The reproductive biology and venom system of the parasitoid wasp *Nasonia vitripennis*

A thesis presented in partial fulfilment of the
requirements for the
degree of

Doctor of Philosophy
in
Genetics

at Massey University, Manawatu
New Zealand.

Ngonidzashe Faya

2019

ABSTRACT

Nasonia vitripennis (*Nasonia*) is a parasitoid wasp that uses a wide range of fly species as hosts to support the growth and development of its offspring. This project focused on two important aspects of *Nasonia* reproductive biology: (1) The role of the endosymbiont *Wolbachia pipientis* (*Wolbachia*) in fertility and (2) the functional characteristics of the venom system. The first part of the project focused on a bacterial endosymbiont *Wolbachia*, which is found in high concentrations within the ovary and sperm cells of *Nasonia*. As *Wolbachia* is maternally inherited, it manipulates host reproduction in ways that favour the production of female offspring. In *Nasonia, Wolbachia* infection results in the generation of reproductive cytoplasmic incompatibilities between infected and uninfected wasps. As very little is known about the molecular mechanism used by *Wolbachia* to interfere with *Nasonia* reproduction. RNA-seq was used to identify 84 and 58 genes differentially expressed in the ovary and testis, respectively, in response to *Wolbachia* infection. RNA interference targeting differentially expressed genes was unable to directly identify functional roles for these genes in cytoplasmic incompatibility, suggesting that this mechanism may be regulated by co-expressed gene networks. Unexpectedly, this study identified a *Wolbachia* gene (*gene1092*) that has been incorporated into the *Nasonia* genome possibly by lateral gene transfer.

The second part of the project looked at the *Nasonia* venom system that has previously been shown to contain a mixture of 79 peptides. The venom is introduced into the host before egg laying and plays a role in altering the host's physiology in ways that favour development of the parasitoid's progeny. However, the individual roles venom peptides play in conferring these dramatic changes in the host have yet to be identified. Therefore, RNA interference was used to target venom genes with the goal of better understanding their function roles. Using this approach, it was shown that *venom X* and *venom Z* were likely involved in arresting development in the host, while *venom Y* was likely to be involved in modulating the host's immune response. As the molecular pathways affected by *Nasonia* venom underlie important cellular pathways, it is predicted that these venom peptides could be developed into drugs to combat diseases such as cancer, hypertension and diabetes.

ACKNOWLEDGEMENTS

First and foremost, I want to thank GOD ALMIGHTY for giving me the grace and strength to pursue this PhD. It wasn't an easy one, each day had its ups and downs therefore to be at this stage of putting all the work on paper, it can only be by grace.

To my supervisors Dave, Helen, Patrick and Gill, I am still in awe of your patience and the support you have given me from day one up to the end. Dr. David Wheeler, thank you so much for giving me the opportunity to come to New Zealand. The support you gave me throughout the application process as well as in the project, taking me into your home when I first moved, helping me to get that balance between work and life, the eggs and nature trips. The list goes on and on…… Thank you for your continual support even from Aussie.

Dr. Helen Fitzsimons, thank you so much for taking over the supervisory role at a time where I wasn't sure of the direction I was going with the PhD. I really appreciate your support throughout, the patience and above all the encouragement and enthusiasm in the progress meetings. I always came out of those meetings believing that its possible.

Associate Professor Patrick Biggs, thank you so much for helping me get back on my feet each time I fell. I really want to appreciate the time you took to teach me some of the bioinformatics software as well as helping me to get that balance of life. You were always available when I need help.

Associate Professor Gill Norris, thank you so much for the support and wise advice you have given me throughout the PhD period. You have been there from the beginning till the end, thank you for believing in me. Protein work was daunting, but you helped me believe I could do it.

To Trevor Loo, thank you so much for helping me get started with the protein work, your support and your wealth of knowledge is highly appreciated. The X-lab team (Fareeda Barzak, Jennifer, Sean, Dell, Brittany), thank you so much for answering all of my questions and being my lab mates too, each time I needed a chat, I knew where to find you. My mates in the Boffin lab (Nikolai, Hannah, Asmat, Elva, Ellie and Alysha), thanks for the chats and encouragements that we shared.

To the MMIC team, Niki and Pani, thank you so much for helping me with sample preparation and teaching how to use the TEM. Raul for helping me understand the TEM methodology and how to analyze the results.

To Trish McLenachan, thank you so much for helping in various ways in the lab, from finding chemicals, to using equipment and ordering lab materials.

To my family for always supporting me and believing in me. Being away from home, you have always checked up on me and encouraged me. Members of Forward in Faith church, thank you so much for the support, the prayers, the packed lunches and dinners.

This project was sponsored by Massey University under the Graduate Assistance Scholarship. Massive thanks go to the School of Fundamental Sciences for supplying all the things required for the project and also thanks a lot to Ann Truter and Cynthia Cresswell for the massive support and great chats.

Table of Contents

List of Figures

List of Tables

ABBREVIATIONS

ATP	Adenosine triphosphate
BCAA	Branched chain amino acid
BLAST	Basic local alignment search tool
Bp	Base pair
℃	Degrees Celsius
cDNA	Complementary deoxyribonucleic acid
CI	Cytoplasmic incompatibility
CN	Cinnabar
COI	Cytochrome oxidase
Da	Dalton
DE	Differential expression
DNA	Deoxyribonucleic acid
DsbC	Disulfide bond isomerase
dsRNA	Double stranded ribonucleic acid
EDTA	Ethylenediaminetetracetic acid
fbpA	Fructose bisphosphate aldolase
FDR	False discovery rate
g	Gram
GH19	Glycoside hydrolase family 19
KD	Knock down
L	Litre
LB	Luria Bertani
LGT	Lateral gene transfer
M	Molar
MGS	Massey Genome Service
mRNA	Messenger ribonucleic acid

NCBI	National Center for Biotechnology Information
Nr	non-redundant
NTC	No template control
NZGL	New Zealand Genomics Limited
OGBF	Otago Genomics and Bioinformtics Facility
Ov	Ovary
PCA	Principal component analysis
PCR	Polymerase chain reaction
pI	Isoelectric point
PO	Phenoloxidase
RIN	RNA integrity number
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	Ribonucleic acid interference
RNA-seq	RNA sequencing
RP49	Ribosomal protein 49
RT	Reverse transcriptase
RT-PCR	Reverse transcriptase polymerase chain reaction
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
sl-CSD	Single locus complementary sex determination
SP97	Serine protease 97
TEM	Transmission electron microscopy
Tes	Testis
TIR	Translation initiation regions
VG	Venom G
VQ	Venom Q
VX	Venom X

VY	Venom Y
VZ	Venom Z
wgs	Whole genome sequencing
Wol-	*Wolbachia* negative
Wol+	*Wolbachia* positive

1. INTRODUCTION & LITERATURE REVIEW

1.1. Overview of *Nasonia vitripennis*

Nasonia vitripennis (*Nasonia*) is a parasitoid wasp that uses a wide range of fly species in the families of Sarcophagidae, Muscidae and Calliphoridae as hosts to support the growth and development of its young [1]. Parasitoid wasps are a group of venomous organisms that are free-living as adults, whereas larvae development is dependent on other organisms (the 'host'). As many of the fly species targeted are agricultural pests, *Nasonia* has potential as a biocontrol system [2]. *Nasonia* has also recently become important as a model genetic system, especially useful in the study of comparative developmental genetics, olfaction, venom, as well as the genetic basis of learning and behaviour [3–5].

Nasonia is a pupal ectoparasitoid in that it targets host flies in the pupal stage of development and the wasp offspring feed from the outside of the envenomated (stung) insect. At the pupal stage the host flies are protected by a hard chitin rich puparium, *Nasonia* females use their ovipositor to drill a hole through this protective layer before envenomating the host and typically laying between 25-40 eggs. *Nasonia* venom does not kill the host immediately, rather the venom functions to manipulate host physiology in ways that favour the growth and development of the offspring (Section 1.3.2.1). Venom injection renders the hosts incapable of completing normal development thus ensuring the constant nutritional supply and safe environment for the feeding *Nasonia* larvae. The *Nasonia* offspring complete development in approximately two weeks when at 25 ℃ (Fig 1.1) [6]. *Nasonia* eggs go through a period of embryogenesis for 24 hours after which the larva emerges and starts to feed using their mandibles to access the host haemolymph. At the end of larval development (5 days) the *Nasonia* pupate and go through a metamorphosis period of 7 to 9 days before emerging as adults wasp when stored at 25 ℃ [6]. Adult males emerge first by chewing an escape hole in the host puparium, where they fight for access to females (often their sisters) as they emerge. Only female *Nasonia* are capable of flight due to the males having vestigial wings, a feature that allows the wasps to be differentiated by sex from the pupal developmental stage (Section 1.1.1).

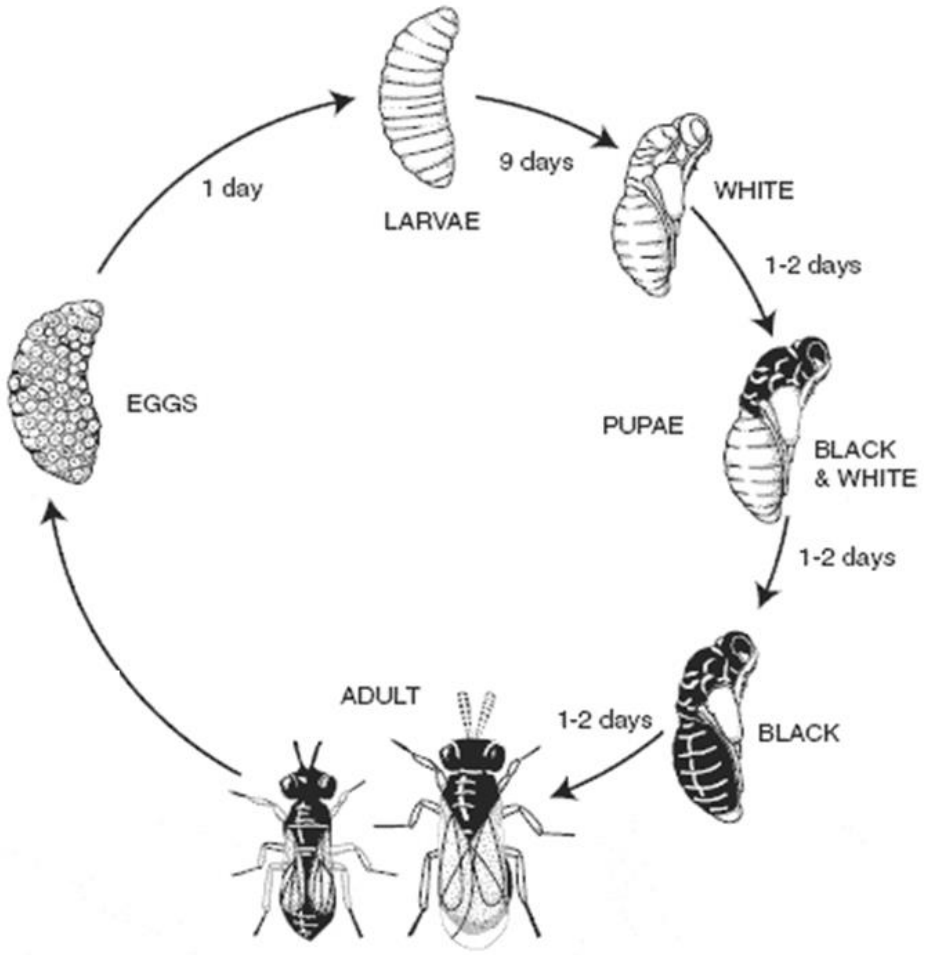


Figure 1.1. Life cycle of *Nasonia vitripennis* wasps. At 25 ℃ the life cycle completes after ~14 days with males being smaller than females and have vestigial wings. Figure modified from Eveline C. Verhurst webpage (*http://www.evelineverhulst.nl/research.html*).

The *Nasonia* genus is made up of four closely related sister species (Fig 1.2): *N. vitripennis, Nasonia giraulti, Nasonia longicornis and Nasonia oneida*. *N. vitripennis* (generally referred to as '*Nasonia'*) targets a wide range of host species from the families of Sarcophagidae, Muscidae and Calliphoridae. In contrast *N. giraulti* and *N. longicornis* are specialist parasitoids that target bird blowflies of the genus *Protocalliphora*. Whilst *N. vitripennis* is thought to be globally distributed [1], *N. longicornis and N. oneida* have only been identified in the north west and north east of the United States, respectively [1,7]. Phylogenetic analysis using nine nuclear markers showed that the three sister

species *N. longicornis*, *N. giraulti* and *N. oneida*, diverged from the *N. vitripennis* lineage less than 1 million years ago (Fig 1.2) [8].

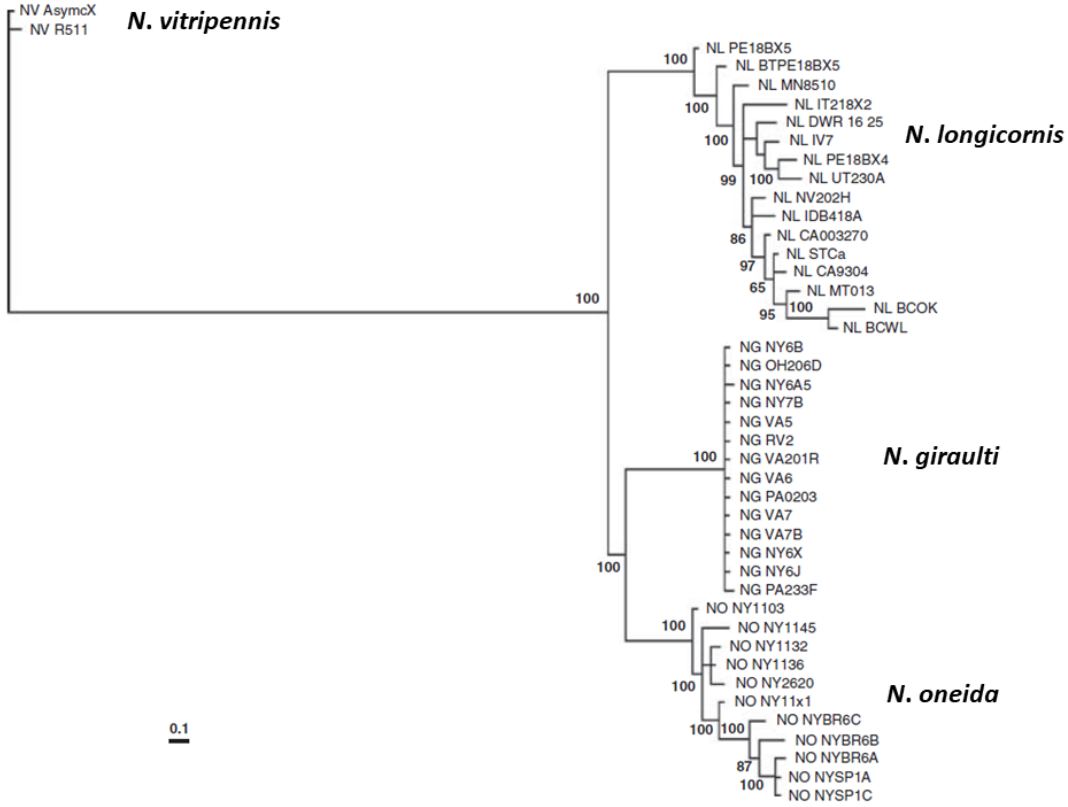


Figure 1.2. Bayesian phylogenetic tree based on nine concatenated nuclear genes showing the phylogenetic relationships of the *Nasonia* sister species (tree taken with permission from reference [8]**).** The names assigned on the tree branches represent multiple strains from each of the *Nasonia* sister species. Permission obtained by licensed content publisher Springer Nature through the copyright clearance center (Springer Nature license number: 4631630447457).

Reproductive incompatibilities mediated by the intracellular bacterium *Wolbachia* is thought to have created cytoplasmic incompatibility and pre-mating isolation driving the divergence within the *Nasonia* lineage [9]. Cytoplasmic incompatibility is the inability of sperm and egg to fuse and is described in depth in the later section (Section 1.2.2.1). Indeed, curing *Nasonia* of *Wolbachia* infection with antibiotics allows the production of viable and fertile hybrid offspring between sister species under laboratory conditions [9]. Pre-mating isolation occurs when females from one species favour mating with males of similar species. In *N. oneida*, females have been found to strongly discriminate against

males of other species [8]. This stereotypic display is a resultant of females that are receptive to the pheromones that are released from the males mandibular region [10]. In comparative biology, discrimination across the *Nasonia* sister species is made possible by differences in male wing sizes and antennae. The antennal structures were found to be different in males; *N. oneida* and *N. giraulti* have an angulate antennal scape whilst *N. vitripennis* have a spindle shaped scape and *N. longocornis* have cylindrical scape. Also the wings can be easily distinguished as *N. vitripennis* has the smallest wings, while *N. longicornis* has triangular forewings and *N. oneida* has broad and rounded forewings [1,8].

1.1.1. *Nasonia* as a model genetic system

The rapid life cycle of *Nasonia*, relative ease of lab maintenance using standard *Drosophila* consumables, and the availability of a high quality genome sequence make this species a powerful model system [11,12]. Although no reliable transgenic system is currently available for *Nasonia*, RNA interference (RNAi) works well and has been successfully used to characterized a number of molecular mechanisms, such as sex determination, embryonic axial patterning, and pheromone production [13–15] (RNAi is described further in Section 1.4).

Another key advantage of *Nasonia* is that genetic lines can be maintained over long periods of time (1-2 years) by inducing diapause, thereby removing the requirement for regular 'turning over' of lines to maintain their viability in the laboratory [16]. The production of diapause larvae can be initiated by storing female wasps/hosts at 15 degrees under a 8/16 hour light/dark cycle [17]. *Nasonia* lines can also be maintained over short periods of time where at 4 ℃ incubation, further development or maturation of wasps can be paused and started again when they are returned to a higher temperature within a certain timeframe. These stages include the embryo that can be stored for up to 48 hours, the early pupal stage that can be stored for up to 3 months, the late pupae that go for one month in cold storage and the adult wasp that can be stores at 4 ℃ for approximately three weeks [11].

Due to their propensity for inbreeding, genetic crosses can easily be made and this has been used to map quantitative trait loci affecting wing size and mate discrimination [4,5].

These genetics studies have shown that only a few genes linked to the OR123 locus on linkage group IV [4] are responsible for the different wing morphology between *N. vitripennis* and *N. giraulti*. Genetic crosses between *N. vitripennis* and *N. giraulti* are possible when antibiotics are used to remove *Wolbachia* [3–5]. Also, discrimination of males by females has a relatively genetic bias which consists of quantitative trait loci and their interactions controlling female mate discrimination [3]. Mate discrimination can also be a resultant of behavioural differences where males stimulate females by positioning their body on top of the immobile females with the forefeet displaying ordered stereotype movements on her head that differ in other species [18].

Genetic studies are aided by the fact that *Nasonia* males and females can be easily differentiated at the pupal stage based on differing wing morphology and the presence of an ovipositor on the abdomen of females. This ability to distinguish sexes when the *Nasonia* are immobile pupae makes it simple to collect virgins for genetic crosses before the insects eclose [1] (Fig 1.3). *Nasonia* like all Hymenoptera has a haplodiploidy sex determination (Section 1.2.1), in which females are generated from fertilised eggs, whilst males are derived from unfertilised eggs. Taking advantage of this system large numbers of males can be collected by exposing hosts to virgin females [12].

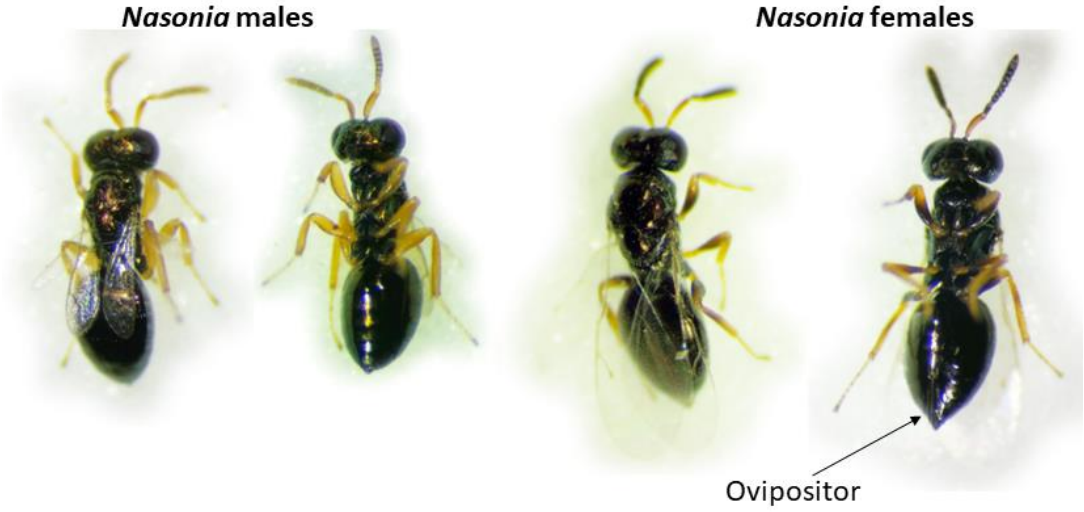


Figure 1.3. The differences between *Nasonia* males and females. Males are smaller and have shorter wings while females have long wings and an ovipositor that can be readably identified on the ventral side of the abdomen.

1.2. The *Nasonia* reproductive system

Nasonia testis are attached to four seminal vesicles where spermatozoa is stored [19,20]. Spermatogenesis occurs during a short period 1 day before eclosion [19]. Females have two ovaries attached to a number of ovarioles filled with eggs [20]. In the ovaries, females are able to store sperm in the spermathecae and use it to produce daughters throughout their entire lifetime [21]. Females can selectively fertilise their eggs to control the number of male offspring depending on environmental conditions. For example, if a female *Nasonia* comes in contact with a host that has already been parasitized by another female, she will lay an excess of unfertilised eggs to bias the production of males and thus increase her chance of passing on her genetic material [12].

1.2.1. Haplodiploidy sex determination

Haplodiploidy is a mode of reproduction that occurs in a wide group of invertebrates mainly the insect orders Hymenoptera and Thysanoptera where males are haploid and females are diploid (Fig 1.4) [22]. Diploid females produce haploid eggs and haploid males produce haploid sperm, so that if the eggs are unfertilized, haploid males are produced. Diploid females are only produced if the haploid sperm fertilizes the haploid eggs [23]. All members of the insect order Hymenoptera have haplodiploid sex determination and thus it is the number, rather than a specific type, of chromosomes that determine the sex of progeny. Within the Hymenoptera group of insects, one mechanism of sex determination is called the single locus complementary sex determination (sl-CSD), which is based on the allelic state of a single locus [24]. In *Nasonia* haplodiploidy sex determination does not appear to be controlled by sl-CSD, with research suggesting that it occurs through the regulation of a network of sex determination genes [25,26]. As part of the *Nasonia* genome project, two genes were identified (*doublesex* and *transformer*) that have been described in *Drosophila sp*. to play an important role in sex determination [27,28]. The *Doublesex* gene controls sexual differentiation into either male or female *Nasonia* by acting as a functional switch gene and is specifically spliced by the *transformer* gene. A signal that is specific for the either male or female sex in turn regulates the splicing pattern of the *transformer* gene. RNA interference (RNAi) of *Nasonia* female pupae's *transformer* gene caused sex reversal of fertilized eggs resulting in diploid male offspring instead of females [15].

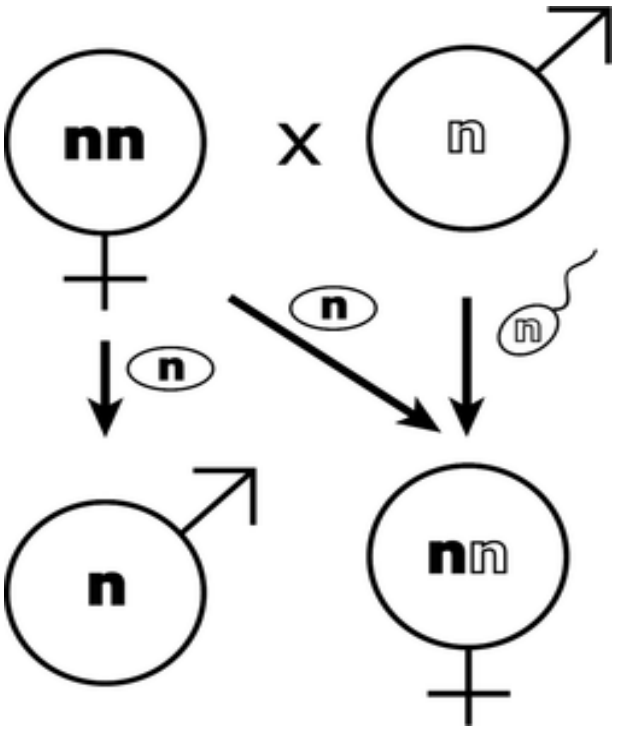


Figure 1.4. Haplodiploid sex determination in *Nasonia*. Males are haploid, and females are diploid.

1.2.2. *Wolbachia*

Like many arthropods, *Nasonia* and its sister species are infected by the intracellular gram-negative bacteria *Wolbachia* (order Rickettsiales). *Wolbachia* is maternally inherited and is found in high concentrations in the *Nasonia* reproductive tissues [29,30]. Multi-locus sequence typing showed *Wolbachia* can be divided into 8 subgroups (A-H) that together infect up to 66% of all arthropod species [31]. Comparisons between the *Wolbachia* phylogeny and that of their hosts suggest that *Wolbachia* is able to move between distantly related species through horizontal transmission. Experiments involving microinjection have been able to induce artificial infections of *Wolbachia* between distantly related species [32,33]. This ability of *Wolbachia* to infect species of distant relationships has been used to great advantage in the production of dengue and zika virus resistant mosquitoes [34,35]. Furthermore, phylogenetic analysis of the *Wolbachia ftsZ* gene between infected hosts and parasitoids revealed the occurrence of horizontal transmission involving both *Nasonia* and its blowfly host as both were infected by the same strain of *Wolbachia* [32].

Naturally occurring *N. vitripennis* populations (LABII strain which will be referred to as Wol+ in this thesis) are coinfected by *Wolbachia* from the A and B super groups, which differ by only 2% in their 16S rDNA sequences [36]. Coinfected hosts present an opportunity for genetic recombination between the different strains of *Wolbachia*. Phylogenetic and PCR-based classification of the *Wolbachia* surface protein gene *wsp* provided sequence evidence for genetic recombination between strains A and B [37,38]. Apart from transfer of genes among the *Wolbachia* strains, genomic DNA of a number of eukaryotic hosts known to carry the *Wolbachia* B-subdivision strain were sequenced and unique sequences from each of the *Wolbachia* strains were obtained revealing possible genetic exchanges between the bacteria and respective hosts [37].

1.2.2.1. Effects of *Wolbachia* on the host.

Reproductive system manipulation

As *Wolbachia* is maternally transmitted, it uses a number of mechanisms to manipulate the host reproductive system in ways that favour the generation of female offspring. These reproductive manipulations can lead to a rapid increase in the frequency of *Wolbachia* infections in the population [39]. The most common forms of reproductive manipulations include cytoplasmic incompatibility, male killing, feminization and parthenogenesis [9,40–42]. Cytoplasmic incompatibility (CI) is the most common reproductive manipulation by *Wolbachia*, where *Wolbachia* infected sperm is incompatible with eggs from females that are either uninfected or infected with a different strain of *Wolbachia* (Fig 1.5) [43]. During the fertilisation process, although sperm and egg fuse normally, the *Nasonia* chromosomes from both male and female are disrupted before first mitotic stage resulting in failure of fusion of the two nuclei [44]. CI occurs in two major forms called unidirectional and bidirectional. In unidirectional incompatibility, the sperm from *Wolbachia* infected male is incompatible with egg from uninfected females, while in the reciprocal cross sperm from uninfected males produces viable offspring after fusion with an egg from an infected female. Bidirectional incompatibility occurs when males and females are both infected by different strains of *Wolbachia* [45]. In this case, sperm from males infected by one strain is incompatible with eggs from females infected by another *Wolbachia* strain (Fig 1.5b). The mechanism that underlies cytoplasmic incompatibility

is still unclear; however, it is believed to involve a two-component system with bacterial mediated modification of the sperm, followed by rescue of the fertilized egg. Based on this hypothesis, successful fertilisation can only occur when *Wolbachia* mediated modifications to the sperm are rescued by the same strain of bacteria found in the egg. If rescue does not occur, the sperm and egg will be incompatible [46]. *Nasonia* displays a classic CI phenotype between sister species as they are mainly infected by different strains of *Wolbachia* [9]. However, as described previously, feeding infected wasps with antibiotics results in successful fertilization of egg [4,36].

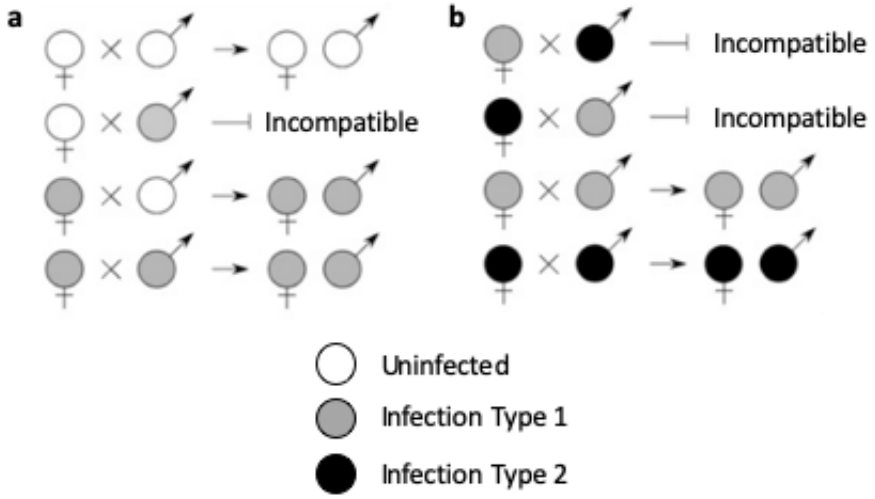


Figure 1.5. Examples of cytoplasmic incompatibility in *Nasonia*. a) CI occurring between *Wolbachia* infected and uninfected individuals. *Wolbachia* infected females are compatible with both infected and uninfected males while their counterparts are incompatible with infected males. b) CI occurring between individuals infected with different strain of *Wolbachia*. CI occurs when male and female have different infection types. Figure modified with permission from reference [47].

The other types of reproductive manipulation, feminization, male killing and parthenogenecity, are less commonly observed in the arthropods and are absent in *Nasonia*. *Wolbachia* induced feminization results in males developing as females and is common in butterflies (Fig 1.6A) [48]. In the isopod, *Armadillidium vulgare*, the individuals female offspring produced are functional [49]. The mechanism in which

feminization occurs remain unclear however experimental evidence showed that this manipulation is achieved by preventing the differentiation of the androgenic gland during post-embryonic sexual differentiation [50]. The androgenic gland produces a hormone that causes male sex differentiation and the development of this gland is dependent on the gene that is on the W chromosome in the isopod [51]. The infection of *Wolbachia* to *A. vulgare* results in the decline and eventual elimination of the W sex chromosome in the infected population leading to sex determination in the isopod to be controlled by *Wolbachia* [51]. Since *Wolbachia* infects a wide range of insects, the mechanism through which feminization occurs is widely different and in others insects such as *Drosophila*, it is suggested that key genes (*doublesex* and *transformer*) that control sex determination may be regulated by *Wolbachia* for the manipulation to occur [52].

Pathogenic induction results in unfertilised eggs developing as females instead of males [53,54], thus further increasing the *Wolbachia* transmission rate (Fig 1.6B). Evidence suggest that this manipulation is restricted to species that have an arrhenotokous type of development (haplodiploidy sex determination) such as hymenopterans, mites and thrips [53–55]. Instead of the normal production of males from unfertilized eggs, *Wolbachia* manipulates the host to produce females from unfertilized eggs and the female offspring is able to transmit the bacteria to their offspring. The most accepted mechanism underlying parthenogenesis is that the manipulation is caused by the disruption of the cell cycle during embryogenesis resulting in unfertilized eggs producing diploid females [56]. For example, in the wasps *Trichogramma pretiosum* and *Leptopilina clavipes*, experimental evidence suggested that during the first embryonic division, anaphase is aborted resulting in one diploid nucleus instead of the normal two haploid nuclei [57,58].

Male killing by *Wolbachia* results in the death of infected male offspring likely in order to free up resources for the surviving infected female embryos (Fig 1.6C) [59,60]. This manipulation has been reported in different arthropod orders such as Coleoptera, Diptera, Lepidoptera and Pseudoscorpiones [59–62] where male offspring are killed early in development. The mechanism in which male killing is achieved remains unclear and since this manipulation is observed in a wide range of insects, the interaction between *Wolbachia* and the host in these different groups is likely to be different too. The male

killing phenotype has been widely studied in the *Drosophila willistoni* where evidence suggested that male death can occur either before or after gastrulation [63]. Before gastrulation, embryonic deaths in males is a resultant of abnormalities in the mitotic processes, in particular the achromatic spindles which show abnormal cleavage patterns [63]. After gastrulation, embryonic death is associated with the breakdown of internal structures and pycnosis of the nuclei, which is supported by the blackening of the embryo instead of the normal brown coloration [63]. Further research is clearly warranted to describe the mechanism in which males are selectively killed.

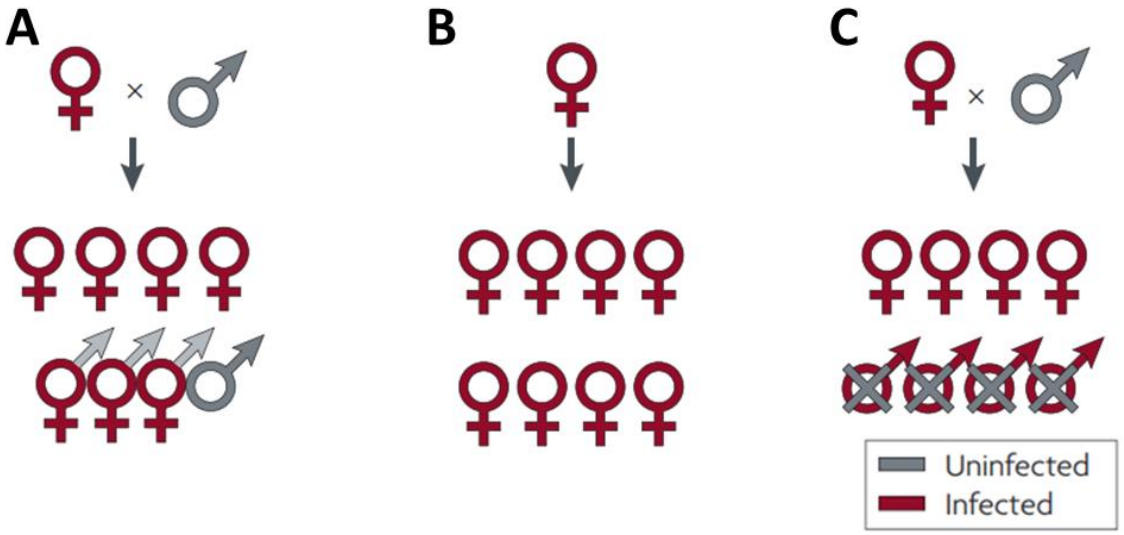


Figure 1.6. *Wolbachia*-induced phenotypes in arthropods. A: Feminization manipulation, which results in genetic males developing as females. B: Parthenogenesis induction, which is development of females from unfertilised eggs. C: Male killing, which is the killing of all infected males favouring the development of infected females. Figure modified with permission from reference [56] (Springer Nature License number: 4631640668751).

Host fitness

Evidence from both theoretical and experimental approaches suggests that despite the reproductive manipulations, *Wolbachia* increases the fitness or overall wellbeing of the host [64–66]. For instance, when the bacterium is introduced in mosquitoes, they are protected by *Wolbachia* from infection by other pathogens such as dengue [67]. The exact mechanism in which the protection happens remains unknown, however immune genes were observed to be upregulated suggesting that the bacteria primes the innate immune

system of the mosquitoes fighting off the infection [67]. In addition, pathogen interference protection is suggested to arise from limiting the cellular resources that are required for pathogen replication [67]. Cholesterol and fatty acids are examples of key resources that are required by insects for development and since they do not produce cholesterol, they obtain it from their diet [68]. This is important for *Wolbachia* and other related bacteria as they lack the ability to biosynthetically synthesize cholesterol therefore they obtain it from the host [69] creating competition for these resources with other bacteria thereby reducing the populations of bacteria in the host.

Lateral gene transfer

Lateral gene transfer (LGT) also referred as horizontal gene transfers, is defined as the nonsexual transfer of genetic material from the genome of one species to another species that is not related. LGTs spread widely between bacteria and archaea, and as such have likely acted as a powerful source for functional innovation and diversity with regard to cellular structures, metabolic properties and lifestyles [70–72]. The diversity among the narrow taxonomic group of enteric bacteria for instance is remarkable, as each species (*Escherichia coli*, *Shigella flexneri*, *Salmonella Enterica*, *Klebsiella pneumoniae*, *Serratia Marrcescens*, *Yersinia pestis*) possess a unique phenotypic profile despite the fact that they all arose from a common ancestor (Table 1.1). LGTs can play an integral role in the evolution of organisms, with a good example provided by the pathogenic bacteria *E. coli* and *S. enterica*. These two species have different lifestyles, where for instance *E.coli* utilise lactose and lysine decarboxylase, *S. enterica* does not (Table 1.1) [70]. It is known that the majority of the archaea species are thermophiles and this lifestyle has been acquired by bacterial thermophiles through LGT events from archaea to bacteria. Comparative analysis of archaeal and *Aquifex aeolicus* genomes showed that at least 20% of the genes were conserved between archaea and bacteria [73] supporting the suggestion that the ability to withstand high temperatures is a resultant of LGTs providing a selective advantage to the recipient organism, even though in most cases the advantages might not be clear. Other bacterial genomes (*E. coli*, *B. subtillus*, *B. burgdorferi* and *Synechocystis sp*.) that lack the ability to resist high temperatures were compared to the archaeal

genomes and had less than 5% reliable hits to archaeal genes suggesting that the heat resistance genes in *Aquifex aeolicus* resulted from gene transfers [73].

Table 1.1. Phenotypic profiles of representative enteric bacteria [70]**.** Genes acquired through LGTs contribute to bacterial metabolism.

Organism	**Lifestyle**	**Lactose utilisation**	**Citrate utilisation**	**Produce H_2S**	**Produce indole**	**Produce urease**	**Lysine decarboxylase**
E. coli	Mammalian commensal	Yes	No	No	Yes	No	Yes
S. flexneri	Primate pathogen	No	No	No	Yes	No	No
S. enterica	Mammalian pathogen	No	Yes	Yes	No	No	Yes
K. pneumoniae	Soil	Yes	Yes	No	No	Yes	Yes
S. marrcescens	Soil	No	Yes	No	No	Yes	Yes
Y. pestis	Mammalian pathogen	No	No	No	No	No	No

The recent completion of a wide range of plant and animal genome sequences has identified unexpected frequencies of LGTs in eukaryotes [74,75]. Although less common than in prokaryotes, LGTs between eukaryotes have been described. One such example is the transfer of P transposable elements from *D. willistoni* to *D. melanogaster*. Another example is the transfer of lectin-antifreeze genes between fish and the acquisition of carotenoid biosynthesis genes by pea aphids from fungi [76–78]. These LGTs show that the acquisition of foreign genes can be an important evolutionary mechanism that allows recipient organisms to evolve novel mechanisms that enhance fitness. Most of the described LGTs occur within a single domain of life e.g. bacteria to bacteria or eukaryote to eukaryote, however evidence of interdomain LGTs has been presented. The most studied example of interdomain LGT is between Eubacteria *Agrobacterium tumefaciens* and plants where 10-30 kb of the bacteria T-DNA from its Ti plasmid can be intentionally transferred to the plant genome [79]. The soil bacterium *A. tumefaciens* is a pathogen that causes crown gall tumours in plants and the oncogenicity in the bacterium is caused by the Ti plasmids [80]. Therefore, introduction of the Ti plasmid ensures success of tumorigenicity to the infected plants.

Endosymbiont to host LGTs are reasonably common, likely due to the close association of cells from both organisms. Given that *Wolbachia* is an intracellular parasite of many insects, there is an increased opportunity for LGT events between this endosymbiont and its host [81,82]. In the bean beetle *Callosobruchus chinensis*, at least 30% of *Wolbachia* genome was transferred to the X chromosome [83,84]. A low level transcription was observed on approximately 50% of the transferred genes so it is possible these LGTs are functional [85]. Interestingly almost the entire *Wolbachia* genome has been incorporated into *D. ananassae* 2L chromosome [75]. Furthermore, four different lines of *D. ananassae* from Asia and the Pacific were observed to have incorporated the bacterial genes revealing that the LGT is widely distributed and has been maintained over evolutionary time, suggesting it provides some kind of fitness increase to its host [75].

LGTs in *Nasonia* have been experimentally verified through PCR amplification and sequencing with the parasitoid's genome containing 13 ankyrin repeat genes previously observed only in Pox viruses [86]. Pox viruses do not infect *Nasonia* and the ankyrin genes are widely distributed in diverse *Wolbachia* strains suggesting that *Nasonia* acquired them from the bacteria through LGT events [86]. Furthermore, fragments of *Wolbachia* sequences (<500 bp) were experimentally verified through PCR coupled with sequencing, and confirmed that they are integrated in the *Nasonia* genome [75]. In this scenario, rather than the complete gene sequence, only traces of *Wolbachia* sequences or gene fragments were transferred into the *Nasonia* genome. The incorporation of *Wolbachia* DNA fragments into existing *Nasonia* genes suggests the possibility that they could potentially introduce new function domains. The function of these newly acquired ankyrin genes has not yet been determined. Another example of an LGT involving *Nasonia* is the acquisition of a glycoside hydrolase family 19 (GH19) chitinase from microsporidia. This GH19 is expressed and protein from this gene has been identified in the venom of *Nasonia*, strongly suggesting that it is functional [87]. The current hypothesis for the role of GH19 chitinase in *Nasonia* venom is that it modifies the hosts immune system as RNAi knock down of this gene results in an increase in expression of immune response genes in the fly host [87].

1.3. The venom system of *Nasonia*

Venoms are toxic substances that mostly function to cause harm or kill other organisms. These toxins range from small organic molecules to large proteins and their effects on the victim vary from mild discomfort to death. Organisms that release venom include but are not limited to snakes, scorpions, spiders, shellfish, insects such as bees, wasps and ants. Snakes represent the most studied group of venomous animals due to the dangers they pose to humans as well as the relative ease of venom harvesting via milking [88]. Depending on the snake species involved, the main effects of snake venom on vertebrates are localised swelling, blistering, necrosis, anti-haemostatic defects, paralysis and shock [88,89]. These effects are caused by various biologically active proteins and peptides including but not limited to proteases, phospholipases and esterases [89]. Antivenom is the principal therapy for venom neutralizing and is made up of antibodies from an animal immunised with the appropriate venom [90]. The animal is successively inoculated with increasing doses of venom thereby producing neutralizing antibodies. The animal is then bled to extract the antivenom rich hyperimmune serum. The most common venomous arthropods that are a danger to humans are scorpions and spiders. A sting from these species generally cause pain, cramping, and at higher venom dosage can lead to death [91]. These effects are caused by a combination of proteins, polypeptides, inorganic and organic ions that make up the venom component [92].

1.3.1. Hymenoptera venom

Despite being one of the largest group of venomous animals, apart from bees, members of the Hymenoptera (bees, ants, wasps) represent one of the least well-understood venom systems. This is because stings from these species are rarely lethal to humans and cause discomfort rather than specific harm. The amounts of venom produced by these insects is also very small, this complicates our ability to study these molecules. In the majority of cases where humans have died from insect stings, these are actually caused by a secondary hypersensitivity reaction to components in the venom, rather as a result of a specific venom function [93]. In addition, death can result from massive envenomation from large swarm attacks that together release a lethal dose. Africanized bees are the most abundant types of bees spreading across the whole globe and are termed killer bees because of their aggressiveness attacks [94]. The growth rate of African bees is higher than other colonies

perhaps explaining their increased swarm production as well as larger distribution [95]. They increase on an average of 16-fold per year whereas the European colonies can only reach a maximum of 6-fold growth rate per year [96,97].

The European honey bee venom is the best characterized Hymenoptera venom, primarily consisting of peptides, enzymes and amines [98]. The venom dose in Hymenoptera is primarily made up of proteins or peptide elements [99] and the main component in bee venom is melittin which has been found to hydrolyse cell membranes thereby causing local pain [99]. Furthermore, melittin induces the release of catecholamine, an important factor in the reaction that causes intravascular haemolysis [100]. Phospholipase A2, which is also present in the venom has been found to be the major allergenic component and helps catecholamine to cause intravascular haemolysis. Another peptide in bee venom is the neurotoxin apamin which has excitatory neurotoxic effects on the central nervous system leading to hypermotility in mice for example when lethal doses are injected [98]. Some of the bee venom peptides have not had their functions experimentally characterized. Instead, functions have been predicted based on sequence homology with other venom proteins [101]. The ant *Myrmecia* (bull ants) also release venom which in comparison with other insect venoms, they are allergenic [102]. The ant venom component is a mixture of peptides, enzymes, hydrocarbons, alkaloids and formic acid [103].

In contrast to the more generalised functions of most animal venoms, parasitoid venoms produce highly specific effects that create a favourable environment for the development of the insect's offspring on the envenomated host. These venom effects often include paralysis/developmental arrest and/or the interference with the hosts' immune system [99]. Heterogeneity of parasitoid insects in terms of phylogeny, host ranges and parasitic habits is reflected by the functional diversity of these insects' venom. Research into these venoms has highlighted general and unique biochemical properties of parasitoid venom proteins. It had been suggested that unlike the venom of bees and other social wasps, parasitoid venom was only made up of higher molecular masses proteins [104]. However, recent proteomic studies are showing that the complex mixtures of venom proteins and peptides have a range of molecular masses (less than 10 kDa to greater than 100kDa)

depending on the species [105–108]. Host-parasitoid relationships are specific due to the complexity and variations in venom components among parasitoids that allows the successful parasitism as well as the parasite to be specific to its target [108,109].

Protein properties such as charge, and size are important for protein separation and characterization, however less information has been reported regarding the charge (acidic, basic or neutral) of venom proteins in parasitoid organisms. The venoms of bees have been reported to possess venom proteins that are either neutral or basic in nature [104]. In contrast, in the few studies carried out on parasitoid venoms it has been shown that they are mostly acidic. For example, the venom of the parasitoid wasp *A. tabida* has venom proteins with acidic isoelectric points ranking from 5.1 to 6.5 [104]. In addition, isolated individual venom proteins from *C. curvimaculatus* and *C. rubecula* were reported to have isoelectric points of 4.9 and 3.99 respectively [110,111]. A few parasitoid venom proteins have been found to be heavily glycosylated and to contain potential glycosylation points even though the exact nature of the polysaccharidic residues remains unknown [106,108,112]. However, in the parasitoids *Aphidius ervi* and *N. vitripennis*, the isolated proteins were observed to lack glycosylation [113,114], suggesting that the post-translational modification is not an obligatory feature for venom proteins.

Signal peptides are important features for secretory proteins including venom sequences and they have been reported to comprise three typical domains: the N-terminal positively charged region, hydrophobic middle domain and the polar C-terminal domain [115]. Interestingly, the signal peptides of parasitoid venoms were observed to have contrasting features with those of other known signal peptides. For example, in *Chelonus sp*. the signal peptide was observed to be negatively charged on both ends and the other venom signal peptide sequences lack basic residues at the C-terminal region [111,116]. Based on these observations, it seems parasitoids may have evolved specialised protein processing and secretory pathways depending on the target host [116]. In *Chelonus sp*. this evolutionary pattern allows its venom proteins to be secreted after biosynthesis through a different mechanism to that used by other venoms that have classical signal peptides with basic residues [116]. There are at least five classes of signal peptides that have been predicted in parasitoid venoms that have contrasting features to the general signal peptide

architecture [104]. The first class which is the regular architecture of signal peptides structures where the N-terminal is positively charged followed by a hydrophobic middle region and the polar C-terminal end [111]. The second class was observed to follow the definition of classical signal peptides however negatively charged amino acid residues that do not interfere with the polar region were observed at the C-terminal [116]. The third class is rare where the negatively charged residues are located at the positively charged N-terminal end [117,118]. Classes 4 signal peptides are composed of neutral residues at both ends while class 5 signal peptides have basic residues at the C-terminal end [119,120].

1.3.2. *Nasonia* venom

Nasonia venom comprises 79 peptides that have been identified using proteomic methods [121,122]. The venom component consists of a combination of high and low molecular mass peptides [121]. Envenomation of *Sarcophaga bullata* hosts by *Nasonia* results in immune suppression, developmental arrest, altered metabolism and eventually death [123]. However, the specific venom peptides that control these processes remain unknown. The diverse and potent effects of *Nasonia* venom highlights the potential of these molecules to be developed as novel therapeutic drugs. Indeed, recently *Nasonia* venom was found to have anti-inflammatory activity on mammalian cells [124].

1.3.2.1.Effects of venom on host

Immune system

The case study using *Manduca sexta* as a host showed that the humoral and cell-mediated immune responses of the host are affected by *Nasonia* venom. The humoral response in insects involves the activation of a phenol oxidase cascade and antibacterial peptide synthesis, while phagocytosis and encapsulation are involved in a cell-mediated response [125,126]. For the humoral response, *Nasonia* venom has been shown to influence the phenoloxidase (PO) cascade that controls the melanisation of pathogens. PO converts phenols to quinones that in turn polymerise to form melanin, part of the first steps in the innate immunity pathway [127]. In the cell mediated response, *Nasonia* venom inhibits the coagulation process of the host based on the fact that when *Nasonia* feeds, there is no evidence of melanisation or clotting observed at the host's wound site until several hours

later [128]. Preventing coagulation of host haemolymph is important as it makes it easier for parasitoid young to feed. Moreover, *Nasonia* venom has been found to downregulate immune response genes in *Muca domestica* haemocytes via RNA-seq studies [129].

Developmental arrest

Developmental arrest in envenomated pupae is most easily identified by the failure to produce eye pigment and to form bristles [123]. The purpose for arresting development remains unknown; however, the parasitoid larva feeds on the host haemolymph therefore the reason could be to stop hardening of the host's ectoderm. Another suggestion is that halting normal development in the host frees up resources that can now be used by the feeding *Nasonia* larvae [130].

Apoptosis

Nasonia envenomation also triggers apoptotic pathways through an unknown mechanism. Apoptosis is more likely to be collateral damage from other specific cellular manipulations such as formation of membrane blebs after envenomation, which disrupt the cytoskeletal cell membrane. Calcium ion intake is regulated by the cytoskeletal membrane therefore unregulated calcium release from the cytosol leads to apoptosis [2]. Additionally, phenoloxidase cascade is critical in the toxin-antitoxin pathway that leads to cell death therefore interference with phenoloxidase cascade after envenomation also causes cell death [131].

Host metabolism

Envenomated hosts also display an altered metabolism as revealed by metabolomics experiments in *S. bullata* [123]. Trehalose is an important sugar reserve in the haemolymph of insects and glycogen levels were found to significantly drop after envenomation. This would indicate hydrolysis of trehalose to glucose which will in turn be available to developing wasp larvae as an energy source [132]. Envenomation is also associated with accumulation of lipid in the fat body of *S. bullata* [123]. Microinjection of *Nasonia* venom into *S. bullata* resulted in an influx of sodium ions and caused a

subsequent release of calcium ions in the mitochondria. This in turn stimulated fatty acid synthesis in the host, repressing the likely source of the lipid found in the fat bodies [123].

1.3.2.2.*Nasonia* venom peptides

Most of the 79 *Nasonia* venom peptides fall into nine bioinformatically assigned functional groupings; proteases, peptidases, protease inhibitors, carbohydrate processing enzymes, DNA processing enzymes, enzymes involved in glutathione metabolism, esterases, recognition/binding proteins, immune related proteins. Several venom proteins currently have no characterized functions and these are indicated by single letter names: venoms D-Z [121]. Serine proteases are the most common enzymes found in *Nasonia* venom with 16 different serine protease venom proteins [121]. These all contain a catalytic triad in which serine is an essential residue. Serine proteases are one of the most common types of genes found in insect genomes and they function in many processes, thus it is difficult to predict their venom specific function based on sequence homology alone. A metalloprotease found in *Nasonia* venom has sequence similarity with a metalloprotease found in *Eulophus pennicornis* venom and injection of this protein into its host resulted in growth retardation [130]. Metalloproteases control development in insects such as *Drosophila melanogaster* by activating the Notch signalling pathway during embryo growth [133]. Another *Nasonia* venom peptidase has homology to angiotensin-converting enzymes which are predicted to be involved in peptide processing and western blot analysis of a *P. hypochondriaca* venom targeting angiotensin-converting enzyme cross reacted with another venom protein Therefore the cross reaction suggest that it could be involved in processing peptide precursors in the venom sac [105].

Both serine and cysteine rich protease inhibitors are found in *Nasonia* venom. Two of the serine protease inhibitors belong to the large family of Kazal-type serine protease inhibitors that are characterized by 6 cysteine residues in a conserved motif of 40-60 amino acids [134]. The type represents a small serine protease inhibitor that shows resemblance to the *Schistocerca gregaria* chymotrypsin inhibitor [121,135]. The *S. gregaria* small serine protease inhibitor has been found to inhibit insect proteases and it may play a similar role in *Nasonia* venom [135]. The cysteine-rich protease inhibitors found in *Nasonia* venom are from the family classes kunitz, pacifastin and trypsin

inhibitor like. These classes are distinguished by the number of disulphide bridges that are formed in the functional domains [121]. Generally, cysteine-rich proteins are single chain polypeptides that block smooth muscle contractions and cyclic nucleotide-gated ion channels [136].

Three venom peptides are predicted to be involved in carbohydrate metabolism, these include a chitinase, trehalase and glucose dehydrogenase. Chitinases are enzymes that degrade the insoluble chitin found in the exoskeleton and gut linings of most insects [137]. Glucose dehydrogenase belongs to the family of oxidoreductases and is known to catalyse the first step of the pentose phosphate pathway [138]. Trehalose, also known as mycose, is a ubiquitous carbohydrate storage molecule that is synthesized as a stress responsive factor [139]. Moreover, trehalose is thought to prevent protein denaturation and stabilize aggregation-prone model proteins. Therefore, the presence of trehalase in the venom mixture may protect the venom components from attack by the host immune response. Alternatively, venom trehalase could convert host trehalose into glucose for use by the feeding *Nasonia* larvae.

A group of venom peptides have been suggested to play a role in DNA metabolism and these are endonucleases, apyrases and inosine uridine-preferring nucleosides. Their specific functions in the host are unknown, but in general, endonucleases are enzymes that cleave the phosphodiester bonds of any foreign DNA [140] while apyrases catalyse the hydrolysis of adenosine triphosphate to adenosine monophosphate and phosphate in the presence of calcium. The increase in extracellular adenosine monophosphate inhibits platelet aggregation suggesting that apyrases are likely to be involved in weakening the host's immunity system. Inosine uridine-preferring nucleosides ensures hydrolysis of purine and pyrimidine nucleosides into ribose and its associated base. [141].

The enzymes (γ-glutamyl transpeptidase and γ-glutamyl cyclotransferase) are thought to be involved in the inhibition of glutathione metabolism and have been found in *Nasonia* venom. The prediction was based on sequence homology with the γ-glutamyl transpeptidase of *Aphidus ervi* which was found to induce apoptosis in the germarial and ovariole sheath of the host aphids [142].

Esterases which are a class of hydrolases are also found in *Nasonia* venom. They include acid phosphatases, arylsulphatase, carboxyl-esterases and lipases. Lipases have been experimentally shown to alter the lipid metabolism of the host [143] while the effects of the other esterases on the host remain unknown. Arylsulphatase could influence the development of the host based on the observation that lysosomes in the southern armyworm produce homologous arysulphatases that have been shown to regulate insect moulting hormones [144].

Finally, *Nasonia* venom has been shown to inhibit the host cellular immune response [128] and two immune related venom proteins that may play a role are calreticulin and C1q-like venom protein. Calreticulin was reported to affect host's innate immune response and reduces host bleeding during feeding [145]. The c1q-like venom protein has a C1 domain, which has been shown to regulate the activation of serum complement system in higher organisms which a series of dissolved proteins that affords animals protection against a variety of pathogens [146].

1.3.3. Characterization of venom proteins

Advancements in genomics, proteomics, and metabolomics technologies have made the study of wasps' venom possible. For instance, liquid chromatography coupled with mass spectrometry (LCMS), a high throughput proteomic method that can obtain protein sequences from the very small samples of venom (5-10 ng), was used to identify and list the venom proteins in *Nasonia* [121]. The elucidation of the proteins that make up the venom mixture has opened possibilities for research to target individual proteins for characterization. Three primary methods of venom functional analysis have been successfully applied to *Nasonia* venom: metabolomics, RNAi and molecular cloning and these will be described further below.

