the most stable gene, but the most stable combination of two genes. Unfortunately, there is no option for generation of a combination of more than two genes, so generation of a larger normalisation factor was dependant on the usage of the raw stability output for each housekeeping gene.

Gene name	Stability value
Actin	0.070
aTub	0.144
JHE	0.157
EF-1a-f1	0.179
GAPDH	0.180
TIF	0.213
gst	0.224
SAM	0.259
ubq	0.277
Cyt-C	0.297
rbp18	0.299
Cox-1	0.301
sdha	0.507

Best gene	Actin
Stability value	0.070
Best combination of two genes	Actin and JHE
Stability value for best combination of two genes	0.047

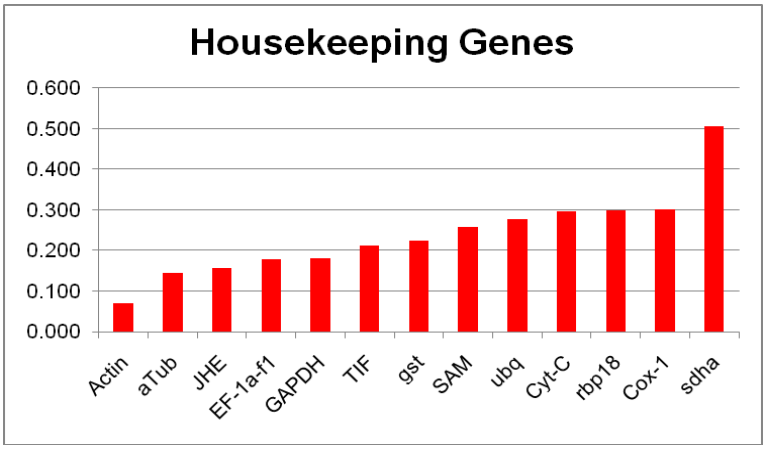

Figure 22. Normfinder selection of appropriate candidate genes. **A** Stability values for individual reference candidates **B** Most stable gene and best gene combination for two gene normalisation factor generation **C** Graph displaying the relative stability values of each candidate gene. The specific genes of interest were excluded from this analysis; however, preliminary analysis did include these genes. Whilst actin was still the most stable gene, actin and Dnmt1b were listed as the best gene combination. This immediately indicated that Dnmt1b was unlikely to have undergone a significant level of differential expression, as its expression was stable enough to qualify as a suitable housekeeping gene.

Before proceeding to statistical analysis, the rt-qPCR data was normalised using both the two gene normalisation factor suggested by the Normkeeper analysis (Actin and JHE), and a five gene normalisation factor comprised of the genes with the lowest stability value (Actin, aTub, JHE, EF-1a-f1, GAPDH, and TIF). This would allow direct comparison between the residual variance produced by each approach.

3.2.3 Statistical Analysis:

The two sets of normalised data were individually exported into the statistical package Genstat. During data exportation, control/experimental, hive number, and sample number were all listed as experimental factors. Analysis of Variance (ANOVA) was then conducted in order to determine the possibility of significant different gene expression between the experimental and control hives. Alongside direct comparison of expression levels between treated and non-treated hives, larval weight was included as a co-variate. This was in order to determine whether changes in larval weight resulted in differential gene expression. If differences in larval weight do result in differential gene expression, it is possible that this could mask any differential expression caused by imidacloprid exposure. More importantly, changes in expression due to larval weight could be mistaken for a real effect of imidacloprid exposure, producing a 'false-positive' result.

A separate ANOVA was performed for each individual gene of interest (Dnmt1a, 1b, 3, MDB, SAM) with the exception of Dnmt 2, which did not produce replicable expression levels. The results of these ANOVA's can be found in table 4. As can be quite clearly seen from the probability values generated during each ANOVA, none of the genes indicated a significant level of differential gene expression. For many genes the standard error was substantial, and in some cases, larger than the difference between the averages of the control and experimental hive groups respectively. Samples from the same hives were characterised by a low amount of variation, but samples from different hives (including within both treatment groups) displayed a much higher level of variance. There was a noticeable significance between data normalised with the two gene factor and data normalised using the five gene factor.

	2 Gene Normalisation Factor				5 Gene Normalisation Factor			
Gene of Interest	P-value	S.E.D	Mean (control)	Mean (treatment)	P-value	S.E.D	Mean (control)	Mean (treatment)
Dnmt 1a	0.442	0.026	0.199	0.176	0.583	0.023	0.257	0.243
Dnmt 1b	0.205	0.030	0.280	0.231	0.886	0.088	0.606	0.619
Dnmt 3	0.522	0.180	0.734	0.604	0.677	0.778	1.710	2.070
MBD	0.451	0.060	0.290	0.239	0.971	0.192	0.653	0.660
SAM_TM6	0.614	0.069	0.262	0.301	0.227	0.166	0.616	0.868

Table 5. ANOVA Output for Normalised Data. As can be seen above, no genes of interest achieved a critical P-vale of 0.05. However, widespread fluctuations in p-value did occur between the 2 gene normalisation data and the 5 gene normalisation data.

	2 Gene Normalisation Factor	5 Gene Normalisation Factor
Gene of Interest	Co-variate P-Value	Co-variate P-Value
Dnmt 1a	0.442	0.587
Dnmt 1b	0.205	0.796
Dnmt 3	0.314	0.885
MBD	0.545	0.952
SAM_TM6	0.421	0.192

Table 6. ANOVA output for larval weight co-variate. As can be seen above, no genes of interest achieved a critical P-vale of 0.05 for the larval weight co-variate. This indicates that differences in larval weight do not result in differential gene expression. Again, widespread fluctuations in p-value did occur between the 2 gene normalisation data and the 5 gene normalisation data

Due to the lack of any significant differential expression, combined with the high variance between hives, a power calculation was performed. This was in order to determine whether the number of hives used within each experimental group was adequate to overcome the variance between individual hives, and thus the probability that the ANOVA would be able to detect a real differential response (i.e. the 'power' of the ANOVA to find a significant change). The power calculation indicated that in order for the ANOVA to have a 95% (i.e. significant) chance of finding a real difference between the two treatment groups, between 11 and 12 hives would be required for each treatment group.

