Probing Insecticide Biology Using
*Drosophila melanogaster*

Shane Denecke

Submitted in total fulfilment of the requirements of the degree of Doctor of Philosophy

March, 2017

School of Biosciences

The University of Melbourne
Abstract

Insecticides are often used to control insect pests, but resistance to these chemicals arises quickly, leading to agricultural losses and public health concerns. Understanding how insects cope with insecticides is necessary when designing rational pest management strategies, but much still remains unknown regarding the fate of insecticides once inside the body. Furthermore, the genetic variation that governs an insect’s ability to survive insecticide exposures has not been fully described.

Here, a 3-pronged approach is applied to study insecticide biology using the model insect *Drosophila melanogaster*. First, an acute, sub-lethal insecticide response assay was developed, which provided information complementary to that obtained from more common toxicology assays. In particular, behavioural response observed in a hyper-resistant target site mutant suggests additional target sites for the insecticide spinosad. This bioassay was then applied in a forward genetics approach to describe the genetic basis of resistance to the insecticide imidaclorpid. This approach identified a variety of neuronal genes and the previously identified drug metabolizing enzyme *Cyp6g1*, which was explored through genetic manipulation. Finally, a reverse genetics approach was employed in order to study the effect of an ABC transporter protein *Mdr65* on insecticide resistance. Removing the gene made the insects more susceptible to a subset of the insecticides tested, and this was confirmed with genetic and chemical complementation tests.

These data provide information both on the genetics and kinetics of insecticide biology. Such information will help to better understand insecticide resistance and design rational resistance management strategies.
Declaration

This is to certify that:

i. The thesis comprises only my original work towards this Ph.D. except where indicated in the Preface;

ii. Due acknowledgement has been made in the text to all other material used;

iii. The thesis is less than 100,000 words in length, exclusive of tables, maps, bibliographies and appendices.

Shane Denecke

March 2017
Preface

I would like to acknowledge the following people for contributing directly to this work:

Chapter II:

Trent Perry: For assisting in the revision of the Chapter II Paper
Alexandre Fournier-Level: For assisting in the analysis of Wiggle Index data
Cameron Nowell: For writing the Wiggle Index script

Chapter III:

Roberto Fusetto: For performing all HPLC analysis and generating the PCFD4 Plasmid
Paul Battlay: For assisting in the transcriptomic analysis of the DGRP
Llewellyn Green: For assisting in the genomic analysis of the DGRP
Jenny Luong: For assisting the Wiggle Index screen of the DGRP

Chapter IV:

Roberto Fusetto: For performing all HPLC analysis
Acknowledgements

I don’t want to make this too sentimental, so I’ll simply thank everyone who helped me either in my academic or personal life. This work would not have attained the same quality (or may not have been completed without you). I acknowledge the full magnitude of your contribution, while taking complete responsibility for any errors or shortcomings present in the work as my own.

Although there are many who fit such a description, several merit individual acknowledgement:

- Everyone in the Batterham lab, for making this a fun place to work and for good scientific discussions.
- Chris Lumb, Tinna Yang, and Jason Somers for teaching most of the techniques I used during my project.
- Harshini Weerasingha and Llewellyn Green specifically and housemates in general for good scientific and non-scientific discussions.
- Alexandre Fournier-Level and Trent Perry for their expertise in data science and biology respectively
- My supervisor Phil Batterham. He conned the University of Melbourne enrolment system in order to take a chance on me, at a time when we had only exchanged a brief correspondence. And for that, I thank him.
- Roberto Fusetto, Felipe Martelli, and Jenny Luong with whom I had productive collaborations.
- Lastly, my family. My parents Chuck and Amy, my siblings Lauren and Aidan and my aforementioned partner Jenny. Couldn’t have done it without you.
Table of Contents

Chapter I: Introduction........................................................................................................... 1
  1.1) The Impact of Insecticide Resistance......................................................................... 2
    1.1.1) Agricultural Pests............................................................................................... 2
    1.1.2) Vector Borne Diseases.......................................................................................... 2
    1.1.3) Overview of this Thesis......................................................................................... 2
  1.2) The History of Pest Management Strategies.............................................................. 3
    1.2.1) Non-Chemical Methods of Managing Resistance................................................ 3
    1.2.2) Naturally Occurring Pesticides............................................................................. 3
    1.2.3) Synthetic Pesticides.............................................................................................. 4
  1.3) Insecticide Resistance................................................................................................. 6
    1.3.1) History of Insecticide Resistance......................................................................... 6
    1.3.2) Exaptation............................................................................................................ 6
    1.3.3) The Mono v. Polygenic Debate............................................................................. 7
    1.3.4) Evolution at the Gene Level................................................................................. 9
  1.4) Common Resistance Mechanisms............................................................................... 10
    1.4.1) Target Site Mediated Resistance......................................................................... 10
      1.4.1.1) Modes of Action............................................................................................. 10
      1.4.1.2) Types of Target site Modification.................................................................... 10
      1.4.1.3) Insecticide:Receptor Crosstalk....................................................................... 11
      1.4.1.4) nAChRs......................................................................................................... 11
    1.4.2) Drug Metabolizing Enzymes (DMEs).................................................................... 13
      1.4.2.1) The Structure and Function of DMEs............................................................. 13
      1.4.2.2) Cytochrome P450s Biochemistry................................................................... 13
      1.4.2.3) Cytochrome P450 Evolution and Diversity.................................................... 15
      1.4.2.4) Cytochrome P450 Biology............................................................................ 15
      1.4.2.5) The Molecular Genetics of Cytochrome P450 Mediated Resistance............ 17
    1.4.3) Xenobiotic Transport............................................................................................. 17
      1.4.3.1) Overview........................................................................................................ 17
      1.4.3.2) The Structure Function of ABC Transporters.............................................. 17
      1.4.3.3) ABC Transporter Phylogeny........................................................................ 18
      1.4.3.4) P-glycoprotein.............................................................................................. 18
      1.4.3.5) P-glycoprotein in Insects.............................................................................. 20
  1.5) Insecticide Biology and Insecticide Resistance.......................................................... 20
    1.5.1) Introduction........................................................................................................ 20
    1.5.2) Mortality assays.................................................................................................. 22
      1.5.2.1) Dosing Methods and Duration.................................................................... 22
      1.5.2.2) Life stage and Scoring Mortality.................................................................. 22
    1.5.3) Sub-lethal Traits.................................................................................................. 23
      1.5.3.1) Why Look at Sub-lethal Traits?..................................................................... 23
      1.5.3.2) Behavioural Assays...................................................................................... 24
      1.5.3.3) Sub-lethal Biochemical Changes................................................................. 24
  1.6) Forward Genetics: From Phenotype to Genotype....................................................... 25
    1.6.1) Overview and history.......................................................................................... 25
      1.6.1.1) What is Forward Genetics............................................................................. 25
      1.6.1.2) Sources of Genetic Variation....................................................................... 25
      1.6.1.3) Existing Genomic Variation......................................................................... 26
1.6.2) Association Studies........................................................................................................26
  1.6.2.1) Mapping Crosses........................................................................................................26
  1.6.2.2) The Genome Wide Association Study (GWAS)..........................................................26
  1.6.2.3) Transcriptomic Association Studies..........................................................................27

1.7) Reverse Genetics: From Genotype to Phenotype.................................................................28
  1.7.1) The Basics of Reverse Genetics......................................................................................28
    1.7.1.1) What is Genetic Manipulation....................................................................................28
    1.7.1.2) Brief History of Genetic Manipulation.........................................................................28
    1.7.1.3) Transformation in Insects and Microinjection...........................................................28
  1.7.2) Genetic Biotechnology in D. melanogaster.......................................................................29
    1.7.2.1) Integrase and Recombinases......................................................................................29
    1.7.2.2) Nucleases...................................................................................................................29
  1.7.3) Applications of the genetic toolkit..................................................................................32
    1.7.3.1) Manipulating Gene Expression...................................................................................32
    1.7.3.2) Disrupting Gene Function........................................................................................34
    1.7.3.3) Recent Advances in CRISPR..................................................................................34

1.8) Outline of This Work............................................................................................................34

Chapter II: The Wiggle Index: An Open Source Bioassay to Assess Sub Lethal
Insecticide Response in Drosophila melanogaster.................................................................36

2.1) Introduction..........................................................................................................................38
2.2) Methods...............................................................................................................................39
  2.2.1) Fly Stocks.......................................................................................................................39
  2.2.2) Obtaining Third Instar Larvae......................................................................................40
  2.2.3) Insecticide Dilution........................................................................................................40
  2.2.4) Experimental Design....................................................................................................40
  2.2.5) Filming..........................................................................................................................40
  2.2.6) Video Processing and Analysis of Larval Movement.....................................................41
  2.2.7) Visualization and Statistics............................................................................................42

2.3) Results..................................................................................................................................43
  2.3.1) The Response of Wild Type to Different Insecticides.....................................................43
  2.3.2) Set One: Imidacloprid Resistant nAChR Alleles............................................................43
  2.3.3) Set Two: Spinosad Resistant Dδ6 Alleles........................................................................44
  2.3.4) Set Three and Four: Cyp6g1 Overexpression.................................................................44

2.4) Discussion............................................................................................................................45
  2.4.1) Armenia14 Dose Response Profiles..............................................................................45
  2.4.2) Target Site Mutants.......................................................................................................46
  2.4.3) Cyp6g1..........................................................................................................................49
  2.4.4) The Utility of the WI.....................................................................................................50

2.5) Conclusion...........................................................................................................................51

Chapter III: Using the Drosophila Genetic Reference Panel to Describe the
Genetic Basis of Imidacloprid Response...........................................................................55

3.1) Introduction..........................................................................................................................56
  3.1.1) Imidacloprid Resistance: What is known.......................................................................56
  3.1.2) Imidacloprid Resistance: What is Unknown.................................................................60
  3.1.3) Forward Genetics to Address the Unknowns of Imidacloprid Resistance.....................61
3.2) Materials and Methods..........................................................................................................62
  3.2.1) Fly Genotypes...............................................................................................................62
  3.2.2) The Wiggle Index Bioassay............................................................................................64
  3.2.3) Analysis of the Wiggle Index and Candidate Gene Selection.......................................64
  3.2.3) The Deletion of Cyp6g1................................................................................................65
Figures List

Chapter I: Introduction

Figure 1.1: IRAC Mode of Action Poster ................................................................. 5
Figure 1.2: Response to Selection ........................................................................... 8
Figure 1.3: Nicotinic Acetylcholine Receptors ....................................................... 12
Figure 1.4: Drug Metabolizing Enzymes ................................................................. 14
Figure 1.5: Cypomes Among Arthropods ............................................................... 16
Figure 1.6: ABC Transporters ............................................................................... 19
Figure 1.7: P-glycoprotein ..................................................................................... 21
Figure 1.8: AttB Integration ................................................................................... 30
Figure 1.9: CRISPR-Cas9 ..................................................................................... 31
Figure 1.10: GAL4-UAS ...................................................................................... 33

Chapter II: The Wiggle Index: An Open Source Bioassay to Assess Sub-Lethal Insecticide Response in Drosophila melanogaster

Figure 2.1: The Wiggle Index ................................................................................ 41
Figure 2.2: Wiggle Index Output and RMR Calculation ........................................ 43
Figure 2.3: Armenia14 Dose Response Curves ...................................................... 44
Figure 2.4: Armenia14 Short time Point Dose Response Curves ............................ 45
Figure 2.5: Imidacloprid Resistance in nAChR Mutants ....................................... 46
Figure 2.6: Spinosad Resistance in nAChR Mutants ............................................. 47
Figure 2.7: Cyp6g1 Expression in the Digestive Tissues ......................................... 48
Figure 2.8: Cyp6g1 Expression in the Central Nervous System ............................. 48

Chapter III: Using the Drosophila Genetic Reference Panel to Describe the Genetic Basis of Imidacloprid Response

Figure 3.1: Cyp6g1 Structural Variation ................................................................. 58
Figure 3.2: Imidacloprid Metabolism .................................................................. 59
Figure 3.3: RAL_517 Crossing Scheme ................................................................. 67
Figure 3.4: Phenotypic Spread of the DGRP ......................................................... 69
Figure 3.5: Correlation of Wiggle Index Response Between RMR values ............. 69
Figure 3.6: QQ Plots ............................................................................................. 70
Chapter IV: Using CRISPR-Cas9 to Examine the Role of the P-gp Homologue Mdr65 in Resistance to Multiple Insecticides

Figure 4.1: ABC Transporter Phylogeny in D. melanogaster ...........................................93
Figure 4.2: Transporters at the Blood Brain Barrier .........................................................96
Figure 4.3: Expression of the ABCB Full Transporters .....................................................96
Figure 4.4: Mdr65KO Schematic .....................................................................................101
Figure 4.5: Mdr65KO Crossing Scheme .........................................................................102
Figure 4.6: The Relative Susceptibility of Mdr65KO to a Panel of Insecticides ..............103
Figure 4.7: Mdr65KO crossed to Mdr65 Deficiencies .....................................................104
Figure 4.8: Synergism with the ABC Inhibitor Verapamil .............................................106
Figure 4.9: HPLC-MS of Nitenpyram in Mdr65KO .......................................................107
Tables List

Chapter II: The Wiggle Index: An Open Source Bioassay to Assess Sub-Lethal Insecticide Response in Drosophila melanogaster

Table 2.1: Genotypes Used in the Present Study

Chapter III: Using the Drosophila Genetic Reference Panel to Describe the Genetic Basis of Imidacloprid Response

Table 3.1: Genotype List for Chapter III
Table 3.2: Insecticide Response Correlations
Table 3.3: GWAS candidates
Table 3.4: TWAS candidates
Table 3.5: HPLC-MS Significance

Chapter IV: Using CRISPR-Cas9 to Examine the Role of the P-gp Homologue Mdr65 in Resistance to Multiple Insecticides

Table 4.1: List of Genotypes Used in Chapter IV
Table 4.2: Ratios between Heads and Bodies of Wxac and Mdr65KO
Chapter I: Introduction
1.1) The Impact of Insecticide Resistance

1.1.1) Agricultural Pests

The United Nations estimates that the global population will continue to grow from its current figure of roughly 7 billion to 9 billion by 2050. In order to sustain this increase, food production must grow proportionally, a challenging goal which has been increasingly discussed in scientific literature (Stefanis, 2014). One of the main obstacles to increased food production is crop loss due to destructive pest species such as insects. Estimates of crop losses due to animal pests vary widely from 7-15% (Oerke, 2006) to 30-40% (Flood, 2010). Given the demand for increased food production, even the lower estimates are alarming. As several aspects of more industrialized farming facilitate the flourishing of pest species, pest management will become increasingly important as many regions of the world develop economically (Bebber et al., 2014).

1.1.2) Vector Borne Diseases

Another challenge associated with an increasing global population is posed by vector borne diseases such as Malaria, Dengue, and, most recently, Zika. Each of these diseases must complete part of their life cycle in insect hosts. These diseases impact millions of people, especially developing countries. Although still unacceptably high, recent figures indicate that per capita rates for some of these diseases (for example malaria) are declining. While many factors likely contribute to this decrease, the effective control of insect vectors is likely one of the most significant (Lindblade et al., 2015). Vector control was deemed to be so critical in preventing vector born disease that in 2006 the World Health Organization (WHO) lifted its ban on DDT, arguing that the costs of vector born disease outweighed the environmental and public health hazards of DDT use (WHO, 2010).

1.1.3) Overview of this Thesis

Despite critical economic and public health concerns, efforts to control pest populations often fall short of their goal. Although a variety of approaches are used (Section 1.2), synthetic pesticides have been far and away the most successful way to mitigate the damage caused by pest populations. However, these pesticides apply selection pressure to the insect populations they target, meaning that the introduction of a pesticide is quickly followed by the appearance of resistance phenotypes caused by various physiological mechanisms (Section 1.3). Understanding these mechanisms, and their genetic basis, is paramount for effective pest management, but the current
understanding of the interface between insecticides and the biology of insects lacks depth. The aim of the current work is to better understand the way an insect responds to insecticide exposure by using alternative insecticide bioassays (Chapter II, Section 1.5), forward genetic screens (Chapter III, Section 1.6) and precise genetic manipulation (Chapter IV, Section 1.7) in the model organism *Drosophila melanogaster*. Doing so will help to better understand how insects cope with synthetic pesticides and allow for more rational pest management strategies.

1.2) The History of Pest Management Strategies

1.2.1) Non-Chemical Methods of Managing Resistance

The desire to eliminate or manage troublesome insects is nothing new; societies have tried to control pests since ancient times. Many of the earliest pest control strategies did not involve the use of chemicals. In the Chinese Western Jin Dynasty, natural predators were introduced into an area to consume a damaging pest species (Fishel, 2013). More mechanical methods have also been employed. The spatial alternation of crops species can inhibit a monophagous pest species by preventing a single outbreak from spreading across an entire farm. Netting and other physical barriers have also been used to separate pest from plant. More recently, sophisticated physiological techniques have been adopted to combat pest populations. Many insects harbour symbiotic, intra-cellular Wolbachia bacteria that can be inherited vertically or horizontally, and which can cause cytoplasmic incompatibility. Infecting pest insects such as *Aedes aegypti* with the certain strains of Wolbachia has been used to crash field populations (Hoffmann et al., 2011). Meiotic drive can also be used to promote the spread deleterious alleles through a pest population (Lindholm et al., 2016).

1.2.2) Naturally Occurring Pesticides

Another way to combat pest species is through pesticidal secondary metabolites produced as defence compounds by plants and microbes. Around the time the infamous Roman general Marius was elected to his first consulship (circa 100 B.C), the use of sulphur and hellebore to ward off rats and insects was commonplace (Fishel, 2013). Their Persian contemporaries in Ctesiphon were also using dried flowers from the genus Pyrethrum to deter a wide variety of insect species (Casida and McLaughlin, 1973). Jumping forward into more recent times, nicotine extracted from tobacco plants was sprayed on crops in the 18th through 20th centuries as a potent pesticide, effective in killing many agricultural pests (Roberto and O’Hair, 2015). While nicotine and pyrethrum are not very insect specific and pose health risks to humans, more recently developed naturally occurring toxins are some of the most effective in the pest control toolkit. A potent insecticide known as Bt
toxin is a protein product of the bacterial species *Bacillus thuringiensis*, and can be transgenically expressed in plants to discourage pest species from eating them (Heckel, 2012). Other spray based insecticides such as the macrolactones are by-products of fungal and bacterial species and are now mass produced as organic insecticides (Kirst et al., 1992).

### 1.2.3) Synthetic Pesticides

Today the most widespread method to control agricultural and epidemiological insect pests is through the use of synthetic insecticides. These compounds are synthesized from defined chemical reactions in laboratory settings. DDT was the first synthetic pesticide to achieve widespread use and popularity. Originally patented by Paul Müller while working for the Geigy chemical company in the 1940s, DDT was soon thought of as a miracle chemical because of its potency as an insecticide and low acute toxicity towards humans (Jarman and Ballschmiter, 2012). It was only years later that the chronic toxicity of DDT to non-target species was fully understood (Carson, 1962). Between the commercialization of DDT and the early 1990s there was an explosion of new insecticidal compounds that can be categorized into classes based on their chemical structure and modes of action (MoA) based on their biological activity (IRAC; Fig 1.1). Interestingly, some of these classes bear structural resemblance to naturally occurring plant secondary metabolites (PSMs). For example, pyrethroids, which target the sodium gated *para* ion channel mimic the naturally occurring chemical pyrethrum (Casida and McLaughlin, 1973).

Also derived from a PSM, in this case nicotine, are the neonicotinoids (Jeschke and Nauen, 2008). As their name suggests, they were originally discovered using nicotine as a lead compound but improve upon nicotine by acting far more selectively on insects compared to mammals. The first neonicotinoid to be released was imidacloprid in 1991, shortly followed by nitenpyram (1995), acetamiprid (1995), thiacloprid (2000) and clothianidin (2002) among others. Although similar in terms of chemical structure, there are key differences between members of the neonicotinoid family which are not yet fully understood. For instance, certain honeybee P450s are capable of metabolizing thiacloprid but not imidacloprid (Nauen unpublished). Compared to other insecticides, neonicotinoids, and in particular imidacloprid, are widely used in both agricultural and academic settings. This makes imidacloprid an ideal model insecticide for use in the current work.
Figure 1.1- IRAC Mode of Action Poster:
Modern synthetic insecticides act on a variety of different tissues and target proteins. Here, insecticides are grouped according to their mode of action (Moa) as classified by the Insecticide Resistance Action Committee (IRAC). Each Box represents a distinct Moa and colours correspond to the tissue that the target site is primarily found. This image can be obtained at http://www.irac-online.org/modes-of-action/.
1.3) Insecticide Resistance

1.3.1) History of Insecticide Resistance

There will be variation among individuals in all insect populations that determines the likelihood that an individual insect will be killed by a pest management strategy. Consequently, each strategy described in Section 1.2 applies a selection pressure, which tends to eliminate the individuals that are the most susceptible to the treatment. Survivors of the treatment will then have an increased tendency to pass on their resistance genes to the following generation, which leads to the gradual emergence of a resistant populations. Thus, effective pest management strategies often engender evolutionary responses, which obviate the strategies themselves.

This is especially true with synthetic insecticides due to the strong selective pressures they impose (Forgash, 1984); the introduction of an insecticide to the market is often quickly followed by the detection of resistant individuals in the areas where it is deployed. DDT is a good example of this phenomenon. In a laboratory setting, significantly higher levels DDT resistance was observed in Musca domestica after only 14 generations of rearing in an insecticide treated environment (Lindquist and Wilson, 1948). These findings match field data which observe the appearance of DDT resistant populations of Musca domestica shortly after the widespread use of DDT (Metcalf and March, 1949). A similar pattern was observed for several other insecticides where a relatively small time gap exists between the introduction of a chemical insecticide and the appearance of resistance (Forgash, 1984).

There are, however, some insecticides that do not follow this trend and elicit a much slower evolutionary response. Insect growth regulators like cyromazine and dicyclanil were used for 30 years to control the Australian sheep blowfly Lucilia cuprina, before relatively low levels of field resistance were reported recently (Levot, 2012). This may be due to the ability relative fitness costs of the resistance mechanisms used to overcome insect growth regulators (McKenzie and Batterham, 1998). Laboratory selection experiments have shown that, while achieving high levels of resistance is possible, the phenotype is unstable without insecticidal selection (Levot, 2013).

1.3.2) Exaptation

One explanation for why insecticide resistance can evolve so rapidly is exaptation, an evolutionary principle by which traits that originally evolved in response to one selective pressure provide a fitness advantage to a separate selective pressure. Because insects are constantly exposed to environmental toxic chemicals, usually PSMs, xenobiotic resistance mechanisms predate the
introduction of synthetic pesticides (Gassmann et al., 2009). Furthermore, synthetics such as neonicotinoids and pyrethroids are structurally similar to PSMs such as nicotine and pyrethrum, which may provide a basis for cross resistance. For example, *Myzus persicae nicotianae*, a species that feeds on nicotine rich tobacco plants, was shown to have higher levels of neonicotinoid resistance compared to a closely related species that did not consume nicotine (Bass et al., 2013).

Such cases of exaptation are usually due to the upregulation of one or more genes from superfamilies involved in metabolizing or excreting xenobiotics (Section 1.4.2). Polyphagous insects, which generally encounter a wider range of PSMs in their diet, usually contain increased expression and copy numbers of these same superfamilies (Claudianos et al., 2006; Dermauw et al., 2013; Rane et al., 2016). Importantly, the relatively recent introduction of synthetic insecticides is far too short a time span to create proteins structurally tailored to act on synthetics. This means that the genes which commonly confer resistance in pest populations may not act most efficiently on synthetic insecticides, instead evolving in an environment where PSM efficiency was favoured. Taken together these data indicate that resistance can evolve rapidly through the upregulation of existing mechanisms rather than having to develop *de novo*.

**1.3.3) The Mono v. Polygenic Debate**

From a genetic perspective, there is a long standing debate over whether the evolution of insecticide resistance proceeds mono or polygenically (Roush and McKenzie, 1987). The basis for this difference is thought to be dependent on the dose used for the selection (McKenzie & Batterham 1994, Fig 1.2). With all quantitative traits, there is a distribution of phenotypes within a population controlled by a variety of loci many of which contribute only slightly. Insecticide resistance is no different and there is substantial variation between individuals within a population in terms of their resistance level. Applying a dose within the population’s range is therefore capable of selecting individuals towards the more resistant end of the distribution, along with the myriad of loci (polygenic) which determine that individuals place on the resistance spectrum. However, no locus of small effect would be important in determining survival if a dose above the LC$_{100}$ of a population was applied. In this case, it would be far more probable that a single allele (monogenic) of large effect would evolve owing to the difficulties of many alleles of small effect arising without an immediate fitness advantage.

Although monogenic and polygenic resistance are more commonly observed in field and laboratory settings respectively, these rules are far from clear cut. The lower doses that have been used in some laboratory selection regimes tend to select for polygenic resistant strains that can attain extremely high resistance levels. Following many generations of selection on DDT,
Figure 1.2- Response to Selection:

The theoretical genetic response to selection is shown in two scenarios. A) The selection dose is within the normal distribution of the population. The most resistant portion of the population is able to propagate, leading to a polygenic response. B) The selection dose is above the normal distribution of the population, which tends to select for an extremely rare variant in a monogenic response. Adapted from McKenzie and Batterham (1998).
the 91R strain of *Drosophila* evolved resistance high level resistance underpinned by a mixture of mechanisms such as transport, metabolism and cuticular thickness (Seong et al., 2016; Steele et al., 2015; Strycharz et al., 2013). In contrast, the high doses of insecticides used in field populations often produce monogenic resistance such as the *kdr* mutation involving an amino acid replacement at the L1014 position in the *para* sodium channel. The independent origin of this mutation in several populations suggests the relative ease with which it can evolve in the field (Du et al., 2016).

Even in cases where most of the observed resistance can be explained by variation in a single gene, variation in many other genes still influence the overall resistance level, even if they do so in a very subtle way. ‘Monogenic’ resistance evolves in a polygenic resistance background. Therefore a mutant resistant allele conferring several hundred-fold resistance compared to the susceptible allele may explain the majority of the overall difference between two strains, but there is still a distribution of resistance levels in populations where the resistance allele has gone to fixation. Such an intra-population distribution in resistance levels may be explained by polygenic variation. The reality is that single gene variation is easier to study, and there is substantially more economic interest in doing so. Hence, the understanding of the nature of the genes that underpin the polygenic response to insecticide, and the scale of the contribution that they make to so called ‘monogenic’ resistance, is rudimentary.

### 1.3.4) Evolution at the Gene Level

Regardless of how many genes are contributing to the insecticide resistance phenotype, the evolution of such a phenotype can only be achieved by a limited number of ways at the gene level. Changing the coding sequence of the open reading frame (ORF) of a gene can occur via the replacement of a single base with another or the removal of all or part of a gene. Transposable element insertions can also take place within the ORF, disrupting normal gene function (Bellen et al., 2004). These types of changes will alter the amino acid sequence of the protein and thus how it performs its function. Individual genes can also be modified by changing the magnitude, location, or timing of their expression pattern. Such changes, mostly the upregulation of genes encoding metabolic enzymes, are frequently observed in insecticide resistant populations (Hemingway et al., 1998). Among other possible mechanism is that transposable elements can cause these changes by inserting into the promoter region of a gene and regulating expression (Rostant et al., 2012; Zhang and Saier, 2009).
1.4) Common Resistance Mechanisms

1.4.1) Target Site Mediated Resistance

1.4.1.1) Modes of Action

Theoretically a myriad of different physiological resistance mechanisms could evolve in an insect population, but a limited number have been described genetically. One of the most common mechanisms observed is the modification of the insecticide target site. Although possible, secondary effects of pesticides have not been systematically investigated, and it is likely that pesticides exert the majority of their toxic effect through binding to a specific target protein(s). Upon binding, the insecticide either activates the receptor (an agonist) or prevents the receptor from activating (an antagonist) depending on its MoA. For example, the neonicotinoid chemical class of insecticides bind nicotinic acetylcholine receptors (nAChRs) at their endogenous ligand binding site and constitutively activate these receptors. Conversely, α Bungarotoxin blocks these same receptors by binding at the same site (Buckingham et al., 1997; Young et al., 2003). Allosteric modulators can activate or repress receptors by binding a site distinct from the endogenous ligand binding site. Spinosad, a part of the macrolactone chemical class, functions in this way and activates nAChRs by binding to an allosteric site (Salgado, 1998; Somers et al., 2015).