Using metabolomics, the overall effects of *Nasonia* venom on the host *S. bullata* was studied through comparison of the metabolic changes that occurred before and after venom injection [147]. Data obtained from this metabolomics study coupled with pathway analysis revealed that the biochemical changes facilitated by the introduction of venom into the host included chitin synthesis along with the metabolism of sugars and amino acids [147]. The second approach used in characterizing *Nasonia* venom is RNAi and this

has allowed new insights into venom gene function [145]. RNAi knock down of a specific venom encoding genes does not disrupt other venom components and thus allows the effects to be studies in isolation. The third approach used to study the *Nasonia* venom was generation of recombinant venom proteins where two recombinant Kazal-type serine proteases tested on the host, *Musca domestica*, showed that the proteins have an influence on the phenoloxidase and prophenoloxidase activation of haemolymph [148,149]. The two methods within the scope of this research are RNAi and characterization of recombinant proteins will be discussed further in sections 1.3.3.1 and 1.3.3.2 respectively.

1.3.3.1.RNA interference (RNAi)

First tested in the nematode worm *Caenorhabditis elegans*, RNAi is a sequence-specific gene silencing method that is initiated by the introduction of double-stranded RNA (dsRNA) into an organism [150]. dsRNA can be introduced in the target organism through feeding, microinjections or by expression plasmids [145,150]. The rate of transcription has not been found to be affected by the introduction of dsRNA and gene silencing is caused by a major loss of the corresponding messenger RNAs (mRNA) after successful transcription [151]. Understanding of the mechanism underlying the use of dsRNA for RNAi was derived from the genetic studies on *C. elegans* [150]. The silencing of mRNAs is carried by an RNA-induced silencing complex (RISC) and occurs through a two-step process (Fig 1.6). The initiation of silencing occurs when dsRNA is recognized by the RISC complex and an enzyme (Dicer) is released to cleave dsRNA into small interfering RNAs (siRNAs) of approximately 22 nucleotide base pairs [152]. In the second step the RISC incorporates the double-stranded siRNAs and unwinds them in an ATP-dependent process before directing the siRNA to the homolog mRNA which is subsequently degrade [153].

In *Nasonia*, RNAi is performed by injecting double-stranded RNA that is homologous to the targeted gene into the late stage larvae using a microinjector [6,145]. The silencing is systemic in that it spreads from the injection site through the entire body of the wasp [145,150]. Parental RNAi was the first method to be applied in *Nasonia*, where dsRNA was injected in female pupae against embryo genes involved in early development and the offspring were examined for any changes in phenotype [153]. RNAi technique has

now expanded and has been adapted to all other stages in the development of the *Nasonia* life cycle (embryo, larvae and adult) [145,154,155]. A number of questions have been addressed using the RNAi technique, including the identification of novel systems for embryonic axial patterning and uncovering genes involved in sex determination in *Nasonia* [13,15]. *Nasonia* venom components have also been investigated using RNAi with probably the best example involving the functional characterization of the venom gene *calreticulin*. *Calreticulin* is a calcium binding protein and molecular chaperone that is found in the mitochondria, endoplasmic reticulum and is secreted in *Nasonia* venom. The protein has been shown to function in regulation of steroid hormones, signal-induced calcium release immune regulation and developmental arrest in envenomated fly hosts [131,156,157]. The calreticulin gene was down regulated in *Nasonia* through RNAi and the wild type and knock down wasps were exposed to hosts to allow them to sting and lay eggs. Eggs were removed on the hosts and RNA-sequencing was performed on the unstung hosts and from the envenomated hosts, the study showed that calreticulin may play a role in manipulating the immunity system of the host based on the clotting genes that had expression upregulated [145].

1.3.3.2.The generation of recombinant venom proteins *in vitro*

The extraction of venom proteins directly from the wasps' venom glands is difficult due to the small size of *Nasonia* (1-2 mm), hence the synthesis of recombinant protein offers an alternative way for studying wasp proteins. This method is relatively cheap to prepare culture mediums as compared to expensive liquid chromatography / mass spectrophotometry extraction methods which requires more starting material and expensive machinery. The recombinant expression systems requires optimization to ensure that all the genetic elements essential for the production of a folded protein are incorporated in the plasmid [158]. The essential genetic elements to be included are the origin of replication (ori), transcriptional promoters, antibiotic resistance marker, translation initiation regions (TIRs) and translational terminators (Fig 1.7) [159]. The pET expression system is a popular system that has been developed for a number of expression applications and includes multiple cloning sites, hybrid promoters and high number of genetic background modifications to both promote expression and to prevent breakdown of the recombinant protein in the cell [160,161]. A very important factor in recombinant

protein expression is the choice of heterologous host to be used. The host cell systems for expression of foreign genes can be either prokaryotic (bacteria) or eukaryote (filamentous fungi, yeasts, insect and mammalian cells) systems [162–164]. The chosen expression system must maintain plasmid and the productivity, bioactivity and the physicochemical properties of the protein of interest.

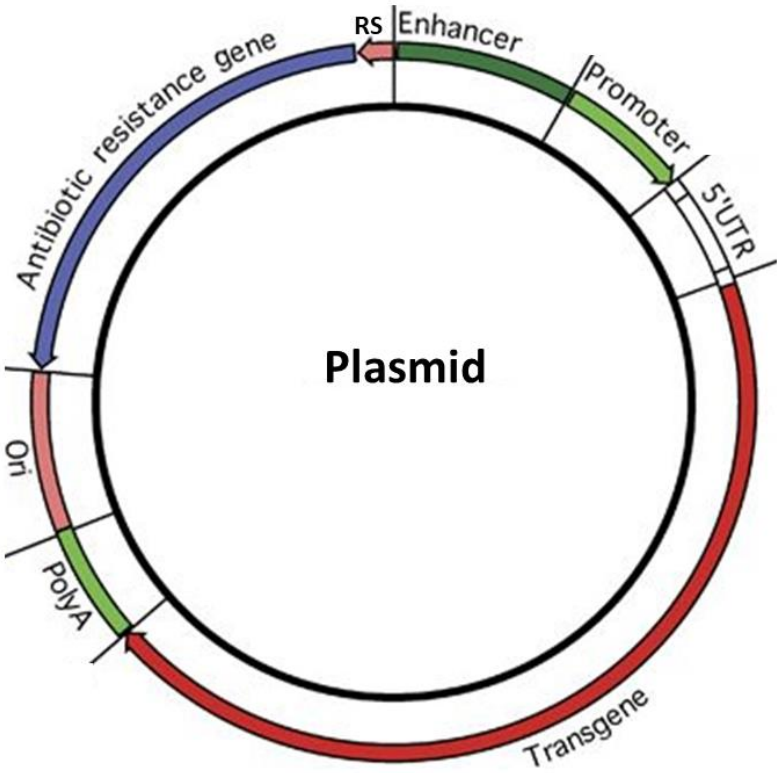


Figure 1.7. Genetic elements required for successful transgenesis. RS represents unique restriction sites.

Prokaryotic expression systems

The successful expression of mammalian peptide hormone in *E. coli* has revolutionised protein characterization as it enabled *in vitro* expression of foreign genes in prokaryotic cells. Since then *E. coli* has been the most popular and extensively used host cell in a lot of individual applications [165–169]. The advantages of *E. coli* include its rapid growth rate, low cost, and our strong knowledge of its physiological and metabolic systems. The availability of different strains and high cell density is also an advantage [170–172]. The most common strains of *E. coli* used for recombinant protein expression are BL21 (DE3)

and its derivatives (Table 1.2) due to their lack of Lon and OmpT proteases, enhanced permeability and produces low acetate at high glucose levels [173]. Several growth conditions of *E. coli* expression systems can be modified in order to optimise the production of recombinant protein. Some of these important growth conditions include: (i) composition of growth medium that can be used with different carbon sources and other supplements, (ii) temperature profile which is very important as it influences recombinant protein toxicity, stability and aggregation, (iii) cell density at the point of induction, which is at mid-log phase and A_{600} of ~0.5 to 0.8 and finally (iv) length of the induction period which plays a role in influencing the recombinant protein expression levels, toxicity and stability [173].

Three different forms of recombinant protein are expressed in *E. coli* expression systems based on the experimental design strategy, namely fusion proteins, secreted proteins or inclusion bodies. Fusion proteins are formed when the gene of interest has been fused or tagged with other sequences encoding a FLAG epitope tag or a polyhistidine and this is very important as tagging facilitates protein enrichment, purification or protection against degradation [174,175]. Secreted proteins are preferred as there is reduced potential degradation because they accumulate either in the periplasm or extracellular medium [176]. In addition, secreted proteins accumulating in the periplasm are easier to purify and it is more likely they will correctly fold, which increases the activity and stability of the recombinant proteins [176]. Most of the untagged eukaryotic proteins expressed in *E. coli* are expressed as inclusion bodies and they are largely protected from protein degradation by the host cell enzymes. However, they are a challenge to deal with as they are typically insoluble, misfolded and therefore requiring subsequent solubilization and refolding steps [176].

Table 1.2. Common bacterial strains for protein expression.

Bacterial Strain	Features	Growth requirements
BL21(DE3)	DE3 lysogen contains T7 polymerase upon IPTG induction.	No tight control over the expression, thus addition of 1% glucose to the growth medium should be considered
BL21(DE3)-pLysS	The pLysS plasmid produces T7 lysozyme to reduce basal level expression of the gene of interest.	Chloramphenicol 34 μg/mL
Origami	Origami host strains are K-12 derivatives that have mutations in both the thioredoxin reductase (*trx*B) and glutathione reductase (*gor*) genes, which greatly enhances disulfide bond formation in the cytoplasm.	Kanamycin 15 μg/mL Tetracyclin 12.5 μg/mL
Origami B	Origami B host strains carry the same *trx*B/*gor* mutations as the original Origami strains, except that they are derived from a *lac*ZY mutant of BL21.	Kanamycin 15 μg/mL Tetracyclin 12.5 μg/mL
Origami B pLysS	Same as original Origami but the plasmid reduce basal level expression of the gene of interest.	Kanamycin 15 μg/mL Tetracyclin 12.5 μg/mL Chloramphenicol 34 μg/mL
Rosetta	Rosetta host strains are BL21 *lac*ZY derivatives designed to enhance the expression of eukaryotic proteins that contain codons rarely used in *E. coli*.	Chloramphenicol 34 μg/mL
Rosetta pLysS	In Rosetta (DE3) pLysS the rare tRNA genes are present on the same plasmids that carry the T7 lysozyme.	Chloramphenicol 34 μg/mL
Rosetta-gami-pLysS	Rosetta-gami host strains are Origami derivatives that combine the enhanced disulfide bond formation resulting from *trx*B/*gor* mutations with enhanced expression of eukaryotic proteins that contain codons rarely used in E. coli.	Kanamycin 15 μg/mL Tetracyclin 12.5 μg/mL Chloramphenicol μg/mL
BL21trxB	BL21*trxB* strains possess the same thioredoxin reductase mutation (*trx*B) as the AD494 strains in the protease deficient BL21 background. The *trx*B mutation enables cytoplasmic disulfide bond formation.	Kanamycin 15 μg/mL
C41(DE3)	Effective in expressing toxic and membrane proteins from all classes of	None

	organisms and the strain C41(DE3) was derived from BL21(DE3)	
C43(DE3)	Same as C41	None

*pLysS plasmids are suitable for the expression of toxic genes

Despite the popularity and success of *E. coli* as a host expression cell system, various situations arise that impede the goal of getting the proteins of interest after protein production. No or low production of the desired protein is a common problem that is often caused by the recombinant protein harming the host cell [173, 174]. Rare codons can also play a role in slowing translation resulting in low production of the target protein. This latter problem can be rectified by using strains that supplement rare codons (Table 1.2) or through genetic engineering of the template DNA to alter these codons [177,178].

Inclusion bodies formation is another challenge that arises when the host cellular environment differs in terms of pH, osmolarity, cofactors, folding mechanisms and redox potential, from that in which the protein normally occurs [179]. These factors cause an unbalanced equilibrium between protein aggregation and solubilization, leading to protein instability and aggregation formation. Inclusion bodies can be prevented by tagging the protein of interest to another protein that can act as a solubility enhancer or through chaperone co-expression aiding the nascent polypeptides to reach their final structure [180,181]. Disulfide bond formation is vital for proteins to attain biologically active three-dimensional structures therefore erroneous formation of disulfide bonds can lead to misfolding resulting in formation of inclusion bodies. Strain that promote disulfide formation (Table 1.2) can be used to prevent this error. Further, disulfide bond isomerase (DsbC), a chaperone that both assist in proper folding proteins and promotes the correction of mis-oxidized proteins to attain their correct form can be used as well to avoid errors in disulfide bonds formation [182]. Although the formation of inclusion bodies can be problematic, in other situations they allow the recombinant protein to be more easily extracted and refolded in a single step process [180].

Molecular cloning using the prokaryotic expression system has been successfully applied on the *Nasonia* venom genes Kazal-type serine protease inhibitors and pacifastin protease [148,183]. These experiments utilizing *E. coli* expression system and a pGEX-4T-2 vector with the recombinant products fused to glutathione S-transferase. The recombinant Kazal-

type serine protease inhibitors and pacifastin protease were tested on the host *M. domestica* and both the proteins showed inhibitory effects on prophenoloxidase [148,183]. A second study utilized pGem-T Easy vector to express a recombinant defensin antimicrobial peptide from the venom of *Nasonia* [184] indicating that different cloning vectors can be used for studying the functions of *Nasonia* venom genes.

Eukaryotic expression system

The eukaryotic expression system is often used to produce recombinant proteins that are a challenge to get using a prokaryote system due to failure of forming a three-dimensional structure or the need for post-translational modifications on the protein of interest. Yeast is the most common eukaryotic expression systems because they have high yield, methods are available to assist protein folding, and they contain pathways that allow many posttranslational modifications to be added to the recombinant protein [182]. The two most utilized strains of yeast are *S. cerevisiae* and *P. pastoris*, which have both been used extensively in the recombinant expression of eukaryotic proteins because of the molecular, genetic and biochemical features that are shared with higher eukaryotes [185,186]. The *P. pastoris* is a popular heterologous expression system for the production of up to 2000 prokaryotic and eukaryotic proteins [187]. These yeast have secondary messenger signaling pathways that are very similar to those in higher eukaryotes which aids in the study of eukaryotic membrane proteins structure-function [188]. Despite these advantages, the usefulness of *S. cerevisiae* can be limited as higher eukaryotes contain sialylated O-linked chains as compared to O-linked oligosaccharides that contain only mannose in yeast [189]. Furthermore, products from *S. cerevisiae* often have immunological problems caused by the reduction in both activity and receptor binding due to over-glycosylation of N-linked sites [189]. *P. pastoris* lacks the capability to synthesize chaperonins that are important for protein folding therefore making it difficult to use this strain for the production of eukaryotic proteins that require folding assistance [190]. Filamentous fungi are also an attractive expression system mainly because of their ability to secrete large amounts of bioactive protein [191]. Despite the high level production of bioactive proteins, filamentous fungi contain about 80 protease genes that increase the possibility that the recombinant protein will be degraded [189].

Insect and mammalian cell expression systems are able to perform complex post-translational modifications that cannot be accomplished with fungi systems [192]. The main advantages of these expression systems is that they are capable of performing additional post-translational modifications, proper protein folding, high expression levels, and efficient cleavage of signal peptides [189]. In insect expression systems, baculovirus is the most commonly used vector system as it contains a gene encoding the protein polyhedrin, which is a viral promoter facilitating the high level expression of foreign genes [193]. Apart from high levels of productivity, mammalian cells are also capable of adding fatty acids to amino acids and phosphorylating tyrosine, threonine and serine hydroxyl groups [189]. Like any expression system, insect cells have drawbacks including the need for proper determination of post-translational processing patterns, inefficient secretion of recombinant proteins, inclusion bodies and errors in glycosylation [194]. The drawbacks for using mammalian cells include the high cost, risk for a potential product contamination by viruses and more demanding culture conditions [195,196].

1.3.4. Expression of *Nasonia* venom genes

Nasonia venom is made up of 79 peptides, with 56 having predicted functions based on sequence homology to proteins in other organisms [121]. The remaining 23 venom peptides have no homology to sequences found in the NCBI database. Attributing functions to *Nasonia* venom peptides has been a challenge because the classical biochemical methods that rely on purification and biological testing could not be used due to the extremely small volume of *Nasonia* venom that can be obtained from each insect. For this reason, high-throughput transcriptomic analysis that can use smaller volumes represents a practical approach to study parasitoid venoms. For example, *Nasonia* venom gene RNA interference coupled with RNA-seq in the envenomated host showed calreticulin suppress expression of innate immune cells [100]. The experiment showed that knocking down *Nasonia* venom gene allows the identification of host responsive genes and respective phenotypes on the host. Since there are 79 venom peptides, randomly targeting venom peptides using RNA interference presents a daunting task. In a study by Sim and Wheeler [197], RNA-seq was used to identify a list of the venom genes with the highest relative expression levels in the venom gland versus the ovary. This list of genes represents a short list of candidates for further functional characterization using RNAi

based on the hypothesis that high expressed loci are likely to be performing important roles in venom function (Table 1.3) [197].

Table 1.3. Top 10 *Nasonia* most differentially expressed genes in the venom gland versus the ovary [197]**.**

Gene Id	Annotation	Accession number
Nasvi2EG007167	*Serine protease 97* (SP97)	NP_001155042
Nasvi2EG009662	*Venom protein Z*	XP_001606832
Nasvi2EG006243	*Gamma-Glutamyl transpeptidase-like venom protein 1*	XP_001602724
Nasvi2EG013868	*Venom protein Y*	XP_001603046
Nasvi2EG009035	*Endonuclease-like venom protein*	XP_001604839
Nasvi2EG036525	*Aminotransferase-like venom protein 2*	XP_001600149
Nasvi2EG009647	*Venom protein Q*	XP_001607488
Nasvi2EG009648	*Venom protein G*	XP_001608198
Nasvi2EG001168	*Immunoglobulin-like venom protein*	lcl|hmm408134
Nasvi2EG014072	*Venom protein X*	XP_001603886

A *serine protease 97* (SP97) was identified as the highest differentially expressed gene between the ovary and the venom gland (Table 1.3 and Table 1.4). Based on total normalised read counts the gene encoding *venom Y* was found to be the highest venom expressed overall. Interestingly, the top five highest expressed venom genes could not be annotated using GO terms, highlighting the potentially important role played by novel genes in *Nasonia* venom function.

Table 1.4. The ten most highly expressed genes in the venom gland of *Nasonia* [197]**.**

Gene Id	Annotation	Accession number
Nasvi2EG013868	*Venom protein Y*	XP_001603046
Nasvi2EG009648	*Venom protein G*	XP_001608198
Nasvi2EG009662	*Venom protein Z*	XP_001606832
Nasvi2EG009647	*Venom protein Q*	XP_001607488
Nasvi2EG014072	*Venom protein X*	XP_001603886
Nasvi2EG007167	*Serine protease 97* (SP97)	NP_001155042
Nasvi2EG003930	*Kazal type serine protease inhibitor-like*	XP_001602179
Nasvi2EG014069	*Venom protein L*	XP_001606165
Nasvi2EG006243	*Gamma-Glutamyl transpeptidase-like venom protein 1*	XP_001602724
Nasvi2EG012348	*Venom protein H*	XP_001605133

1.4. Aims and Hypothesis

This project aims to:

I. Identify *Wolbachia* gene candidates that control cytoplasmic incompatibility (CI). In *Nasonia*, *Wolbachia* is maternally transmitted and manipulates the *Nasonia* reproductive system in ways that favour the generation of female offspring in CI. However, the mechanism of *Wolbachia* induced cytoplasmic incompatibility remains unknown thus, it is hypothesised in this study that genes that are differentially expressed when *Wolbachia* infects *Nasonia* ovaries and testes could provide clues to the underlying mechanism of *Wolbachia* induced CI.

II. To functionally characterize six venom genes. Previous work has shown that the wasp venom causes many changes that favour the growth and development of their offspring in the host (Section 1.3.2.1). The changes in the host include developmental arrest, immune system suppression, increase in the production of lipids and carbohydrates. The individual venoms that are responsible for each of these specific envenomation responses remains unknown. Therefore, a hypothesis is drawn that individual venom proteins play specific roles in the manipulation of

the host and this study aims to identify and functionally characterize key *Nasonia* venom encoding genes that are key to successful parasitism proteins using the host as a natural assay system.

1.4.1. Research objectives

Aim 1

- Generate a *Wolbachia* free strain of *Nasonia*.
- Identify *Nasonia* genes that are differentially expressed after *Wolbachia* infection in female and male *Nasonia* wasps.
- Synthesis of dsRNA targeting highly differentially expressed venom genes for functional characterization via RNAi.
- Search for lateral genes transfers through searching *Wolbachia* transcripts in the *Nasonia* transcriptome and genome.

Aim 2

- Establish effectiveness of RNAi in *Nasonia* using dsRNA targeting *E. coli LacZ* gene (negative control) and *Nasonia cinnabar* gene (positive control).
- Generate double stranded RNA targeting selected highly expressed venom genes for knock down using RNAi assay to observe phenotypes in the *L. sericata* host after envenomation.
- Compare development of cells in unstung and stung *L. sericata* hosts via transmission electron microscopy.
- Generate recombinant protein for candidate venom proteins of interest for structural characterization.

2. MATERIALS AND GENERAL METHODS

2.1. Culturing *Nasonia* wasps

The *Nasonia* LABII strain used in this project was a generous gift from Prof. Jack Werren (Werren Lab, University of Rochester). Four to six female wasps were carefully placed in 30 mL vial, PS V-base (LabServ) plugged with cotton wool and *Lucilia sericata* hosts (Biosuppliers Ltd, NZ) were supplied at a ratio of two hosts per wasp. The vial was left at room temperature, natural humidity and lighting for approximately 14 days and after that period, four to six female offspring were selected to repeat the process.

2.2. Ultrapure Distilled Water

Unless stated otherwise, UltraPure™ distilled water (Invitrogen) was used in all protocols and throughout this thesis; it is referred as 'pure water'.

2.3. Tissue Extraction

All equipment and bench were cleaned using RNAZap (Ambion). A working buffer was prepared by mixing 1 M of KH_2PO_4 and 1 M of Na_2HPO_4 at a ratio of 1:4. Approximately 20 µL of working buffer was pipetted on a concave microscope slide that was already placed at the microscope stage. One day old live wasp was placed on its back in the buffer and held between the thorax and the abdomen using a pair of iris forceps. Dissection was done on the abdomen section to extract the ovaries and testes in females and males respectively. The tissues were immediately placed in a 1.5 mL microcentrifuge tube containing Trizol (Ambion) placed in ice.

2.4. Single fly genomic DNA prep

Materials: Squishing buffer (10 mM Tris-HCl pH8, 1 mM EDTA, 25 mM NaCl) and 20 mg/mL Proteinase K stock. Proteinase K was added to the squishing buffer to make up a working solution with the enzyme final concentration of 200 µg/mL.

The fly was placed in a 1.5 mL microfuge and stored at -20 ℃ until ready to process. 50 µL of the working solution was pipetted into a yellow tip, which was then used to smash the fly against the walls of the microfuge tube without depressing the pipette plunger. Sufficient liquid escaped from the tip and was used to grind up the fly several times. The remaining buffer was expelled and mixed thoroughly by pipetting several times. The mixture was incubated at 37 ℃ for 30 minutes and the proteinase K in the solution was inactivated by heating to 95 ℃ for two minutes. The 1.5 mL microfuge was centrifuged at 8000 X g for 5 minutes and the supernatant was transferred to a new PCR tube. The extracted DNA was stored at -20 ℃ until ready for use.

2.5. RNA Extraction

All equipment and benches used for RNA extraction were first cleaned with 95 % ethanol followed by a thorough clean up using RNA-Zap (Ambion). *Nasonia* whole body or tissue were placed in 1.5 mL microfuge containing 150 µL of Trizol (Ambion). These were stored in -80 ℃ until ready for extraction. For RNA extraction, the wasp tissue or whole body was homogenised using a clean plastic pestle before the addition of Trizol (Ambion) to make up to 1mL. Following a 5-minute incubation at room temperature, 200 µL of chloroform was added and the solution was shaken vigorously for 15 seconds. Following a 5-minute incubation at room temperature, 200 µL of chloroform was added and the solution was shaken vigorously for 15 seconds. The extraction tube was centrifuged for 15 minutes at 12000 X g, then the aqueous layer carefully transferred to a new tube, and 0.5 X volume of 100% ethanol added to precipitate the RNA. The sample was then added to a spin column placed in a 2 mL collection tube and centrifuged at 12000 X g for 15 seconds. The flow-through was discarded and 700 µL of RW1 solution from the QIAGEN RNEASY mini kit (Qiagen) was added to the spin column. After, centrifugation at 12000 X g for 30 seconds was performed and the flow through was discarded. To wash the membrane, 500 µL of RPE solution was added to the spin column and centrifuged at 12000 X g for 30 seconds. The membrane was dried by centrifuging again at 12000 X g for 60 seconds. To elute the RNA, 30 µL of water was added to the spin column and centrifuged at 12000 X g for 60 seconds.

2.6. Polymerase Chain Reaction (PCR)

Polymerase chain reactions were employed to amplify specific DNA sequences for diagnostic tests. Primers were designed based on specific DNA sequences that are publicly available in the National Center for Biotechnology Information (NCBI) database. Unless stated otherwise, *Taq* DNA polymerase (Roche Applied Science) was used for all amplification reactions. The standard reaction set-ups are set as below:

Table 2.1. Standard PCR set-up for 25 µL PCR mixture.

Component	**Volume (µL)**
10 X Reaction Buffer	2.50
50mM MgCl	0.50
10mM dNTPs	0.50
10 µM Forward Primer	0.20
10 µM Reverse Primer	0.20
Taq polymerase (Roche)	0.20
Pure water	19.90

A master mixture was prepared in the event where more than one reaction was performed. The reaction mixture was prepared on ice and 1 µL of DNA or cDNA was added into the PCR mixture. The Biometra TGradient thermocycler was set up for PCR reactions as follows:

Table 2.2. Thermal profile used for amplification of DNA fragments.

Reaction detains:	**Temperature (℃)**	**Time (seconds)**
step 1: Initial denaturation	94	120
step 2: Denaturation	95	35
step 3: Annealing	50	30
step 4: Elongation	72	45
step 5: Final elongation	Go to step 2 for 35X	

The results of all PCR reactions were analysed using agarose gel electrophoresis.

2.7. Agarose gel electrophoresis

Agarose gel electrophoresis was used to separate negatively charged DNA fragments according to their sizes as they migrate in an electric field towards the positively charged anode. DNA migration is directly proportional to the length size of the DNA. Longer nucleotides travel slower through the agarose gel matrix. The concentration of agarose can be chosen based on the specific range of sizes to be separated. Unless stated otherwise, agarose gel concentration (%w/v) of 1% was used for analysis.

Materials: Agarose powder (Merck), 1 X TAE buffer (40 mM Tris, 20 mM Acetic acid, 20 mM EDTA) pH=8, ethidium bromide, 6 X DNA loading buffer (0.025% Bromophenol blue, 0.25% Xylene cyanol, 15% Ficoll), gel casting tray, electrophoresis tank, power supply and tanks.

Agarose powder was dissolved in 1 X TAE buffer through microwaving to make up a 1 % gel concentration. The boiling agarose was poured into the gel-casting tray and after solidifying, the samples mixed with loading dye were loaded into the gel wells. Electrophoresis was performed at a constant voltage of 90 V for approximately 45 minutes. Ethidium bromide in distilled water (2 µg/mL) was used to stain the gel. The GelDoc gel imaging system (Bio-Rad) was used to view the gel the DNA-ethidium, bromide complex exposed to UV light.

2.8. PCR purification

The QIAquick start protocol was followed for PCR purification. Briefly, five volumes of binding buffer (PB) was mixed with one volume of PCR reaction mix. The mixture was then applied to the QIAquick column that was centrifuged at 8000 X g for 60 seconds. The flow through was discarded and 750 µL of wash buffer (PE) applied to the column and centrifuged for 60 seconds at 8000 X g. The flow through was discarded and centrifugation was repeated at the same speed for two minutes to dry the column. To elute the DNA, 50 µL of elution buffer (EB: 10 mM Tris.Cl pH=8.5) was added to the column and centrifuged for one minute at a speed of 8000 X g.

2.9. Reverse Transcriptase-Polymerase Chain Reaction

The manufacturer's full protocol for the transcriptor first strand cDNA synthesis kit (Roche) was used to perform the RT-PCR. The reaction components were added into a microcentrifuge tube as described in Table 2.3. The microcentrifuge tube was incubated at 25 ℃ for 10 minutes followed by a 30 minutes incubation at 55 ℃. The transcriptase was inactivated by heating the reaction mixture for five minutes at 85 ℃. The cDNA was stored at -20 ℃ until ready to use.

Table 2.3. Reaction components for RT-PCR.

Component	Amount
Total RNA	1 µg total RNA of final concentration
Primer	1 µL
Water	To make final volume of 13 µL
Transcriptor reverse transcriptase reaction buffer	4 µL
Protector RNase inhibitor	0.5 µL
Deoxynucleotide	2 µL
Transcriptor reverse transcriptase	0.5 µL

2.10. Quantification of Nucleic Acids

The quality and quantity of prepared RNA and DNA was determined using the Nanodrop spectrophotometer (Thermo Fisher) at absorbance ratios of 260/280 and 260/230. At an absorbance ratio of 260/280, a ratio of ~1.8 is accepted as pure for DNA while for RNA, a ratio of ~2.0 is accepted as pure.

2.11. DNA Sequence Analysis

All DNA sequencing were carried out at the Massey Genome Service centre (MGS; Massey University, Palmerston North). Information about the equipment and chemicals used, please refer to http://www.massey.ac.nz/massey/learning/departments/centres-research/genome/massey-genome-service-home.cfm. The samples were prepared

according to the instructions given by MGS with the following requirements for a 20 µL final volume:

Table 2.4. Massey Genome Service template and primer concentration requirements. Final volume required is 20 µL.

Template Type	**Template total quantity (ng)**	**Primer total quantity (pmol)**
~~PCR product (bp)~~:		
100-200	1.25-5	4
200-500	5-12.5	
500-1000	12.5-25	
1000-2000	25-50	
>2000	50-125	
Single-stranded plasmid	62.5-125	4

The results obtained were then analysed and visualised using Geneious v. 9.1.8 package.

2.12. Double stranded RNA synthesis

Double stranded RNA was synthesized using the components from the MEGAscript RNAi kit (Ambion) following the manufacturer's instructions. The components were mixed as follows:

Table 2.5. Double stranded RNA synthesis reaction set up.

MEGAscript Component	**Volume (µL)**
ATP solution	2
CTP solution	2
GTP solution	2
UTP solution	2
Water	Up to 20 µL
10 X Reaction Buffer	2
Linear template DNA	0.1 – 1 µg
Enzyme mix (T7 enzyme mix, T7 RNA polymerase, RNase Inhibitor protein in 50 % glycerol)	2

2.13. Cultivation of bacterial cells

Luria Bertani (LB) medium was the liquid media used for bacterial growth and was prepared by adding 10 g of peptone, 5 g of yeast and 10g of NaCl in 1.0 L of distilled water. To make solid media, 1.6% (w/v) Bacto-Agar (Oxoid) was added. These media were sterilized through autoclaving at 121 ℃ and 2 X 10_5 Pa for 20 minutes. After autoclaving, the filter sterilised antibiotics ampicillin and/or chloramphenicol were added to LB and solid media to a final concentration of 100 µg/mL for ampicillin and 34 µg/mL for chloramphenicol, before pouring into plates.

The plates were stored at 4 ℃ until needed for use.

2.14. Restriction Digest

Plasmids were subjected to restriction enzymes and the reactions set up as guided by the kit manufacturer (Roche) (Tables 2.6 and 2.7). The reaction mixture was then incubated at 37 ℃ for 3 hours. To test the success of restriction digest, DNA from all the reactions was run on agarose gel and upon confirmation; the restriction digest mixture was cleaned up using the Qiaquick protocol (Section 2.8). The DNA was then stored at -20 ℃ until when needed for ligation.

Table 2.6. Restriction digest protocol for gene of interest and plasmid (Roche).

Component	**Volume (µL)**
10X Buffer B	10
DNA	50
Restriction enzyme 1 (*Bam*H1) Restriction enzyme 2 (*Eco*R1)	1 1
Pure water	38

Table 2.7. Restriction digest protocol for controls (Roche).

Component	Uncut vector (µL)	*Bam*H1 (µL)	*Eco*R1 (µL)
10X Buffer B	1	1	1
DNA	1	1	1
Restriction enzyme	-	0.2	0.2
Pure water	8	7.8	7.8

2.15. Ligation and Transformation

The cut plasmid and gene of interest were ligated following the manufacturer's instructions (Roche). As approximately 100 ng of vector was required for the ligation process, the insert was added at a vector: insert ratio of 1:3. No insert was added in the control reaction. The reaction mixtures were incubated at 4 ℃ overnight and the following day, 5 µL of the ligation mixture was mixed with 50 µL of competent cells (XL1 Blue) before being left on ice for 30 minutes. Heat shock was then applied to the reaction mixture by placing it in a water bath at 42 ℃ for 30 seconds before cooling on ice for five minutes. Approximately 700 µL of LB was added to the mixture which was then incubated at 37 ℃ for one hour with shaking (200 rpm). 100 µL of transformed cells were then plated on agar and incubated at 37 ℃ overnight.

2.16. Gene Expression

The plasmids containing the genes of interest were transformed into Rosetta B cells using heat shock as described above. A small-scale expression trial using 5 mL LB (with ampicillin and chloramphenicol) was undertaken at 37 ℃. Cells were induced by adding 1M of IPTG when the OD_{600nm} reached 0.6 -1.0. The cells were incubated at 25 ℃ while being shaken for either 3 or 8 hours. After this time, the cells were pelleted by centrifugation at 4000 X g for 20 minutes and the supernatant was discarded. The pellet was resuspended in minimal volume (~1 mL) lysis buffer (10 mM Tris-HCl pH 8.0, 5 mM DTT, 0.1 M NaCl, 2 mM EDTA). Cells were lysed by sonication using an energy of ~160 J and an amplitude of 15. A small fraction of the total solution (~ 1 mL) was pipetted into a 1.5 mL microcentrifuge labelled whole cell solution while the rest was centrifuged

at 4000 X g for 10 minutes at 4 ℃. The supernatant was collected into a well labelled tube (soluble proteins) and the pellet was resuspended in 8 M Urea and the solution was labelled as insoluble proteins.

2.17. SDS-PAGE

Denaturing polyacrylamide gel electrophoresis was used to separate proteins based on their size Tables 2.8 and 2.9. The molecular weight of proteins was then predicted through comparison to a protein standard (BioRad, Fermentas) containing proteins of known molecular weight.

Table 2.8. Separating gel preparation (16%).

Separating gel components	**Volume (mL)**
Distilled Water	3.5
1.5 M Tris-HCl pH=8.8	2.5
10 % (w/v) SDS stock	0.1
Acrylamide/Bis (40% stock)	4.0
10% Ammonium persulfate	0.1
TEMED	0.005

Table 2.9. Stacking gel preparation (5.2%).

Stacking gel components	**Volume (mL)**
Distilled Water	6.1
0.5 M Tris-HCl pH=6.8	2.5
10 % (w/v) SDS stock	0.1
Acrylamide/Bis (40% stock)	1.3
10% Ammonium persulfate	0.1
TEMED	0.01

3. TRANSCRIPTOMICS OF *WOLBACHIA*-FREE AND *WOLBACHIA*-INFECTED *NASONIA VITRIPENNIS*

3.1. Introduction

Nasonia infected with the intracellular parasite *Wolbachia* display a form of cytoplasmic incompatibility (CI) that is manifested by dramatic changes to the wasps' reproductive biology. In the case of *Nasonia* CI, *Wolbachia* infected sperm is incompatible with eggs from females that are either uninfected by the bacteria or infected by a different *Wolbachia* strain [43]. As uninfected females are at a reproductive disadvantage, rates of *Wolbachia* infection can rapidly increase in the host population. For instance, the introduction of *Wolbachia* into the north eastern Australian mosquito population, *Aedes aegypti*, resulted in up to 90% infection across the population within 5 weeks of the initial introduction [198]. The mechanism underlying CI remains unknown and thus the experiments in this chapter are based on the hypothesis that CI mechanisms require expression changes in both the *Wolbachia* and the wasp (Chapter 1, Section 1.6). To explore this hypothesis, dual RNA-seq was used to simultaneously measure gene expression in both *Wolbachia* and the reproductive tissues of the wasps that are directly exposed to this bacterium. This method therefore allows us to identify putative *Wolbachia* molecular signals that triggers CI, as well as the molecular responses to the induction of CI in the host *Nasonia*. Characterization of the molecular changes that occur during CI will help us better understand this important mechanism, which is important given *Wolbachia* mediated CI is already being using to control problematic insect populations (such as the Dengue virus vector *Aedes aegypti*).

The method of RNA-seq has largely replaced microarrays in gene expression studies due to its increased sensitivity, specificity, and the ability to detect novel transcripts [199]. The high sensitivity and specificity of RNA-seq opens up the possibility to measure gene expression simultaneously in two species (Dual RNA-seq), thus providing us with a tool that allows us to explore interactions between host and pathogens at the molecular level. This method has previously been applied in studies that have explored interactions

between *Salmonella* and their host human cells, where it was shown that bacterial PhoP activated human small RNAs in order to control expression of virulence genes required for intracellular survival [200]. The method has also been used to show that *Wolbachia* is able to up-regulate its purine biosynthesis genes to benefit its nematode host *Brugia malayi*, which lacks these important metabolic pathways [201].

The experiments in this chapter identified differentially expressed genes in the *Wolbachia* infected (Wol+) and *Wolbachia* free (Wol-) *Nasonia* reproductive tissue. Based on their differential expression in response to *Wolbachia* infection, seven *Nasonia* genes were selected as candidates to target for molecular characterization using RNA interference (RNAi). RNAi in *Nasonia* is induced by injecting double stranded RNA (dsRNA) complimentary to the target gene sequence into the wasps at the larvae stage of development. The dsRNA induces sequence specific degradation of target transcripts, and thus allows us to investigate the possible functions of the gene by studying observing phenotypic changes that occur when the gene product is lost. Importantly, *Nasonia* RNAi is systemic, that is the silencing signal spreads throughout all tissues of the animal and remains stable into adulthood. The RNAi treated *Nasonia* were then mated to test for the changes to the typical CI induced reproductive incompatibilities, unfortunately none of the seven candidates were shown to have significant effects on CI.

3.2. Methods

3.2.1. Generating a *Wolbachia* free line of *Nasonia* (Wol-)

Wildtype (Wol+) *Nasonia* were fed a 10 µg/ml of solution of the antibiotic tetracycline (TET) in honey (Arataki) (50% v/v honey in water) daily to ‘cure’ the *Wolbachia* infection. Cured *Nasonia* wasps (Wol-) were maintained through the provision of blow fly (*Lucilia sericata*) hosts while stored in a 30 mL vial, PS V base (LabServ). Every two weeks, when new offspring emerged, 4-6 female wasps were transferred into new 30 mL vial, PS V base (LabServ) and provided with hosts. Testing for *Wolbachia* (Section 3.2.1.1) was regularly performed to ensure that the Wol- wasps had not been reinfected.

3.2.1.1. Test for *Wolbachia*

Quick DNA extraction ("Squish" prep)

The DNA extraction solution consisted of squishing buffer (10 mM Tris-Cl pH 8.2, 1 mM EDTA, 25 mM NaCl) with 200 µg/mL Proteinase K. A single *Nasonia* wasp was crushed in 50 µL of DNA extraction solution using a pipette tip, then incubated at 25 ℃ for 30 minutes. Proteinase K was then inactivated by heating the extraction solution to 95 ℃ for 2 minutes.

Assay for *Wolbachia*

To test for the presence of *Wolbachia*, PCR primers (fbpA_F: 5'-GCTGCT CCRCTTGGYWTGAT-3' and fbpA_R: 5'-CCRCCAGARAAAAYYACTATTC-3') were designed to target the *Wolbachia fructose bisphosphate aldolase gene* (fbpA) [31]. A second set of positive control primers (COI_F: 5'-CATAGATCAGGCAAGTAAGG-3' and COI_R: 5'-AATATGATCAGGAGTATTAGGG-3') targeted the *Nasonia cytochrome oxidase* gene (COI). PCR was carried out using the following reaction conditions: 30 cycles of 95 ℃ for 30 seconds, 60 ℃ for 15 seconds, 72 ℃ for 30 seconds. The resulting products were visualized on a 1% w/v agarose gel with a 1kb ladder.

3.2.1.2. Wol- and Wol+ cross to test Cytoplasmic Incompatibility

Reciprocal crosses between virgin Wol- females and Wol+ males were set up with an overnight mating period. Triplicate crosses of five females and four males per replicate were set up for a 4-hour mating period. At the end of the mating period, female wasps were moved to individual tubes and provided with fresh *Lucilia sericata* hosts for a 24-hour period to allow stinging and egg laying. The stung hosts were incubated at 27 ℃ for 10 days. At the end of this period the number and sex of *Nasonia* pupae that resulted from each cross and replicate was recorded.

3.2.2. RNA extraction

Before each extraction all benches and equipment used were cleaned with RNA-Zap (Ambion). The dissection of *Nasonia* wasps was done in PBS under a compound microscope and the dissected ovaries and testes tissue were immediately placed in a 1.5 mL microfuge containing Trizol (Ambion). RNA extraction was performed as per

protocol in Section 2.5. The final RNA yield and quality was assessed using a Nanodrop spectrophotometer and all RNA samples were also tested on a 2100 Bioanalyzer Eukaryote Total RNA Nano gel to further check the quality of the RNA. Samples were only submitted for sequencing if they had an RNA integrity number above seven, and the OD_{260}/OD_{280} and OD_{260}/OD_{230} values were above 1.8 and 1.3 respectively. The RNA Integrity Number (RIN) assigned to a eukaryote total RNA sample indicates the presence or absence of degraded RNA.

3.2.3. Transcriptome sequencing and RNA-seq analysis

NZGL Ltd (Otago Genomics and Bioinformatics Facility (OGBF), University of Otago, Dunedin) conducted the preparation of TruSeq stranded total RNA libraries and performed the paired-end sequencing on an Illumina HiSeq2500 platform. The majority of the rRNA were depleted using the RiboMinus™ technology (Ambion). The FastQC v.0.11.5 package (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/) was used to perform quality control checks to ensure that the raw data had no biases that may affect downstream analysis such as adapter dimer contamination. The data was pre-processed using fastq-mcf (http://code.google.com/p/ea-utils/wiki/FastqMcf) to remove adapters and primers. SolexaQA++ v.3.1.4 [202] was used to quality trim the reads to a minimum Phred score threshold of >20 and any reads less than 50 bp were removed.

The quality filtered short read data was mapped to the *Nasonia* and *Wolbachia* genomes separately using Tophat2 v. 2.1.1 [203] and Bowtie2 [204] respectively with default settings, except for the library type option set to fr-firststrand, and maximum intron length reduced to 10000 for Tophat2. The *Nasonia* (GCA_000002325.2) and *Wolbachia* (GCA_000204545.1) reference genomes were downloaded from NasoniaBase (http://hymenopteragenome.org/nasonia/?q=tools) and NCBI (http://www.ncbi.nlm.nih.gov/) respectively. Gene based read counts were generated using HTSeq-count v. 0.6.1 [205] with the default settings, except that the strand-specific assay was set to reverse using the flag –s reverse. Taxonomic classification of unmapped reads was done using the Kaiju webserver (http://kaiju.binf.ku.dk/). Differential expression analysis was performed using DESeq2 [206] with a false discovery rate significance threshold set at 0.05. Graphical representation of data counts was

transformed using the DESeq variance stabilizing transformation method. The transformed data was then used for visualizing the overall effect of *Wolbachia* in all samples through principal component analysis (PCA) and heatmap plots. Functional annotation of genes was done using Blast2Go [207].

3.2.4. RNA interference of differentially expressed genes

3.2.4.1. Generation of double stranded RNA for RNA interference

Primers were designed to amplify coding regions of target genes and the resulting PCR product was used to generate dsRNA. The targets included the negative control *lacZ* that is not found in the *Nasonia* genome, and the visual eye color marker *cinnabar*, which generates wasps with bright red eyes when RNAi is successful. The full list of primers used to generate dsRNA templates is provided in (Table 3.1).

Table 3.1. Primers for amplifying Wol- male and female, *lacZ* and *cinnabar* gene fragments.

Gene	Primers	Size (bp)
lacZ	F: TAATACGACTCACTATAGGGAGACCACGTTTTACAACGTCGTGACTGG R: TAATACGACTCACTATAGGGAGACCACGGGCCTCTTCGCTATTACG	150
cinnabar (CN)	F: TAATACGACTCACTATAGGGAGACCACCGGTGAAGGTGAGGTCGTCG R: TAATACGACTCACTATAGGGAGACCACCTCGCATCGGGGTAGGGAAG	534
NV_16517-RA (F1)	F: TAATACGACTCACTATAGGGAGACCACGAAAGATATACCTGTACGGC R: TAATACGACTCACTATAGGGAGACCACTTTCTCTATATCAGCGGACT	314
NV_10329-RA (F2)	F: TAATACGACTCACTATAGGGAGACCACAGGATTTTGATATTGACGCA R: TAATACGACTCACTATAGGGAGACCACCTAATTTAGGTTGCTTCGGA	486
NV_15393-RA (M1)	F: TAATACGACTCACTATAGGGAGACCACCCATTCCAGGTGTCTCTACA R: TAATACGACTCACTATAGGGAGACCACGGGAAAATTTCTCGGCCATA	458
NV_18293-RA (M2)	F: TAATACGACTCACTATAGGGAGACCACTACATTGGATCAAACATCCC R: TAATACGACTCACTATAGGGAGACCACTGAGGTTTCAAGATTTCGTT	309
NV_11725-RA (M3)	F: TAATACGACTCACTATAGGGAGACCACCAGAATCATTTGTCGTTCAC R: TAATACGACTCACTATAGGGAGACCACGACTGATCCAATCAACGTAA	474
NV_16404-RA (M4)	F: TAATACGACTCACTATAGGGAGACCACCAATTAAAATCTCGCGATCC R: TAATACGACTCACTATAGGGAGACCACAATCATTACCTGCCAGAATC	466
NV_17088-RA (M5)	F: TAATACGACTCACTATAGGGAGACCACCGCAAAAGTCAACATGAAAG R: TAATACGACTCACTATAGGGAGACCACATGCTATGCACTTTATTCGG	376

The RNAi off-target prediction tool on the WaspAtlas website (http://waspatlas.com/tools/RNAi) was used to predict if the designed dsRNA sequences had any RNAi off-target effects. The genes of interest (Table 3.1) were supplied to the

RNAi off-target prediction tool as query. The PCR reaction was set up as follows: 94 °C for 2 min, a 1st step of 6 cycles: 95 °C for 30 seconds, 55 °C for 30 seconds and 72 °C for 45 seconds; a 2nd step of 30 cycles: 94 °C for 30 seconds, 72 °C for 45 seconds, and finally 72 °C for a 5-minute extension. The expected PCR product sizes were confirmed using agarose gel electrophoresis before the product was purified using the QIAquick PCR purification kit (Qiagen). The purified PCR product (1 µg) was used as a template to generate dsRNA using the Megascript T7 transcription kit (Ambion), as described by the manufacture.

3.2.4.2. RNA interference in Wol- *Nasonia*

For increase of the number of larvae that could be collected from each host, *Nasonia* females were conditioned to sting by pre-exposure to hosts at a ratio of one female per two hosts for a period of 24 hours. The preconditioned females were then given fresh hosts for an overnight period (~18 hours). The females were removed, and the stung hosts were incubated for a further 5 days at 27 °C. After this time, larvae were removed from the host and placed on 1.5% PBS agar plates (5 mM NaCl, 0.56 mM KCl, 1 mM Na_2HPO_4, 1.5% v/v agar) for injection. Microinjection capillary needles were prepared using a Sutter Instrument Co. needle-pulling machine (model P-2000) under these settings: Heat - 270, Velocity - 255, Pull - 4, Delay - 85, Filament - 70. They were back loaded with 5 µL of 1 µg/µL dsRNA dissolved in 10% (v/v) Hansells blue food colouring. The dsRNA was injected into the posterior region of the larvae using a FemtoJet express transjector (Eppendorf) with the regulator set at 600 psi and the pressure of each injection set at 1000psi Successful injection was visually confirmed by the presence of a blue dye stain on the posterior end of the larvae. The injected larvae were left to develop to late stage pupae on the agar plates stored at 27 °C for six days. Knocked down female and male wasps were separated at the pupae stage, transferred to a 30 mL vial, PS V base (LabServ) until eclosion.

3.2.4.3. Training and reciprocal crosses

RNAi knocked mating tests using virgin *Nasonia* (males or females) were set-up for mating with both Wol+ and Wol- (females or males respectively) for a period of 4 hours (Fig 3.1). Mated female Wol- wasps were then provided the host's pupae in a 30 mL vial,

PS V base (LabServ) for 24 hours to allow them to learn how to sting and lay eggs. The host's pupae were carefully removed from the wasps, placed in a new 30 mL vial, PS V base (LabServ), and stored at 27 °C. The wasps were then re-hosted overnight and carefully separated from host's pupae which were incubated at 27 °C for 12 days. The host pupae were cracked open to count the number of *Nasonia* offspring.

Figure 3.1. Experimental design and expected results of a cytoplasmic incompatibility test on the knocked down wasps. Normally infected males are incompatible with uninfected females while uninfected males are compatible with both infected and uninfected females. Males: square, females: circles, KD: knock down, grey: infected, white: uninfected. A: CI causative gene will be identified if Wol- KD male becomes incompatible with Wol- female or B: Wol- KD female becomes compatible with Wol+ male.

3.3. Results

3.3.1. Generation of *Wolbachia* free wasps

The aim of this chapter was to identify the genes that are differentially expressed (DE) in Wol- and Wol+ wasps and to determine whether knock down of DE genes results in alterations to the typical CI induced reproductive incompatibilities. These experiments were performed by a comparative analysis of transcriptomic data of *Wolbachia* infected (Wol+) and *Wolbachia* free (Wol-) *Nasonia*. As there was only a Wol+ stock available at

the start of this study, Wol- wasps were first generated by feeding Wol+ wasps with tetracycline (10µg/ml) dissolved in honey over 3 generations. The *Wolbachia* status of the antibiotic treated *Nasonia* was determined using PCR and CI based test as described in section 3.3.1.1 and 3.3.1.2, respectively.

3.3.1.1. Amplification of fructose bisphosphate A and cytochrome oxidase 1 in *Nasonia* genomic DNA

Fructose bisphosphate A (*fbpA*) is a *Wolbachia* housekeeping gene that is conserved among *Wolbachia* strains and is used as a diagnostic for the presence of the bacteria [31]. Genomic DNA was extracted from the antibiotic treated wasps (Wol-) and the antibiotic-free Wol+ wasps for use as template DNA in an amplification of *fbpA* to confirm the presence/absence of *Wolbachia*. The size of *fbpA* is 524bp and as expected amplification of the correct size product was observed in the Wol+ genomic DNA, yet no amplification was observed in the antibiotic treated wasps. This result confirms the antibiotics had successfully 'cured' the Wol- line of *Nasonia* (Fig 3.2.). A second set of primers that target the *Nasonia* housekeeping *COI* were used to ensure that the absence of an amplicon in the Wol- genomic DNA was not caused by degraded template in this sample [31]. As expected, the *COI* primer set amplified a product of the expected size in both the Wol+ and Wol- lines of *Nasonia* (Fig 3.2.). This result confirms the generation of a *Wolbachia* free *Nasonia* Wol- line.

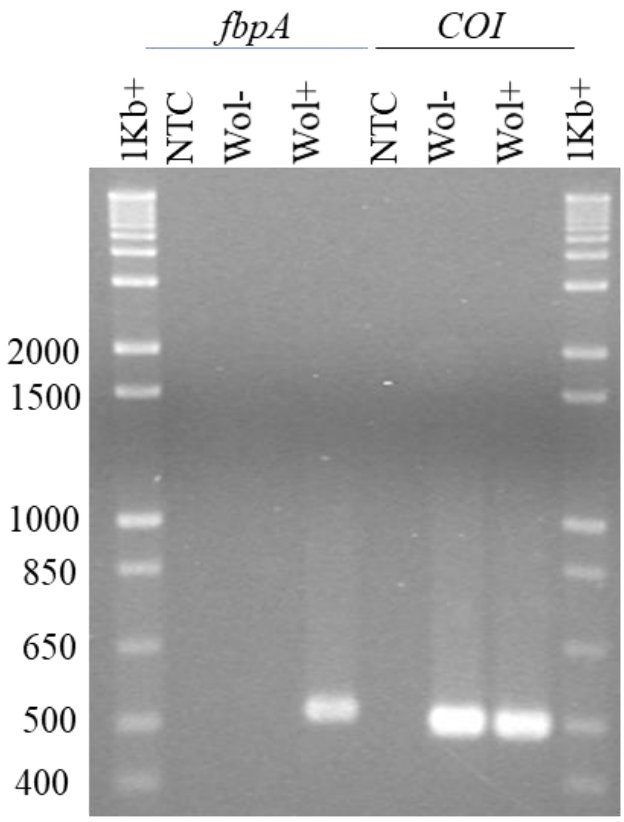


Figure 3.2. Test for the presence of *Wolbachia* using two sets of primers that target *Wolbachia fructose bisphosphate aldolase* (*fbpA*) and *Nasonia cytochrome oxidase 1* (*COI*). Wol- represent genomic DNA extracted from antibiotic wasps; Wol+, DNA from *Wolbachia* infected wasps, NTC, no template (DNA) control. Ladder is 1kb plus (ThermoFisher Scientific) and the units shown are in bp. The expected sizes of the amplified products 524bp (*fbpA*) and 486bp (*COI*). No amplification of *fbpA* in the Wol- sample confirms the absence of *Wolbachia* genomic DNA.

3.3.1.2. Reciprocal crosses between Wol+ and Wol- confirms cytoplasmic incompatibility

To further test for the absence of *Wolbachia* in the cured wasps, a test for the expected CI was conducted. A cross between Wol- females and Wol+ males should result in no female offspring, as any fertilised diploid eggs will die during early embryogenesis due to CI. In contrast, a reciprocal cross between Wol+ females and Wol- males will result in the successful production of female offspring from fertilised eggs. Consistent with the

expectations of CI, the crosses between Wol- females and Wol+ males produced no female progeny, with the observed male progeny (Table 3.2) resulting from eggs that are unfertilised by *Wolbachia* infected sperm (and thus not subject to CI). Wol+ females and Wol- males produced an average of 81 female offspring. These results are consistent with the findings from the PCR test (Section 3.3.1.1.) and confirm the successful generation of a *Wolbachia* free line (Wol-) that could be used for further experiments to identify *Nasonia* DE genes that may play a role in the molecular mechanism of CI.

Table 3.2. Progeny counts from reciprocal crosses between *Wolbachia* free (Wol-) and *Wolbachia* infected (Wol+) *Nasonia* lines.

Cross	Number of hosts	Progeny
Wol- females and Wol+ males	8	39 males, 0 females
	8	26 males, 0 females
	8	28 males, 0 females
Wol+ females and Wol- females	8	44males, 71 females
	8	37 males, 85 females
	8	23 males, 84 females

3.3.2. Transcriptome sequencing and RNA-seq analysis

3.3.2.1. RNA extraction

Wolbachia is concentrated mainly in the reproductive tissues of *Nasonia* where it induces CI; therefore, the ovaries and testes of both Wol- and Wol+ one-day old virgin wasps were dissected for total RNA extraction. The RNA was assessed for quality and quantity as a requirement for sequencing (RNA-seq) purposes. The quality and quantity of the RNA was initially assessed on a Nanodrop spectrophotometer at an absorbance of 260nm (Table 3.3) and the quality of the quantified RNA was analysed by measuring the $A_{260/280}$ and $A_{260/230}$ ratios with the Nanodrop spectrophotometer. The $A_{260/280}$ and $A_{260/230}$ ratios predict the presence of contaminants in the sample, with a low $A_{260/280}$ ratio indicating the presence of residual phenol or proteins while a low $A_{260/230}$ ratio indicates residual phenol or carbohydrates. The NZGL sample submission requirements for RNA-seq includes $A_{260/280}$ and $A_{260/230}$ absorbance values above 1.8 and 1.3, respectively. The RNA Integrity

Number (RIN) assigned to a eukaryote total RNA sample indicates the presence or absence of degraded RNA and ranges from 1-10, where 10 indicates the absence of degraded products and 0 represents completely degraded RNA. All samples exceeded these minimum sample quality criteria. Further, the total RNA samples were then analysed on a Bioanalyzer 2100 machine (Agilent) to determine the integrity of the extracted RNA and all the samples had a RIN above 7.7 (Table 3.3), passing the RNA-seq quality check minimum requirement number of 7. Based on these quality criteria, 12 total RNA samples, 3 from each tissue (Wol+ testes and ovaries, Wol- testes and ovaries) were submitted to the OGBF (NZGL) where TruSeq stranded total RNA sequencing libraries were prepared. Using a total RNA library preparation method ensures that both the *Nasonia* and bacterial transcriptomes will be sequenced (dual RNA-seq). The rRNA depletion kit was used to eliminate rRNAs from the library.