3.3 Transcriptome Sequencing:

3.3.1 EdgeR Analysis:

High-throughput AB SOLiD platform sequencing was conducted by Genomnia, a sequencing laboratory in Milan, Italy, on the six previously mentioned larval samples. This sequencing resulted in a total of 327 identifiable tags for known microRNA transcripts across both experimental groups (control and treated). Known sequence tags for honey bee microRNA transcripts are limited to those profiled by Chen et al. (2010), which were also derived through the use of SOLiD platform sequencing. These tag identifiers were organised as a table of counts for each transcript library with each library representing a single hive (C1, C2, C3, E1, E2, E3). These libraries were read into edgeR utilising a tab-delimited plain-text 'targets file', as shown in figure 23. The following analysis was then performed by myself.

files	group	description
C1.txt	C	Control Hive
C2.txt	C	Control Hive
C3.txt	C	Control Hive
E1.txt	E	Exp Hive
E2.txt	E	Exp Hive
E3.txt	E	Exp Hive

Figure 23. Tab-delimited plain-text targets file. This file organises the libraries by group, to which descriptions can be assigned.

This allows production of a Differential Gene Expression (DGE) List object for the use of any subsequent functions, as shown by figure 24. This summarises the amount of candidate tags, along with individual library size. From these possible tags, quality control was performed to filter out and low expressers. This was conducted in order to remove any tags that could not possible achieve any statistical significance (i.e. tags with less than 5 counts) for differential expression (DE). This action reduced the amount of candidate tags from 327 candidates to 291 across all 6 libraries (Robinson et al., 2011).

Due to the possibility of a small number of genes being highly expressed in one sample, but not in another, normalisation was carried out. This high expression can result in the rest of the DGE library being under sampled. To account for this, normalisation is utilised to estimate any bias and build it into the library size. The resulting normalisation factors were all very close to one, indicating a similar composition in all six libraries (Robinson et al., 2011).

```
files group  description lib.size norm.factors
1 C1.txt     C Control Hive  1028966          1
2 C2.txt     C Control Hive  1219608          1
3 C3.txt     C Control Hive  1021440          1
4 E1.txt     E     Exp Hive  1367681          1
5 E2.txt     E     Exp Hive   739228          1
6 E3.txt     E     Exp Hive   767586          1

$counts
                  1     2     3     4     5     6
ame-mir-1-2   35892 45221 36757 49717 28863 32468
ame-mir-1-2    474   493   500   653   282   358
ame-mir-10-1  1460  1606  1338  1826  1447  1110
ame-mir-10-1   121   157   163   198   234   122
ame-mir-10-2  1432  1581  1327  1760  1432  1096
308 more rows ...
```

Figure 24. DGE Readout. This display confirms that the tables of counts contained within our target files have been correctly read into the R session. The function readDGE calculates the sizes of individual count libraries and produces a DGEList object from which can be utilised for subsequent functions.

With each library fed into edgeR and subject to normalisation, it was now possible to conduct analysis of the DGE data. The first step of this analysis was to estimate the dispersion parameter for each tag. The most frequently used method for estimating these when analysing DGE data is through the use of a common distribution. For the vast majority of DGE data, such an analysis is adequate, preventing the requirement of tagwise dispersions calculation (i.e. calculating every gene dispersion parameter individually). The generation of a common distribution is highly dependent on the generation of pseudocounts (i.e. the required counts for there to be a significant

difference in expression level between the two groups). Calculation of these pseudocounts relies on the assumption of the alternative hypothesis being correct, that there is a significant difference in the level of gene expression between the two datasets. A quantile-to-quantile method is then used to calculate these pseudocounts, which are then utilised as the count data for obtaining an estimate of the common distribution. The size of the pseudocount libraries generated for the microRNA data was close to that of the original count data. It is optimum to produce psuedocount libraries of a similar size to the original counts, as this indicates the content of these libraries has not been significantly altered. Furthermore, each pseudocount library closely corresponds with the common library size of the original counts (Robinson et al., 2011).

The square root of the common distribution gives the coefficient of variation of biological variation, when using a negative binomial model. The coefficient of variation of biological variation in this investigation was shown to be 0.0905, indicating that there is a much greater level of variation than it would be possible to account for using a Poisson model. In this case of a library containing 291 tags, the given estimate of the tags variation would be over 30 times greater than that given under a Poisson distribution (Robinson et al., 2011).

With the common distribution for the microRNA data now generated, it can be tested in order to determine any differential expression. The edgeR package utilises a test which strongly parallels with Fisher's exact test, in order to compute p-values that be used to assess any differential expression. This allows a direct pair-wise analysis of the respective groups to be conducted. Once analysis is complete, a top-tags table can be produced in order to identify any DGE candidates. The results of this exact test are shown in figure 25 9 (Robinson et al., 2011).

	logConc	logFC	P.Value	FDR
ame-mir-3734	-15.64	3.1808	4.11E-11	1.29E-08
ame-mir-3801	-15.72	-2.1013	1.24E-05	1.94E-03
ame-mir-3783	-17.59	2.4528	5.67E-04	5.92E-02
ame-mir-3720	-15.39	1.3464	3.05E-03	1.78E-01
ame-mir-92a	-16.19	1.4812	3.16E-03	1.78E-01
ame-mir-9a	-16.79	1.67	3.41E-03	1.78E-01
ame-mir-282	-12.62	0.9996	7.17E-03	3.20E-01
ame-mir-971	-15.6	1.1559	1.07E-02	4.02E-01
ame-mir-14	-13.23	-0.9609	1.16E-02	4.02E-01

Figure 25. Top Tags List. These are the candidate tags that have the highest probability of displaying differential gene expression. The log-fold change (logFC) indicates the direction of this differential expression.

As the p-values in this figure indicate, we have 7 strong microRNA candidates (P<0.01) that indicate differential gene expression between the experimental and control groups (E-C). Figure 26 indicates that for the raw count data, there are very large differences between individual groups for these differential genes. This suggests that the expression is truly differential, as otherwise would indicate the possibility of a false-positive result. Further edge R analysis indicates that of these 7 candidate genes, 6 are up-regulated in experimental larval samples, whereas only one transcript is down-regulated. Whilst 7 of these tags were judged to be strong candidates (P<0.01), adjustment for multiple testing under the Benjamini and Hochberg (1995) method indicated two transcripts that also displayed significant differentiation of expression between the experimental and control groups. One of these samples was up-regulated in the experimental group, whilst the other was down-regulated. The properties of these micro-RNA's are listed in table 7.

	1	2	3	4	5	6
ame-mir-3734	8	3	10	5	3	134
ame-mir-3801	3	126	7	9	8	8
ame-mir-3783	7	0	0	16	9	9
ame-mir-3720	18	11	18	55	31	22
ame-mir-92a	13	8	5	8	10	39
ame-mir-9a	2	4	10	7	7	27
ame-mir-282	102	131	134	184	206	203
ame-mir-971	17	14	13	39	17	30
ame-mir-14	171	117	179	109	76	32

Figure 26. Raw tag count comparisons across individual libraries.