1.4.1.2) Types of Target site Modification

Resistance via target site modification had been observed in the field for every insecticide class for which the target protein is known. Substitutions of single amino acids caused by single nucleotide polymorphisms (SNPs) have been shown to evolve in both lab and field selected populations (Ffrench-Constant et al., 1993; Liu et al., 2005; Perry et al., 2008). Such point mutations can be located in the binding sites of the receptors and are thus a useful tool for defining the insecticide binding site. For agonist insecticides, mutations that remove entire target protein have been shown to confer very high levels of resistance, even if such changes are less likely to reach high frequencies in the field due to fitness costs (Perry et al., 2015; Somers et al., 2017). Although rarer, changing the expression of receptors has also been shown to confer resistance. Reduced expression of the *Nilaparvata lugens* nAChR α8 subunit, a hypothesised target for imidacloprid, was associated with increased imidacloprid tolerance (Zhang et al., 2015). Each of these genetic modifications can potentially incur fitness costs. One strategy to circumvent such costs is to establish permanent heterozygosity via the inheritance of a gene duplication containing one mutant copy and one wild type copy of the gene encoding the target. This was observed for *Rdl*, the target of the insecticide dieldrin, in *D. melanogaster* (Remnant et al., 2013).
1.4.1.3) Insecticide:Receptor Crosstalk

While insecticide mediated toxicity is usually thought to be due to a single chemical binding to a single receptor, this is not always the whole story. Many pesticides bind more than one receptor, and receptors can interact with each other in vivo. For example, preliminary studies on spinosad suggested that it was able to activate GABA gated chloride channels in addition to its nAChR primary target (Watson, 2001). More recently, the widely used neonicotinoid imidacloprid, which is thought to act primarily through nAChR subunits, was also found to antagonize the GABA gated chloride channel (Taylor-Wells et al., 2015). Receptors can also interact with each other in vivo in ways that can alter insecticide binding. For instance, imidacloprid and spinosad target distinct nAChR subunits in Drosophila, but there is evidence that changes in the expression of one subunit influences the expression of others. A genotype carrying a mutant copy of the imidacloprid target site, nAChR subunit Dα1, increases susceptibility to spinosad, possibly due to an upregulation of spinosad’s target site, nAChR subunit Dα6 (Ghazali and Perry unpublished). Insecticide/receptor interactions are thus likely more complicated than they first appear.

1.4.1.4) nAChRs

Particularly relevant to this thesis are the nAChRs, which are the targets for several widely used insecticides and are members of the larger pentameric ligand gated ion channel superfamily (Dent 2010). Functional nAChRs are made up of 5 subunits and bind ligands which function as neurotransmitters in the nervous system (Fig 1.3). These proteins are typically expressed on the postsynaptic neural membrane and allow an influx of cations into the postsynaptic neuron upon ligand binding, propagating the neural signal (Jones and Sattelle, 2010). There are 10 genes encoding distinct nAChR subunits in Drosophila, which come together in various combinations to form unique nAChR subtypes. In the extracellular domain of α (but not β) subunits, there is a cysteine loop motif that is involved in binding the natural ligand for nAChRs, acetylcholine. Thus, while all nAChR subunits form pentameric channels, α subunits are able form homomeric receptors, but β subunits can only be assembled into heteromeric receptors with α subunits. Critically, functional receptors behave differently depending on the specific nAChR subunit composition.

Neonicotinoid insecticides activate certain nAChR subtypes by binding acetylcholine binding site at the interface between two subunits, but cannot be degraded by the cellular machinery as acetylcholine is. This means that neonicotinoids trigger a constitutive activation of nAChRs that leads to the death of the organism. Conversely, spinosad binds the α6 nAChR subunit allosterically (Tricoire-Leignel and Thany, 2010). Although the precise binding site of spinosad has not been determined, high levels of amino acid conservation between the first and third transmembrane
domains of various insect α6 orthologues indicates that this region is likely involved (Perry et al., 2015).

**Figure 1.3- Nicotinic Acetylcholine Receptors:**
The structure of a nAchR is shown. A) A cartoon of a single subunit is shown with a large extracellular ligand binding domain at the N-terminus, 4 transmembrane (TM) domains, and an intracellular loop between TM3 and TM4. B) A fully functional nAchR is shown made up of 5 individual subunits arranged in a co-planar manner in the plasma membrane. Adapted from the thesis of Jason Somers (2015).
1.4.2) Drug Metabolizing Enzymes (DMEs)

1.4.2.1) The Structure and Function of DMEs

Apart from modifying the insecticide target site, another common resistance mechanism is through changing the structure of an insecticide so that the chemical has a lower affinity for its receptor or is more easily excreted (Fig 1.4A). This process is mediated by drug metabolizing enzymes (DMEs), which are traditionally classified as Phase I or II enzymes. Phase I enzymes, represented by carboxylesterases and cytochrome P450s (P450s), often make a xenobiotic more polar (Feyereisen, 1999; Hemingway and Karunaratne, 1998). Phase II enzymes, including glutathione S-transferases and uridine 5’-diphospho-glucuronosyltransferases, attach bulky side groups onto foreign compounds (Enayati et al., 2005; McGurk et al., 1998). Finally, Phase III metabolism consists of excretion of metabolized compounds out of the body by transporters. This terminology, however, implies a linear progression that has not been found to occur in insects.

A more general model classifies proteins into functionalization DMEs, which add or modify functional groups, conjugation DMEs, which conjugate xenobiotics with bulky side groups and transporter proteins (Testa and Krämer, 2008). The Testa and Krämer model does not imply any linear progression and thus more accurately reflects the reality of insecticide metabolism. Although DMEs are often associated with making insecticides less toxic or more easily excreted, the opposite is also possible whereby DMEs make insecticides more toxic or less easily excreted (Fig 1.4B). Likewise transporters are considered to move insecticides out of the body, but also import insecticides into a tissue depending on their sub-cellular expression pattern.

1.4.2.2) Cytochrome P450s Biochemistry

By far the most widely cited DMEs implicated in insecticide resistance are members of the P450 superfamily. Structurally, P450s are hemoproteins containing a conserved iron-protoporphyrin IX domain and a variable catalytic domain (Testa and Krämer, 2008). In terms of biological function, P450s can be classified as functionalization DMEs, adding an oxygen to their substrates at locations determined by the P450 and substrate in question. For example, D. melanogaster CYP6G1 converts imidacloprid into 5’ hydroxyl imidacloprid (IMI-5OH) by changing the methyl group at the 5’ position to a hydroxyl group (Hoi et al., 2014; Joussen et al., 2008). Conversely, the nitrogen atom of the pyridine ring of nicotinamide, a form of vitamin B3, can be oxidized by another P450 (Real et al. 2013). P450s catalyse these reactions in combination with their reductase cofactors by a well conserved redox reaction (Feyereisen, 1999).
**Figure 1.4- Drug Metabolizing Enzymes:**

A schematic for metabolism mediated insecticide resistance is shown. A) A toxic insecticide is converted into a less toxic or more easily excretable metabolite. B) A toxic insecticide is converted into a more toxic or less easily excretable metabolite. Not shown is a scenario where insecticide and metabolite are of equal toxicity and ability to be excreted.
1.4.2.3) **Cytochrome P450 Evolution and Diversity**

The cytochrome P450 superfamily is extremely diverse and has undergone expansions and contractions in various insect lineages. Broadly, P450s can be broken into 4 clans, CYP2, CYP3, CYP4 and the mitochondrial P450 clan (Feyereisen, 2006). Within each clan is a family designated by a number and then a subfamily designated by letter. P450s nomenclature ignores clan and focuses exclusively on family and subfamily (for example Cyp6g1; family 6, subfamily g). The number of P450s per insect species is highly variable ranging from only 46 in *Apis mellifera* (honeybee) to 180 in *Culex pipens*, but each family is usually represented by at least one gene in each species. Expansion of the number of P450s in a lineage is often due to blooms of one particular family driven by successive rounds of recent gene duplication events and genetic hitchhiking (Fig1.5; Feyereisen 2011). Indeed, copy number for a given P450 gene can vary greatly within a specie or among closely related species (e.g. species of the genus *Drosophila*; Good et al. 2014). This makes the annotation of CYPomes difficult. Beyond phylogenetic classifications, P450 genes can also grouped into either “stable” or “unstable” families based upon their conservation among closely related species. Compared to most other gene families, P450s are highly divergent, and only about 1/3 of *D. melanogaster* P450s have unambiguous 1:1 orthologues in all other *Drosophila* species (Chung et al., 2009).

1.4.2.4) **Cytochrome P450 Biology**

Although there are many exceptions to this rule, the more stable families are usually involved in vital developmental functions while the unstable lineages are responsible for more adaptive functions such as drug metabolism (Thomas, 2007). For example, stable P450s such as the Halloween genes are involved in the production of the hormone ecdysone, which regulates larval development (Rewitz et al., 2006). In contrast, members of unstable families, like the Cyp6 family, have been implicated in insecticide resistance and drug metabolism in several species (Bass et al., 2013; Dabor, 2002; Zhu et al., 2010). These rules are not axioms. The recent sequencing of a *D. melanogaster* ring gland transcriptome revealed a role for members of the Cyp6 family in development, and RNAi knockdown of several Cyp6 family members produced lethal phenotypes (Christesen et al., 2017; Chung et al., 2009). Furthermore, members of the mitochondrial P450 clan have been implicated in xenobiotic metabolism in houseflies and *D. melanogaster* (Bogwitz et al., 2005; Guzov et al., 1998).
Figure 1.5- Cypomes among arthropods:

The total number of P450s in the genome’s of a variety of arthropod species is shown. Numbers range from 36 P450s in the body louse Pediculus humanus to 180 in the mosquito Culex pipens. Adapted from Feyereisen (2011).
1.4.2.5) The Molecular Genetics of Cytochrome P450 Mediated Resistance

Genetically there are various mechanisms by which P450s can be modified to enhance their ability to metabolize insecticides. The less common way is for changes to take place in the coding sequence of a gene, thereby structurally changing the P450 protein. A point mutation in *D. melanogaster* Cyp6a2 was shown to increase the ability of the CYP6A2 enzyme to metabolize DDT (Amichot et al., 2004). A chimeric P450 in *Helicoverpa armigera* was able to metabolize to fenvalerate, while the two parental genes that gave rise to the chimera could not (Joussen et al., 2012). A far more common mechanism of cytochrome P450 mediated insecticide resistance is to change the expression of a protein already capable of metabolizing insecticides so that it is expressed at high levels in active metabolic tissues. This occurs most frequently by increasing the copy number or modifying the promoter region of a particular P450 (Bass et al., 2013; Chung et al., 2007; Schmidt et al., 2010). Transcription factors such as DHR96 or Cap’n’Collar can also help to regulate inducible metabolic responses by detecting the presence of xenobiotics and subsequently upregulating a large number of genes, including P450s (Chung et al., 2011; King-Jones et al., 2006; Misra et al., 2013, 2011).

1.4.3) Xenobiotic Transport

1.4.3.1) Overview

By virtue of the fact that the targets of insecticides are not expressed ubiquitously, a significant obstacle to insecticidal efficacy is how the pesticide can reach its target tissue within the organism. Transporter proteins can regulate the ability of an insecticide to access a tissue by actively effluxing the chemical from the tissue. These transporters can be classified broadly as primary (directly use ATP) or secondary (make use of proton gradients). Although a myriad of different transporter proteins acts on an even larger number of substrates, only five superfamilies have been so far implicated in xenobiotic transport. Of these the ATP-binding cassette transporters (ABCs) are primary, and the Major Facilitator Superfamily (MFS), small multidrug resistance (SMR), Multidrug and Toxic Compound Extrusion (MATE) and resistance-nodulation-cell division (RND) are all secondary transporters (Putman et al., 2000). The ABC transporters are more often implicated in drug transport in animals, while secondary transporters are sometimes cited in bacteria and fungi. Hence ABC transporters are far more commonly implicated in insecticide resistance.

1.4.3.2) The Structure Function of ABC Transporters

The ABC superfamily is present in all forms of life from “microorganisms to man” (Higgins 1992), and is characterized by a conserved nucleotide binding domain (NBD). The high
conservation of the NBD is contrasted by the more highly variable transmembrane domains (TMDs), which are thought to be involved in substrate binding (Wilkens 2015; Fig 1.6A). In eukaryotic cells, ABC transporters are only able to efflux substrates into the extracellular matrix, but this can result in the importation of substrates into or out of tissues depending on whether the transporter localizes to the inward facing or outward facing membrane. Hence, these transporters can facilitate transport between tissues. Another feature of the ABC transporter superfamily is the ability of some members to dimerize. Half transporters encompass one NBD and one TMD and must form homo or heterodimers to become functional. Full transporters, on the other hand, include 2 copies of each domain in a single polypeptide making them relatively self-contained. It is generally accepted that upon ATP-hydrolysis the protein undergoes a conformational change, which effluxes the substrate out of the cell (Wilkens 2015; Fig 1.6B). Exactly how this is accomplished is still poorly understood, although several biochemical models have been proposed (Higgins & Linton 2004; Senior et al. 1995; Siarheyeva et al. 2010).

1.4.3.3) ABC Transporter Phylogeny

ABC transporters can be divided into 8 families (Labelled A-H) based on sequence similarity in the highly conserved NBD. Families A and C are composed of full transporters, while families D, E, F, G and H are composed of only half transporters. The B family is composed of both half and full transporters (Dean and Annilo, 2005). Furthermore, families E and F are, somewhat confusingly, not actually transporters at all. Although they share a conserved nucleotide binding domain, they lack the TM domains necessary for transmembrane transport (Xiong et al., 2015). So far, only members of the families B, C, and G have been implicated in xenobiotic transport. For example, the full transporter multidrug resistant protein (MRPs) and the half transporter breast cancer resistance protein (BCRP) have each been implicated in drug resistance in cancer cell lines and come from families C and G respectively (Cole et al., 1992; Doyle et al., 1998).

1.4.3.4) P-glycoprotein

The ABCB family is the most thoroughly researched ABC transporter family involved in drug transport. In particular, ABCB1 is one of the best studied genes in the history of genetics (Gottesman and Ling, 2006). This gene was originally discovered in a multi-drug resistant cancer cell line via labelling of cell surface carbohydrates, hence its alternative nom du guerres P-glycoprotein (P-gp) and MDR1 (Juliano and Ling, 1976; Kartner et al., 1983). Eventually it was found that increased expression of this gene was able to confer resistance to a wide range of drugs (Ueda et al., 1987). Further work revealed that the endogenous expression of P-gp is enhanced in the blood brain barrier (BBB) and the kidney suggesting a role for P-gp in keeping toxins out of
the brain and excreting them from the body (Theibaut et al., 1987). One of the primary reasons why P-gp is so commonly implicated in drug resistance is its broad substrate specificity. The crystal structure of Mus musculus P-gp revealed a binding site capable of selectivity based on hydrophobic and aromatic interactions, which may account for the polyspecificity of P-gp (Aller et al. 2009; Li et al. 2014; Fig 1.7).

**Figure 1.6: ABC Transporters:**

A) A cartoon depicting the structure of ABC transporters are shown. Full ABC transporters have 12 transmembrane (TM) domains and two nucleotide binding domains (NBD) arranged in a symmetrical pattern. B) The mechanism of ABC mediated efflux is shown. Upon substrate binding ATP is hydrolysed to ADP which provides the energy for a conformational shift in the protein, which effluxes the substrate. Adapted from Dermauw et. al (2014).
1.4.3.5) P-glycoprotein in Insects

Substantially less is known about the role of the role of P-gps in insects. Most insect species have between 1-4 orthologues of P-gp although the roles of these genes are poorly understood (Merzendorfer 2014). The *D. melanogaster* orthologue *Mdr49* has been shown to transport germ cell attractant during development (Ricardo and Lehmann, 2009). Likewise, another P-gp *Drosophila* orthologue, the gene *Mdr65*, has been implicated in combating polyglutamine expansive neurodegenerative diseases (Yadav and Tapadia, 2013). More in keeping with its history, insect P-gp has also been both associated with drug transport in many studies (Dermauw and Van Leeuwen, 2014; Merzendorfer, 2014). Critically though, many such studies rely on indirect evidence such as synergism with P-gp inhibitors, increased expression of P-gp in resistant strains, or upregulation of P-gp in response to insecticide exposure. A smaller number of studies have done functional studies on P-gp and tested their role in xenobiotic transport (Aurade et al., 2010; Mayer et al., 2009). More recently, overexpression of certain orthologues of *Mdr49* in *D. melanogaster* was shown to confer resistance to DDT (Seong et al., 2016). However, systematic testing of insect P-gp orthologues for their substrate specificity has not been performed.

1.5) Insecticide Biology and Insecticide Resistance

1.5.1) Introduction

Thus far various factors centred around the concept of insecticide resistance have been discussed. At this point a distinction must be made between insecticide resistance and insecticide biology. Insecticide resistance is the concept that individual insects or populations can better survive exposure to concentrations of insecticides that would normally be lethal. Insecticide biology is a term that describes all of the different biological mechanisms by which an insect can interact with an insecticide. Processes that are involved in insecticide resistance are by definition involved in insecticide biology. For example, a receptor can be involved in insecticide resistance if mutated and is most certainly involved in interacting with the insecticide (insecticide biology). However, by focusing on biology rather than resistance, the discussion is oriented away from the narrow subset of resistance mechanisms that are frequently described, and towards the myriad of physiological mechanisms by which resistance could potentially evolve. Failing to do so leaves us in the position of the superhero Captain Hindsight from the popular TV series *South Park*, diagnosing how resistance has previously evolved, while not being able to predict how it may evolve in the future. Probing insecticide biology is best accomplished by using many different insecticide related phenotypes that can highlight the different aspects of the insecticides biology. Therefore, an
understanding of the various types of insecticide phenotypes is crucial to understand insecticide biology and predict resistance mechanisms.

Figure 1.7- P-glycoprotein:
A ribbon model showing P-glycoprotein, the most common ABC transporter implicated in drug transport from A) front and B) back. The two halves of the protein are shown in blue and yellow and TMDs and NBDs are labelled. Adapted from Aller et. al (2009).
1.5.2) Mortality assays

1.5.2.1) Dosing Methods and Duration

One of the most common ways to assess insecticide response is by measuring mortality following an insecticide exposure. There are many types of mortality assay which may differ in their capacity to detect and quantify differences in resistance level. Assays vary in the exposure method, duration and dose and in the life stage exposed. Exposure methods can be divided into contact assays whereby the insect is placed in an insecticide treated environment, direct application whereby drops of insecticide solution are applied to each insect individually and feeding assays whereby insecticides are mixed into the diet of the insect. The length of the exposure can be sometimes be varied so that acute or chronic effects can be observed. A dose must also be chosen so that some, but not all, of the individuals exposed will survive the treatment.

Each of these variables can elicit very different conclusions about the genetic basis of resistance even considering identical individuals and insecticides. For example, *Blattella germanica* showed resistance to deltamethrin with direct application but not with a contact assay (Scott et al., 1986). Similar discrepancies were observed between two different contact assays (slide dip and leaf dip) when dicofol was tested against the spider mite (Dennehy et al., 1983). Changing the dose has also been shown to alter conclusions from experiments with identical exposure methods. Battlay et al. (2016) found that the candidate genes obtained from a genome wide association study varied substantially, based on which of 5 doses was chosen. Each variation of an insecticide mortality assay highlights different aspects of the insecticides biology of insecticides and thus will have a different set of genes underpinning the overall phenotype.

1.5.2.2) Life stage and Scoring Mortality

Of similar consequence to exposure conditions is the life stage that is tested and how mortality is scored. Holometabolous insects complete metamorphosis having larval and adult stages while hemimetabolous insects progressing through a series of nymph stages that resemble the final adult stage (Labandeira and Phillips, 1996). Differences in resistance levels between larval and adult stages in species such as *Lucilia cuprina* and *Bemisia tabaci* attest to the wide range of physiological differences between life stages that can determine resistance levels (Arnold and Whitten, 1975; Nauen et al., 2008). Such temporal effects are also not strictly limited to those between distinct life stages. The age of Anopheles *stephensi* was negatively correlated with their level of resistance and metabolism of malathion (Rowland and Hemingway, 1987). How mortality is scored represents another variable that can influence results. Determining if and when an insect is dead is complicated by the reality that many insecticides cause “knock down”, a phase where the
insect is stunned before actually dying. Further complicating this issue, in acute feeding or contact assays immobilized insects may be exposed to far less insecticide than mobile ones, which may mask differences in tolerance. There are several examples where mortality resistance does not match up with related response phenotypes thought to be predictors of resistance. Scott et. al (1986) found that the time needed for an insecticide treatment to immobilize cockroaches was not a great predictor of overall mortality. More recently, various species of mosquitoes were used to show that knockdown time was not highly correlated with 24-hour mortality (Owusu et al., 2015). These studies highlight that other phenotypes, which can be measured that may reflect other elements of insecticide biology masked by mortality studies.

1.5.3) Sub-lethal Traits

1.5.3.1) Why Look at Sub-lethal Traits?

The sub-lethal response of insects to insecticides is of particular interest in the context of this thesis. Here, sub-lethal response is defined as any phenotype that can be measured under conditions of insecticide exposure while the insect is still living. This definition merits further expansion. Doses which will kill an insect after some time may be used in sub-lethal exposures if the trait is measured before death is achieved. Likewise lengths of exposures usually deemed sufficient to cause lethality can be considered sub-lethal if the dose is very low. The difference between lethal and sub-lethal exposures may be achieved by changing any one of the experimental components listed previously (Section 1.5.2) as long as the insect is living while the measurements are being taken.

Sub lethal responses are relevant to insecticide resistance for many reasons, two of which are relevant here. Firstly, sub-lethal response can influence fitness and thus can be selected for in field settings (Guedes et al., 2016). Insecticides degrade over time, meaning that the dose to which an insect may be exposed can be far below the high concentrations initially deployed. Further, some non-pest insects that are not the intended targets of insecticides are, nonetheless, exposed. This is exemplified by the potential contribution of neonicotinoid exposures to colony collapse disorder in the honeybee, A. mellifera (Fairbrother et al., 2014). Secondly, focusing on sub-lethal effects is biologically relevant because the genetic basis for sub-lethal phenotypes may be different from those of mortality phenotypes, highlighting different aspects of the insecticide’s biology (Section 1.5.1). Understanding genes and mechanisms underpinning sub-lethal response may be critical in understanding insecticide biology and, thus, how resistance may evolve.
1.5.3.2) Behavioural Assays

Perhaps the most common way to assess sub-lethal responses is through behavioural assays during or after insecticide exposures. This is especially pertinent due to the negative impact on behaviour for non-pest species such as honeybees due to sub-lethal insecticide exposures (Desneux et al., 2007). For example, the foraging behaviour of *A. mellifera* was found to be altered after a sub-lethal exposure to neonicotinoids (Schneider et al., 2012). Similar changes in feeding behaviour were observed when *Sitobion avenae* was exposed to the same insecticides (Miao et al., 2014). Learning can also be affected by insecticide exposures. Insects such as *A. mellifera* can be trained to extend their proboscis in response to olfactory cues, but this learning is disrupted when individuals are dosed with imidaclorpid concentrations far below their LD$_{50}$ (Decourtye et al., 2005). Other systems can use mobility as a general proxy for sub-lethal insecticidal effect. For example, Risse et al. (2015) developed software that is able to track fluorescent protein expressing larvae in a variety of environments. The movement of adults too can be assessed using a slightly different apparatus. The *Drosophila* Activity Monitor (DAM) or Trikinetic system (Waltham, MA, USA) was recently used to assess the DDT response of *D. melanogaster* sampled from Australian populations (Fournier-Level et al., 2016). Advances in dealing with large datasets and the ability to quantify complicated behavioural data will likely increase the prevalence of such behavioural studies.

1.5.3.3) Sub-lethal Biochemical Changes

Moving into the molecular realm, there are several phenotypes that can be measured biochemically. Adding a substrate of a particular DME superfamily to homogenized tissue can be used to compare the activity of that superfamily in resistant and susceptible backgrounds, thus implicating it in the observed resistance phenotype (Sawicki et al., 1980). Similar exposure methods can also be used in binding studies with insecticide targets (Liu et al., 1993). DMEs or transporter genes can also be expressed heterologously for characterization in controlled genetic backgrounds (Papke and Smith-Maxwell, 2009; Yun et al., 2006). Finally, transcriptional and translational changes in insects undergoing insecticide exposure often points to the genes involved in their biology. The activity of transcription factors such as DHR96 and Cap ‘n’ Collar are both induced by xenobiotics and regulate a variety of DMEs and transporters (King-Jones et al. 2006, Misra et al. 2011). Measuring the transcription of genes downstream of these pathways can be used as a way to assess the impact of the insecticide on the insect. Receptors can also be regulated as is the case with the downregulation of $\alpha_6$ in response to sub-lethal spinosad exposures (Nguyen and Perry unpublished).
1.6) Forward Genetics: From Phenotype to Genotype

1.6.1) Overview and history

1.6.1.1) What is Forward Genetics

One of the central goals of biology in general, and genetics in particular, is to explain the link between genotype and phenotype. One way to accomplish this is with forward genetics whereby a researcher begins with an observed phenotype and attempts to describe its genetic basis by associating it with genetic variation. Because forward genetics does not necessarily involve \textit{a priori} assumptions regarding which genetic variants underlie a trait, an unbiased genetic description of the trait can be achieved. Genetic variation can take one of two forms. Quantitative trait loci (QTLs) are locations on the chromosome, which contribute to a phenotype in a quantitative way. Changing a QTL through a SNP, insertion or deletion can influence the overall trait. Measuring transcription level is another form of genetic variation. Here, the expression pattern and quantity of a given transcript can be measured and associated with a phenotype. Most traits are to some extent polygenic meaning that the contribution of individual QTLs or transcripts may be small enough to be missed unless very high resolution techniques are used. Despite these drawbacks forward genetics remains a powerful technique for explaining phenotype with genotype.

1.6.1.2) Sources of Genetic Variation

In order to undertake a forward genetic study, one must first obtain genetic and phenotypic variation that can be correlated. Field populations are one of the most common and relevant sources of variation, relying on the genetic diversity that has arisen as a result of, and been sieved by, evolutionary processes. However, these processes are extremely slow, and populations are subject to a myriad of different evolutionary pressures, which may confound results. Selection experiments in laboratories can partially overcome this problem by keeping all other factors constant and precisely defining the selection pressure (Swallow and Garland, 2005). Still, field populations and selection experiments are both likely to purge any variant which, although biologically interesting, is too deleterious to overall fitness to persist at appreciable frequencies. Mutagenesis agents such as radiation, chemicals, or transposable elements can speed up the process and generate lots of genetic diversity in a single generation. Such methods have been frequently used to generate genetic variation in insecticide resistance studies (Perry et al., 2008, 2007). Using these techniques, essential processes can be studied and a more complete picture of the biology of a trait can be observed.
1.6.1.3) Existing Genomic Variation

One way to implicate a particular QTL in a trait is by observing variation at a given locus in a field population that has undergone recent selection. Insecticide resistance is a paradigm for the study of such selective pressures because the selection applied is recent and the fitness cost is high. Therefore, evidence of recent selection at a particular locus can suggest potential resistance locus. Selective sweeps are one such line of evidence. When a QTL has a large enough impact on an evolutionarily important trait it will tend to spread through the population rapidly and outpace the ability of recombination to dissociate it with its surrounding variants. Decreased genetic variation around a locus can thus be interpreted as evidence for recent adaptation (Nielsen, 2005; Schlenke and Begun, 2004). This principle can also be applied to coding sequences of DNA by looking at the ratio of synonymous to non-synonymous changes (dN/dS values). However, this method is thought to be more effective when observing variation between populations as opposed to within them (Kryazhimskiy and Plotkin, 2008).