Table 3.3. Summary of the RNA concentrations and quality obtained by Nanodrop and Bioanalyzer machines. Sample names are derived from the source of RNA. Wol-Ov: *Wolbachia* free ovaries, Wol+Ov: *Wolbachia* infected ovaries, Wol-Tes: *Wolbachia* free testes and Wol+Tes: *Wolbachia* infected testes.

Sample	Concentration (ng/µl)	RNA Integrity Number (RIN)	$A_{260/280}$	$A_{260/230}$
Wol-Ov1	455	10	2.17	2.22
Wol-Ov2	572	10	2.16	1.75
Wol-Ov3	595	10	2.17	2.36
Wol+Ov1	224	10	2.17	2.12
Wol+Ov2	214	10	2.19	2.16
Wol+Ov3	566	10	2.15	2.15
Wol-Tes1	63.6	8.2	2.10	1.53
Wol-Tes2	58	9.4	2.16	1.72
Wol-Tes3	44	7.7	2.08	1.43
Wol+Tes1	62.2	8.1	2.09	1.34
Wol+Tes2	56.7	9.0	2.02	1.39
Wol+Tes3	54.7	9.0	2.05	1.58

3.3.3. Illumina sequencing and RNA-seq

The libraries were run on an Illumina HiSeq 2500 machine (Illumina Inc.) in the high output mode generating 125bp paired end sequences. The run generated 978,085,344 reads (Table 3.4), with an average Phred quality of 35.76. The Phred score is a metric used to assess the accuracy of a sequencing platform and to each nucleotide base call in the automated sequencer traces [208]. The average of 35.76 is equivalent to a probability of 1 in 3767 incorrect base calls indicating that the data was of good quality and could be used confidently for downstream analysis. The first step in the transcriptome analysis was to remove the adapters and primers attached to the reads during sequencing and to remove the reads that were less than 50bp after the trimming process. Paired reads increase the mapping confidence of reads in regions with repeat content. After trimming and removing reads shorter than 50bp, 80% of the data remained as paired end (Appendix 1), and this data was used for read mapping the reference genomes.

Table 3.4. Overview of *Nasonia* (Nas) and *Wolbachia* (Wol) transcriptome coverage in ovary and testes. 'Assigned' refers to reads that were mapped to a specific gene while 'No feature' represented reads that were mapped to a section of the genome that had no gene annotation.

Sample	**Total Reads**	**Mapped (%)**		**Multiple Alignments (%)**		**Assigned (%)**	**No Feature (%)**
		Nas	Wol	Nas	Wol	Nas	Nas
Wol-Ov1	85,591,020	92.63	0.0	3.6	0.0	83.21	10.24
Wol-Ov2	87,231,102	86.48	0.0	3.5	0.0	85.41	8.86
Wol-Ov3	83,138,074	90.1	0.0	3.8	0.0	84.09	9.36
Wol+Ov1	84,142,140	92.3	0.6	3.9	8.5	82.46	9.61
Wol+Ov2	76,626,942	93.2	0.4	3.4	8.1	82.14	10.43
Wol+Ov3	85,517,880	92.1	0.6	3.2	8.65	84.11	10.55
Wol-Tes1	84,807,058	91.3	0.0	3.9	0.0	81.13	12.36
Wol-Tes2	77,767,570	91.9	0.0	4.0	0.0	84.13	10.02
Wol-Tes3	80,082,656	87.5	0.0	4.6	0.0	83.82	9.42
Wol+Tes1	69,804,208	38.1	0.3	4.0	9.8	85.53	9.43
Wol+Tes2	78,153,148	87.6	1.4	5.5	10.1	84.79	9.27
Wol+Tes3	85,223,546	88.5	1.1	3.5	9.7	86.80	8.98

The high-quality trimmed reads were mapped to both the *Nasonia* and *Wolbachia* genomes to quantify the expression of genes in the bacteria and host. An average of 86.05% of the reads successfully mapped to the published *Nasonia* genome [86] (Table 3.4). A lower mapping percentage of 38.1% was observed in the Wol+ *Nasonia* male sample (Wol+Tes1; Table 3.4). In order to investigate the reasons for this lower mapping percentage, the Kaiju webserver was used to classify the taxonomic distribution of unmapped reads from the Wol+Tes1 sample. This analysis revealed an unexpected large percentage (~80%) comprised of unclassified and bacterial hits (Fig 3.3). *Wolbachia* is grouped under the Proteobacteria phylum and as expected a significant percentage of reads (~25%) were taxonomically classified in this group. Despite the low mapping percentage of Wol+Tes1 on the *Nasonia* genome, the sample was used for further analysis because read depth normalization occurs during the RNA-seq analyses (although there is an expectation for this sample to be subject to greater statistical noise).

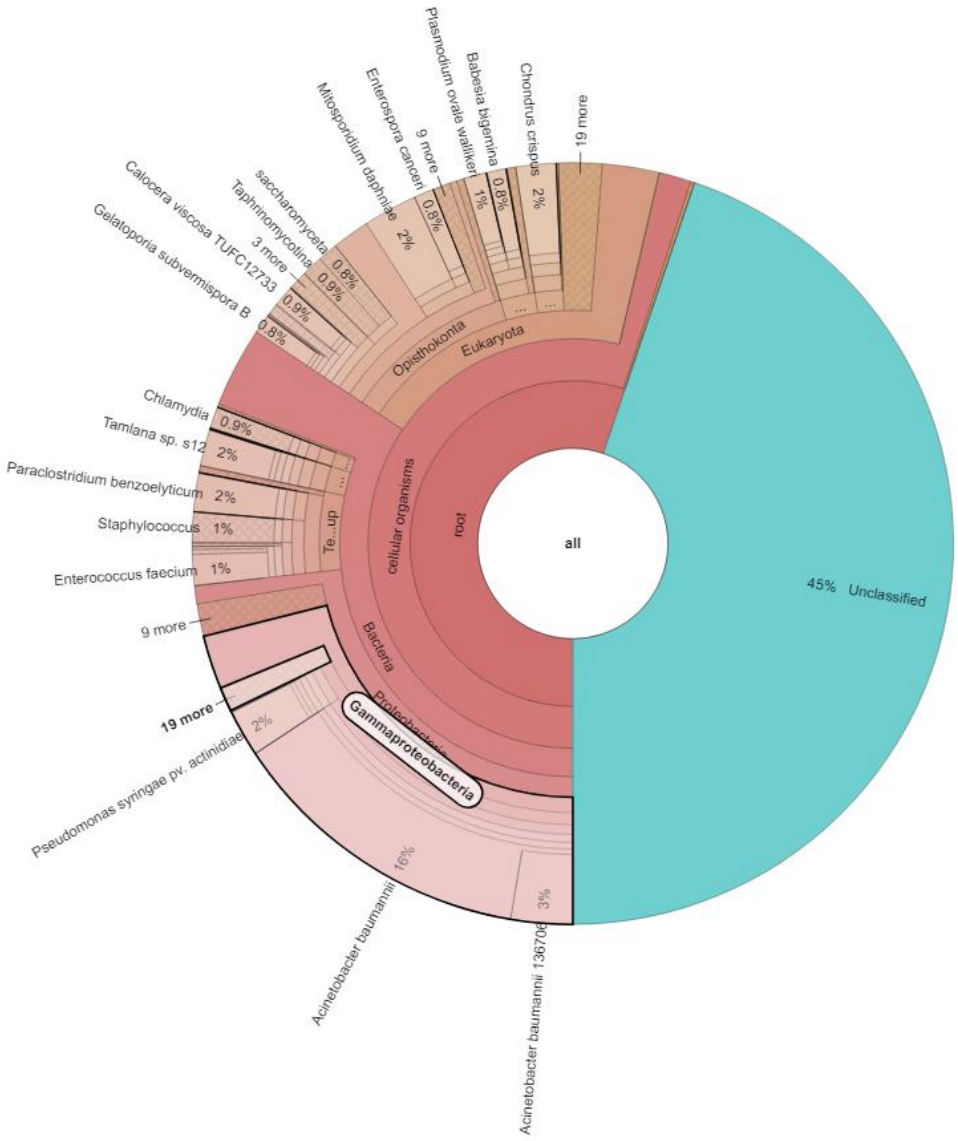


Figure 3.3. Taxonomic classification of Wol+Tes1 unmapped reads predicted using the Kaiju webserver. Different colours are used to differentiate classification and more indicates more species of the same order classification.

The number of reads mapped to the *Wolbachia* genome was very low with an average mapping percentage of 0.5% and 0.9% in Wol+ females and males, respectively (Table 3.4). The purpose of performing dual RNA-seq was to profile the expression of both host and pathogen but the low mapping from the *Wolbachia* transcriptome dataset meant that the data could not be used with confidence for downstream analysis. Unexpectedly, an average of 5065 high quality reads from the Wol- lines mapped to the *Wolbachia* genome.

This result was not consistent with the PCR findings (Section 3.3.1) or the CI test experiment, which both indicated that the line was *Wolbachia* free. Therefore, further analyses was carried out to try and better understand this unexpected result. This analysis suggested that expression from a *Wolbachia* lateral gene transfer (LGT) in the *Nasonia* genome was causing the positive mapping to the *Wolbachia* genome. The LGT results are described in Chapter 4.

An average of 9.8% of reads mapped to the *Nasonia* genome were labelled as "no feature" and could not be assigned to any known genes in the *Nasonia* genome (Table 3.4). This result reveals a possibility of uncharacterized genes, in the *Nasonia* genome, that are being expressed during the CI response however this was not investigated further due to time constraints.

3.3.4. Differential Expression analysis

Gene based read counts generated from the short read mapping data was used with the R package DESEQ2 [206] to detect significant differential expression between genes in the Wol+ and Wol- *Nasonia* tissues. Two independent pairwise comparisons were made between Wol+ vs Wol- ovaries and Wol+ vs Wol- testes, with positive fold change genes 'up regulated' in the *Wolbachia* infected tissue.

The first step of DESEQ2 normalizes the read counts based on the overall number of reads sequenced for each sample. This was done by dividing each count over its geometric mean across all samples. A negative binominal distribution was then used to calculate differential expression between loci under the two conditions and p-values were calculated. To reduce the effect of multiple testing, false discovery rate (FDR) adjusted p-values were calculated and a conservative cut-off threshold was set to 0.05 [206].

Gene expression comparison between Wol+ and Wol- females

The DESEQ2 analysis identified 84 *Nasonia* genes in the *Wolbachia* infected ovaries (Table 3.5, Fig 3.4) that were differentially expressed at an adjusted p-value of less than 0.05. The range of log2 fold changes between the two ovary samples was from -3 to 1.8, indicating the effect of down regulation in the Wol+ tissue is stronger than the up regulation. Also, the number of down-regulated genes is higher (62 genes) as compared

to the up-regulated genes (22 genes). BLAST2GO was used to assign GO terms to 84 of the differentially expressed genes and the assigned annotations showed the genes to be mainly involved in membrane and transport functions. BLAST2GO could not assign annotations to 11 differentially expressed genes (Table 3.5).

Table 3.5: GO terms of *Nasonia* genes manipulated by *Wolbachia* in the ovaries. Log2 fold change is the log-ratio of a gene's expression values in two different conditions (uninfected and *Wolbachia* infection). padj: p-adjusted value estimates how correct the call is for the accepted differential expressed genes based on the cut-off.

Gene	**log2 Fold Change**	**padj**	**description**
Nasvi2EG000722	-3.57755	1.07E-28	venom acid phosphatase acph-1-like
Nasvi2EG013996	-2.42072	1.36E-18	---NA---
Nasvi2EG003896	-2.42382	3.91E-17	intraflagellar transport protein 20 homolog
Nasvi2EG016026	-2.7133	2.53E-16	facilitated trehalose transporter tret1-like
Nasvi2EG011462	-2.48989	3.60E-16	multiple inositol polyphosphate phosphatase 1-like
Nasvi2EG014000	-1.95302	3.60E-16	---NA---
Nasvi2EG015664	-2.70144	3.60E-16	novel protein
Nasvi2EG033025	-2.0825	9.72E-15	gag-pol polyprotein
Nasvi2EG013999	-1.9721	2.53E-14	---NA---
Nasvi2EG010520	-1.85155	2.45E-10	PREDICTED: hypothetical protein LOC100120806 [*Nasonia vitripennis*]
Nasvi2EG021225	1.315167	7.62E-09	uncharacterized protein cg3556-like isoform 2
Nasvi2EG014580	-2.15444	7.72E-09	collagen alpha-1 chain-like
Nasvi2EG024518	-1.8182	7.72E-09	copia protein
Nasvi2EG013998	-1.74124	1.03E-07	---NA---
Nasvi2EG010897	-1.09523	5.76E-07	organic cation transporter
Nasvi2EG011624	-1.70817	4.13E-06	retroelement pol polyprotein
Nasvi2EG029947	1.486028	6.49E-06	---NA---
Nasvi2EG028057	1.838419	1.08E-05	ankyrin repeat family a protein 2-like
Nasvi2EG000723	-1.81152	1.64E-05	venom acid phosphatase acph-1-like
Nasvi2EG034664	-1.21503	2.28E-05	aaa family atpase
Nasvi2EG016480	-1.54473	3.87E-05	PREDICTED: hypothetical protein LOC100679471 [*Nasonia vitripennis*]
Nasvi2EG012043	-1.49153	5.71E-05	uncharacterized protein loc100867799
Nasvi2EG010355	1.219266	7.17E-05	PREDICTED: hypothetical protein LOC100116530 [*Nasonia vitripennis*]
Nasvi2EG036951	-1.27159	9.05E-05	solute carrier family 28 member 3-like
Nasvi2EG030600	1.009885	0.000436	udp-glucose 4-epimerase
Nasvi2EG018100	-1.12222	0.000471	probable cytochrome p450 6a13

Nasvi2EG006396	1.58662	0.000526	PREDICTED: hypothetical protein LOC100678308 [*Nasonia vitripennis]*
Nasvi2EG014067	-0.9241	0.000792	uncharacterized protein loc584740
Nasvi2EG016037	-1.1889	0.002024	mpv17-like protein 2-like
Nasvi2EG007239	0.865235	0.002552	speckle-type poz protein
Nasvi2EG025132	1.459018	0.00284	scavenger receptor class b member 1-like
Nasvi2EG029946	1.026829	0.002938	transposase domain-containing protein
Nasvi2EG013632	1.411261	0.004135	PREDICTED: hypothetical protein LOC100120435 [*Nasonia vitripennis*]
Nasvi2EG013403	-1.28747	0.00521	monocarboxylate transporter 9
Nasvi2EG001403	-1.38225	0.005816	PREDICTED: hypothetical protein LOC100679784 [*Nasonia vitripennis*]
Nasvi2EG002308	-1.34421	0.006349	haemolymph lipopolysaccharide-binding
Nasvi2EG006230	-0.88753	0.006864	probable 2-oxoglutarate dehydrogenase e1 component mitochondrial-like
Nasvi2EG019480	0.6882	0.007175	PREDICTED: hypothetical protein LOC100678380 [*Nasonia vitripennis*]
Nasvi2EG013081	1.319317	0.00796	si:dkey- protein
Nasvi2EG028056	-1.04008	0.008114	nuclease harbi1-like
Nasvi2EG003403	-1.30096	0.00901	proton-coupled amino acid transporter 4-like
Nasvi2EG031855	-0.74383	0.009532	PREDICTED: hypothetical protein LOC100678913 [*Nasonia vitripennis*]
Nasvi2EG027700	1.126953	0.010118	transmembrane protein 135-like
Nasvi2EG029118	0.958317	0.011452	transposase domain-containing protein
Nasvi2EG014637	-0.92108	0.011986	retrotransposon ty1-copia subclass
Nasvi2EG003549	0.740269	0.014274	protein distal antenna-like
Nasvi2EG004139	-1.25284	0.016697	plasma alpha-l-fucosidase
Nasvi2EG004115	-1.26119	0.01791	retrovirus-related pol polyprotein from transposon tnt 1-94
Nasvi2EG010609	-0.88186	0.022457	PREDICTED: hypothetical protein LOC100121890 [*Nasonia vitripennis*]
Nasvi2EG013664	-1.23372	0.022457	feline leukemia virus subgroup c receptor-related protein 2-like
Nasvi2EG025530	0.945369	0.028689	---NA---
Nasvi2EG004520	-1.01084	0.029437	PREDICTED: hypothetical protein LOC100680511 [*Nasonia vitripennis*]
Nasvi2EG003498	-1.25193	0.030769	xanthine dehydrogenase
Nasvi2EG015593	-0.69465	0.031643	PREDICTED: hypothetical protein LOC100119140 [*Nasonia vitripennis*]
Nasvi2EG000134	-1.2465	0.031857	---NA---
Nasvi2EG009218	1.071654	0.031857	oxidase peroxidase
Nasvi2EG012630	-0.72773	0.035498	juvenile hormone-inducible protein
Nasvi2EG013276	1.042454	0.036588	hymenoptaecin
Nasvi2EG019943	-1.04823	0.036588	---NA---
Nasvi2EG013601	-1.13011	0.038274	ornithine mitochondrial-like

Nasvi2EG015724	-1.11068	0.038274	haemolymph lipopolysaccharide-binding
Nasvi2EG010309	0.734569	0.03848	transketolase-like protein 2-like isoform 1
Nasvi2EG002058	-1.19964	0.041582	---NA---
Nasvi2EG002385	0.734146	0.041582	polycomb protein
Nasvi2EG002431	-0.9339	0.041582	zinc finger cchc domain-containing protein 24-like
Nasvi2EG004461	-0.93165	0.041582	transitional endoplasmic reticulum atpase-like
Nasvi2EG007942	-0.92772	0.041582	protease m50 membrane-bound transcription factor site 2 protease
Nasvi2EG018621	0.610057	0.041582	uncharacterized protein loc100879586
Nasvi2EG019585	-0.90105	0.041582	tpa_inf: venus kinase receptor
Nasvi2EG029760	-1.2131	0.041582	hypothetical protein DAPPUDRAFT_72494 [Daphnia pulex]
Nasvi2EG031971	-0.89019	0.041582	PREDICTED: hypothetical protein LOC100678210 [*Nasonia vitripennis*]
Nasvi2EG000662	-1.09812	0.041766	fibroblast growth factor 18-like
Nasvi2EG000244	1.156981	0.043671	PREDICTED: hypothetical protein LOC100121611 [*Nasonia vitripennis*]
Nasvi2EG004171	-1.0015	0.043671	uncharacterized protein loc100882767
Nasvi2EG006332	-1.19606	0.043671	carboxypeptidase b-like
Nasvi2EG010665	-1.17971	0.043671	amylase-related protein
Nasvi2EG034405	-1.09011	0.043671	rrna promoter binding protein
Nasvi2EG000441	-0.85454	0.045735	set and mynd domain-containing protein 4-like
Nasvi2EG013304	-0.84434	0.045735	---NA---
Nasvi2EG029761	-1.07299	0.045735	---NA---
Nasvi2EG011707	-1.02358	0.046556	gag-pol polyprotein
Nasvi2EG021557	-0.86575	0.047265	uncharacterized protein loc100869430
Nasvi2EG007726	1.093501	0.048849	uncharacterized protein loc100865432

A PCA analysis was used as a control parameter to show the relationships between the sample replicates used for differential expression calculations and this showed that two of each Wol+ and Wol- ovary samples were in two different PC1 levels (Fig 3.4b). Interestingly, the lack of a clear separation of the treatment groups (Wol+Ov3 and Wol-Ov1 samples) along the PC1 axis would suggest that *Wolbachia* infection only has subtle effects on the *Nasonia* transcriptome. One outlier from each group was observed at PC1 level and this was confirmed by the sample to sample distance plot (Fig 3.5). Two clusters of 3 samples each were observed where two Wol+ and one Wol- (Wol-Ov1) formed the top cluster while the bottom cluster was made up of two Wol- and one Wol+ outlier (Wol+Ov3).

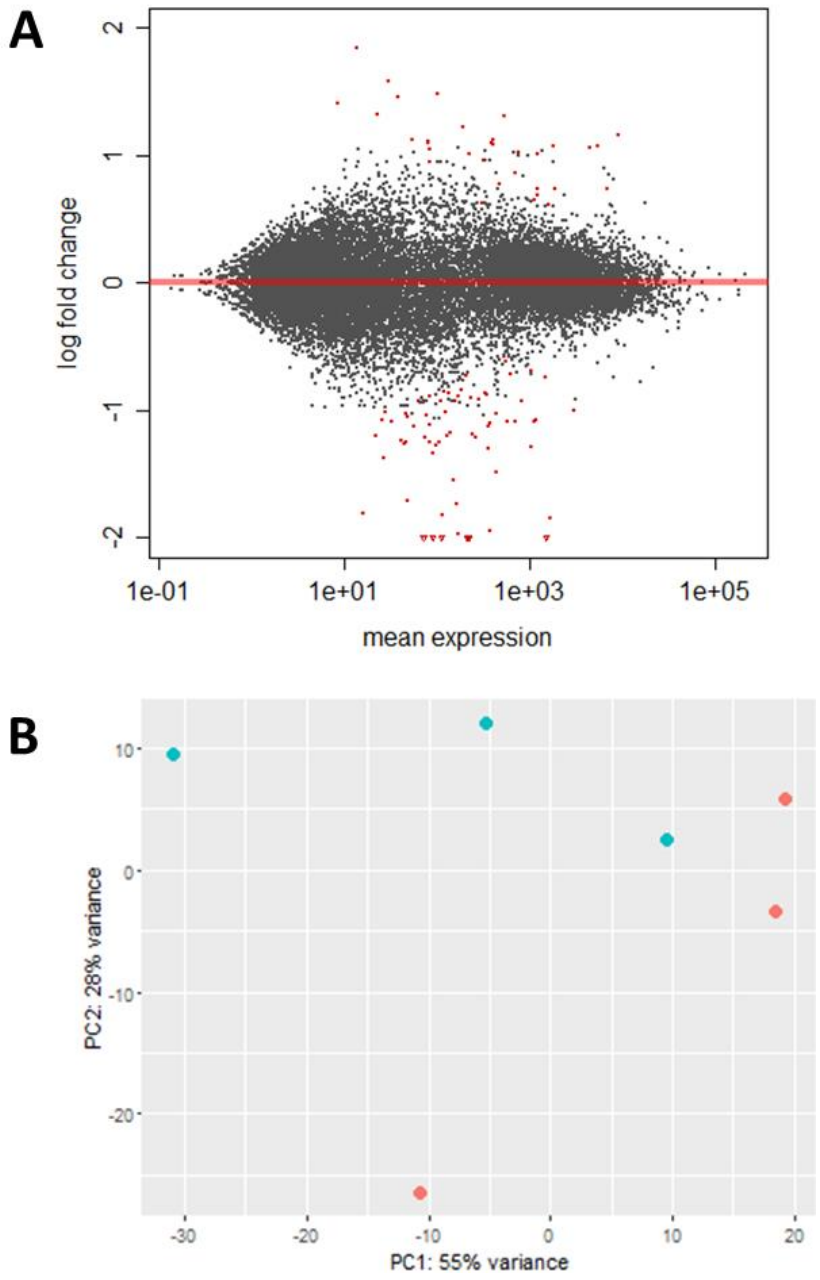


Figure 3.4. Differential expression analysis of Wol+ and Wol- female *Nasonia* ovaries. A: log fold change vs mean expression data where the red dots represent genes that are significantly differentially expressed at an adjusted p-value of 0.05. Red triangles are significant genes whose fold change is greater than the y-axis limit. B: Principal component analysis showing clustering of samples where red represents Wol- and blue Wol+ line.

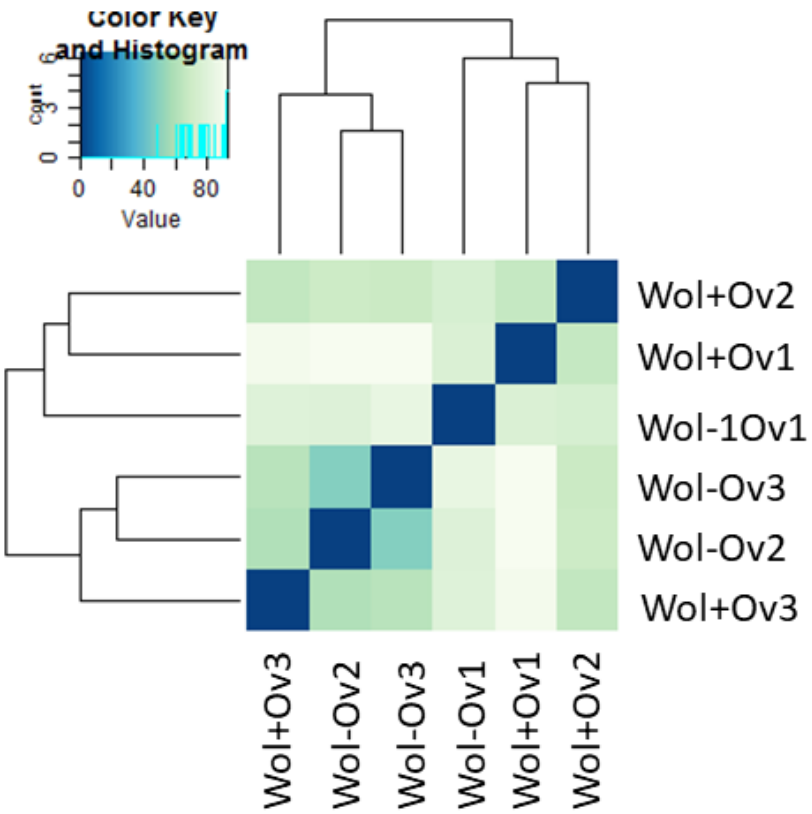


Figure 3.5. Sample to sample distance. Heat map showing the Euclidean variance stabilizing transformation. Wol+: *Wolbachia* infected, Wol-: *Wolbachia* free, Ov: Ovary sample. Colour range from white to blue- blue means identical while white is not similar.

Gene expression comparison between Wol+ and Wol- *Nasonia* males

In the *Nasonia* testes tissue, the comparison of gene expression between Wol+ and Wol- wasps showed that 58 genes were differentially expressed. In the testis tissue, 55 genes were down-regulated in the presence of *Wolbachia* whereas only three were up-regulated at an adjusted p-value of 0.05 (Table 3.6, Fig 3.6a). BLAST2GO analysis showed annotations for 25 of the differentially expressed genes, mostly with terms related to catalytic activity. The PCA analysis between the Wol- and Wol+ samples showed a general separation of the treatment groups along the PC1 access, but once again this clustering was reasonably weak (Fig 3.6b). Two testes samples, one from Wol+ and one from Wol- (Wol-Tes3 and Wol+Tes2) were placed together at both the PC1 and PC2 levels of the PCA plot (Fig 3.6b). This showed that the two samples were closely related

as confirmed by the sample to sample distances plot that clustered the two samples as a clade (Fig 3.7). Wol+Tes1 sample that had the lowest mapping percentage was observed to be clearly separated from all other Wol+ samples in the PCA plot.

Table 3.6. GO terms of *Nasonia* genes manipulated by *Wolbachia* in the testes. Log2 fold change is the log-ratio of a gene's expression values in two different conditions (uninfected and *Wolbachia* infection). padj: p-adjusted value estimates how correct the call is for the accepted differential expressed genes based on the cut-off.

Gene	log2FoldChange	padj	description
Nasvi2EG007347	-3.56747	7.41E-24	serine protease 50 precursor
Nasvi2EG010293	-3.49108	1.49E-23	trypsin-1-like
Nasvi2EG017845	-3.49909	1.49E-23	haemolymph lipopolysaccharide-binding
Nasvi2EG003848	-3.01263	1.35E-20	serine protease
Nasvi2EG013755	-3.32138	1.85E-19	maltase 1
Nasvi2EG011613	-3.03045	4.14E-17	mite allergen der f 3-like
Nasvi2EG002671	-2.71072	1.74E-15	maltase 1-like
Nasvi2EG006310	-2.8259	1.97E-15	PREDICTED: hypothetical protein LOC100115777 [*Nasonia vitripennis*]
Nasvi2EG015981	-2.59131	3.68E-14	PREDICTED: hypothetical protein LOC100117656 [*Nasonia vitripennis*]
Nasvi2EG012873	-2.68682	1.78E-12	uncharacterized protein loc100883880
Nasvi2EG002914	-2.13698	3.49E-09	PREDICTED: hypothetical protein LOC100114783 [*Nasonia vitripennis*]
Nasvi2EG015703	-2.21307	3.49E-09	sodium potassium calcium exchanger 5
Nasvi2EG007567	-2.18394	3.90E-09	esterase fe4
Nasvi2EG009660	-2.41731	7.41E-09	PREDICTED: hypothetical protein LOC100680432 [*Nasonia vitripennis*]
Nasvi2EG015589	-1.9108	1.71E-08	laccase 1
Nasvi2EG001978	-1.87931	1.75E-08	facilitated trehalose transporter tret1-like
Nasvi2EG022158	-2.2311	3.91E-08	myrosinase 1-like
Nasvi2EG010034	-1.95588	1.72E-07	speckle-type poz
Nasvi2EG016112	-1.97509	1.72E-07	esterase fe4
Nasvi2EG006334	-2.19888	1.90E-07	low quality protein: coagulation factor ix
Nasvi2EG008702	-1.83109	5.69E-07	zinc carboxypeptidase a 1-like
Nasvi2EG019613	-1.97782	1.02E-06	cytochrome p450 9e2
Nasvi2EG000442	-2.15551	1.11E-06	pheromone-binding protein 3-like
Nasvi2EG011002	-2.14806	1.26E-06	pancreatic triacylglycerol lipase-like
Nasvi2EG011000	-2.12528	1.55E-06	pancreatic triacylglycerol lipase
Nasvi2EG015590	-2.08396	1.55E-06	laccase 1
Nasvi2EG001187	-1.92787	3.15E-06	aminopeptidase n-like

Nasvi2EG007214	-2.01322	4.04E-06	deoxyribonuclease i
Nasvi2EG015820	-1.97062	1.10E-05	mite allergen der f 3-like
Nasvi2EG003845	-1.8686	2.06E-05	tripartite motif-containing protein 2-like
Nasvi2EG009872	-1.75809	8.43E-05	venom protein r-like protein
Nasvi2EG013386	-1.71354	8.43E-05	dopamine n isoform a
Nasvi2EG003947	-1.83954	0.000134	inorganic phosphate cotransporter-like
Nasvi2EG009670	-1.83609	0.000134	serine protease 122 precursor
Nasvi2EG006338	-1.55716	0.000163	serine protease 49 precursor
Nasvi2EG016622	-1.40642	0.000202	beta-hexosaminidase b
Nasvi2EG007159	1.445425	0.000225	---NA---
Nasvi2EG008530	-1.45573	0.000369	neutral alpha-glucosidase ab-like
Nasvi2EG013070	-1.76264	0.000369	ejaculatory bulb-specific protein 3
Nasvi2EG010147	-1.52003	0.000472	udp-glucuronosyltransferase 2c1-like
Nasvi2EG023377	-1.62954	0.000682	uncharacterized protein loc100865493
Nasvi2EG008726	-1.54884	0.000864	PREDICTED: hypothetical protein LOC100123346 [*Nasonia vitripennis*]
Nasvi2EG009665	-1.55577	0.001588	pacifastin-related serine protease inhibitor precursor
Nasvi2EG010997	-1.61128	0.00185	---NA---
Nasvi2EG013873	-1.41512	0.002346	pao retrotransposon peptidase superfamily
Nasvi2EG013754	-1.57058	0.003127	maltase 1
Nasvi2EG013436	1.227564	0.003743	dynein heavy chain
Nasvi2EG002291	-1.54912	0.004188	phosphatidylethanolamine-binding protein
Nasvi2EG011377	-1.54659	0.005163	transthyretin-like periplasmic protein
Nasvi2EG018534	-1.40275	0.005421	---NA---
Nasvi2EG008018	-1.3201	0.009419	transmembrane protease serine 9-like
Nasvi2EG015664	-1.42306	0.010298	novel protein
Nasvi2EG009601	-1.47234	0.012803	PREDICTED: hypothetical protein LOC100677926 [*Nasonia vitripennis*]
Nasvi2EG023376	-1.44847	0.013181	uncharacterized protein loc100883975
Nasvi2EG012422	-1.43703	0.018276	uncharacterized protein loc100879344
Nasvi2EG019597	-1.34815	0.034078	carbonic anhydrase 11
Nasvi2EG010298	-1.378	0.034908	trypsin-1- partial
Nasvi2EG016364	-1.26192	0.038858	probable maltase h-like

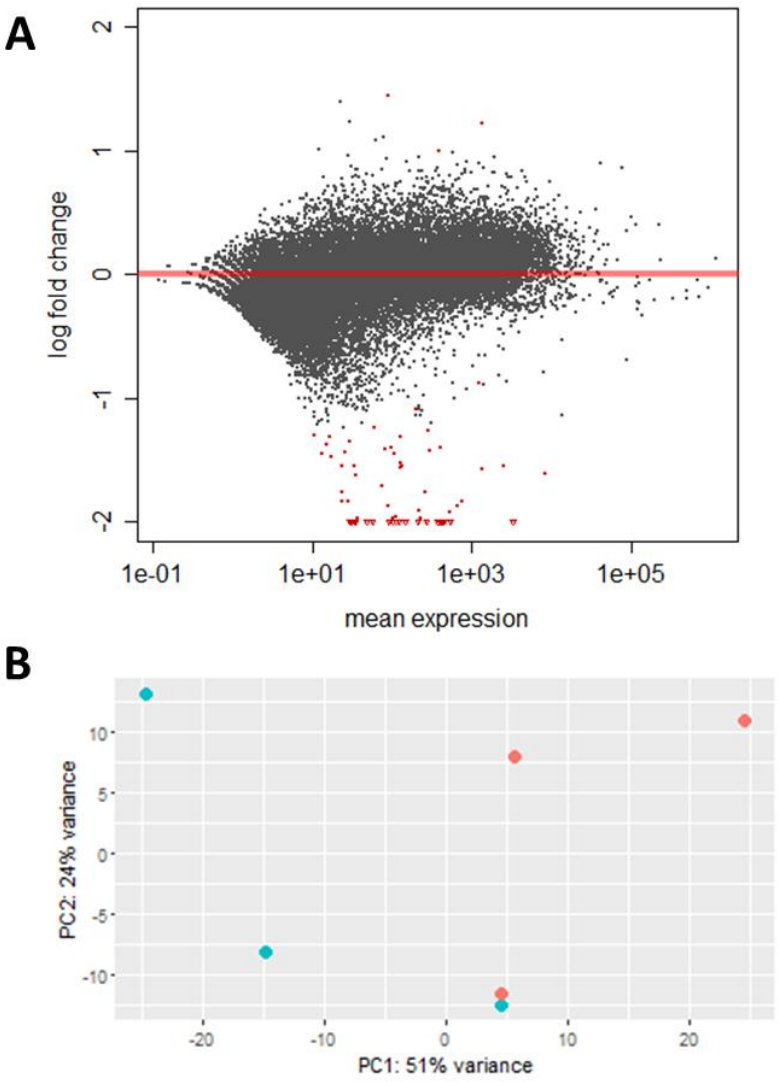


Figure 3.6. Differential expression analysis of Wol+ and Wol- male testes. A: log fold change vs. mean expression data where red dots represent genes that are significantly differentially expressed at an adjusted p-value of 0.05. Red triangles are significant genes whose fold change is greater than the y-axis limit. B: Principal component analysis showing PC1 and PC2 levels where red represents Wol- and blue Wol+ line.

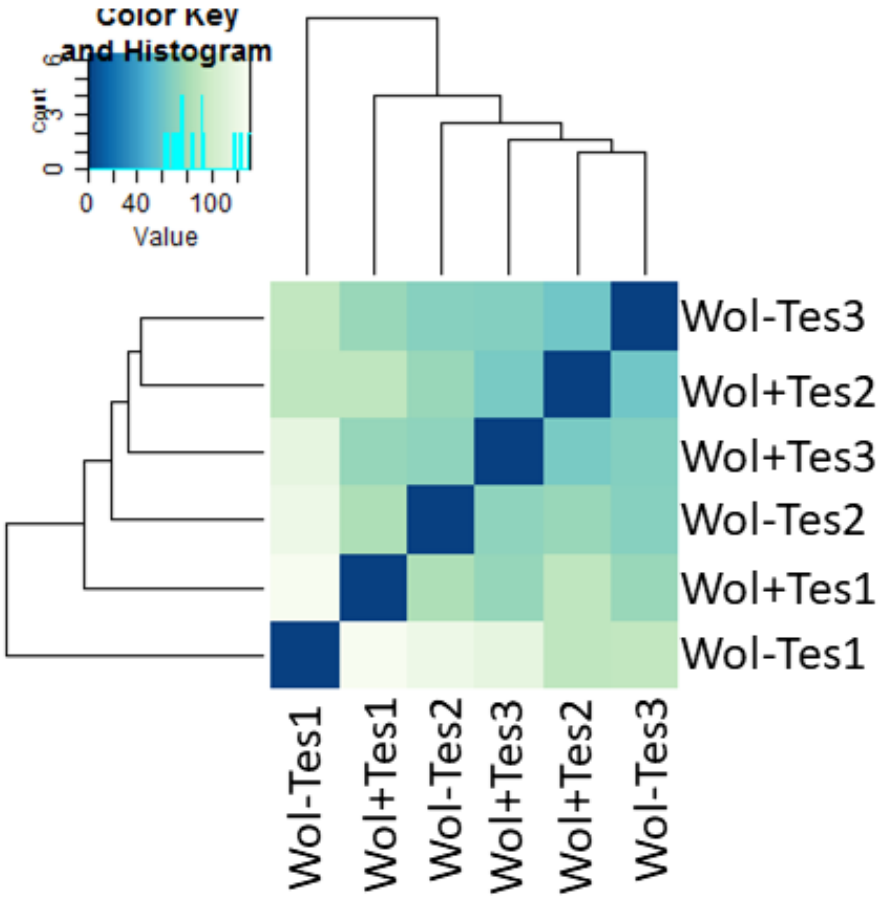


Figure 3.7. Sample to sample distance. Heat map showing the Euclidean variance stabilizing transformation. Wol+: *Wolbachia* infected, Wol-: *Wolbachia* free, Ov: Ovary sample. Colour range from white to blue- blue means identical while white is not similar.

3.3.5. Semi-quantitative RT-PCR

RNA-seq analysis revealed genes that were differentially expressed between Wol+ and Wol- reproductive tissues, and here semi qRT-PCR was used as an independent experiment to validate the changes in gene expression. cDNA was synthesized from RNA that was extracted from the ovary and testis tissues of both Wol- and Wol+ female and male wasps. An experiment that agrees with RNA-seq analysis would result in amplification of target genes after 36 PCR cycles and lower amplification intensity in Wol+ wasps' cDNA when compared to Wol- wasps' cDNA at fewer PCR cycles. The control gene which was shown not to be down-regulated in the RNA-seq experiment showed bands of similar intensity after amplification for 25 and 36 cycles. Based on the

RNA-seq analysis *NV_15393-RA* and *NV_10329-RA* genes were down-regulated in testis and ovary, respectively, were selected for testing using semi qRT-PCR (Fig 3.8). RP49 was used as a template control for the integrity of cDNA and in both female and male experiments, RP49 amplification was similar in Wol+ and Wol- showing that similar amounts of template cDNA were found in both samples. At 25 cycles, in the Wol+ ovary cDNA sample, less amplification of *NV_10329-RA* was observed compared to the Wol- sample, confirming down-regulation of the *NV_10329-RA* gene in *Wolbachia* infected wasps. However, in cDNA synthesized from male testis RNA, the amplification of *NV_15393-RA* could not be distinguished between the two samples.

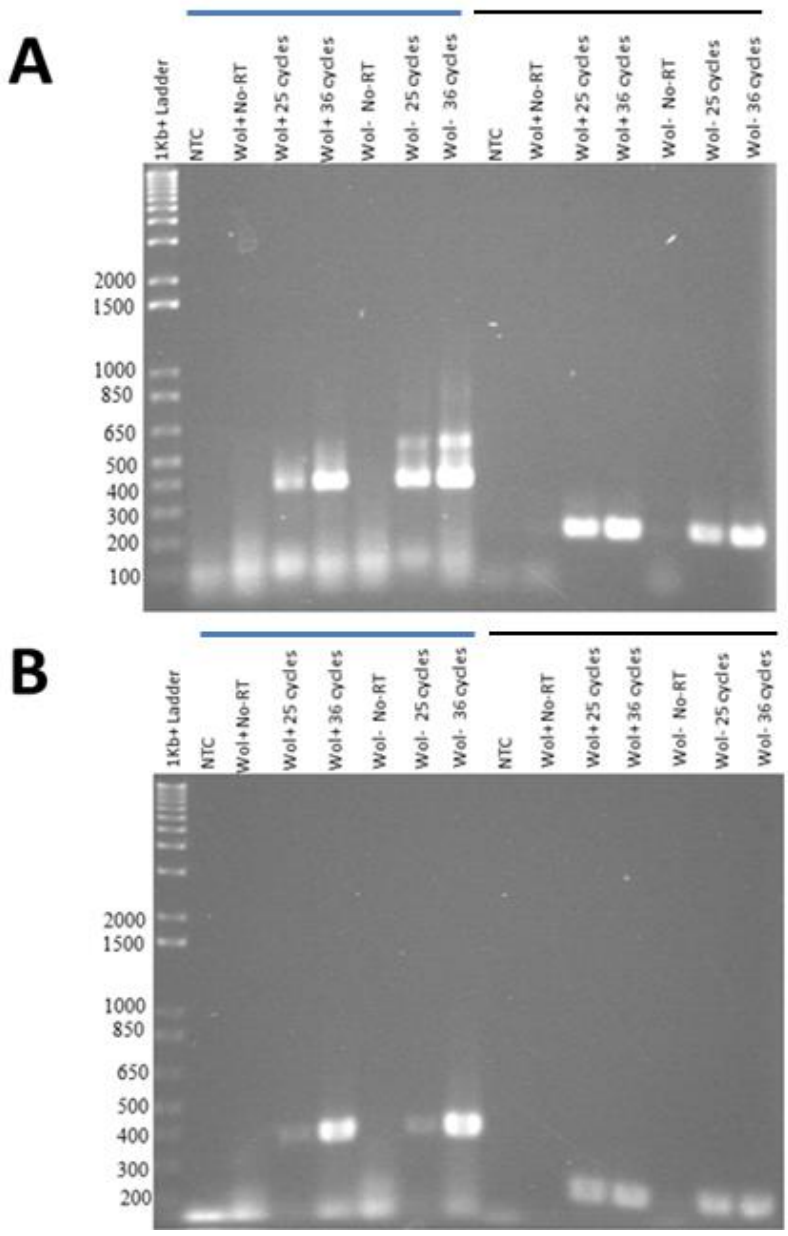


Figure 3.8. Semi qRT-PCR targeting *Nasonia* Wol+ and Wol- genes and ribosomal protein 49 (RP49) as validation for RNA-seq. The lines represent, blue: *Nasonia* gene samples and black: RP49. A: *Nasonia* gene expressed in females - *NV_10329-RA* (F2) and B: *Nasonia* gene expressed in males *NV_15393-RA* (M1). NTC = no template control, no-RT = no reverse transcriptase added. Ladder is 1kb plus (ThermoFisher Scientific) and the units shown are in bp.

3.3.6. RNA interference of down-regulated genes

Transcriptome analysis revealed 84 and 58 genes in the ovaries and testes respectively that were down-regulated by *Wolbachia* which could potentially be involved in cytoplasmic incompatibility (CI). Therefore, to study this phenomenon, the differentially expressed genes based with the smallest adjusted p-values (Table 3.5 and 3.6) were targeted for RNA interference (RNAi). In females, the top two differentially expressed genes (*NV_16517-RA* and *NV_10329-RA*) and five genes in the males (*NV_15393-RA*, *NV_18293-RA*, *NV_11725-RA*, *NV_16404-RA* and *NV_17088-RA*) were targeted for RNAi analysis (Table 3.7). The log fold change values of over 2.5-fold showed that the genes were all down-regulated by *Wolbachia* in both male and female reproductive tissues. The annotations of the down-regulated genes were predicted using the BLAST2GO package and all were predicted to be hydrolytic enzymes except for a female gene *NV_10329-RA* that had no homology to any known genes.

Table 3.7. *Nasonia* reproductive tissue genes targeted for RNAi analysis. The negative sign on the fold change shows that all the genes were down-regulated in *Wolbachia* infected tissue.

Sex	**Gene**	**Log fold change**	**Annotation**
Female	*NV_16517-RA (F1)*	-3.57755	venom acid phosphatase
	NV_10329-RA (F2)	-2.48072	No annotation
Male	*NV_15393-RA (M1)*	-3.56747	serine protease 50 precursor
	NV_18293-RA (M2)	-3.49108	trypsin-1-like
	NV_11725-RA (M3)	-3.49909	haemolymph lipopolysaccharide-binding
	NV_16404-RA (M4)	-3.01263	serine protease
	NV_17088-RA (M5)	-3.32138	maltase 1

Since the top differentially expressed genes were down-regulated, RNAi knock down was an appropriate method to assess the function of these genes in CI as RNAi is expected to mimic the results of *Wolbachia* induced CI in Wol- knock down wasps. The gene sequences to be used for RNAi were examined for RNAi off target probability and all the genes had no off-target prediction except for *NV-18293* which also targeted a gene called *NV-18294* (Appendix 2). Despite this similarity, *NV-18293* was used in the downstream

processes because *NV-18294* was not expressed in the testis tissue in this experiment. RNA was extracted from whole body Wol- wasps and subjected to reverse transcription and the resulting cDNA was used as template for double stranded RNA (dsRNA) synthesis. An agarose gel was used to confirm the successful synthesis of dsRNA of the correct size (Fig 3.9).

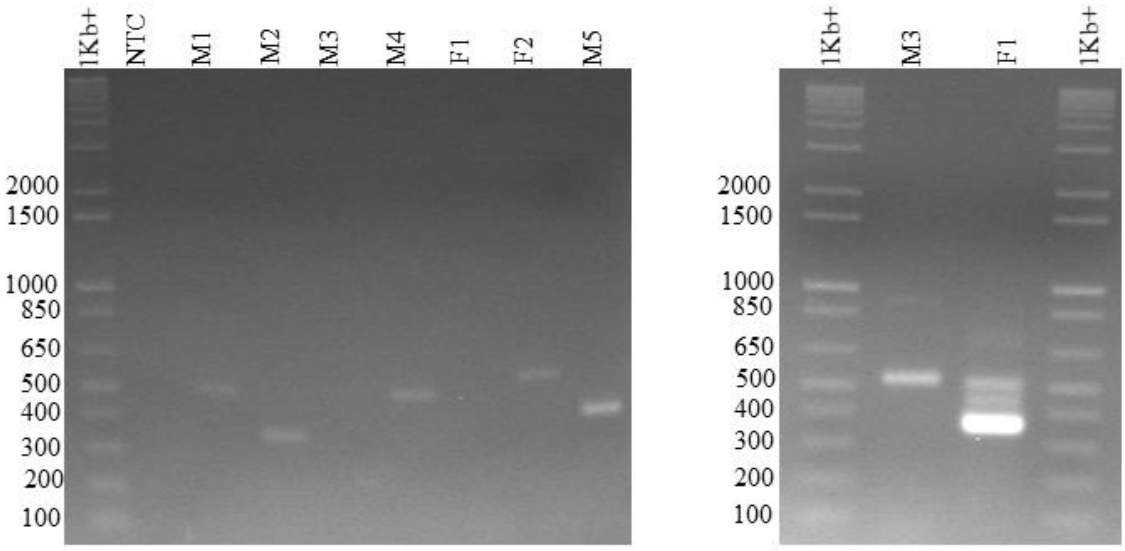


Figure 3.9. cDNA amplification of *Wolbachia* down regulated genes in *Nasonia*. The resulting purified PCR products were used for dsRNA synthesis. M1 - *NV_15393-RA*, M2 - *NV_18293-RA*, M3 - *NV_11725-RA*, M4 - *NV_16404-RA*, M5 - *NV_17088*-RA, F1 - *NV_16517-RA*, F2 - *NV_10329-RA (F2)*. The experiment was repeated for M3 and F1. NTC = no template control. Ladder is 1kb plus (ThermoFisher Scientific) and the units shown are in bp.

For RNAi injection process, fourth instar predefacation larvae were reared on PBS agar plates and dsRNA diluted with food colouring (as a visual marker) was injected posterior of the gut. To ensure that PBS agar had no negative impact on the development of *Nasonia*, fresh larvae with no injection (negative control) were also reared on a PBS agar plate and incubated at 27 ºC. The number of adult offspring were counted and compared to the number of wasps that emerged from the dsRNA injected larvae (Table 3.8). *LacZ*, which is an *E. coli* gene absent in *Nasonia*, was targeted as a control to test for the effect of injecting dsRNA in *Nasonia. Nasonia* larvae injected with dsRNA targeting LacZ developed into adults and were able to produce healthy offspring, suggesting that the dsRNA itself has no negative effects on normal development (Table 3.8). However, 31 out of 120 larvae died following *LacZ* dsRNA microinjection, likely due to accidental

penetration of the gut tissue by the needle or miss handling during transfer between PBS agar plates. The lethality rates were higher in the *LacZ* injected animals suggesting that these observations are due to technical limitations of the methodology, rather than any negative effects of knocking down itself. Significant rates of larval death following microinjection are known to be an issue with this invasive method in *Nasonia* (David Wheeler personal communication) (Table 3.8). Another control used in the RNAi was to target the eye colour gene *cinnabar* (CN) that plays a role in development of the wasp eye. Wild type eyes are grey whereas CN mutant eyes are red. Thus, successful knock down of CN gene results in adult wasps that have red eyes instead of the normal of grey eyes. As expected, injection of dsRNA targeting CN resulted in nearly all the surviving wasps (99/102) having the red eye cinnabar phenotype (Table 3.8; Fig 3.10). These results confirm that RNAi is a robust method for gene knock down in *Nasonia*.

Semi quantitative RT-PCR using the wasp whole-body cDNA was performed to confirm the knock down of CN gene and amplification was conducted for 25 and 36 cycles. At 25 cycles, no amplification of CN was observed in the CN knock down (CN_KD) cDNA as well as the wild types (WT) (Fig 3.11). At 36 cycles, no amplification was observed in CN_KD cDNA while amplification was observed in WT. To make sure that the result observed was not caused by a loss of cDNA integrity, the ribosomal protein 49 (RP49) which is a *Nasonia* housekeeping gene, was also targeted for PCR and amplification was expected for both samples. Amplification of the RP49 transcripts at both 25 and 36 cycles in both CN_KD and WT samples showed that the absence of amplification was not caused by degraded cDNA. The RT-PCR result together with the red eye phenotype confirmed the efficiency of RNAi in *Nasonia*.

Table 3.8. Summary of the RNAi control experiment of *Nasonia* offspring that developed from larvae to pupae stage after dsRNA injections.

dsRNA	Injected wasps	Survived wasps	Dead wasps	Red eyes	Grey eyes
Negative control	120	115	5	0	115
Lac Z	120	89	31	0	89
Cinnabar	120	102	18	99	3

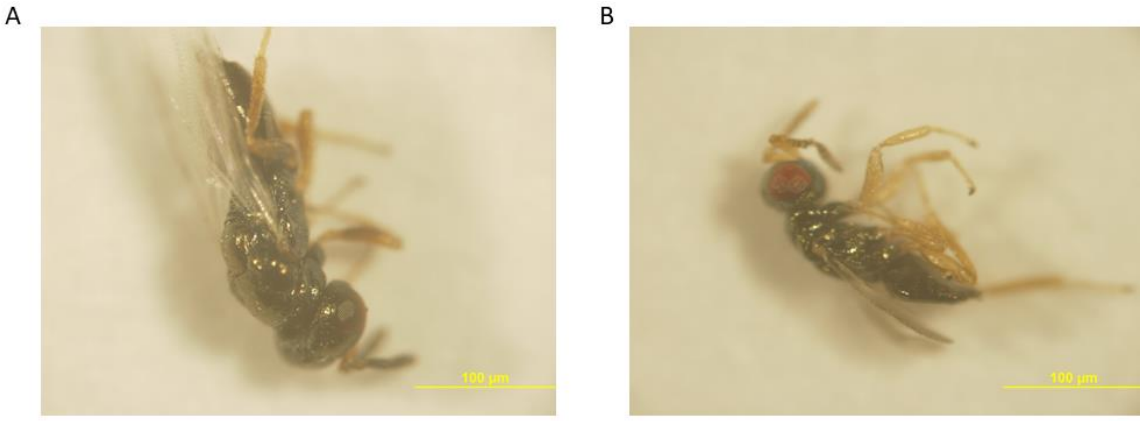


Figure 3.10. The effects of knocking down the *cinnabar* gene in *Nasonia*. A represents the wild type and B: the wasp knock down of the *cinnabar* gene.

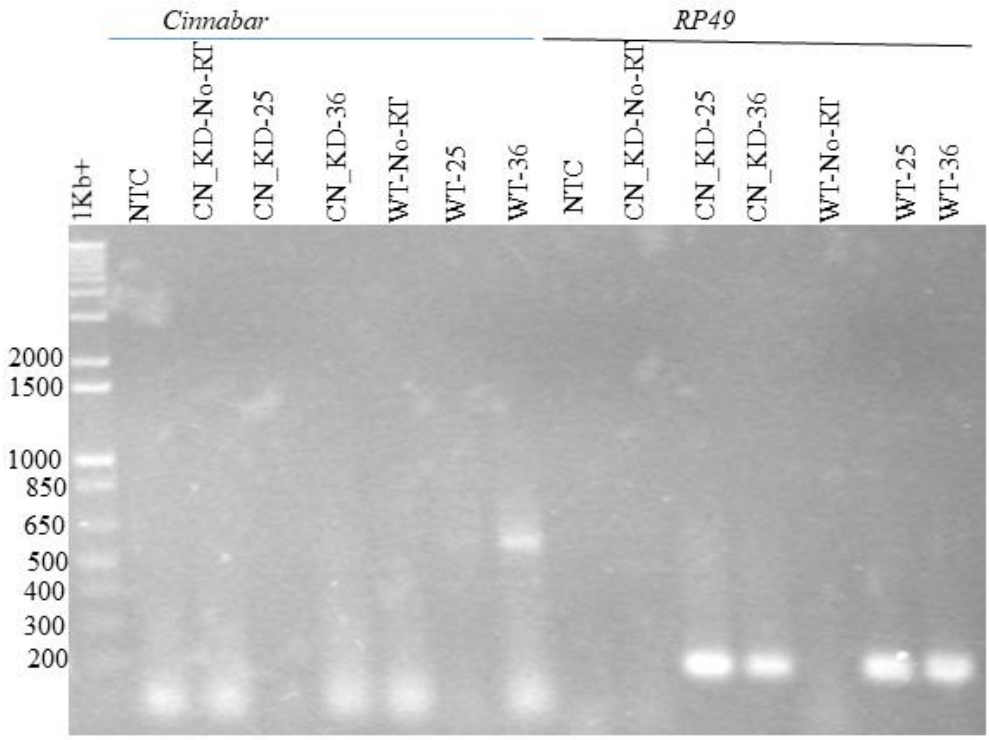


Figure 3.11. Semi-quantitative PCR diagnostic test for *cinnabar*. RNA was extracted from adult wasps, *cinnabar* kockdown (CN) and wild type (WT). RT-PCR was performed for 25 and 36 cycles. The expected product size of CN gene is 534bp. RP49 was used as a control for the cDNA used and the expected size was 196bp. NTC = no template control and No-RT means no reverse transcriptase added. Ladder is 1kb plus (ThermoFisher Scientific) and the units shown are in bp.

Wol+ females can produce offspring with either Wol+ or Wol- males, while Wol- females cannot produce offspring with Wol+ males due to CI. RNA-seq showed that F1 (*NV_16517-RA*) and F2 (*NV_10329-RA*) genes are down-regulated by the bacteria and could potentially play a role in blocking CI in infected females. Therefore, Wol- females were knocked down of the F1 and F2 genes to mimic what *Wolbachia* does to the infected females and were mated with Wol+ and Wol- males with an expectation that they would reproduce with both lines of *Nasonia* males if CI is being blocked by the RNAi knock down. Unfortunately, RNAi targeting both F1 and F2 did not block CI in the knock down wasps, suggesting that these genes may not be directly involved in the CI response (Table 3.9).

Table 3.9. Progeny counts from crosses between Wol- KD females.