4: Discussion

4.1 Investigative Aims:

This investigation was conducted with two major aims:

1. To determine whether low level Imidacloprid exposure results in differential transcript expression in honey bee larvae.

2. To conduct a bioinformatic analysis of transcriptomic data generated by SOLiD platform sequencing in order to draw conclusions about the effects of larval imidacloprid contamination.

These aims were accomplished through the use of direct larval RNA extraction, in order to analyse the effects of chronic low-level imidacloprid contamination on the DNA Methyltransferase enzyme family. Quantitative real-time multiplex PCR was performed in order to ascertain whether or not this contamination resulted in differential levels of gene expression. Of the four Dnmt homologues present in the honey bee, three (Dnmt1a, 1b, and 3) were analysed for differential gene expression, alongside two related genes of interest (MBD, SAM). After this data was subjected to an Analysis of Variance (ANOVA), it was concluded that none of these candidate genes indicated a significant difference ($P<0.05$) in the level of gene expression between treated and non-treated hives. Alongside this conclusion, it was also ascertained that differences in weight between individual larvae do not result in a significant amount of differential gene expression.

Concurrently with this experimental validation, a bioinformatic study was conducted utilising transcriptomic data generated from both treated and non-treated larvae. This data was organised as a series of tag identifiers of known microRNA sequences.

Bioinformatic analysis was then conducted through the use of the R-based software package Edge-R. This analysis presented seven strong gene candidates for differential gene expression ($P<0.01$), and two possible candidates ($P<0.05$). These candidates and their respective probabilities are presented in Table 8. From this analysis it can be concluded that exposure to chronic levels of imidacloprid toxicity does result in some differential expression of known microRNAs.

Whilst these investigations have drawn differing conclusions, it is possible that imidacloprid contamination may result in some differential gene expression. Although rt-qPCR has shown that this level of contamination does not result in differential expression of the Dnmts, the differential candidates identified during the transcriptome analysis may be equally important for maintaining normal colony function. As a result, any change in their expression may be damaging to overall colony fitness. However, of these nine microRNA's, only one has been significantly documented. This is miR-9a which has been previously investigated within genetic model organisms.

4.1.1 MiR-9a:

The MiR-9 family has been shown to be directly correlated with the onset of glioma tumour development in the mouse brain. Any suppression of miR-9 results in enhanced SOX2 protein expression, which in turn induces increased expression of ABCC3/ABCC6 (Jeon et al., 2011). This process results in the proliferation of ATP-binding cassette transporters, allowing glioma stem-like cells to exhibit highly chemoresistant behaviour. However, it is unknown as to whether this pathway is retained across such a wide evolutionary gap. Real-time PCR analysis has also identified miR-9 as a potential biomarker in Biliary Tract cancer (Shigehara et al., 2011), due to its higher levels of expression in malignant tissues. Broadly speaking, differential expression of this transcript is correlated with many cancer phenotypes, often acting as a contributing factor in further tumour proliferation. For example, in

breast cancer stem cells, miR-9 is induced by Myc, targeting E-cadherin (Ma et al., 2010). E-cadherin is the major epithelial adherens junction protein, and its removal primes the cells for epithelial-mesenchymal transition, resulting in the stimulation of angiogenesis within tumours. These newly established blood vessels can then contribute to metastasis, as well as vastly increasing the growth and regenerative abilities of the tumour.

Most notably, the micro-RNA generated by this edgeR analysis implicates miR-9a, which has no established background with cancer development. A more relevant role for miR-9a has been documented in *Drosophila melanogaster*, a model organism displaying a greater degree of evolutionary similarity with Apis. Knocking out this gene resulted in a significant loss in wing tissue through apoptosis. This loss was mainly traced to the posterior wing margin, and was a consequence of de-repressing *Drosophila* LIM-only (dLMO)(Biryukova, et al., 2009). A subtle increase in sensory bristles was also noted. The existence of this pathway has not been proven in Apis, but its existence may indicate that differential expression of the microRNA could have a significant effect on the honey bee phenotype. However, our transcriptomics data found that this particular transcript was up regulated in the exposed larvae.

Another important regulatory role for miR-9a regards neurogenesis. Experiments performed by Li et al. (2006) found that loss of miR-9a function in the Drosophila peripheral nervous system caused widespread ectopic production of sensory organ precursors (SOPs). More relevantly, over expression of miR-9a results in a severe loss of SOPs. This loss is due to the inhibitory effect that miR-9a has on the expression of the transcription factor senseless (sens). A high level of this transcription factor is required for proneural development; non-SOP cells see a down regulation in this factor as a result. This is accomplished in order to generate a precise number of neural precursor cells during development. As SOP development is integral for the development of the peripheral nervous system, it is possible that this differential

expression could be damaging to the detection and information processing capabilities of developing larvae.

4.2 Limitations:

As this was part of the first investigation conducted by the lab in this area, there were inevitably some limitations to the investigation. From the start there were a few limitations to the project that in particular affected both its formulation and implementation. Due to the many factors required for a honey bee colony to function, it is almost impossible to replicate these conditions in a laboratory setting. As a result only a limited amount of control could be exerted over the degree of environmental stimuli that each hive came into contact with. The hives utilised for this investigation were located in a dense area of arable land, and there remains the slight possibility that other pesticides may have been used in the local area. Whilst hives were screened visually for any obvious signs of disease or parasites, financial reasons prevented any assessment of colony parasite weight.

Most importantly, the hives from which these larvae were selected were not directly owned by the department conducting the investigation, instead being donated by a local beekeeper. Due to these bees also being used in a production capacity, the level of contamination allowed was also limited to prevent any significant damage to hive performance. The number of hives used within the investigation was also limited as a result, though this number was largely selected due to similar investigations employing a similar number of hives (Beliën et al., 2009).

Within the larger investigation, this project was also originally intended to focus on the full larval transcriptome sequencing data generated by both the University of Nottingham and the Genomnia institute in Italy. However, due to the amount of

time for the whole transcriptome to be sequenced, its focus was switched exclusively onto the micro-RNA data generated by Genomnia.