1.6.2) Association Studies

1.6.2.1) Mapping Crosses

The earliest form of forward genetics was mapping crosses which made use of the principles of independent assortment and genetic linkage. These two principles reveal that two genes are inherited independently unless they are on the same chromosome, whereby the physical distance between the two loci is inversely proportional to their linkage (Mendel, 1865; Sturtevant, 1913). Reversing this logic, the frequency by which two variants are inherited together can predict how far apart they are on the chromosome. Mapping crosses use this principle by crossing uncharacterised mutants to genotypes with known, penetrant markers that can implicate specific genes in a trait. To improve resolution, molecular markers are typically used instead of the far more limited array of visible markers (Darvasi, 1998). These molecular markers can take the form of restriction fragment length polymorphism (RFLPs), variable number tandem repeats or SNPs among others (Jones et al., 1997). For instance, mapping crosses using such markers were used to map lufenuron resistance to the Cyp12a4/Cyp12a5 locus (Bogwitz et al., 2005). Although these strategies have proved useful in dissecting QTLs, mapping studies are labour intensive and still do not provide enough resolution for the efficient implication of specific genes and QTLs in a trait.

1.6.2.2) The Genome Wide Association Study (GWAS)

The recent sequencing boom has helped to resolve this issue by using a much denser array of molecular markers taken from whole genome sequences. Here, the only limit on the number of markers available is the amount of variation annotated from assembled genome sequences, and
variants can be called systematically using high throughput methods. The first such genome wide association study (GWAS) was published in 2005, suggesting that a single QTL was highly associated with age related macular degeneration (Klein et al., 2005). Although this kind of study has become commonplace, such strategies are not without their drawbacks. Scoring such large numbers of markers is likely to throw up spurious associations through multiple testing error. Statistical corrections such as the Bonferroni threshold and false discovery rate have been used to try to correct these biases, but false positives remain a problem (Bland and Altman, 1995; Storey and Tibshirani, 2003). Additionally, false negatives may also confound results because many genetic variants that exert the strongest influences on a trait tend to appear infrequently and are thus difficult to associate with a trait in a statistically meaningful way. The most successful and reliable GWAS to date have identified common variants which have large effect sizes on a given trait (Korte and Farlow, 2013). Furthermore, the environmental heterogeneity observed when looking at phenotypes outside a controlled laboratory environment is another source of error. Despite these limitations GWASs remain a powerful tool for connecting genes to phenotype.

1.6.2.3) Transcriptomic Association Studies

Another way to describe the basis of a trait is to use transcriptomic studies, that allow mRNA levels for genes to be compared among genetic backgrounds or environmental conditions. Individual transcripts can be measured using quantitative “real-time” PCR (Higuchi et al., 1992), but this can be systematized to encompass a broader perspective by looking at the whole transcriptome. The first such studies used microarrays to measure hybridization intensity between cDNA libraries and pre-treated cDNA chips (Pollack et al., 1999). The decreasing cost of DNA sequencing has made RNA-seq, whereby the total cDNA of a sample is sequenced, a popular technique that uses read depth to estimate transcript abundance (Nagalakshmi et al., 2008; Wang et al., 2009). The effect of multiple testing in transcriptomic studies is far less because the number of distinct genomic variants is far greater than the number of distinct transcripts. Furthermore, measuring transcript level allows for the consideration of tissue specific expression patterns not present when considering genomic variants. As a result, comparing transcriptomes between insecticide resistant and susceptible genotypes (Bariami et al., 2012) and between insecticide exposed and unexposed insects (Bretschneider et al., 2016) has become commonplace.
1.7) Reverse Genetics: From Genotype to Phenotype

1.7.1) The Basics of Reverse Genetics

1.7.1.1) What is Genetic Manipulation

The logical counterpart to forward genetics is reverse genetics whereby a defined genetic change is made in an organism, which is then assessed for phenotypic differences. There are many kinds of manipulations that can be made, paralleling those which occur spontaneously in nature (Section 1.3.4). DNA can be inserted, deleted or replaced down to the level of single nucleotides at precise locations in the genome. Likewise, the magnitude or pattern of a transcript's expression can be modified. Two primary advantages of reverse genetics are that it can 1) assess the impact of alleles that would be unlikely to occur in field populations 2) assess the impact of genetic changes observed in the field in controlled genetic backgrounds. In the context of insecticide resistance, it means variation that may not exist yet in the field can be tested for an ability to confer resistance. Reverse genetics can thus more easily explain how a wild type genotype copes with insecticides, rather than limiting the scope to the explanation of hyper-resistant genotypes.

1.7.1.2) Brief History of Genetic Manipulation

Somewhat counter-intuitively, the first genetic transformation predates the discovery of DNA as the mechanism of inheritance. Griffith (1928) showed that heat killed pathogenic bacteria were capable of “transforming” non-pathogenic bacteria into virulent ones using mice as a host. It was not until years later that DNA was established as the basis of this transformation (Avery et al., 1944). These early experiments, however, provided no way to control genetic transformation in any meaningful way. It was not until the discovery of restriction endonucleases and plasmids that the first recombinant transformation was accomplished in Escherichia coli (Cohen et al., 1973). Multicellular organisms proved to be more challenging to transform due to their lack of plasmids and the need to target the germline cells. A genome altering plant pathogen Agrobacterium tumefaciens, helped solve this problem in plants and relatively high efficiency genome engineering became prevalent very quickly (Chilton et al., 1977; Zambryski et al., 1983). Animal models proved more difficult but, Gordon et. al (1980) successfully integrated DNA into the mouse genome using microinjection into embryos.

1.7.1.3) Transformation in Insects and Microinjection

This microinjection technique has proven to be virtually the only means of transformations in the Insects. D. melanogaster was the first species to be transformed using transposable elements to stably integrate into the Drosophila genome (Spradling and Rubin, 1982). Transposable elements need other factors to help them work such as integrases, which can be expressed endogenously
(direct injection) or injected with the exogenous DNA (co-injection). Transformation can also proceed by crossing two transgenic lines which each express a component of the integration/recombination system eliminating the need for injection (although the transgenic lines used in this cross must be initially generated by injection; Venken et al. 2016). However, transformation in non-Drosophila species have been so far limited by 2 barriers: low survival of injectees and low integration efficiency of vectors. P element mediated integration is Drosophila specific and other transposable element vectors usually have integration efficiencies of under 5% (Atkinson et al., 2001). Sensitivity to desiccation and hydrocarbon oil toxicity contributes to low survivorship of injectees and also reduces the practicality of using other insect species. Furthermore, non-Drosophila insect species are often difficult to inbreed, increasing genetic variability in a system that demands a certain degree of homogeneity. Due to these factors transformation is very difficult apart from D. melanogaster.

1.7.2) Genetic Biotechnology in D. melanogaster

1.7.2.1) Integrases and Recombinases

One of the principle ways to genetically manipulate D. melanogaster is to introduce exogenous DNA into the genome. While transposable elements are often used to insert DNA, the randomness by which these elements integrate into the genome limits their utility in targeting specific genes or genomic loci. Microbial recombinases such as Cre and FLP help solve this problem by catalysing recombination between specific DNA sequences such as LoxP and FRT sites respectively. The LoxP or FRT sites in the genome can be induced to recombine with LoxP or FRT sites in the exogenous DNA which catalyses the precise integration of the exogenous DNA (Venken et al., 2016). Integrases such as φ31 perform a similar reaction which results in the integrating of exogenous DNA adjacent to corresponding AttB (Attachment B) DNA sequences in the genome (Fig 1.8; Venken et al. 2016). As there are no AttB sites in wild type Drosophila, inserting these AttB sites into the genome creates a landing for future integrations (Bischof et al., 2007).

1.7.2.2) Nucleases

Still the integration of DNA into the genome is spatially limited to loci with either pre-existing recombinase or integrase recognition sites or those generated by transposable element mediated insertions. In order to precisely define the location of genomic insertions, nucleases must be used to create double stranded breaks. The most basic of these nucleases are the so-called homing endonucleases or meganucleases. They are conceptually similar to restriction enzymes in that they recognize specific (18-30 nucleotide) recognition sites and create double stranded breaks in the DNA (Taylor and Stoddard, 2012). These enzymes are of limited utility as they only
recognize a very limited number of sites across the *Drosophila* genome. Far more useful are the programmable nucleases such as zinc finger nucleases (ZNF), Transcription Activator-Like Effector Nucleases, (TALENs) and Cas9. ZNF and TALENS use sequential amino acid motifs to target specific nucleotide sequences and can thus be “programmed” to target a large number of genomic locations (Gaj et al., 2013). CRISPR-Cas9 goes one step further and allows the directing of the nuclease Cas9 to genomic loci dictated by customized RNA sequences limited only by the presence of a di-guanine (GG) sequence (Jinek et al. 2012; Fig 1.9).

**Figure 1.8- AttB Integration:**
A schematic showing AttB integration is shown. Corresponding AttB sites in the genome and the recombinant plasmid guide the integration of the recombinant plasmid into the genome. Adapted from Venken (2016).
Figure 1.9- CRISPR-Cas9:
The CRISPR-Cas9 system relies on the endonuclease Cas9 to be guided to the appropriate genomic location by a complementary guide RNA. Once there the Cas9 protein creates double stranded breaks in the DNA adjacent to a protospacer adjacent motif (PAM) sequence. Adapted from Jinek et. al (2012)
1.7.3) Applications of the genetic toolkit

1.7.3.1) Manipulating Gene Expression

One of the major applications of genetic engineering technology in *Drosophila* has been to regulate the expression of individual genes. Most expression systems have made use of the yeast transcription factor GAL4, which can bind an upstream activation sequence (UAS) and drive the expression of the gene directly downstream (Fig 1.10; Brand & Perrimon 1993). By altering the promoter of the GAL4 construct, and thereby its spatiotemporal expression pattern, one can alter the magnitude and location of gene expression. This system has been used frequently to implicate individual P450s in insecticide resistance by driving their expression in the metabolically relevant tissues of the midgut, malpighian tubules or fat body (Chung et al., 2007; Daborn et al., 2007). A similar system can be used to reduce the expression of a gene by driving an RNAi hairpin complementary to a given gene under the control of a GAL4 promoter. Stocks centres like the Vienna Drosophila Resource Center and the Transgenic RNAi Project contain UAS-RNAi hairpin lines for the majority of known *Drosophila* genes (Dietzl et al., 2007; Perkins et al., 2015) although inconsistencies have been reported with the former (Green et al., 2014). These libraries allow for the quick ordering and knockdown of genes of interest, which has previously helped implicate several nAChR subunits in imidacloprid resistance (Mitchell unpublished). Although the level of overexpression is dependent on where the UAS construct is located in the genome, this problem has been partially alleviated by the generation of precise integration sites using the φ31 integrase/attB system (Bischof et al., 2007).

Tagging genes, or more specifically the proteins that they encode, is another way to characterize and manipulate the expression of a gene. This involves adding in a short cDNA sequence usually at the N or C terminal of a gene that encode for a known protein sequence which displays some sort of marker and does not interfere with the proteins endogenous function. These can take the form of small (8 amino acid) FLAG tags, which aid in protein purification and antibody staining, or the more easily visualized fluorescent proteins such as GFP (Tsien, 1998). This technique has also been used to describe insecticide biology. Tagging of the Dα6 nAChR subunit with YFP showed that the translation of this protein is reduced in the presence of its insecticidal ligand spinosad (Nguyen and Perry unpublished).
Figure 1.10- GAL4-UAS:
The GAL4-UAS system relies on a cross between a GAL4 driver line, expressing GAL4 in a subset of tissues, and a UAS line, containing a promoter induced by GAL4 upstream of a gene of interest. The resulting cross will drive the gene in the tissues dictated by the GAL4 line. Adapted from Brand and Perrimon (1993).
1.7.3.2) **Disrupting Gene Function**

Yet another way to interrogate the functions of a gene is to create a null allele or a non-functional copy. Traditionally this was accomplished by inserting P-elements in the coding regions of a gene or creating frameshift mutations in untargeted mutagenesis screens (Bellen et al., 2004). More recently these methods have begun to be replaced by procedures that specifically target only the genes in question. Ends out targeting allows for the replacement of targeted DNA with a visible marker but is labour intensive and relatively low throughput (Gong and Golic, 2003). This technique was used to remove the Dα1 nAChR subunit conferring high levels of neonicotinoid resistance (Somers et al., 2017). More recently, “Trojan exons” have be inserted into intronic locations that create an extra exon within a stop codon early in the gene which creates a truncation (Diao et al., 2015). Finally, CRISPR was recently demonstrated in *D. melanogaster* (Gratz et al., 2013), allowing for the clean deletion of part or all of the coding sequence of a gene.

1.7.3.3) **Recent Advances in CRISPR**

On this last point (CRISPR) the surface has so far only been scratched of what is possible with genetic manipulation. It has recently been shown that single base pair changes, which produce single amino acid changes in proteins, can be made with relatively high efficiency to recapitulate alleles originally described in mutagenesis (Sternberg et al., 2016). This principle was applied to insecticide resistance by regenerating a mutation in the Dα6 nAChR subunit originally discovered in a mutagenesis screen for spinosad resistance (Somers et al., 2015). CRISPR can also be used to direct previously randomly integrating DNA cassettes. Zhang et. al (2014) recently developed a CRISPR guided Trojan exon approach to knock out genes without disrupting their surrounding genomic region. All these strategies are helped by the increasing efficiency with which Cas9 nuclease activity can be achieved. Transgenic expression of both Cas9 and sgRNAs was recently shown to be highly effective in generating both germline and somatic knockouts of genes making it a versatile tool (Port et al., 2014; Port and Bullock, 2016). These methods can likely be applied to a variety of biological settings to answer scientific questions that have so far sat beyond the scope of what was testable.

1.8) **Outline of This Work**

Using the model insect *D. melanogaster*, this study seeks to understand the genetics and physiology of insecticide toxicology. First, a bioassay that is able to detect motility changes in response to insecticides and compare them across different genetic backgrounds was developed
(Chapter II). This assay was useful in understanding how sub-lethal response can be different from insecticide induced mortality. The technique was then used in a forward genetic screen to implicate individual QTLs and transcripts in imidacloroprid response, one of which was tested via gene knockout (Chapter III). Finally, a reverse genetic approach was employed to manipulate an ABC transporter and characterize its role in resistance to multiple insecticides. By doing this, a better understanding of the genetic basis of insecticide biology and resistance was achieved.
Chapter II: The Wiggle Index: An Open Source Bioassay to Assess Sub Lethal Insecticide Response in *Drosophila Melanogaster*
RESEARCH ARTICLE

The Wiggle Index: An Open Source Bioassay to Assess Sub-Lethal Insecticide Response in *Drosophila melanogaster*

Shane Denecke¹, Cameron J. Nowell², Alexandre Fournier-Level¹, Trent Perry¹, Phil Batterham¹* *

1 School of BioSciences, Bio21 Molecular Science and Biotechnology Institute, University of Melbourne, Parkville, Victoria 3010, Australia, 2 Faculty of Pharmacy and Pharmaceutical Sciences, Monash University, Parkville, Victoria 3052, Australia

* p.batterham@unimelb.edu.au

Abstract

Toxicological assays measuring mortality are routinely used to describe insecticide response, but sub-lethal exposures to insecticides can select for resistance and yield additional biological information describing the ways in which an insecticide impacts the insect. Here we present the Wiggle Index (WI), a high-throughput method to quantify insecticide response by measuring the reduction in motility during sub-lethal exposures in larvae of the vinegar fly *Drosophila melanogaster*. A susceptible wild type strain was exposed to the insecticides chlorantraniliprole, imidacloprid, spinosad, and ivermectin. Each insecticide reduced larval motility, but response times and profiles differed among insecticides. Two sets of target site mutants previously identified in mortality studies on the basis of imidacloprid or spinosad resistance phenotypes were tested. In each case the resistant mutant responded significantly less than the control. The WI was also able to detect a spinosad response in the absence of the primary spinosad target site. This response was not detected in mortality assays suggesting that spinosad, like many other insecticides, may have secondary targets affecting behaviour. The ability of the WI to detect changes in insecticide metabolism was confirmed by overexpressing the imidacloprid metabolizing Cyp6g1 gene in digestive tissues or the central nervous system. The data presented here validate the WI as an inexpensive, generic, sub-lethal assay that can complement information gained from mortality assays, extending our understanding of the genetic basis of insecticide response in *D. melanogaster*.

Introduction

Insect pests are major vectors of human infectious diseases and also impose massive production losses and control costs in agriculture [1]. Chemical insecticides are often used as a weapon of choice to control or eliminate pests. While insecticides are potent in the short term, long term efficacy is often stymied by the evolution and spread of resistance [2,3]. There are a
limited number of chemicals, targeting an even smaller number of insect target proteins, meaning it is critical to manage or even prevent resistance to current generation insecticides [4]. A deeper understanding of the insect’s biological response to an insecticide, and thus the array of resistance options available to insect pests, underpins rational resistance management strategies. The vinegar fly Drosophila melanogaster, while not a pest species, has proven to be a powerful model that can be manipulated to interrogate the insect: insecticide interface [5–7].

There are many genetically determined biological mechanisms that influence how an insect responds to an insecticide. Among these, two that are particularly intensively studied because of their large contribution to insecticide resistance both in field and laboratory settings. The overexpression of genes encoding drug metabolizing enzymes (DMEs) such as cytochrome P450s (CYPs), carboxylesterases (CoEs) glutathione S transferases (GSTs) and glucuronosyltransferases (UGTs), which have the capacity to modify the chemical structure of insecticides, can cause resistance [8]. For example, the overexpression Cyp6cm1 in Bemisia tabaci and Cyp6g1 in D. melanogaster is sufficient to confer resistance to a range of insecticide classes [9–11]. The modification or loss of insecticide target sites is another commonly found and well-studied resistance mechanism. In D. melanogaster a variety of mutations in Da1 or Dfj2 niconitin acetylcholine receptor (nAChR) subunit genes confer imidacloprid resistance [12], while knockouts and amino acid substitutions in the Da6 subunit confer resistance to spinosad [13,14]. These findings were mirrored by studies in other insects, which found nAChR modifications linked to imidacloprid resistance in Nilaparvata lugens [15] and M. persicae [16] and to spinosad resistance in Plutella xylostella and Franklinella occidentalis [17,18].

Much of what is currently known about the mechanisms by which insects respond to insecticides comes from mortality assays [19], but sub-lethal insecticide response is an economically relevant phenotype that can describe additional information about insecticide biology [20,21]. Pest populations are frequently subjected to sub-lethal exposures in field settings [22,23], which have been shown to provide selection pressure [24]. For example, sub-lethal exposures to insecticides have been shown to significantly alter the fecundity [25], feeding behaviour [26] and locomotion [27] of targeted pest species. Each of these phenotypes have fitness costs and are often elicited by concentrations far below the LC50 value of a compound [28]. Considering these sub-lethal effects on pests may also give new insights into the mechanisms that influence insecticide resistance. Following the ingestion of an insecticide there is a cascade of biological processes which determine the response of the insect, with death as one possible endpoint. Observing the impact of sub-lethal exposures on living insects could focus on processes earlier in this cascade and the potential mechanisms of resistance that may arise from them. These points can also be extended to non-pest insects, which often undergo the same sub-lethal exposures [20]. Special attention has recently been given to the sub-lethal neonicotinoid exposures in honeybees that have been suggested to contribute to colony collapse disorder [29], although this is controversial [30].

Several methods to measure the sub-lethal effects of xenobiotics have been developed. For example, the crawling speed and pattern of D. melanogaster larvae from sensitive and control backgrounds were found to differ during acute exposure to ethanol [31]. More recently, several video tracking softwares such as Ethovision (Nodulus) and VideoTrack (ViewPoint Life Sciences) have been used to more precisely measure insect motility. These kinds of software have been used to describe sub-lethal insecticide response in mosquitos [27] and beetles [32] with a range of different insecticides. Other studies have measured alterations in sub-lethal phenotypes discussed above such as fecundity, feeding behaviour and locomotion after sub-lethal xenobiotic exposures [25–27].

Here, we describe an additional sub-lethal behavioural assay the Wiggle Index (WI; 38), which uses an open source ImageJ macro to quantify the temporal motility response of
D. melanogaster larvae to sub-lethal insecticide exposures. Larvae from a wild type susceptible strain were tested with the WI with insecticides from four mode of action (MoA) classes. The response curve of the larvae differed depending on the insecticide used. The capacity of WI to detect insecticide resistance was tested using several pairs of resistant and matching control (susceptible) genotypes previously characterized using mortality assays. In each case resistant and susceptible genotypes were distinguishable. Additional biological information was also gained by observing motility response to the insecticide spinosad in a highly resistant target site mutant, suggesting additional target(s) for this compound. Using the wealth of genetic resources available in D. melanogaster, the WI can be used to investigate the role of individual genes in sublethal insecticide response.

Methods

Fly Stocks

All stocks were maintained at room temperature and under constant light. Armenia14 is an iso-female line derived from Armenia60 (Drosophila Genomics Resource Center #103394), and was used as a susceptible wild type for all dose response experiments. Three additional sets of genotypes consisted of lines previously associated with either imidacloprid or spinosad resistance in D. melanogaster, and had genetically matched control backgrounds (Table 1). Additionally, each of these sets had relatively well characterized resistance mechanisms that have been elaborated on in other insect species [17] or with in vitro studies [33,34]. Set one contains nAChR subunit mutants Da1144 and Dβ21,351Q that were generated with EMS mutagenesis of Armenia14 and are resistant to imidacloprid in mortality assays [12]. Da1144 produces a truncated Da1 protein product that is cut short in the transmembrane M4 domain, and Dβ21,351Q contains a single amino acid replacement in the Dβ2 subunit. Set two includes two genotypes also generated with EMS mutagenesis and selection of Armenia14 carrying mutations in the D06 nAChR subunit (Dα6305, Dα6W357); each confers spinosad resistance [7]. Dα6W357 is truncated 13 residues after the third transmembrane domain, while Dα6305 has no detectable Dα6 expression. Set three used the HR_GAL4 to drive the expression of Cyp6g1 in the midgut, Malpighian tubules, fat body (referred to as digestive tissues here) by crossing it to the UAS_Cyp6g1 line [6]. A matched control for this genotype was generated by crossing the HR_GAL4 driver line to the F86FB genotype, which has the same genetic background as the UAS_Cyp6g1 line [35]. Previous studies linked overexpression of Cyp6g1 in these digestive tissues with resistance to multiple classes of insecticide [36]. Finally, in set four the overexpression of Cyp6g1 was achieved in the central nervous system (CNS) using the Elav-GAL4 (Bloomington Stock

<table>
<thead>
<tr>
<th>Name</th>
<th>Background</th>
<th>Stock #</th>
<th>Resistance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Armenia14</td>
<td>Armenia</td>
<td>DGRC (103394)</td>
<td>Susceptible</td>
</tr>
<tr>
<td>Da1144</td>
<td>Armenia14</td>
<td>Batterham Lab Stock</td>
<td>Imidacloprid</td>
</tr>
<tr>
<td>Dβ21,351Q</td>
<td>Armenia14</td>
<td>Batterham Lab Stock</td>
<td>Imidacloprid</td>
</tr>
<tr>
<td>Da6305</td>
<td>Armenia14</td>
<td>Batterham Lab Stock</td>
<td>Spinosad</td>
</tr>
<tr>
<td>Dα6W357</td>
<td>Armenia14</td>
<td>Batterham Lab Stock</td>
<td>Spinosad</td>
</tr>
<tr>
<td>F86FB</td>
<td>F86FB</td>
<td>Biehoff et. al 2007</td>
<td>Susceptible</td>
</tr>
<tr>
<td>UAS_Cyp6g1</td>
<td>F86FB</td>
<td>Batterham Lab Stock</td>
<td>Imidacloprid</td>
</tr>
<tr>
<td>Elav-GAL4</td>
<td>Elav-GAL4</td>
<td>Bloomington (458)</td>
<td>Susceptible</td>
</tr>
</tbody>
</table>

All genotypes used in the current study are displayed along with their backgrounds and previously reported resistance status.

doi:10.1371/journal.pone.0145051.001
2.2.2) Obtaining Third Instar Larvae

For all isogenic lines, sixty 2–5 day old females and twenty males were collected and transferred onto maize meal medium (S1 Table) in vials sprinkled with dried yeast. For all crosses, the same procedure was followed with virgin females. In each case flies were left undisturbed for one day at 25°C for oviposition and then cleared from the vial. The vials were kept at 25°C for 68 additional hours to generate a population of early third instar larvae, which were recovered from the food using sucrose extraction [37]. Briefly, 30mL of 20% w/v sucrose (non-Analytical Reagent) in distilled H₂O was poured onto the food. The top layer of the food was then gently disrupted with a metal rod in order to release the larvae, which float in the sucrose solution. The solution was then carefully poured onto a fine cloth mesh to isolate the larvae, which were then dried and transferred onto grape juice agar plates (S1 Table). Third Instar larvae (~5mm in length) were then individually picked, 25 per well, into a NUNC cell culture treated 24 well plate (Thermo-Scientific) preloaded with 200µL Analytical Reagent 5% w/v sucrose (Chem Supply) in distilled H₂O. This volume of 5% sucrose was chosen because it is the minimum amount needed to cover the 1.8cm² bottom of the well. In this environment the larvae were able to translocate, but stationary gyrations represented the majority of their activity. To dose the larvae, 50µL of 5x insecticide stock solution was added to each well in order to bring the final concentration to the desired dosage. After mixing, 50µL of the mixed solution was removed in order to bring the final volume back to 200µL.

2.2.3) Insecticide Dilution

Chlorantraniliprole (10 gL⁻¹ Coragen®, Du Pont), Imidacloprid (200 gL⁻¹ Confidor®, Bayer Crop Science), and Spinosad (10 gL⁻¹ Success®, Yates) were all purchased commercially and diluted to 5,000ppm stocks using distilled water. On the day of exposure, 5x stocks were generated for each dose being used (Chlorantraniliprole: 60, 30, 15, 7.5, 3.75, 0 ppm; Imidacloprid: 240, 120, 60, 30, 15, 0 ppm; Spinosad: 240, 60, 30, 15, 7.5, 3.75, 0 ppm) by diluting the 5,000ppm stock in 5% Analytical Reagent sucrose (Chem Supply). A similar procedure was followed for Ivermectin (Sigma), and a 10,000ppm stock was generated using DMSO as a solvent. 5x stocks were generated (120, 60, 30, 15, 7.5, 0 ppm) by dissolving the original stock in 5% sucrose, and the highest concentration of DMSO used in dosing was added to the 0ppm solution in order to control for solvent effects, none of which were observed.

2.2.4) Experimental Design

To assess the effect of different insecticides on movement, the susceptible control Armenia14 was tested with a range of 5 doses for imidacloprid (3ppm-48ppm), ivermectin (1.5ppm-24ppm), spinosad (1.5ppm-24ppm), and chlorantraniliprole (75ppm-12ppm). Pilot studies were used to determine the best range of doses to be used for each insecticide. Set one mutants (D6T and D672L) were tested at 12 and 48ppm imidacloprid and the same doses were used for set two mutants (D6T and D672L) with spinosad. Set three and four genotypes expressing Cyp6g1 in the three digestive tissues or CNS were tested at 48ppm imidacloprid.