Cross	Expected phenotypes	Progeny
F1 females and Wol+ males	No female offspring if CI occurs Female offspring if CI is blocked	45 males, 0 females 51 males, 0 females 47 males, 0 females
F1 females and Wol- males	Female offspring produced	44 males, 57 females 31 males, 63 females 35 males, 59 females
F2 females and Wol+ males	No female offspring if CI occurs Female offspring if CI is blocked	40 males, 0 females 51 males, 0 females 47 males, 0 females
F2 females and Wol- males	Female offspring produced	19 males, 35 females 21 males, 37 females 39 males, 42 females

Wol- males can mate with both Wol+ and Wol- females and produce offspring however; Wol- females do not produce female offspring if they mate with Wol+ males. In this experiment, Wol- males were knocked down with genes that were found to be down regulated in the bacteria infected tissue (Table 3.7). The expectation was to observe no female offspring from the reciprocal crosses of knocked down Wol- males and Wol- females; however, as all knock down wasps produced female progeny in crosses with Wol- females, this experiment also provides no direct evidence that these genes are involved in CI (Table 3.10).

Table 3.10. Progeny counts from crosses between Wol- KD males. No female offspring are produced if CI is blocked so this experiment shows that knocking down these male genes does not block CI.

Cross	Progeny
M1 males and Wol+ females	22 males, 54 females 31 males, 47 females 27 males, 61 females
M1 males and Wol- females	31 males, 49 females 21 males, 65 females 27 males, 58 females
M2 males and Wol+ females	11 males, 45 females 19 males, 52 females 19 males, 49 females
M2 males and Wol- females	18 males, 47 females 29 males, 52 females 9 males, 35 females
M3 males and Wol+ females	17 males, 51 females 23 males, 47 females 15 males, 49 females
M3 males and Wol- females	14 males, 46 females 15 males, 45 females 19 males, 41 females
M4 males and Wol+ females	19 males, 53 females 18 males, 59 females 21 males, 54 females
M4 males and Wol- females	22 males, 51 females 13 males, 47 females 14 males, 39 females
M5 males and Wol+ females	9 males, 39 females 15 males, 43 females 14 males, 48 females
M5 males and Wol- females	18 males, 49 females 19 males, 55 females 13 males, 46 females

3.4. Discussion

The aim of performing dual *Wolbachia*/*Nasonia* RNA-seq was to identify the *Nasonia* genes regulated by *Wolbachia* in the reproductive tissues and also to compare the

Wolbachia genes expressed in *Nasonia* ovary vs testis tissues. Only the *Nasonia* genes expressed in the reproductive tissues were analysed in this thesis because there was not enough sequencing for the bacterial genes.

Genes that were found to be differentially expressed in the ovary of *Nasonia* infected with *Wolbachia* (Wol+) were mainly found to have GO term annotations related to intracillary transport, centriole, transmembrane transport, catalytic activity, nucleic/DNA binding, membrane and cell adhesion. The observation of genes involved in cell division is interesting given the known effects *Wolbachia* has on cell division in the germline. Confocal microscope imaging of *Wolbachia* infected ovaries in *Drosophila* showed that *Wolbachia* is present in female germline stem cells [209]. During stem cell mitosis, *Wolbachia* is partitioned between the renewing cell and the cytoblast to allow transfer to the daughter cell as well as to be retained in the paternal cell [210]. The mechanism underlying this partitioning mechanism remains unknown and moreover 50% of the genes differentially expressed in this study have no functional annotation perhaps (unsurprisingly) suggesting that the pathways that control these events might be novel. In addition, it has been shown that manipulating the microtubule network disrupts the partitioning of *Wolbachia* suggesting that the bacteria is anchored to this network by host factors [209]. This previous experiment showed the need for *Wolbachia* to manipulate the structures involved in cell division to ensure transmission as well as survival. The RNA-seq experiment in this chapter showed that genes involved in cell division were regulated by the bacteria thereby complementing the observation in the previous studies.

Unlike in ovary tissues, genes involved in cell division were not predominantly differentially expressed in the testis tissue, rather genes with catalytic roles were affected. As *Wolbachia* is not transferred through the male germline, this could be a reason why manipulating cell division is not important for the CI mechanism. In addition, *Wolbachia* have been found to be excluded from the mature sperm, which ensures that no transmission occurs through the male germ line [211]. In *Nasonia* testis, *Wolbachia* is known to modify the sperm in a way that only allow successful fertilization if the same *Wolbachia* strain is present in the egg [44]. Based on GO term annotations, genes involved in carbohydrate metabolic processes are downregulated by *Wolbachia* in the *Nasonia*

testis. Previous studies showed that sperm uses carbohydrate-mediated signalling which triggers binding and exocytosis of the acrosomal contents when it contacts the eggs extracellular coat [212]. Therefore, the down regulation of genes involved in carbohydrate-mediated signalling reveals the possibility of improper binding between the sperm and egg or the failure of acrosomal reaction in CI.

Also found to be down regulated in this study were a number of enzymes such as serine proteases, hydrolases, esterases, sulphatases and phosphatases. These enzymes have been found to form part of the acrosome contents and their roles are mainly to assist the sperm with binding and penetration [212]. Proteases are released at the site of sperm-egg binding to regulate the penetration of acrosomal contents therefore the down regulation observed further confirms the possibility of *Wolbachia* having a negative activity on sperm-egg binding and acrosomal reaction. Regulation of catalytic function related genes is a strategy to weaken the sperm and make it dependent on the egg for rescue.

Previous studies have shown that aside from the classical manipulation of the host's reproduction system in CI, *Wolbachia* influences gene expression in *Drosophila*, such that it impacts physiology and immunity of the host fly [213,214]. Principal component analysis (PCA) which is a clustering method to group related data into linearly uncorrelated variables could not significantly distinguish between Wol+ and Wol- ovary samples. Thus, the overall changes in gene expression differences were not enough to separate these Wol+ and Wol- ovary tissue samples. In contrast, for the testis data the PCA plot reveals that changes in overall gene expression can be observed when *Wolbachia* is found in the tissue, suggesting that the bacterial induced change might be more dramatic in sperm cells. *Wolbachia* has been found to interact with the host's nuclei manipulating the development of the sperm in the testis [215] indicating that apart from catalytic genes that have been observed to be regulated by the bacteria, genes involved in development are likely to be regulated increasing the overall expression differences between the infected and non-infected tissues. A suggestion for this likely gene expression differences in the testis tissue is that that there is a smaller background of genes being expressed in testis therefore the *Wolbachia* signal is stronger.

Semi qRT-PCR approach could not clearly validate the RNA-seq result that *NV_15393* gene was knocked down in Wol+ testis tissue. This is most likely caused by the fact qRT-PCR is not sensitive enough compared to the RNA-seq approach. The RNA-seq-RNAi approach to study gene function in *Nasonia* has been used on calreticulin to show that the gene is involved in suppressing the immunity system of the wasps' hosts [145]. In this study, the approach was also applied on the top differentially expressed genes in the male and female reproductive tissues. The observation on the RNA-seq data was that the genes with the lowest adjusted p-value were down-regulated therefore using the RNAi approach, *Wolbachia* influence on *Nasonia* genes was mimicked. However, using this approach, the CI manipulation could not be rescued, indicating that the knocking down of one gene was not enough to pinpoint the CI causative genes.

3.5. Conclusions

Using the RNA-seq technology, a comprehensive dataset of Wol+ and Wol- expression information was generated, and genes likely to be regulated by *Wolbachia* in *Nasonia* to cause cytoplasmic incompatibility were shown. In the ovaries, genes that are involved in cell division were regulated while in the testes, genes with catalytic responsibilities were manipulated. Overall, less than 100 out of the 33,000+ genes in *Nasonia* were significantly differentially expressed, suggesting that *Wolbachia* has very specific effects on the reproductive tissue at the transcriptome level. This list is important in providing an unbiased selection targeting genes for CI studies in *Nasonia*. Notably, this analysis of knocking down genes individually to test CI leads to a speculation that a network of genes rather than an individual gene causes this manipulation.

4. *WOLBACHIA-NASONIA* LATERAL GENE TRANSFER

4.1. Introduction

Lateral gene transfer (LGT), the movement of genes between distantly related organisms, facilitates the acquisition of novel gene functions. Benefits such as evolution of antibiotic resistance and improving metabolic pathways constitutes the consequences of LGTs. Most of the LGTs described occur within a single domain of life (bacteria to bacteria) however, there are examples of transfers involving multicellular eukaryotes have been found [71,216]. An endosymbiont *Wolbachia pipientis* infects at least 20% of insect species and is present in the gametes therefore providing a conducive circumstance for LGTs with its hosts [217]. *Wolbachia*-host transfer have been described in the bean beetle, filarial nematodes and *Nasonia* [85,86,218].

In this study, additional *Wolbachia* sequences were identified to be putative LGTs in the wasp *Nasonia*. A widespread scan for *Wolbachia* genes in the *Nasonia* genome showed nine putative LGTs including three already described LGTs. However, only one gene (*gene1092*) was confirmed to be part of the current assembled *Nasonia* genome while the other five still need further analysis.

4.2. Methods

4.2.1. A Search for LGT events in the *Nasonia* genome and trace repositories

Nasonia spp. (*N. vitripennis*: GCA_000002325.2, *N. giraulti*: GCA_000004775.1 and *N. longicornis*: GCA_000004795.1) and *Wolbachia* (GCA_000204545.1) genomes and gene model files were downloaded from NCBI while whole genome shotgun sequencing reads (wgs) were downloaded from the NCBI trace archive (https://www.ncbi.nlm.nih.gov/Traces/trace.cgi?view=registered_spieces). Based on the *Wolbachia* gene models, gene sequences were extracted from the genome and used as query for BLASTn search against the databases made from *Nasonia* spp. genomes and whole shotgun sequence reads. Importantly, the *Nasonia* spp. genomes were assembled from antibiotic-cured insects to avoid the possibility of false positive LGT identifications resulting from *Wolbachia* contamination of sequenced insect DNA. The putative LGT

events were identified based on high levels of sequence homology between *Wolbachia* sequences and regions in the *N. vitripennis* genome.

4.2.2. RNA-seq analysis and data acquisition

Nasonia transcriptomic data from three independent sources were analysed to search for *Wolbachia* gene transcription. The datasets analysed were the CI-based sets prepared in Chapter 3, as well as two others independently generated datasets downloaded from the NCBI SRA archive (https://www.ncbi.nlm.nih.gov/sra/) (dataset 1 (SRR1566022, SRR1566023 and SRR1566024) [219] and dataset 2 (SRR988302 and SRR988303) [220]. The protocol of the RNA-seq dataset analysis was described in section 3.2.3 where the overall quality of the raw RNA-seq data was assessed using the FastQC v.0.11.5 package [221]. Pre-processing and trimming of the reads were performed using fastq-mcf and SolexaQA++ v. 3.1.4 [202], respectively. The high quality filtered short read data was mapped to the *Nasonia spp.* genomes using Tophat2 v.2.1.1 [203] with default settings, except for the library type option which was set to fr-firststrand. Gene based read counts were generated using HTSeq-count v.0.6.1p [205] using the default settings, except for the strand-specific assay which was set to reverse using the flag –s reverse. The DESEQ2 [206] package was then used to normalise the read counts.

4.2.3. cDNA synthesis, amplification of *Wolbachia* genes and Comparative genomics

All the equipment used and benches were cleaned with RNA-Zap (Ambion). Complementary DNA was generated from 1μg of RNA extracted from dissected ovary tissues using the transcriptor first strand cDNA synthesis kit (Roche), with anchored-oligo (dT) primers as described by the manufacturer. cDNA and genomic DNA was used for PCR to detect *Wolbachia* genes using the primers described in Table 4.1

Table 4.1. Primer sequences used for targeting *Wolbachia* genes.

Gene	Primers
gene1092	F: AGGAAGCGACAACAGGTAAA
	R: GCTGTTGCACTTCTAGTTCG
gene1120	F: TCTTACTTAGCGGAAATGCTTG
	R: GCAATATGCGCACAACATCC
gene710	F: CGCGAAAAACCTTACCACTT
	R: CGTACAAAGCCTAAGGAGGT
gene786	F: CATATCGACGACACCGTTTG
	R: ACCCAACTCACGTAACACTT

The *Wolbachia* gene sequences expressed in Wol- wasps were extracted from the *Wolbachia* B genomic scaffolds (GCA_000204545.1) using a python script that extracts the gene coordinates from the gene annotation file then uses the corresponding coordinates to extract the sequences from the *Wolbachia* genome. The resulting potentially transferred *Wolbachia* gene sequences were then used as queries for BLASTn against the *Nasonia sp*. genomic scaffolds at NCBI (GCA_000002325.2) (http://www.ncbi.nlm.nih.gov/) and *N. vitripennis* raw assembly data to determine their loci in the *Nasonia* genome.

4.2.4. Bioinformatics characterization of *Wolbachia* genes

Open reading frames were predicted using the ExPASy bioinformatics resource portal webpage tool (https://web.expasy.org/translate/) while homologous sequences were searched using NCBI BLAST tools (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The DNA paired reads dataset (SRR846562) was downloaded from the trace archives and mapped to the *Nasonia* genome using the Bowtie2 [222] package. The integrative genomics viewer (IGV) package [223] was used to visualize the reads-genome alignments, and phylogenetic trees were calculated using Mega v.7.0.20 [223].

4.3. Results

4.3.1. LGT events detection

BLASTn searches of the *Wolbachia* gene models against the databases of *Nasonia* sister species genomes and the whole shotgun sequencing (wgs) reads were used to identify LGT events. The BLASTn approach was preferred for this search because it reduces the likelihood of false hits between non-LGT genes that contain very ancient domains shared between bacteria and insects. DNA based searches reduce the risk of these false positive hits because the genetic code degeneracy allows DNA to evolve much more rapidly than the proteins these regions encode. Since the wgs reads were extracted from antibiotic-cured *Nasonia* [28], any *Wolbachia* sequence found in these reads are unlikely to be a result of *Wolbachia* contamination of the sequenced DNA. Using this approach 9 putative *Wolbachia* to *Nasonia* LGTs could be identified (Table 4.2). The search showed that three putative LGT events (*gene60, gene1092 and gene1120*) were common in the sister species suggesting that the LGT event occurred in the common ancestor of these species. However, there were also species-specific LGT sequences identified, four in *N. vitripennis* and one in *N. giraulti* (Table 4.2). Three of the genes identified (*gene283, gene786* and *gene617*) had already been described as LGTs [224].

Table 4.2. *Wolbachia* inserts in *Nasonia spp*. identified through a BLASTn search against the trace repository database. Insert size refers to the length of *Wolbachia* gene sequence inserted into the *Nasonia* fragments; identity refers to the sequence similarity between *Nasonia* and *Wolbachia* sequences while gap is the number of gaps that were introduced in the alignment to increase matching.

Organism	*Wolbachia* gene	Insert size (bp)	Identity (%)	Gap (bp)
N. vitripennis	*gene22*#	118	83.1	4
	gene60±	56	94.6	0
	*gene283**#	122	100	0
	*gene617**#	102	84.3	1
	gene710#	697	80.2	26
	*gene786**	874	82.2	35
	gene1092±	266	98.9	0
	gene1120±	254	99.6	0
N. giraulti	*gene60*±	56	94.6	0
	*WVB_RS06960**#	446	91.0	0
	gene1092±	257	94.2	0
	gene1120±	254	99.6	0
N.longicornis	*gene60*±	77	94.8	1
	*gene786**	43	97.7	0
	gene1092±	257	94.1	0
	gene1120±	254	99.6	0

*Indicates LGT events that have already been described [224]
±Indicates LGT events that are common to the 3 sister species
#Indicates LGT events that are unique to the organism

4.3.2. Expression of *Wolbachia* genes in *Nasonia* Wol- line

To determine whether the nine putative LGTs described in Table 4.2 were expressed in *Nasonia*, the Wol- ovary and testis RNA-seq reads (Described in Chapter 3) were mapped to the *Wolbachia* B genome. RNA-seq reads from the *Nasonia* ovary successfully mapped

to 18 *Wolbachia* genes, while reads from the testis samples mapped to 61 *Wolbachia* loci (Appendix 3, Table A3.1, Table A3.2). However, only five genes showed evidence of expression (based on read mapping) across all replicate samples. Therefore to show that this result was not only specific to the RNA-seq data obtained in this study, two independent transcriptomic datasets were downloaded from NCBI SRA archive. The first (SRR1566022, SRR1566023 and SRR1566024) [219] was from a study of gene expression and epigenetic sexual dimorphism in two jewel wasp species *Nasonia vitripennis* and *Nasonia giraulti*, while the second (SRR988302 and SRR988303) was from a study of transcriptome sequencing of infected and uninfected control of *Nasonia* vitripennis samples [220]. As before the resulting reads were mapped to the *Wolbachia* B genome and these results confirmed the expression of five *Wolbachia* genes in Wol-*Nasonia* RNA-seq generated in this study (Appendix 3, Table A3.3 and Table A3.4). Interestingly, in these two data sets, mapping of reads to the genome was only to the five genes. *Gene1076* was eliminated from the study because pairwise alignment showed that it was identical to *gene710* (Fig 4.1).

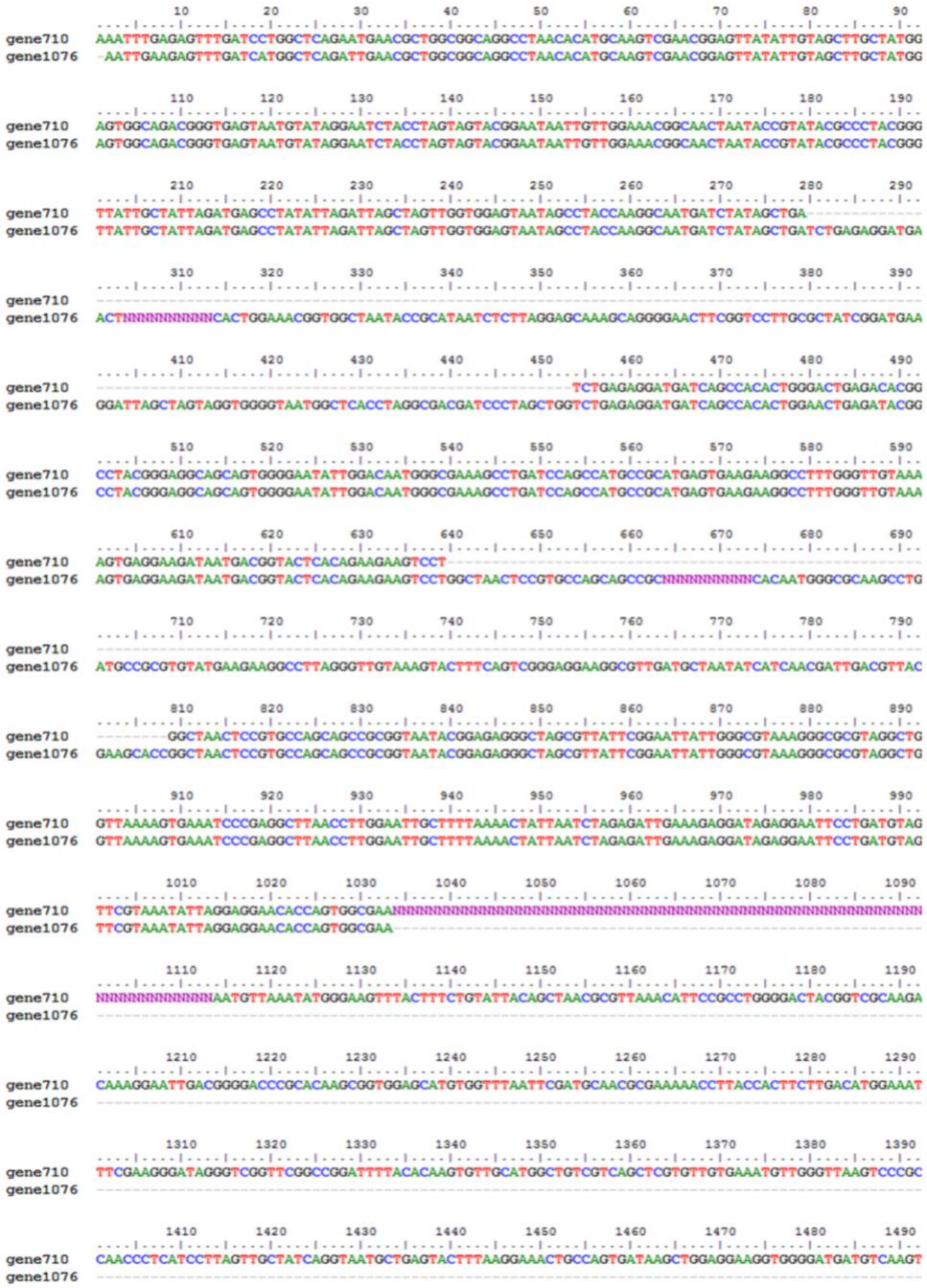


Figure 4.1: Pairwise alignment of *gene710* and *gene1076*. The alignment shows 100% identity in the aligned regions sugesting that *gene1076* is part of *gene710*.

4.3.3. PCR screen for LGTs and comparative genomics

To confirm the presence of the *Wolbachia* LGTs in *Nasonia* genome, PCR was performed using genomic DNA. Additionally, RT-PCR was performed to confirm expression of the suggested LGT. As expected, no amplicons were obtained in the PCR and RT-PCR negative controls (Fig 4.2). The absence of amplification in no-RT samples ruled out genomic DNA contamination. *Gene1092* and *gene1120* were successfully amplified from both Wol- cDNA and gDNA indicating both the presence and expression of the genes in the Wol- *Nasonia*. The products were sequenced and confirmed the correct identity of the products. Together, these data plus the RNA-seq results suggest lateral gene transfer has occurred and that these genes are being actively expressed. However, *gene710* and *gene786* were only present and expressed in Wol+ wasps, which was not consistent with the BLASTn searches and transcriptomic analysis that showed the genes to be present in Wol- samples.

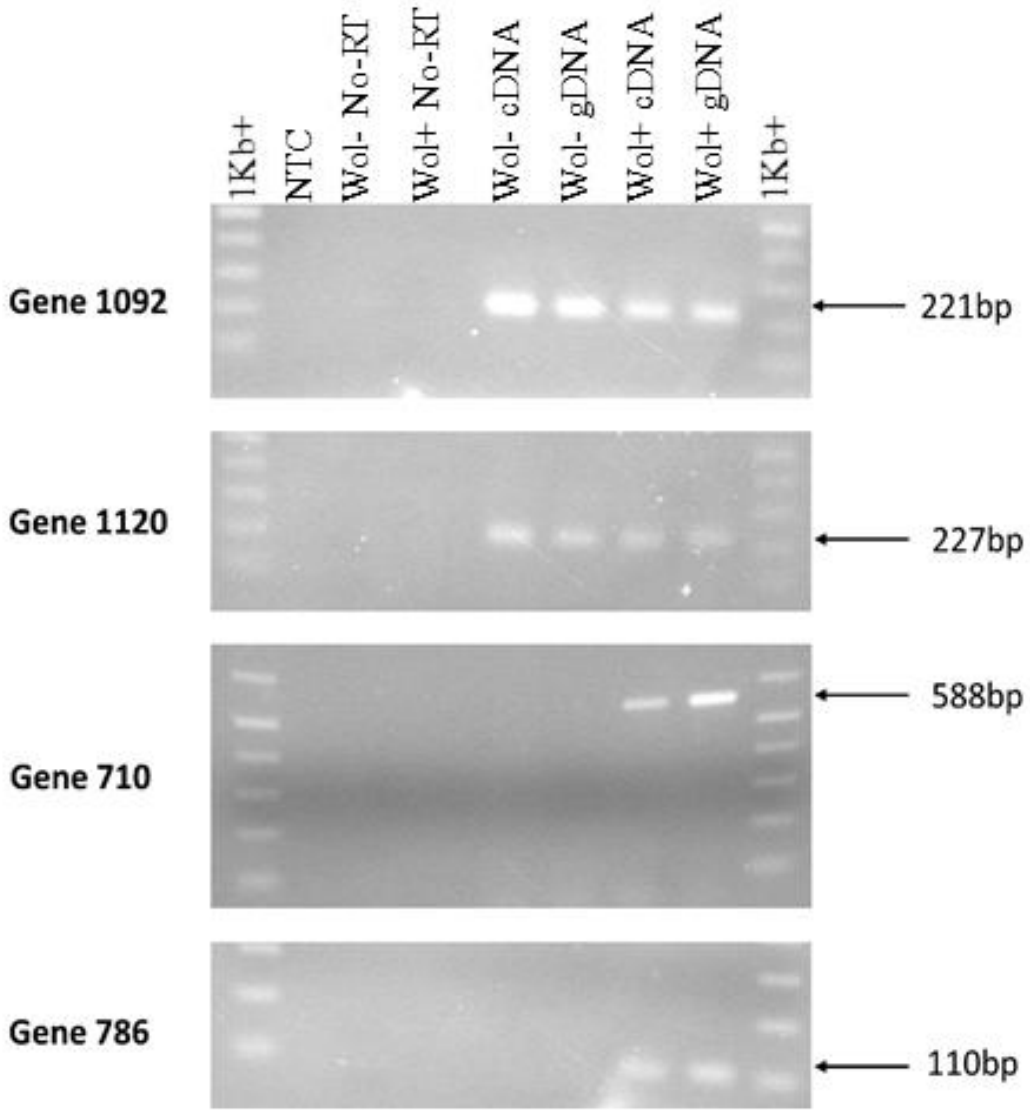


Figure 4.2. Genomic and cDNA amplification of four putative LGT genes in Wol+ and Wol- *Nasonia* samples. Amplification of Wol+ gDNA showed that all four genes were present in the genome while only *gene1092* and *gene1120* were present in Wol-. NTC – no template control, No-RT – reverse transcriptase was not added in the reaction mixture. Ladder is 1kb plus (ThermoFisher Scientific) and the units shown are in bp.

4.3.4. *Gene1092* was inserted into the coding region of the *Nasonia* gene *Nasvi2EG014694*

The putative LGT sequence *gene1092* contained no open reading frame suggesting that it was a pseudogene in *Wolbachia* and BLASTn searches using this insert DNA sequence returned no hits at an e-value cut-off of 1e-05. The BLASTn analysis showed that

g*ene1092* is located on scaffold NC_015870.2 of *N. vitripennis* (Table 4.3) and I was able to show that this flanking *Nasonia* DNA sequence restored the open reading frame of this gene. To verify this observation, the RNA-seq reads that mapped to the scaffold NC_015870.2 at the region of the putative LGT were visualized and this showed that RNA-seq reads do indeed span the newly predicted open reading frame (g*ene1092* and the flanking *Nasonia* DNA). Further examination of this region revealed that the g*ene1092* insert was expressed as the 5' part of the *Nasonia* gene *NASVI2EG014694* (Fig 4.4A). To rule out the possibility of a hybrid assembly artefact in this region, paired genomic reads from the SRA archive (SRR846562, SRR8464346 and SRR8456518) were mapped to the *Nasvi2EG014694* region in *Nasonia* scaffold NC_015870.2 (Fig 4.4B). The continual mapping of reads across the NC_015870.2 scaffold suggests that the insertion of gene *gene1092* into *Nasvi2EG014694* was not a genome assembly artefact.

Table 4.3. *Wolbachia* inserts in hosts found by BLASTn. Identity refers to the sequence similarity between *Nasonia* and *Wolbachia* sequences while gap is the number of gaps that were introduced in the alignment to increase matching.

Host	Position of insertion	*Wolbachia* gene	Identity (%)	Gap (bp)	e-value
N. vitripennis	NC_015870	*gene1092*	91.13	1	3.00E-110
N. longicornis	GL277994.1	*gene1092*	94.54	0	2.00E-127
N. giraulti	GL275759.1	*gene1092*	93.13	0	1.00E-119

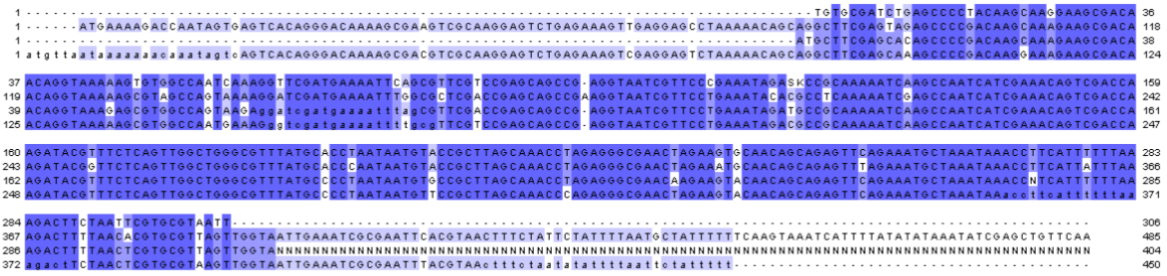


Figure 4.3. Multiple sequence alignment of *gene1092* and the insert regions in the *Nasonia* sister species. The sequences are arranged from top to bottom in the following order: *gene1092*, *N. vitripennis* insert, *N. longicornis* insert and *N. girault*i insert. A high level of conservation shows that the *Wolbachia* gene is incorporated into the genomes of the *Nasonia* sister species.

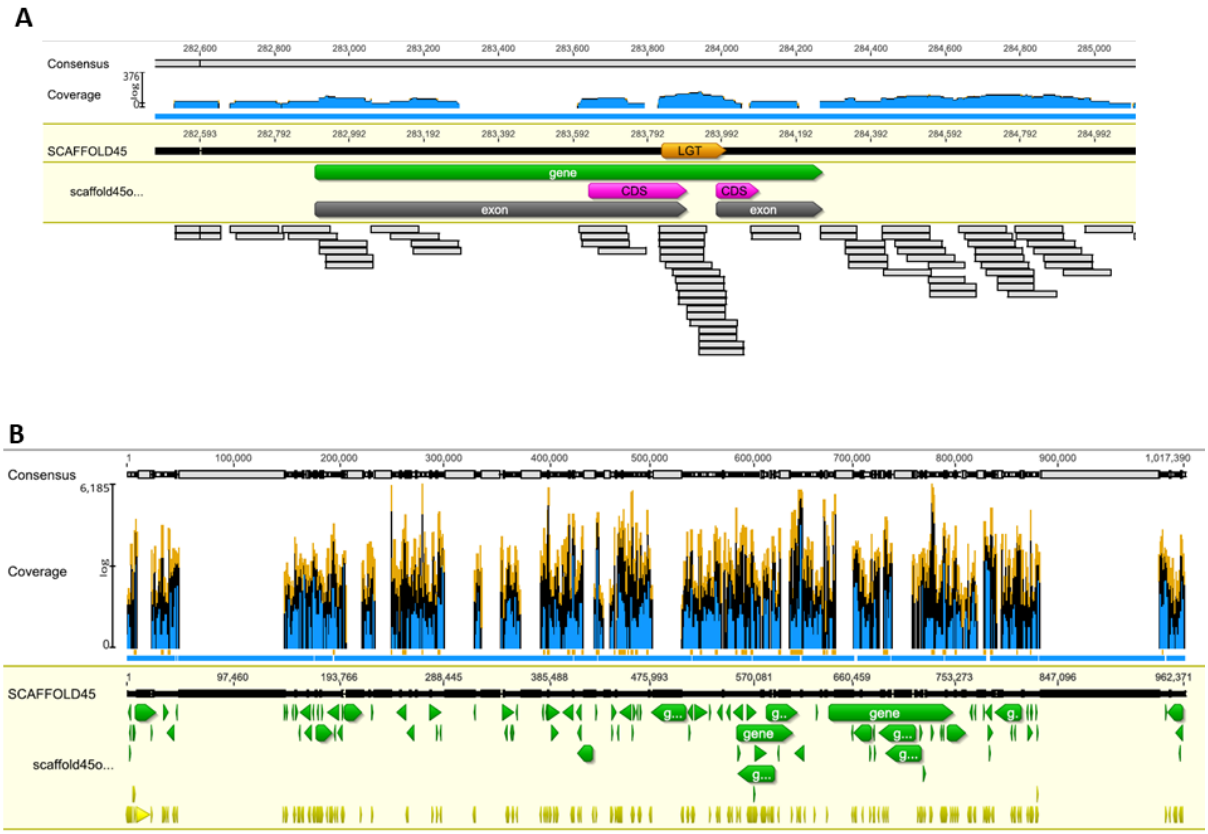


Figure 4.4. Visualization of the alignments between the sequencing reads and *Nasonia* genome. A: RNA-seq reads mapped (grey rectangular blocks) to the region of *gene1092* insert (LGT - orange block) showing the insert is part of the coding region of a *Nasonia* gene (gene - green block). **B**: DNA-seq reads mapped to *Nasonia* scaffold45 that contains the insert showing the neighbouring genes (green blocks) well aligned with reads (yellow bars), this point out the accuracy of the assembly on the scaffold.

4.3.5. Comparative genomics showed *gene1120* to be a *Nasonia* gene

PCR based screens provided evidence for gene *gene1120* fragment being present in the *Nasonia* genome (Fig 4.2). Using the *gene1120* fragment sequence, BLASTx gave hits to eukaryotic hypothetical protein sequences from *Trachymyrmex cornetzi*, *Carpathonesticus biroi* and *Lasius niger* as top hits all having 100% query coverage and sequence identities of 58%, 59% and 62%, respectively (Table 4.4). A BLASTn search of the *Nasonia* genome failed to identify any sequences homologous to the *gene1120* fragment. One possible explanation for this result is that raw reads containing this fragment were labelled as 'contamination' and were subsequently removed before the genome assembly process was carried out. To explore this possibility, the NCBI non-redundant database was searched using the raw Sanger read (937 bp) as the query. It was expected that any homology to eukaryotic sequences would suggest that the read is a *Nasonia* sequence that had been left out in the assembly. A BLASTn search resulted in all the hits arising from a eukaryotic origin therefore suggested this *gene1120* fragment is of eukaryotic in origin and unlikely to represent an LGT between *Nasonia* and *Wolbachia* (Appendix 3, Table A3.5). Additionally, phylogenetic analysis also confirmed that the *gene1120* is more closely related to eukaryote sequences than it is to bacterial loci, once again suggesting that it is not an LGT (Fig 4.5).

Table 4.4. BLASTx output against the nr database using gene1120 sequence as the query.

Organism	**Accession number**	**Query cover (%)**	**Identity (%)**	**e-value**
Trachymyrmex cornetzi	KYN27960.1	100	58	3.00E-33
Carpathonesticus biroi	EZA545576.1	100	59	5.00E-33
Lasius niger	KMQ87194.1	100	62	1.00E-30
Cyphomymex costatus	KYM99014.1	100	58	4.00E-30
Camponotus floridanus	XP_011251810.1	100	56	5.00E-28
Microplitis demolitor	XP_008553228.2	100	55	9.00E-25

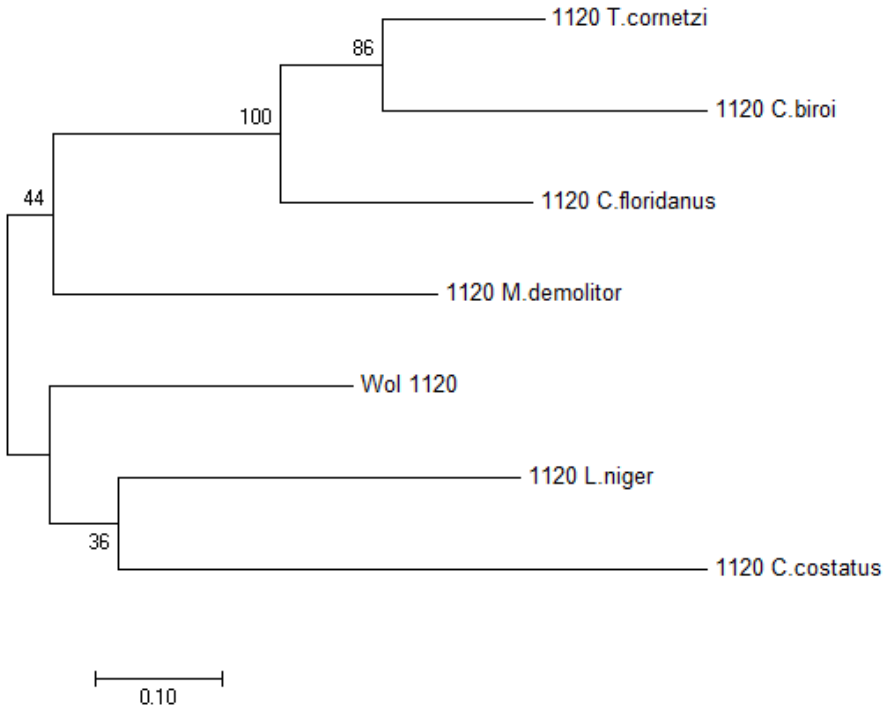


Figure 4.5. Maximum Likelihood Phylogenetic analysis of *gene1120* conducted using MEGA7 [225]**.** The evolutionary history was inferred by using the Maximum Likelihood method based on the JTT matrix-based model [226]. The numbers are the bootstrap values estimating the reliability of the branching. The higher the bootstrap value the more reliable the branching estimation is. 1120 represent *gene1120* sequence from the respective species.

4.4. Discussion

LGTs have been found to be a common tactic used by bacteria to increase their fitness [70], with examples being the evolution of antibiotic resistance, pathogenicity and metabolic pathways in bacteria [227]. *Wolbachia* is maternally inherited and found in developing gametes, therefore circumstances ideal for transferring its genes to the host are provided [81]. LGT has been described in the human parasitic nematode *Onchocerca volvus*, where genome sequencing results have suggested that the insertion of *Wolbachia* genes allowed this species to transition from bovine to human hosts [218]. In *Wolbachia*-cured *Nasonia* wasps, six *Wolbachia* inserts were verified by PCR and sequencing and these inserts (LGTs) showed a nucleotide sequence similarity above 96% when compared to the *Wolbachia* sequences [224]. The results presented in this chapter provided evidence

for LGT events having occurred in the wasp *Nasonia,* with *Wolbachia* acting as the most likely source of this genetic material.

A *Wolbachia* gene sequence (*gene1092*) was shown to be incorporated in the *Nasonia sp.* genome in this study. The data suggests that the LGT event involving *gene1092* occurred before the separation of the *Nasonia* sister species. The high level of conservation in this region among the sister species suggests that this gene could be under purifying selection and thus be functional. A BLASTn search in the nucleotide collection (nr/nt) database could not identify homologous sequences to *gene1092* in the entire NCBI database, however RNA-seq and RT-PCR confirmed that this gene was expressed in both the testis and ovaries of *Nasonia*. In addition, the molecular evidence presented here suggests that the *gene1092 Wolbachia* fragment is incorporated in the coding region of a *Nasonia* gene, *NASVI2EG014694* and transcribed as part of one transcript. It has been shown that *Wolbachia* manipulates the reproductive tissues of its host [42] therefore the transfer of its functional genes reinforces reproductive isolation. In addition, based on the normalized read counts, *gene1092* is down regulated in both the testis and ovaries of Wol+ compared to Wol- suggesting the genes' potential involvement in reproductive isolation.

Because *Wolbachia* is an intracellular parasite, it cannot be cultured independently, which poses a major challenge in assembling the genome because its DNA is isolated from a heterogeneous mixture of eukaryotic, bacterial and viral DNA [228], thereby increasing the chances of assembly errors. This is the likely case with *Wolbachia gene1120* that has been shown to be present in the *Nasonia* genome through PCR screening. Comparative genomics analysis of the *Nasonia* raw assembly data showed Sanger reads that had been excluded from the final assembly, but which have *Gene1120* incorporated in them. Homology-based searches and phylogenetic analysis provided strong evidence that the gene is a eukaryotic gene and therefore has been mistakenly included in the *Wolbachia* genome. Transcriptomic analysis showed that *gene1120* is downregulated in Wol+ wasps and this was also observed in an independent transcriptomic analysis of *N. vitripennis* [220].

4.5. Conclusion

The availability of next generation sequencing data has allowed us to identify many examples of LGTs between intracellular parasites and their hosts. This study shows that *gene1092* from *Wolbachia* has been inserted into the *Nasonia* gene *NASVI2EG014694* and became part of the open reading frame. The *gene1092* LGT has also been identified in the sister species of *N. vitripennis* indicating that the transfer event occurred before separation of this group from a common ancestor. Phylogenetic analysis and BLAST search predicted that the *Wolbachia gene1120* is a eukaryotic sequence. Considering that the DNA sequences used for the *Wolbachia* genome assembly were extracted from *Nasonia*, chances are that the *Nasonia* sequence was mistakenly added into the *Wolbachia* genome. Despite RNA-seq showing that *gene710* and *gene786* are expressed in the Wol-*Nasonia* reproductive tissues, semi qRT-PCR could not validate this observation. Therefore, a more sensitive approach qRT-PCR could be used instead to confirm these LGTs.

5. FUNCTIONAL CHARACTERIZATION OF *N. VITRIPENNIS* VENOM GENES

5.1. Introduction

Nasonia vitripennis (Nasonia) venom proteins once injected into the host larvae, collectively manipulate it in ways that support the growth of *Nasonia* offspring. These manipulations include developmental arrest, suppression of the host immunity system and regulation of its metabolites [2,143,229,230]. However, the functions of the individual venom proteins are still yet to be characterized. Challenges to predict the function of the venom proteins have been caused mainly by functional redundancy as well as the fact that over 20% of the venom genes have no sequence homology to any of the known genes. Previous studies ranked the venom genes on the number of copies of each transcript in the venom reservoir [197]. On this basis, the protein products of the top six most highly expressed genes in the venom gland were chosen for biochemical and possible structural characterization [197]. Other studies using recombinant techniques on the 7th most highly expressed gene (Kazal type serine protease inhibitor-like venom protein 2) suggested the venom protein has a role in modifying host's immunity [148]. Furthermore, studies showed that calreticulin reduces melanisation on the stung host [148]. In this study bioinformatic techniques were used to predict and characterize the functions of the venom proteins.

The venom proteins targeted in this chapter are, except for *serine protease 97* (SP97), novel and have no homology to any known proteins in the NCBI or Swiss-Prot databases. The other five venom proteins that are unknown studied in this chapter are *venom Y* (VY), *venom Z* (VZ), *venom X* (VX), *venom G* (VG) and *venom Q* (VQ).

5.2. Methods

5.2.1. Identification and *in silico* analysis of the venom genes

The *Nasonia* venom genes characterized were selected based on the expression levels in the venom gland as ranked by Sim and Wheeler [231]. The top six most highly expressed genes in the venom gland were selected for characterization (Table 5.1) and their translated protein sequence subjected to protein-protein BLAST and HMMER analyses to assess if there was any sequence homology with other proteins in the non-redundant protein, Swiss-Prot and Pfam databases [232,233]. SignalP v4.1 [234] default settings that use both neural network and hidden Markov model methods were used to predict any signal peptides in the venom protein sequences. The molecular weight and isoelectric points of the hypothetical mature proteins were calculated using the ExPASy ProtParam web tool [235]. The Protein Homology/analogY Recognition Engine v2.0 (Phyre2) webserver was used to predict the secondary structures of the venom proteins using the intensive mode [236].

Table 5.1. List of the venom genes selected for characterization.

Classification	**Gene Loci (length bp)**	**mRNA ID (length bp)**	**Peptide (length aa)**
Venom Y (VY)	LOC100120682 (712)	NM_001161696.1 (466)	NP_001155168.1 (103)
Venom G (VG)	hmm716344 (914)	XM_008205495.2 (460)	XP_008203717.1 (102)
Venom Z (VZ)	LOC100123138 (799)	NM_001161697.1 (584)	NP_001155169.1 (154)
Venom Q (VQ)	LOC100123037 (864)	NM_001161689.1 (665)	NP_001155161.1 (164)
Venom X (VX)	LOC100123123 (708)	NM_001161695.1 (475)	NP_001155167.1 (112)
Serine protease 97 (SP97)	SP97 (5442)	NM_001161570.1 (1244)	NP_001155042.1 (293)

5.2.2. Double stranded RNA generation for RNA interference

Extraction of RNA from whole *Nasonia* females was carried out using a two-step protocol involving the use of Trizol and Qiagen RNeasy mini kit. First, five females were homogenized inside a 1.5 mL Eppendorf tube using a pestle in 150µL of Trizol (Ambion) before the addition of extra Trizol to make a total of 1mL. 200µL of chloroform was then added to the sample, which was then vortexed and incubated for three minutes at room temperature. The sample was then centrifuged for 15 minutes at 8000 X g and the aqueous layer removed and subjected to purification using the RNeasy mini kit (Qiagen) following the manufacturer's specifications. The concentration of RNA obtained was measured using the Nanodrop and the remaining sample was stored at -80 C°. cDNA was then synthesized using the QuantiTect Reverse Transcription Kit (Qiagen) following the manufacturer's instructions.

The primers tabulated below (Table 5.,2) were used to amplify the gene fragments, from the selected *Nasonia* genes and the control sample, *E. coli lacZ* to be used for RNAi. Prior to confirming the primers for PCR, the Wasp Atlas (http://waspatlas.com/tools/RNAi) was used to check for RNAi off-target effects using the gene fragments to be amplified as the query. The PCR reaction was set as follows: 94 °C for 2 minutes, 1st step for 6 cycles: 95 °C - 30 seconds, 55 °C - 30 seconds and 72 °C – 45 seconds; 2nd step for 30 cycles: 94 °C – 30 seconds, 72 °C for 45 seconds and finally 72 °C for 5 minutes extension. The expected PCR product sizes were confirmed using agarose gel electrophoresis at 90 V for 60 minutes before the product was purified using the QIAquick PCR purification kit (Qiagen) kit. The purified PCR product was then used as a template to generate dsRNA using the Megascript T7 transcription kit (Ambion).

Table 5.2. Primers for amplifying *Nasonia* and *E. coli lacZ* gene fragments.

Gene	Primers
Lac Z	F: TAATACGACTCACTATAGGGAGACCACGTTTTACAACGTCGTGACTGG R: TAATACGACTCACTATAGGGAGACCACGGGCCTCTTCGCTATTACG
Cinnabar	F: TAATACGACTCACTATAGGGAGACCACCGGTGAAGGTGAGGTCGTCG R: TAATACGACTCACTATAGGGAGACCACCTCGCATCGGGGTAGGGAAG
VY	F: TAATACGACTCACTATAGGGAGACCACAGCCTTTACCGTGCTACTGG R: TAATACGACTCACTATAGGGAGACCACAATGGCGGCAACTCTTGACC
VZ	F: TAATACGACTCACTATAGGGAGACCACAATATGAAGCCTTGAAAGAGG R: TAATACGACTCACTATAGGGAGACCACTCAAACTTGTTGTTCCAGCC
VX	F: TAATACGACTCACTATAGGGAGACCACTATCGGAAAAACGGTAACTA R: TAATACGACTCACTATAGGGAGACCACGCTACTAGAGGTAATAAGGC
SP97	F: TAATACGACTCACTATAGGGAGACCACCGGAAGAGATTACGGAATCC R: TAATACGACTCACTATAGGGAGACCACTCCTGGAGCCAGTGTAGCAC
VG	F: TAATACGACTCACTATAGGGAGACCACAAATCGCACTTTTCTTAGCC R: TAATACGACTCACTATAGGGAGACCACACTTATTACACAGTTATGGATG
VQ	F: TAATACGACTCACTATAGGGAGACCACCACTAAAGTGTTGATTCTCGT R: TAATACGACTCACTATAGGGAGACCACTAGTTCGATACCTTTGTCCT

5.2.3. RNAi Microinjection

Nasonia females were conditioned to sting by pre-exposure to hosts at a ratio of one female per two hosts for a period of 24 hours. The preconditioned females were then provided fresh hosts for an overnight period (~14 hours). The females were removed and the stung hosts were incubated for a further 5 days at 27 °C. Larvae were removed from the hosts and placed on an 1.5% PBS agar plates (5mM NaCl, 0.56 mM KCl, 1mM Na_2HPO_4, 1.5% v/v agar) before being injected with dsRNA injections.

Microinjection capillary needles were prepared using a Sutter Instrument Co. needle-pulling machine (model P-2000) under these settings: Heat - 270, Velocity - 255, Pull - 4, Delay - 85, Filament - 70. Needles were back loaded with 5uL of 1µg/µL dsRNA

dissolved in 10% (v/v) Hansel's blue food colouring. The dsRNA was injected into the posterior region of the larvae using a FemtoJet microinjector with the pressure of each injection set at 1000 psi. Successful injection was visually confirmed by the presence of a blue dye stain in the posterior end of the larvae. The injected larvae were allowed to develop to late stage pupae on the agar plates stored at 27 °C for six days. Pupae were then transferred to a 30 mL v-bottom vial tube until eclose.

5.2.4. Training and re-hosting experimental and control wasps

Wasps were then provided pupae of *L. sericata* in a v-bottom vial tube for 24 hours to allow them to learn how to sting and lay eggs. Host pupae were carefully separated from wasps and stored at 27 °C in a v-bottom vial tube. The wasps were then supplied with fresh pupae and left to sting overnight. To minimise the sting surface area, the hosts were wrapped in a foil paper or positioned in such a way that only the anterior section of the pupae was exposed to the wasp for stinging. The following day, the pupae were removed and cracked open to remove eggs allowing the host's response to venom to be analysed without the interference of developing larvae. The Fisher exact statistical test was used to measure the stinging rate of the knock down wasps in comparison to the wild type. This test was chosen because the grouped data was non-parametric from a small sample. This method returns a *p*-value for a test of the null hypothesis that the data in each one group category is significantly different while the alternative hypothesis simply means the data is not significantly distributed. The hypotheses are as follows:

H_0: The stinging rate is affected

H_1: The stinging rate is not affected

At 95% significance level, if the *p*-value is less than 0.05, then the H_0 hypothesis is rejected.

5.2.5. Functional characterization of knock down wasps

The cracked open hosts were counted as stung if a sting site or eggs were observed. Eggs were carefully removed using a soft brush and hosts were laid on the top of a moist paper towel to avoid desiccation for 5 days. The paper towel was changed every 2 days to avoid

fungal growth while developmental cues, eye pigmentation and bristle formation on the host were recorded and photographed.

5.2.6. Transmission electron microscopy

After photographs have been taken, cracked open *L. sericata* hosts at day 1 and day 3 were prepared for viewing under the transmission electron microscope (TEM). Sample preparation and cutting were done at the Manawatu Microscopy and Imaging Centre (MMIC, Massey University, Palmerston North). Whole body host pupae placed into Modified Karnovsky's Fixative (3% Gluteraldehyde (v/v) 2% Formaldehyde (w/v) in 0.1M Phosphate Buffer (pH7.2) for at least 24 hours. The host pupae were then washed in 0.1M sodium phosphate (pH7.2) buffer three times for 45 min each wash. Thee samples were then dehydrated through a graded acetone series (25%, 50%, 75%, 95%, 100%, 100%, 100%) for 45min each before being put into 50:50 resin:acetone and stirred overnight. This 50:50 resine:acetone was replaced by fresh 100% resin (Procure 812, ProSciTech Australia) for 8 hours on the stirrer. This step was repeated four more times (overnight in 100% resin, 8 hours in 100% resin, overnight in 100% resin, 8 hours in 100% resin). Whole body pupae were embedded in moulds with fresh resin and cured in a 60 °C oven for 48 hours. Light microscope sections were cut at 1 micron using a glass knife on the ultramicrotome (Leica EM UC7, Germany) and heat fixed onto glass slides. These were stained with 0.05% Toluidine Blue for approximately 12 seconds and viewed under the light microscope. The block was then trimmed down to the selected area and cut using a Diamond Knife (Diatome, Switzerland) at 100nm. These were stretched with chloroform and mounted on a grid using a Quick Coat G pen (Daido Sangyo, Japan).

Grids were stained in Saturated Uranyl Acetate in 50% Ethanol for 6.5 min, washed with 50% ethanol and MilliQ water and then stained in Lead Citrate (Venable and Coggeshall, 1965) for a further 6.5 minutes. This was followed by a wash in MilliQ water and samples were viewed and imaged with an FEI Tecnai G_2 Spirit BioTWIN, with 120kV accelerating voltage (Czech Republic) (Camera: Veleta, Olympus SIS Germany).

5.2.7. Recombinant protein expression

RNA extracted from whole body wasps was reverse transcribed into cDNA using the transcriptor first strand cDNA synthesis kit (Roche). Primers that incorporated the restrictions site at the 5' and 3' ends of the forward and reverse primers respectively were designed and synthesized (Table 5.3).

Table 5.3. Primers designed to amplify the venom genes. *Bam*H1 restriction sites were added to the forward primers (Fwd-*Bam*H1) and *Eco*R1 restriction sites were added to the reverse primers (Rev-*Eco*R1).

Gene	Primers	Length of fragment
VY	Fwd-*Bam*H1: AATTGGATCCACTGTACCTTTCCAACTAGCTACG Rev-*Eco*R1: GCGGGAATTCTTACTGGCTGTGATATATATTCCTGC	263
VZ	Fwd-*Bam*H1: TATTGGATCCGCCGAAGTCCAATCCAAAT Rev-*Eco*R1: CGAGGAATTCTTAAGATTTGATAACAACTTCGGCTTTC	416
VX	Fwd-*Bam*H1: TATTGGATCCTCAGAAGATGACCGACGATCTC Rev-*Eco*R1: CGGAGAATTCTTAATTCTCGGAGGACTTGATGG	304

Amplification of the genes of interest and purification of the PCR products were performed as per protocols in sections 2.8 and 2.10.The vectors pPROEX HTb, lab-edited pET-OmpA HT MCS and pET-DsbC HT Bam plasmids kindly donated by Trevor Loo (Massey University) were used as expression vectors (Figure 5.1). pET-OmpA HT MCS contains an *OmpA* signal peptide downstream from a T7 promoter which exports attached polypeptide sequences into the periplasm of the bacteria. The vector contains the gene encoding *disulfide bond isomerase C* (DsbC), which is a periplasmic chaperone that has a signal sequence directing the protein to the periplasm. The multiple cloning site contains sites for *Bam*H1 and *Eco*R1 (Roche), and contains a His Tag and tobacco etch virus (TEV) protease site 5' to the gene of interest (Fig 5.1). The genes of interest and plasmid were subjected to restriction digest by *Bam*H1 and *Eco*R1 followed by ligation following the manufacturer's ligation kit instructions (Roche).

The plasmid was transformed into XL1-Blue cells (Stratagene, USA) using the heat shock protocol (Section 2.17) and incubated overnight at 37 ℃ to get increase the plasmid stock Colony PCR was used to test for the presence of the correct insert in the plasmids. In this type of PCR, six colonies were picked from the agar plate and each mixed with 15 µL of pure water, 5 µL of this solution was added to 1 mL of LB and incubated at 37 ℃, while the remaining 10 µL was boiled for five minutes. The boiled mixture was used as DNA template for PCR to confirm the presence of insert of interest in the plasmid. The products of each PCR reaction were analysed by gel electrophoresis and colonies that had the correct insert were cultured overnight in LB with shaking at 37 °C before the plasmids extracted using the Gene Elute HP plasmid miniprep kit (Roche) following the manufacturer's instructions then sent for sequencing to confirm the identity of the insert. The plasmids were transformed into the Rosetta™ (DE3) (Novagen) expression cells and the genes of interest were expressed and harvested as described in section 2.18 and 2.19.

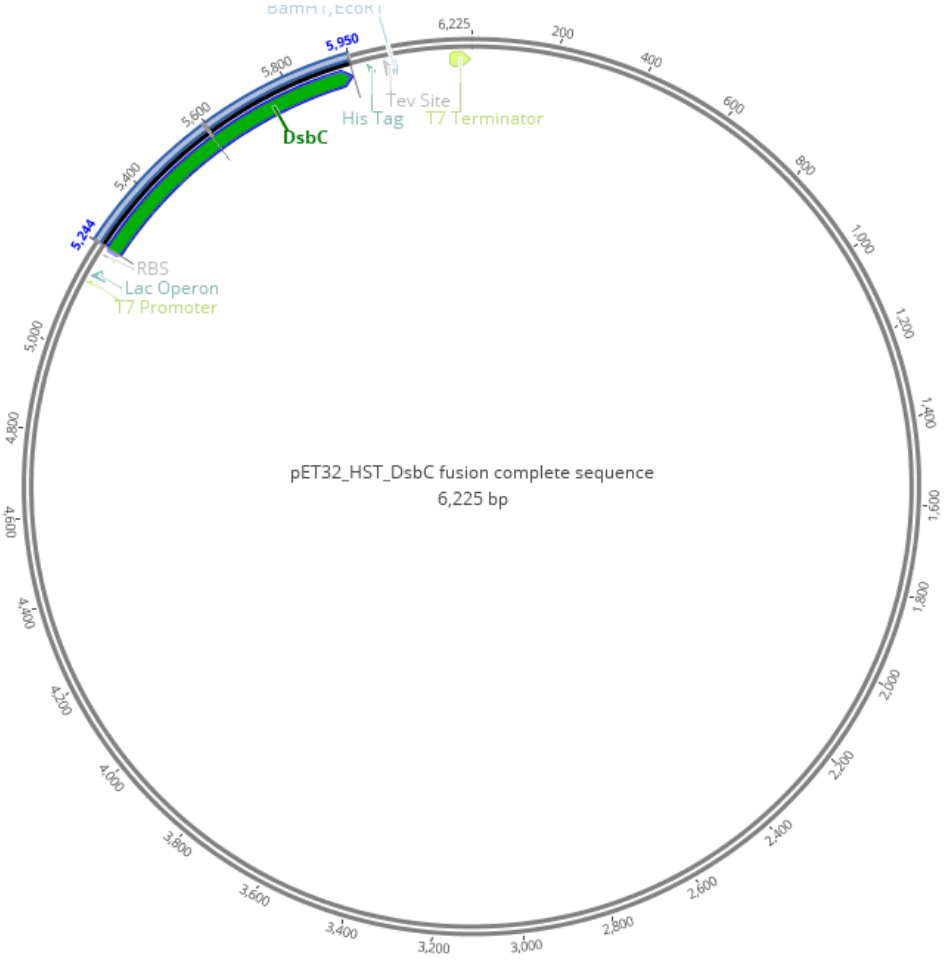


Figure 5.1. Edited pET-DsbC HT Bam vector. *Disulfide bond C isomerase* added downstream of the ribosomal binding site (RBS).