4.3 Further Work:

Any further studies into the effects of imidacloprid contamination should take into account some of the problems and limitations encountered by this investigation. Due to a low number of replicates being conducted, it may be beneficial to repeat the investigation using a larger number of hives within both experimental groups. Power calculations have indicated that in order for a one-way ANOVA to have significant chance (i.e. power) of finding a real difference between our two groups, each group would have to consist of at least 12 replicates. However, it must be considered that repeating such an investigation may be difficult in terms of finance. As such, it may be prudent to seek funding for a new investigation, in which individual hives would be exposed to a range of different imidacloprid concentrations. Whilst studies have been performed looking at the acute affects of imidacloprid contamination concerning both overall colony mortality (Beliën et al., 2009) and its effects on reflexes and learning processes (Lambin et al., 2001), the actual molecular effects of acute exposure have not been investigated in any depth. As such, it may be possible to perform an investigation comparing the transcriptome wide effects of both acute and chronic treatment. Analysis of the full larval transcriptome is currently being conducted, and may reveal more differential candidate genes. If the transcriptome of the larval samples receiving an acute level of imidacloprid contamination were then to be sequenced, the two transcriptomes could be compared alongside the previously sequenced control larvae. A possible hypothesis for this experiment would be that ‘acute exposure to imidacloprid causes a greater amount of differential gene expression than chronic exposure’.

However, there are some limitations to this approach, mostly related to some of those previously mentioned regarding this investigation. As the hives utilised by this

investigation were owned by a local beekeeper rather than the department, the possibility of acute imidacloprid exposure damaging colony productivity could risk commercial losses. The proposed increase in the number of hives may amplify this possibility. Based off these problems, any lab willing to conduct this investigation would require the financial means to independently purchase and maintain a large enough number of colonies to produce reliable results. However, this may cost less than any actual sequencing..

Another possible direction to take with any further work would be to focus on the differential candidates found within the transcriptome analysis. Whilst our documented differential gene candidate, miR-9a, has been shown to have significant regulatory roles in *Drosophila*, it is not known if this microRNA performs the same role in *Apis*. In order to confirm whether or not this role is conserved, knockouts could be performed in order to produce miR-9a mutants. Whilst the functional roles of the other differential candidates produced by the transcriptome analysis are not currently documented, knockouts could similarly be produced. Any changes in phenotype may garner more information as to their overall regulatory role, and shed light as to whether disruption of this role could have any significant differential effect on colony health/performance.

It is also possible to draw some conclusions about the number of sensory organ precursors (SOPs) within the peripheral nervous system using the level of *sens* expression. A possible avenue of work would be to compare this level of expression between control and experimental larvae, in order to ascertain whether the downregulation of miR-9a is having a significant effect on the number of SOPs in the peripheral nervous system. A potential working hypothesis for this investigation would be 'Chronic exposure to imidacloprid does result in differential expression of *senseless* in the Honey bee.'

References:

1. Aikin, R. C. (1897) Bees evaporated; a new malady*, Gleanings in Bee Culture*, **25**: 479-480.

2. Aizen, M. A. and Harder L.D. (2009) The global stock of domesticated honey bees is growing slower than agricultural demand for pollination, *Current Biology*, **19**: 1-4.

3. Anderson, J. (1930) The 'Isle of Wight' disease: in bees, *Bee World*, **11**: 37-42.

4. Arias, M. C. and Sheppard, W. S. (2005) Phylogenetic relationships of honey bees (Hymenoptera : Apinae : Apini) inferred from nuclear and mitochondrial DNA sequence data, *Molecular Phylogenetics and Evolution,* **37**(1): 25-35.

5. Bai, D. L., Lummis, S. C. R., Leicht, W., Breer, H. and Sattelle, D. B. (1991) actions of imidacloprid and a related nitromethylene on cholinergic receptors of an identified insect motor-neuron, *Pesticide Science,* **33**(2): 197-204.

6. Bailey, L. and Gibbs, A. J. (1964) Acute infection of bees with paralysis virus, *Journal of Insect Pathology,* **6**(4): 395-407.

7. Barchuk, A. R., Cristino, A. S., Kucharski, R., Costa, L. F., Simoes, Z. L. P. and Maleszka, R. (2007) Molecular determinants of caste differentiation in the highly eusocial honeybee Apis mellifera, *Bmc Developmental Biology,* **7**: 70.

8. Barker, R. J. and Lehner, Y. (1972) Free amino-acids in thoraces of flown honey bees, apis-mellifera l (hymenoptera-apidae), *Comparative Biochemistry and Physiology,* **43**(1B): 163-170.

9. Beliën, T., Kellers, J., Heylen, K., Keulemans, W., Billen, J., Arckens, L., Huybrechts, R. and Gobin, B. (2009) Effects of sublethal doses of crop protection agents on honey bee (Apis mellifera) global colony vitality and its potential link with aberrant foraging activity, *Communications in agricultural and applied biological sciences,* **74**(1): 245-253.

10. Benjamini, Y. and Hockberg, Y. (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing, *Journal of the Royal Statistical Society*, **57**(1): 289-300.

11. Bestor, T. H. (2000) The DNA methyltransferases of mammals. *Human Molecular Genetics,* **9:** 2395–2402.

12. Bestor, T. H. and Ingram, V. M. (1983) 2 DNA methyltransferases from murine erythroleukemia-cells - purification, sequence specificity, and mode of interaction with DNA, *Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences,* **80**(18): 5559-5563.

13. Bestor, T.H. (1990) DNA methylation: evolution of a bacterial immune function into a regulator of gene expression and genome structure in higher

eukaryotes, *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, **326**(1235): 179-187.

14. Betts, A. D. (1919) Fungus diseases of bees. *Bee World*, **1**: 132.

15. Beuhne, R. (1910) Bee mortality, *Journal of the Department of Agriculture of Victoria*, **7**: 149-151.

16. Biesmeijer, J. C., Roberts, S. P. M., Reemer, M., Ohlemuller, R., Edwards, M., Peeters, T., Schaffers, A. P., Potts, S. G., Kleukers, R., Thomas, C. D., Settele, J. and Kunin, W. E. (2006) Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands, *Science,* **313**(5785): 351-354.

17. Biryukova, I., Asmar, J., Abdesselem, H., and Heitzler, P. (2009) Drosophila miR-9a regulates wing development via fine-tuning expression of the LIM only factor, dLMO, *Developmental Biology*, **327**(2): 487-496.

18. Blackiston, H., The bee as a pupa, accessed 10th November 2011, <http://www.dummies.com/how-to/content/tracking-the-life-cycle-of-a-honey-bee.html>

19. Blackiston, H., The honey bee lifecycle, accessed 10th November 2011, <http://www.dummies.com/how-to/content/tracking-the-life-cycle-of-a-honey-bee.html>

20. Bortolotti, L., Montanari, R., Marcelino, J., Medrzycki, P., Maini, S. and Porrini, C. (2003) Effects of sub-lethal imidacloprid doses on the homing rate and foraging activity of honey bees, *Bulletin of Insectology,* **56**(1): 63-67.