2.2.5) Filming

Ten second videos were taken at 5, 10, 15, 30, 45, 60, 120 and 240 minutes after the addition of insecticide and compared to videos taken immediately before the addition of the insecticide.
(time = 0). Each video captured larval activity in 4 individual wells that were processed separately. All filming was done using a Panasonic 3CCD Ultra-Compact Digital Palmcorder against a LED light box (Huion L4S Led Light Pad). As some of the insecticides examined here are light sensitive [38,39], larvae were kept in darkness except for a brief period before and during filming. 15 seconds was allowed for larvae to acclimate to light conditions on the LED pad, before filming began. Because each 24 well plate required 6 separate videos to be taken, larvae were exposed to the light for approximately 120 seconds for each time point.

**Video Processing and Analysis of Larval Movement**

Raw videos were named in bulk via a renaming script written in R (R: A language and Environment for Statistical Computing). Named videos were split into jpeg image sequences using the free software Video Jpg Converter (DVDFvideoSoft), and a ten second video produced 250 frames for analysis. Image Sequences were then uploaded to a server on the NeCtar research cloud for further processing. There, the Fiji distribution of ImageJ [40] was used to run the WI script (Fig 1; S1 Script).

The WI script analysed individual image sequences and calculated the total motility of the larvae in a given well. Firstly, the script crops one of the four wells from the image sequence and converts it into stacks of tiff format (Fig 1A). An ImageJ algorithm was then performed similar to that of Preston et. al [41], which is briefly outlined here. Prior to any calculations a Gaussian Blur of 2 pixels was applied to each tiff stack in order to minimize the impact of artefacts due to video compression. Then for each pixel, a series of frame window standard deviations (\(\sigma_{FW}\)) was calculated by measuring the standard deviation of the pixel's light intensity for

\[
\sigma_{FW} = \sqrt{\sum_{i=1}^{n} (I(i) - \mu)^2 / n}
\]

where \(I(i)\) is the intensity of pixel \(i\), \(\mu\) is the mean intensity of the pixel window, and \(n\) is the number of frames in the window.

**Fig 1. The Wiggle Index.** The Wiggle Index measures total motility in a given well. A) Individual wells are cropped out of a sequence of 250 jpeg images B) \(\sigma_{FW}\) values are calculated for each pixel by calculating the standard deviation of all \(\sigma_{FW}\) values over the entire sequence. C) \(\sigma_{FW}\) from each pixel are filtered based on a threshold value and then averaged to yield the Wiggle Index (WI) value.

\[\text{WI Value} = \frac{\text{Sum of } \sigma_{FW} \text{ values above threshold}}{\text{Total number of pixels above threshold}}\]

\[\text{doi:10.1371/journal.pone.0145051.g001}\]
150 frame windows in a rolling manner starting with frame 1 (F1-F150) and continuing until the end of the stack (F101-F250). The formula is:

$$\sigma_{FW} = \left( \frac{1}{F} \sum_{i=1}^{W} (x_{fi} - \mu_{FW})^2 \right)^{1/2}$$

where W denotes the number of windows, F denotes the frame window size, $\mu_{FW}$ denotes the average frame window intensity and $x_{fi}$ denotes the frame intensity value. A second intermediate variable ($\sigma_s$) was then calculated for each pixel by taking the standard deviation of the series of previously calculated series of $\sigma_{FW}$ values (Fig 1B). The formula for these calculations is:

$$\sigma_s = \frac{1}{W} \sum_{i=1}^{W} (-\mu_{FW})^2$$

where $\mu_{FW}$ is the frame window intensity average.

Using the $\sigma_s$ values for each pixel, WI values were calculated, which represent the total motility of the larvae in the ten second video. This was accomplished by averaging the $\sigma_s$ values for every pixel above a cut-off threshold (Fig 1C). The formula for WI value is:

$$WI\ Value = \frac{\Sigma_{\sigma_s \geq T} \sigma_s}{TA}$$

where T is the threshold cut off, TA is the number of pixels above the threshold and $\Sigma_{\sigma_s \geq T} \sigma_s$ is the sum of values above the threshold. The cut-off threshold, in this case 30, was applied to remove any background noise due to global movement of the plate or fluctuations of the liquid.

**Visualization and Statistics**

The WI script generates heat maps (Fig 1C) and reports numeric estimates of total motility (WI values) for each well at a given time. Two independent corrections were applied after the calculation of WI values to account for differences in initial motility between genotypes and to isolate the organisms’ responses to the insecticide.

The first correction divided WI values at a given time point by the WI value of the same larva prior to the addition of the insecticide and generated relative motility ratio values (RMR values; Fig 2). The RMR correction normalized the motility of larvae to those same larvae in the absence of insecticide. For example, a genotype with a WI value of 10 before the addition of an insecticide and 3 at a later time point would have a RMR value of 3/10 = .3 at that time point. Mean corrected RMR values were plotted together with their associated 95% confidence intervals to visualize the response of different genotypes for a given dose or the response to different doses for a given genotype. Visualizing the data allowed for the description of several aspects of the insecticide response. The Response Time was defined as the time needed for a treatment to significantly and irreversibly reduce the RMR value less than one based on a 95% confidence interval. The End Point RMR value was defined as the final RMR value recorded (240 minutes) and was compared among doses or genotypes using Tukey’s Honestly Significant Difference (HSD) pairwise test (P ≤ .05).

The second correction fits a generalized linear model (GLM) to the response curve based on WI values from a particular well over time. In the case of spinosad, no transformation of the data was needed, but for imidacloprid, a log_{10} transformed time scale was used as it better fit the data based on adjusted R squared values. The inverse of the slopes of these regression lines ($\beta$ values) were used as quantitative measurements of overall response. $\beta$ values of each genotype were compared to control lines using a Student’s t test (P ≤ .05) and mean $\beta$ values were
Results

The Response of Wild Type to Different Insecticides

To observe the effect of different classes of insecticides on motility, dose response profiles were generated for *Armenia* with four insecticides from distinct (MoA) classes (*Fig 3; Fig 4*). Regardless of dose or insecticide, each chemical reduced the motility of larvae decreased as exposure time increased, but two distinct response types were observed depending on the insecticide used, *fast acting* and *slow acting*. Fast Acting responses were characterized by rapid, dose dependent Response Times and dose dependent End Point RMR values, while the Slow Acting response displayed delayed, dose dependent Response Times but with similar End Point RMR values. Imidacloprid and chlorantraniliprole were both Fast Acting, with Response Times of ten minutes or less at the highest doses tested followed by End Point RMR values that were significantly different among doses (*Fig 3A and 3B; Fig 4A and 4B*). Spinosad and Ivermectin were Slow Acting, with Response Times of 45 minutes or greater. The End Point RMR values for these Slow Acting insecticides were not significantly different as determined by Tukey’s HSD (*Fig 3C and 3D; Fig 4C and 4D*).

Set One: Imidacloprid Resistant nAChR Alleles

The motility of previously characterized imidacloprid resistant *DaeI* and *DaeII* mutants was examined at two different concentrations of imidacloprid (12, 48ppm). *Armenia* showed a greater response to imidacloprid than either of the resistant mutants in the GLM analysis (*Fig 5B and 5C*) and End Point RMR values (*Fig 5B and 5C*). *DaeI* had the longest Response Time and the highest End Point RMR value. *DaeII* displayed an intermediate Response.

---

Fig 2. Wiggle Index Output and RMR Calculation. The Wiggle Index produced heat maps and numeric estimates of total larval motility in each well at a given time. RMR values were calculated by dividing the WI value at time = x from the WI value from the same well at time = 0.

doi:10.1371/journal.pone.0145051.g002

plotted with 95% confidence intervals. An interactive R script *Wl_Analysis* was written that accepts raw data from the WI script and generates a series of descriptive graphs including RMR curves and GLM analysis (available on GitHub https://github.com/shanedenecke/Wl_Analysis).
2.3.3) Set Two: Spinosad Resistant Da6 Alleles

2.3.4) Set Three and Four: Cyp6g1 Overexpression

**Fig 3. Armenia**

*Fig 3. Armenia* was tested using the Wiggle Index at 5 doses of 4 insecticides. Chlorantraniliprole (A) and Imidaclorpid (B) each show rapid response times and dose dependent long term RMR values. Ivermectin (C) and Spinosad (D) responded differently, displaying delayed response times and similar long term RMR values. All points correspond to the mean RMR values with 95% confidence intervals. Boxes in each plot indicate values between 0 and 30 minutes magnified in *Fig 4.*

doi:10.1371/journal.pone.0145051.g003

Time and End Point RMR Value compared to the other genotypes. In the absence of imidaclorpid, there were no significant differences among genotypes in the GLM or RMR analysis (*Fig 5A and 5A*).

**Set Two: Spinosad Resistant Da6 Alleles**

*Da6* mutant genotypes, *Da6<sup>Ar</sup>* and *Da6<sup>W337</sup>* were screened at 2 concentrations of spinosad (12, 48ppm). At each dose both *Da6* mutants showed delayed Response Times compared to Armenia<sup>14</sup>. At 12ppm *Da6<sup>W337</sup>* had a larger End Point RMR value when compared to Armenia<sup>14</sup>, but the *Da6<sup>Ar</sup>* End Point RMR value was not statistically different. At 48ppm there were no differences in End Point RMR values among any of genotypes (*Fig 6B and 6C*). GLM analysis indicated small but significant differences between Armenia<sup>14</sup> and each *Da6* mutant at 12ppm, but at 48ppm only *Da6<sup>W337</sup>* was significantly different from the control (*Fig 6B’ and 6C’*). Each genotype displayed near identical response profiles at 0ppm, indicating that there were no motility differences in the absence of insecticide (*Fig 6A and 6A’*).

**Set Three and Four: Cyp6g1 Overexpression**

*Cyp6g1* was overexpressed in the midgut, fat body and Malpighian tubules and tested at 48ppm imidaclorpid. Lines that overexpressed *Cyp6g1*, showed a slightly delayed Response Time to
imidacloprid than the matched control and showed a significantly higher End Point RMR value, which was supported by GLM analysis (Fig 7B and 7B'). No significant difference was observed at 0ppm (Fig 7A and 7A').

Cyp6g1 overexpression in the CNS (Elav x UAS_Cyp6g1 c.f. Elav x D86FB) was tested at 48ppm imidacloprid. Elav x UAS_Cyp6g1 displayed a 5 minute delay in Response Time and had an End Point RMR value significantly higher than the control (Fig 8B). GLM analysis also found differences between overexpression and control genotypes (Fig 8B'). Overexpression of Cyp6g1 in the CNS had less of an impact on all phenotypes than overexpression in the digestive tissues.

Discussion

Armenia\textsuperscript{4} Dose Response Profiles

Insect pests impose significant costs on agriculture and human health, but lack of descriptive phenotypes have hindered a deeper understanding of how insects respond and interact with insecticides. Here, a novel method of assessing sub lethal insecticide response in the model insect D. melanogaster is described, in part, by testing a wild type strain with insecticides of different MoA classes. This strain (Armenia\textsuperscript{4}) responded in two distinct ways depending on which of the 4 insecticides was used. Although each of these insecticides belongs to a distinct MoA class, characteristics common between classes may partially explain the types of response
profiles observed. Imidacloprid and spinosad target distinct nicotinic acetylcholine receptor subtypes found in the CNS, ivermectin targets ligand gated chloride channels in the same tissue [42], and chlorantraniliprole binds to the ryanodine receptor in muscle cells [43]. Because, imidacloprid, ivermectin, and spinosad all target receptors in the CNS, they must therefore cross the blood-brain barrier. The molecular weight of xenobiotics has been shown to influence their rates of diffusion [44] and transport [45] into the CNS. This could explain the differences in Response Time observed in the WI. Spinosad and ivermectin have molecular weights of 732 Da and 875 Da respectively, while imidacloprid is relatively smaller with a weight of 255 Da. Chlorantraniliprole, though of intermediate size (430 Da), would not be affected by this phenomenon due to its targeting of the more easily accessible ryanodine receptor in muscle cells [43].

Target Site Mutants

For each set of target site mutants used in this study, significant differences were detected in insecticide response for resistant mutants compared to their matched controls using the WI. Following the addition of imidacloprid, there was always a reduction in RMR values. Dael and
D$	ext{J}$62 mutants displayed significantly higher End Point RMR values and lower $\beta$ values than did Armenia$^{14}$ (Fig 5). In contrast, D$	ext{Oo}$6 mutants tested with spinosad displayed similar End Point RMR values to Armenia$^{14}$, but their Response Time was significantly delayed (Fig 6). These same mutants were previously found to be more resistant in mortality assays, and it was hypothesized that this was due to a decreased ability of the insecticide to bind its primary target [7,12]. While both mortality and motility are complex phenotypes, these data do link specific alleles to insecticide response phenotypes. Resistance levels observed in mortality assays, however, do not correlate perfectly with the data presented here. For example, D$	ext{Oo}$1$^{562}$ showed less response to imidacloprid than D$	ext{J}$62$^{3550}$ as measured by all three WI phenotypes even though the latter has higher resistance in mortality assays [46]. The partial overlap in the mortality and motility phenotypes for these mutants suggest that different aspects of insect/insecticide interactions are being detected by each assay. These differences could be partially explained by differences in the concentrations of insecticides, stages of development and exposure methods in the two assays. The mortality assays discussed here, were performed on first instar larvae reared
Fig 7. *Cyp6g1* Expression in the Digestive Tissues. *Cyp6g1* overexpression was achieved in the midgut, Malpighian tubule and the fat body using the GAL4-UAS system. At 0ppm (A, A'), no significant differences were observed between genotypes. At 48ppm imidacloprid the overexpression line responded less as measured by Response Time, End Point RMR values and GLM analysis. All plots display mean RMR or $\beta$ values with 95% confidence intervals.

doi:10.1371/journal.pone.0145051.g007

Fig 8. *Cyp6g1* Expression in the Central Nervous System. *Cyp6g1* overexpression was achieved in the CNS using the GAL4-UAS system. At 0ppm (A, A'), no significant differences were observed between genotypes. At 48ppm imidacloprid the overexpression line responded less as measured by Response Time, End Point RMR values and GLM analysis. All plots display mean RMR or $\beta$ values with 95% confidence intervals.

doi:10.1371/journal.pone.0145051.g008
to adulthood on solid media dosed with insecticide with much of the mortality occurring at the first instar stage. In contrast, the WI used third instar larvae, exposed for a short time (240 minutes) in liquid media.

The \textit{Dea6} mutant, which has no detectable \textit{Dea} expression, still responds to spinosad in the WI and has similar End Point RMR values when compared to Armenia\textsuperscript{14} and \textit{Dea6}\textsuperscript{W327} even though the two \textit{Dea} mutants have far higher mortality LC\textsubscript{50} value (Fig 5; Fig 8). To date \textit{Dea} orthologues are the only genes for which a role in spinosad target site resistance has been demonstrated [17,47]. The \textit{Dea6}\textsuperscript{W65} data presented here suggest that spinosad may also bind to at least one additional target impacting motility. Early mode of action studies indicated that other ligand gated chloride channels may also be spinosad targets [48]. No mutations were found in the \textit{Dea} orthologue in a spinosad resistant strain of \textit{Musca domestica} [49], but there is evidence that spinosad can bind to \textit{M. domestica} \textit{a6} [7]. The WI has allowed us to make this significant observation, not detected using classical mortality studies.

There is the potential that baseline differences in motility in the absence of insecticide may confound any motility responses measured by the WI. In the current study the only genotypes to show such differences were the imidacloprid target site mutants (\textit{Dea1}\textsuperscript{M4} \textit{Db2}\textsuperscript{35325}) with each mutant moving significantly less than Armenia\textsuperscript{14} at time point 0 (S1 Fig). Despite lower initial WI values, in the longer term the WI values of the mutants were higher than Armenia\textsuperscript{14}. Therefore, even in the absence of any correction for motility differences in the insecticide free control, the WI was still able to differentiate target site mutants. However, only matched genetic backgrounds were tested in the current study. It is therefore possible that comparing field populations may restrict the WI to describing alleles of large effect as is the case with mortality testing [24]. However, quantitative genetics approaches including genome wide association studies may help overcome this obstacle (See section 4.4).

\textit{Cyp6g1}

The impact of imidacloprid on motility was reduced significantly when \textit{Cyp6g1} was overexpressed in the digestive tissues (Fig 6). These data correlate with those previously reported for mortality assays [10]. As imidacloprid targets nACHRs in the CNS, the response of larvae may be an indicator of the levels of imidacloprid and/or toxic metabolites in the CNS at a given time. Hence, the data presented here could be explained if the overexpression of \textit{Cyp6g1} reduces the effective concentration of imidacloprid in the CNS. This hypothesis is supported by data from metabolic studies conducted under the same exposure conditions employed here (3rd instar larvae exposed to imidacloprid in 5% sucrose solution). Hoi et al. [50] used Twin Ion Mass Spectrometry to show that increased \textit{in vivo} expression of \textit{Cyp6g1} in the digestive tissues leads to a large increase in the metabolism of imidacloprid to toxic metabolites that are rapidly excreted. As a consequence, when \textit{Cyp6g1} is overexpressed in the metabolic tissues, significantly lower concentrations of imidacloprid and metabolites are retained in the body with access to the CNS.

The impact of imidacloprid on motility was also significantly reduced when \textit{Cyp6g1} was overexpressed in the CNS (Fig 7). While the capacity of \textit{Cyp6g1} to provide neuroprotection has not been previously reported, a role has been proposed for a Cytochrome P450, \textit{Cyp6cm1}, in \textit{Tribolium castaneum} [9].

One advantage of using the WI in combination with the GAL4-UAS system is that it provides the capacity to determine the route(s) that an insecticide follows to arrive at its target. It is theoretically possible that under the conditions used for the WI, that the insecticides tested may have entered the body orally, reaching the CNS via the digestive system, or through the cuticle, gaining direct access to the CNS via the hemolymph. Cuticular thickness has been cited
as a resistance mechanism in adult *D. melanogaster* [51]. The large difference in RMR values at early time points (5 minutes; Fig 7B) when *CytoG1* is overexpressed in digestive tissues indicates that a significant portion of imidacloprid reaches the CNS via the digestive system, probably due to oral ingestion. The path that insecticides and their metabolites follow in larvae could be further studied by using the WI in combination with Twin Ion Mass Spectrometry [50] and the GAL4-UAS system to examine the role of different families of genes such as those encoding transporter or cuticular proteins.

**The Utility of the WI**

While only the response of *D. melanogaster* was tested in this study, this method offers several insights into insecticide biology in pest species. *D. melanogaster* has been frequently used to elaborate on resistance mechanisms in matched genetic backgrounds. For example, the contribution of individual pest genes to resistance has been measured by driving the expression of genes encoding metabolic enzymes and targets in *D. melanogaster* [6,9,52]. Beyond the transgenic expression of pest genes in *D. melanogaster*, there is also the possibility that the methods used in this study may be modified for use in some other insect species. A similar exposure method has already been used to study drug responses in the parasitic nematode *Haemonchus contortus* (38). While the exposure method described here (24 well plates, chemical concentrations etc.) may be *D. melanogaster* specific, the ImageJ algorithm itself only measures the average change in light intensity of pixels in a video. Thus, in theory a variety of different exposure conditions may be used so long as the organism under investigation can be differentiated from its background.

Several technical aspects of the WI make it an attractive option for assessing an insecticide response. Analysing motility is a relatively cheap, high-throughput process when compared with other methods. The WI setup requires only an inexpensive video camera, a source of even light and free downloadable software (See Materials and Methods). Typical mortality assays are measured in adults by analysing how many adults survive insecticide exposure after a given interval or in larvae by measuring what proportion of individuals develop to adulthood [19,53]. The WI measures insecticide response earlier in the life cycle of the organism and significant differences can be observed between genotypes in as little as 5 minutes. The ease of sampling (filming) allows for a nearly unlimited number of time points that can be analysed in a single run. While in this study three parameters were used to quantify insecticide response, many other features of the response profile could be analysed. These factors make the WI suitable for screening large numbers of genotypes.

One of the distinct advantages of using *D. melanogaster* is the ability to implicate individual genes in a phenotype by using previously generated genotypes from stock centres such as Vienna *Drosophila* RNAi Centre [54], the Bloomington Deficiency Kit [55] and the MiMic library [56]. Each stock in these collections contains a defined genetic change in a common background, much like the genotypes discussed in the current study. However, other *Drosophila* specific resources use genetic variation similar to those observed in field settings to implicate alleles in complex traits. The *Drosophila* Genetic Reference Panel is a collection of fully sequenced homozygous stocks derived from one field population [57,58]. By phenotyping the collection and performing a subsequent genome wide association study, alleles of potentially very small effect can be associated with a complex trait, such as sub-lethal insecticide response, as has been done for other traits [59,60]. A similar resource, the *Drosophila* Synthetic Population Resource, relies on a synthetic populations derived from several fully sequenced founder individuals [61] and has previously been used to describe the genetic basis for nicotine resistance [62]. The use of these resources, in combination with the WI may lead to the
identification of previously undescribed genes influencing insecticide ingestion, metabolism, transport, excretion and those encoding targets.

Conclusion

The WI quantifies insecticide response by measuring motility and is a valuable addition to the current array of bioassays. By measuring the motility of larvae as they are being exposed, a more complete understanding of this response can emerge. Using insecticides of distinct classes and validation with previously characterized metabolic and target site resistant *D. melanogaster*, we have demonstrated that this assay is both sensitive and rapid, while providing novel biological information.

Supporting Information

S1 Fig. WI Value Plot. The uncorrected responses (WI values) for imidacloprid resistant alleles during exposure to 48ppm imidacloprid over A) 240 minutes and B) 30 minutes. Despite lower starting values, the WI Value prior to correction are still clearly capable of discriminating between the known resistant alleles and susceptible strain (Armenia).

(S1 Fig)

S1 Script. The Wiggle Index Script. Written in imagej macro language, this script will iterate over a folder of image sequences and will return a sub-folder full of heat maps and a table of WI values. These images and values correspond to the motility of larvae in one well in each image sequence in the folder.

(S1 Script)

S1 Table. Media Recipes. Recipes for both maize meal media (A) and grape juice plates (B) used in this study.

(S1 Table)

Acknowledgments

The authors would like to thank the Basler lab, the Bloomington stock centre and the Drosophila Genome Resource Center for the provision of *D. melanogaster* stocks and the Australian Drosophila Biomedical Research Support Facility for handling the quarantine of imported flies. This research was supported by funding from the Australian Research Council (Discovery Project Grant DP130102414) awarded to PB.

Author Contributions

Conceived and designed the experiments: SD PB TP. Performed the experiments: SD. Analyzed the data: SD AFL. Contributed reagents/materials/analysis tools: SD CJN. Wrote the paper: SD TP PB.

References


2. Palumbi SR. Evolution—Humans as the world's greatest evolutionary force. Science (80-). 2001; 293: 1786–1790. PMID: 11546683


Chapter III: Using the Drosophila Genetic Reference Panel to Describe the Genetic Basis of Imidacloprid Response
3.1) Introduction

Because of its widespread agricultural use, imidacloprid is one of the most important insecticides in the arsenal of global food producers. Despite frequent studies dealing with imidacloprid resistance in various insect species, much still remains uncertain about the mechanisms and genes that underpin this trait and those that have the potential to do so. While certain QTLs and transcripts have been shown to influence imidacloprid resistance, the magnitude of the contribution of any genetic variant has not so far been quantified. Furthermore, there are undoubtedly many genes which contribute to imidacloprid resistance which have yet to be identified. These may be members of gene families with well annotated roles in insecticide resistance (such as P450s) or they may act in pathways that have not yet been associated with resistance. Identifying and quantifying the contribution of the QTLs and transcripts involved in imidacloprid resistance is paramount to managing agricultural pest populations. The examination of imidacloprid resistance in *D. melanogaster* has proven to be a valuable tool in understanding how resistance can and does evolve in economically important pest species (Perry et al., 2011).

3.1.1) Imidacloprid Resistance: What is known

One of the best understood facets of imidacloprid’s biology is its target site; it is thought to primarily bind to a subset of insect nAChR subunits (Bai et al., 1991). The affinity of imidacloprid for these subunits is high compared to related molecules like nicotine and is the primary cause of the specificity of this insecticide for insects compared to mammals (Tomizawa and Casida, 2009). Although initial studies suggested that binding occurred at the interface of two subunits, the specific subunits were at first unknown. Some light was shed on this question by the discovery of a conserved Y151S substitution in the brown planthopper, *Nilaparvata lugens*, α1 and α3 subunits, which was associated with resistance (Liu et al., 2005). Additionally, *D. melanogaster* mutagenesis studies that found that truncations or amino acid substitutions in the α1 or β2 subunits conferred imidacloprid resistance (Perry et al., 2008). Further, systematic manipulation of nAChR subunits has also pointed to a role for the α3 subunit (Chen unpublished). More recent work indicates that imidacloprid likely also binds to non-nAChR proteins such as the RDL, a GABA gated chloride channel subunit (Taylor-Wells et al., 2015). These data suggest that imidacloprid is capable of binding an array of ligand gated chloride channel subunits, and that modifications of these subunits can confer resistance.

Another known factor is the ability of imidacloprid resistance to be conferred via the overexpression of P450s. This mechanism is far more commonly reported in field resistance to
imidacloprid than is target site mediated resistance (Bass et al., 2015) possibly due to a fitness cost associated with target site resistance (Somers et al., 2017). In Bemisia tabaci, P450 mediated imidacloprid resistance was first described in 2002; the gene Cyp6cm1 was later identified as the primary P450 involved (Karunker et al., 2008; Nauen et al., 2002; Rauch and Nauen, 2003). Similar cases of individual P450s conferring imidacloprid resistance have been reported in N. lugens with Cyp6a1 (Ding et al., 2013) and Myzus persicae with Cyp6cy3 (Puinean et al., 2010). Most relevant to the current work, overexpression of the D. melanogaster P450 Cyp6g1 has been shown to confer resistance to imidacloprid and insecticides of other chemical classes (Daborn et al., 2001; Daborn, 2002).

The evolution of Cyp6g1 merits further discussion as it is the best characterized insect DME. In all Drosophila species so far tested, the Cyp6g1 gene has the capacity to confer resistance to at least one class of insecticides, suggesting that this function was present 40 million years before the use of modern synthetic insecticides began (Harrop et al., 2014). The expression of Cyp6g1 is highly variable in field populations of D. melanogaster due to the insertion of the long terminal repeat of the retroviral element Accord, copy number variation and further transposable element (TE) insertions (Battlay et al., 2016; Catania et al., 2004; Schmidt et al., 2010). The Accord insertion contains tissue specific enhancer sequences that drive the overexpression of Cyp6g1 in the midgut, Malpighian tubules and fat body (Chung et al 2007). The ancestral M haplotype contains a single copy of Cyp6g1 and expresses low levels of the gene compared to the more derived AA haplotype, which contains an Accord insertion, several partial repeats of Cyp6g1 and Cyp6g2 and a subsequent duplication of Accord-Cyp6g1 cassette (Schmidt et. al 2010). Further modifications of the AA haplotype resulted from the insertions of the TE HMS-Beagle and a P element upstream of Cyp6g1, creating the BA and BP haplotypes, respectively (Fig 3.1). These derived haplotypes have been associated with increased levels of Cyp6g1 expression and resistance to insecticides such as DDT and azinphos-methyl (Schmidt et. al 2010, Battlay et. al 2016). Furthermore, they appear to have arisen in the last 100 years (Schmidt et. al 2010). In one population, Cyp6g2 expression appears to correlate with Cyp6g1 expression, but the contribution of this gene to resistance has not been shown (Battlay et. al 2016).