5.3. Results

The primary goal of this chapter was to elucidate the possible functions of individual *Nasonia* venom proteins in the host *Lucilia sericata*. Six venom genes (Table 5.1) were chosen for this study because they are highly expressed in the *Nasonia* venom producing gland [231]. The individual functions of the venom proteins transcribed from these genes remain unknown. Previous work had shown that five of the selected genes (Venom Y, G, X, Z, Q) have no homology to any known protein while one, SP97 contained a structural motif identified or belonging to the serine protease family [121]. To analyse the functions of the venom proteins, RNA interference (RNAi) was used to individually knock down expression from these genes. Attempts were also made to clone and recombinantly

produce the venom proteins in a heterologous host in order to structurally characterize them.

5.3.1. *In silico* analysis of *Nasonia* venom proteins

The amino acid sequences of the six venom proteins were retrieved from the NCBI database (Appendix 4) and used as a query to search against the protein databases for proteins that have homology to the selected sequences. None of the three databases searched (NCBI nucleotiode (nr/nt), SwissProt and Pfam) returned any hit similar to the venom proteins except for SP97, which contains homology to members of the serine protease family (Table 5.4). *Nasonia* venom is synthesized in the venom gland where it is discharged to the reservoir for storage while awaiting delivery into the host. Since these are secretory proteins, SignalP v4.1 was used to identify possible signal peptide sequences. As expected all the six venom proteins were found to have potential signal peptides varying from 17 to 23 amino acids (Table 5.4). When these sequences were removed, the mature proteins of venoms Y, Z, X, Q and G ranged from 240 to 441 amino acids, corresponding to predicted molecular weights of 19.6 to 33.9 kDa. All six proteins had theoretical isoelectric points (pI) of ~5. Multiple sequence alignment (MSA) of the signal peptide sequences showed that the branched chain amino acids (BCAA) are the most represented group and could potentially play an important role in the translocation of venom proteins to their targets (Fig 5.2).

Table 5.4. Homology search and physicochemical properties of the *Nasonia* venom proteins.

Protein	Signal peptide	Length	Molecular weight (kDa)	pI
VY	Yes (23aa)	243	20.4	5.34
VG	Yes (22aa)	240	19.6	5.12
VZ	Yes (23aa)	396	31.3	5.25
VQ	Yes (18aa)	441	33.9	5.24
VX	Yes (20aa)	279	23.0	5.11
SP97	Yes (17aa)	831	68.9	5.08

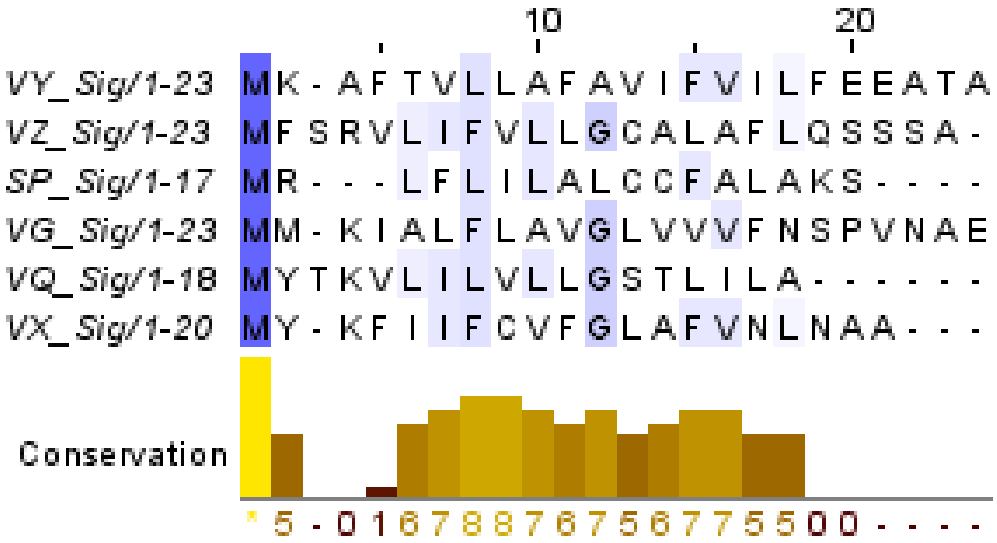


Figure 5.2. Multiple sequence alignment of the signal peptides found in the six most highly expressed venom genes. The colour intensity shows the level of conservation, blue – highly conserved and light blue – weaker sequence conservation. The numbers under the histogram are a numerical index reflecting the conservation of physico-chemical properties in the alignment ranging from 0 – lowest to 11 or * highest.

Structural prediction can provide tentative information regarding the functions of a protein. An assumption is made that proteins that share the same functions often have similar folds therefore structural predictions rely mainly on fold matches. The mature venom protein sequences were submitted to the Phyre2 v2.0 webserver for structural prediction. Table 5.5 summarises the fold predictions of the six venom proteins and graphical views of the secondary predictions are shown in Appendix 5. Average confidence refers to the probability that the sequence is homologous to the predicted template. Phyre2 predicted that VY was predicted to be a hydrolase with a 58.4 % confidence showing that 50% of the protein in random coil, while the rest is made up of β-strands with a small amount of α helix as shown in Fig 5.3A. VZ was predicted to be a transport protein with an average confidence of 80.1%. The suggested model is mainly α-helical, (96%) as shown in Fig 5.3B. Although VX is predicted to be a transcription factor the confidence in the prediction was low. Despite this, Phyre2 predicted its structure would be mainly α-helical as shown in Fig 5.3C.

A BLASTp search assigned SP97 to the serine protease family (Table 5.5) and this was supported by the analysis using Phyre2 that classified this protein as a protease with an average confidence of 100%. Phyre2 predicted that SP97 contains approximately three α-helices and 10 β-strands (Fig 5.3D). VG was predicted to belong to the cysteine family of proteases and the model generated was mainly α-helical (Fig 5.3E). VQ was predicted to function as an α-helical lipid-binding protein with high confidence (Fig 5.3F).

Table 5.5. Consensus secondary structure predictions for the venom proteins through the Phyre2 webserver.

Protein	Predicted Function	α-Helix	β-Strand	Coil	Average confidence
VY	Hydrolase	4%	44%	52%	58.4%
VZ	Transport protein	96%	0%	4%	80.1%
VX	Transcription	72%	0%	28%	32.1%
SP97	Serine Protease	14%	37%	49%	100%
VG	Cysteine protease	91%	0%	9%	28.2%
VQ	Lipid binding protein	97%	0%	3%	96.3%

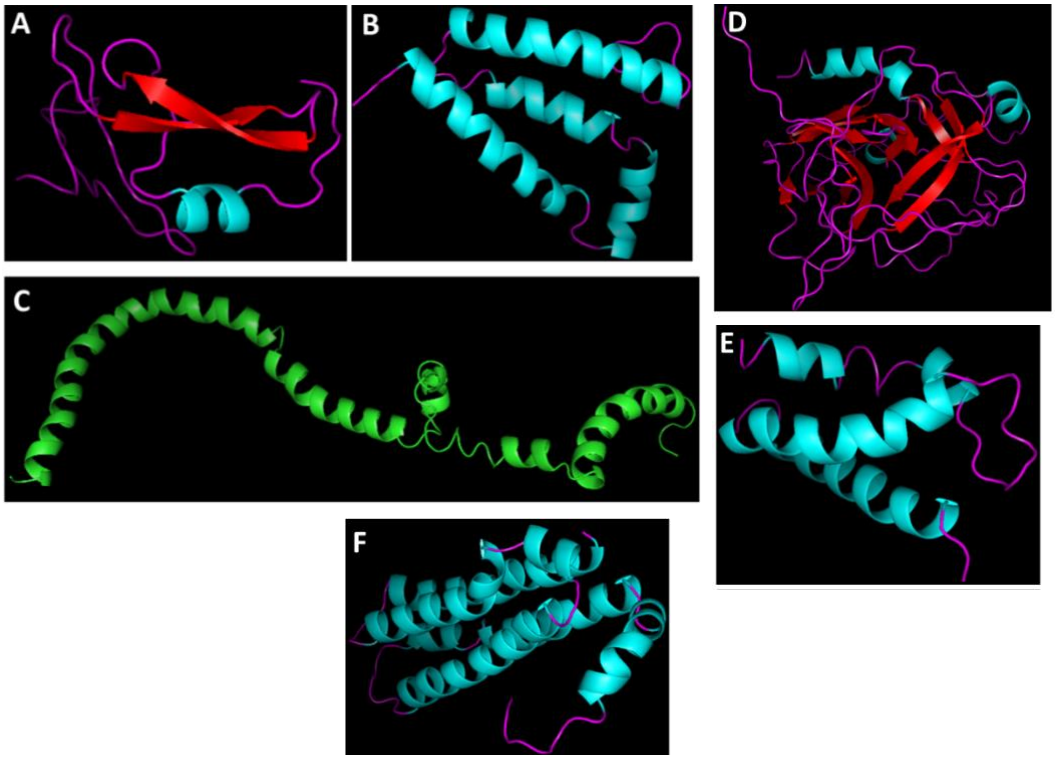


Figure 5.3. Secondary structure predictions of the venom proteins. A: *Venom Y*, B: *Venom Z*, C: *Venom X*, D: *Serine Protease 97*, E: *Venom G*, F: *Venom Q*. The coil at the top or left represents the N-terminal.

5.3.2. Functional characterization of venom genes through RNA interference

To investigate the function of *Nasonia* venom in the host, a combination of knock down (KD) of wasp venom genes by RNA interference (RNAi) and phenotype examination of the envenomated hosts was used. RNAi is a technique used to knock down gene function in which *Nasonia* larvae were injected with double stranded RNA (dsRNA) designed to target mRNA for degradation, therefore reducing the amount of mRNA available for translation into protein. KD wasps were then exposed to hosts for envenomation and phenotype produced by removal of a single venom gene product was examination. Eggs were removed from the hosts and host development was observed for a period of five days. Envenomation of the host has been shown to result in immune response suppression, developmental arrest, an increase in lipid levels and nutrient release [122]. Therefore, changes in development on the host pupae were used to evaluate the effects of individual gene knock downs.

Cinnabar (CN) is an eye colour gene in *Nasonia* responsible for the change in the red eye colour of pupae to grey in the adult. Knock down of CN results in adult wasps that have a red eye. To test the efficacy of the RNAi knock down method, the experiment in Section 3.3.6 was repeated where CN was targeted using RNAi with the expectation that the CN KD wasps would have red eyes if the KD was successful. As expected, the wild type adult wasps developed a grey eye colour while the CN KD wasps could be identified by a red eye phenotype and a semi-quantitative RT-PCR was carried out where wild type and CN KD cDNA was amplified using three different numbers of PCR cycles. Wild type cDNA showed amplification over 25, 30 and 36 PCR cycles, confirming the expression of CN (Section 3.3.6, Fig 3.10). However, no amplification of the cDNA synthesized from CN KD RNA was observed even using the highest number of cycles, confirming there was no detectable expression of CN. This result confirmed the efficacy of the RNAi strategy to study gene function in *Nasonia*, validating this same approach being used to study the function of venom genes.

5.3.3. Venom gene knock downs

In order to KD each gene, primer sets were synthesized based on the quality check of RNAi off target prediction. The PCR product amplified from these reactions would be used as template to generate dsRNA. The expected PCR sequence was submitted to the WaspAtlas webserver to ensure that each dsRNA was specific to only one target gene. The analysis showed that all the synthesized dsRNAs should be specific to the target genes. For the KD experiments, dsRNA was diluted with water and food colouring to make a concentration of 1µg/µL. The food colouring was added to enable visibility of the injected dsRNA to the *Nasonia* larvae. A successful injection of dsRNA could thus be confirmed by the appearance of blue colour on the posterior side of the larvae. These larvae were then incubated on PBS agar plates until they developed into adults. Recently eclosed adult *Nasonia* (<1 day) were trained to sting by exposing them to *L. sericata* pupae (the host) for a period of approximately 14 hours. The trained wasps were then provided with fresh hosts with only their anterior part exposed, for a period of six hours at a ratio of one host per wasp. It was important to restrict the stinging to allow the

identification of the sting site and the visualization of the laid eggs. The anterior end of the host that was exposed to the wasp was cracked open and the *Nasonia* eggs were removed to ensure that any phenotypic change observed would be due to the action of the venom alone. The cracked open hosts were incubated at 25℃ whilst being examined daily for physiological changes.

Table 5.6 and Appendix 6 summarize the RNAi of the six venom genes and the host phenotypes after a period of five days. Initially three sets of controls (negative, positive and RNAi) were set up to give an outline of results that were used for phenotype comparisons. The negative control group included hosts that were not exposed to wasps and therefore not stung. These were cracked open and after a period of five days, and signs of development such as the eye pigment could be clearly observed (Fig 5.4). The positive control group comprised wasps that were injected with water only. Fifty-four of the 60 positive control wasps stung the host and as expected development was stopped in these stung hosts as evidenced by the lack of bristle formation and eye pigmentation (Fig 5.4). To ensure that introduction of dsRNA into *Nasonia* did not affect the ability of the wasp to sting, an RNAi control assay was performed in which dsRNA targeting *LacZ* was injected. *Nasonia* wasps do not possess the *LacZ* gene therefore *LacZ* dsRNA should have no effect on the wasp (Table 5.6, Fig. 5.4). *LacZ*-KD wasps stung the hosts and the phenotype observed was similar to the positive control indicating that introduction of dsRNA does not affect stinging or envenomation (Fig 5.4). Fisher's exact statistical test showed that the RNAi assay does not affect the stinging rate of the KD wasps because all the p-values for the KD wasps were above 0.05 meaning the null hypothesis could not be rejected and there was no significant difference in the stinging rate between the wild type and the KD wasps.

Table 5.6 Summary of the RNAi in *Nasonia* targeting venom genes.

Gene	Hosts exposed	Stung hosts	Fishers exact *p*-value	Host Phenotype (Number showing the phenotype)
Unstung (-ve control)	60	-	-	Developed with eye pigment observed
WT (+ve control)	60	54	-	No development
LacZ (RNAi control)	60	58	0.7943	No development
VY	75	63	0.8009	Increased melanisation (39)
VZ	57	31	0.0872	No eggs laid + Host slowly developing (27)
SP97	88	74	0.8072	No development
VG	75	69	1	No development
VQ	77	69	1	No development
VX	111	92	0.7265	Developed with eye pigments observed (65)

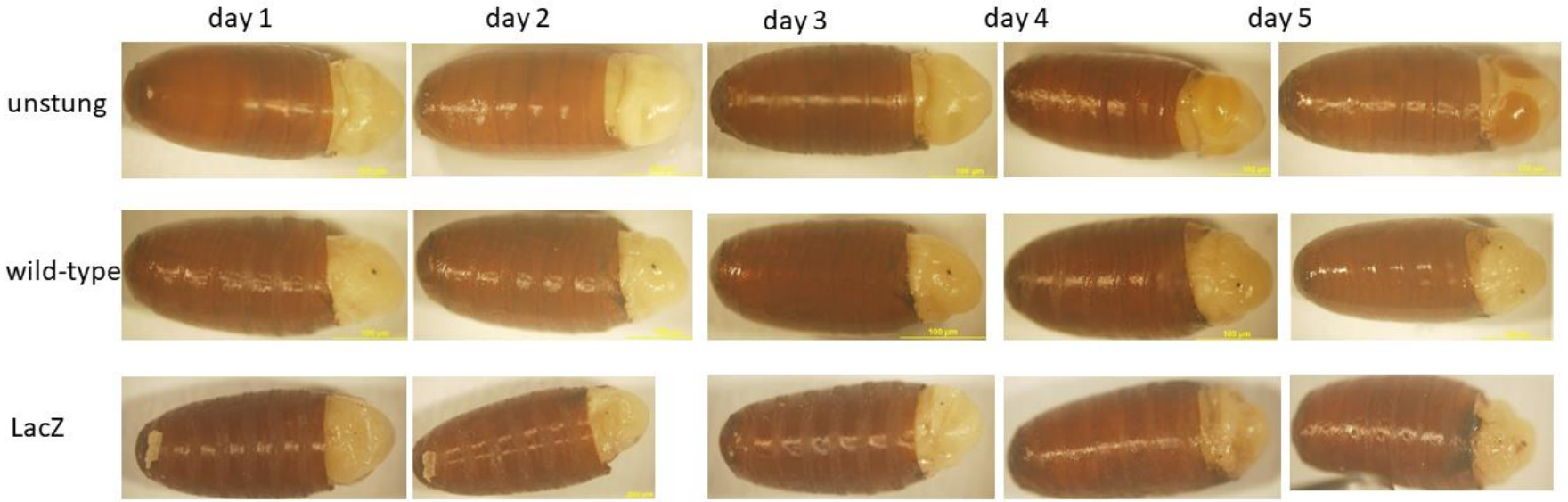


Figure 5.4. ***L. sericata*** **pupae hosts cracked open on the anterior side**. These hosts were stung by *Nasonia* wasps (black spot) except the KD control - unstung – negative control. The unstung hosts developed as shown by the appearance of eye pigment at day 5. wild-type – positive control which is hosts stung by wildtype *Nasonia* (injecting all the 79 peptide venoms), *LacZ* – RNAi control which means wasps that were injected dsRNA mixture which had no effect in the wasp.

Injection of dsRNA targeting VY mRNA did not significantly affect the stinging rate (Table 5.6). Interestingly, 39 of the 65-stung hosts showed a qualitative increase in melanisation at the sting sites which was not observed in the control (Fig 5.5A). Melanisation is part of the host immune defence mechanism suggesting that when VY is removed from venom (by KD) the hosts immune system is not supressed in the same way as when envenomated by wild-type venom. This result suggests that VY is involved in suppressing the immune response in the host, although more work needs to be carried out to verify this result. Less than 50% of the these VY stung hosts had a melanisation level at the sting site similar to the hosts stung by the wild type wasps, suggesting that the dsRNA injected may not have been enough to knock down VY expression in *Nasonia* that stung these hosts or there was individual variability in the immune response of the stung hosts. To confirm that the increased melanisation phenotype observed was due to the knock down of VY gene, a semi-quantitative RT-PCR experiment was carried out. RNA was extracted from whole body wasps of wild type and VY-KD wasps and used to synthesize cDNA which was then used as a template for a PCR reaction for 25 and 36 cycles. Amplification was observed for both 25 and 36 cycles for both wild type and VY-KD samples, although the intensity of the wild type cDNA amplicon was greater compared to that for VY-KD, supporting at least partial knock down of VY. Ribosomal protein 49 (RP49) was targeted for amplification as a positive control for the semi-quantitative RT-PCR. Amplification of this gene produced bands of the same intensity for both wild type and VY-KD suggesting that knock down of VY was successful (Fig 5.6).

Venom Z knock down appears to affect the stinging rate of the wasps as only 54% of the KD wasps stung the hosts. Furthermore, none of the stung hosts had eggs on them indicating that VZ-KD also affects *Nasonia* egg laying. Although the stung hosts developed eye pigment at day two, the pupae failed to develop further and at day five they began to decay, suggesting that the hosts had died. Another interesting observation was that the hosts had an “oily” appearance that developed at day 3 potentially indicating an increase in the production of lipids. This suggests that VZ may play a role in suppressing development, control of lipid production and possibly apoptosis of the host pupae (Fig 5.5B). Semi-quantitative RT-PCR confirmed the VZ gene was indeed knocked down and

the RP49 positive controls support the observation that phenotypic changes observed are a result of the knocked down gene (Fig 5.7).

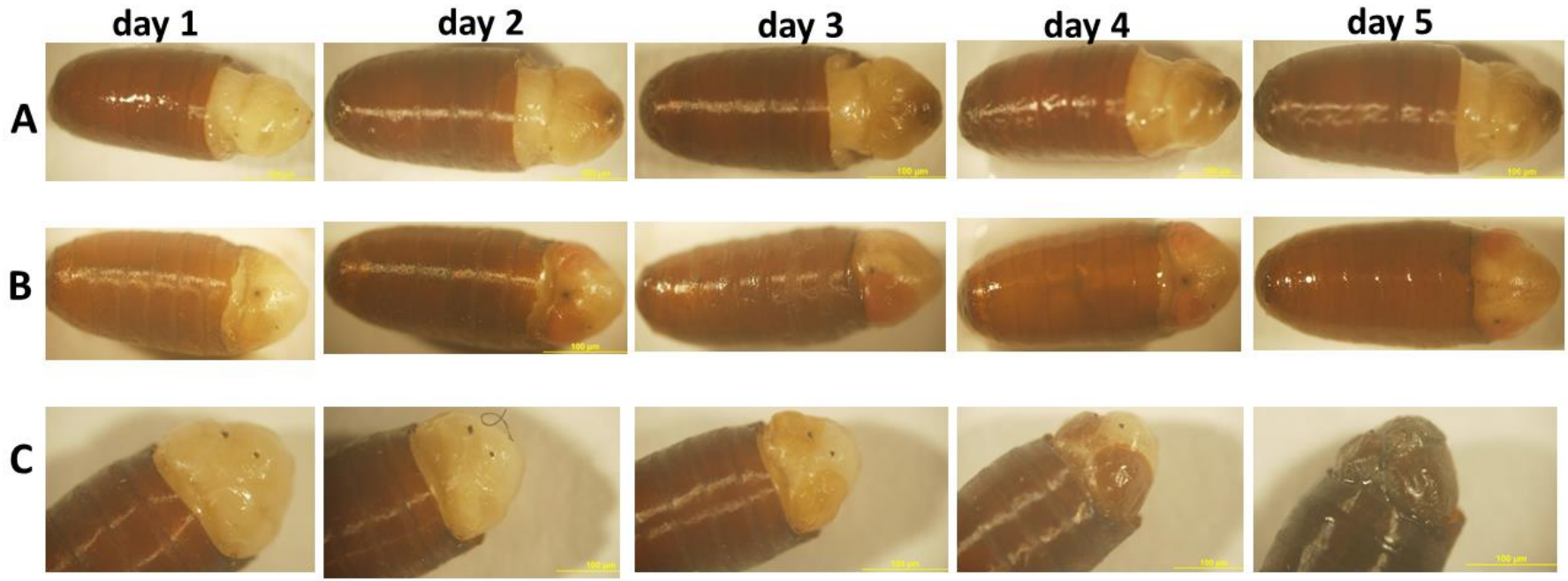


Figure 5.5. *L. sericata* pupae hosts cracked open after a stinging assay using KD wasps. A: VY-KD, B: VZ-KD, C: VX-KD. Development is observed when the eye pigment develops in which hosts stung by VZ-KD and VX-KD wasps showed signs of development.

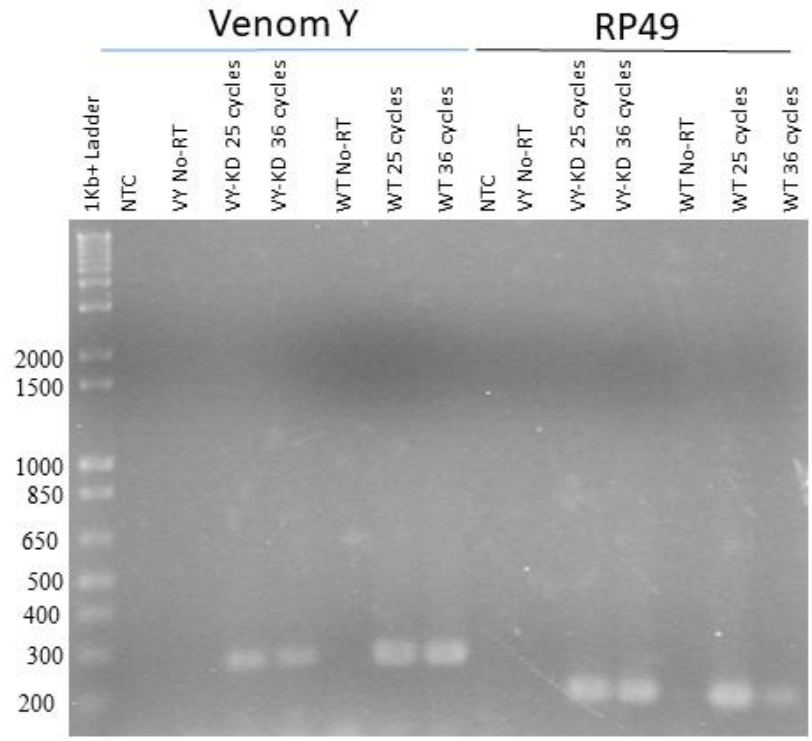


Figure 5.6. Semi-quantitative RT-PCR targeting VY and RP49 in wild type and VY-KD wasps. RP49 is a loading control, which is ubiquitously expressed across *Nasonia* tissue and amplicon intensity is similar for VY-KD and wild type as expected. Ladder is 1kb plus (ThermoFisher Scientific) and the units shown are in bp.

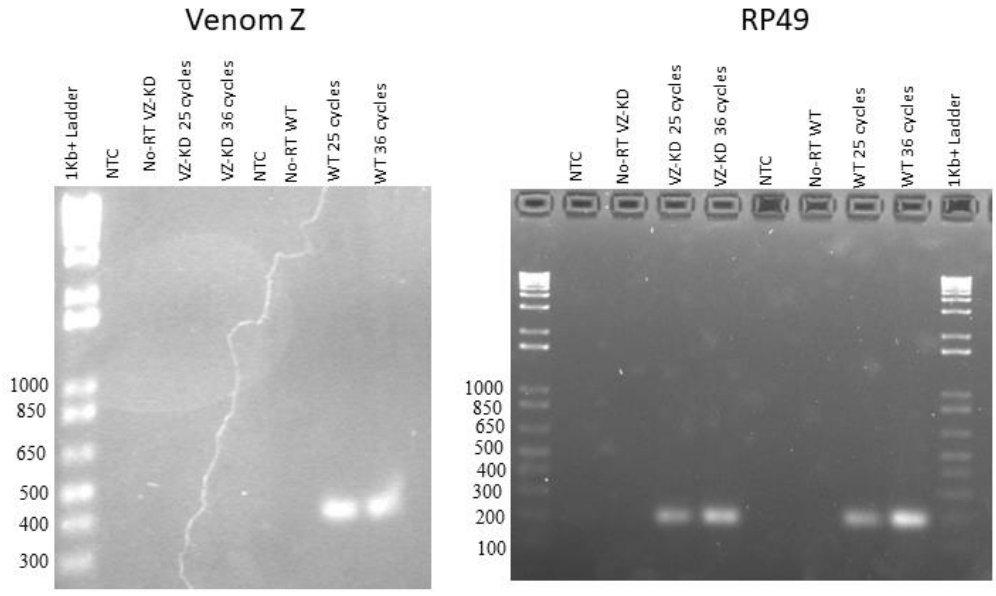


Figure 5.7. Semi-quantitative RT-PCR targeting VZ and RP49 in wild type and VZ-KD wasps. RP49 is a loading control, which is ubiquitously expressed across *Nasonia* tissues and its amplicon intensity is similar for VZ-KD and wild type as expected. Ladder is 1kb plus (ThermoFisher Scientific) and the units shown are in bp.

The stinging assay for VX-KD resulted in 71% of the stung hosts successfully developing to generate eye pigment and form bristles (Fig 5.5C). The phenotype was similar to the unstung hosts suggesting that VX may be involved in suppressing development in stung hosts. Semi-quantitative RT-PCR confirmed that VX-had been successfully knocked down, as there were no amplicons after both 25 and 36 cycles in contrast to wild type cDNA (Fig 5.8). Amplification of the positive control RP49 was observed from both VZ-KD and wild type cDNAs confirming that knock down of VX was successful.

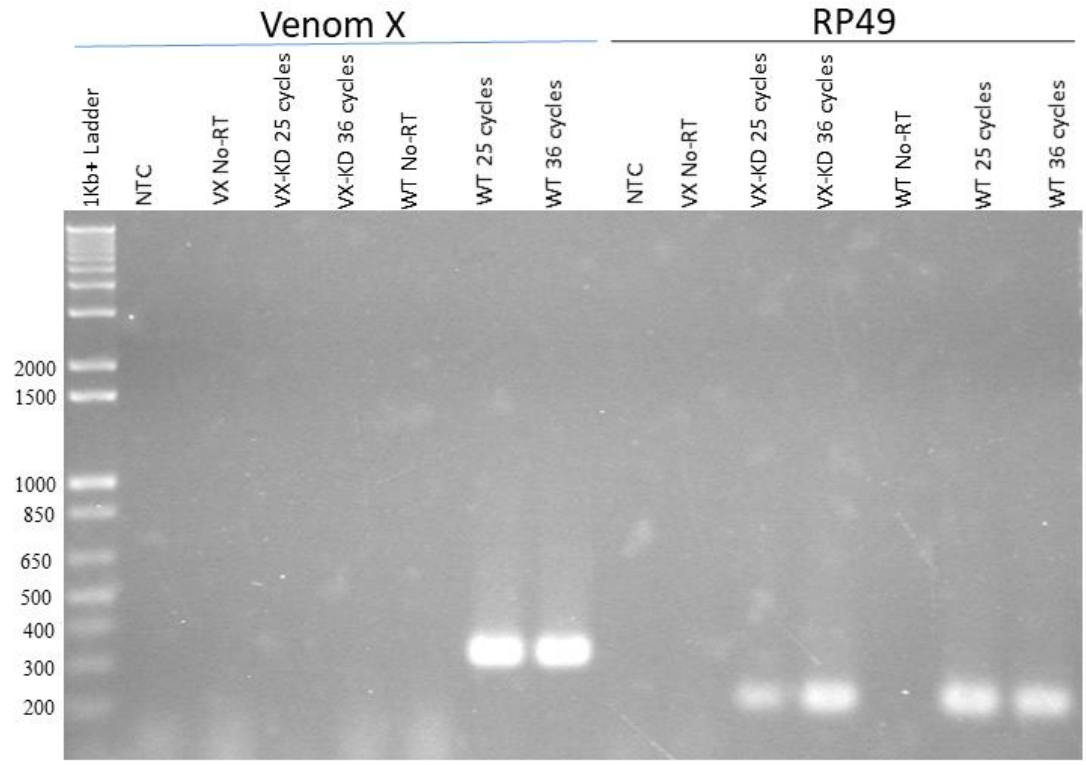


Figure 5.8. Semi-quantitative RT-PCR targeting VX and *RP49* in wild type and VX-KD wasps. *RP*49 is a loading control, which is ubiquitously expressed across *Nasonia* tissue and is equally amplified from VX-KD and wild type cells as expected. Ladder is 1kb plus (ThermoFisher Scientific) and the units shown are in bp.

The stinging assay carried out using SP-KD, VG-KD and VQ-KD wasps did not significantly affect their stinging rate and they all laid eggs (Appendix 6). As the phenotypes of these three gene knock downs were similar to the positive control (Fig 5.9), functional predictions of these genes could not be made based on the visual assays used

in this experiment (see section 5.3.4). Therefore, the SP97, VG and VQ venoms were not examined further in the subsequent experiments. Semi-quantitative RT-PCR confirmed that the genes were knocked down of their expression, as the intensity of the amplicons of the KD wasps was lesser than the wild type (Fig 5. 10).

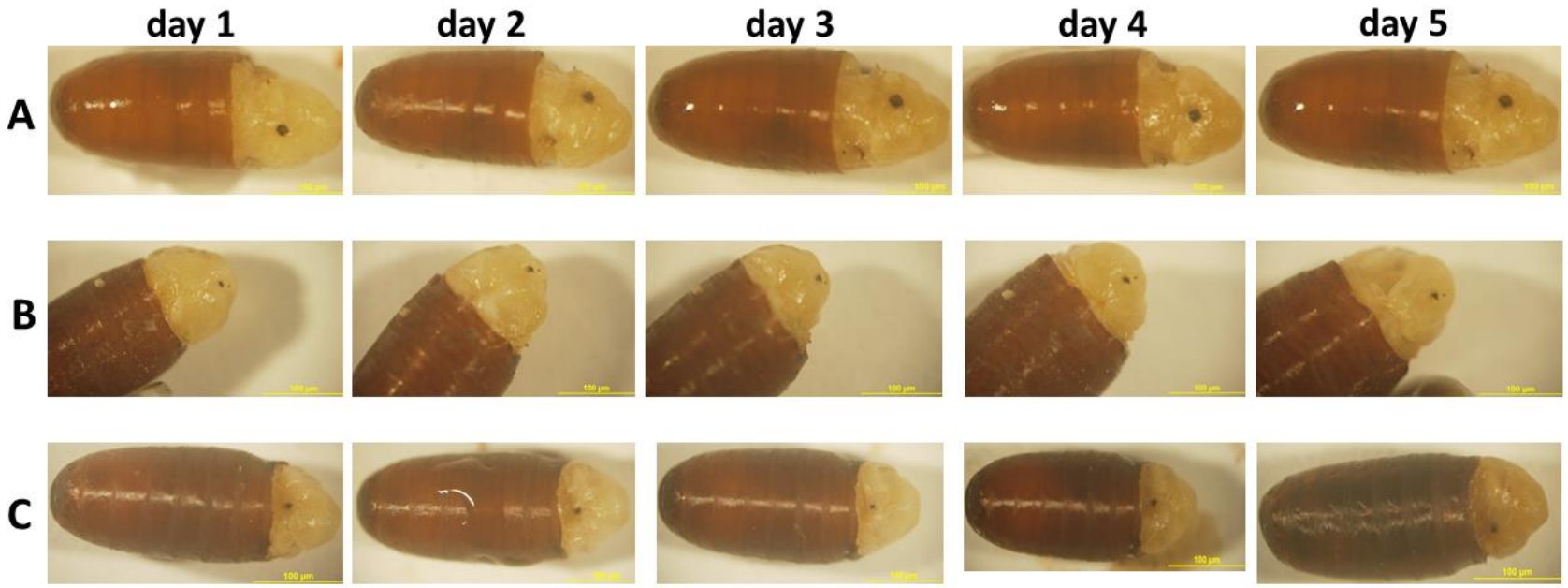


Figure 5.9. *L. sericata* pupae hosts cracked open after a stinging assay using KD wasps. A: VG-KD, B: VQ-KD, C: SP-KD. No development observed on the hosts as evidenced by no eye pigment appearance.

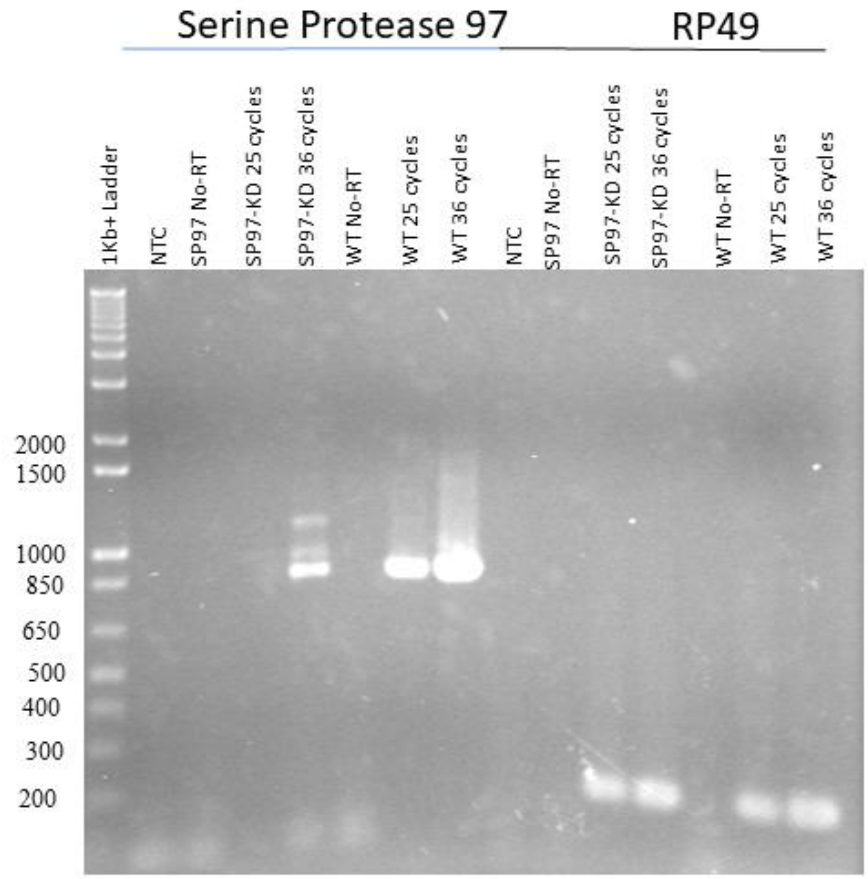


Figure 5.10. Semi-quantitative RT-PCR targeting SP97 and *RP49* in wild type and SP97-KD wasps. RP49 is a loading control, which is ubiquitously expressed across *Nasonia* tissue and its amplicon intensity is similar for SP97-KD and wild type as expected. Ladder is 1kb plus (ThermoFisher Scientific) and the units shown are in bp.

5.3.4. Transmission electron microscope analysis of the larvae stung by VY, VZ and VX knock down wasps

In an effort to elucidate the changes happening at the cellular level, the whole body pupae of *L. sericata* hosts was prepared for examination under a transmission electron microscope (TEM). Preliminary images were focused on the quantity and size of fat globules and oil droplets in the host cells as they develop after being stung by the wasps. The negative control in which hosts that were not exposed to wasps showed that after day 3 of development the number of fat globules and oil droplets decreased although the size of the larvae remains almost the same (Fig 5.11). Phenotypically, the unstung hosts show signs of development based on the appearance of eye pigment (Fig 5.4).

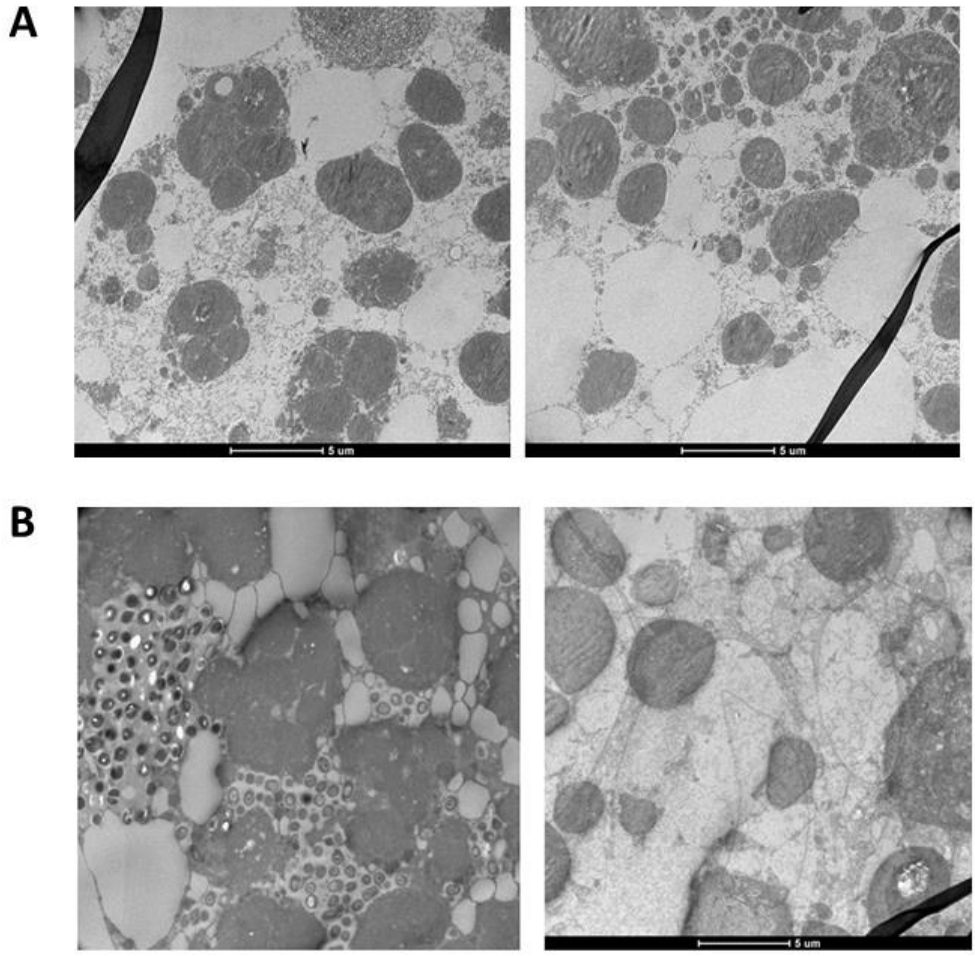


Figure 5.11. Transmission electron microscopy image of sections cut from whole body of unstung *L. sericata* pupae. A: day 1, B: day 3. Presumably, the circular black spots represent oil droplets while the circular white spots represent the fat globules.

In hosts that were stung by wild-type *Nasonia* that had all the 79 venom peptides represented it was observed that at day 3 the number of presumed fat globules had increased (Fig 5.12). A comparison of hosts that were stung by VY-KD wasps at day 1 and day 3 showed that the number and size of both the presumed fat globules and oil droplets had increased, suggesting that the hosts were not developing (Fig. 5.13). Hosts stung by VZ-KD wasps showed two phenotypes, which are development as shown by the appearance of eye pigment as well as increase of lipid production as shown by the oily cuticle of the host (Fig. 5.5.) Observing the sections from the VZ-KD stung hosts under TEM showed an increase in the number of fat globules and large oil droplets validating the observation of any oily surface (Fig 5.14). The oil droplets showed signs of

degradation as shown by the disappearance of uniform colour in the droplets (Fig 5.14). TEM of VX-KD stung hosts showed a significant decrease of fat globules after day 3 suggesting development of the host (Fig 5.15). In addition, deformation of the oil droplet structure was observed in the micrograph at day 3.

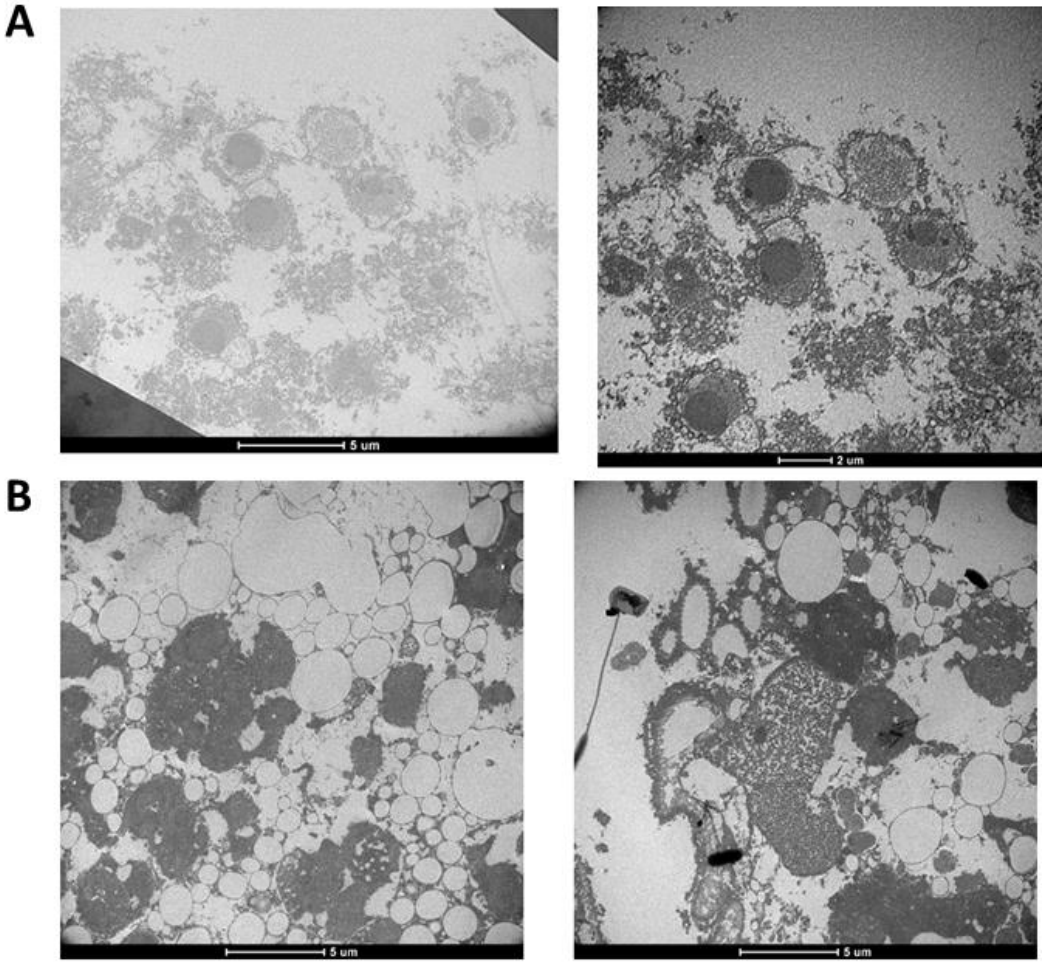


Figure 5.12. Transmission electron microscopy image of sections cut from whole body of *L. sericata* pupae stung by wild type *Nasonia* (injected with venom containing all the 79 peptides). A: day 1, B: day 3. At day 1 the same area of focus is viewed at different magnifications. Presumably the circular black spots represent oil droplets while the circular white spots represent the fat globules.

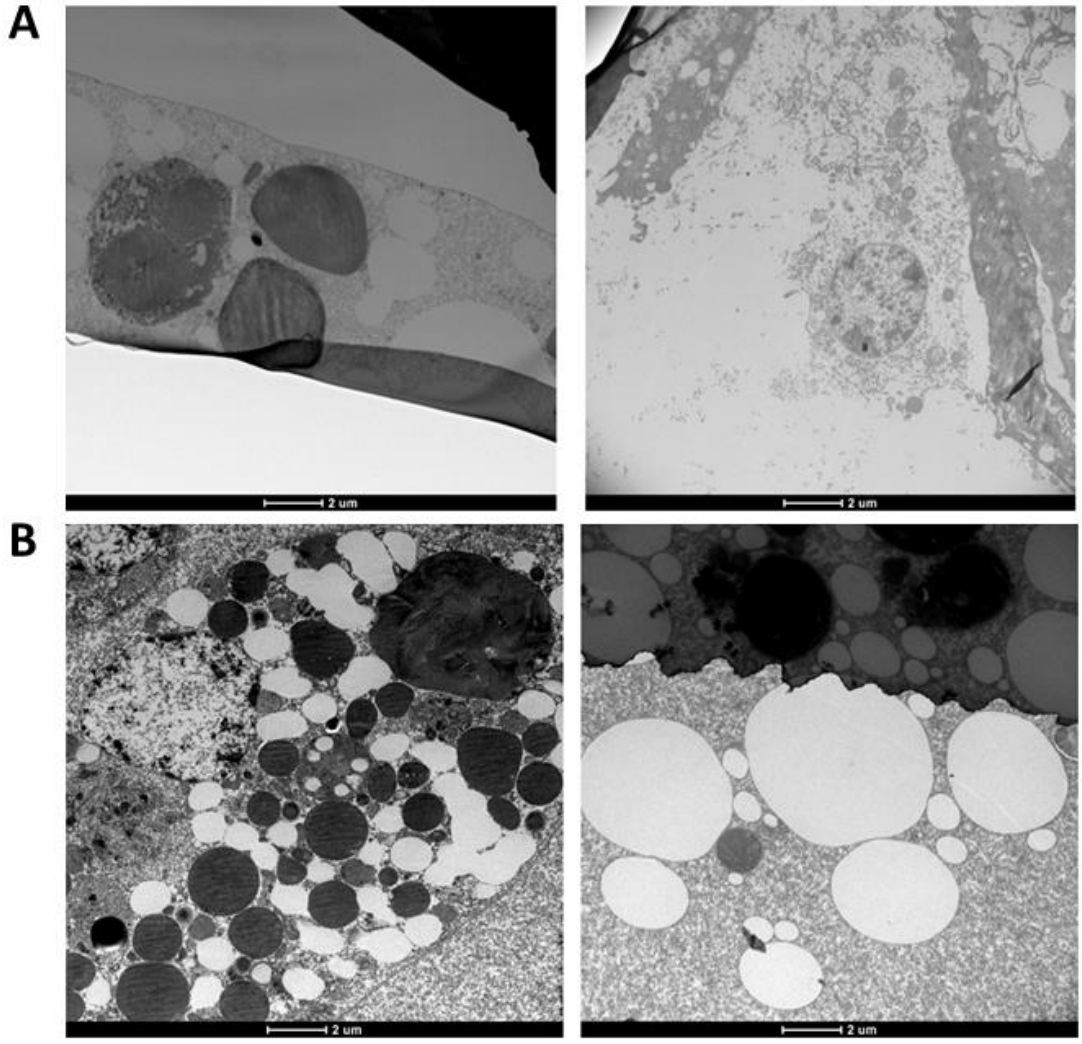


Figure 5.13. Transmission electron microscopy image of sections cut from whole body of *L. sericata* pupae stung by VY-KD wasps. A: day 1, B: day 3. Presumably, the circular black spots represent oil droplets while the circular white spots represent the fat globules.

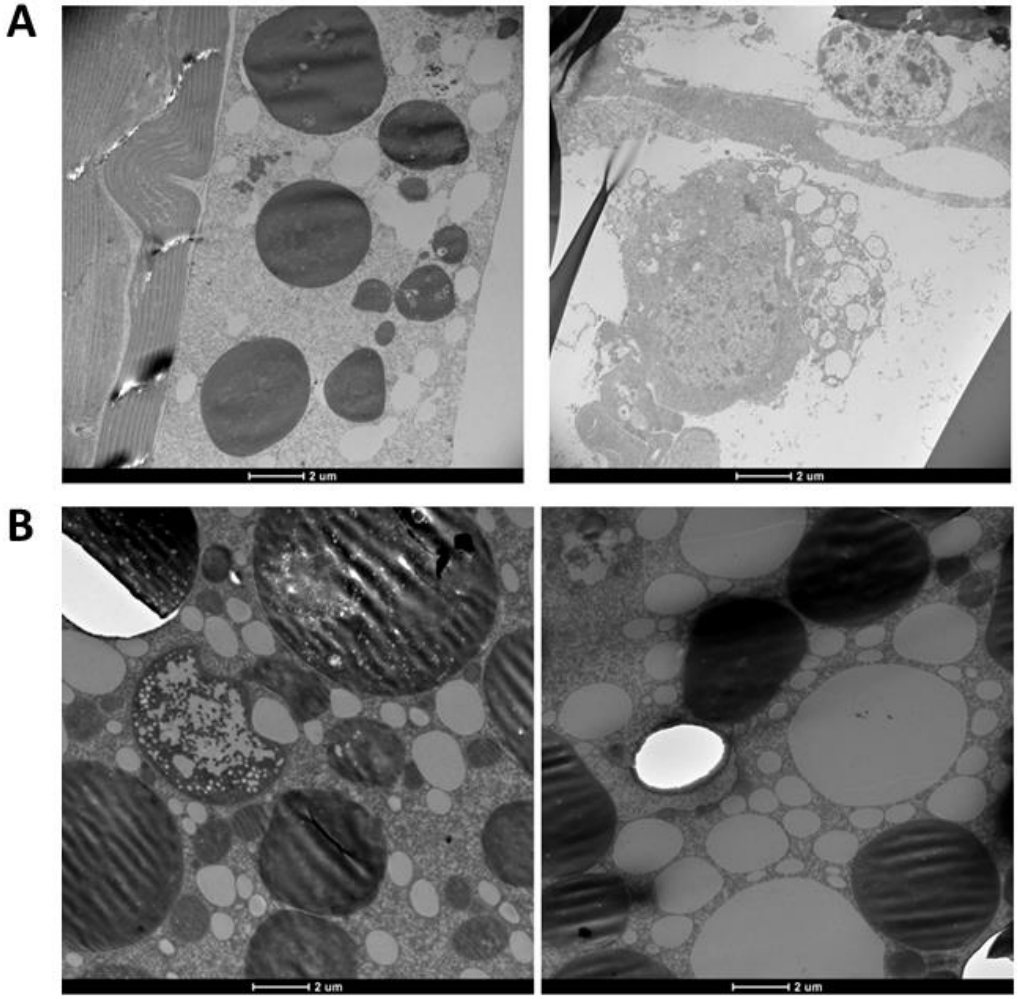


Figure 5.14. Transmission electron microscopy image of sections cut from whole body of *L. sericata* hosts stung by VZ-KD wasps. A: day 1, B: day 3. Presumably, the circular black spots represent oil droplets while the circular white spots represent the fat globules.

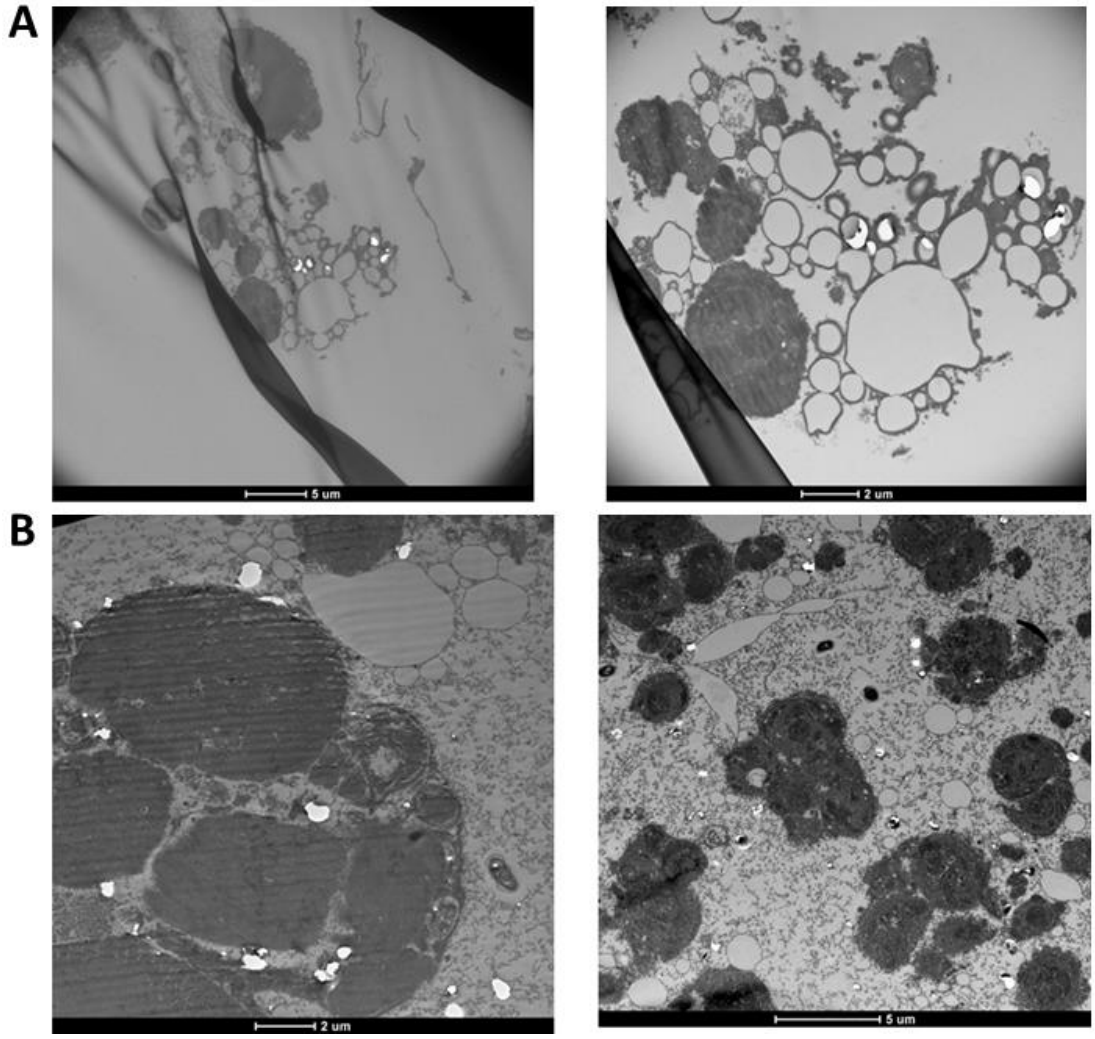


Figure 5.15. Transmission electron microscopy image of sections cut from whole body of *L. sericata* pupae stung by VX-KD wasps. A: day 1, B: day 3. Presumably, the circular black spots represent oil droplets while the circular white spots represent the fat globules.