21. Breer, H. and Sattelle, D. B. (1987) Molecular-properties and functions of insect acetylcholine-receptors, *Journal of Insect Physiology,* **33**(11): 771-790.

22. Breese, C. R., Lee, M. J., Adams, C. E., Sullivan, B., Logel, J., Gillen, K. M., Marks, M. J., Collins, A. C. and Leonard, S. (2000) Abnormal regulation of high affinity nicotinic receptors in subjects with schizophrenia, *Neuropsychopharmacology,* **23**(4): 351-364.

23. Brooks, R. W. and Engel, M. S. (1999) A revision of the augochlorine bee genus Chlerogas Vachal (Hymenoptera : Halictidae), *Zoological Journal of the Linnean Society,* **125**(4): 463-486.

24. Bruegel, P. (1568) The bee-keepers, accessed 10th November 2011, <http://amycrehore.blogspot.com/2010/07/bee-keepers-1568-by-pieter-bruegel.html>

25. Buckingham, S. D., Balk, M. L., Lummis, S. C. R., Jewess, P. and Sattelle, D. B. (1995) Actions of nitromethylenes on an alpha-bungarotoxin-sensitive neuronal nicotinic acetylcholine-receptor, *Neuropharmacology,* **34**(6): 591-597.

26. Bullamore, G. W. (1922) *Nosema apis* and *Acarapis (Tarsonemus) woodi* in relation to Isle of Wight disease, *Parasitology* **14**: 53-62.

27. Burnside, C. E. (1930) Fungous Diseases of the Honeybee, *U.S. Department of Agriculture Technical Bulletin*, **149**: 1-43.

28. Cano Lozano, V., and Armengaud, C. (2001) Memory impairment induced by cholinergic antagoinsts injected into the mushroom bodies of the honeybee, *Journal of Comparative Physiology*, **187**(4): 249-254.

29. Capaldi, E. A., Smith, A. D., Osborne, J. L., Fahrbach, S. E., Farris, S. M., Reynolds, D. R., Edwards, A. S., Martin, A., Robinson, G. E., Poppy, G. M. and Riley, J. R. (2000) Ontogeny of orientation flight in the honeybee revealed by harmonic radar, *Nature,* **403**(6769): 537-540.

30. Cardinal, S. and Danforth, B. N. (2011) The Antiquity and Evolutionary History of Social Behavior in Bees, *Plos One,* **6**(6).

31. Chapman, N. C., Nanork, P., Gloag, R. S., Wattanachaiyingcharoen, W., Beekman, M. and Oldroyd, B. P. (2009) Queenless colonies of the Asian red dwarf honey bee (Apis florea) are infiltrated by workers from other queenless colonies, *Behavioral Ecology,* **20**(4): 817-820.

32. Cho, S. C., Huang, Z. Y., Green, D. R., Smith, D. R. and Zhang, J. Z. (2006) Evolution of the complementary sex-determination gene of honey bees: Balancing selection and trans-species polymorphisms, *Genome Research,* **16**(11): 1366-1375.

33. Colhoun, E. H. and Smith, M. V. (1960) Neurohormonal properties of royal jelly, *Nature,* **188**(4753): 854-855.

34. Colony collapse disorder, accessed 10th November 2011, < http://bee-rapture.blogspot.com/2009/04/found-cause-of-colony-collapse-disorder.html>

35. Cox, R. L. and Wilson, W. T. (1987) The behaviour of insecticide-exposed honey-bees, *American Bee Journal,* **127**(2): 118-119.

36. Danka, R. G. and Villa, J. D. (2000) Inheritance of resistance to Acarapis woodi (Acari : Tarsonemidae) in first-generation crosses of honey bees (Hymenoptera : Apidae), *Journal of Economic Entomology,* **93**(6): 1602-1605.

37. de Miranda, J. R., Cordoni, G. and Budge, G. (2010) The Acute bee paralysis virus-Kashmir bee virus-Israeli acute paralysis virus complex, *Journal of Invertebrate Pathology,* **103**: S30-S47.

38. Di Prisco, G., Pennacchio, F., Caprio, E., Boncristiani, H. F., Jr., Evans, J. D. and Chen, Y. (2011) Varroa destructor is an effective vector of Israeli acute paralysis virus in the honeybee, Apis mellifera, *Journal of General Virology,* **92**: 151-155.

39. Engel, M. S. (1999) The taxonomy of recent and fossil honey bees (Hymenoptera: Apidae; Apis), *Journal of Hymenoptera Research,* **8**(2): 165-196.

40. Engel, M. S. (2000) A new interpretation of the oldest fossil bee (Hymenoptera, Apidae), *American Museum Novitates,* **3296**: 1-11.

41. European honey bee extracts nectar, accessed 10th November 2011, <http://en.wikipedia.org/wiki/File:European_honey_bee_extracts_nectar.jpg >

42. Fantham, H. B. and Porter, A. (1912) The structure and homology of the microsporidian spore, as seen in Nosema apis, *Proceedings of the Cambridge Philosophical Society,* **16**: 580-583.

43. Faucon, J. P., Mathieu, L., Ribiere, M., Martel, A. C., Drajnudel, P., Zeggane, S., Aurieres, C. and Aubert, M. F. A. (2002) Honey bee winter mortality in France in 1999 and 2000, *Bee World,* **83**(1): 14-23.

44. Field, L. M., Lyko, F., Mandrioli, M. and Prantera, G. (2004) DNA methylation in insects, *Insect Molecular Biology*, **13**(2): 109-115.

45. Fries, I., Martin, R., Meana, A., Garcia-Palencia, P. and Higes, M. (2006) Natural infections of Nosema ceranae in European honey bees, *Journal of Apicultural Research,* **45**(4): 230-233.

46. Gade, G., Hoffmann, K. H. and Spring, J. H. (1997) Hormonal regulation in insects: Facts, gaps, and future directions, *Physiological Reviews,* **77**(4): 963-1032.

47. Gary, N. E., Witherell, P. C. and Marston, J. M. (1975) Distribution of foraging honey bees from colonies used for honeydew melon pollination, *Environmental Entomology,* **4**(2): 277-281.

48. Giauffret A., Quiot, J. M., Vago, C. and Poutier, F. (1967) In vitro culture of bee cells, *Comptes Rendus Hebdomadaires Des Seances De L Academie Des Sciences Serie D,* **265**(11): 800-811.

49. Goll, M. G. and Bestor, T. H. (2005) Eukaryotic cytosine methyltransferases, *Annual Review of Biochemistry,* **74**: 481-514.