The structural modifications of imidacloroprid by DMEs such as CYP6G1 have also been well studied in various organisms including insects. Broadly, there are two structural changes that imidacloprid can undergo, nitroreduction and oxidation (Fig 3.2; Fusetto unpublished). Metabolites from both pathways have been detected in plants, animals and insects, but soil bacteria produce exclusively the nitroreduction metabolites (Ford and Casida, 2008, 2006; Pandey et al., 2009). The low levels of nitroreduction metabolites produced in animals and plants may be due to bacteria
living within other organisms. This was the case in *D. melanogaster* where nitroreduction metabolites were linked to the presence of symbiotic bacteria (Fusetto unpublished).

Figure 3.1- *Cyp6g1* structural variation:
The 6 different alleles of Cyp6g1 are shown. Of these the M (singe copy of *Cyp6g1*; No TE insertion), AA (Two copies of *Cyp6g1*; Accord insertion), and BA (Two copies of *Cyp6g1*; Accord insertion; HMS Beagle Insertion) are present within the DGRP. Adapted from Schmidt et. al (2010).
Figure 3.2- Imidacloprid Metabolism:
A schematic for the known pathways of imidacloprid (IMI) metabolism is shown. IMI can be directly converted to IMI-NNO, IMI-de, or IMI-5OH and from there further metabolized.
Insect P450s are thought to produce oxidative metabolites exclusively, and the metabolites formed by CYP6G1 have been the best characterized. Heterologous expression of CYP6G1 in tobacco cells produced the metabolites IMI-5-OH, IMI-Olefin and IMI-diol (Joussen et al., 2008). These results were replicated when driving the in vivo expression of Cyp6g1 in D. melanogaster (Hoi et al. 2014). Interestingly though, while increased Cyp6g1 expression caused increases in IMI-Olefin levels, Cyp6g1 was found to be unable to catalyse the conversion of IMI-5-OH to IMI-Olefin reaction, meaning that at least one other gene may be involved in imidacloprid metabolism (Fusetto unpublished).

### 3.1.2) Imidacloprid Resistance: What is Unknown

If the conversion of IMI-5-OH to IMI-Olefin is catalysed by an enzyme, then the gene encoding that enzyme has not been identified, highlighting how little is known about imidacloprid metabolism. Cyp6g1 is so far the only Drosophila P450 linked to imidacloprid resistance, but it is not known which, if any, of the other 89 P450s in Drosophila may contribute to resistance. The trend is similar in other many other pest species where only a small fraction (usually only one) of the total P450s have been linked to imidacloprid resistance in field populations. While it is possible that no other P450 is capable of metabolizing imidacloprid, this appears unlikely because most other insecticides are not so exclusively acted upon by insect metabolic machinery. For example DDT resistance can be conferred by Cyp6a2, Cyp6g1, and Cyp12d1 (Amichot et al., 2004; Daborn et al., 2007). If it is assumed that some other P450 is capable of metabolizing imidacloprid in D. melanogaster, a failure to detect such a mechanism in field populations may be because a mutation that leads to overexpression of the gene may not have arisen yet. Alternatively, these genes are prevented from being upregulated in resistant strains due to high associated fitness costs (Bogwitz et al., 2005). Finally, it is possible that these P450s are playing a role, albeit a less prominent one, that is masked by other QTLs and transcripts which have larger effects. Higher resolution techniques are needed to determine the diversity of P450s which contribute to resistance phenotypes. This is a challenging problem to address in D. melanogaster and even more difficult in pests.

Even in cases where individual QTLs or transcripts (such as Cyp6g1) have been identified as contributing to imidacloprid resistance the magnitude of the contribution of a variant to insecticide resistance has never yet been directly measured. Statistical models which use either genomic polymorphism or transcript expression to predict imidacloprid resistance can be used to estimate the effect of a QTL or transcript, but this is a statistical approximation rather than a direct measurement. Indeed, a large gap often exists between resistance levels reported in field populations (often
attributed to a single gene or transcript) compared to laboratory genotypes transgenically overexpressing a single gene (Bass et al., 2013; Zhu et al., 2010). Direct measurement can only take place by deleting portions of the genome or by reducing the expression of a given transcript, a process made easier by recent advances in genome engineering (Venken et al., 2016).

Laying even further beyond the horizon are physiological mechanisms that have yet to be identified as contributing to imidacloprid resistance. It is possible to speculate about known physiological processes that, if altered in certain ways, could impact resistance levels. For example, nAChRs rely on a host of accessory proteins to make sure that they are properly trafficked and maintained at the neuronal membrane (Jones et al., 2010). Alterations of these proteins could have a dramatic impact on the amount or condition of nAChRs, which could change how much imidacloprid can bind to its target. Further changes in any process downstream of insecticide binding (synaptic transmission, neural connectivity) could have similar influences on how the insect responds to imidacloprid.

3.1.3) Forward Genetics to Address the Unknowns of Imidacloprid Resistance

Clearly, much of the genetic basis for imidacloprid resistance is unknown. One of the best ways to describe the genetic basis of any trait is to use forward genetic techniques (Section 1.6). The primary advantage of this is that neither the loci nor the physiological mechanisms that influence a trait are assumed a priori. Instead, forward genetic screens associate measured phenotypic and genetic variation in an unbiased way. While attempts have been made before to characterize imidacloprid in this way (Ilias et al., 2015; Kalajdzic et al., 2012), decreased cost of sequencing and computational analysis has increased the resolution.

Applying these techniques in model organisms further enhances detection capabilities by minimizing genetic heterogeneity. Only in model organisms is the genetic toolkit available to cheaply and accurately manipulate implicated QTLs and transcripts. The Drosophila Genetic Reference Panel (DGRP) exemplifies this concept. The DGRP is a collection of 201 fully sequenced inbred Drosophila stocks, which represents a snapshot of the genetic diversity from a single population in North Carolina, USA (Mackay et al., 2012). A GWAS using the DGRP can be accomplished by testing the associations of the ~2.5 million genetic variants (most commonly SNPs) across the DGRP genomes with phenotypic data on any trait. Further sequencing of the DGRP male and female transcriptomes allowed for similar association studies to be performed with transcript expression in a transcriptome wide association study (TWAS; Huang et al. 2015). So far the DGRP has been used to understand the genetic basis of a wide variety of traits including insecticide resistance (Battlay et. al 2016, Green unpublished, Schmidt unpublished).
Increasing the depth of study in a single population has the effect of decreasing the pool of both genetic and phenotypic variation in contrast to inter-population comparisons. This provides the rare opportunity to study QTLs and transcripts which may contribute very slightly to the overall phenotype. Usually, these genetic factors of small effects are undetectable due to the presence of genetic factors of very large effect. However, by deeply analysing one population there is a high chance that a particular QTL or transcript which is extremely influential has already been fixed and thus will not be a factor influencing the trait within that population. Furthermore, by considering smaller changes in phenotypic variation across a continuous gradient, as is present in most populations, more information on the loci of small effects can be gathered. This provides for a better opportunity to study the polygenic basis of imidacloprid resistance and thus give a better understanding of insecticide biology.

In this chapter, the genetics of imidacloprid response is described using the DGRP population and the Wiggle Index bioassay (Chapter II). Many QTLs and transcripts were associated with imidacloprid resistance via the GWAS and TWAS, implicating several genes involved in CNS development and function as well as $Cyp6g1$. Quantification of imidacloprid and its metabolites via HPLC-MS also showed differences between resistant and susceptible subsets of the DGRP, indicating that differences in overall metabolism make a significant contribution to the observed phenotypic differences. Furthermore, the deletion of $Cyp6g1$ from laboratory strains showed no significant differences in any imidacloprid phenotype, while the same deletion caused large decreases in resistance and metabolism when removed from a resistant DGRP line. This allowed for the direct measurement of the contribution a single haplotype of $Cyp6g1$ in this DGRP line. These data suggest that imidacloprid resistance within the DGRP is likely driven by both neurological and metabolic factors.

3.2) Materials and Methods

3.2.1 Fly Genotypes

All lines used in this study were ordered from the Bloomington Drosophila Stock Center (BDSC, Bloomington, Indiana) or generated in this study (Table 3.1). 178 of the 201 living DGRP genotypes were considered for phenotypic analysis. A subset of 9 of the 12 most imidacloprid resistant and susceptible DGRP lines (highlighted in Table 3.1) were chosen to perform imidacloprid metabolite analysis. The Actin-Cas9 genotype (Bloomington #54590) and a genotype containing an AttB landing site on chromosome 2 (Bloomington #25709) were used for used for gene knockouts in three genetic backgrounds: Canton-S (Bloomington #64394), Wxac (Actin-Cas9
with the X chromosome substituted from w\textsuperscript{1118} genotype) and a DGRP genotype, RAL\_517 (Bloomington #25197).

<table>
<thead>
<tr>
<th>Bloomington Stock Number</th>
<th>Genotype</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>25709</td>
<td>25709v</td>
<td>AttB landing site line</td>
</tr>
<tr>
<td>54590</td>
<td>Actin-Cas9</td>
<td>Expresses Cas9 under actin promoter</td>
</tr>
<tr>
<td>NA</td>
<td>Wxac</td>
<td>Actin-Cas9 with w\textsuperscript{1118} X chromosome</td>
</tr>
<tr>
<td>64394</td>
<td>Canton-S</td>
<td>Common laboratory strain</td>
</tr>
<tr>
<td>NA</td>
<td>RAL_517-Cyp6g1KO</td>
<td>RAL_517 with both copies of Cyp6g1 removed</td>
</tr>
<tr>
<td>NA</td>
<td>Wxac-Cyp6g1KO</td>
<td>Wxac with Cyp6g1 removed</td>
</tr>
<tr>
<td>NA</td>
<td>Canton-S-Cyp6g1KO</td>
<td>Canton-S with Cyp6g1 removed</td>
</tr>
<tr>
<td>28134</td>
<td>RAL_83</td>
<td>Susceptible DGRP Line</td>
</tr>
<tr>
<td>28274</td>
<td>RAL_85</td>
<td>Resistant DGRP Line</td>
</tr>
<tr>
<td>28164</td>
<td>RAL_280</td>
<td>Susceptible DGRP Line</td>
</tr>
<tr>
<td>28165</td>
<td>RAL_287</td>
<td>Resistant DGRP Line</td>
</tr>
<tr>
<td>55023</td>
<td>RAL_397</td>
<td>Resistant DGRP Line</td>
</tr>
<tr>
<td>28278</td>
<td>RAL_409</td>
<td>Resistant DGRP Line</td>
</tr>
<tr>
<td>25195</td>
<td>RAL_486</td>
<td>Susceptible DGRP Line</td>
</tr>
<tr>
<td>28206</td>
<td>RAL_509</td>
<td>Susceptible DGRP Line</td>
</tr>
<tr>
<td>25197</td>
<td>RAL_517</td>
<td>Resistant DGRP Line</td>
</tr>
<tr>
<td>55025</td>
<td>RAL_528</td>
<td>Resistant DGRP Line</td>
</tr>
<tr>
<td>55027</td>
<td>RAL_559</td>
<td>Susceptible DGRP Line</td>
</tr>
<tr>
<td>55032</td>
<td>RAL_634</td>
<td>Resistant DGRP Line</td>
</tr>
<tr>
<td>28218</td>
<td>RAL_703</td>
<td>Resistant DGRP Line</td>
</tr>
<tr>
<td>25202</td>
<td>RAL_730</td>
<td>Resistant DGRP Line</td>
</tr>
<tr>
<td>28223</td>
<td>RAL_738</td>
<td>Susceptible DGRP Line</td>
</tr>
<tr>
<td>28229</td>
<td>RAL_776</td>
<td>Susceptible DGRP Line</td>
</tr>
<tr>
<td>28247</td>
<td>RAL_843</td>
<td>Susceptible DGRP Line</td>
</tr>
<tr>
<td>28252</td>
<td>RAL_857</td>
<td>Susceptible DGRP Line</td>
</tr>
</tbody>
</table>

**Table 3.1- Genotype List:**

A list of all genotypes used in chapter III is provided along with their Bloomington stock numbers and short descriptions. There are no stock numbers for lines generated in this study or for the Batterham Lab stock Wxac.
3.2.2 The Wiggle Index Bioassay

The response of *D. melanogaster* larvae to imidacloprid was measured using the Wiggle Index motility assay (Denecke et. al 2015; Chapter II). 2-5 day old adults (60 females, 20 males) were added to a vial sprinkled with dried yeast and were incubated for one day at 25° for oviposition. The flies were then cleared, and third instar larvae were recovered via extraction with 20% sucrose in distilled water. Larvae of each genotype were picked, 25 per well, into a NUNC cell culture treated 24 well plate (Thermo-Scientific) preloaded with 200µL 5% w/v sucrose (AR Grade; Chem Supply) in distilled H2O. To dose the larvae, 50µL of 5x imidacloprid (200 gL⁻¹ Confidor®; Bayer Crop Science) stock solution was added to each well and removed after mixing to return the total volume to 200µL. The final concentrations for each well were 25 and 100ppm imidacloprid. Larvae were filmed at two time points, 0 (before the addition of insecticides) and at 1 hour. Each film lasted for 30 seconds, and the Wiggle Index ImageJ script was used to quantify the total motion in each video. This yielded WI values estimating the total motility of larvae within one well at one time point. Comparing WI values before and after the addition of insecticide yielded relative movement ratio values (RMR values) which were averaged to estimate the acute imidacloprid response of the genotype.

3.2.3 Analysis of the Wiggle Index and Candidate Gene Selection

Prior to any genetic analyses, the Wiggle Index data observed in the DGRP was analysed in several ways. Broad sense heritability $H^2$ was measured by comparing variance within genotypes to variance between genotypes. Data were also compared to other insecticide phenotypes elicited from the DGRP, by linear regression. Each dose of imidacloprid used in the Wiggle Index presented here (25 and 100ppm) were correlated with a Chlorantraniliprole 1st instar toxicology assay (Green unpublished), DDT adult survival (Schmidt unpublished), Imidacloprid adult survival (Battlay unpublished), and Azinphos-methyl 1st instar toxicology assay (Battlay et. al 2016). The association of initial motility values (at 0 minutes) with final RMR value was also tested in order to test the any confounding effects of starting motility on imidacloprid response. The correlation between 25 and 100ppm RMR values was tested to observe the relationship between the two phenotypes. All these associations were tested by assessing the fit to a linear model using Pearson's correlation test.

In order to implicate QTLs in imidacloprid resistance, phenotypic data from the Wiggle Index was fed into the Mackay lab DGRP GWAS pipeline (http://dgrp2.gnets.ncsu.edu/, Huang et. al 2014), which tested the association of all annotated SNPs in the DGRP appearing in at least 4 DGRP genotypes. Each association was tested by assessing the fit of a linear model using binary
genetic variants to predict the phenotype. In this case, the phenotype was the mean RMR value from the Wiggle Index. Two separate GWASs were performed, one at 25ppm and one at 100ppm.

A transcriptome wide association study (TWAS) was also performed by associating the phenotypic Wiggle Index data with transcriptomic data recently generated for male and female mated 3-5 day old adults in the DGRP (Huang et. al 2015). Analysis was performed using a modified pipeline from Battlay et. al (2016) and was accomplished by testing the fit of a linear model predicting the mean RMR value of a genotype with the amount of transcript measured in that genotype. This analysis was repeated for each transcript reported by Huang et. al (2015). Transcript levels for each gene were averaged between males and females as larvae were not sexed before use in the Wiggle Index. Furthermore, the expression quantitative trait loci (eQTLs) of each significant transcript (reported by Huang et. al 2015) were cross examined with significant SNPs from the GWAS. This was done in order to test if significantly associated genomic loci regulate the expression of any significantly associated transcripts.

Candidate genes from GWAS and TWAS were chosen based on their significance of association with either the 25 or 100ppm phenotype. Significant GWAS variants were assigned to their nearest annotated gene according to the software SnpEff (built into the Mackay pipeline). Any gene eliciting a p-value of less than $10^{-5}$ and any transcript with a p-value of less than $10^{-3}$ was considered as a candidate gene in the current study. Bonferroni and false discovery thresholds were applied to each output in order to bring the statistical significance of such findings into focus. Candidate genes were analysed by testing enrichment in third instar larvae tissues using the modEncode transcriptome datasets. Any gene with an expression level 2 fold higher than the average for all third instar tissues combined was considered enriched.

3.2.3 The Deletion of Cyp6g1

The P450 gene Cyp6g1 was deleted in 3 distinct genetic backgrounds by crossing a Cyp6g1-sgRNA genotype generated in this study with a Cas9 expressing genotype reported previously (Port et. al 2014). Cyp6g1-sgRNA plasmids were made by first cutting the PCFD4 plasmid (Addgene #49411) with the restriction enzyme Bsb1. A separate fragment was generated by amplifying a portion of the PCFD4 plasmid with the Cyp6g1_PCFD4 primer set (Appendix 1), introducing Cyp6g1 sgRNAs into the PCR product (60 degree annealing, 1 minute extension). The cut plasmid and PCR product were then reassembled using the Gibson assembly kit (New England Biolabs) to make a circular PCFD4 plasmid with two Cyp6g1 sgRNAs under the control of two U6 promoters. Verification of this modification was accomplished by sequencing the plasmid using the PCFD4_seq primer (Appendix 1). This plasmid was then injected into a genotype expressing phi-31
integrase and which contained an attB landing site (Bloomington # 25709), both of which facilitated the integration of PCFD4 into the germline. Transgenic flies were identified by scoring the visible marker vermilion eyes which was restored to wild type upon successful PCFD4 integration.

Of the three genotypes chosen for Cyp6g1 knockout, one was a genotype from the DGRP and two were laboratory strains. The DGRP genotype RAL_517 was chosen due to its high level of imidacloprid resistance and because it carries the BA haplotype (two copies of Cyp6g1, each with an Accord insertion, with one Accord insertion containing an HMS-Beagle insertion) at the Cyp6g1 locus. The two additional genotypes were Canton-S and Wxac, which are both laboratory strains naive to insecticide selective pressures and carry an M haplotype (one copy of Cyp6g1, no Accord insertion) at the Cyp6g1 locus. Each deletion was created by first crossing both Cas9 and sgRNA expressing chromosomes into one of the three genotypes. Subsequent backcrossing using balancer chromosomes ensured that the only difference between background and control was the Cyp6g1 locus (Fig 3.3).

3.2.5 Quantification of Imidacloprid and its Metabolites

Levels of imidacloprid and its metabolites were quantified by Roberto Fusetto using High Pressure Liquid Chromatography coupled with Mass Spectrometry (HPLC-MS) following a recently described method (Hoi et al., 2014). Briefly, quadruplicates of 200 third instar larvae from a given genotype were placed into 200µL of 5% analytical reagent sucrose and were exposed to a 50:50 mix of 12C6:13C6 imidacloprid at a final concentration of 25ppm for 1 hour. These chemicals were of 99% and >97% total purity and were purchased from AK Scientific (Union City, CA, USA) and IsoSciences (King of Prussia, PA, USA) respectively. Larvae were then recovered from the solution and washed 3 times with 3mL of dH2O to remove imidacloprid from the cuticle. Larvae and media were collected separately and the metabolites produced and retained within the larvae and excreted into the exposure media were identified and quantified using High Pressure Liquid Chromatography (HPLC) coupled with an Agilent 6520 Q-TOF Mass Spectrometer (Agilent Technologies, Inc., Santa Clara, CA, USA). These measurements were taken in order to compare 9 of the most resistant and susceptible DGRP genotypes as well as RAL_517-Cyp6g1KO and its background control. The levels of imidacloprid and metabolites produced were averaged among the four biological replicates. Because the HPLC-MS exposure conditions are nearly identical to those described in the Wiggle Index, direct comparisons can be made between data for these two assays. For the DGRP subset, the effect of Wiggle Index RMR value on metabolite levels was assessed using linear regression.
Figure 3.3- RAL\_517-Cyp6g1KO Crossing scheme:

The crossing scheme that created the RAL\_517-Cyp6g1KO genotype is shown. The 2nd chromosome from RAL\_517, where Cyp6g1 is to be deleted, is crossed into a background expressing Cas9 and sgRNAs targeting Cyp6g1. The presence of these two factors creates the deletion which is then backcrossed into the RAL\_517 background.
3.3) Results

3.3.1 Non-Genetic Analyses on Imidacloprid Resistance in the DGRP

Screening of 171 of the 205 DGRP genotypes using the Wiggle Index at doses of 25 and 100 ppm imidacloprid yielded significant variation between individual genotypes (Fig 3.4). The mean RMR values of the population were correlated between the two doses (Adjusted $R^2 = .18$; p-value $< 3.6 \times 10^{-9}$, Figure 3.5A), but not between initial motility value and final RMR value (Adjusted $R^2 = .001$; p-value = .06, Figure 3.5B). This suggests that the imidacloprid response was independent of larval motility measured in the absence of imidacloprid. The 25 ppm dose produced a slightly left skewed distribution of RMR values and had a disproportionate amount of genotypes (92.4%) with RMR values above 0.5. The presence of a few very susceptible genotypes breaks the continuity of the distribution at this dose. The 100 ppm dose produced a more continuous distribution of RMR values with 52.6% of genotypes with RMR values above .5. Broad sense heritability was estimated $H^2_{25ppm} = .628$, $H^2_{100ppm} = .699$.

Imidacloprid 25 and 100 ppm Wiggle Index phenotypes for each of the DGRP lines tested were separately compared with the response phenotypes for other insecticides, previously measured in those genotypes in other studies. No significant correlations were observed with the exception of DDT adult survival and 25 ppm imidacloprid wiggle response (p-value = .03; Table 3.2). The azinphos-methyl phenotype also showed some degree of correlation with the 25 ppm dose, but it failed to meet the $\alpha = .05$ significance threshold (p-value = .07).

3.3.2 A GWAS Implicates Neuronal Genes in Imidacloprid Response

Candidate genes for involvement in acute imidacloprid response were identified by correlating phenotypic differences in the DGRP with the annotated genomic variants (usually in the form of SNPs) in the same population. The distribution of the significance of genomic variants matched closely the predicted distribution assuming normality. This suggests that the p-values measuring the strength of associations were not artificially inflated (Fig 3.6). Generation of Manhattan plots revealed a scattered distribution of significant variants across all chromosomes (Fig 3.7). At each dose there was a group of variants which showed some degree of linkage disequilibrium (LD). While the 100 ppm LD block was in an intergenic region on chromosome 2, the 25 ppm GWAS contained several genes including CG17544 of which seems to function as an oxidase and *fondue* which has been implicated in muscle attachment (Faust et al., 2012; Green et al., 2016).
Figure 3.4- Phenotypic Spread of the DGRP:
Imidacloprid response in the DGRP was assessed using the Wiggle Index at 2 doses A) 25ppm and B) 100ppm. At each dose there was a spectrum of phenotypic responses ranging from susceptible (blue) to resistant (red). Error bars represent 95% confidence intervals.

Figure 3.5- Correlation of Wiggle Index Response Between RMR Values:
A) For each genotype the 25 and 100ppm mean RMR values were plotted and a linear regression line was fit, showing significant correlation. B) Raw Wiggle Index values at time 0 (x axis) were compared to RMR values after 60 minutes (y axis), suggesting no correlation. All models were fit with 95% confidence intervals.
Figure 3.6- QQ Plots:
Quantile Quantile (QQ) plots were made for the GWAS using A) 25ppm and B) 100ppm RMR values. Each plot shows the expected versus the observed distribution of p-values. The red line indicates where the observed values should fall if they were to match the expected values exactly.

Figure 3.7- Manhattan Plots:
Manhattan plots were made for the GWAS using A) 25ppm and B) 100ppm RMR values. Each plot shows the significance of a P-value on the Y axis and the position of the genetic variant on the X axis. Genome wide significance thresholds are shown at $10^{-5}$ and $2.65 \times 10^{-7}$ and any variants below the latter are highlighted in red.
<table>
<thead>
<tr>
<th>Treatment</th>
<th>Wiggle_25</th>
<th>Wiggle_100</th>
</tr>
</thead>
<tbody>
<tr>
<td>Azinphos-methyl .25ppm</td>
<td>0.13</td>
<td>0.2</td>
</tr>
<tr>
<td>Azinphos-methyl .5ppm</td>
<td>0.05</td>
<td>0.28</td>
</tr>
<tr>
<td>Azinphos-methyl .1ppm</td>
<td>0.07</td>
<td>0.27</td>
</tr>
<tr>
<td>Azinphos-methyl 2ppm</td>
<td>0.25</td>
<td>0.25</td>
</tr>
<tr>
<td>Azinphos-methyl .LD50</td>
<td>0.07</td>
<td>0.31</td>
</tr>
<tr>
<td>Chlorantraniliprole Dose 1</td>
<td>0.5</td>
<td>0.51</td>
</tr>
<tr>
<td>Chlorantraniliprole Dose 2</td>
<td>0.45</td>
<td>0.3</td>
</tr>
<tr>
<td>Chlorantraniliprole Dose 3</td>
<td>0.25</td>
<td>0.7</td>
</tr>
<tr>
<td>Chlorantraniliprole Dose 4</td>
<td>0.39</td>
<td>0.92</td>
</tr>
<tr>
<td>Chlorantraniliprole Dose 5</td>
<td>0.31</td>
<td>0.85</td>
</tr>
<tr>
<td>Chlorantraniliprole Dose 6</td>
<td>0.07</td>
<td>0.26</td>
</tr>
<tr>
<td>DDT Adult</td>
<td><strong>0.04</strong></td>
<td>0.18</td>
</tr>
<tr>
<td>Imidacloprid Adult</td>
<td>0.92</td>
<td>0.65</td>
</tr>
</tbody>
</table>

*Table 3.2- Insecticide Response Correlations:*

A table showing the significance (p-value) of correlation between each phenotype presented in the current thesis (25 and 100ppm Wiggle Index Response to Imidacloprid) and a panel of different insecticide response phenotypes gathered from the DGRP. Pearson’s correlation test was used for significance testing. Significant correlations are highlighted in yellow. Data for correlations was obtained from Green (unpublished), Schmidt (unpublished) and Battlay et. al (2016).
Annotation of significant (P≤10^{-5}) genomic variants with their nearest gene implicated a high proportion of candidates having roles in the development and function of the CNS (Table 3.3). 52.6% (10/19) of the genes showed at least two fold enrichment in the third instar larvae CNS compared to the average expression across all tissues at the same life stage. This number is only 19.6% when all genes in the genome are considered. A high proportion of the genes appear to have annotated functions in the CNS, with 7 out of the 15 candidates having at least one annotated “Biological process” GO terms relating to the CNS (e.g. axon guidance, olfactory detection etc.). Several of the remaining genes appear to play developmental roles (e.g. Notch ligand, cytoskeletal organization etc.) and are upregulated in the nervous system (Table 3.3).

Notable among this list are several variants located near genes showing very low p-values (high significance) and having well annotated functions. A SNP near the *Sickie* gene was the only variant with a p-value below the Bonferroni threshold (p=.04) at 25ppm and was significantly associated at both 25 and 100ppm (Table 3.3). This gene is thought to act in the Cofilin-slingshot pathway in order to regulate actin polymerization and ultimately axon guidance in the CNS (Abe et al., 2015). Also notable is the gene *Tweek*, which is a key player in the vesicle recycling necessary for synaptic transmission (Verstreken et al., 2009). Furthermore, the genes *Kirre* and *beat-Vc* both have predicted roles in cell adhesion via an immunoglobulin like fold, a process which underlies axon guidance (Pipes et al., 2001; Strünkelnberg et al., 2001). These genes are far from an exhaustive list, but rather highlight some of the processes which may underpin imidacloprid response in the CNS.