5.3.5. Molecular cloning of VY, VZ and VX genes

L. sericata hosts that were stung by *Nasonia* wasps knocked down of venoms VY, VZ and VX showed different phenotypes when compared to the hosts stung by wild-type wasps (Table 5.6). Therefore, these three venom genes were selected as candidates for recombinant protein production. Most of the time was spent on preliminary experiments to find the expression vector and expression cells to use for recombinant production. Firstly, genes of interest were ligated into the pPROEX HTb plasmid and transformed into *E. coli* BL21 (DE3), C41 (DE3) and Rosetta™ (DE3) cells. However, inducing these

cells for overexpression of the genes of interest resulted in no growth as the venom was toxic to the cells. To avoid toxicity to the expression cells, the vector pET-OmpA HT MCS which contains a signal peptide that directs polypeptides to the periplasm was used. The use of pET-OmpA HT MCS helped to reduce the toxicity challenge however the proteins expressed were observed to be insoluble using both C41 (DE3) and Rosetta™ (DE3) expression cells. This insolubility challenge led to change of plasmid of choice to use the pET-DsbC HT Bam plasmid which fuse the recombinant protein to the chaperone, disulfide bond C (DsbC). DsbC catalyse the formation of protein disulfide bonds in the periplasm of *E. coli*. Therefore, fusing the gene of interest to DsbC has potential to improve solubility of recombinant proteins and the pET-DsbC HT Bam plasmid was used for small scale expression experiments.

RNA was successfully extracted from the venom reservoir and cDNA was synthesized using the primers that added *Bam*H1 and *Eco*R1 restriction sites to the end of the PCR products (Table 5.3). *Bam*H1 and *Eco*R1 restriction digests were carried out on the pET-DsbC HT Bam plasmid and the venom VZ, VX and VY cDNA. The restriction enzymes were shown to have cut the DNA templates based on the vector controls, which showed that when the plasmid was not cut it had a bigger size compared to the cut plasmid (Fig 5.16A). The VY, VZ, and VX cDNAs were then ligated into the cut vector and transformed into XL1 Blue cells. Transformation was confirmed through colony PCR targeting the genes of interest and the presence of amplicons of the expected size confirmed insertion of venom gene into the plasmid (Fig 5.16B). The colonies that were confirmed as carrying the cloned venom genes were cultured in Luria Bertani (LB) medium overnight under the selection of ampicillin. Plasmids extracted from the cells were then transformed into expression cells.

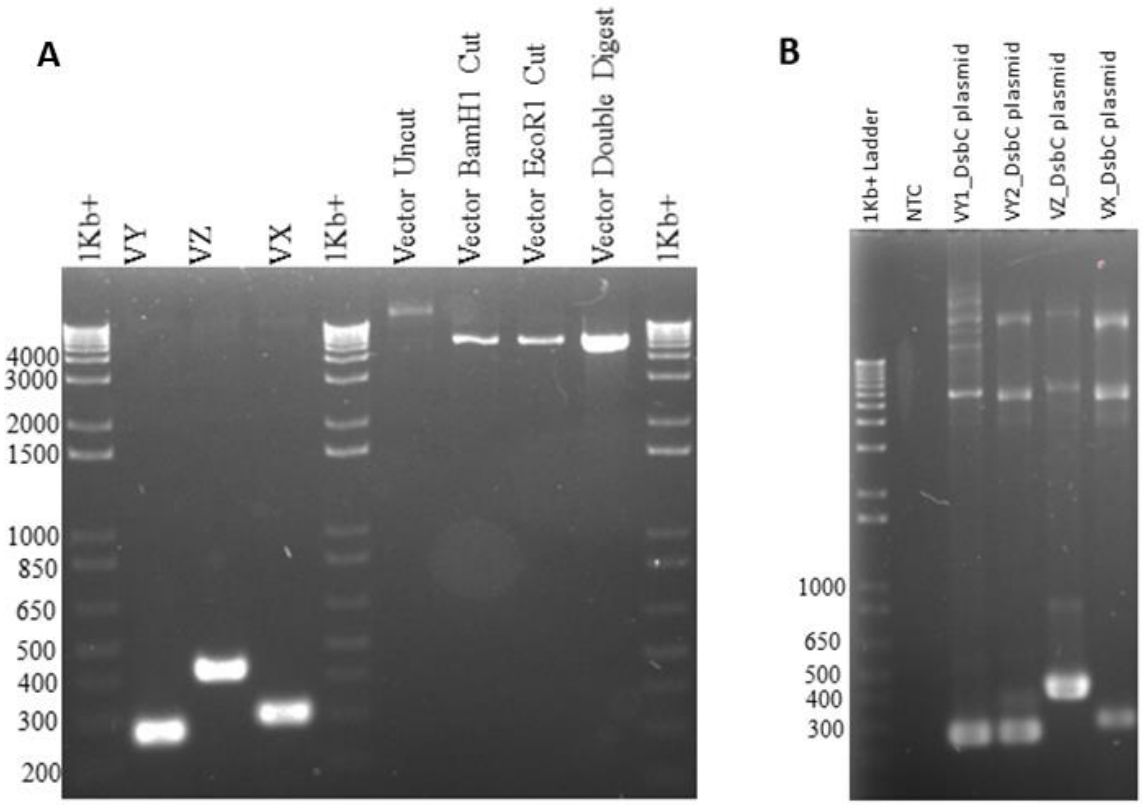


Figure 5.16. Amplification of venom genes and pET-DsbC HT Bam plasmid. A: Restriction digest confirming that both the restriction enzymes *Bam*H1 and *Eco*R1 cut the vector. B: Colony PCR targeting the genes of interest through PCR of whole colony and this showed the *Nasonia* genes were amplified from the colony DNA suggesting that the genes of interest were ligated into the plasmids. Expected sizes were VY – 263bp, VZ – 416bp and VX – 304bp. Ladder is 1kb plus (ThermoFisher Scientific) and the units shown are in bp.

The *E. coli* expression strain BL21 (DE3) that does not have a tight control over the expression was used as the first choice for expressing the recombinant proteins. However, immediately after IPTG induction, the cell density remained very low (OD_{600} of 1.2) which was four times smaller than that of the control cells that lacked the recombinant gene. This result indicated that the venom proteins were toxic to the expression cells therefore the C41 (DE3) strain of *E. coli* was trialled as an expression strain. These cells were developed to enable the expression of toxic proteins. To determine the best expression time, the *E. coli* C41 (DE3) transformed cells were induced for three and eight hours (Fig 5.17). Unfortunately, production of VZ recombinant protein still adversely affected cell growth and the OD_{600} dropped on induction, indicating cell death. The

remaining cells were pelleted by centrifugation then lysed and fractionated once more by centrifugation into soluble and insoluble fractions. SDS PAGE analysis of these showed that the VZ recombinant protein was only partially soluble.

When VY and VX were expressed in *E. coli* C41 (DE3) cells, they did grow, albeit slowly compared to the control cells. Although there was no overproduction of recombinant protein, tentative identification of recombinant VY and VX proteins in the insoluble fraction was based on their expected sizes (Fig 5.17). To counteract the insolubility challenge, the expression strain was changed to *E. coli* Rosetta™ (DE3) which is designed to enhance the expression of proteins that contain codons rarely used in *E. coli*. The transformed cells were only grown for three hours after induction as once more, the cells failed to thrive after induction, indicating the recombinant proteins were toxic to the cells in contrast to the control cells that contained the empty plasmid. Unfortunately, when the soluble and insoluble induced whole cell fractions were analysed by SDS PAGE there was no band of the expected size for the fusion DsbC-VY protein (37.1 kD). In the control cells on the other hand, DsbC was overexpressed after induction as seen by the intense band around the expected size of 32 kDa (Figure 5.18). When parallel experiments were carried out with plasmids containing the VX and VZ proteins similar results were obtained. The expected size of the DsbC-VZ chimera was 45 kDa and a slightly intense band is seen on the gel suggesting the protein to be partially soluble (Fig 5.19). The expected size of DsbC-VX was 40.1 kDa and a band of this size was also not obvious on the gel (Fig 5.20). Furthermore, the expected DsbC product was not represented in the control cells suggesting induction using IPTG was not successful for these cells. The exact reasons these proteins were not successfully produced by the heterologous host *E. coli* were unable to be determined.

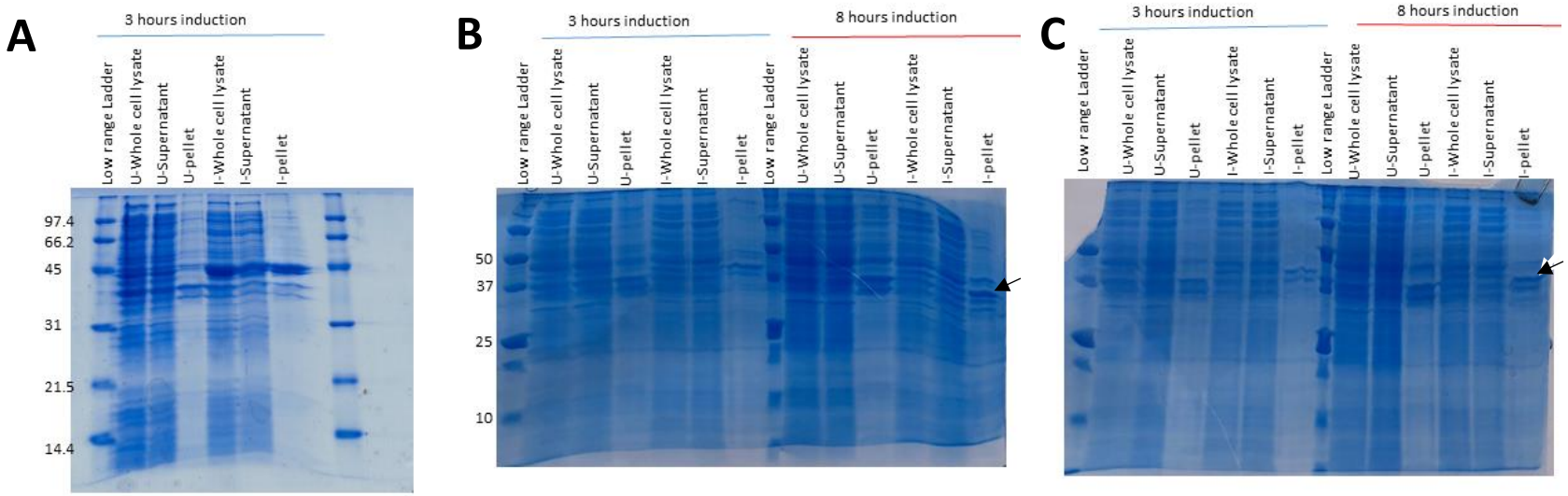


Figure 5.17. Expression time trials of *Nasonia* venom recombinant proteins (VY, VZ and VX) using C41 (DE3) expression cells. A: VZ production after 3 hours induction with IPTG. The expected size of the product was 45 kDa and was produced as a partially soluble protein. **B**: VY expression continued to 8 hours after induction. Insoluble protein was produced around the expected size of 37.1 kDa (black arrow). **C**: VX expression continued to 8 hours after induction however, the no clear overexpression was observed even though the expected protein size of 40.1 kDa (black arrow) was represented on the SDS PAGE image. U – uninduced and I – induced.

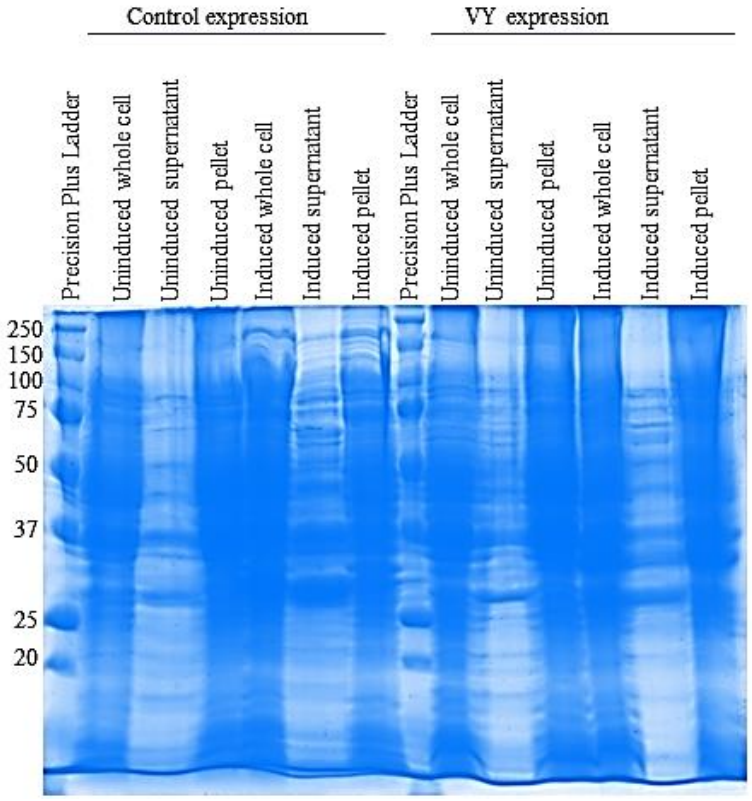


Figure 5.18. SDS-PAGE of VY expression using Rosetta cells. Control expression represent vector that had no VY insert. The expected size of vector only product (DsbC) is 32 kDa while vector and insert (DsbC-VY) is 37.1 kDa. The bands in the soluble fractions of both control and induced cells are similar indicating there was no overexpression of DsbC-VY protein.

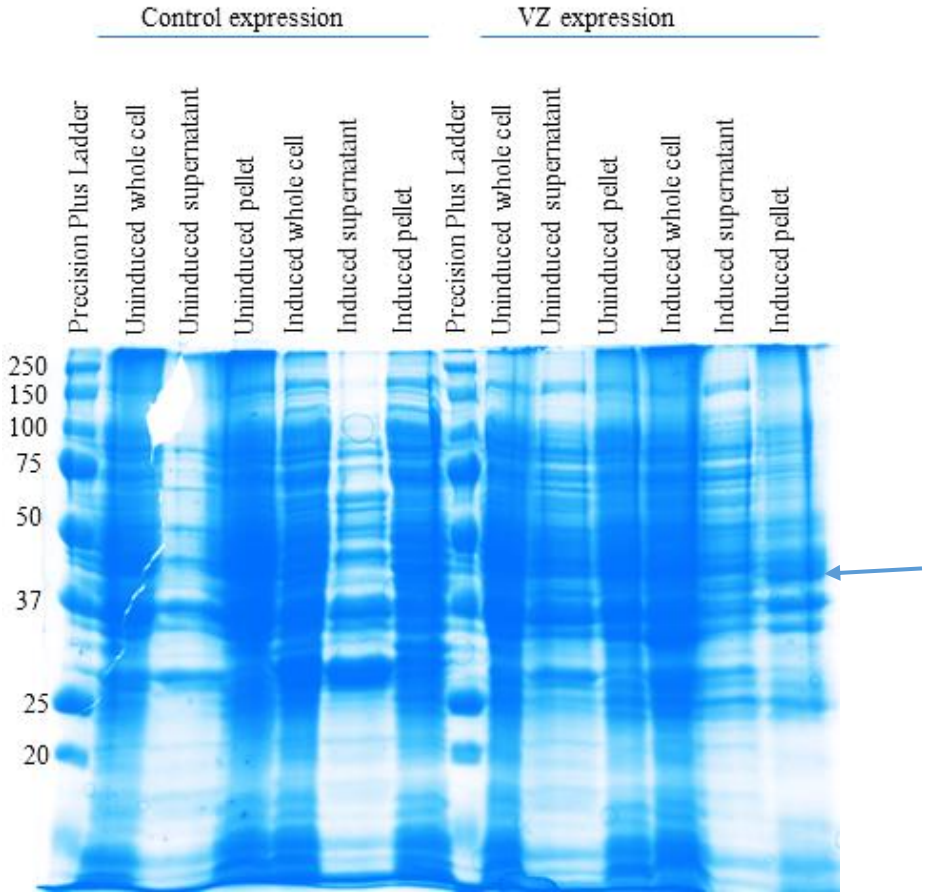


Figure 5.19. SDS-PAGE of VZ expression using Rosetta cells. Control expression vector that had no VY insert. The expected size of vector only product (DsbC) is 32 kDa while vector and insert (DsbC-VZ) is 45 kDa. The arrow shows the overexpressed protein with the expected DsbC-VZ product size as insoluble.

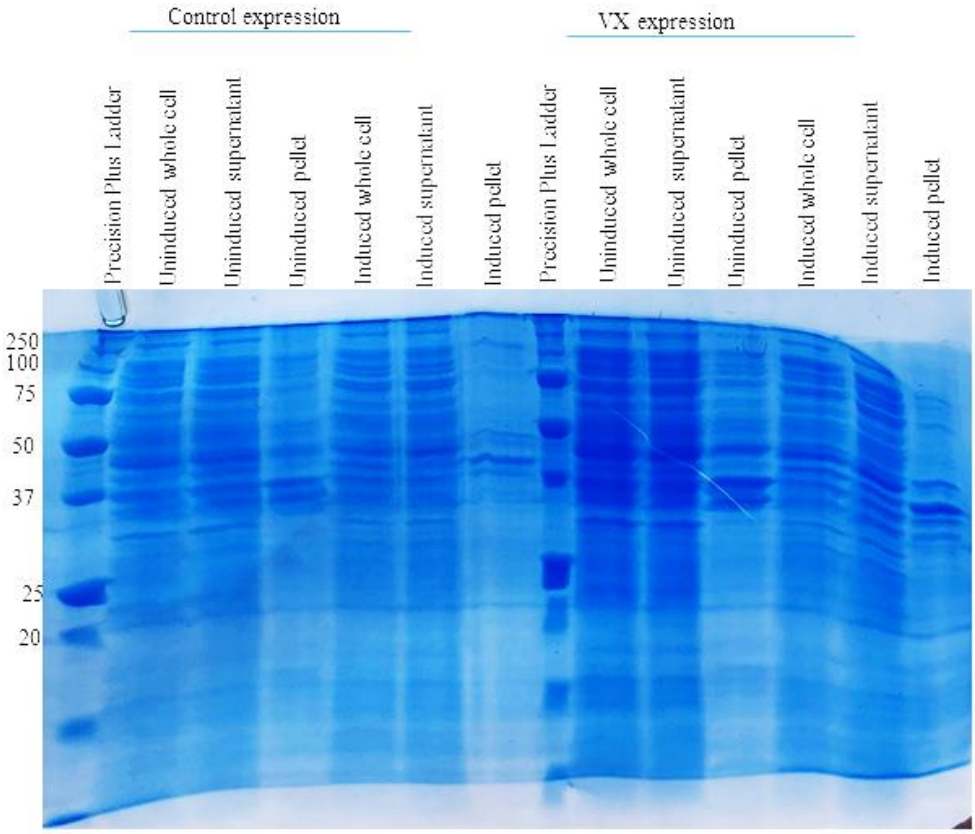


Figure 5.20. SDS-PAGE VX expression using Rosetta™ cells. Control expression represents a vector that had no VX insert. The expected size of vector only product (DsbC) is 32 kDa while vector and insert (DsbC-VX) is 40.1 kDa. This gel shows that there was no overexpression of proteins in both the control and experimental cells suggesting induction was not successful.

5.4. Discussion

Nasonia venom is made of a mixture of peptides and proteins that collectively manipulates the host's physiology. The host is not killed by the venom, rather is kept alive to act as a food supply for the developing offspring [122]. Previous studies attributed that manipulation is achieved through arresting the host's development, changes in cell defence pathways as well as increase of several metabolites [143,229,237,238]. The venom proteome comprises 79 proteins, with over 25% having no known homology to any other proteins [121] indicating that they are novel . Of the

venom proteins homologous to known proteins, pacifastin protease inhibitors and the kazal-type serine protease inhibitors have been found to inhibit the activation of prophenoloxidase in host haemolymph, which is a key enzyme in melanisation cascade and wound healing [148,183]. Calreticulin is proposed to suppress expression of the innate immune cell response [145]. To date there have been no reports that focus on characterizing the unknown venom proteins. A transcriptomic study that ranked the venom proteins in order of transcript concentration in the venom gland showed that the top five highly expressed genes are the genes with no homology to any known genes [231]. In this research, the top six most highly expressed genes in the venom gland were chosen for characterization.

The lack of homology of the six highly expressed *Nasonia* venom encoding genes was confirmed by searches against the NCBI, Pfam and Swiss-Prot databases. The exception to this finding was SP97, which had sequence homology to other insect serine proteases. As signal peptidases are necessary to direct proteins to specific organelles [239], the presence of the signal peptides in the venom proteins confirmed the six proteins in question were secreted proteins and in the case of *Nasonia*, most likely transferred to the venom reservoir after synthesis in the venom gland. Signal peptides are generally rich in hydrophobic residues, and alanine and serine residues usually make up the cleavage site [115], as observed for these six proteins. The branched chain hydrophobic amino acid residues have been found to control protein synthesis through translation initiation [240] and their presence in the venom proteins suggests that they are likely to play the same role in the *Nasonia* venom proteins. The pI is an intrinsic value of molecules and is defined as the pH at which the net charge of the molecule is zero [241]. The calculated pI values of the venom proteins were all below seven speculating that they are soluble in acidic environment. The venom gland (also referred to as the acid gland) is a reservoir with an acid environment thereby supporting the predicted calculations of the isoelectric points [242].

Because the unknown proteins had no sequence similarity to any other proteins in the data base, their structures were threaded using bioinformatic algorithms in order to try and predict the function of these proteins [243]. Secondary structures were predicted using the Phyre2 webserver

which utilizes three independent prediction programs, mainly Psi-Pred, SSPro and JNet [244–246]. The consensus prediction is generated by the scores given by the programs at each position of the query sequence and the consensus structure is scanned against the Phyre fold library, made up of solved protein structures. Interestingly, the predicted folds were assigned to protein families that allowed the functions of the unknown venom proteins to be predicted. *Venom Y* (VY) was assigned into the hydrolase family of proteins. These are hydrolytic enzymes including esterases, proteases, glycosidases and lipases. In venoms, hydrolases have been shown to process lipids, suppress the host's development and immunity [229,247] and *Nasonia* venom is known to suppress host development, immunity and increase lipid production. Therefore, based on the prediction, VY potentially plays an important role in host manipulation.

Venom Z (VZ) is predicted to be a transport protein. Newly synthesized proteins have to be transported from the endoplasmic reticulum and Golgi to organelles by vesicles such as endocytes [249]. In both pathways, there are signal and cognate-receptor-dependent transport steps and based on the prediction, VZ might be involved in transporting the venom proteins to the venom reservoir or through the stinging apparatus. Previous studies have shown that the venom genes are also expressed in low amounts in the ovaries of the wasp suggesting [231] that VZ might have other functions than simply acting as a venom protein. In bacteria, transport proteins have been found to serve as virulence factors and are involved in transport of other proteins across the outer membranes of chloroplasts and mitochondria [250–253].

The secondary structure prediction of *Venom X* (VX) classified the protein as a transcription factor or DNA-binding protein. DNA-binding proteins are involved in a variety of functions in cells such as transcription, repair and replication. A small DNA binding protein from *C. rubella* venom is thought to interfere with the transcription of immune-protective proteins [110] but little is known about the roles of DNA-binding proteins in venoms. However, based on secondary structure prediction, VX plays a role in manipulating the expression of genes in ways that supports successful parasitism. The secondary structure prediction for *Venom Q* (VQ) suggests that it is a lipid binding protein (LBP). Members of the LBP group have been found to transport lipophilic

ligands such as retinoids and long-chain fatty acids in the cytoplasm [254,255] as well as serving as regulators of retinoic acid metabolism during development [256]. The hosts stung by *Nasonia* appear to have increased lipid production, [147], and based on secondary structure prediction, VQ could have a role in the regulation of lipid metabolism in the host.

Serine protease 97 (SP97) and *Venom G* (VG) proteins were predicted to have folds similar to members of the protease family and were classified as serine and cysteine proteases respectively. Serine proteases have been found to have several roles in insects including the activation of the toll signalling cascade, the innate immune response, embryonic development as well as prophenoloxidase signalling [110,247,257]. The serine protease family is the most represented family in the *Nasonia* venom constituents with 15 members found out of the 79 venom proteins [121]. This overrepresentation of serine proteases suggests they play a significant role in successful host manipulation and poses a challenge in attributing function to SP97, as there is a higher possibility of shared functions with other venom proteins belonging to this family. The secondary structure prediction of VG as a cysteine protease suggests that this protein might have a role in tissue and cell invasion. In *Plasomodium*, *T. cruzi* and *E. histolytica*, cysteine proteases have been implicated in cell and tissue invasion [258–260]. Cysteine proteases also potentially interfere with the host immunity as they have been found to degrade immunoglobulin A [261].

Functional characterization of *Nasonia* venom proteins was complicated by the fact that they share little to no sequence homology to other known proteins [121]. Therefore, to counteract this challenge, genetic knock down of the genes encoding these proteins was achieved using RNA interference (RNAi). RNAi in *Nasonia* is systemic, meaning extracellular application of double stranded (dsRNA) *via* injection into larvae results in global knock down of the target gene [6]. The RNAi method has previously been successfully used to characterize the venom protein calreticulin as an inhibitor of the innate immune cell response [145]. The top six highly expressed genes in the venom gland were selected for RNAi and the dsRNA was confirmed to be specific to the target gene *via* RNAi off target prediction. This was important in reducing the possibility of any off-targeting effects that may produce an erroneous result. Previously, *Nasonia* wasps that had had the

calreticulin venom gene knocked down were observed to display a low stinging rate [145]. Fisher's exact tests on the stinging rate of the wild-type and KD wasps showed no significant difference suggesting that the KD did not affect this aspect of wasp behaviour.

Nasonia venom is known to cause observable phenotypes on the host which include suppression of development, immunity, and melanisation [122,248]. In this thesis, the venom genes that were knocked down were characterized on the basis of observable phenotypes in the hosts stung by the KD wasps. Hosts that were stung by wasps knocked down of SP97, VG and VQ showed a no development phenotype similar to hosts stung by wild-type wasp. Two possible reasons for no observable phenotypes after knocking down these genes are, firstly, the proteins encoded by these particular genes may not play a role in manipulating the host. Secondly, the overrepresentation of serine proteases in *Nasonia* venom strongly suggest that there is functional redundancy and hence knocking down one serine protease may not affect the overall function of the venom on the host. The same can be said for VG, which was also predicted to be a member of the protease family. Based on the structural prediction, VQ was predicted to be a lipid binding protein suggesting that the protein might play a role in the regulation of metabolites in the targeted host. Although metabolomic studies have shown that *Nasonia* venom influences a number of biochemical pathways in the host *S. bullata*, such as the sorbitol pathway, the TCA cycle and chitin biosynthesis [147], no observable phenotype was seen in hosts stung by VQ.

Hosts that were stung by wasps knocked down of VY gene showed a marked increase in melanisation at the sting site. Melanisation in insects is due to phenoloxidase cascade activity of the innate immune system [131]. Successful parasitism relies on the ability of the venom to suppress the host's immune system and parasitoid venoms are known to achieve this by inhibiting both host encapsulation response and phenoloxidase activity [131,229,262]. VY was bioinformatically predicted to be a hydrolase and hydrolases have been shown to play a role in suppressing the host immune system [229]. It is therefore quite possible that VY plays a role in inhibiting the melanisation process in the host limiting the innate immune response in *L. sericata*.

To confirm this, more work needs to be carried out to quantify the levels of increased melanisation in hosts envenomated by *Nasonia* lacking VY.

Over 60 % of hosts stung by wasps in which VX was knocked down developed eye pigments and bristles after five days indicating that VX might play a role in suppressing the development of stung hosts. As phyre2 predicted a structure for VX that had some similarity to a DNA binding protein, it is possible that VX might regulate transcription. Other venoms have been found to disrupt target development by disrupting hormonal or neurohormonal pathways [263] raising the possibility that VX could play a role in regulating the expression of peptide hormones. Based on the eye pigment development phenotype of the hosts stung by VZ-KD wasps, VZ plays an essential role in limiting the development of *L. sericata*. The venom protein VZ was predicted to be a transport protein and could be involved in lipid transport, immunoregulation and modulation of cell growth and differentiation [264]. Based on the slight development phenotype shown by VZ-KD stung hosts, there is a suggestion that VZ might have a role in the modulation of cell growth. Furthermore, the hosts stung by VZ-KD wasps seemed to have an increased oil production suggesting that VZ may be involved in the regulation of lipid metabolism. Sequence analysis predicted that VZ is a transport protein. A previous study showed that apolipoprotein E, a plasma transport protein, transports cholesterol and other lipids between various cells [264]. The phenotype observed for this knockdown support the functional prediction that VZ is involved in the regulation of lipid production. Another interesting observation was that although VZ-KD wasps could sting hosts, they did not lay eggs. Previous studies had shown that VZ is also expressed in the ovaries as well as the venom gland [231] and therefore it is not unreasonable to suggest that the VZ gene plays a role in either egg formation or the process of laying eggs. The RNAi effect is systemic in *Nasonia*, so dsRNA should also target VZ expressed in other tissues. Despite the signs of development observed in the hosts stung by VZ-KD wasps, all pupae did not survive longer than 5 days after envenomation strongly implying that the removal of VZ is not enough to prevent the action of other venom induced processes.

The three venom KD experiments that showed phenotype changes were further analysed by examining the sections of the host's whole body pupae using TEM microscopy (Fig 5.5). As it had previously been shown that *Nasonia* venom is able to regulate lipid metabolism in its target pupae [265], the experiments were focused on observing the fat globules and oil droplets around the sting site that are normally spherical in structure [266–268]. The negative control (unstung hosts) and the positive control (wild type stung hosts) were speculated to have decreased and increased numbers of fat globules and oil droplets respectively based on the images observed on the same magnification. These preliminary analyses predict a direct relationship between the decrease in lipid content and pupae development. Although hosts stung by the wasps lacking the VY gene showed no development phenotype there was an increase in the number of fat globules and oil droplets around the sting site after three days as seen in the TEM images. Hosts that were stung by VX-KD wasps on the other hand showed a decrease in lipid content which was a similar to what was seen in unstung hosts suggesting that host development was taking place. In addition, the oil droplets appeared to become deformed and degraded three days after being stung, suggesting that VX is a peptide that supresses development of stung hosts. Interestingly, hosts stung by VZ-KD wasps were characterized by a contrasting phenotype *i.e.* red eye pigment consistent with development and an oily surface showing uncontrolled lipid production. TEM analysis of the sections cut from whole body of the host's pupae also showed the size of fat globules and oil droplets had increased, suggesting a change in lipid metabolism. Furthermore, the oil droplets appeared to be degrading after day 3 of the hosts being stung also supporting the phenotype observed. As the VZ-KD stung hosts did not develop to adults, this is most likely to be a result of uncontrolled lipid production in the host, although this study can't rule out the lipid build up was not due to other processes, such as cell death. Based on these findings, the VZ protein may be involved in regulating lipid metabolism of the stung hosts. Overall, preliminary TEM analysis showed changes in the number of fat globules and oil droplets suggesting that lipid content can be used as a marker for host development.

To validate the functions of the venom proteins VY, VZ and VX that caused change of phenotype in the host, trials to produce recombinant protein were conducted. Recombinant technology has

been used before to characterize Hymenoptera venoms including those of *Nasonia* [148]. Kazal-type serine protease inhibitors have been characterized in previous studies through molecular cloning and have been found to inhibit the phenoloxidase activity in *Musca domestica* [148]. Unfortunately, despite much effort, soluble recombinant protein could not be produced in this study. Many different *E. coli* expression strains were trialled including BL21 (DE3), C41 (DE3) and Rosetta™ cells and different expression vectors including pProEX HTb, pET-OmpA HT BAM and pET-DsbC HT MCS. OmpA and pET32 HTC contain a leader sequence for export of the recombinant product into the periplasm. The pET32 vector produces the cloned protein as a fusion with the disulfide isomerase (DbsC) in the *E. coli* expression strains used. Transforming the plasmid into the C41 (DE3) expression system partially solved the toxicity challenge as the cells grew, albeit slowly. The resultant recombinant proteins produced, were however insoluble. These results show that *E. coli* are most likely not the best heterologous host for the expression of the venom proteins VY, VZ and VX. An alternative way of producing recombinant protein would be to use a eukaryotic expression system which have different strategies for processing secreted proteins. Yeast (*Pichia pastoris*) has been successfully used to produce a recombinant spider venom neurotoxin, which was purified, and had biocidal activity against the tomato moth (*Lacanobia oleracea*) [269]. The same yeast expression system has been successfully used to produce active recombinant snake and snail venom proteins that had blood coagulation and insecticidal activities [270,271]. The baculovirus expression system is another alternative that could be used to express the *Nasonia* venom genes as it has been found to produce biologically active proteins from any foreign gene. Although this appears to be attractive, the fact that the recombinant protein is expressed in insect cells probably excludes it from contention as a heterologous host. Future experiments should look at the yeast systems with the aim of producing enough pure recombinant protein for structural studies, bearing in mind these sequences have no homology to any other protein. Because of the size of VX, VY and VZ, it should be possible to solve their structures using NMR methods. To do this requires special culture media, which in itself can be a problem for culturing cells. Thus, crystallisation should be trialled in the first instance to see if diffraction quality crystals can be obtained.

5.5. Conclusion

This chapter describes functional characterization of six highly expressed venom encoding genes in *Nasonia*. I successfully managed to knock down all the target genes based on RT-PCR results and showed that RNAi does not affect the stinging rate of the wasps. The bioinformatics predictions on the functions of the venom proteins VY, VZ and VX were somewhat supported through RNAi experimental assays. Putative functions assigned to venom proteins in this thesis are: (1) VY plays a significant role in suppressing the innate immunity of the host (2) VX is directly involved in developmental suppression of the host and (3) VZ has a number of roles which are: suppressing development and regulation of lipid production in the host and also has an effect on the ability of wasps to lay eggs after envenomation. Nevertheless, transcriptomic studies to determine gene changes due to the knock down will be useful to pin point the direct pathways affected by these individual venom proteins. TEM analysis showed a relationship between lipid content with development thereby suggesting host lipid content can be used as a marker for host development. Increase in lipid content shows no development while a decrease is associated with development. VX and VZ were observed to be involved in suppressing development through the analysis of lipid content.

Bioinformatic analyses suggested that venoms SP97, VG and VQ may have roles in suppressing the immunity system and lipid production in the host. As RNAi knockdowns could not confirm these predictions it is likely that functional redundancy within the venom proteins results in upregulation of the other genes with the same function as the impaired gene as a response to a lack of function of that gene. However, functional redundancy and lack of observable phenotype do not rule out the possibility that these genes have other functions that are not phenotypic, such as the regulation of metabolites. Therefore, high-throughput analysis of temporal changes of metabolites in the host might reveal other hidden functions of these venom proteins.

This thesis investigation has revealed a number of possible venom functions thereby opening up avenues for future research on the applications of the venom proteins especially in drug research. For instance, in cancer research, VX could prove to be useful in suppressing the growth and spread

of cancer cells. The effects of *Nasonia* venom has on host metabolism could be relevant diabetes research as it has been found that sorbitol levels are increased without changing the glucose level. Further, the approach used in predicting function could be applied elsewhere to characterize function of individual proteins.

6. GENERAL DISCUSSION, CONCLUSIONS AND FUTURE WORK

6.1. General Discussion

Nasonia has emerged as a model organism for answering various biology questions due to its important features such as short generation time, ease of handling, RNAi and availability of visible and molecular markers [272]. In the research carried out in this PhD project, these *Nasonia* characteristics were used to examine the mechanism that underlies cytoplasmic incompatibility as well as characterizing the selected venom proteins. The aims of this project were firstly, to identify the genes regulated by *Wolbachia* to manipulate the reproductive system of the wasp through cytoplasmic incompatibility (CI). Secondly, this study also aimed to characterize functionally the top six highly expressed genes in the venom gland of *Nasonia*.

6.1.1. Aim 1 – Characterization of the cytoplasmic incompatibility mechanism

The bacteria *Wolbachia* manipulates the *Nasonia* reproductive system through CI so that it can be spread rapidly in the population. However, the mechanism through which it causes CI remains unknown. In this study, it was predicted that infection of *Nasonia* with *Wolbachia* results in changes in gene expression that directly lead to the development of CI. RNA-seq showed that 84 *Nasonia* ovary and 58 *Nasonia* testis genes are regulated by *Wolbachia infection,* the first such study to do so. Annotation of these genes showed that genes involved in cell division and catalytic activities in female and male wasps were mainly affected.

A previous study that investigated the influence of *Wolbachia* on *Drosophila mauritana* showed that flies that are infected by the bacteria produce four times more eggs as compared to their uninfected counterparts [273]. In addition, an increase in mitotic activity of germline stem cells and decrease of programmed cell death in the germarium was observed in *Wolbachia*-infected *Drosophila* [273]. This result suggested that development of germline stem cells was influenced by *Wolbachia* infection and the results in Chapter 3 suggest that the bacteria cause alterations in gene expression that may regulate cell division in *Nasonia* ovaries. In the testis, genes that played

a role in catalytic activity were differentially regulated in *Wolbachia*-infected *Nasonia*. In a previous study in *Drosophila* larvae, microarrays and qRT-PCR were used to show that the expression of many genes annotated as having 'catalytic activity' were disrupted after in testis infected with *Wolbachia* [274]. Furthermore, it has been shown that in *Wolbachia* infected males, there are multiple altered sperm structures at the late stage of spermatogenesis, suggesting *Wolbachia* causes genes expression changes that lead to sperm modification [275]. The findings of this PhD project together with the evidence suggested in previous studies, show that *Wolbachia* plays a huge role in influencing spermatogenesis in the host.

A wide range of invertebrates including up to 66% of insect species as well as nematodes, spiders and mites are infected by *Wolbachia* [276]. Some of the insects that harbour *Wolbachia* are pests of agricultural crops and vectors (mosquitoes) or causative agents (filarial nematodes) of devastating infectious diseases [277–279]. Therefore, understanding the mechanism underlying CI is very important as *Wolbachia* could be used to combat some of the agricultural and human diseases. Personally, this PhD project has been of great importance as many people in my home country (Zimbabwe, Africa) die due to vector-borne diseases such as malaria, dengue fever and schistosomiasis. Therefore, personally the understanding of CI mechanism may not only close the research gap, but also actually help to find health solutions with the potential to save many lives including those of my family and friends.

Traditional methods used to control the spread of these vector borne viruses such as insecticides, are becoming less effective indicating that there is need to develop alternative methods. Since *Wolbachia* infects some of the insects that are causative agents for these diseases, it has become a choice for alternative approach to control the populations of causative agents such as mosquitoes. The principle of CI used by the bacteria is that infected males cannot produce viable offspring with uninfected females, therefore releasing infected mosquito males in the environment will significantly reduce the population of mosquitoes [280]. Experiments where *Wolbachia* infected mosquito males were released into a controlled environment resulted in massive reduction of the mosquito population as well as the reduction of the incidence of dengue fever and Zika virus infection in the same area [281–283]. Therefore, CI offers a potential huge role in controlling

vector-borne related viral infections in humans and understanding the mechanism in which the bacteria achieve this phenomenon will be key to successful control.

In this study RNA-seq was used to identify a set of genes that are differently regulated by *Wolbachia* infection in both the ovary and testes. Five and two of the top differentially expressed genes in the testes and ovaries, respectively, were individually targeted for RNAi knock down (KD) experiments. All of these genes were down-regulated following *Wolbachia* infection, so the idea was to mimic this down regulation in uninfected wasps creating an RNAi induced CI. The expected results were to generate only male progeny when KD uninfected males mate with uninfected females or expect female progeny when infected males mate with KD uninfected females. Unfortunately, RNAi experiments could not mimic the CI effects caused by the bacteria in the wasps and in this study. A number of experimental limitations could have been responsible for this result.

I. The *Nasonia* genome assembly official gene set that was used to map the RNA-seq reads was generated through comparisons with the eukaryotic genomes of *Apis mellifera*, *Tribolium castaneum*, *Drosophila melanogaster*, *Pediculus humanus*, *Daphnia pulex*, and *Homo sapiens* [86]. Therefore, as these species are not subject to CI it is possible genes involved in the *Nasonia* CI response were not annotated in the official gene set and thus their expression could not be assayed using our method. Previously, 472 *Nasonia* putative novel genes were identified through RNA-seq of the venom gland and ovary tissues, supporting the fact that the official gene set model is not complete, and also suggests that genes expressed only in the reproductive organs could be underrepresented [231]. Future work would involve testing to see if regions where RNA-seq reads mapped outside of the official gene set could represent functional loci. It would also be possible to investigate the role of such genes using packages such as Cufflinks, as this software does not require gene models; rather they generate their own based on the transcriptome data itself.
II. While RNAi remains a reliable technique for studying gene functions, the method also comes with technical and biological limitations that can complicate the interpretation of

KD phenotypes. Technical limitations include injection of varying amounts of the dsRNA into each wasp, potentially causing a range of KD phenotypes due to the amount of dsRNA introduced into each individual. Biological limitation could have arisen because RNAi is systemic in *Nasonia*; when injected, dsRNA targets mRNA throughout the whole body, which may change wasp behaviour in ways that complicate the identification of venom or CI KD phenotypes. In the future, RT-qPCR could be used to quantify the number of transcripts of specific genes, thereby clearly comparing knock down between uninfected and infected wasps.

III. Functional redundancy occurs when two or more independent genes perform the same function in an organism. When there is redundancy the KD phenotype will be masked by the second gene [284]. This is perhaps the most likely reason for not identifying the genes involved in CI. Future studies should include analyses of gene networks and their regulation which may help to characterize the CI mechanism..

6.1.2. Identification of lateral gene transfers

The RNA-seq data was generated from RNA extracted from reproductive tissues of infected and uninfected *Nasonia*. As a control measure, the uninfected sample reads were mapped to the *Wolbachia* genome with an expectation that there would be no mapping of reads. Interestingly, four *Wolbachia* genes were shown to be expressed in all uninfected *Nasonia* samples suggesting that the observation was either due to contamination or to the transfer of bacterial genes from *Wolbachia* to the wasp as lateral gene transfers (LGTs). To rule out contamination, independent RNA-seq data from uninfected *Nasonia* wasps was downloaded from the NCBI SRA repository, which confirmed that these genes are indeed expressed. Further, a BLASTn search was carried out using the *Wolbachia* genes as a query to search against the database of assembly reads (raw data) that were used to assemble the *Nasonia* genome. Using this approach eight *Wolbachia* genes (including the four from RNA-seq data) were shown to be present in the *Nasonia* genome. The assembly reads were independently generated from the DNA of uninfected *Nasonia* wasps which would tend to rule out contamination as a source of these bacterial gene sequences. Of the eight LGTs, two have already been described, while only one was found incorporated in the current

Nasonia assembly. This suggests that the rest of the reads that contain the other seven LGTs were discarded from the assembly.

LGTs can be an important source of new genes and functions to the recipient organism. Glycoside hydrolase family 19 (GH19) chitinase for instance was recruited as a venom in parasitoid wasps including *Nasonia* [87]. GH19 has been observed to regulate immune genes therefore as a *Nasonia* venom protein; it suppresses the immune system of the host thereby bringing a novel function to the wasp venom. This study has identified five potential new LGTs from *Wolbachia* to *Nasonia* and one artefact in the bacterial genome where a wasp gene has been mistakenly included. In the future, more experiments to investigate the expression and function of these genes should be conducted in order to investigate the function and benefits of these LGTs to the bacteria and/or the host. Besides the time constraints, the methods used to investigate the LGTs were also associated with limitations which are listed below.

I. The approach used can only detect bacterial genes that are expressed (RNA-seq) or that are currently incorporated in the *Nasonia* genome (BLASTn).
II. *Wolbachia* cannot be independently cultured. As a result harvesting of the bacterial DNA has to be done together with the host DNA thereby increasing the chances of contamination. The presence of a eukaryotic gene in the *Wolbachia* genome shows that artefacts create issues with the confident identification of LGTs.
III. BLASTn search results are totally dependent on the information that was entered in the database being searched against. If a sequence is added to the database with an incorrect annotation then any hits to this sequence will also be given this annotation. This means that bacterial sequences are sometimes labelled as eukaryotic (and *vice versa*) because they are sequenced as part of eukaryotic genome projects.

6.1.3. Aim 2 – Functional characterization of *Nasonia* venom peptides

Venoms and toxins are known to have high specificity and potency for their targets therefore have been used as drug targets for various human diseases. Spider venom peptides for instance, have been found to selectively modulate the activity of therapeutic targets such as purigenic receptors,

which are important for antithrombotic drugs and also they have antimalarial and antimicrobial activity [285]. Snake venoms also act as a rich source of bioactive molecules with pharmacological activities with the snake bile and blood being used in Chinese medicine; for example disintegrins have been shown to be powerful platelet aggregates [286]. A previous study revealed that *Nasonia* venom has anti-inflammatory activity on mammalian cell lines through prolonging c-Jun N-terminal kinase (JNK), a major signal transduction pathway involved in inflammation, activation [124]. The specific venom peptide that has this therapeutic activity is, however, unknown, and in general, the *Nasonia* venom peptides are still to be functionally characterized. In this PhD project, six top expressed venom genes in the venom gland identified through deep transcriptome sequencing in a previous study were selected for characterization. [231].

RNAi experiments showed that genes VZ and VX play a role in inhibiting growth of the envenomated hosts while gene VY supresses the immune system (Chapter 5). Knock down of VG, VQ and SP97 did not show any visible phenotype changes indicating that the venom proteins could be involved in other cellular activities such as regulating calcium intake or facilitating nutrient production that is vital for *Nasonia* offspring. In addition, since this was not a complete knockout, the residual protein might have been sufficient for normal function. For the first time, this study has identified the *Nasonia* venom protein (VX) that is directly involved in stopping host larval development.

The venom proteins that caused observable host phenotypic changes (Chapter 5) were expressed in *E. coli* so that the recombinant protein could be purified. Expressing the venom genes VZ, VX and VY in *E. coli* resulted in decline of growth of the cell population due to the toxicity of the recombinant venom proteins, suggesting that the *Nasonia* venom peptides have antimicrobial activity. While the antimicrobial activity phenotype observed could play a major role in pharmaceutical therapy, for this study it hindered the progress and ability to produce soluble recombinant proteins. All the cloned produced proteins that were insoluble in *E. coli* therefore, in the future, choosing eukaryotic cells as expression hosts might counteract the challenge faced in this study. The methods used in the chapter had limitations which are described below.

I. While RNAi has been efficiently used to identify the functions of genes, it still has limitations and, in this study, not all the hosts stung by KD wasps showed the same phenotype. For instance, 39 out of the 63 VY-KD stung hosts showed increased melanisation while the remaining ones did not show any differences in comparisons to the positive controls. This result could have been due to differing amounts of dsRNA injected into each *Nasonia*, which is a technical limitation of the method when used on these small insects. In addition, the assay method used is only useful when the targeted gene generate visible phenotypes in the envenomated host.

II. Gene functional redundancy is another major limitation in this project as knocking down one gene can be compensated by the presence of other genes performing the same function.

III. Expressing *Nasonia* venom genes in a bacteria cells was very detrimental to the growth of the cells. Changing the vectors did not efficiently solved the toxicity challenge suggesting that not all venom proteins were being exported to the periplasm (the system was leaky).

IV. Soluble proteins could not be obtained despite trying out several vectors and expression cells suggesting that possibly, post-translational modifications that are important to eukaryotic venom proteins could not be carried out in the bacterial cells and this could have caused insoluble proteins to be obtained.

6.2. **Future Work**

Several areas of future work can be suggested from this thesis. These are indicated below.

- Previous studies together with this project have shown that the *Nasonia* genome annotation and assembly is still a work in progress. Therefore, RNA-seq data will be very useful in the annotation of genes as well as improving the current *Nasonia* gene set taking into account alternative isoforms and novel genes not found in other insects lacking in CI.

- As qRT-PCR is the benchmark for the detection and quantification of RNA targets, validating RNA-seq results using qRT-PCR in this project will help to eliminate some of the doubt about the results obtained. Semi-qRT-PCR could be replaced with qRT-PCR to

quantify the number of transcripts instead of relying on the intensity of the amplicons on the agarose gel image. RNA-seq data showed that the gene *NV_15393-RA* is down regulated in the testis. While semi-qRT-PCR could not validate this result, measuring the actual number of transcripts could be a better validation method.

- Considering the fact that targeting individual genes to understand the CI mechanism was unsuccessful, possibly because a group of genes works together to cause the phenotype, gene co-expression network analysis could be a new direction to follow.

- This study has identified five more LGTs whose functions were not described. Further experiments on gene functions using RNAi or recombinant technology would explain the importance of these LGTs to the wasp or the bacteria.

- RNAi of VZ gene in *Nasonia* revealed that the wasps lost the ability to lay eggs when they sting the host so further analysis on the cause of this phenotype is necessary. Initially, dissections to extract the ovaries and examine them microscopically to see if the wasps' ovaries are producing the eggs and to examine if there are deformations in the reproductive system.

- The RNAi method chosen to characterize venom genes SP97, VG and VQ did not show any phenotype, therefore in the future combining RNAi with RNA-seq of hosts will help in determining whether there are changes in gene expression in the host. This will involve analysis of differential gene expression for unstung hosts, wild type stung hosts and KD stung hosts to show the group of genes affected by knock down of the individual venom gene.

- Expressing some of the venom protein in *E. coli* had a negative impact on the growth of *E. coli* showing that there is need to express them in vectors that tolerate toxic recombinant

proteins. Eukaryotic systems using yeast and baculovirus have been successfully used to express bee recombinant venom proteins [287,288] and this system can be adopted too.

6.3. **Conclusions**

Nasonia has previously been successfully used as a model organism to answer various biological questions, and in this study it was used to investigate the mechanism underlying CI as well as to characterize venom genes. Eighty-four and 58 genes in the ovaries and testes, respectively, were found to be regulated by *Wolbachia* during CI. The individual genes that are directly involved in CI unfortunately could not be determined using RNAi, however, these findings have opened new research avenues to be pursued in answering questions about the mechanism of CI. Based on the GO terms annotation, the genes in the testes regulated by the bacteria are mainly involved in catalytic roles supporting the speculation that *Wolbachia* regulates spermatogenesis through changes in gene expression.

Enzymes play an important role in ensuring successful fertilisation of eggs. Regulating these genes will therefore allow *Wolbachia* to manipulate the wasps' reproductive system. In the ovaries, the genes that were found to be regulated by the bacteria had roles mainly in cell division supporting the findings of previous studies that *Wolbachia* regulates mitotic activities in other species such as *Drosophila*. Despite the evidence from previous studies that genes involved in catalytic roles and cell division play a significant role in CI, targeting these genes individually was not enough to determine the genes involved in CI. Based on these findings the project opens future research to use gene network analysis in characterising the mechanism of CI.

The RNA-seq analysis revealed *Wolbachia* genes that were being expressed in *Nasonia* leading to discovery of additional bacterial LGTs in the wasp. Further analysis of the LGTs showed an artefact in the *Wolbachia* genome where *gene1120* was mistakenly incorporated in the genome. BLASTn and phylogenetic analysis provided evidence that *gene1120* is a eukaryotic gene. The LGTs proteins found are novel and present a challenge in functional characterization as well as genome assembly and annotation as this process relies heavily on comparisons with other known genomes and genes. Due to time constraints, LGTs could not be functionally characterized but in

the future experimental approaches such as RNAi would be a good starting point to characterize the genes. In addition, these observations show that *Nasonia* can also be used as a model organism to study host-parasite LGTs.

The putative functions of three *Nasonia* venom genes were determined in this study where VX and VZ were shown to be involved in suppressing stung host development whereas VY is likely to supress the immune response. These experiments showed for the first time the venom peptides that are directly involved in suppressing the host development. The venoms SP97, VG and VQ, however did not show any visible phenotypes after RNAi suggesting that these genes could be involved in other cellular activities that do not lead to visible phenotypes. Expression of venom recombinant proteins showed that *Nasonia* venom has antimicrobial activity. Despite the challenges faced in trying to characterize the venom peptides, this study has helped to close the research gap significantly and furthermore, this method could be applied to other venom proteins as well as other proteins from other species.

References

[1] D.C. Darling, J.H. Werren, Biosystematics of *Nasonia* (Hymenoptera: Pteromalidae): Two New Species Reared from Birds Nests in North America, Ann. Entomol. Soc. Am. 83 (1990) 352–370.

[2] D.B. Rivers, T. Crawley, H. Bauser, Localization of intracellular calcium release in cells injured by venom from the ectoparasitoid *Nasonia vitripennis* (Walker) (Hymenoptera: Pteromalidae) and dependence of calcium mobilization on G-protein activation, J. Insect Physiol. 51 (2005) 149–160.

[3] B.J. Velthuis, W. Yang, T. Van Opijnen, J.H. Werren, Genetics of female mate discrimination of heterospecific males in *Nasonia* (Hymenoptera, Pteromalidae), Anim. Behav. 69 (2005) 1107–1120.

[4] R.F. Weston, I. Qureshi, J.H. Werren, Genetics of a wing size difference between two *Nasonia* species, J. Evol. Biol. 12 (1999) 586–595.

[5] J. Gadau, R.E. Page, J.H. Werren, Mapping of hybrid incompatibility loci in *Nasonia*, Genetics. 153 (1999) 1731–1741.

[6] J.A. Lynch, C. Desplan, A method for parental RNA interference in the wasp *Nasonia vitripennis*., Nat. Protoc. 1 (2006) 486–494.

[7] R. Raychoudhury, L. Baldo, D.C.S.G. Oliveira, J.H. Werren, Modes of acquisition of *Wolbachia*: Horizontal transfer, hybrid introgression, and codivergence in the *Nasonia* species complex, Evolution (N. Y). 63 (2009) 165–183.

[8] R. Raychoudhury, C. a Desjardins, J. Buellesbach, D.W. Loehlin, B.K. Grillenberger, L. Beukeboom, T. Schmitt, J.H. Werren, Behavioral and genetic characteristics of a new species of *Nasonia*., Heredity (Edinb). 104 (2010) 278–288.

[9] J.A. Breeuwer, R. Stouthamer, S.M. Barns, D.A. Pelletier, W.G. Weisburg, J.H. Werren,

Phylogeny of cytoplasmic incompatibility micro-organisms in the parasitoid wasp genus *Nasonia* (Hymenoptera: Pteromalidae) based on 16S ribosomal DNA sequences., Insect Mol. Biol. 1 (1992) 25–36.

[10] J. Van Den Assem, F. Jachmann, P. Simbolotti, Courtship behaviour of *Nasonia vitripennis* (Hym ., Pteromalidae): Some Qualitative , Experimental Evidence for the Role of Pheromones, Behaviour. 75 (1980) 301-307.

[11] J.A. Lynch, The expanding genetic toolbox of the wasp *Nasonia vitripennis* and its relatives, Genetics. 199 (2015) 897–904.

[12] B.Y.D.A. Levin, of BIOLOGY, 48 (2015) 3–15. doi:10.1177/0959354311432905.

[13] J.A. Lynch, A.E. Brent, D.S. Leaf, M.A. Pultz, C. Desplan, Localized maternal orthodenticle patterns anterior and posterior in the long germ wasp *Nasonia*, Nature. 439 (2006) 728–732.

[14] O. Niehuis, J. Buellesbach, J.D. Gibson, D. Pothmann, C. Hanner, N.S. Mutti, A.K. Judson, J. Gadau, J. Ruther, T. Schmitt, Behavioural and genetic analyses of *Nasonia* shed light on the evolution of sex pheromones, Nature. 494 (2013) 345–348. doi:10.1038/nature11838.

[15] E.C. Verhulst, Maternal control of haplodiploid sex, Science (80-.). 328 (2010) 620–623.

[16] J.H. Werren, D.W. Loehlin, The parasitoid wasp *Nasonia*: An emerging model system with haploid male genetics, Cold Spring Harb. Protoc. 4 (2009) 1–11.

[17] D.S. Saunders, Larval Diapause of Maternal Origin-II. The effect of photoperiod and temperature on *Nasonia vitripennis*. Insect Phys, 12 (1966) 569–581.

[18] J. Van Den Assem, L.W. Beukeboom, A review of *Nasonia* (Chalcidoidea , Pteromalidae) courtship and mating behaviour , with some additional , new observations, Proc. Neth. Entomol. Soc. 15 (2004) 123–132.

[19] M. Chirault, L. Van De Zande, C. Chevrier, C. Bressac, The spatio-temporal partitioning of sperm by males of the prospermatogenic parasitoid *Nasonia vitripennis* is in agreement with its gregarious lifestyle, J. Insect Physiol. 92 (2016) 10–17.

[20] L.W. Beukeboom, A. Kamping, M. Louter, L.P. Pijnacker, V. Katju, P.M. Ferree, J.H. Werren, Haploid females in the parasitic wasp *Nasonia vitripennis*., Science. 315 (2007) 206.

[21] A. R. Whiting, The biology of the parasitic wasp *Mormoniella vitripennis*, Chicago Press. 42 (1967) 333–406.

[22] M.J.D. White, Chromosomal mechanisms in animal reproduction, Bolletino Di Zool. 51 (1984) 1–23.

[23] L. Beukeboom, L. Van De Zande, Genetics of sex determination in the haplodiploid wasp *Nasonia vitripennis* (Hymenoptera: Chalcidoidea), Genetics. 89 (2010) 333–339.