50. Goll, M. G., Kirpekar, F., Maggert, K. A., Yoder, J. A., Hsieh, C. L., Zhang, X. Y., Golic, K. G., Jacobsen, S. E. and Bestor, T. H. (2006) Methylation of tRNA(AsP) by the DNA methyltransferase homolog Dnmt2, *Science,* **311**(5759): 395-398.

51. Gordon, D. M. (1989) Caste and change in social insects, in *Oxford surveys in evolutionary biology* (ed. P. H. Harvey & L. Partridge), pp. 67-72. Oxford University Press.

52. Gundelfinger, E. D. (1992) How complex is the nicotinic receptor system of

insects, *Trends in Neurosciences,* **15**(6): 206-211.

53. Hagenguth, H. and Rembold, H. (1978) Identification of juvenile hormone-3 as the only jh homolog in all developmental stages of the honey bee, *Zeitschrift Fur Naturforschung C-a Journal of Biosciences,* **33**(11-1): 847-850.

54. Honey bee castes, West Mountain Apiary, accessed 10th November 2011, <http://www.recycle.co.uk/blog/save-the-bees-%E2%80%93-a-global-crisis/>

55. Hour, K.F. (2011) Homecoming queen bee? Accessed 10th November 2011, <https://podcast.ad.edenpr.org/groups/ephsenrichedbiology2011/weblog/289c2/Homecoming_Queen_Bee.html>

56. Huber, F. (1814) New Observations Upon Bees, Translated by C.P. Dadant, 1926, *American Bee Journal*, Hamilton, IL.

57. Jeon, H., Sohn, Y., Oh, S., Kim, S., Beck, S., Kim, S. and Kin, H. (2011) ID4 imparts chemoresistance and cancer stemness to glioma cells by derepressing miR-9-mediated suppression of SOX2. *Cancer Research*, **71**: 3410-3421.

58. Johnson, W. B. (1995) Bee miles, *American Bee Journal,* **135**(7): 445-445.

59. Jurkowski, T. P., Meusburger, M., Phalke, S., Helm, M., Nellen, W., Reuter, G. and Jeltsch, A. (2008) Human DNMT2 methylates tRNA(Asp) molecules using a DNA methyltransferase-like catalytic mechanism, *Rna-a Publication of the Rna Society,* **14**(8): 1663-1670.

60. Kagabu, S. (1997) Discovery of the chloronicotinyl insecticides, *Abstracts of Papers of the American Chemical Society,* **214**: 4-AGRO.

61. Kagabu, S. and Medej, S. (1995) Chloronicotinyl insecticides .6. stability comparison of imidacloprid and related-compounds under simulated sunlight, hydrolysis conditions, and to oxygen, *Bioscience Biotechnology and Biochemistry,* **59**(6): 980-985.

62. Kamakura, M. (2011) Royalactin induces queen differentiation in honeybees, *Nature,* **473**(7348): 478-483.

63. Klungness, L. M. and Peng, Y. S. (1984) A histochemical-study of pollen digestion in the alimentary canal of honeybees (*Apis mellifera*), *Journal of Insect Physiology,* **30**(7): 511-21.

64. Kotthoff, U., Wappler, T. and Enge, M. S. (2011) Miocene honey bees from the Randeck Maar of southwestern Germany (Hymenoptera, Apidae), *ZooKeys,* **96**: 11-37.

65. Kucharski, R., Maleszka, J., Foret, S. and Maleszka, R. (2008) Nutritional control of reproductive status in honeybees via DNA methylation, *Science,* **319**(5871): 1827-1830.

66. Kulincevic, J. M., Rothenbuhler, W. C. and Rinderer, T. E. (1984) Disappearing disease .3. a comparison of 7 different stocks of the honey bee (*Apis mellifera*), *Ohio Agricultural Research and Development Center Research Bulletin,* **1160**: 1-21.

67. Lambin, M., Armengaud, C., Raymond, S. and Gauthier, M. (2001) Imidacloprid-induced facilitation of the proboscis extension reflex habituation in the honeybee, *Archives of Insect Biochemistry and Physiology,* **48**(3): 129-134.

68. Langstroth, L. L. (1853) The hive and the honey-bee, a bee-keeper's manual. Northampton, Hopkins, Bridgman.

69. Larvae in royal jelly, accessed 10th November 2011, <http://en.wikipedia.org/wiki/File:Weiselzellen_68a.jpg>

70. Li, Y., Wang, F., Lee, J. A. and Gan, F. B. (2006) MicroRNA-9a ensures the precise specification of sensory organ precursors in *Drosophila*, *Genes and Development*, **20**: 2793-2805.

71. Littleton, J. T. and Ganetzky, B. (2000) Ion channels and synaptic organization: Analysis of the Drosophila genome, *Neuron,* **26**(1): 35-43.

72. Liu, M. Y. and Casida, J. E. (1993) High-affinity binding of H-3 imidacloprid in the insect acetylcholine-receptor, *Pesticide Biochemistry and Physiology,* **46**(1): 40-46.

73. Liu, M. Y., Lanford, J. and Casida, J. E. (1993) Relevance of H-3 imidacloprid binding-site in house-fly head acetylcholine-receptor to insecticidal activity of 2-nitromethylene-imidazolidines and 2-nitroimino-imidazolidines, *Pesticide Biochemistry and Physiology,* **46**(3): 200-206.

74. Ma, L., Young, J., Prabhala, H., Pan, E., Mestdagh, P., Muth, D., Teruya-Feldstein, J., Reinhardt, F., Onder, T. T., Valastyan, S., Westermann, F. Speleman, F., Vadesompele, J. and Weinberg, R. A. (2010) miR-9, a MYC/MYCN-activated microRNA, regulates E-cadherin and cancer metastasis, *Nature Cell Biology*, **12**: 247-256.

75. Martin, L. F., Kem, W. R. and Freedman, R. (2004) Alpha-7 nicotinic receptor agonists: potential new candidates for the treatment of schizophrenia, *Psychopharmacology,* **174**(1): 54-64.

76. Matsuda, K., Buckingham, S. D., Kleier, D., Rauh, J. J., Grauso, M. and Sattelle, D. B. (2001) Neonicotinoids: insecticides acting on insect nicotinic acetylcholine receptors, *Trends in Pharmacological Sciences,* **22**(11): 573-580.