### 3.3.3) A TWAS Implicates Cyp6g1 and Cyp6g2 in Imidacloprid Response

The transcriptome datasets for 185 DGRP genotypes (Huang et. al 2015) were also used to associate the expression of specific genes with the acute-sub-lethal imidacloprid response phenotype. Unlike the GWAS candidate list, the TWAS candidate list does not appear to be enriched in any particular third instar tissue (Table 3.4). Furthermore, far less of a pattern emerges with regard to any conserved processes or trends that run through the list. Manual analysis of the individual genes, however, did reveal several noteworthy candidates. Two P450s *Cyp6g1* and *Cyp6g2* have been previously implicated in insecticide resistance previously (Fig 3.8). Furthermore, the second most significantly associated gene in the 25ppm TWAS was *CG30345* a transporter in the Major Facilitator Superfamily (MFS), which is highly upregulated in the third instar fat body. Also present are *Jon65Aiii*, a predicted serine protease highly upregulated in the L3 digestive system, and *nbs* which has previously been implicated in caffeine response (Li et al., 2013; Ross et al., 2003). There was no evidence that any variants from the GWAS influenced the expression of
any significantly associated transcripts as no eQTL for any transcript was present among the significantly associated GWAS variants.

Table 3.3- GWAS candidates:

A table of all significant data from the GWAS is presented. This table includes the gene name, ID of the SNP found in the association, the SNPs minor allele frequency and the significance of the association with either 25 or 100ppm imidacloprid response. Enrichment ratios compare the genes predicted expression in the CNS to the average expression in all third instar tissues.

<table>
<thead>
<tr>
<th>Gene Name</th>
<th>Variant ID</th>
<th>Minor Allele Frequency</th>
<th>P-value</th>
<th>CNS Enrichment Ratio</th>
<th>GWAS Dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>CG14431</td>
<td>X 6819592 SNP</td>
<td>0.07059</td>
<td>4.18E-07</td>
<td>10.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>Msp-300</td>
<td>2L 5105071 SNP</td>
<td>0.1824</td>
<td>1.86E-06</td>
<td>1.19</td>
<td>100ppm</td>
</tr>
<tr>
<td>CG33232</td>
<td>3L 2495753 SNP</td>
<td>0.2722</td>
<td>3.58E-06</td>
<td>1.25</td>
<td>100ppm</td>
</tr>
<tr>
<td>wry</td>
<td>2L 1836932 SNP</td>
<td>0.1479</td>
<td>3.80E-06</td>
<td>23.75</td>
<td>100ppm</td>
</tr>
<tr>
<td>cenG1A</td>
<td>2L 13896989 SNP</td>
<td>0.05848</td>
<td>3.88E-06</td>
<td>3.72</td>
<td>100ppm</td>
</tr>
<tr>
<td>Intergenic</td>
<td>2R 18329255 SNP</td>
<td>0.2645</td>
<td>4.75E-06</td>
<td>NA</td>
<td>100ppm</td>
</tr>
<tr>
<td>Intergenic</td>
<td>2R 14899318 SNP</td>
<td>0.2321</td>
<td>5.11E-06</td>
<td>NA</td>
<td>100ppm</td>
</tr>
<tr>
<td>Intergenic</td>
<td>2R 18329249 SNP</td>
<td>0.25</td>
<td>5.31E-06</td>
<td>NA</td>
<td>100ppm</td>
</tr>
<tr>
<td>lilli</td>
<td>2L 2925754 SNP</td>
<td>0.08824</td>
<td>5.43E-06</td>
<td>1.86</td>
<td>100ppm</td>
</tr>
<tr>
<td>dpr10</td>
<td>3L 10159678 SNP</td>
<td>0.1856</td>
<td>5.47E-06</td>
<td>15.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>beat-Vc</td>
<td>3R 8576914 SNP</td>
<td>0.4969</td>
<td>6.07E-06</td>
<td>25.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>Intergenic</td>
<td>2R 18329251 SNP</td>
<td>0.2532</td>
<td>6.32E-06</td>
<td>NA</td>
<td>100ppm</td>
</tr>
<tr>
<td>Intergenic</td>
<td>2R 12858731 SNP</td>
<td>0.06667</td>
<td>6.32E-06</td>
<td>NA</td>
<td>100ppm</td>
</tr>
<tr>
<td>wry</td>
<td>2L 1836962 SNP</td>
<td>0.1479</td>
<td>6.39E-06</td>
<td>23.75</td>
<td>100ppm</td>
</tr>
<tr>
<td>Intergenic</td>
<td>2L 14249499 SNP</td>
<td>0.1938</td>
<td>7.09E-06</td>
<td>NA</td>
<td>100ppm</td>
</tr>
<tr>
<td>sick</td>
<td>2L 19928580 SNP</td>
<td>0.1242</td>
<td>7.99E-06</td>
<td>5.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>sick</td>
<td>2L 19928580 SNP</td>
<td>0.122</td>
<td>2.18E-08</td>
<td>5.00</td>
<td>25ppm</td>
</tr>
<tr>
<td>Intergenic</td>
<td>2L 19375473 DEL</td>
<td>0.0578</td>
<td>6.68E-08</td>
<td>NA</td>
<td>25ppm</td>
</tr>
<tr>
<td>Intergenic</td>
<td>2L 19376418 SNP</td>
<td>0.07602</td>
<td>6.94E-07</td>
<td>NA</td>
<td>25ppm</td>
</tr>
<tr>
<td>tweek</td>
<td>2L 16917294 SNP</td>
<td>0.09859</td>
<td>1.37E-06</td>
<td>2.22</td>
<td>25ppm</td>
</tr>
<tr>
<td>Intergenic</td>
<td>3L 22083995 DEL</td>
<td>0.05952</td>
<td>2.86E-06</td>
<td>NA</td>
<td>25ppm</td>
</tr>
<tr>
<td>Intergenic</td>
<td>2L 3227580 SNP</td>
<td>0.05917</td>
<td>3.35E-06</td>
<td>NA</td>
<td>25ppm</td>
</tr>
<tr>
<td>CG34356</td>
<td>3L 9910712 DEL</td>
<td>0.06433</td>
<td>4.16E-06</td>
<td>25.00</td>
<td>25ppm</td>
</tr>
<tr>
<td>Intergenic</td>
<td>2R 10886753 SNP</td>
<td>0.06024</td>
<td>4.29E-06</td>
<td>NA</td>
<td>25ppm</td>
</tr>
<tr>
<td>CG30151</td>
<td>2R 16416362 SNP</td>
<td>0.05161</td>
<td>4.48E-06</td>
<td>0.00</td>
<td>25ppm</td>
</tr>
<tr>
<td>Hs3st-A</td>
<td>2R 14614239 SNP</td>
<td>0.08721</td>
<td>5.32E-06</td>
<td>17.50</td>
<td>25ppm</td>
</tr>
<tr>
<td>CG42272</td>
<td>3L 5797410 SNP</td>
<td>0.1314</td>
<td>5.39E-06</td>
<td>0.00</td>
<td>25ppm</td>
</tr>
<tr>
<td>CG17350</td>
<td>2L 19380943 SNP</td>
<td>0.1091</td>
<td>8.22E-06</td>
<td>0.00</td>
<td>25ppm</td>
</tr>
<tr>
<td>CG4593</td>
<td>X 6908817 DEL</td>
<td>0.106</td>
<td>8.27E-06</td>
<td>1.30</td>
<td>25ppm</td>
</tr>
<tr>
<td>kire</td>
<td>X 2949168 SNP</td>
<td>0.3879</td>
<td>8.32E-06</td>
<td>5.00</td>
<td>25ppm</td>
</tr>
<tr>
<td>CG17544</td>
<td>2L 19391228 SNP</td>
<td>0.05202</td>
<td>8.65E-06</td>
<td>1.82</td>
<td>25ppm</td>
</tr>
<tr>
<td>Acsl</td>
<td>2R 4562894 SNP</td>
<td>0.05172</td>
<td>8.72E-06</td>
<td>1.73</td>
<td>25ppm</td>
</tr>
<tr>
<td>CG17544</td>
<td>2L 19391829 SNP</td>
<td>0.05233</td>
<td>9.45E-06</td>
<td>1.82</td>
<td>25ppm</td>
</tr>
</tbody>
</table>
Figure 3.8- Cyp6g1 in the DGRP:
The relationship between Cyp6g1 expression and imidacloprid response is explored. A linear model is presented predicting mean RMR value with adult Cyp6g1 expression for each genotype at A) 25ppm and B) 100ppm. The 95% confidence interval for each model is displayed. Each point is represented by a letter which corresponds to its haplotype at the Cyp6g1 locus.
<table>
<thead>
<tr>
<th>Gene Name</th>
<th>P-value</th>
<th>False Discovery Rate</th>
<th>Dose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cyp6g1</td>
<td>2.62E-05</td>
<td>0.48</td>
<td>100ppm</td>
</tr>
<tr>
<td>CG30345</td>
<td>7.96E-05</td>
<td>0.72</td>
<td>100ppm</td>
</tr>
<tr>
<td>CG3876</td>
<td>2.24E-04</td>
<td>1.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>CG13026</td>
<td>2.36E-04</td>
<td>1.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>CG13088</td>
<td>3.04E-04</td>
<td>1.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>CG18343</td>
<td>3.43E-04</td>
<td>1.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>RagA-B</td>
<td>4.66E-04</td>
<td>1.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>CG1764</td>
<td>5.02E-04</td>
<td>1.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>nbs</td>
<td>6.02E-04</td>
<td>1.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>Cyp6g2</td>
<td>8.47E-04</td>
<td>1.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>ORY</td>
<td>9.22E-04</td>
<td>1.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>CG31600</td>
<td>9.72E-04</td>
<td>1.00</td>
<td>100ppm</td>
</tr>
<tr>
<td>Cyp6g1</td>
<td>5.02E-06</td>
<td>0.09</td>
<td>25ppm</td>
</tr>
<tr>
<td>Jon65Aiii</td>
<td>1.49E-04</td>
<td>1.00</td>
<td>25ppm</td>
</tr>
<tr>
<td>CG11504</td>
<td>2.71E-04</td>
<td>1.00</td>
<td>25ppm</td>
</tr>
<tr>
<td>T3dh</td>
<td>2.96E-04</td>
<td>1.00</td>
<td>25ppm</td>
</tr>
<tr>
<td>CG13151</td>
<td>6.09E-04</td>
<td>1.00</td>
<td>25ppm</td>
</tr>
<tr>
<td>Atf3</td>
<td>8.71E-04</td>
<td>1.00</td>
<td>25ppm</td>
</tr>
<tr>
<td>snoRNA:Or-aca3</td>
<td>9.92E-04</td>
<td>1.00</td>
<td>25ppm</td>
</tr>
</tbody>
</table>

**Table 3.4- TWAS candidates:**

A table of all significant data from the TWAS is presented. This includes the transcript name, the significance of correlation with imidacloprid response. A false discovery rate correction is applied to p-values in order to correct for multiple testing.
3.3.4) **Imidacloprid Metabolism is Elevated in Resistant DGRP Lines**

To estimate the contribution of insecticide metabolism to the observed differences in Wiggle Index response, the amount of imidacloprid and its metabolites recovered from both larvae and the exposure media was quantified for resistant and susceptible subsets of the DGRP. The quantity of imidacloprid in larval bodies was significantly correlated with RMR value at 25ppm among genotypes (Fig 3.9A; Table 3.5). Additionally, the quantities of IMI-5-OH and IMI-Olefin recovered from the media were significantly associated with RMR value (Fig 3.9E,F; Table 3.5). However, RMR value did not significantly predict the level of either metabolite in the body (Fig 3.9B,C; Table 3.5). These data suggest that imidacloprid metabolism is higher in imidacloprid resistant genotypes, and that these metabolites are preferentially excreted from the body.

3.3.5) **The Removal of Cyp6g1 Produces Background Specific Effects**

In order to investigate the contribution of the top transcriptomic candidate gene, Cyp6g1, CRISPR-Cas9 was employed to remove this gene from two susceptible lab genotypes and from one of the most imidacloprid resistant DGRP genotypes. Crossing the Cyp6g1-sgRNA genotype to a Cas9 genotype resulted in nearly identical deletions in each background which removed the majority of the Cyp6g1 open reading frame (Fig 3.10). The validity of the deletion was confirmed by sequencing across the deletion using the Cyp6g1KO_sequence primer set (Appendix 1). The knockout sequence introduced a stop codon shortly after the Cyp6g1 transcriptional start site and was assumed to be a null allele. While Wxac and Canton-S have only one copy of Cyp6g1, RAL_517 has two. This lead to concerns that only one copy of the gene would be removed. However, failure to amplify a full length Cyp6g1 gene in RAL_517-Cyp6g1KO combined with the ability to do so in heterozygotes indicated that each of the two predicted Cyp6g1 copies in this background had been successfully removed.

The imidacloprid response of these knockouts was compared to their matched controls in Wiggle Index response assay mimicking the conditions used in the initial GWAS screen. RAL_517 showed a moderate response at the higher dose of 100ppm imidacloprid after 1 hour, while the knockout RAL_517-Cyp6g1KO showed a significantly greater response (Fig 11A). The RAL_517 control and knockout also showed significant differences at the lower GWAS screening dose of 25ppm (Fig 11B). These findings were not replicated when testing knockouts in the Canton-S and wxac backgrounds, where no significant differences exist between knockout and control using a discriminatory dose of 5ppm (Fig 3.12A,B). This lower dose was used because 25ppm elicits too a strong response in laboratory strains, making the detection of susceptibility more unlikely.
Figure 3.9- HPLC-MS in a DGRP Subset:
The amount of Imidacloprid (A,D) IMI-5-OH (B,E) and IMI-Olefin (C,F) recovered from larval bodies (A-C) or the media (D-F) is reported from a subset of the most susceptible and resistant DGRP lines. The more susceptible lines are in blue and the more resistant lines in red. No data is presented for imidacloprid in the media, due to the relative abundance of this molecule in the media, which makes detecting changes impossible. Error bars represent standard error of the mean.
Table 3.5- HPLC-MS Significance:
The significance of linear models predicting the amount of imidacloprid or metabolites in larvae or media predicted by RMR value for each genotype. Significant results are highlighted in red.

<table>
<thead>
<tr>
<th>Figure Position</th>
<th>HPLC Measurement</th>
<th>Predicting Variable</th>
<th>Model Significance</th>
<th>Adjusted R squared</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Imidacloprid Body</td>
<td>RMR Value</td>
<td>0.001061502</td>
<td>0.5717</td>
</tr>
<tr>
<td>B</td>
<td>5-OH Larvae</td>
<td>RMR Value</td>
<td>0.3350665</td>
<td>0.0006591</td>
</tr>
<tr>
<td>E</td>
<td>5-OH Media</td>
<td>RMR Value</td>
<td>0.001606329</td>
<td>0.5425</td>
</tr>
<tr>
<td>C</td>
<td>Olefin Larvae</td>
<td>RMR Value</td>
<td>0.1082343</td>
<td>0.134</td>
</tr>
<tr>
<td>F</td>
<td>Olefin Media</td>
<td>RMR Value</td>
<td>0.003965641</td>
<td>0.4722</td>
</tr>
</tbody>
</table>

Figure 3.10- Cyp6g1KO Schematic:
A representation of the Cyp6g1 deletion is shown that was approximately common to each of the three backgrounds. A) The Wild type Cyp6g1 gene is shown. B) The Cyp6g1 knockout allele is shown which truncates the gene and introduces a stop codon.
**Figure 3.11: RAL_517-Cyp6g1KO in the Wiggle Index:**

The effect of the removal of Cyp6g1 from RAL_517 on Wiggle Index Response is shown at A)25ppm and B)100ppm. Significant differences exist between RAL_517 (dark blue) and RAL_517-Cyp6g1KO (light blue) at each dose. Error bars represent 95% confidence intervals.
Figure 3.12: Laboratory Strain Cyp6g1KO Genotypes in the Wiggle Index:
The effect of the removal of Cyp6g1 from laboratory strains on the Wiggle Index Response is shown in A) Wxac and B) Canton-S. Darker colours represent wild type and lighter colours represent Cyp6g1 knockout flies. No significant differences were found in either genotype between wild type (dark) and knockout (light). 5ppm was used as a discriminatory dose for each genotype Error bars represent 95% confidence intervals.
HPLC-MS analysis of imidacloprid and metabolite levels showed a large reduction in imidacloprid metabolism in RAL_517-Cyp6g1KO larvae compared to RAL_517. Significantly less imidacloprid and more metabolites were found in both media and larval body samples of RAL_517 compared to RAL_517-Cyp6g1KO (Fig 3.13 A-C,E,F). This differs from the DGRP subset in that significant differences were found in larval bodies as well as media samples. The lesser amount of IMI and higher amounts of metabolites indicate that IMI metabolism is higher in the RAL_517 genotype compared to RAL_517-Cyp6g1KO.

3.4) Discussion

Insecticide resistance offers a unique opportunity to study evolution in response to a strong recent selection. Furthermore, model organisms such as *D. melanogaster* facilitate both the identification of the loci that contribute to a phenotype and the quantification of their contribution. Here the DGRP is used to dissect the genetics of imidacloprid resistance under assay conditions that permit both acute imidacloprid response and metabolism to be examined.

3.4.1) Candidate Genes

One of the most salient features of the set of candidates derived from the GWAS was the high proportion of genes showing enriched expression in the third instar CNS (Table 3.2). Upon closer examination many of these candidates have annotated roles in various CNS functions such as synaptic transmission and axon guidance. While it is not possible to propose a precise mechanistic link between these genes and imidacloprid resistance, less specific hypotheses can be suggested. Synaptic transmission may influence the response following the binding of an insecticide to its target site. Upon binding a nAChR subunit, imidacloprid switches the neuron to a constant “on” position, whereby it continually propagates its signal to other neurons until the organism dies. If the machinery governing synaptic transmission is altered, the “on” switch may be dampened, which may protect the insect.

Neurodevelopment and in particular axon guidance may also play a role. Firstly, a nervous system that has undergone alterations in neurodevelopment may present more or less target at the synaptic membrane, increasing or decreasing the levels of imidacloprid binding, respectively. Changes in developmental genes have also been shown to drastically alter transcriptional levels of paradigmatic resistance genes such as P450s (Dialynas et al., 2015). Although none of these GWAS candidates were found among the eQTLs thought to govern the expression of the TWAS candidates, it is possible that variation in nervous system development can engender transcriptional changes, which may influence imidacloprid resistance. This appears to be the case for chlorantraniliprole.
Figure 3.13: HPLC-MS with RAL_517-Cyp6g1KO:

The amount of Imidacloprid (A,D) IMI-5-OH (B,E) and IMI-Olefin (C,F) recovered from larval bodies (A-C) or the media (D-F) is reported in RAL_517 (red) and RAL_517-Cyp6g1KO (blue). No data is presented for imidacloprid in the media, due to the relative abundance of this molecule in the media, which makes detecting changes impossible. Error bars represent standard error of the mean.
resistance within the DGRP where modifications of muscle development appear to underpin the regulation of the Cnc gene, which intern influences P450 expression (Green unpublished).

The most statistically significant GWAS candidate was a SNP in an intron of the gene Sickie. This candidate was the only variant to be significantly associated in both the 25 and 100ppm GWAS and was unique in its association below the Bonferroni threshold. Sickie is orthologous to mammalian NAV2, which has been shown to regulate neuronal development (Marzinke et al., 2013). Although originally identified as a regulator of Relish and posited to have a role in innate immune response (hence the name Sickie, Foley & O’Farrell 2004), this gene has also been implicated in mushroom body development in D. melanogaster. Its expression is also highly enriched in the third instar larvae CNS (Table 3.3). Furthermore, genotypes carrying hypomorphic Sickie alleles showed axon growth defects were “unable to extend their axons to the lobe terminus” (Abe et al., 2015). Although the pathway in which Sickie acts has been thoroughly characterized (Fig 3.14), how this pathway relates to insecticide response can only be said to fit in with the general hypotheses advanced above.

From the list of 18 candidate transcripts extracted from the TWAS, there was no clear trend in terms of tissue enrichment as there was in the GWAS candidates, but nonetheless several individual candidates are interesting in their own right. One must be cautious in interpreting these results as the expression data was taken on sexed adults, rather than unsexed larvae. However, averaging the two sexes mitigates some bias, and larvae have been used previously in conjunction with a DGRP TWAS (Battlay et. al 2016). Additionally, the implication of Cyp6g1, a known imidacloprid metabolizing enzyme as the most significantly associated candidate provides an unbiased verification of the TWAS because a role for Cyp6g1 in imidacloprid metabolism has already been well established.

Another prominent candidate from the TWAS was an unnamed member of the MFS, a group of transporter proteins known to transport xenobiotics. While ABC transporters are far more commonly implicated in drug transport in animals, members of the MFS have been implicated in xenobiotic transport in Drosophila (Chahine et al., 2012). Furthermore, in fungi upregulation of MFSs are frequently the cause of drug resistance (Cannon et al., 2009). Also present appearing in the TWAS candidate list is Cyp6g2. This, P450 gene is normally expressed in the ring gland in non-metabolic tissues (Chung et al., 2009); it has not been characterized as having expression in the midgut, Malpighian tubules or fat body. Nonetheless, transgenic overexpression of this gene in these tissues has been shown to confer resistance to the neonicotinoid nitenpyram (Daborn et al 2007). Indeed, CYP6G2 has the capacity to metabolize imidacloprid and produces the same metabolites (Fusetto unpublished). The role of Cyp6g2 in the imidacloprid resistance observed in
the DGRP lines in the DGRP is somewhat ambiguous, and it is unknown if it Cyp6g2 contributes to imidacloprid metabolism or if it is simply co-regulated with Cyp6g1. However, recent work measuring the expression of Cyp6g2 in the metabolic tissues of AA haplotype DGRP genotypes suggests that Cyp6g2 is enriched in these tissues and may contribute to imidacloprid resistance (Martelli unpublished).

![Cellular Regulation Diagram](image)

Fig3.14: Sickie Regulates Axonal Growth:
The top candidate from the GWAS acts in the Cofilin (Tsr) dependent pathway to regulate axonal growth. Disruption of the Sickie gene prevents activation of Slingshot which prevents axonal growth. Adapted from Abe et. al (2014).
The candidate list generated by both the GWAS and TWAS lacked large numbers of statistically robust candidates. In the two GWAS, only 19 genes from a total of 33 unique variants were associated with imidacloprid response below the $p<10^{-5}$ threshold. Similarly, the TWAS was only able to associate 18 unique transcripts below the $p<10^{-3}$ threshold. It is important to note that while these p-values appear extremely low, they have not been corrected for multiple testing. False discovery and Bonferroni corrections were applied here in order to correct for multiple testing, and only one SNP near the candidate *Sickie* cleared either threshold. In other GWAS, the confidence of an association can be buttressed by LD peaks, groups of highly associated variants that are located in a continuous genomic locus and are inherited together. LD peaks can be useful as they support the involvement of a locus in a trait even if no individual genetic variant is associated below a statistical threshold. There were two such peaks in the current study, however, they were comprised of a small number of variants at a low significance level. Furthermore, in the region of each of these peaks there was only one gene, *fondue*, which had any annotated function. This was a common theme in that many of the candidate genes from the GWAS and TWAS had either no annotated functions or had not been tested experimentally.

The observations do not invalidate the findings of this study nor are they unique. Other GWAS on the DGRP have also failed to produce candidates below the Bonferroni threshold or clearly significant LD peaks (Ivanov et al., 2015; Montgomery et al., 2014; Swarup et al., 2013). The lack of prominent LD peaks could be attributed to low LD present in many regions of the *Drosophila* genome (Mackay et. al 2012). Furthermore, the distribution of the phenotypes across the DGRP also influences the significance of associated genetic variants. Skewed distributions with a small number of either very resistant or susceptible genotypes have tended to draw out more significant candidates compared to more even distributions (Battlay et. al 2016, Green unpublished). This is consistent with the Imidacloprid Wiggle dataset, where the more skewed 25ppm phenotype elicited more significant candidates than did the 100ppm, even though both datasets were relatively evenly distributed. While this prevented the association of a large number of significant candidates, it allows for the understanding the polygenic basis for imidacloprid resistance. Skewed distributions are more likely to positively implicate a single QTL or transcript of large effect, but they also prevent the detection of the many QTLs or transcripts which contribute in more subtle ways. The primary advantage then of having an even distribution of phenotypes is to better understand imidacloprid biology by implicating these genetic variants of small effect even if there is uncertainty regarding any one individual variant. Imidacloprid response is a complex phenotype likely determined by a host of different genes and processes.
3.4.2) Cyp6g1

Many genes which clearly have the potential to influence imidacloprid resistance, may not appear in the GWAS or TWAS because there is not sufficient variation within the DGRP at that locus. However, the fact that Cyp6g1 was the most associated transcript was unsurprising as there is significant variation in the amount of Cyp6g1 expressed among DGRP different genotypes (Battlay et. al 2016). The DGRP is far from unique in this respect. Previous population surveys show that the Cyp6g1 locus contains structural variation at varying frequencies depending on the population (Catania et. al 2004, Schmidt et. al 2010). Most DGRP genotypes carry an AA or BA haplotype, which have very similar expression levels well above that observed in the ancestral M haplotype. Only a handful of M haplotypes are present in the DGRP, and they express far less Cyp6g1. However, residual variance of Cyp6g1 expression levels within haplotypes, suggest that other factors may influence Cyp6g1 expression. This may also explain why no genetic variants around Cyp6g1 were significantly associated in the GWAS.

While Cyp6g1 had been previously shown to be involved in imidacloprid metabolism and in Wiggle response (Joussen et. al 2008; Fig 2.7), the loss of function of this allele has not previously been reported. The removal of Cyp6g1 in laboratory strains did not have any discernible effect, while the deletion in RAL_517 showed a significant difference in Wiggle response and metabolite levels (Fig 3.11; 3.12; 3.13). There are many differences between the two laboratory strains and the RAL_517 strain, but the most obvious source of the difference in knockout effect is at the Cyp6g1 locus itself. Wxac and Canton-S both have a low expression M haplotype while RAL_517 has a high expression BA haplotype. Because it is unlikely that Cyp6g1 is playing any essential role in how the insecticide kills the insect, as for instance the insecticide target site would, removing the gene in a background with low Cyp6g1 expression may have an undetectably small effect.

This is unsurprising in the case of Cyp6g1, but validation of candidate genes from forward genetic screens often rely on removal or knockdown of a gene in laboratory strains to validate a candidate’s contribution to a phenotype. This strategy is often employed in DGRP studies as a quick and cheap way to verify candidates (Garlapow et al., 2015; Montgomery et al., 2014). However, as the case of Cyp6g1 shows, it has the capacity to throw up false negatives, even when the gene is capable of playing a major role in the phenotype. For Cyp6g1, only overexpression of the gene reveals its role in imidacloprid resistance (Daborn et. al 2002, Chapter II).

The ability to remove Cyp6g1 from a genotype that formed a part of the initial forward genetic screen allowed for the direct measurement of a the BA haplotype of Cyp6g1 to the imidacloprid resistance. To the authors knowledge this is the first time that the contribution of a
genetic variant influencing insecticide resistance has been directly measured in a natural population. While a myriad of QTLs have been associated with imidacloprid resistance (Kalajdzic et. al 2012), measuring their effect has been limited to genetic manipulation in laboratory strains or by modelling such an effect statistically. In this case, the removal of Cyp6g1 from laboratory strains would have given a false negative. Furthermore, the linear model which predicted RMR value with Cyp6g1 expression was qualitatively accurate but quantitatively inaccurate, predicting a negative RMR value (an impossibility) in the absence of Cyp6g1 expression (Fig 3.8).

This model is obviously imperfect as it assumes linearity and and ignores other factors besides Cyp6g1 that are involved in imidacloprid resistance. As the statistician George E. Box said “All models are wrong, some models are useful” (Box, 1976). Other factors clearly contribute to imidacloprid resistance. Removal of the gene from RAL_517 reduced the resistance of the genotype, but the knockout genotype still had a resistance level which would have put it in the middle of the DGRP resistance distribution. This highlights the polygenic basis for imidacloprid resistance in the DGRP. It is interesting that Cyp6g1 only appears to play a role when expressed above a certain threshold level. It may be that Cyp6g1 has a higher affinity for endogenous substrates, and only when there is more than enough CYP6G1 to metabolize that substrate is CYP6G1 free to metabolize insecticides. Future studies will need to address these speculative ideas.