[24] E. Van Wilgenburg, G. Driessen, L.W. Beukeboom, Single locus complementary sex determination in Hymenoptera: An “unintelligent” design?, Front. Zool. 3 (2006) 1–15.

[25] A. Pane, M. Salvemini, P.D. Bovi, C. Polito, G. Saccone, The transformer gene in Ceratitis capitata provides a genetic basis for selecting and remembering the sexual fate, Development. 129 (2002) 3715–3725.

[26] L. Sánchez, Sex-determining mechanisms in insects, Int. J. Dev. Biol. 52 (2008) 837–856.

[27] D.C.S.G. Oliveira, J.H. Werren, E.C. Verhulst, J.D. Giebel, A. Kamping, L.W. Beukeboom, L. Van De Zande, Identification and characterization of the doublesex gene of *Nasonia*, Insect Mol. Biol. 18 (2009) 315–324.

[28] J. H. Werren, S. Richards, C. A.Desjardins, O. Niehuis et al., Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species, Science. 327 (2010) 343–349.

[29] S.L. O ’neill, R. Giordano, A.M.E. Colbert, T.L. Karrf, H.M. Robertson, 6S rRNA phylogenetic analysis of the bacterial endosymbionts associated with cytoplasmic incompatibility in insects, Microbiology. 89 (1992) 2699–2702.

[30] F. Rousset, D. Bouchon, B. Pintureau, P. Juchault, M. Solignac, Wolbachia endosymbionts responsible for various alterations of sexuality in arthropods., Proc. Biol. Sci. 250 (1992) 91–98.

[31] L. Baldo, J.C.D. Hotopp, K.A. Jolley, S.R. Bordenstein, S.A. Biber, R.R. Choudhury, C. Hayashi, M.C.J. Maiden, H. Tettelin, J.H. Werren, Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*, Appl. Environ. Microbiol. 72 (2006) 7098–7110.

[32] J. H. Werren, W. Zhang, L. R. Guo, Evolution and phylogeny of *Wolbachia*: reproductive parasites of arthropods, Roy. Society. 261 (1995) 55-71.

[33] L. Boyle, S.L. O’Neill, H.M. Robertson, T.L. Karr, Interspecific and intraspecific horizontal transfer of *Wolbachia* in *Drosophila*., Science. 260 (1993) 1796–1799.

[34] H.L.C. Dutra, M.N. Rocha, F.B.S. Dias, S.B. Mansur, E.P. Caragata, L.A. Moreira, *Wolbachia* Blocks Currently Circulating Zika Virus Isolates in Brazilian *Aedes aegypti* Mosquitoes, Cell Host Microbe. 19 (2016) 771–774.

[35] A.A. Hoffmann, B.L. Montgomery, J. Popovici, I. Iturbe-Ormaetxe, P.H. Johnson, F. Muzzi, M. Greenfield, M. Durkan, Y.S. Leong, Y. Dong, H. Cook, J. Axford, A.G. Callahan, N. Kenny, C. Omodei, E.A. McGraw, P.A. Ryan, S.A. Ritchie, M. Turelli, S.L. O’Neill, Successful establishment of *Wolbachia* in *Aedes* populations to suppress dengue transmission, Nature. 476 (2011) 454–459.

[36] M.J. Perrot-Minnot, L.R. Guo, J.H. Werren, Single and double infections with *Wolbachia* in the parasitic wasp *Nasonia vitripennis*: Effects on compatibility, Genetics. 143 (1996) 961–972.

[37] F.M. Jiggins, J.H. von Der Schulenburg, G.D. Hurst, M.E. Majerus, Recombination

confounds interpretations of *Wolbachia* evolution., Proc. Biol. Sci. 268 (2001) 1423–7.

[38] J.H. Werren, J.D. Bartos, Recombination in *Wolbachia*, Curr. Biol. 11 (2001) 431–435.

[39] M. Turelli, Evolution of incompatibility-inducing microbes and their hosts, Orga. Evol. 48 (1994) 1500–1513.

[40] T. Rigaud, P. Juchault, J.P. Mocquard, Experimental study of temperature effects on the sex ratio of broods in Terrestrial Crustacea Armadillidium vulgare Latr. Possible implications in natural populations, J. Evol. Biol. 4 (1991) 603–617.

[41] R. Stouthamer, M. van Tilborg, J.H. de Jong, L. Nunney, R.F. Luck, Selfish element maintains sex in natural populations of a parasitoid wasp., Proc. Biol. Sci. 268 (2001) 617–622.

[42] G.D.D. Hurst, C. Bandi, L. Sacchi, A.G. Cochrane, D. Bertrand, I. Karaca, M.E.N. Majerus, Adonia variegata (Coleoptera: Coccinellidae) bears maternally inherited Flavobacteria that kill males only, Parasitology. 118 (1999) 125–134.

[43] H.L. Harris, H.R. Braig, Sperm chromatin remodelling and *Wolbachia*-induced cytoplasmic incompatibility in *Drosophila*., Biochem. Cell Biol. 81 (2003) 229–240.

[44] C.W. Lassy, T.L. Karr, Cytological analysis of fertilization and early embryonic development in incompatible crosses of *Drosophila simulans*, Mech. Dev. 57 (1996) 47–58.

[45] H. Mercot, B. Llorente, M. Jacques, A. Atlan, C. Montchamp-Moreau, Variability within the Seychelles cytoplasmic incompatibility system in *Drosophila simulans*, Genetics. 141 (1995) 1015–1023.

[46] J.H. Werren, Biology of *Wolbachia*, Annu. Rev. Entomol. 42 (1997) 587–609.

[47] C.L. Brelsfoard, S.L. Dobson, *Wolbachia*-based strategies to control insect pests and

disease vectors, Science (80-.). 17 (2009) 55–63.

[48] M. Hiroki, Y. Kato, T. Kamito, K. Miura, Feminization of genetic males by a symbiotic bacterium in a butterfly, Eurema hecabe (Lepidoptera: Pieridae), Nature 89 (2002) 167–170.

[49] J. Li, R. Wang, C. Tang, Somatostatin and Octreotide on the Treatment of Acute Pancreatitis - Basic and Clinical Studies for Three Decades, Curr. Pharm. Des. 17 (2011) 1594–1601.

[50] T. Rigaud , P. Juchault and J. Mocquard, The evolution of sex determination in sopod crustaceans, BioEssays. 19 (2005) 409-416.

[51] P. Grève, C. Braquart-Varnier, J.M. Strub, C. Félix, A. Van Dorsselaer, G. Martin, The glycosylated androgenic hormone of the terrestrial isopod *Porcellio scaber* (Crustacea), Gen. Comp. Endocrinol. 136 (2004) 389–397.

[52] E.C. Verhulst, L. van de Zande, L.W. Beukeboom, Insect sex determination: It all evolves around transformer, Curr. Opin. Genet. Dev. 20 (2010) 376–383.

[53] A. R. Weeks, J. A. Breeuwer, Wolbachia-induced parthenogenesis in a genus of phytophagous mites., Proc. Biol. Sci. 268 (2001) 2245–2251.

[54] R. Stouthamer, R.F. Luck, W.D. Hamilton, Antibiotics cause parthenogenetic *Trichogramma* (Hymenoptera/Trichogrammatidae) to revert to sex., Proc. Natl. Acad. Sci. U. S. A. 87 (1990) 2424–2427.

[55] N. Arakaki, T. Miyoshi, H. Noda, *Wolbachia*-mediated parthenogenesis in the predatory thrips *Franklinothrips vespiformis* (Thysanoptera: Insecta), Proc. R. Soc. B Biol. Sci. 268 (2001) 1011–1016.

[56] J.H. Werren, L. Baldo, M.E. Clark, *Wolbachia*: master manipulators of invertebrate biology, Nat. Rev Microbiol. 6 (2008) 741–751.

[57] B.A. Pannebakker, L.P. Pijnacker, B.J. Zwaan, L.W. Beukeboom, Cytology of *Wolbachia* -induced parthenogenesis in *Leptopilina clavipes* (Hymenoptera: Figitidae), Genome. 47 (2004) 299–303.

[58] F. Gómez, Estudio de casos prácticos: aplicación del modelo de constelaciones familiares en universidades de Perú, Ecuador y España, 73 (2012) 317–327.

[59] K.A. Dyer, J. Jaenike, Evolutionarily stable infection by a male-killing endosymbiont in *Drosophila innubila*: Molecular evidence from the host and parasite genomes, Genetics. 168 (2004) 1443–1455.

[60] R.F. Fialho, L. Stevens, Male-killing *Wolbachia* in a flour beetle., Proc. Biol. Sci. 267 (2000) 1469–1473.

[61] R. F. Fialho, L. Stevens, Male-killing Wolbachia in a flour beetle, Roy. Society. 267 (2000) 1469–1473.

[62] D.W. Zeh, J.A. Zeh, M.M. Bonilla, *Wolbachia*, sex ratio bias and apparent male killing in the harlequin beetle riding pseudoscorpion, Heredity (Edinb). 95 (2005) 41–49.

[63] S.J. Counce, D.F. Poulson, Developmental effects of the sex ratio agent in embryos of Drosophila willistoni, J. Exp. Zool. 151 (1962) 17–31.

[64] A.R. Weeks, M. Turelli, W.R. Harcombe, K.T. Reynolds, A.A. Hoffmann, From parasite to mutualist: Rapid evolution of *Wolbachia* in natural populations of *Drosophila*, PLoS Biol. 5 (2007) 0997–1005.

[65] M. Riegler, M. Sidhu, W.J. Miller, S.L. O'Neill, Evidence for a global *Wolbachia* replacement in *Drosophila melanogaster*, Curr. Biol. 15 (2005) 1428–1433.

[66] A.A. Hoffmann, M. Hercus, H. Dagher, Population dynamics of the *Wolbachia* infection causing cytoplasmic incompatibility in *Drosophila melanogaster*, Genetics. 148 (1998) 221–231.

[67] L.A. Moreira, I. Iturbe-Ormaetxe, J.A. Jeffery, G. Lu, A.T. Pyke, L.M. Hedges, B.C. Rocha, S. Hall-Mendelin, A. Day, M. Riegler, L.E. Hugo, K.N. Johnson, B.H. Kay, E.A. McGraw, A.F. van den Hurk, P.A. Ryan, S.L. O'Neill, A *Wolbachia* Symbiont in *Aedes aegypti* Limits Infection with Dengue, *Chikungunya*, and *Plasmodium*, Cell. 139 (2009) 1268–1278.

[68] E.J. Blitzer, I. Vyazunova, Q. Lan, Functional analysis of AeSCP-2 using gene expression knockdown in the yellow fever mosquito, *Aedes aegypti*, Insect Mol. Biol. 14 (2005) 301–307.

[69] J. Li, R. Wang, C. Tang, Somatostatin and Octreotide on the Treatment of Acute Pancreatitis - Basic and Clinical Studies for Three Decades, Curr. Pharm. Des. 17 (2011) 1594–1601.

[70] H. Ochman, J.G. Lawrence, E. a Groisman, Lateral gene transfer and the nature of bacterial innovation., Nature. 405 (2000) 299–304.

[71] W. Ford Doolittle, Lateral genomics, Trends Biochem. Sci. 24 (1999) 5–8.

[72] E. V Koonin, K.S. Makarova, L. Aravind, Horizontal gene transfer in prokaryotes: Quantification and classification, Annu. Rev. 55 (2001) 709-742.

[73] L. Aravind, R.L. Tatusov, Y.I. Wolf, D.R. Walker, E. V. Koonin, Evidence for massive gene exchange between archaeal and bacterial hyperthermophiles, Trends Genet. 14 (1998) 442–444.

[74] J.C. Dunning Hotopp, Horizontal gene transfer between bacteria and animals, Trends Genet. 27 (2011) 157–163.

[75] J.C. Dunning Hotopp, M.E. Clark, D.C.S.G. Oliveira, J.M. Foster, P. Fischer, M.C. Muñoz Torres, J.D. Giebel, N. Kumar, N. Ishmael, S. Wang, J. Ingram, R. V. Nene, J. Shepard, J. Tomkins, S. Richards, D.J. Spiro, E. Ghedin, B.E. Slatko, H. Tettelin, J.H. Werren, Widespread lateral gene transfer from intracellular bacteria to multicellular

eukaryotes, Science (80-.). 317 (2007) 1753–1756.

[76] N.A. Moran, T. Jarvik, Lateral Transfer of Genes from Fungi, Science (80-.). 328 (2010) 624–627.

[77] S.B. Daniels, K.R. Peterson, L.D. Strausbaugh, M.G. Kidwell, A. Chovnik, Evidence for horizontal transmission of the P transposable element between *Drosophila* species, Genetics. 124 (1990) 339–355.

[78] L.A. Graham, S.C. Lougheed, K.V. Ewart, P.L. Davies, Lateral transfer of a lectin-like antifreeze protein gene in fishes, PLoS One. 3 (2008).

[79] M. Naseem, M.Q. Zaman, H. Nazih, K. Ouguerram, I. Rabbani, H. Zaneb, T. Yaqoob, H.U. Rehman, J. Michel, S.K. Tahir, M.S. Yousaf, K.A. Majeed, M.S. Hussain, The effects of ginkgo biloba leaf extract on metabolic disturbances associated to alloxan-induced diabetic rats, J. Anim. Plant Sci. 26 (2016) 627–635.

[80] A. Hoekema, P.R. Hirsch, P.J.J. Hooykaas, R.A. Schilperoort, Binary Plant Vector Strategy Based on Separation of Vir- and T-Region of the Agro Ti Plasmid.Pdf, Nature. 303 (1983) 179–180.

[81] R. Stouthamer, J.A.J. Breeuwer, G.D.D. Hurst, *Wolbachia Pipientis* : Microbial Manipulator of Arthropod Reproduction, Annu. Rev. Microbiol. 53 (1999) 71–102. doi:10.1146/annurev.micro.53.1.71.

[82] J.H. Werren, Biology of Wolbachia., Annu. Rev. Entomol. 42 (1997) 587--609.

[83] N. Kondo, N. Nikoh, N. Ijichi, M. Shimada, T. Fukatsu, Genome fragment of *Wolbachia* endosymbiont transferred to X chromosome of host insect, Proc. Natl. Acad. Sci. 99 (2002) 14280–14285.

[84] N. Nikoh, K. Tanaka, F. Shibata, N. Kondo, M. Hizume, M. Shimada, T. Fukatsu, *Wolbachia* genome integrated in an insect chromosome: Evolution and fate of laterally

transferred endosymbiont genes, Genome Res. 18 (2008) 272–280.

[85] N. Kondo, N. Nikoh, N. Ijichi, M. Shimada, T. Fukatsu, Genome fragment of *Wolbachia* endosymbiont transferred to X chromosome of host insect, Proc. Natl. Acad. Sci. 99 (2002) 14280–14285.

[86] N. Nikoh, K. Tanaka, F. Shibata, N. Kondo, M. Hizume, M. Shimada, T. Fukatsu, *Wolbachia* genome integrated in an insect chromosome: Evolution and fate of laterally transferred endosymbiont genes, Genome Res. 18 (2008) 272–280.

[87] E.O. Martinson, V.G. Martinson, R. Edwards, Mrinalini, J.H. Werren, Laterally transferred gene recruited as a venom in parasitoid wasps, Mol. Biol. Evol. 33 (2016) 1042–1052.

[88] S.K. Sharma, F. Chappuis, N. Jha, P.A. Bovier, L. Loutan, S. Koirala, Impact of snake bites and determinants of fatal outcomes in Southeastern Nepal, Am. J. Trop. Med. Hyg. 71 (2004) 234–238.

[89] F.S. Markland, Review Paper Snake Venoms and the Hemostatic System, Toxicon. 36 (1998) 1749–1800.

[90] D.A. Warrell, Snake bite, Lancet. 375 (2010) 77–88.

[91] T. Junghanss, M. Bodio, Medically important venomous animals: biology, prevention, first aid, and clinical management., Clin. Infect. Dis. 43 (2006) 1309–17.

[92] P. Escoubas, S. Diochot, G. Corzo, Structure and pharmacology of spider venom neurotoxins.pdf, 82 (2000).

[93] R.S. Vetter, P.K. Visscher, Bites and stings of medically important venomous arthropods, Int. J. Dermatol. 37 (1998) 481–496.

[94] R. Munoz-Arizpe, L. Valencia-Espinoza, L. Velasquez-Jones, C. Abarca-Franco, J.

Gamboa-Marrufo, P. Valencia-Mayoral. Africanized Bee Stings and Pathogenesis of Acute Renal Failure, Nephron 61 (1992) 478.

[95] S. Scott Schneider, G. DeGrandi-Hoffman, D.R. Smith, The African honey bee: Factors Contributing to a Successful Biological Invasion, Annu. Rev. Entomol. 49 (2004) 351–376.

[96] P. Rosenkranz, P. Aumeier, B. Ziegelmann, Biology and control of Varroa destructor, J. Invertebr. Pathol. 103 (2010) S96–S119.

[97] S. Scott Schneider, G. DeGrandi-Hoffman, D.R. Smith, Factors Contributing to a Successful Biological Invasion, Annu. Rev. Entomol. 49 (2004) 351–376.

[98] K.T. Fitzgerald, A.A. Flood, Hymenoptera Stings, Clin. Tech. Small Anim. Pract. 21 (2006) 194–204.

[99] B.E.C. Banks, R.A. Shipolini, Chemistry and Pharmacology of Honey-bee Venom, (1986).

[100] C.S. Seibert, M.R.L. Oliveira, L.R.C. Gonçalves, M.L. Santoro, I.S. Sano-Martins, Intravascular hemolysis induced by *Lonomia obliqua* caterpillar bristle extract: An experimental model of envenomation in rats, Toxicon. 44 (2004) 793–799.

[101] N. Peiren, F. Vanrobaeys, D.C. De Graaf, B. Devreese, J. Van Beeumen, F.J. Jacobs, The protein composition of honeybee venom reconsidered by a proteomic approach, Biochim. Biophys. Acta - Proteins Proteomics. 1752 (2005) 1–5.

[102] T. Wanandy, N. Gueven, N.W. Davies, S.G.A. Brown, M.D. Wiese, Pilosulins: A review of the structure and mode of action of venom peptides from an Australian ant *Myrmecia pilosula*, Toxicon. 98 (2015) 54–61.

[103] N.W. Davies, M.D. Wiese, S.G.A. Brown, Characterisation of major peptides in “jack jumper” ant venom by mass spectrometry, Toxicon. 43 (2004) 173–183.

[104] S.J.M. Moreau, S. Guillot, Advances and prospects on biosynthesis, structures and functions of venom proteins from parasitic wasps, Insect Biochem. Mol. Biol. 35 (2005) 1209–1223.

[105] M.P. Dani, E.H. Richards, R.E. Isaac, J.P. Edwards, Antibacterial and proteolytic activity in venom from the endoparasitic wasp *Pimpla hypochondriaca* (Hymenoptera: Ichneumonidae), J. Insect Physiol. 49 (2003) 945–954.

[106] S. Asgari, G. Zhang, R. Zareie, O. Schmidt, A serine proteinase homolog venom protein from an endoparasitoid wasp inhibits melanization of the host hemolymph, Insect Biochem. Mol. Biol. 33 (2003) 1017–1024.

[107] Y. Nakamatsu, T. Tanaka, Venom of ectoparasitoid, *Euplectrus* sp. near plathypenae (Hymenoptera: Eulophidae) regulates the physiological state of *Pseudaletia separata* (Lepidoptera: Noctuidae) host as a food resource, J. Insect Physiol. 49 (2003) 149–159.

[108] N. Parkinson, E.H. Richards, C. Conyers, I. Smith, J.P. Edwards, Analysis of venom constituents from the parasitoid wasp *Pimpla hypochondriaca* and cloning of a cDNA encoding a venom protein, Insect Biochem. Mol. Biol. 32 (2002) 729–735.

[109] F. Libersat, Wasp uses venom cocktail to manipulate the behavior of its cockroach prey, J. Comp. Physiol. A Neuroethol. Sensory, Neural, Behav. Physiol. 189 (2003) 497–508.

[110] S. Asgari, R. Zareie, G. Zhang, O. Schmidt, Isolation and characterization of a novel venom protein from an endoparasitoid, *Cotesia rubecula* (Hym: Braconidae), Arch. Insect Biochem. Physiol. 53 (2003) 92–100.

[111] A. Krishnan, P.N. Nair, J. D, Isolation, Cloning,and Characterization of New Chitinase Stored in Active Form in Chitin-linedVenom Reservoir, J Biol Chem. 33 (1994) 20971–20976.

[112] W. Gnatzy, W. Volknandt, Venom gland of the digger wasp Liris niger: Morphology, ultrastructure, age-related changes and biochemical aspects, Cell Tissue Res. 302 (2000)

271–284.

[113] F.K. Nielsen, N. Egund, A. Jørgensen, A.G. Jurik, Risk factors for joint replacement in knee osteoarthritis; A 15-year follow-up study, BMC Musculoskelet. Disord. 18 (2017) 1041–1050.

[114] D.B. Rivers, F. Uckan, E. Ergin, Characterization and biochemical analyses of venom from the ectoparasitic wasp *Nasonia vitripennis* (Walker) (Hymenoptera: Pteromalidae), Arch. Insect Biochem. Physiol. 61 (2006) 24–41.

[115] G. von Heijne, The Signal peptide, J Membr Biol. 115 (1990) 195–201.

[116] D. Jones, G. Sawicki, M. Wozniak, Sequence, structure, and expression of a wasp venom protein with a negatively charged signal peptide and a novel repeating internal structure, J. Biol. Chem. 267 (1992) 14871–14878.

[117] H. Andersson, G. Von Heijne, Position-specific Asp-Lys pairing can affect signal sequence function and membrane protein topology, J. Biol. Chem. 268 (1993) 21389–21393.

[118] C. Labrosse, K. Stasiak, J. Lesobre, A. Grangeia, E. Huguet, J.M. Drezen, M. Poirie, A RhoGAP protein as a main immune suppressive factor in the *Leptopilina boulardi* (Hymenoptera, Figitidae)-*Drosophila melanogaster* interaction, Insect Biochem. Mol. Biol. 35 (2005) 93–103.

[119] S. Inouye, X. Soberon, T. Franceschini, K. Nakamura, K. Iakura, M. Inouye, Role of positive charge on the amino-terminal region of the signal peptide in protein secretion across the membrane, Proc. Natl. Acad. Scie. 79 (1982) 3438-3441.

[120] M. Chen, V. Nagarajan, Effect of alteration of charged residues at the N termini of signal peptides on protein export in *Bacillus subtilis*, J. Bacteriol. 176 (1994) 5796–5801.

[121] D.C. De Graaf, M. Aerts, M. Brunain, C.A. Desjardins, F.J. Jacobs, J.H. Werren, B.

Devreese, Insights into the venom composition of the ectoparasitoid wasp *Nasonia vitripennis* from bioinformatic and proteomic studies, Insect Mol. Biol. 19 (2010) 11–26.

[122] E.L. Danneels, D.B. Rivers, D.C. de Graaf, Venom proteins of the parasitoid wasp *Nasonia vitripennis:* Recent discovery of an untapped pharmacopee, Toxins (Basel). 2 (2010) 494–516.

[123] D.B. Rivers, D.L. Denlinger, Developmental fate of the flesh fly, *Sarcophaga bullata*, envenomated by the pupal ectoparasitoid, *Nasonia vitripennis*, J. Insect Physiol. 40 (1994) 121–127.

[124] E.L. Danneels, S. Gerlo, K. Heyninck, K. Van Craenenbroeck, K. De Bosscher, G. Haegeman, D.C. De Graaf, How the venom from the ectoparasitoid wasp *Nasonia vitripennis* exhibits anti-inflammatory properties on mammalian cell lines, PLoS One. 9 (2014) 1–13.

[125] M.R. Kanost, H. Jiang, X.Q. Yu, Innate immune responses of a lepidopteran insect, *Manduca sexta*, Immunol. Rev. 198 (2004) 97–105.

[126] M.J. Williams, *Drosophila* hemopoiesis and cellular immunity., J. Immunol. 178 (2007) 4711–4716.

[127] K. Söderhäll, L. Cerenius, Role of the prophenoloxidase-activating system in invertebrate immunity, Curr. Opin. Immunol. 10 (1998) 23–28.

[128] D.B. Rivers, L. Ruggiero, M. Hayes, The ectoparasitic wasp *Nasonia vitripennis* (Walker) (Hymenoptera: Pteromalidae) differentially affects cells mediating the immune response of its flesh fly host, *Sarcophaga bullata* Parker (Diptera: Sarcophagidae), J. Insect Physiol. 48 (2002) 1053–1064.

[129] C. Qian, Y. Liu, Q. Fang, Y. Min-Li, S.S. Liu, G.Y. Ye, Y.M. Li, Venom of the ectoparasitoid, *Nasonia vitripennis*, influences gene expression in *Musca domestica* hemocytes, Arch. Insect Biochem. Physiol. 83 (2013) 211–231.

[130] D.R.G. Price, H.A. Bell, G. Hinchliffe, E. Fitches, R. Weaver, J.A. Gatehouse, A venom metalloproteinase from the parasitic wasp *Eulophus pennicornis* is toxic towards its host, tomato moth (*Lacanobia oleracae*), Insect Mol. Biol. 18 (2009) 195–202.

[131] M. Abt, D.B. Rivers, Characterization of phenoloxidase activity in venom from the ectoparasitoid *Nasonia vitripennis* (Walker) (Hymenoptera: Pteromalidae), J. Invertebr. Pathol. 94 (2007) 108–118.

[132] B. Tang, X. Chen, Y. Liu, H. Tian, J. Liu, J. Hu, W. Xu, W. Zhang, Characterization and expression patterns of a membrane-bound trehalase from *Spodoptera exigua*., BMC Mol. Biol. 9 (2008) 51.

[133] S. Sotillos, F. Roch, S. Campuzano, The metalloprotease-disintegrin Kuzbanian participates in Notch activation during growth and patterning of *Drosophila* imaginal discs., Development. 124 (1997) 4769–79.

[134] D. Krowarsch, T. Cierpicki, F. Jelen, J. Otlewski, Canonical protein inhibitors of serine proteases, Cell. Mol. Life Sci. 60 (2003) 2427–2444.

[135] Z. Gáspári, A. Patthy, L. Gráf, A. Perczel, Comparative structure analysis of proteinase inhibitors from the desert locust, *Schistocerca gregaria*., Eur. J. Biochem. 269 (2002) 527–37.

[136] Y. Yamazaki, T. Morita, Structure and function of snake venom cysteine-rich secretory proteins, Toxicon. 44 (2004) 227–231.

[137] K.J. Kramer, S. Muthukrishnan, Mini-review: molecular biology and potential use as biopesticides, Insect Biochem. Mol. Biol. 27 (1997) 887–900.

[138] T. Bak, Studies on glucose dehydrogenase of Aspergillus oryzae. Biochim. Biophys, Acta. 139 (1967) 277–293.

[139] N.K. Jain, I. Roy, Effect of trehalose on protein structure, Protein Sci. 18 (2009) 24–36.

[140] M. Jinek, F. Jiang, D.W. Taylor, S.H. Sternberg, E. Kaya, E. Ma, C. Anders, M. Hauer, K. Zhou, S. Lin, M. Kaplan, A.T. Iavarone, E. Charpentier, E. Nogales, J.A. Doudna, Structures of Cas9 endonucleases reveal RNA-mediated conformational activation., Science. 343 (2014) 1247997.

[141] A.Y. Bustos, G.F. de Valdez, R. Raya, A.M. de Almeida, S. Fadda, M.P. Taranto, Proteomic analysis of the probiotic *Lactobacillus reuteri* CRL1098 reveals novel tolerance biomarkers to bile acid-induced stress, Food Res. Int. 77 (2015) 599–607.

[142] P. Falabella, L. Riviello, P. Caccialupi, T. Rossodivita, M. Teresa Valente, M. Luisa De Stradis, A. Tranfaglia, P. Varricchio, S. Gigliotti, F. Graziani, C. Malva, F. Pennacchio, A-glutamyl transpeptidase of *Aphidius ervi* venom induces apoptosis in the ovaries of host aphids, Insect Biochem. Mol. Biol. 37 (2007) 453–465.

[143] D.B. Rivers, D.L. Denlinger, Redirection of metabolism in the flesh fly, *Sarcophaga bullata*, following envenomation by the ectoparasitoid *Nasonia vitripennis* and correlation of metabolic effects with the diapause status of the host, J. Insect Physiol. 40 (1994) 207–215.

[144] C.F. Wilkinson, Biochem. J. (1973) 136, 817-820, (1973) 817–820.

[145] A.L. Siebert, D. Wheeler, J.H. Werren, A new approach for investigating venom function applied to venom calreticulin in a parasitoid wasp, Toxicon. 107 (2015) 304–316.

[146] D.C. De Graaf, M. Brunain, B. Scharlaken, N. Peiren, B. Devreese, D.G. Ebo, W.J. Stevens, C.A. Desjardins, J.H. Werren, F.J. Jacobs, Two novel proteins expressed by the venom glands of *Apis mellifera* and *Nasonia vitripennis* share an ancient C1q-like domain, Insect Mol. Biol. 19 (2010) 1–10.

[147] Mrinalini, A.L. Siebert, J. Wright, E. Martinson, D. Wheeler, J.H. Werren, Parasitoid venom induces metabolic cascades in fly hosts, Metabolomics. 11 (2015) 350–366.

[148] C. Qian, Q. Fang, L. Wang, G.Y. Ye, Molecular cloning and functional studies of two

Kazal-type serine protease inhibitors specifically expressed by *Nasonia vitripennis* venom apparatus, Toxins (Basel). 7 (2015) 2888–2905.

[149] C. Qian, D. Liang, Y. Liu, P. Wang, S. Kausar, G. Wei, B. Zhu, L. Wang, C. Liu, Identification of a small pacifastin protease inhibitor from *Nasonia vitripennis* venom that inhibits humoral immunity of host (*Musca domestica*), Toxicon. 131 (2017) 54–62.

[150] L. Timmons, a Fire, Specific interference by ingested dsRNA., Nature. 395 (1998) 854.

[151] A. Fire, S. Xu, K. Montgomery, S. A. Kostas, S. E. Driver, C. C. Mello, Potent and specific genetic inteference by double-stranded RNA in *Caenorhabditis elegans*, Nature. 391 (1998) 806–811.

[152] E. Bernstein, a a Caudy, S.M. Hammond, G.J. Hannon, Role for a bidentate ribonuclease in the initiation step of RNA interference., Nature. 409 (2001) 363–366.

[153] S.M. Hammond, E. Bernstein, D. Beach, G.J. Hannon, An RNA-directed nuclease mediates post-transcriptional gene silencing in *Drosophila* cells, Nature. 404 (2000) 293–296.

[154] M.I. Rosenberg, A.E. Brent, F. Payre, C. Desplan, Dual mode of embryonic development is highlighted by expression and function of *Nasonia* pair-rule genes, Elife. 2014 (2014) 1–24.

[155] M. Abdel-latief, L.A. Garbe, M. Koch, J. Ruther, An epoxide hydrolase involved in the biosynthesis of an insect sex attractant and its use to localize the production site, Proc. Natl. Acad. Sci. USA. 105 (2008) 8914–8919.

[156] N. Mesaeli, K. Nakamura, E. Zvaritch, P. Dickie, E. Dziak, K.H. Krause, M. Opas, D.H. MacLennan, M. Michalak, Calreticulin is essential for cardiac development, J. Cell Biol. 144 (1999) 857–868.

[157] Z. Prohaszka, [Heat shock proteins as chaperones of the immune response. The optimal

stress of life], Orv Hetil. 144 (2003) 1331–1339.

[158] J.G. Geske, D.T. Huynh, J.I. Seiferas, A note on almost-everywhere-complex sets and separating deterministic-time-complexity classes, Inf. Comput. 92 (1991) 97–104.

[159] H.P. Sørensen, K.K. Mortensen, Advanced genetic strategies for recombinant protein expression in *Escherichia coli*, J. Biotechnol. 115 (2005) 113–128.

[160] J.W. Dubendorfft, W. Studier, Controlling basal expression in an inducible T7 expression system by blocking the target T7 promoter with *lac* repressor, Mol. Biol. 219 (1991) 45–59.

[161] F.W. Studier, Use of bacteriophage T7 lysozyme to improve an inducible T7 expression system, Mol. Biol. 13, (2014) 27.

[162] A.L. Demain, P. Vaishnav, Production of Recombinant Proteins by Microbes and Higher Organisms, Compr. Biotechnol. Second Ed. 3 (2011) 333–345.

[163] A.R. Gomes, S.M. Byregowda, B.M. Veeregowda, V. Balamurugan, An Overview of Heterologous Expression Host Systems for the Production of Recombinant Proteins, Adv. Anim. Vet. Sci. 4 (2016) 346–356.

[164] G.L. Rosano, E.A. Ceccarelli, Recombinant protein expression in *Escherichia coli*: Advances and challenges, Front. Microbiol. 5 (2014) 1–17.

[165] R. Chen, Bacterial expression systems for recombinant protein production: *E. coli* and beyond, Biotechnol. Adv. 30 (2012) 1102–1107.

[166] D. Drew, L. Fröderberg, L. Baars, J.W.L. De Gier, Assembly and overexpression of membrane proteins in *Escherichia coli*, Biochim. Biophys. Acta - Biomembr. 1610 (2003) 3–10.

[167] D. Drew, D.J. Slotboom, G. Friso, T. Reda, P. Genevaux, M. Rapp, N.M. Meindl-Beinker, W. Lambert, M. Lerch, D.O. Daley, others, A scalable, GFP-based pipeline for membrane

protein overexpression screening and purification, Protein Sci. 14 (2005) 2011–2017.

[168] J. Wei, S.H. Zhang, C.N. Wu, X.G. Jian, Relationship between interfacial polymerization monomer structure and separation properties of PPESK based composite membranes, Acta Polym. Sin. (2006) 298–302.

[169] M. Saidijam, G. Psakis, J.L. Clough, J. Meuller, S. Suzuki, C.J. Hoyle, S.L. Palmer, S.M. Morrison, M.K. Pos, R.C. Essenberg, M.C.J. Maiden, A. Abu-Bakr, S.G. Baumberg, A.A. Neyfakh, J.K. Griffith, M.J. Stark, A. Ward, J. O'Reilly, N.G. Rutherford, M.K. Phillips-Jones, P.J.F. Henderson, Collection and characterisation of bacterial membrane proteins, FEBS Lett. 555 (2003) 170–175.

[170] K.R. Andersen, N.C. Leksa, T.U. Schwartz, Optimized E. coli expression strain LOBSTR eliminates common contaminants from His-tag purification, Proteins Struct. Funct. Bioinforma. 81 (2013) 1857–1861.

[171] G. Sezonov, D. Joseleau-Petit, R. D'Ari, *Escherichia coli* physiology in *Luria-Bertani* broth, J. Bacteriol. 189 (2007) 8746–8749. doi:10.1128/JB.01368-07.

[172] J. Shiloach, R. Fass, Growing E. coli to high cell density - A historical perspective on method development, Biotechnol. Adv. 23 (2005) 345–357.

[173] I. Ahmad, N. Nawaz, N.M. Darwesh, S. ur Rahman, M.Z. Mustafa, S.B. Khan, S.G. Patching, Overcoming challenges for amplified expression of recombinant proteins using *Escherichia coli*, Protein Expr. Purif. 144 (2018) 12–18.

[174] K. Janiyani, T. Bordelon, G.L. Waldrop, J.E. Cronan, Function of *Escherichia coli* Biotin Carboxylase Requires Catalytic Activity of Both Subunits of the Homodimer, J. Biol. Chem. 276 (2001) 29864–29870.

[175] N. Rosenfeld, J.W. Young, U. Alon, P.S. Swain, M.B. Elowitz, Gene regulation at the single-cell level, Science (80-.). 307 (2005) 1962–1965.

[176] N. Rosenfeld, J.W. Young, U. Alon, P.S. Swain, M.B. Elowitz, Gene regulation at the single-cell level, Science (80-.). 307 (2005) 1962–1965.

[177] N.J. Claassens, M.F. Siliakus, S.K. Spaans, S.C.A. Creutzburg, B. Nijsse, P.J. Schaap, T.E.F. Quax, J. Van Der Oost, Improving heterologous membrane protein production in *Escherichia coli* by combining transcriptional tuning and codon usage algorithms, PLoS One. 12 (2017).

[178] O. Schlesinger, Y. Chemla, M. Heltberg, E. Ozer, R. Marshall, V. Noireaux, M.H. Jensen, L. Alfonta, Tuning of Recombinant Protein Expression in *Escherichia coli* by Manipulating Transcription, Translation Initiation Rates, and Incorporation of Noncanonical Amino Acids, ACS Synth. Biol. 6 (2017) 1076–1085.

[179] I. Ahmad, N. Nawaz, N.M. Darwesh, S. ur Rahman, M.Z. Mustafa, S.B. Khan, S.G. Patching, Overcoming challenges for amplified expression of recombinant proteins using *Escherichia coli*, Protein Expr. Purif. 144 (2018) 12–18.

[180] R.R. Burgess, Chapter 17 Refolding Solubilized Inclusion Body Proteins, 1st ed., Elsevier Inc., 2009.

[181] C. Biology, Molecular Chaperones in the Cytosol : from Nascent Chain to Folded Protein, Science (80-.). 1852 (2012) 1852–1859.

[182] J. Lobstein, C.A. Emrich, C. Jeans, M. Faulkner, P. Riggs, M. Berkmen, SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm, Microb. Cell Fact. 11 (2012) 1–16.

[183] C. Qian, D. Liang, Y. Liu, P. Wang, S. Kausar, G. Wei, B. Zhu, L. Wang, C. Liu, Identification of a small pacifastin protease inhibitor from *Nasonia vitripennis* venom that inhibits humoral immunity of host (Musca domestica), Toxicon. 131 (2017) 54–62.

[184] P.D. Shirk, R.B. Furlong, A. Dolan, J.H. Werren, Functional characterization of the transcriptional regulatory elements of three highly expressed constitutive genes in the

jewel wasp, *Nasonia vitripennis*, Insect Mol. Biol. 26 (2017) 743–751.

[185] R. Kumar, J. Singh, Expression and secretion of a prokaryotic protein streptokinase without glycosylation and degradation in *Schizosaccharomyces pombe*, Yeast. 21 (2004) 1343–1358.

[186] J.M. Zhou, Y.X. Tang, D.Y. Fang, J.J. Zhou, Y. Liang, H.Y. Guo, L.F. Jiang, Secreted expression and purification of dengue 2 virus full-length nonstructural glycoprotein NS1 in *Pichia. pastoris*, Virus Genes. 33 (2006) 27–32.

[187] J.M. Cregg, J.L. Cereghino, J. Shi, D.R. Higgins, Recombinant Protein Expression in *Pichia pastoris*, Mol. Biotechnology. 16 (2000) 23-52.

[188] R.M. Bill, Yeast - A panacea for the structure-function analysis of membrane proteins?, Curr. Genet. 40 (2001) 157–171.

[189] A.L. Demain, P. Vaishnav, Production of Recombinant Proteins by Microbes and Higher Organisms, Compr. Biotechnol. Second Ed. 3 (2011) 333–345.

[190] T. Gerngross, Production of Complex Human Glycoproteins in Yeast, 301 (2005) 139–139.

[191] M. Ward, C. Lin, D.C. Victoria, B.P. Fox, J.A. Fox, D.L. Wong, H.J. Meerman, J.P. Pucci, R.B. Fong, M.H. Heng, N. Tsurushita, C. Gieswein, M. Park, H. Wang, Characterization of humanized antibodies secreted by *Aspergillus niger*, Appl. Environ. Microbiol. 70 (2004) 2567–2576.

[192] G. Britain, A.R. Reserved, P. Press, Production Scale Insect Cell Culture, Biotechnol. Adv. 9 (1991) 51–58.

[193] E.D.J. Stewart, Brown Coal in Victoria., Conf. Ser. - Australas. Inst. Min. Metall. 23 (1982) 567–575.

[194] J. Yin, G. Li, X. Ren, G. Herrler, Select what you need: A comparative evaluation of the advantages and limitations of frequently used expression systems for foreign genes, J. Biotechnol. 127 (2007) 335–347.

[195] M.L. Nagpal, Y. Chen, T. Lin, Effects of overexpression of CXCL10 (cytokine-responsive gene-2) on MA-10 mouse Leydig tumor cell steroidogenesis and proliferation, J. Endocrinol. 183 (2004) 585–594.

[196] X. Ren, J. Glende, M. Al-Falah, V. de Vries, C. Schwegmann-Wessels, X. Qu, L. Tan, T. Tschernig, H. Deng, H.Y. Naim, G. Herrler, Analysis of ACE2 in polarized epithelial cells: Surface expression and function as receptor for severe acute respiratory syndrome-associated coronavirus, J. Gen. Virol. 87 (2006) 1691–1695.

[197] A.D. Sim, D. Wheeler, The venom gland transcriptome of the parasitoid wasp *Nasonia vitripennis* highlights the importance of novel genes in venom function, BMC Genomics. 17 (2016) 571.

[198] A.A. Hoffmann, B.L. Montgomery, J. Popovici, I. Iturbe-Ormaetxe, P.H. Johnson, F. Muzzi, M. Greenfield, M. Durkan, Y.S. Leong, Y. Dong, H. Cook, J. Axford, A.G. Callahan, N. Kenny, C. Omodei, E.A. McGraw, P.A. Ryan, S.A. Ritchie, M. Turelli, S.L. O'Neill, Successful establishment of *Wolbachia* in Aedes populations to suppress dengue transmission, Nature. 476 (2011) 454–459.

[199] A.J. Westermann, S.A. Gorski, J. Vogel, Dual RNA-seq of pathogen and host, Nat. Rev. Microbiol. 10 (2012) 618–630.

[200] A.J. Westermann, K.U. Förstner, F. Amman, L. Barquist, Y. Chao, L.N. Schulte, L. Müller, R. Reinhardt, P.F. Stadler, J. Vogel, Dual RNA-seq unveils noncoding RNA functions in host-pathogen interactions, Nature. 529 (2016) 496–501.

[201] A. Twaddle, E. Ghedin, T. Ding, T.R. Unnasch, S. Lustigman, A. Grote, D. Voronin, Defining *Brugia malayi* and *Wolbachia* symbiosis by stage-specific dual RNA-seq, PLoS

Negl. Trop. Dis. 11 (2017) e0005357.

[202] M.P. Cox, D.A. Peterson, P.J. Biggs, SolexaQA: At-a-glance quality assessment of Illumina second-generation sequencing data., BMC Bioinformatics. 11 (2010) 485.

[203] D. Kim, G. Pertea, C. Trapnell, H. Pimentel, R. Kelley, S.L. Salzberg, TopHat2: accurate alignment of transcriptomes in the presence of insertions, deletions and gene fusions., Genome Biol. 14 (2013) R36.

[204] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2., Nat. Methods. 9 (2012) 357–9.

[205] S. Anders, P.T. Pyl, W. Huber, HTSeq-A Python framework to work with high-throughput sequencing data, Bioinformatics. 31 (2015) 166–169.

[206] M.I. Love, S. Anders, W. Huber, Differential analysis of count data - the DESeq2 package, 2014.

[207] A. Conesa, S. G??tz, J.M. Garcia-Gomez, J. Terol, M. Talon, M. Robles, Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research, Bioinformatics. 21 (2005) 3674–3676.

[208] E. Brent, G. Phil, Base-Calling of Automated Sequencer Traces Using Phred. II. Error Probabilities, Genome Res. 8:186-194 (1998) 186–194.

[209] P.M. Ferree, H.M. Frydman, J.M. Li, J. Cao, E. Wieschaus, W. Sullivan, *Wolbachia* utilizes host microtubules and Dynein for anterior localization in the *Drosophila* Oocyte, PLoS Pathog. 1 (2005) 0111–0124.

[210] L.R. Serbus, C. Casper-Lindley, F. Landmann, W. Sullivan, The genetics and cell biology of *Wolbachia*-host interactions., Annu. Rev. Genet. 42 (2008) 683–707.

[211] C. Bressac, F. Rousset, The reproductive incompatibility system in *Drosophila simulans*:

DAPI-staining analysis of the *Wolbachia symbionts* in sperm cysts., J. Invertebr. Pathol. 61 (1993) 226–230.

[212] D.R.P. Tulsiani, A. Abou-Haila, C.R. Loeser, B.M.J. Pereira, REVIEW The Biological and Functional Significance of the Sperm Acrosome and Acrosomal Enzymes in Mammalian Fertilization, Exp. Cell Res. 240 (1998) 151–164.

[213] J.C. Brownlie, B.N. Cass, M. Riegler, J.J. Witsenburg, I. Iturbe-Ormaetxe, E.A. McGraw, S.L. O'Neill, Evidence for metabolic provisioning by a common invertebrate endosymbiont, wolbachia pipientis, during periods of nutritional stress, PLoS Pathog. 5 (2009).

[214] N. Kremer, D. Charif, H. Henri, F. Gavory, P. Wincker, P. Mavingui, F. Vavre, Influence of *Wolbachia* on host gene expression in an obligatory symbiosis, Bmc Microbiol. 12 (2012) S7.

[215] A. Heddi, A.M. Grenier, C. Khatchadourian, H. Charles, P. Nardon, K. Hatchadourian, H.U.C. Harles, A.B.H. Eddi, A.N.N.E.A.G. Renier, C. Haque, Four intracellular genomes direct weevil biology: Nuclear, mitochondrial, principal endosymbiont, and *Wolbachia*, Proc. Natl. Acad. Sci. U. S. A. 96 (1999) 6814–6819.

[216] J.O. Andersson, Lateral gene transfer in eukaryotes, Cell. Mol. Life Sci. 62 (2005) 1182–1197.

[217] D. DeWayne Shoemaker, C.A. Machado, D. Molbo, J.H. Werren, D.M. Windsor, E.A. Herre, The distribution of *Wolbachia* in fig wasps: correlations with host phylogeny, ecology and population structure, Proc. R. Soc. B Biol. Sci. 269 (2002) 2257–2267.

[218] K. Fenn, C. Conlon, M. Jones, M.A. Quail, N.E. Holroyd, J. Parkhill, M. Blaxter, Phylogenetic relationships of the Wolbachia of nematodes and arthropods, PLoS Pathog. 2 (2006) 0887–0899.

[219] X. Wang, J.H. Werren, A.G. Clark, Genetic and epigenetic architecture of sex-biased

expression in the jewel wasps *Nasonia vitripennis* and *giraulti*, Proc. Natl. Acad. Sci. 112 (2015) E3545–E3554.

[220] T.B. Sackton, J.H. Werren, A.G. Clark, Characterizing the infection-induced transcriptome of *Nasonia vitripennis* reveals a preponderance of taxonomically-restricted immune genes, PLoS One. 8 (2013).

[221] E. Aronesty, Comparison of Sequencing Utility Programs, Open Bioinforma. J. 7 (2013) 1–8.

[222] B. Langmead, S.L. Salzberg, Fast gapped-read alignment with Bowtie 2., Nat. Methods. 9 (2012) 357–9.

[223] IGV (Integrative Genomic Viewer), Integrative Genomics Viewer, Broad Inst. 29 (2013) 24–26.

[224] J.C.D. Hotopp, M.E. Clark, D.C.S.G. Oliveira, J.M. Foster, P. Fischer, M.C.M. Torres, J.D. Giebel, N. Kumar, N. Ishmael, S. Wang, J. Ingram, R. V Nene, J. Shepard, J. Tomkins, S. Richards, D.J. Spiro, E. Ghedin, B.E. Slatko, H. Tettelin, J.H. Werren, Widespread Lateral Gene Transfer from Intracellular Bacteria to Multicellular Eukaryotes, Science, 317 (2007) 1753-1756.

[225] S. Kumar, G. Stecher, K. Tamura, MEGA7: Molecular Evolutionary Genetics Analysis version 7.0 for bigger datasets., Mol. Biol. Evol. 33 (2016).

[226] D.T. Jones, W.R. Taylor, J.M. Thornton, PHYLIP (Phylogenetic Inference Package) version 3.6. The rapid generation of mutation data matrices from protein sequences, Comput Appl Biosci. 8 (1992) 275–282.

[227] Y. Boucher, C.J. Douady, R.T. Papke, D.A. Walsh, M.E.R. Boudreau, C.L. Nesbø, R.J. Case, W.F. Doolittle, C.L. Nesbo, R.J. Case, W.F. Doolittle, Lateral gene transfer and the origins of prokaryotic groups, Annu Rev Genet. 37 (2003) 283–328.

[228] B.N. Kent, L. Salichos, J.G. Gibbons, A. Rokas, I.L.G. Newton, M.E. Clark, S.R. Bordenstein, Complete bacteriophage transfer in a bacterial endosymbiont (*Wolbachia*) determined by targeted genome capture, Genome Biol. Evol. 3 (2011) 209–218.

[229] G. Zhang, Z.Q. Lu, H. Jiang, S. Asgari, Negative regulation of prophenoloxidase (proPO) activation by a clip-domain serine proteinase homolog (SPH) from endoparasitoid venom, Insect Biochem. Mol. Biol. 34 (2004) 477–483.

[230] E.L. Danneels, E.M. Formesyn, D.C. de Graaf, Exploring the Potential of Venom from *Nasonia vitripennis* as Therapeutic Agent with High-Throughput Screening Tools., Toxins (Basel). 7 (2015) 2051–70. doi:10.3390/toxins7062051.

[231] A.D. Sim, D. Wheeler, The venom gland transcriptome of the parasitoid wasp *Nasonia vitripennis* highlights the importance of novel genes in venom function, BMC Genomics. 17 (2016) 571. doi:10.1186/s12864-016-2924-7.

[232] S.F. Altschul, W. Gish, W. Miller, E.W. Myers, D.J. Lipman, Basic local alignment search tool., J. Mol. Biol. 215 (1990) 403–10.

[233] R.D. Finn, J. Clements, W. Arndt, B.L. Miller, T.J. Wheeler, F. Schreiber, A. Bateman, S.R. Eddy, HMMER web server: 2015 Update, Nucleic Acids Res. 43 (2015) W30–W38.

[234] T.N. Petersen, H. Nielsen, SignalP 4.0: discriminating signal peptides from transmembrane regions, Nature Methods. 8 (2011) 785-786.

[235] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, M.R. Wilkins, R.D. Appel, A. Bairoch, Protein Identification and Analysis Tools on the ExPASy Server, Proteomics Protoc. Handb. (2005) 571–607.

[236] L.A. Kelly, S. Mezulis, C. Yates, M. Wass, M. Sternberg, The Phyre2 web portal for protein modelling, prediction, and analysis, Nat. Protoc. 10 (2015) 845–858.

[237] S. Asgari, D.B. Rivers, Venom Proteins from Endoparasitoid Wasps and Their Role in

Host-Parasite Interactions, Annu. Rev. Entomol. 56 (2011) 313–335.

[238] D.B. Rivers, M.M. Rocco, A.R. Frayha, Venom from the ectoparasitic wasp *Nasonia vitripennis* increases Na+ influx and activates phospholipase C and phospholipase A2 dependent signal transduction pathways in cultured insect cells, Toxicon. 40 (2002) 9–21.

[239] N.J. Patron, R.F. Waller, Transit peptide diversity and divergence: A global analysis of plastid targeting signals, BioEssays. 29 (2007) 1048–1058.

[240] S.R. Kimball, L.S. Jefferson, Regulation of global and specific mRNA translation by oral administration of branched-chain amino acids, Biochem. Biophys. Res. Commun. 313 (2004) 423–427.

[241] S. Guo, X. Zhu, D. Jańczewski, S.S.C. Lee, T. He, S.L.M. Teo, G.J. Vancso, Measuring protein isoelectric points by AFM-based force spectroscopy using trace amounts of sample, Nat. Nanotechnol. 11 (2016) 817–823.

[242] P. Wasps, T. Venoms, Evolution of Venomous Animals and Their Toxins, (2015) 187–212.

[243] J. Skolnick, J.S. Fetrow, From genes to protein structure and function: Novel applications of computational approaches in the genomic era, Trends Biotechnol. 18 (2000) 34–39.

[244] G. Pollastri, D. Przybylski, B. Rost, P. Baldi, Improving the prediction of protein secondary structure in three and eight classes using recurrent neural networks and profiles, Proteins Struct. Funct. Genet. 47 (2002) 228–235.

[245] L.J. McGuffin, K. Bryson, D.T. Jones, The PSIPRED protein structure prediction server, Bioinformatics. 16 (2000) 404–405.

[246] C. Cole, J.D. Barber, G.J. Barton, The Jpred 3 secondary structure prediction server., Nucleic Acids Res. 36 (2008) 197–201.

[247] F.L. Cônsoli, H.S. Tian, S.B. Vinson, C.J. Coates, Differential gene expression during wing morph differentiation of the ectoparasitoid *Melittobia digitata* (Hym., Eulophidae), Comp. Biochem. Physiol. - A Mol. Integr. Physiol. 138 (2004) 229–239.

[248] S. Asgari, D.B. Rivers, Venom Proteins from Endoparasitoid Wasps and Their Role in Host-Parasite Interactions, Annu. Rev. Entomol. 56 (2011) 313–335.

[249] S. R. Pfeffer, J.E. Rothman, Biosynthetic protein transport and sorting by the endoplasmic reticulum and golgi. Annu. Rev. 56 (1987) 829-852.

[250] F. Jacob-Dubuisson, R. Fernandez, L. Coutte, Protein secretion through autotransporter and two-partner pathways, Biochim. Biophys. Acta - Mol. Cell Res. 1694 (2004) 235–257.

[251] S.C. Hinnah, K. Hill, R. Wagner, T. Schlicher, J. Soll, Reconstitution of a chloroplast protein import channel, EMBO J. 16 (1997) 7351–7360.

[252] V. Kozjak, N. Wiedemann, D. Milenkovic, C. Lohaus, H.E. Meyer, B. Guiard, C. Meisinger, N. Pfanner, An essential role of Sam50 in the protein sorting and assembly machinery of the mitochondrial outer membrane, J. Biol. Chem. 278 (2003) 48520–48523.

[253] M.R. Yen, C.R. Peabody, S.M. Partovi, Y. Zhai, Y.H. Tseng, M.H. Saier, Protein-translocating outer membrane porins of Gram-negative bacteria, Biochim. Biophys. Acta - Biomembr. 1562 (2002) 6–31.

[254] J.F. Glatz, M.M. Vork, D.P. Cistola, G.J. van der Vusse, Cytoplasmic fatty acid binding protein: significance for intracellular transport of fatty acids and putative role on signal transduction pathways., Prostaglandins. Leukot. Essent. Fatty Acids. 48 (1993) 33–41.

[255] M. Donovan, B. Olofsson, A.L. Gustafson, L. Dencker, U. Eriksson, The cellular retinoic acid binding proteins., J. Steroid Biochem. Mol. Biol. 53 (1995) 459–65.

[256] M. Donovan, B. Olofsson, A.L. Gustafson, L. Dencker, U. Eriksson, The cellular retinoic

acid binding proteins., J. Steroid Biochem. Mol. Biol. 53 (1995) 459–65.

[257] I.-H. Jang, H.-J. Nam, W.-J. Lee, CLIP-domain serine proteases in Drosophila innate immunity, BMB Rep. 41 (2008) 102–107.

[258] X. Que, S.L. Reed, Cysteine proteinases and the pathogenesis of amebiasis, Clin. Microbiol. Rev. 13 (2000) 196–206.

[259] M.N.L. Meirelles, L. Juliano, E. Carmona, S.G. Silva, E.M. Costa, A.C.M. Murta, J. Scharfstein, Inhibitors of the major cysteinyl proteinase (GP57/51) impair host cell invasion and arrest the intracellular development of *Trypanosoma cruzi* in vitro, Mol. Biochem. Parasitol. 52 (1992) 175–184.

[260] J. Inselburg, I.C. Bathurst, J. Kansopon, P.J. Barr, R. Rossan, Protective immunity induced in Aotus monkeys by a recombinant SERA protein of *Plasmodium falciparum:* Further studies using SERA 1 and MF75.2 adjuvant, Infect. Immun. 61 (1993) 2048–2052.

[261] S.L. Reed, J.A. Ember, D.S. Herdman, R.G. DiScipio, T.E. Hugli, I. Gigli, The extracellular neutral cysteine proteinase of *Entamoeba histolytica* degrades anaphylatoxins C3a and C5a, J Immunol. 155 (1995) 266–274.

[262] L. Wang, Q. Fang, C. Qian, F. Wang, X.Q. Yu, G. Ye, Inhibition of host cell encapsulation through inhibiting immune gene expression by the parasitic wasp venom calreticulin, Insect Biochem. Mol. Biol. 43 (2013) 936–946.

[263] N.E. Beckage, D.B. Gelman, Wasp parasitoid disruption of host development: Implications for New Biologically Based Strategies for Insect Control, Annu. Rev. Entomol. 49 (2004) 299–330. doi:10.1146/annurev.ento.49.061802.123324.

[264] N.E. Beckage, D.B. Gelman, Implications for New Biologically Based Strategies for Insect Control, Annu. Rev. Entomol. 49 (2004) 299–330.

[265] Rivers_Venom-Induced Alterations in Fly Lipid Metabolism and Its Impact on Larval Development of the Ectoparasitoid *Nasonia vitripennis*.pdf, (n.d.).