77. Mayer, A. (2009) Save the bees – a global crisis, accessed 10th November 2011, < http://www.recycle.co.uk/blog/save-the-bees-%E2%80%93-a-global-crisis/>

78. McAllister, E. (2011) Varroa on a pupal, accessed 10th November 2011, < http://rbgeherbaljournal.blogspot.com/2011/01/18th-january-2011-chain-bridge-honey.html>

79. Mesarch, M.B., Nakatsu, C.H., Nies, L. (2000) Development of catechol 2,3-dioxygenase-specific primers for monitoring bioremediation by competitive quantitative PCR, Applied and Environmental Biology, **66**(2): 678-683.

80. Mitchell, A. (2006) Africanized killer bees: A case study, *Critical Care Nurse,* **26**(3): 23-26.

81. Momparler, R. L. and Bovenzi, V. (2000) DNA Methylation and cancer, *Journal of Cellular Physiology*, **183**(2): 145-154.

82. Mraz, C. (1977) Bee venom therapy, *American Bee Journal,* **117**(4): 260-260.

83. Munoz, I., Garrido-Bailon, E., Martin-Hernandez, R., Meana, A., Higes, M. and De la Rua, P. (2008) Genetic profile of Varroa destructor infesting Apis mellifera iberiensis colonies', *Journal of Apicultural Research,* **47**(4): 310-313.

84. Nguyen, B. K., Saegerman, C., Pirard, C., Mignon, J., Widart, J., Thirionet, B., Verheggen, F. J., Berkvens, D., De Pauw, E. and Haubruge, E. (2009) Does Imidacloprid Seed-Treated Maize Have an Impact on Honey Bee Mortality?, *Journal of Economic Entomology,* **102**(2): 616-623.

85. Nicot, N., Hausman, J.F., Hoffmann, L., Evers, D. (2005) Housekeeping gene selection for real-time RT-PCR normalization in potato during biotic and abiotic stress, Journal of Experimental Botany, **56**(421): 2907-2914.

86. Nosemoza, accessed 10th November 2011, <http://www.hodowle.eu/298_nosemoza.html>

87. Oertel, E. (1965) Many bee colonies die of an unknown cause, *American bee Journal*, **105**: 48-49.

88. Okano, M., Xie, S. P. and Li, E. (1998) Dnmt2 is not required for de novo and maintenance methylation of viral DNA in embryonic stem cells, *Nucleic Acids Research,* **26**(11): 2536-2540.

89. Oldroyd, B. P. (2007) What's killing American honey Bees?, *Plos Biology,* **5**(6): 1195-1199.

90. Olley, K. (1976) Those disappearing bees, *American Bee Journal,* **116**(11): 520-521.

91. Pain, J. A. and Mugenet, J. (1966) Biochemical and physiological research on pollen stored by bees, *Ann. Abeille*, **9**: 209-236.

92. Parker, R. L. (1926) The utilization of pollen and pollen substitutes by the honeybee, *Journal of Economic Entomology,* **19**: 161-163.

93. Peng, Y. S., Fang, Y. Z., Xu, S. Y. and Ge, L. S. (1987) The resistance mechanism of the asian honey-bee, *Apis cerana fabr*, to an ectoparasitic mite, *Varroa jacobsoni oudemans*, *Journal of Invertebrate Pathology,* **49**(1), 54-60.

94. Pfeifer, G.P., Steigerwald, S.D., Hansen, R.S., Gartler, S.M. and Riggs, A.D. (1990) Polymerase chain reaction-aided genomic sequencing of an X chromosome-linked CpG island: Methylation patterns suggest clonal inheritance, CpG site autonomy, and an explanation of activity state stability, *Proceedings of the National Academy of Sciences,* **87:** 8252–8256.

95. Robertson, K. D., Uzvolgyi, E., Liang, G., Talmadge, C., Sumegi, J., Gonzales, F. A. and Jones, P. A. (1999) The human DNA methyltransferses (Dnmts) 1, 3a and 3b: coordinate mRNA expression in normal tissues and overexpression in tumours, *Nucleic Acids Research*, **27**(11): 2291-2298.

96. Robinson, M., McCarthy, D., Chen, Y., Smyth, G.K. (2011) edgeR: differential expression analysis of digital gene expression data - User's Guide, accessed November 10th, <http://www.bioconductor.org/packages/2.9/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf>

97. Rosov, S. A. (1944) Food consumption by bees, *Bee World*, **25**: 94-95.

98. Satta, R., Maloku, E., Zhubi, A., Pibiri, F., Hajos, M., Costa, E. and Guidotti, A. (2008) Nicotine decreases DNA methyltransferase 1 expression and glutamic acid decarboxylase 67 promoter methylation in GABAergic interneurons, *Proceedings of the National Academy of Sciences of the United States of America,* **105**(42): 16356-16361.

99. Scharlaken, B., de Graaf, D. C., Goossens, K., Brunain, M., Peelman, L. J. and Jacobs, F. J. (2008) Reference gene selection for insect expression studies using quantitative real-time PCR: The head of the honeybee, *Apis mellifera*, after a bacterial bn challenge, *Journal of Insect Science*, 8(33): 1-10.

100. Schumacher, M. J. and Egen, N. B. (1995) Significance of africanized bees for public-health - a Review, *Archives of Internal Medicine,* **155**(19): 2038-2043.

101. Seeley, T.D. (1982) Adaptive significance of the age polyethism schedule in honeybee colonies. Behavioral Ecology and Sociobiology, 11: 287-293.

102. Sen, G. L., Reuter, J. A., Webster, D. E., Zhu, L. and Khavari, P. A. (2010) Dnmt1 maintains progenitor function in self-renewing somatic tissue. *Nature*, **463**: 563-567.

103. Shigehara, K., Yokomuro, S., Ishibashi, O., Mizuguchi, Y., Arima, Y. Kawahigashi, Y., Kanda, T., Akagi, I., Tajiri, T., Yoshida, H., Takizawa, T. and Uchida, E. (2011) Real-time PCR-based analysis of the human bile MicroRNAome identifies miR-9 as a potential diagnostic biomarker for biliary tract cancer. *PLoS One*, **6**(8): e23584.

104. Shiokawa, K., Tsuboi, S., Kagabu, S. and Moriya, K. (1995) Heterocyclic compounds, *Official Gazette of the United States Patent and Trademark Office Patents,* **1175**(4): 2500-2501.

105. Silver, N., Best, S., Jiang, J. and Thein, S. L. (2006) Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR, *Bmc Molecular Biology,* **7**: 33.

106. Smirle, M. J., Winston, M. L. and Woodward, K. L. (1984) Development of a sensitive bioassay for evaluating sublethal pesticide effects on the honey bee (Hymenoptera, Apidae), *Journal of Economic Entomology,* **77**(1): 63-67.