3.4.3) Imidacloprid Metabolism

The ability of imidacloprid metabolism to influence imidacloprid resistance was further supported by HPLC-MS data, used to measure the amount of imidacloprid and metabolites present in larval bodies and exposure media. Importantly, this HPLC data can be directly integrated with RMR data because nearly identical exposure conditions were used among the two assays. The high degree of negative correlation between RMR response and amount of imidacloprid recovered from the body suggests that some kind of drug metabolism as a likely cause (Table 3.4). The further correlation of RMR value with amount of the IMI-5-OH in the media points to a role for P450s as the underlying cause of the variation in imidacloprid response observed in the DGRP. Although IMI-Olefin is not known to be produced by a P450, the metabolite is produced from IMI-5-OH, so its levels are dependent upon Cyp6g1 expression (Fusetto unpublished). Hence, the correlation of RMR value with the amount of IMI-Olefin recovered from the media still supports the hypothesis that RMR value is at least partially dependent on P450 activity.

The HPLC-MS data presented here also posits a role for transportation in imidacloprid resistance within the DGRP. The correlation between RMR value and metabolite level observed in the media was not observed within larval bodies, and no significant correlation was found between
the RMR value of a genotype and the amount of either IMI-5OH or IMI-Olefin recovered from the body (Fig 3.9; Table 3.5). This stands in contrast to the knockout of Cyp6g1 in RAL_517 which showed less metabolites in both larval bodies and excreta, with differences being greater in larval bodies than they were in larval excreta (Fig 3.13). This suggests, that P450s such as Cyp6g1 are producing the metabolites but that there are other factors that determine how much of the toxins can get out of the body. While the genetic basis for this difference was not determined in the current study, several conjectures can be made as to the types of genes responsible for transport. ABC transporters from the ABCB clade have been shown to play roles in transporting other insecticides in ways that effect resistance (Chapter IV). Furthermore, a member of the MFS, such as the 2nd most associated transcript found in the TWAS, could be involved in transporting imidacloprid metabolites out of the body. These proteins may be playing a role at the blood brain barrier (BBB) as Cyp6g1 metabolites generated in the CNS are able to be excreted into the media (Fusetto unpublished).

In further evaluating the importance of the BBB, the consequences of IMI-5-OH and IMI-Olefin reaching the brain need to be considered. It is unknown with what affinity imidacloprid metabolites can bind nAChRs in Drosophila, although structural models predict the strong binding of both IMI-5-OH and IMI-Olefin at the interface of the extracellular domains of the nAChR subunits Dα1 and Dβ2 (Mitchell unpublished). Specific binding of these metabolites to nAChRs has been demonstrated in other insect species. This is especially true for IMI-Olefin which has a greater affinity for nAChRs than does imidacloprid (Nauen et al., 1998). IMI-5-OH is slightly less toxic than imidacloprid and IMI-Olefin is as toxic as imidacloprid in D. melanogaster (Fusetto unpublished). Therefore a genotype that lacks the ability to excrete these metabolites allowing them access to the brain is likely to be extremely susceptible, even if it displays a high level of imidacloprid metabolism.

This phenotype is evident in RAL_509, easily the most susceptible line among the DGRP lines tested. When DGRP genotypes were tested with other insecticides, RAL_509 is typically found towards the susceptible end of the distribution (Schmidt unpublished; Battlay et. al 2016; Green unpublished), but it has only proved to be the most susceptible line when tested with imidacloprid, either with the Wiggle Index or in mortality screens (Fig 3.4, Battlay unpublished). While being the most susceptible line, RAL509 had the highest levels of IMI-OH and IMI-Olefin in larval bodies. This genotype can produce these metabolites, but does not efficiently excrete them. Although the gene(s) responsible have not been identified, it would seem that this excretion defect is somewhat specific to imidacloprid. Identifying such a gene is likely impossible in this GWAS
because RAL_509 is the only genotype to display this phenotype (hyper-susceptibility despite high metabolism) and rare variants cannot be detected by GWAS.

Another insight gleaned from the HPLC-MS data is the contribution of *Cyp6g1* to the metabolism of imidacloprid compared to all other genes. While the removal of *Cyp6g1* from RAL_517 significantly reduced the production of IMI metabolites, it did not eliminate them. P450 metabolites are still produced in the absence of *Cyp6g1*, suggesting that other P450s metabolize imidacloprid in RAL_517. The most logical candidates of the superfamily would be the genes with the closest homology to *Cyp6g1*, which are usually part of the more unstable P450 clades discussed previously (Section 1.4.2.3). Transgenic expression of *Cyp6v1* and *Cyp6t1* conferred small amounts of resistance to imidacloprid (Giang unpublished), so it is feasible that they contribute to imidacloprid metabolism in RAL_517-Cyp6g1KO, albeit in relatively small ways. It is possible that the failure to detect these genes in the GWAS is due to their relatively homogeneous expression across DGRP genotypes. The more likely candidate is *Cyp6g2*, which as been previously shown to metabolize imidacloprid (Fusetto unpublished). Although this gene is normally expressed in tissues that is not metabolically important, the presence of the Accord element may drive the expression of this gene in a similar pattern to *Cyp6g1* (midgut, malpighian tubules, fat body). Further work is needed to confirm any role played within the DGRP by these candidates, but considering a background without *Cyp6g1* will likely be necessary to achieve the appropriate resolution.

### 3.4.4 Limitations and Future Directions

Despite the successes and novelties of this study there are several areas where the work could have been pushed further conceptually. The choice to restrict genetic manipulations to a relatively well understood gene, *Cyp6g1*, instead of any number of neuronal candidates that were implicated in the GWAS was influenced by the accessibility of the former. Knocking out the relatively small (2kb) *Cyp6g1* gene was not lethal, while there is evidence to suggest that the top genomic candidate *Sickie* (166kb) is either lethal or has severe fitness defects (Abe et al., 2015). Genes such as *Sickie* also interact in complex pathways, which makes it difficult to connect any observed phenotypic difference with a single biological mechanism. Xenobiotic metabolising P450s, by contrast, appear to act far more independently.

Furthermore, the measurement of *Cyp6g1* contribution could have been expanded to include different haplotypes. While the removal of *Cyp6g1* from RAL_517 measured the contribution of a BA haplotype of *Cyp6g1* compared to a null allele, it does not measure the difference among *Cyp6g1* haplotype present within the population. To address this problem CRISPR could have been employed to replace one haplotype of *Cyp6g1* with another. While the logistics of doing this in a
specific genetic background is not as straightforward as a deletion, it is possible and could represent the most accurate way of quantifying the contribution of alternative variants to a given phenotype.
Chapter IV: Using CRISPR-Cas9 to Examine the Role of the P-gp Homologue *Mdr65* in Resistance to Multiple Insecticides.
4.1) Introduction

The description of insecticide resistance usually centres around mechanisms involving target site modification or increasing the expression of DMEs capable of chemically transforming insecticides (Perry et al., 2011). However, the low cost and high sensitivity of new genetic techniques provides the opportunity to explore the effect of other mechanisms. The most common of these less well studied resistance mechanisms is drug transport. Changes in the structure or expression of individual transporter proteins can determine which tissues an insecticide is able to access and the rate at which the insecticide is excreted from the body. Recently, transporters have taken an increasingly prominent role in the literature (Porretta et al., 2016); the ABC transporter superfamily is the most frequent group of such proteins implicated in insecticide transport (Section 1.4.3). Of the 56 ABC transporters present in Drosophila only select transporters have been implicated in drug transport and their substrate specificities have not been tested systematically (Fig 4.1). The identification and characterization of insecticide transporters is crucial for a better understanding of insecticide resistance.

4.1.1) The Identification of ABC Transporters that act on Insecticides

The transcription of some genes encoding ABC transporters is induced by an insecticide exposure. Transporters induced by a given insecticide may be candidates for the transport of that insecticide, although such a feedback loop is asserted more often than demonstrated. Transcriptomes of unexposed and exposed insects are often compared to measure which ABCs have been upregulated. Such comparisons with Anopheles stephensi and Helicoverpa armigera found that several ABCs were upregulated following exposures to pyrethroids and various PSMs respectively (Bretschneider et al., 2016; Epis et al., 2014). Several ABCs were also present among the genes downstream of the xenobiotic response modulator Cnc, which can be induced by many forms of chemical stressors in insect species (Misra et. al 2011). However, although it is possible that the transporters induced by an insecticide may also use that insecticide as a substrate, this need not be the case. Evolution proceeds blindly, and the fact that a gene is upregulated in response to a stress does not guarantee a role for that gene in combating the same stress. The relationship between induction and function warrants further research research. In the context of this study, the observed induction of some ABCs transporter genes is merely taken as a line of evidence suggesting their a role in insecticide transport and resistance.
Figure 4.1: ABC Transporter Phylogeny in D. melanogaster:

A phylogeny showing the 56 ABC transporters present in humans and *D. melanogaster* is shown. Of the 8 families (A-H) only B, C and G (highlighted in red) have been linked to xenobiotic transport. I and II represent the alignments of the C or N terminus nucleotide binding domains respectively. Adapted from Dean & Allikmets (2001).
Further evidence for an ABCs involvement in insecticide resistance can be obtained by associating the levels of transporter expression with resistance. This approach is similar to the TWAS presented in Chapter III. Such studies have implicated the ABCB4 gene in pyrethroid resistance, finding it upregulated due to gene amplification in pyrethroid resistant *Aedes aegypti* populations (Bariami et al., 2012). Similar enrichments of ABC transporters were found in pyrethroid resistant bed bugs compared to susceptible ones (Zhu et al., 2013). Many other studies have also found associations between ABC transporter expression and insecticide resistance. However, association studies fail to causally link transporters with resistance phenotypes, and these studies often use small sample sizes (often under 10 total genotypes are used), meaning that results must be interpreted with caution.

The most conclusive evidence for the involvement of ABC transporters in resistance comes from direct genetic manipulations. Individual transporters can be knocked down using RNAi as was accomplished in ciliate *Tetrahymena thermophila*, which implicated the *Abc15* gene in DDT transport (Ning et al., 2015). Null alleles can also be used to study the effect of removing an ABC. This method was used to implicate several ABC transporters in influencing the permeability of the BBB to several dye molecules of *D. melanogaster* (Mayer et al., 2009). Disruption of the *D. melanogaster* ABCC subfamily transporter *Mrp1* also increased susceptibility to methyl mercury and (Prince et al., 2014). However, both RNAi and null alleles have their drawbacks in insects. RNAi has well known drawbacks including off target effects (Liberali et al., 2014) and the null alleles so far used in insects make use of pre-existing strains, which have uncharacterised mutations in heterogeneous backgrounds.

One additional way to implicate ABCs in insecticide resistance takes a physiological approach rather than a genetic one. Rather than identifying individual transporters which may influence resistance phenotypes, chemicals that inhibit many ABC transporters can be co-administered with insecticides. Synergism between inhibitors and insecticides indicates that ABC transporters play a role in resistance much as done with P450s and their inhibitors PBO (Feyereisen, 2015). In the case of ABCB and ABCC transporters, verapamil is the most commonly cited agent. Originally used as a calcium channel blocker, verapamil is also able to directly bind and inhibit certain ABC transporters (Cornwell et al., 1987; Falasca and Linton, 2012). While verapamil is far too toxic to humans to be used as a commercial synergist, it is often used in laboratory settings to test whether drug transport is significantly effecting resistance (Dermauw and Van Leeuwen, 2014).
4.1.2) P-glycoprotein in D. melanogaster

P-glycoprotein (P-gp) is the ABC most commonly implicated in insecticide and drug resistance, and insects contain between 1-4 paralogues of this gene (Section 1.4.3.4-5). Silencing of this gene with RNAi in Aedes aegypti larvae lead to an increase in temphos susceptibility (Figueira-Mansur et al., 2013). More conclusively, the purified P-gp protein from Helicoverpa armigera was able to transport insecticides in a reconstituted membrane (Aurade et al., 2010). The role of P-gp in insecticide transport has also been explored in D. melanogaster. Resistance to the insecticide α-amanitin was mapped to a region containing Mdr65, one of the 4 P-gp orthologues of present in D. melanogaster (Begun and Whitley, 2000). Selection experiments also associated regulation of the same gene with abamectin resistance and found a null allele which was more susceptible to abamectin (Luo et al., 2013). Furthermore, work done with the D. melanogaster 91R strain has revealed roles for both Mdr65 and another P-gp orthologue Mdr49 in DDT resistance (Gellatly et al., 2015; Strycharz et al., 2013). It appears that sequence variation may contribute to the substrate specificity of P-gp as only certain Mdr49 alleles were able to effectively confer resistance to DDT when overexpressed (Seong et al., 2016).

So far, few studies considering the role of ABCs or P-gp as insecticide transporters have considered the tissues in which they act. However, there are a number of tissues where insecticide transport could have a significant impact on resistance. One such tissue is the malpighian tubules, which perform a similar excretory function to the mammalian kidney. Transporters in this tissue have been implicated in regulating osmolarity and effluxing non-insecticidal xenobiotics out of the hemolymph in D. melanogaster (Chahine et al., 2012; O’Donnell, 2009). A role for transportation can also be posited in the D. melanogaster midgut. Besides being targets for the various forms of Bt toxin (Heckel, 2012), ABCs in the midgut could also play a role in importing insecticides into the hemolymph via the midgut. There is precedent for insecticide transport occurring in this tissue, but the overall effect of midgut transport has not been established (Lara, et. al 2015). Finally, transporters embedded in the blood brain barrier (BBB) can be critical in determining how much insecticide reaches the CNS, a tissue containing the target protein for the majority of insecticide. The BBB is made up of elongated glial cells forming tight junctions, which make the barrier difficult to cross for most molecules without active transport (Fig 4.2, Stork et al. 2008; Schirmeier & Klämbt 2015). Furthermore, mammalian transporters in this tissue have proven to be critical in determining the permeability of the CNS to xenobiotics (Saunders et al., 2016). This may explain the localization of a P-gp like protein to the Manduca sexta BBB, which may allow the organism to deal with why a nicotine rich diet (Murray et al., 1994).
**Figure 4.2 - Transporters at the Blood Brain Barrier:**

A cartoon representing xenobiotic transport at the blood brain barrier is shown. Insecticides (red triangles) can be exported or imported into the CNS by ABCs embedded in the outward facing or inward facing glial cell membranes. Tight junctions hinder the diffusion of molecules without active transport.

**Figure 4.3 - Expression of the ABCB Full Transporters:**

The expression of each member of the ABCB full transporter clade is presented for Adult (A,C,E) and adult (B,D,F) life stages in the brain (A-B), midgut (C-D), and malpighian tubules (E-F). All expression data was obtained from ModEncode data on Flybase. Error bars represent standard error of the mean.
The tissue specificity of P-gp is particularly interesting in the case of *D. melanogaster*, which has 4 paralogues of the gene. While each parologue clearly clusters together there in terms of sequence similarity, there are only between 45 and 60% identical residues between the 4 proteins, which may give each different substrate specificities (Fig 4.1). Furthermore, there are non-overlapping expression patterns for each. While *Mdr50* and *Mdr49* are restricted to the midgut and malpighian tubules, *Mdr65* and *CG10226* are enriched in the CNS and adult tubules (Fig 4.3). *Mdr65* also shows enrichment in the tubules but only at the adult stage. Separate analysis of the glial cell transcriptome showed that the expression of both *Mdr65* and *CG10226* is enriched in these cells which form the protective BBB (DeSalvo et al., 2014). There is also functional evidence which implicates *Mdr65* in this protective role, as *Mdr65* null genotypes showed increased permeability to fluorescent dyes (Mayer et. al 2009).

In this chapter, the *Mdr65* gene is analysed for its role in resistance to several insecticides. Using CRISPR-Cas9, a full knockout of the gene was created in a matched genetic background. This means that the knockout genotype only differed from the control line in the absence of the *Mdr65* gene. The knockout mutant (*Mdr65KO*) was viable and showed no obvious signs of fitness cost, but was significantly more susceptible to a variety of different neuroactive insecticides. These findings were confirmed chemically and genetically by testing synergism with verapamil and crossing the knockout to deficiency genotypes respectively. HPLC-MS was also employed to measure the levels of the insecticide nitenpyram in adult bodies and heads following a sub-lethal exposure, revealing significantly more nitenpyram present in *Mdr65KO* flies. These data represent a starting point to consider the function of individual xenobiotic transporting proteins in insects.

### 4.2) Methods

#### 4.2.1) Drosophila Genotypes

Genotypes used in this chapter were alternatively ordered in from stock centres and generated herein (Table 4.1). The gene *Mdr65* was completely deleted from the Actin-Cas9 background to create the *Mdr65KO* genotype, using a transgenic CRISPR strategy (Port et. al 2014). Briefly, sgRNAs were designed to target the 5’ and 3’ UTRs of *Mdr65* were introduced into the PCFD4 plasmid using the Gibson Assembly Kit (NEB) and verified by sequencing using the PCFD4_seq primer (Appendix 1). The method was copied from Section 3.2.1, but the *Mdr65_PCFD4* primers were used for the initial amplification (Appendix 1). This plasmid was then injected into a line expressing φ31 integrase and containing an AttB landing site (Bloomington
Drosophila Stock Center #25709 in vermilion background), leading to the creation of the Mdr65-sgRNA genotype.

<table>
<thead>
<tr>
<th>Bloomington Stock #</th>
<th>Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>25709</td>
<td>25709v</td>
<td>AttB landing site Line</td>
</tr>
<tr>
<td>54590</td>
<td>Actin-Cas9</td>
<td>Expresses Cas9 ubiquitously</td>
</tr>
<tr>
<td>NA</td>
<td>Wxac</td>
<td>Background Control Genotype</td>
</tr>
<tr>
<td>NA</td>
<td>Mdr65KO</td>
<td>Mdr65 knockout in the Wxac background</td>
</tr>
<tr>
<td>7587</td>
<td>Mdr65_Df_Excel</td>
<td>Genotype with deficiency spanning the Mdr65 locus</td>
</tr>
<tr>
<td>24397</td>
<td>Mdr65_Df_BDSC</td>
<td>Genotype with deficiency spanning the Mdr65 locus</td>
</tr>
</tbody>
</table>

Table 4.1-List of Genotypes Used in Chapter IV:

A list of all genotypes used in chapter IV is provided along with their Bloomington stock numbers and short descriptions.

Deficiency lines from the Bloomington and Harvard Exelixis collections were used to verify that Mdr65 was the cause of any phenotypes observed. Mdr65_Df_Excel (BDSC# 7587) had breakpoints at 3L:6218161 and 3L:6263874, and Mdr65_Df_Bloom (BDSC# 24397) had breakpoints at 3L:6173578 and 3L:6256742. All cytological locations were referenced from Release 6 of the D. melanogaster genome, and the background of each deficiency line was w^1118.

4.2.2) Chemicals

Several insecticides were used in this study to test the substrate specificity of MDR65. Chlorantraniliprole (10 gL\(^{-1}\) Coragen1; Du Pont), Imidacloprid (200 gL\(^{-1}\) Confidor1; Bayer Crop Science), and Spinosad (10 gL\(^{-1}\) Success1; Yates) were bought commercially and were diluted in distilled water. Nitenpyram (Sigma), DDT (Supleco) and Ivermectin (Sigma) were purchased as analytical grade chemicals and dissolved in water, acetone and DMSO respectively. The P-gp inhibitor verapamil (Sigma) was diluted in DMSO and used at a concentration of 50ppm in combination with .2 or 1ppm nitenpyram to test for synergism.

4.2.3) Toxicology Bioassays

Larval resistance levels were measured using developmental toxicology assays (Daborn et al., 2007; Pyke et al., 2004). Briefly, 1\(^{st}\) instar larvae were obtained by letting cages of adult flies lay eggs on grape juice plates for 24 hours, at which time plates were removed. Eggs were then washed on to fresh plates and left undisturbed at room temperature for a further 18 hours to generate a population of 1\(^{st}\) instar larvae. 50 larvae were then picked onto corn meal media pre-mixed with a given concentration of insecticide. 50μL of food dye was added per 100mL of fly food, serving as a
visual marker of insecticide mixing. As insecticides were diluted in either water or DMSO, insecticide free control vials were dosed with the same amount of solvent as the insecticide treated vials. Survival was then scored 14 days after picking by counting the number of emerged adults. As ivermectin slows development an additional 4 days were allowed for development in bioassays using this insecticide.

Adult resistance levels were measured using acute contact assays described elsewhere (Daborn et al. 2007, Seong et. al 2016). Briefly, insecticides were first dissolved in acetone at a particular concentration. 300µL of the solution was then added to a scintillation vial, which was then automatically rolled using a hybridization oven, which evenly coated the vial in insecticide. 20 0-1 day old males were anaesthetized with CO2 and aged for 3 days before being transferred into the DDT vials (20/vial) without anaesthetization. Vials were then capped with a cotton stopper moistened with distilled H2O, and mortality was scored after one day by tapping the vials and counting any fly that either moved or twitched as alive.

4.2.4 The Quantification of Nitenpyram in Adult Fly Tissues

In order to determine where MDR65 was functioning endogenously to transport insecticides, amounts of nitenpyram were quantified in adult heads and bodies following a 12 hour 1ug/vial nitenpyram exposure using the adult bioassay conditions described above. All 20 flies in each vial were pooled in triplicates in 1.7mL Eppendorf tubes and immediately frozen in liquid nitrogen. 50 adult heads and bodies were then isolated from each pooled sample by shaking the frozen tubes and sorting. Heads and bodies were transferred into different 2mL cryomill vials (MP Bio), so that levels of nitenpyram in each sample could be measured. This was done by Roberto Fusetto using High Pressure Liquid Chromatography (HPLC) coupled with an Agilent 6520 Q-TOF mass spectrometer (Agilent Technologies, Inc., Santa Clara, CA, USA). Averages for each sample type were calculated from a total of 3 different pooled reps and Students t test was used to assess whether significant differences existed. Levels of nitenpyram were not sufficient for precise quantification so the area under the chromatogram curve was used to describe the relative amount in each sample.

4.2.5 Data Analysis

LC50 values were calculated for each genotype and each insecticide using the Pri-probit software (Sakuma, 1998; SAS equivalent method, Logistic distribution, Model 2: all or nothing natural response rate, Likelihood ratio test). Susceptibility ratios (SR) for each insecticide were calculated by dividing the Wxac (control) LC50 by the Mdr65KO (mutant) LC50. 95% confidence
intervals for SRs were determined by calculating finding which LC$_{50}$ in a given ratio had the highest variation and scaling that 95% confidence intervals to the ratio value. All direct comparisons between two samples were made using the Students t test (unpaired samples, unequal variance). When considering verapamil synergism Abbots correction was used in order to eliminate any effects observed at the control dosage (Hoekstra 1987; Rosenheim & Hoy 1989).

**4.3) Results**

**4.3.1) The Generation of the Mdr65KO Genotype**

In order to characterize the Mdr65 gene, a knockout genotype was created in the Wxac background. Crossing Mdr65-sgRNA genotype with the Actin-Cas9 genotype created a total deletion of the Mdr65 gene on the Actin-Cas9 third chromosome (Fig 4.4). A subsequent crossing scheme transferred the mutated chromosome into the Wxac line (Actin-Cas9 with the X chromosome from w1118), and used balancer chromosomes to prevent recombination (Fig 4.5). This means that the genetic background of the Mdr65KO genotype matched the final Wxac genotype. Originally, this genotype was confirmed by the ability to amplify a PCR product of ~500bp using the Mdr65KO primer set flanking the sgRNA the cut sites (Appendix 1). Sanger sequencing across this region confirmed the that this PCR product contained DNA from either side of the cut sites, confirming the deletion. Because the deleted region contains the transcriptional start site, there is no predicted product of Mdr65 in the Mdr65KO genotype.

**4.3.2) Mdr65KO is Susceptible to a Variety of Insecticides**

In order to test the substrate specificity of MDR65, SRs between Mdr65KO and Wxac genotypes were calculated for several insecticides. Of the insecticides tested here, chlorantraniliprole elicited similar LC$_{50}$ values for Wxac and Mdr65KO. Conversely, Imidacloprid, ivermectin, DDT, spinosad and nitenpyram showed various magnitudes of increased susceptibility (Fig 4.6). A very modest increase in the SR was observed with imidacloprid and ivermectin, and these insecticides were very close to the significance threshold. DDT, spinosad and nitenpyram however, showed SRs of 2.5, 3.5 and 4.5 fold respectively. This suggests that Mdr65 contributes to the wild type level of resistance for several neuroactive insecticides.

**4.3.3) Mdr65 Deficiency Lines Fail to Complement the Mdr65 Loss of Function**

Two deficiency lines missing the region of chromosome 3 that spans the Mdr65 locus were used to confirm that the deletion of Mdr65 was causing the susceptibility observed in Mdr65KO.
Crossing either deficiency line to Mdr65KO would produce progeny without *Mdr65*. When Mdr65KO was crossed to either deficiency line, the resistance level of the progeny to nitenpyram was similar to that observed with a homozygous Mdr65KO genotype in adult resistance assays. However, when Mdr65KO was crossed to the deficiency’s control (w^{1118}), which was wild type at the *Mdr65* locus, no change in nitenpyram resistance was observed compared to wild type (Fig 4.7). This confirms that the *Mdr65* gene was responsible for the observed differences between Wxac and Mdr65KO and also indicates that one copy of the *Mdr65* gene was sufficient for wild type response.

**Figure 4.4- *Mdr65KO* Schematic:**
A representation of the Mdr65 deletion is shown that was approximately common to each of the three backgrounds. A) The Wild type Mdr65 gene is shown. B) The Mdr65 knockout allele is shown which truncates the gene and introduces a stop codon.
Figure 4.5- Mdr65KO Crossing Scheme:

The crossing scheme that created the Mdr65KO genotype is shown. A balanced 2nd chromosome expressing two sgRNAs targeting the 5’ and 3’ UTRs of Mdr65 (green) is crossed to a background containing a Cas9 expression cassette (blue). The presence of these two factors creates the deletion which is then backcrossed into the Wxac background.
Figure 4.6- The Relative Susceptibility of Mdr65KO to a Panel of Insecticides:

The susceptibility ratios between Mdr65KO and its control line (Wxac) are plotted for several insecticides. A susceptibility ratio of 1 indicates no change in resistance and any value above 1 indicates the magnitude of susceptibility. Error bars represent 95% confidence intervals.
Figure 4.7 - Mdr65KO crossed to Mdr65 Deficiencies:

The percent survival of Mdr65KO crossed to genotypes with chromosomal deficiencies at the Mdr65 locus is plotted with appropriate control genotypes. The two blue colours represent crosses with at least one functional copy of Mdr65, while the grey and yellow dots lack this gene. Error bars represent 95% confidence intervals.
4.3.4) Mdr65KO Fails to Synergize with Verapamil

The ABC transporter inhibitor verapamil was also used to chemically verify that a ABC transporter like protein was involved in the change in relative insecticide resistance. When verapamil (50ppm) was used in conjunction with a discriminatory dose of nitenpyram (1ppm) in the Wxac control genotype, significantly more flies died than did when the insecticide was administered without verapamil (Fig 4.8A). In contrast, no difference was observed between verapamil + and verapamil - treatments when a discriminatory dose of .25ppm was used for Mdr65KO (Fig 4.8B). These data shows that verapamil was only able to synergize with nitenpyram when a functional copy of Mdr65 was present in the genome.