[266] A.J. van der Goot, A.T. Tenorio, R.M. Boom, C. V. Nikiforidis, E.W.M. de Jong, Interfacial properties and emulsification performance of thylakoid membrane fragments, Soft Matter. 13 (2016) 608–618.

[267] R. Waninge, E. Kalda, M. Paulsson, T. Nylander, B. Bergenståhl, Cryo-TEM of isolated milk fat globule membrane structures in cream, Phys. Chem. Chem. Phys. 6 (2004) 1518–1523.

[268] S. Gallier, K. Vocking, J.A. Post, B. Van De Heijning, D. Acton, E.M. Van Der Beek, T. Van Baalen, A novel infant milk formula concept: Mimicking the human milk fat globule structure, Colloids Surfaces B Biointerfaces. 136 (2015) 329–339.

[269] E. Fitches, M.G. Edwards, C. Mee, E. Grishin, A.M.R. Gatehouse, J.P. Edwards, J.A. Gatehouse, Fusion proteins containing insect-specific toxins as pest control agents: Snowdrop lectin delivers fused insecticidal spider venom toxin to insect haemolymph following oral ingestion, J. Insect Physiol. 50 (2004) 61–71.

[270] W.K. You, W.S. Choi, Y.S. Koh, H.C. Shin, Y. Jang, K.H. Chung, Functional characterization of recombinant batroxobin, a snake venom thrombin-like enzyme, expressed from Pichia pastoris, FEBS Lett. 571 (2004) 67–73.

[271] C. Bruce, E.C. Fitches, N. Chougule, H.A. Bell, J.A. Gatehouse, Recombinant conotoxin, TxVIA, produced in yeast has insecticidal activity, Toxicon. 58 (2011) 93–100.

[272] J.H. Werren, D.W. Loehlin, The Parasitoid Wasp Nasonia: An Emerging Model System With Haploid Male Genetics, Cold Spring Harb. Protoc. 2009 (2010) 1–31.

[273] D. Desjardins, M.E. Toomey, H.M. Frydman, E.D. Kolaczyk, E.M. Fast, K. Panaram, *Wolbachia* Enhance *Drosophila* Stem Cell Proliferation and Target the Germline Stem Cell Niche, Science (80-.). 334 (2011) 990–992.

[274] Y. Zheng, J.L. Wang, C. Liu, C.P. Wang, T. Walker, Y.F. Wang, Differentially expressed profiles in the larval testes of *Wolbachia* infected and uninfected *Drosophila*, BMC Genomics. 12 (2011).

[275] M.G. Riparbelli, R. Giordano, G. Callaini, Effects of *Wolbachia* on sperm maturation and architecture in *Drosophila simulans* Riverside, Mech. Dev. 124 (2007) 699–714.

[276] K. Hilgenboecker, P. Hammerstein, P. Schlattmann, A. Telschow, J.H. Werren, How many species are infected with *Wolbachia*? - A statistical analysis of current data, FEMS Microbiol. Lett. 281 (2008) 215–220.

[277] K. Pfarr, A. Hoerauf, Antibiotics which Target the *Wolbachia* Endosymbionts of Filarial Parasites: A New Strategy for Control of Filariasis and Amelioration of Pathology, Mini-Reviews Med. Chem. 6 (2006) 203–210.

[278] G.L. Hughes, R. Koga, P. Xue, T. Fukatsu, J.L. Rasgon, *Wolbachia* infections are virulent and inhibit the human malaria parasite *Plasmodium falciparum* in *Anopheles gambiae*, PLoS Pathog. 7 (2011) 3–10.

[279] K. Pfarr, A. Hoerauf, Antibiotics which target the *Wolbachia* endosymbionts of filarial parasites: A new strategy for control of filarisi and amelioration of pathology, Medic. Chem. 6 (2006) 203-210.

[280] H.A. Flores, S.L. O'Neill, Controlling vector-borne diseases by releasing modified mosquitoes, Nat. Rev. Microbiol. 16 (2018) 508–518.

[281] N.M. Ferguson, T.N. Bich Chau, D.T. Hue Kien, H. Clapham, V.T. Long, B. Wills, R. Aguas, J. Popovici, C.P. Simmons, V.T. Trung, L.T. Dui, E.A. McGraw, H.L. Nguyen, S.L. O'Neill, N. Van Vinh Chau, P.A. Ryan, Modeling the impact on virus transmission

of Wolbachia -mediated blocking of dengue virus infection of *Aedes aegypti*, Sci. Transl. Med. 7 (2015) 279ra37-279ra37.

[282] I. Dorigatti, C. McCormack, G. Nedjati-Gilani, N.M. Ferguson, Using Wolbachia for Dengue Control: Insights from Modelling, Trends Parasitol. 34 (2018) 102–113.

[283] H.L.C. Dutra, M.N. Rocha, F.B.S. Dias, S.B. Mansur, E.P. Caragata, L.A. Moreira, *Wolbachia* Blocks Currently Circulating Zika Virus Isolates in Brazilian *Aedes aegypti* Mosquitoes, Cell Host Microbe. 19 (2016) 771–774.

[284] M.A. Nowak, M.C. Boerlijst, J. Cooke, J.M. Smith, Evolution of genetic redundancy, Nature. 388 (1997) 167–170.

[285] S.Y. Er, L.D. Rash, N.J. Saez, S. Senff, G.F. King, J.E. Jensen, V. Herzig, Spider-Venom Peptides as Therapeutics, Toxins (Basel). 2 (2010) 2851–2871.

[286] D.C.I. Koh, A. Armugam, K. Jeyaseelan, Snake venom components and their applications in biomedicine, Cell. Mol. Life Sci. 63 (2006) 3030–3041.

[287] U.R. Müller, Recombinant hymenoptera venom allergens, Allergy Eur. J. Allergy Clin. Immunol. 57 (2002) 570–576.

[288] M.D. Street, G.R. Donovan, B.A. Baldo, Molecular cloning and characterization of the major allergen Myr p II from the venom of the jumper ant *Myrmecia pilosula*: Myr p I and Myr p II share a common protein leader sequence, Biochim Biophys Acta. 1305 (1996) 87–97.

Appendices

Appendix 1

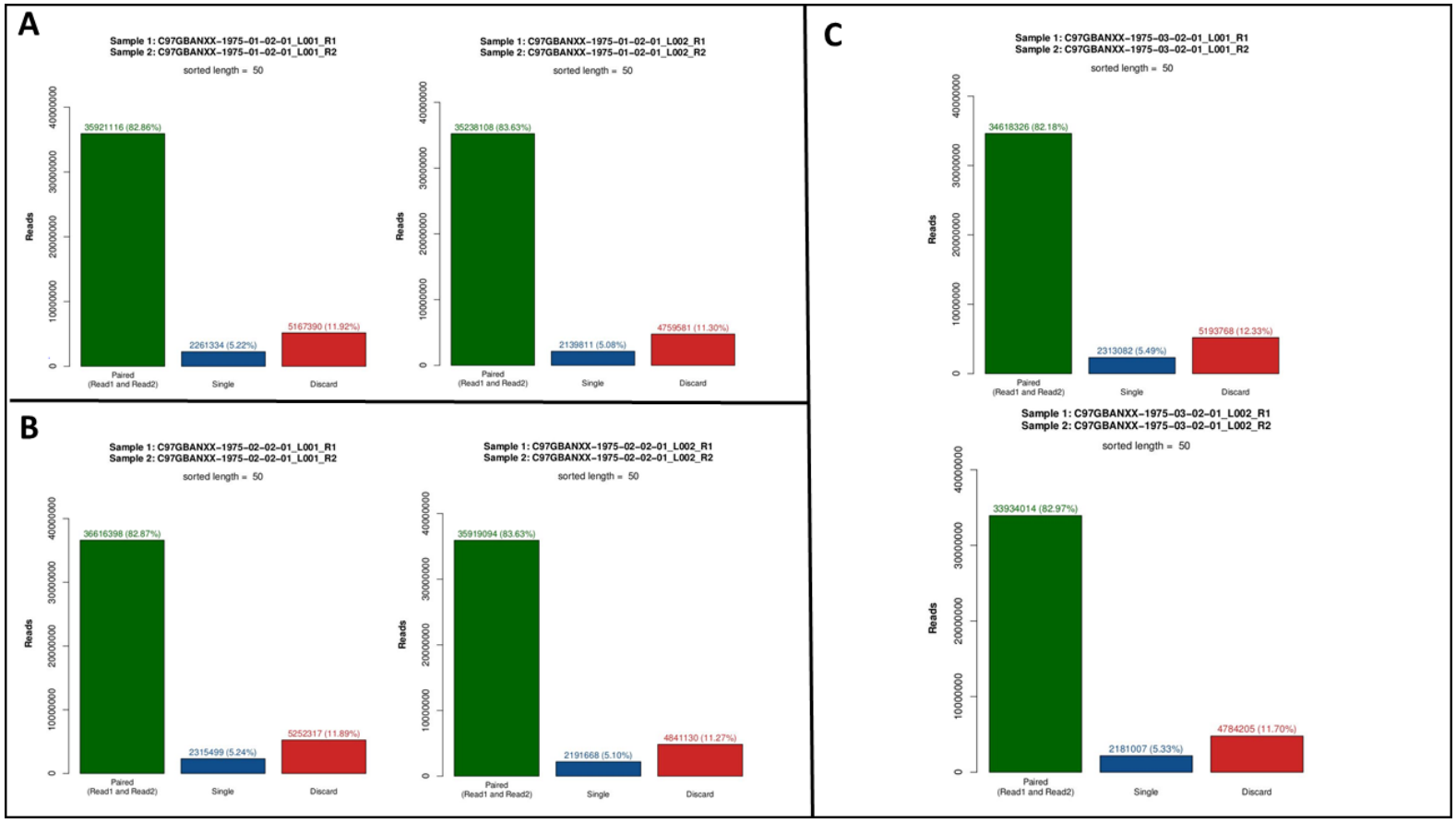


Figure A1.1: Summary of the trimming process of reads in Wol- ovaries replicates using the SolexaQA++ package. Green represent reads with both pairs retained, blue shows the number of reads whose pairs were discarded and red represent the reads the number of reads discarded.

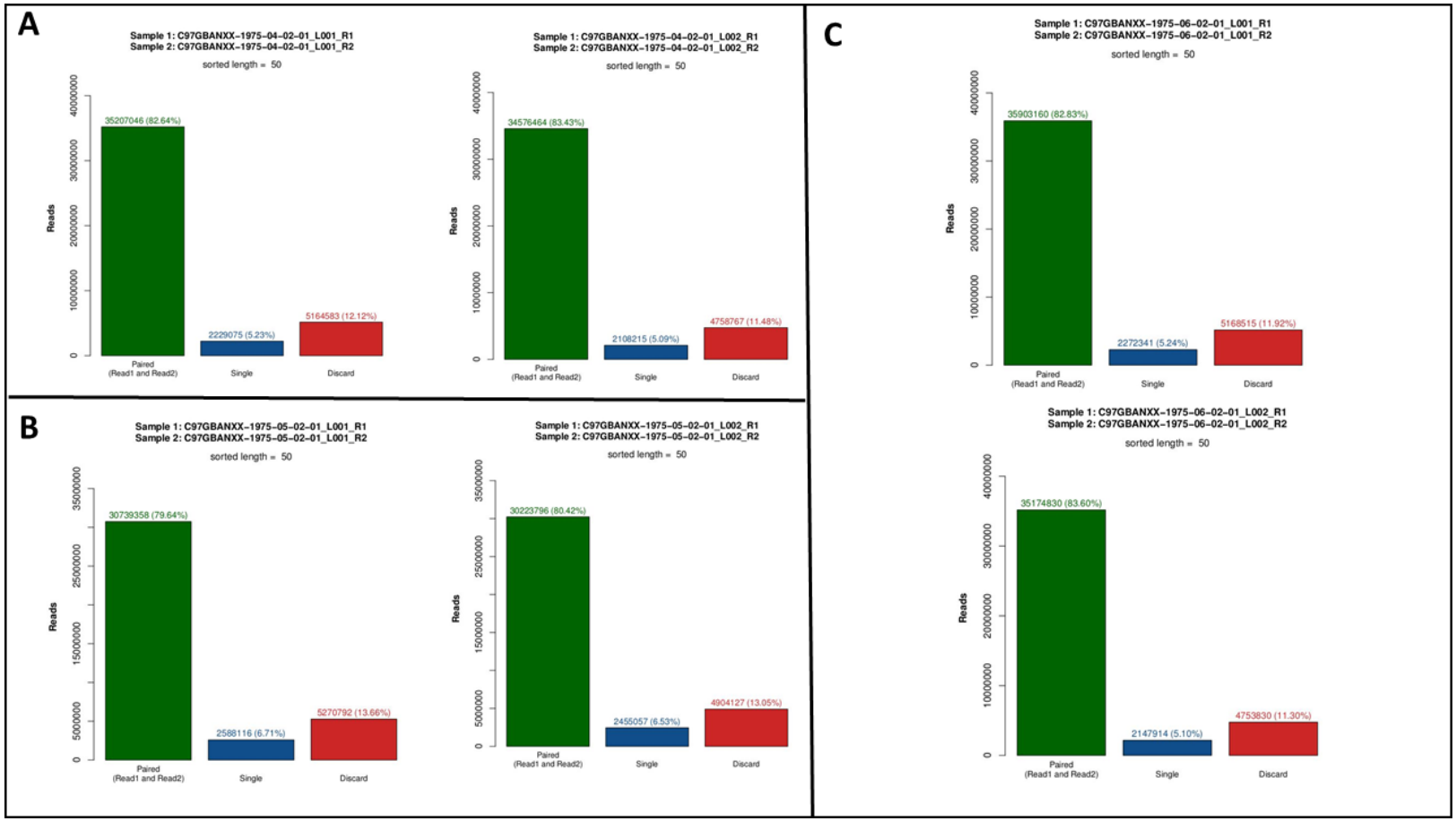


Figure A1.2: Summary of the trimming process of reads in Wol+ ovaries replicates using the SolexaQA++ package. Green represent reads with both pairs retained, blue shows the number of reads whose pairs were discarded and red represent the reads the number of reads discarded.

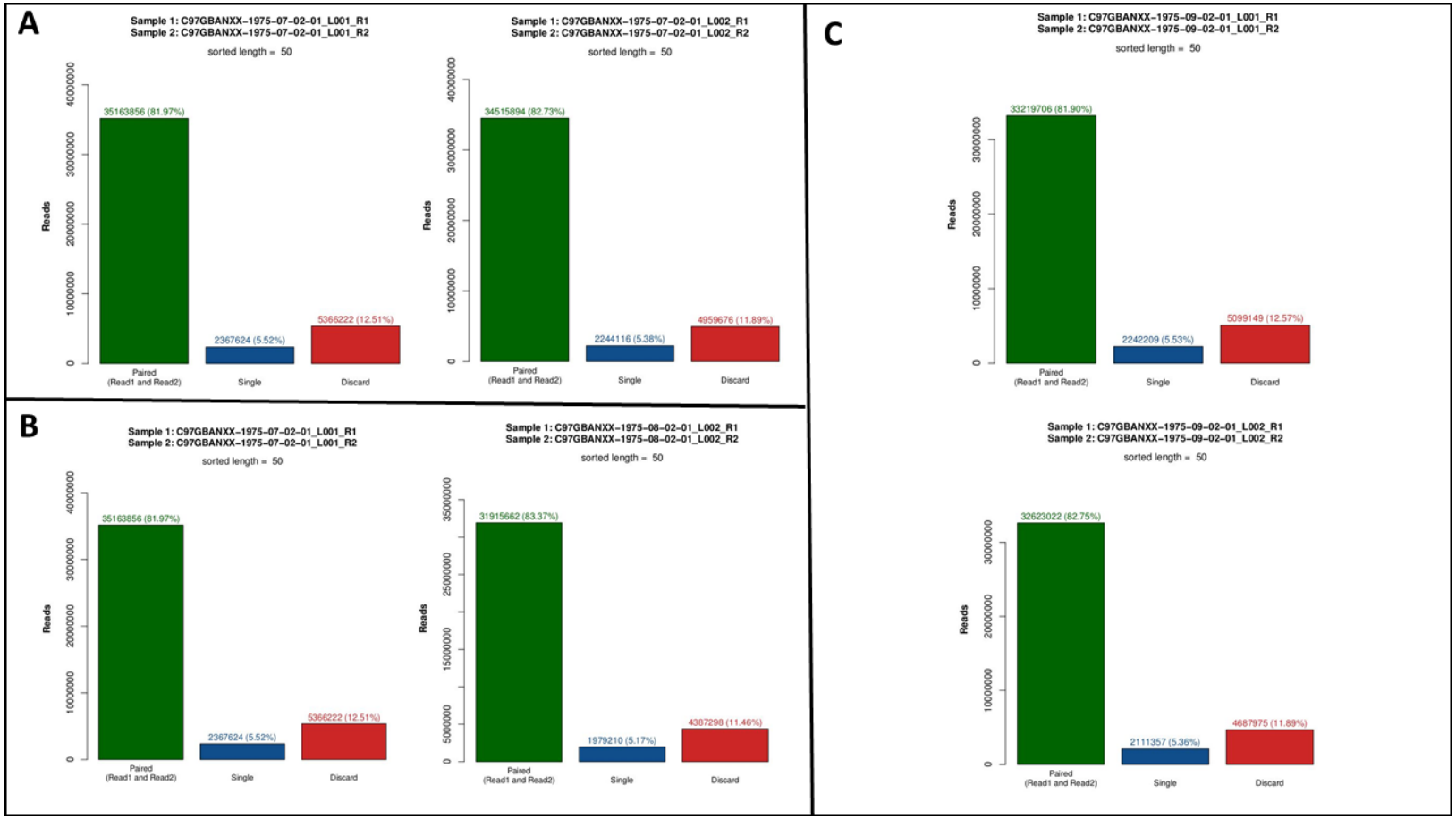


Figure A1.3: Summary of the trimming process of reads in Wol- testes replicates using the SolexaQA++ package. Green represent reads with both pairs retained, blue shows the number of reads whose pairs were discarded and red represent the reads the number of reads discarded.

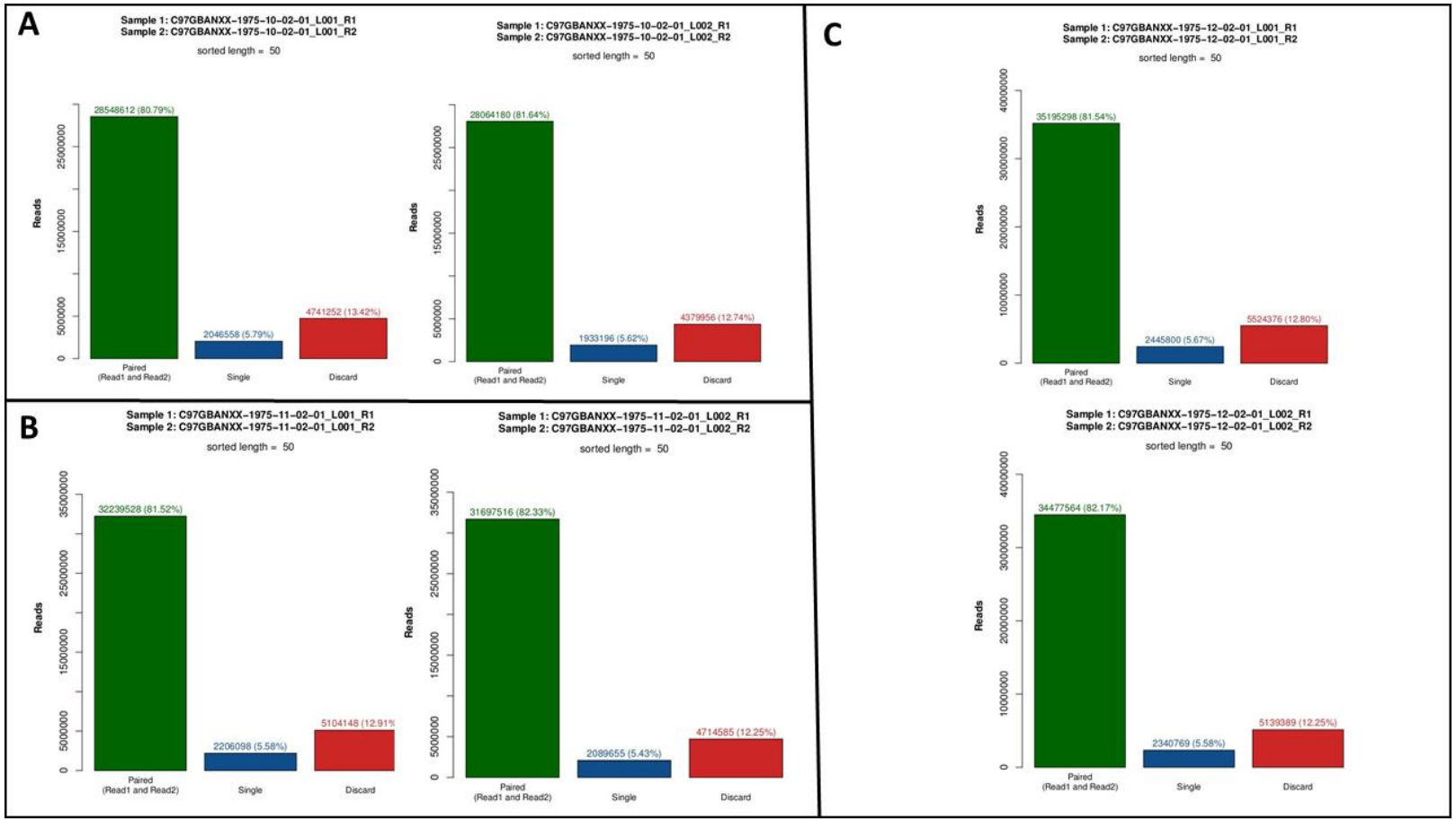
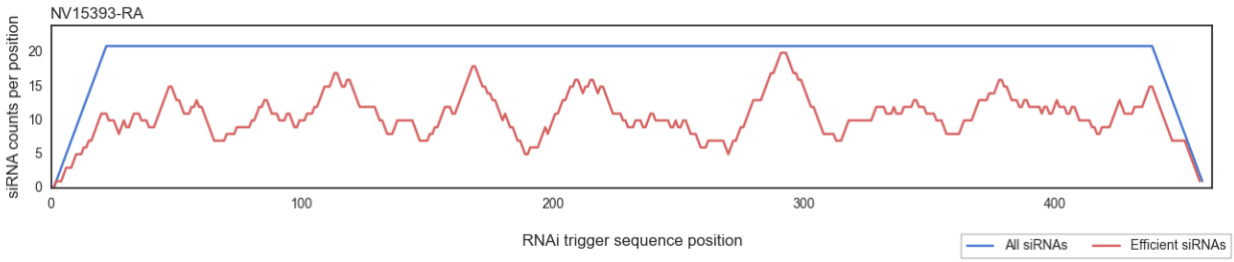


Figure A1.4: Summary of the trimming process of reads in Wol+ testes replicates using the SolexaQA++ package. Green represent reads with both pairs retained, blue shows the number of reads whose pairs were discarded and red represent the reads the number of reads discarded.

Appendix 2

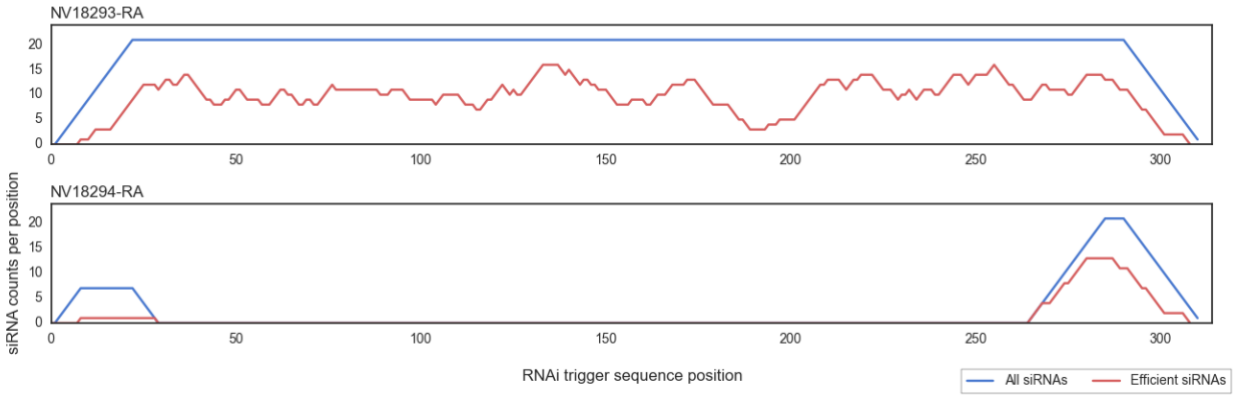
M1

Targets	Total siRNA hits	Efficient siRNA hits
NV15393-RA	438	233



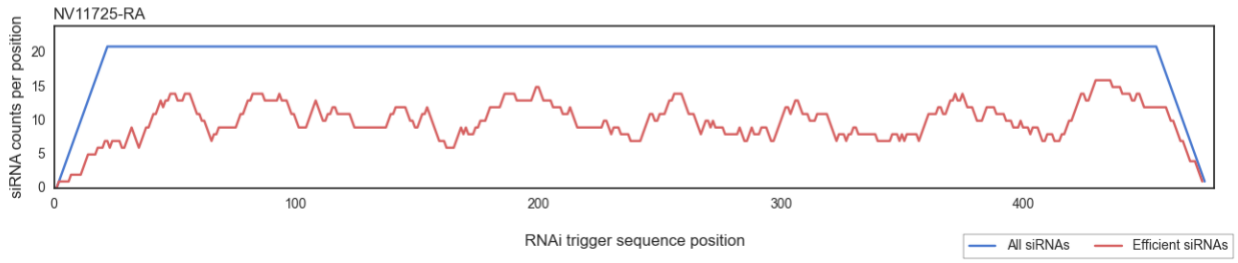
M2

Targets	Total siRNA hits	Efficient siRNA hits
NV18293-RA	289	141
NV18294-RA	33	16



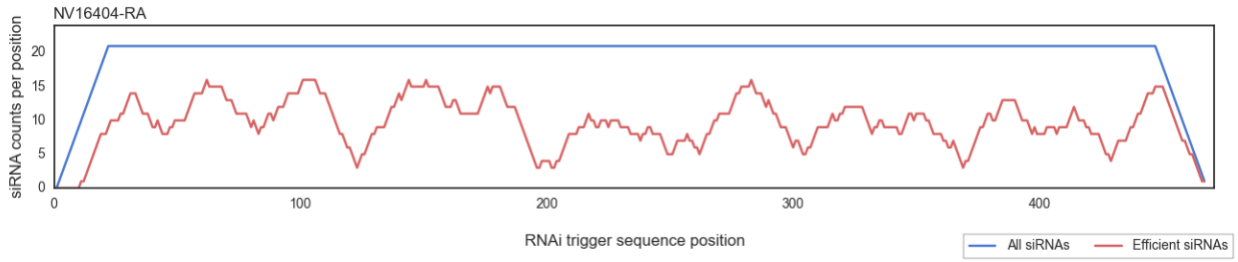
M3

Targets	Total siRNA hits	Efficient siRNA hits
NV11725-RA	454	224



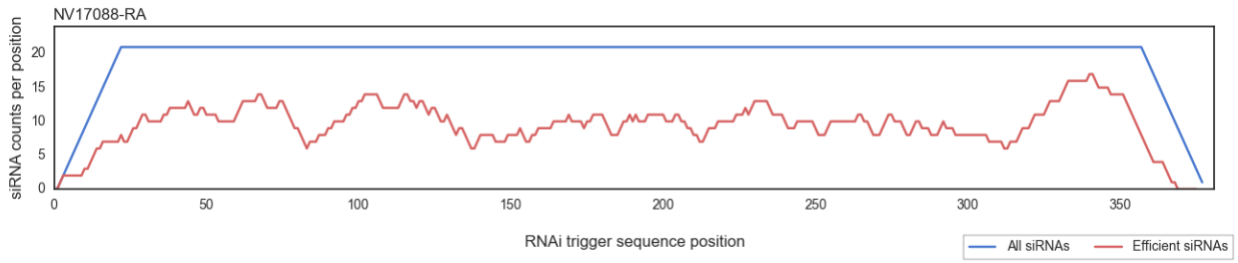
M4

Targets	Total siRNA hits	Efficient siRNA hits
NV16404-RA	446	214



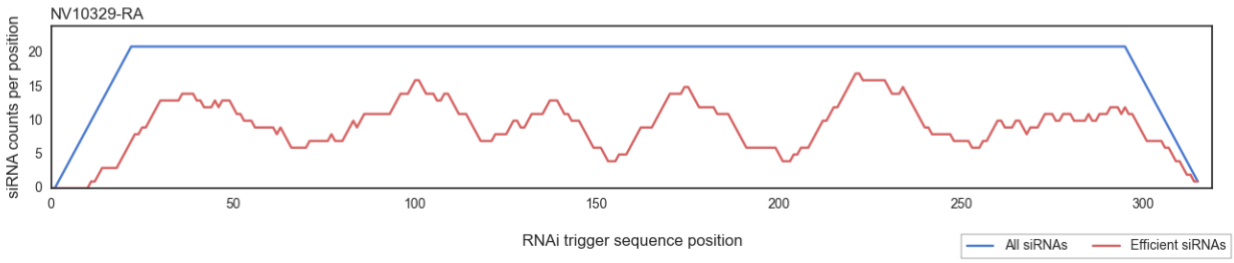
M5

Targets	Total siRNA hits	Efficient siRNA hits
NV17088-RA	356	171



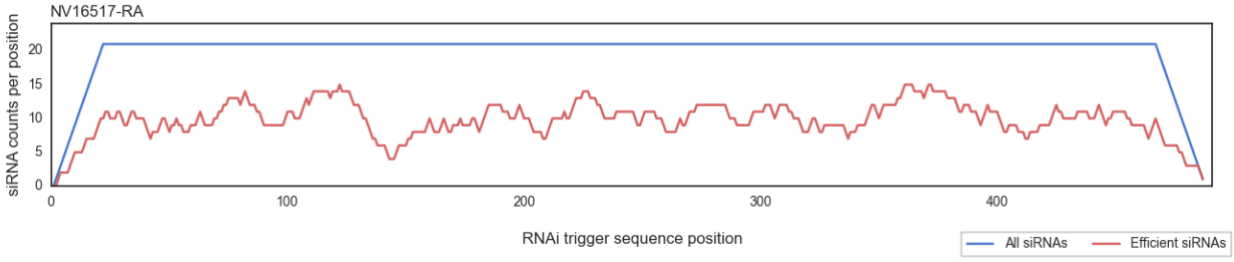
F1

Targets	Total siRNA hits	Efficient siRNA hits
NV10329-RA	294	139



F2

Targets	Total siRNA hits	Efficient siRNA hits
NV16517-RA	466	228



Appendix 3

Table A3.1: Normalized log_{10} counts of ovary tissue reads. Grey rows indicate genes that are expressed in all Wol+ and Wol- ovary tissue replicates.

Gene	Wol+ rep1	Wol+ rep2	Wol+ rep3	Wol- rep1	Wol- rep2	Wol- rep3
gene1038	15.38659268	20.74762147	7.899033435	0	61.29979072	0
gene1076	183.8993722	183.8993722	183.8993722	183.8993722	183.8993722	183.8993722
gene1092	1.109610049	2.011890567	0.811061469	237.9874228	2390.691838	1140.176107
gene1120	2.071272092	2.451991629	1.727913564	205.5345924	2022.893094	1103.396233
gene274	6.065868269	7.104488565	2.644765659	0	0	36.77987443
gene304	29.03479629	63.06019497	27.75240765	0	0	73.55974886
gene328	5.40010224	4.841111677	1.939494817	5.408805064	0	0
gene421	19.56612387	20.68474989	7.47587093	5.408805064	0	0
gene483	3.698700164	4.275267455	1.481068769	0	61.29979072	0
gene550	1.405506062	0.943073703	0.881588553	0	0	36.77987443
gene710	1274.017272	1261.958358	1244.803037	200.1257874	551.6981165	147.1194977
gene721	20.67573392	21.87930992	4.866368813	5.408805064	0	0
gene722	1.627428072	2.263376888	0.634743758	0	0	36.77987443
gene786	2871.226963	3177.152435	5251.658274	508.427676	1103.396233	1397.635228
gene790	12.35365855	13.89461923	4.090570886	5.408805064	0	0
gene805	7.693296341	9.619351775	4.125834428	5.408805064	0	0
gene821	3.217869143	3.017835851	1.093169806	5.408805064	0	0
gene963	2.589090115	1.949018987	0.6700073	0	0	36.77987443

Table A3.2: Normalized log_{10} counts of testes tissue reads. Grey rows:genes that are expressed in all Wol+ and Wol- ovary tissue replicates*indicate expression in two replicates of Wol- testes.

Gene	Wol+ rep1	Wol+ rep2	Wol+ rep3	Wol- rep1	Wol- rep2	Wol- rep3
gene1007	20.64095	10.00881	20.79942	0	17.49545	0
gene1017	4.243933	1.52766	4.4276	0	8.747727	0
gene1019	0	0.158034	0	0	8.747727	0
gene1055	0.771624	1.369626	1.544512	0	17.49545	0
gene1076	519.303	734.4885	1067.052	3215.82	1565.843	3245.891
gene1092	4.629745	4.161557	7.104754	778.779	848.5295	1064.226
gene1100	0.385812	0.158034	0.617805	0	8.747727	0
gene1105	0	0.052678	0.102967	0	8.747727	0
gene1106	0.192906	0.474101	1.132642	0	8.747727	0
gene1107	0.771624	0.158034	0.205935	0	8.747727	0
gene1120	8.680771	5.057082	15.54808	805.6334	743.5568	1210.558
gene1123	5.015557	3.05532	3.089023	0	26.24318	0
gene166	2.89359	1.000881	3.294958	0	8.747727	0
gene179	4.629745	2.159795	4.736503	0	8.747727	0
gene22	5.208463	3.002642	2.677154	0	17.49545	0
gene235	0.578718	0.737491	0.514837	6.713612	0	0
gene251	0	0.105356	0.205935	0	8.747727	0
gene252	1.92906	2.107118	2.677154	0	8.747727	0
gene275	0.192906	0.474101	0.926707	6.713612	0	0
gene279	1.543248	1.738372	1.853414	0	17.49545	0
gene29	1.92906	1.422304	4.015731	0	8.747727	0
gene304	87.57934	58.78858	40.46621	13.42722	17.49545	13.30283
gene311	0.771624	0.368746	0.617805	6.713612	0	0
gene316	4.822651	2.633897	4.942438	0	8.747727	0
gene327	2.507778	0.790169	2.265284	0	8.747727	0
gene34	0.192906	0.052678	0	6.713612	0	0
**gene346*	1.543248	0.579457	0.720772	6.713612	0	13.30283
gene361	0	0.210712	0.102967	0	8.747727	0
gene426	1.543248	0.684813	0.926707	0	8.747727	0

gene431	1.350342	0.632135	0.205935	0	8.747727	0
gene434	0.96453	1.264271	0.82374	0	8.747727	0
gene473	0	0.052678	0	0	17.49545	0
gene481	0.385812	0.210712	0.308902	0	8.747727	0
gene486	5.980087	4.530303	10.19378	0	17.49545	0
gene496	0.96453	0.158034	0.205935	0	8.747727	0
gene498	0	0.158034	0.102967	0	17.49545	0
gene5	2.121966	1.106237	2.162316	0	8.747727	0
**gene520*	4.051027	2.581219	3.089023	6.713612	17.49545	0
gene533	5.594275	2.05444	2.368251	0	17.49545	0
gene534	0.192906	0.316068	0.308902	0	8.747727	0
gene578	0.96453	0.842847	0.617805	0	8.747727	0
gene593	1.543248	0.474101	0.205935	0	8.747727	0
gene64	0.192906	0.105356	0.308902	6.713612	0	0
gene653	1.92906	1.738372	2.986056	0	8.747727	0
**gene687*	3.858121	1.738372	1.441544	13.42722	17.49545	0
gene699	0.96453	0.474101	0.514837	0	8.747727	0
gene7	0	0.105356	0.308902	0	8.747727	0
gene705	0.771624	0.210712	0.617805	0	17.49545	0
gene710	4726.391	4254.744	6603.508	463.2392	1032.232	345.8736
gene716	0.192906	0.052678	0.205935	0	8.747727	0
gene724	0.578718	0.895525	1.544512	0	8.747727	0
gene738	5.015557	2.897287	5.766177	0	8.747727	0
gene784	1.157436	0.737491	0.308902	0	8.747727	0
gene786	16533.59	17585.05	32185.57	711.6428	5458.581	1077.529
gene793	1.350342	1.053559	0.514837	0	8.747727	0
gene804	0.578718	0.26339	0	0	8.747727	0
gene805	11.76727	5.004404	15.85699	6.713612	0	0
gene818	0	0.158034	0.102967	0	8.747727	0
gene836	0.771624	0.26339	0	0	8.747727	0
gene920	0.385812	0.210712	0.617805	0	8.747727	0
gene922	2.121966	1.158915	1.750447	0	8.747727	0

gene976	3.472309	2.317829	5.045405	0	8.747727	0

Table A3.3: Normalized log_{10} counts of Wol- whole body reads.

Gene	Wol- rep1	Wol- rep2
gene1092	121.4106483	337.1010799
gene1120	10.11755402	19.82947529
gene710	222.5861885	158.6358023
gene786	60.70532413	19.82947529
gene1076	283.2915126	475.9074069

Table A3.4: Normalized log_{10} counts of female and male Wol- whole body reads.

Gene	Female rep1	Female rep2	Female rep3	Male rep1	Male rep2	Male rep3
gene1076	0	45.67365217	6.315369189	0	7.76153566	7.200801018
gene1092	5.405818736	2.537425121	3.157684595	0	7.76153566	7.68085442
gene1120	71.04790338	71.04790338	71.04790338	71.0479034	71.0479034	71.04790338
gene710	0	304.4910145	113.6766454	1.1277445	77.0183154	72.00801018
gene786	0	43.13622705	23.68263446	1.1277445	17.9112361	12.96144183

Table A3.5: Gene1120 inserts in the assembly raw Sanger reads found by BLASTn. The entire gene1120 (length = 256bp) was used as query to search the database and was shown to be present in the raw data.

Read name	% identity	Alignment length	e-value
gnl|ti|2131252149	99.606	254	2.65E-129
gnl|ti|2131251788	99.606	254	2.65E-129
gnl|ti|2131244597	99.606	254	2.65E-129
gnl|ti|1507751068	99.606	254	2.65E-129
gnl|ti|1507404775	99.606	254	2.65E-129
gnl|ti|1505803994	99.606	254	2.65E-129
gnl|ti|1505753599	99.606	254	2.65E-129
gnl|ti|1496963786	99.606	254	2.65E-129

gnl|ti|1496904802	99.606	254	2.65E-129

Appendix 4

>NP_001155168.1 venom protein Y precursor [Nasonia vitripennis]
MKAFTVLLAFAVIFVILFEEATATVPFQLATNDKSRCNTPFGQLSHRKRSIMFNGGSNSK
VFYKFIENTNSCVPFYGQELPPFYGFRTEAECDLHCRNIYHSQ

>NP_001155169.1 venom protein Z precursor [Nasonia vitripennis]
MFSRVLIFVLLGCALAFLQSSSAAEVQSKLRVRREQHGEKVAAWWHNLEDKVSKFAQ
NVKDKITNWWDKITGRAAYKAKQQELLKLKQEYEALKEESERAVRAKKEAWQKVLA
ENEKRLAELKNQAESEVEGEINGWNNKFDQLQEKAEVVIKS

>Serine Protease 97 [Nasonia vitripennis]
MRLFLILALCCFALAKSQIIGYYPGIGQPIQPIQPIQPIQPIQPVQPVGQGYPRFEAYAGEY
PYQVAIQVDGHAHCGGTLISKKHVLTAAHCTHDWILQRKDKTTIKVIVGTNDLNNGGT
VMNVARVSQHPQFRWYGPDVPILKHDVAVIRLTEEITESDTVKPISLPAANSEIAANTRL
ILTGFGATYAGGPSSSVLRHIYLYVTDHNTCSINWLNRGKITTDHLCATLAPGYGACNG
DSGGPLVLEDKSTVVGIVSEGIGLGHGCGSGWPDLFTKVSHHLEFINGELALTN

>XP_008203717.1: venom protein G isoform X1 [Nasonia vitripennis]
MMKIALFLAVGLVVVFNSPVNAELTEEDGQKILDAIKDEVAKQSLQCAGIKSYCSHDVP
SEIPAILCPNLLKAYLCDNVNSAAGINQVLVDIAKRVAENQKN

>NP_001155161.1 venom protein Q precursor [Nasonia vitripennis]
MYTKVLILVLLGSTLILAKPTADVEVEADTESSSWWSRAASKVMNWWSSGKNAVERT
LQKIKEQGHYAGELAEDGYLYLKDKGIELAKKSAAALEEEKQKWEQKIQQVKQSTKET
ADTVEYKVKQLQEQAKKAIEEKVREAKKTWNQWKQEAREAYEKEQQKQL

>NP_001155167.1 venom protein X precursor [Nasonia vitripennis]
MYKFIIFCVFGLAFVNLNAASEDDRRSQINQTKSYTKMVVEENNLTCENYLTEYCKSSH
LSIYCQIIMKGYVCDNIRDGEAFDQLYDQFKAALLPLVANRTSDEPPIKSSEN

Appendix 5

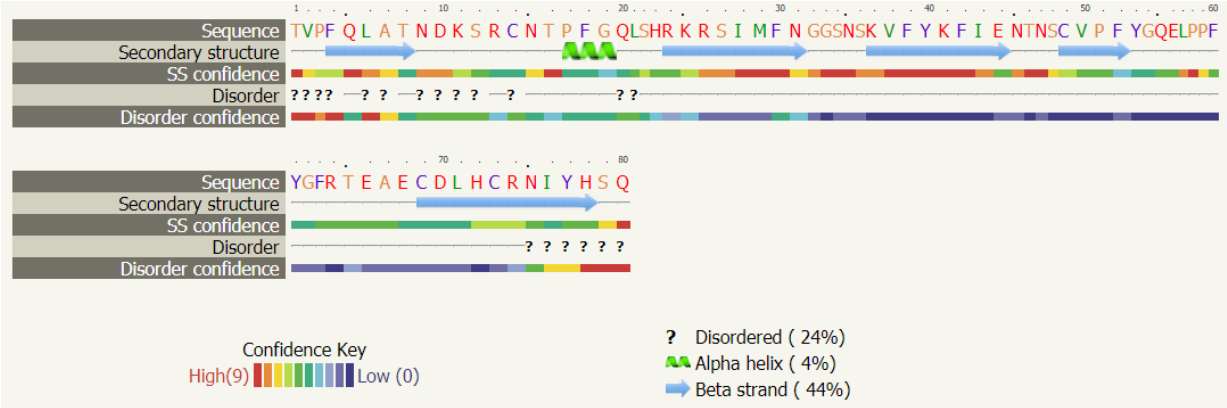


Figure A5.1. Visual representation of the predicted secondary structure of Venom Y

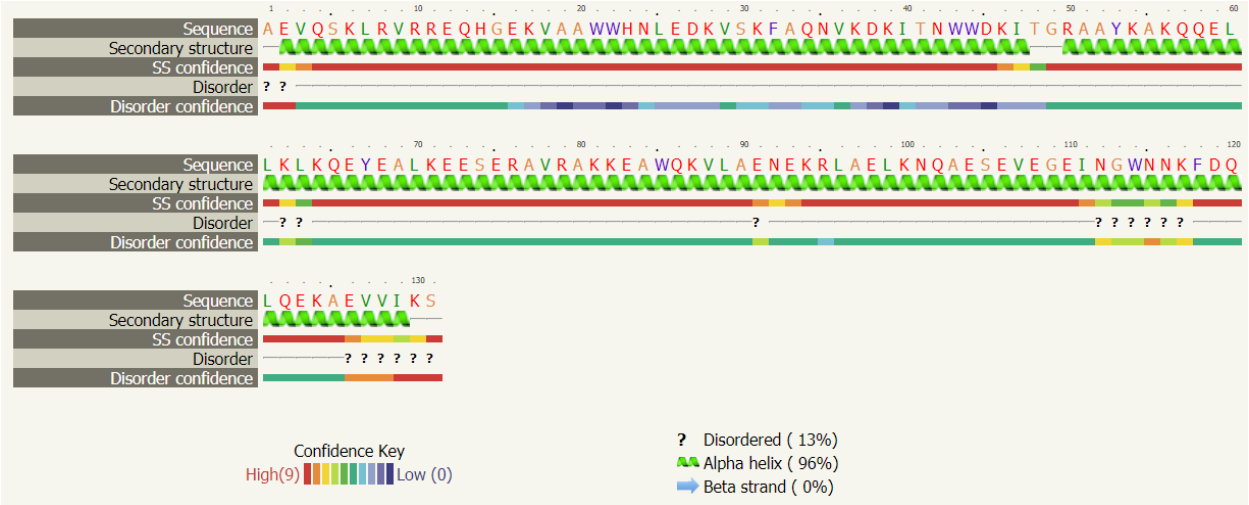


Figure A5.2. Visual representation of the predicted secondary structure of Venom Z

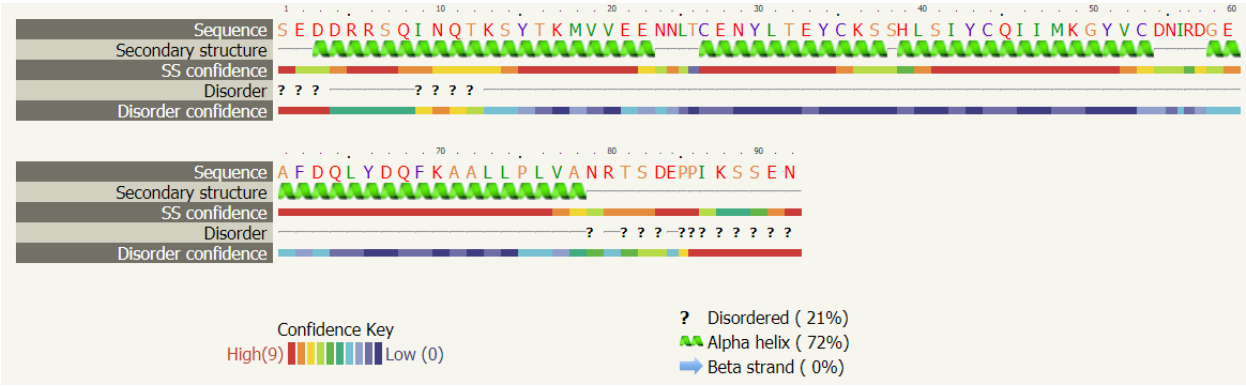


Figure A5.3 Visual representation of the predicted structure of Venom X

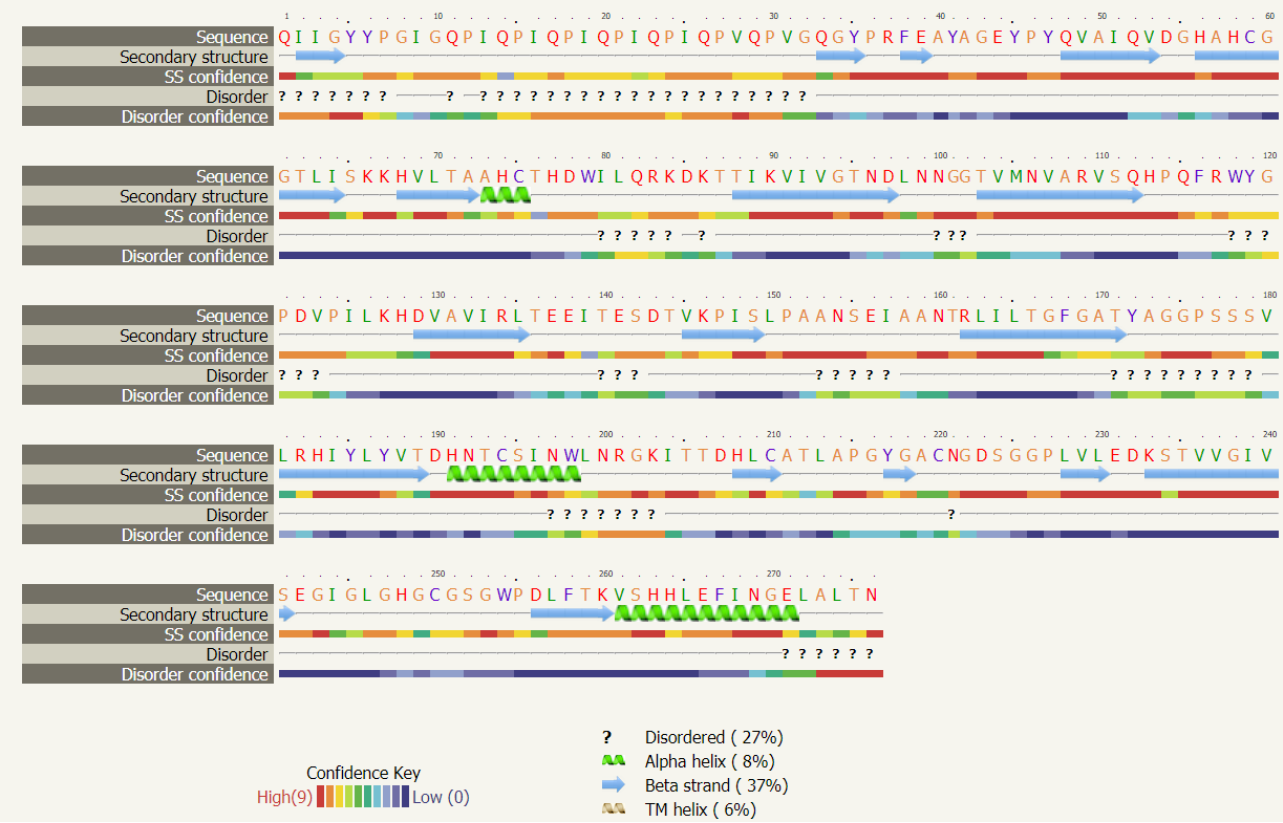


Figure A5.4 Visual representation of the predicted secondary structure of SP97

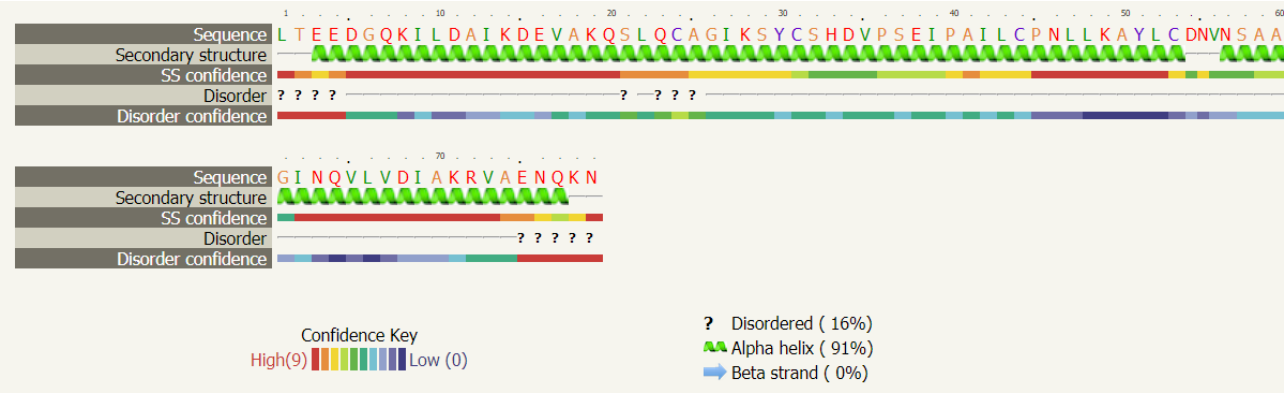


Figure A5.5. Visual representation of the predicted secondary structure of Venom G.

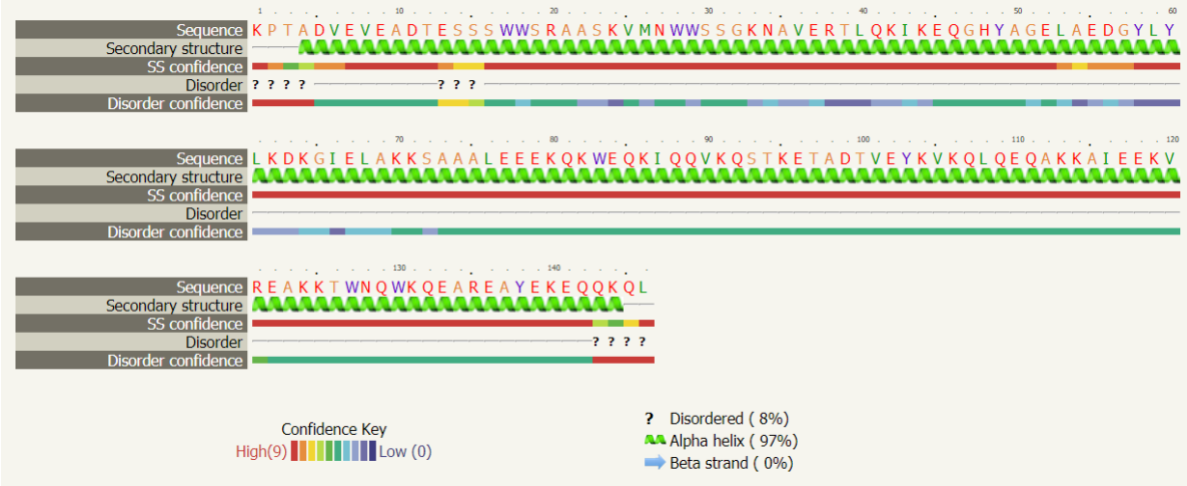


Figure A5.6. Visual representation of the predicted secondary structure of Venom Q.

Appendix 6

Table A6.1. Summary of the VY-KD stinging assay. N is the number of hosts that were exposed to the wasps.

Wasp	N	Unstung hosts	Stung + no eggs	Stung + eggs	Host Phenotype
Unstung	10 10 10	- - -	- - -	- - -	Host developed
Wild Type	10 10 10	0 1 0	0 0 0	10 9 10	No development
VY-KD	19 30 26	4 3 5	0 0 0	15 (9) 27 (13) 21 (17)	Increased melanisation
	75	12	0	63 (39)	

*In brackets are the numbers of hosts that showed the described phenotype

Table A6.2. Summary of the VZ-KD stinging assay. N is the number of hosts that were exposed to the wasps.

Wasp	N	Unstung hosts	Stung + no eggs	Stung + eggs	Host Phenotype
Unstung	10 10 10	- - -	- - -	- - -	Host developed
Wild Type	10 10 10	0 0 0	0 0 0	10 10 10	No development
VZ-KD	15 22 20	3 8 5	12 (10) 14 (8) 15 (9)	0 0 0	No eggs laid + Host dying quicker than the wild type
	57	16	31(27)	0	

*In brackets are the numbers of hosts that showed the described phenotype

Table A6.3. Summary of the SP97-KD stinging assay. N is the number of hosts that were exposed to the wasps.

Wasp	**N**	**Unstung hosts**	**Stung and no eggs**	**Stung + eggs**	**Host Phenotype**
Unstung	10 10 10	- - -	- - -	- - -	Host developed
Wild Type	10 10 10	0 1 2	0 0 0	10 9 8	No development
SP97-KD	21 31 36	2 4 8	0 0 0	19 27 28	No development
	88	14		74	

Table A6.4. Summary of the VG-KD stinging assay. N is the number of hosts that were exposed to the wasps.

Wasp	N	Unstung hosts	Stung and no eggs	Stung + eggs	Host Phenotype
Unstung	10 10 10	- - -	- - -	- - -	Host developed
Wild Type	10 10 10	0 0 0	0 0 0	10 10 10	No development
VG-KD	12 33 30	0 4 2	0 0 0	12 29 28	No development
	75	6	0	69	

Table A6.5. Summary of the VQ-KD stinging assay. N is the number of hosts that were exposed to the wasps.

Wasp	N	Unstung hosts	Stung and no eggs	Stung + eggs	Host Phenotype
Unstung	10 10 10	- - -	- - -	- - -	Host developed
Wild Type	10 10 10	0 1 0	0 0 0	10 9 10	No development
VQ-KD	12 35 30	5 2 1	0 0 0	7 33 29	No development
	77	8	0	69	

Table A6.6. Summary of the VX-KD stinging assay. N is the number of hosts that were exposed to the wasps.

Wasp	N	Unstung hosts	Stung and no eggs	Stung + eggs	Host Phenotype
Unstung	10 10 10	- - -	- - -	- - -	Host developed
Wild Type	10 10 10	0 1 0	0 0 0	10 9 10	No development
VX-KD	21 30 27 33	4 3 5 2	2 0 3 0	15 (7) 27 (21) 19 (13) 31 (24)	Host development
	111	14	5	92 (65)	

*In brackets are the numbers of hosts that showed the described phenotype