107. Smith, D. R., Villafuerte, L., Otis, G. and Palmer, M. R. (2000) Biogeography of Apis cerana F. and A-nigrocincta Smith: insights from mtDNA studies, *Apidologie,* **31**(2): 265-279.

108. Soloway, S. B., Henry, A. C., Kollmeyer, W. D., Padgett, W. M., Powell, J. E., Roman, S. A., Tieman, C. H., Corey, R. A. and Horne, C. A. (1979) Nitromethylene insecticides, H. Geissbühler, G.T. Brooks, P.C. Kearney, Editors *Advances in Pesticide Science*, **Vol. 2** Pergamon Press, Oxford (1979), p. 206.

109. Stein, K., Beekeeping in ancient Egypt, accessed 10th November 2011, < http://www.bee-hexagon.net/en/creativeexpression.htm>

110. Suchail, S., Guez, D. and Belzunces, L. P. (2001) Discrepancy between acute and chronic toxicity induced by imidacloprid and its metabolites in Apis mellifera, *Environmental Toxicology and Chemistry,* **20**(11): 2482-2486.

111. Sullivan, J. P., Fahrbach, S. E., Harrison, J. F., Capaldi, E. A., Fewell, J. H. and Robinson, G. E. (2003) Juvenile hormone and division of labor in honey bee colonies: effects of allatectomy on flight behavior and metabolism, *Journal of Experimental Biology,* **206**(13): 2287-2296.

112. Sullivan, J. P., Jassim, O., Fahrbach, S. E. and Robinson, G. E. (2000) Juvenile hormone paces behavioral development in the adult worker honey bee, *Hormones and Behavior,* **37**(1): 1-14.

113. Takano, J., The Mystery of Vanishing Honeybees (Colony Collapse Disorder), accessed 10th November 2011, <http://www.pyroenergen.com/articles09/colony-collapse-disorder.htm>

114. Tang, L. Y., Reddy, M. N., Rasheva, V., Lee, T. L., Lin, M. J., Hung, M. S. and Shen, C. K. J. (2003) The eukaryotic DNMT2 genes encode a new class of cytosine-5 DNA methyltransferases, *Journal of Biological Chemistry,* **278**(36): 33613-33616.

115. Thellin, O., Zorzi, W., Lakaye, B., de Borman, B., Coumans, B., Hennen, G., Grisar, T., Lgout, A. and Heinen, E. (1999) Housekeeping genes as internal

standards: use and limits, *Journal of Biotechnology*, **75**: 291-295.

116. Thom, C., Gilley, D. C., Hooper, J. and Esch, H. E. (2007) The scent of the waggle dance, *Plos Biology,* **5**(9): 1862-1867.

117. Tyvand, P. A. and Thorvaldsen, S. (2010) Wright-Fisher model of social insects with haploid males and diploid females, *Journal of Theoretical Biology,* **266**(3): 470-478.

118. vanEngelsdorp, D., Speybroeck, N., Evans, J. D., Nguyen, B. K., Mullin, C., Frazier, M., Frazier, J., Cox-Foster, D., Chen, Y., Tarpy, D. R., Haubruge, E., Pettis, J. S. and Saegerman, C. (2010) Weighing Risk Factors Associated With Bee Colony Collapse Disorder by Classification and Regression Tree Analysis, *Journal of Economic Entomology,* **103**(5): 1517-1523.

119. vanEngelsdorp, D., Underwood, R., Caron, D. and Hayes, J. (2007) An estimate of managed colony losses in the winter of 2006-2007: A report commissioned by the apiary inspectors of America, *American Bee Journal,* **147**(7): 599-603.

120. Varroa destructor, accessed 10th November 2011, <http://en.wikipedia.org/wiki/Varroa_destructor>

121. Virgil, Roman beekeeping, accessed 10th November 2011, < http://www.amazon.com/ROMAN-BEEKEEPING-VIRGIL-Medium-28x35cm/dp/B004X0GANY>

122. Visscher, P. K. and Seeley, T. D. (1982) Foraging strategy of honeybee colonies in a temperate deciduous forest, *Ecology,* **63**(6): 1790-1801.

123. von Frisch, K. (1923) The problem of animal colour sense, *Naturwissenschaften,* **11**: 470-476.

124. Waddington, K. D. (1982) Honey bee foraging profitability and round dance correlates, *Journal of Comparative Physiology,* **148**(3): 297-301.

125. Wang, Y., Jorda, M., Jones, P. L., Maleszka, R., Ling, X., Robertson, H. M., Mizzen, C. A., Peinado, M. A. and Robinson, G. E (2006) Functional CpG methylation system in a social insect, *Science*, **314**(5799): 645-647.

126. Wang, Y., Zheng, W., Luo, J., Zhang, D. and Lu, Z. (2006) In situ bisulfite modification of membrane-immobilized DNA for multiple methylation analysis, *Analytical Biochemistry,* **359**(2): 183-188.

127. What is bee pollen used for? Accessed 10th November 2011, < http://www.ehow.com/about_5137591_bee-pollen-used.html>

128. White, J. W. (1975) Composition of honey, in Crane E. (Ed), *Honey: A Comprehensive Survvey*, Heinemann, London pp. 180-194.

129. White, J.W. (1979) Spectrophotometric method for hydroxymethylfurfural in honey, *Journal - Association of Official Analytical Chemists*, **62**: 509.

130. Williams, G. R., Tarpy, D. R., vanEngelsdorp, D., Chauzat, M.-P., Cox-Foster, D. L., Delaplane, K. S., Neumann, P., Pettis, J. S., Rogers, R. E. L. and Shutler, D. (2010) Colony Collapse Disorder in context, *Bioessays,* **32**(10): 845-846.

131. Wilson, D. S. and Sober, E. (1989) Reviving the superorganism, *Journal of Theoretical Biology,* **136**(3): 337-356.

132. Winfree, R., Aguilar, R., Vazquez, D. P., LeBuhn, G. and Aizen, M. A. (2009) A meta-analysis of bees' responses to anthropogenic disturbance, *Ecology,* **90**(8): 2068-2076.

133. Winston, M. L. (1994) The africanized killer bee - biology and public-health, *Quarterly Journal of Medicine,* **87**(5): 263-267.

134. Worker/nurse bees storing nectar in honeycomb (2008), accessed 10th November 2011, <http://www.flickr.com/photos/borderglider/2709909550/>

135. Wyatt, G. R. and Davey, K. G. (1996) Cellular and molecular actions of juvenile hormone .2. Roles of juvenile hormone in adult insects, *Advances in Insect Physiology,* **26**: 1-155.