4.3.5) Mdr65KOs Retain More Nitenpyram than its Control

In order to directly measure levels of insecticide in flies with and without Mdr65, HPLC-MS was used to quantify the amount of nitenpyram in adult heads and bodies following exposure to nitenpyram. Relative quantification of nitenpyram levels in adult heads and bodies revealed that in both genotypes far more nitenpyram was present in the bodies than in the heads. This is unsurprising as the heads are smaller in size relative to the bodies. There was also significantly higher concentrations of nitenpyram in the Mdr65KO genotype in both tissues when compared to the Wxac control (Fig 4.9). Additional information can be gleaned from the ratios between different genotypes in the same tissue and between different tissues in the same genotype. The difference between Wxac and Mdr65KO was most pronounced in the bodies of the two genotypes, and 4.60 times more nitenpyram was detected in the bodies compared to a 2.97 fold difference in the heads (Table 4.2). Comparing tissues within the same genotype, the differences between head and body samples was greatest in the Mdr65KO genotype, which had 37.00 times more nitenpyram in the body compared to the head sample. This difference was only 23.68 fold in the Wxac control genotype. From these data it can be concluded that the Mdr65KO genotype accumulated far more nitenpyram than the Wxac genotype. Furthermore, this accumulation appears to be concentrated in the body of the insects rather than in the head, pointing to a role for Mdr65 in excretion.

4.5) Discussion

ABC transporters such has P-gp have long been thought to be involved in insecticide resistance, but published data typically establish correlation rather than causation. Here, a precise removal of the D. melanogaster P-gp orthologue Mdr65 is described. The knockout genotype showed significantly lower tolerance for several neuroactive insecticides. That this was due to an
ABC transporter was supported by the verapamil synergism experiment and that the transporter involved was *Mdr65* involved by the reinforced by the genetic complementation studies with deficiency mutations. Consistent with the data from toxicological studies, significantly less nitenpyram was detected in Mdr65KO flies compared to their Wxac controls. The findings in this chapter differ fundamentally from those presented in chapter III. While *Cyp6g1* causes resistance if overexpressed and does nothing in its ancestral form, *Mdr65* plays a role in how wild type flies from susceptible laboratory strains cope with insecticides.

![Synergism with the ABC Inhibitor Verapamil](image)

**Figure 4.8: Synergism with the ABC Inhibitor Verapamil:**

The corrected survival of A) Mdr65KO and B) Wxac are plotted at discriminatory doses (1, .25 ppm) of nitenpyram with verapamil (Red) and without verapamil (Black). Error bars represent 95% confidence intervals.
**Figure 4.9- HPLC-MS of Nitenpyram in Mdr65KO:**

The relative amount of nitenpyram recovered from adult heads and bodies following a 12 hour exposure is plotted for Wxac and Mdr65KO. Crossbars represent the mean quantity of nitenpyram calculated by averaging the area under each sample’s chromatogram.
<table>
<thead>
<tr>
<th>Numerator</th>
<th>Denominator</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wxac Bodies</td>
<td>/ Wxac Heads</td>
<td>23.69</td>
</tr>
<tr>
<td>Mdr65KO Bodies</td>
<td>/ Mdr65KO Heads</td>
<td>37.00</td>
</tr>
<tr>
<td>Wxac Bodies</td>
<td>/ Mdr65KO Bodies</td>
<td>4.64</td>
</tr>
<tr>
<td>Mdr65KO Bodies</td>
<td>/ Wxac Bodies</td>
<td>2.97</td>
</tr>
</tbody>
</table>

**Table 4.2- Ratios between Heads and Bodies of Wxac and Mdr65KO:**

A table showing pairwise comparisons of nitenpyram levels in Wxac and Mdr65KO heads and bodies.

Since the first reported case of transporter mediated drug resistance in mammalian cell culture (Kartner et al., 1983), ABC transporters have been understood as a potent resistance mechanism. The leap to insect research was made when Murray et al. (1994) hat a P-gp like protein was enriched at the BBB of the tobacco hornworm *Manduca sexta*, probably contributing significantly to the high nicotine tolerance in this species. It is therefore surprising that so little functional genetics has been done on insect P-gp orthologues. This is the first study to make a clean deletion of an insect ABC transporter gene in a matched genetic background, providing mutant and control genotypes that differ only at the *Mdr65* locus. Much effort was made to ensure that the only modification to the Wxac genome was at the *Mdr65* locus and that the change made was a total removal of the *Mdr65* gene. The preliminary finding that *Mdr65* contributed to insecticide resistance was confirmed by both chemical and genetic means (Fig 4.3, 4.4). Other studies have reported phenotypes for an *Mdr65* hypomorph, but some expression of *Mdr65* was detected, and the control line did not share the same genetic background (Luo et. al 2013, Mayer et. al 2009).

The current study also expanded upon published research by taking a more systematic approach to which insecticides were tested. Rather than focusing on one particular insecticide, a panel of insecticides were considered representing different chemical classes. The chemical properties of these insecticides such as molecular weight or polarity did not appear to influence the SRs obtained from comparing Wxac to Mdr65KO. This is exemplified by comparing the chemically similar neonicotinoid insecticides. Imidacloprid elicited the smallest change in susceptibility tested here while nitenpyram elicited the largest (Fig 4.6). Further, spinosad and ivermectin are both macrolactones, yet removal of *Mdr65* significantly influenced spinosad resistance but not ivermectin resistance. From the data reported here, there does not appear to be any obvious pattern to the substrate specificity of MDR65.

Because of their role in human cancer drug resistance, much biochemical work has been dedicated to understanding the substrate specificity of P-gps. *In vivo* and *in vitro* mammalian models are often used to estimate which molecules will be good substrates for P-gp, but
inconsistencies between models makes it difficult to establish any firm rules (Doran et al., 2005; Feng et al., 2008). The biochemical basis for polyspecificity was better understood following the publication of the crystal structure of P-gp from *Mus musculus*, which revealed a polyspecific binding site that is based on “hydrophobic and aromatic interactions” (Aller et al., 2009; Li et al., 2014). Further in silico modelling has lead to several criteria that are predictors of P-gp substrate status, such as polarity, size and flexibility as well as specific interactions between P-gp residues and the substrate in question (Erić et al., 2014). Despite this no perfect model has yet been described suggesting a need to test substrates in vivo as well as in silico.

The expression pattern of *Mdr65* likely influences its function as a drug transporter. Within the ABCB subfamily there are 4 P-gp orthologues, but *Mdr65* is one of only two that has significant enrichment in the glial cells and is the only one that has been characterized as a putative BBB transporter (Mayer et. al 2009, De Salvo et. al 2014). The efflux of toxins from the CNS would seem to be the primary role of *Mdr65* in larvae because the gene shows a disproportionate enrichment in the BBB during larval stages (Fig 4.3). Since chlorantraniliprole’s primary targets the ryanodine receptor in muscles (Lahm et al., 2009), the first instar toxicology assays conducted here would not have been expected to elicit a large difference in response comparing the Mdr65KO and Wxac control. Conversely, every neuroactive insecticide tested here significantly changed the resistance level of the flies, even if only very slightly in the case of imidacloprid and ivermectin.

In the adult stage, however, *Mdr65* shows enrichment in malpighian tubules in addition to the BBB (Fig 4.3). The Malpighian tubules have been shown to be crucial for effluxing substrates from the hemolymph out of the body, regulating the concentrations of both endogenous chemicals and xenobiotics within an organism (O’Donnell, 2009). This role of the tubules is assumed by the kidneys in mammals, and the *Mdr65* orthologue P-gp is expressed there in addition to the mammalian BBB (Theibaut et. a 1987). It is likely that *P-gp* in insects plays similar roles in transporting xenobiotics out of the body as well as the brain. A role for *Mdr65* playing a role in the malpighian tubules is further supported by the ratios of nitenpyram levels among genotypes and tissues (Fig 4.9). The ratio between amounts of nitenpyram found in Mdr65KO bodies to Mdr65KO heads, was greater than that when the same comparison was made in Wxac, indicating Mdr65KO accumulated nitenpyram primarily in the body (Table 4.2). This would suggest that the primary mechanism of transport is at the tubules rather than the BBB. One explanation for this result would be the size of the tubules relative to the BBB. While only a small amount of insecticide is likely transported across the BBB, a much larger amount can be effluxed from the body across the tubule. It is very possible that, although transporting less insecticide, the BBB is more important in determining resistance due to the importance of the brain as the target tissue. In any case, these data
suggest that ABC mediated resistance is influenced by places other than just the BBB. Similarly, MDR65 has been demonstrated here to be an important transporter in the response to neurotoxic insecticides, but it is unlikely to act alone.

The findings in this chapter represent the first step forward in a more functional, precise approach to the study of ABC transporter mediated resistance. By eliminating individual genes and assessing their resulting changes in insecticide resistance, the prevalence and capacity of ABC transporter mediated resistance to take place in field populations can be better understood. Furthermore, combining a knowledge of the tissue-specific expression of the gene and the outcomes from toxicological and biochemical analysis can inform a discussion on the movement of insecticides around the body of the insect.
Chapter V: Discussion
5.1) Overall Discussion

In the course of this work several fundamental concepts focusing around insecticide resistance specifically, and adaptation to strong selective pressures in general, have been explored. Although there are many ways to organize such ideas, this discussion will be divided into two large sections. Part one will focus specifically on the biology and physiology of insecticide resistance. Here, the various ways that insects are impacted by insects and how insects can tolerate higher concentrations of insecticides will be discussed. Following this will be a discussion of the more abstract genetic concepts touched on in this work. Specifically an emphasis will be put on describing the genetic basis underlying observed phenotypes by both identifying and quantifying the contribution of an allele in a different genetic backgrounds. Along the way, the data from this study will be contextualized in terms of previously published works and future directions will be suggested.

5.2) Insecticide Biology

5.2.2.1) Secondary Target Receptors

Undoubtedly, the primary mechanism through which insecticides affect insects is through their primary target proteins, but these targets do not sit on the proverbial phylogenetic desert island. In fact, most of the targets for insecticides are closely related to other proteins expressed in similar tissues meaning that triangular crosstalk between insecticides, primary targets and secondary targets is not uncommon. This was clearly demonstrated in Chapter II, which featured a genotype lacking spinosad’s primary target site responding to spinosad in an acute motility assay (Fig 2.6). The target of spinosad is the α6 subunit of the nAChR family; this family contains 9 other members with conserved functional domains. While there is no direct evidence that spinosad binds to any of the other nAChR subunits receptors, the high degree of homology between subunits makes it a distinct possibility. There is strong evidence that imidacloprid, also considered in chapter II, binds to the Dα3 subunit in addition to its main targets of Dα1 and Dβ2 (Chen unpublished). This could underpin the moderate level of response observed in the Dα1 or Dβ2 subunit mutants (Fig 2.5).

Widening the circle further, nAChRs are members of an even larger family composed of pentameric ligand gated ion channels, many of which are common targets of other insecticides (Dent 2010). The ability of insecticides to bind non-target receptors is poorly understood, but there is precedent for cross-talk within the pentameric ligand gated ion channel family (Taylor-Wells et al., 2015). It is
probable that secondary target binding affinities are far lower than primary targets, but it may underpin more subtle resistance phenotypes. Off target binding may influence behavioural phenotypes but not mortality as was suggested in Chapter II. If true, distinguishing such difference could be valuable tool in understanding the role of specific receptors in neural circuits.

5.2.2.2) Oxidative stress/ lipids

Not all effects mediated by a xenobiotic result from the binding to any specific target. For instance, some heavy metals are highly damaging to insects not by binding a receptor, but rather by interfering with the general chemistry of the cell (Wang and Shi, 2001). Similar non-target effects may be seen with insecticides. What is less clear is what causes this change and how it affects the insect. Non-target effects may be a downstream consequence of the impact the insecticide has on the target or/and independent of it. Recent studies provide some clues. Higher levels of oxidative stress were detected in insects following sub-lethal exposures to the pyrethroid insecticide permethrin (Brinzer et al., 2015). Also, increased levels of smaller lipid droplets were detected in 3rd instar Drosophila fat bodies following imidacloprid or spinosad exposures, while the opposite was observed in midguts from the same individuals (Martelli unpublished). It is possible that these two processes are linked as the latter result suggests lipid mobilization between tissues is occurring which could be the cause of or result of the oxidative stress. In either case, neither the mechanism by which insecticides trigger this kind of response nor the downstream effects of these changes are well understood. Interestingly, perturbations in both lipid homeostasis and oxidative stress have been shown to influence transcription factors which regulate many drug metabolism proteins such as P450s and ABC transporters (King-Jones et al., 2006; Misra et al., 2011; Sieber and Thummel, 2009). This could mean that non-target effects like lipid mobilization and oxidative stress help regulate of the more widely cited resistance genes. If this were the case, genetic variation underpinning the lipidomic or oxidative response, could influence resistance.

5.2.2.3) The Utility of Alternative Bioassays

The choice of bioassay will influence the insecticide response phenotype and thus influence the genes identified as underpinning the trait (Section 1.5). While increasing the number of bioassays is optimal, limitations on time and resources mean that intelligent choices must be made regarding which tests to perform. One approach is to look at more precise phenotypes that are components of the overall resistance level. This was exhibited in chapters III and IV by considering the levels of insecticide and metabolites in different tissues at a given time point. High concentrations of insecticide in the body will precede insecticide mediated death, but isolates
metabolic genes and removes other confounding variables. Considering such a phenotype could provide valuable information about insecticide metabolism and excretion even in a background with target site mediated resistance. Likewise, a prerequisite for metabolic mediated resistance will be the production of less toxic or more easily excreted insecticide metabolites, which can be assessed by HPLC-MS (Fig 3.9). Thus, by focussing on phenotypes upstream of death it is possible to isolate the many processes which influence resistance without the more subtle ones being masked by the processes or larger effect. This can be taken to its logical conclusion by devising ever more exact methods for examining insecticide related phenotypes. Future studies may attempt measure the concentration of insecticide in isolated tissues at various time points, or directly measure the rate of transport into and out of specific tissues.

Another approach is to simply consider a greater number of complex phenotypes, which are high-throughput enough to be time efficient. Some aspects of insecticide biology may only be apparent when insects are exposed to certain types of assays, and this approach increases the chances that a gene which plays a role in insecticide biology will be identified. As an abstract example, a resistance gene only expressed at the adult stage may confer no resistance in a bioassay targeted at 1st instar larvae. Testing at both life stages would identify this gene, which would have to be further characterized in future studies. This may have been what was occurring in Mdr65KO with regards to nitenpyram accumulation (Fig 4.9). Although functional as a BBB efflux transporter in larvae, examination of Mdr65’s adult expression pattern points to a role in tubule mediated efflux that would be overlooked in any larval assay. Variation in the type of exposure was also shown to influence results. While large differences were observed in 1st instar toxicology between Mdr65KO and Wxac, their behaviour in the Wiggle Index was identical (Denecke unpublished). This is probably due to the higher doses and shorter times necessary for a Wiggle Index assay, which may have overwhelmed the transportation mechanism. Similar discrepancies were observed with Dα6 mutants which showed no mortality response to spinosad, but did show behavioural differences (Fig 2.6).

These data highlight how differences in bioassays can test different aspects of the insecticide’s biology, that are missed when a single assay is considered. These differences are potentially economically consequential as often times field resistance is only assessed via one technique. The array of bioassays used in the current work provides a far deeper understanding of where the insecticide goes once inside the body and what it does when it gets there. Although little is usually understood regarding an insecticide’s biology at the time of its commercial release, the increasing prevalence of field resistance and public health concerns means that a better understanding of insecticide biology is needed to more rationally design and apply insecticides. The
current research points to the need for a range of resistance assays in order to better understand the genetic basis of insecticide resistance. Clearly, there are advantages to conducting these experiments in the *D. melanogaster* model system but the resistance and metabolic assays used here can be employed in pests and will be empowered by the ever increasing availability of resources (e.g. genome sequences) and powerful tools (e.g. CRISPR) in pests.

### 5.2.3) The Evolutionary Response to Selection by Insecticides: Physiology

#### 5.2.3.1) Target site resistance

There are many physiological processes which determine the insect’s response to an insecticide. One of the most common of such physiological changes is the modification of the insecticide’s target site, which was considered in Chapter II. Of particular economic interest is the likelihood of particular subunit being modified in a field population. This can be determined for any mutation by comparing the fitness gain in terms of insecticide resistance against the fitness costs it generates. Field and laboratory studies reveal that single amino acid substitutions, truncations, and null alleles may all evolve in response to insecticidal selection pressures (Ffrench-Constant et al., 1993; Perry et al., 2008, 2007). Interestingly, the fitness costs imposed on the genotypes considered in Chapter II appeared to have much more to do with the particular gene affected rather than the severity of the mutation. *Dα1* and *Dβ2* mutants showed reduced motility measured by the Wiggle Index and were generally unhealthier compared to *Dα6* nulls, which behaved like wild type flies. This occurs despite the fact that *Dα1* and *Dβ2* mutants containing only truncations or amino acid substitutions as oppose to *Dα6* nulls showing no *Dα6* expression. This discrepancy may also explain why *Dα6* mediated spinosad resistance is so common, while P450 mediated resistance is much more common in response to imidacloprid.

The difference in fitness cost between nAChR mutants that confer resistance to spinosad and imidacloprid highlights an import concept in receptor biology. Namely, that a key factor influencing fitness costs of a receptor mutant is the degree of functional redundancy for that receptor. Some nAChR subunits appear to be essential. RNAi knockdown of the *Dα5* subunit is developmentally lethal, and the removal of *Dβ1* appears to be homozygous lethal (Christensen et. al 2017, Yang unpublished). Underpinning this is the role many nAChRs play in essential processes. By way of contrast a loss of function for the *Dα1* subunit is not lethal; it confers a significant level of imidacloprid resistance as well as mating and sleep defects (Somers et al., 2017). Further sleep abnormalities have been found in other nAChR subunits (Luong unpublished). There is evidence that functional redundancy among the nAChR subunits may be due to a level of regulatory compensation. A *Dα1* null mutant was found to be spinosad sensitive, likely to be due to trafficking...
of more Dα6 protein to the membrane (Nguyen and Perry, unpublished). Compensation of nAChR subunits for missing members of the family may partially explain the differences in observed phenotypes between point mutations and full deletions (Perry et al., 2015; Somers et al., 2015).

5.2.3) DMEs

5.2.3.1) Evolutionary paths/fitness costs

The other common resistance mechanism is the upregulation of DMEs, mainly P450s. P450 upregulation tends to exert less of a fitness cost than does target site mutations, but also generally provides lower levels of overall resistance. The reason for this lies in how the two mechanisms of resistance evolve. Insecticides tend to target proteins with critical physiological roles, so their removal often has adverse effects on the fitness of the insect. Conversely, many DMEs have no characterized endogenous role; increasing their expression only costs a small increase in energy expenditure. This dichotomy is not absolute as P450 resistance can evolve through amino acid substitutions and many P450s have essential endogenous roles (Claudianos et al., 1999; Rewitz et al., 2006). Similarly, while the overexpression of P450s does not always cause fitness defects, sometimes this is not the case. Driving the lufenuron metabolizing Cyp12a4 gene ubiquitously lead to lethality (Bogwitz et al. 2005). However, the majority of DMEs involved in insecticide metabolism belong to hyper-variable subfamilies of the P450 superfamily which have less of a critical role in endogenous metabolism (Feyereisen, 2011). Additionally, enhancers such as those in Accord can upregulate the expression of P450s only in metabolic tissue, minimizing any fitness costs incurred by ubiquitous expression. In Chapter III it was shown that, within the DGRP and in relation to imidacloprid, DMEs were playing a large role in determining insecticide resistance and that the primary contributor was a non-essential P450 Cyp6g1 (Fig 3.8-3.12).

5.2.3.2) Substrate Specificity of P450s

The capacity for P450s to be the underlying evolutionary mechanism for resistance appears to be related to their affinity for a particular class of insecticides. For example, resistance mechanisms reported for neonicotinoids such as imidacloprid usually involve P450s (Bass et al., 2015). Conversely, resistance to macrolactones such as spinosad and ivermectin is not commonly due to P450 activity. It is possible is that the propensity for target site resistance to evolve for macrolactones is due to the low fitness costs engendered by their target receptor mutations (Section 5.2.3.1). Alternatively, the discrepancy could be because P450s have a lower affinity for macrolactones compared to neonicotinoids. However, the details of which substrates bind to which P450s is poorly understood and often times closely related chemicals show different substrate
specificities. For example a member of the Cyp9 family in honeybees was recently reported to metabolize only thiacloprid while showing little affinity for any other member of the neonicotinoid family (Nauen unpublished). Furthermore, it is unknown whether the large contribution of \textit{Cyp6g1} to imidacloprid metabolism (Chapter III) is due to that enzyme’s high affinity for imidacloprid compared to other P450s or whether it is simply due to \textit{Cyp6g1}’s relative abundance in metabolically active tissues. Analysis of the DGRP for azinphos-methyl resistance showed a much stronger association for the same gene to that chemical compared to imidacloprid which possibly indicating a higher affinity (Battlay et. al 2016). Although there have some efforts made to determine substrate recognition sites, the substrate binding capacities of CYP6G1 are not perfectly understood.

\textbf{5.2.3.3) P450 Diversity in Metabolism}

In the preceding section the question the question \textit{how many of insecticides can a given DME metabolize?} was considered. Here the question is put into reverse asking \textit{How many P450s in an organism can act on a given insecticide?}. As discussed in Section 3.4.3, not all of the HPLC-MS and Wiggle Index phenotypes could be explained by \textit{Cyp6g1}. On the contrary, there is still a significant amount of imidacloprid metabolism occurring in the RAL\textsubscript{517}-\textit{Cyp6g1KO} genotype (Fig 3.13, Fusetto unpublished). This suggests that other DMEs are involved in metabolizing imidacloprid. P450 activity is suggested because the metabolites produced, IMI-5OH and IMI-Olefin remain the same. However, the identity and contribution of these P450s is uncertain. Several conjectures were made regarding the identity of potential imidacloprid metabolizers (Section 3.4.3). Apart from the previously implicated genes of \textit{Cyp6v1} and \textit{Cyp6t1}, several other P450s implicated in metabolism of other insecticides could be likely candidates. These include all P450s tested by previously implicated in insecticide resistance including \textit{Cyp12d1}, \textit{Cyp6a8}, \textit{Cyp6a2} and \textit{Cyp12a5}. \textit{Cyp6t3} would also be a suitable candidate as it shows a high degree of sequence similarity to \textit{Cyp6g1}.

While noting the potential for other genes to be involved, the best supported candidate is \textit{Cyp6g2}, the gene adjacent to \textit{Cyp6g1} on the 2\textsuperscript{nd} chromosome. The transcripts of this gene were among those most highly associated with the Wiggle phenotype, and \textit{Cyp6g2} has been shown to directly metabolize imidacloprid, yielding the same metabolites as \textit{Cyp6g1} (Fusetto unpublished). It is still not certain whether this association is an artefact of \textit{Cyp6g1} expression; the two transcripts show a significant correlation in terms of magnitude suggesting that they are co-regulated (Battlay et. al 2016). Furthermore, \textit{Cyp6g2} appears to only be expressed in non-metabolic tissues in reference strains (Chintapalli et al., 2007; Chung et al., 2009; ModENCODE Consortium et al.,

117
However, the fact that Cyp6g2 is co-regulated with Cyp6g1 in the DGRP suggests that its expression pattern may be different than in the reference laboratory strains. The Accord transposon responsible for Cyp6g1 expression could be influencing Cyp6g2 expression due to the close proximity of these two genes. Cyp6g2 expression is significantly (p<2x10^{-15}) higher in AA and BA haplotypes compared to M haplotypes within the DGRP. Further, significantly higher levels of Cyp6g2 expression have been measured in the midgut and malpighian tubules of two AA haplotype containing DGRP genotypes compared to two M genotypes (Martelli unpublished).

Despite the capacity of many D. melanogaster P450s to metabolize many different insecticides, it appears that a much smaller subset are commonly identified in field populations. Similar observations are made in other species where usually only one or two P450s that are identified as insecticide metabolizers. If many P450s are capable of metabolizing an insecticide, why then are only one or two usually upregulated? Several different possibilities exist.

One hypothesis is that the discrepancy between the number of P450s to metabolize an insecticide and the number upregulated in the field is due to the fact that evolution proceeds in a random way. This hypothesis can be called The Blind Watchmaker Model (Dawkins, 1987). In this case, a gene such as Cyp6g1 is overexpressed not because it is the best candidate for drug metabolism, but because it was the first. It was by chance that an Accord element inserted in front of the Cyp6g1 gene instead of another gene equally capable of metabolizing insecticides. This hypothesis is supported by population genetics analysis of worldwide Cyp6g1 alleles, which suggest that the Accord insertion happened very recently and only once (Catania et. al 2004, Schmidt et. al 2010). Once established, it was then far more likely for the pre-existing Cyp6g1 overexpression haplotype to spread around the globe than it was for another P450 to become upregulated. The competing hypothesis would assume the opposite, that the upregulated P450 responsible for insecticide resistance is the most suited P450 in terms of the resistance:fitness cost ratio. This could be called the Evolutionary Peak Model. Applied to the example at hand, upregulating Cyp6g1 conferred the greatest fitness benefit compared to all other P450s either due to its high capacity for insecticide metabolism, its low fitness costs compared to other P450s (Section 5.2.3.1) or both. It is possible that Accord could have and did insert upstream of other P450s, which were either poor insecticide metabolizers or had detrimental side effects that prevented the alleles from spreading throughout global populations. It should be noted that the very recent widespread use of the same insecticides is likely to have been a major factor in the global spread of Cyp6g1 resistance. Prior to this D. melanogaster will have been exposed to plant toxins that are unlikely to have been so homogeneously found in habitats around the world as synthetic insecticides have been.
It is most probable that a combination of these two models is what has occurred in D. melanogaster. Cyp6g1 is clearly capable of metabolizing and conferring resistance to a wide range of insecticides in ways that many other P450s cannot. Beyond insecticides, there is an element of sexual selection at play because females tend to experience a fitness benefit from Cyp6g1 overexpression, while males tend to have associated fitness costs (Hawkes et al., 2016). However, it must be stressed that other P450s could play a role similar to Cyp6g1 if by chance they were overexpressed. Studies comparing the transgenic expression of Cyp6g1 to Cyp6g2, found Cyp6g2 to confer higher levels of nitenpyram resistance despite lower levels of transgenic expression (Daborn et al. 2007). Also, anecdotal evidence suggests that in vitro assays reveal CYP6G1 to be a relatively poor metabolizer of insecticides compared to several other P450s tested (Nauen unpublished). The situation may be different in other species, but further analysis of resistant populations is needed. If an upregulated DME evolved several times in independent populations this would support the Evolutionary peak model and if the converse were true the Blind Watchmaker model would be supported.

5.2.4) Transporters

5.2.4.1) Fitness Costs and Prevalence in Field Resistance

The final resistance mechanism explored in the current work centres around of how the insecticide moves between tissues in the insect, and more specifically how it ends up in the tissue expressing its targeted receptor. Until recently, most reviews of insecticide resistance made only passing references of ABC transporters (Perry et al. 2011), although there have been several recent reviews dedicated to transporters (Dermauw et al. 2014, Merzendorfer 2014). This mechanism is far less reported in resistance literature, but the frequency of such reports is rapidly rising. Indeed, 82% of all mentions of transporter mediated resistance have occurred within the last 5 years (Porretta et al. 2016). Such an increase in recent attention likely stems from a detection bias present in most resistance literature studies. Porretta et. al (2016) made a laudable attempt to quantify how heterogeneous the involvement of ABC transporters was in insecticide resistance, but such numbers are difficult to ascertain because most labs fail to publish negative results. The fact that transporters have only recently been studied and that they are less frequently cited as a resistance mechanism likely reflects their more marginal role in insecticide resistance in field populations compared to target site or P450 mediated resistance.

However, this assertion is not true across all taxa, and transport appears to be the primary mechanism for resistance in some species. For instance, the original discovery of P-gp was in
mammalian cell lines, which have since been shown to frequently upregulate members of the ABCB, ABCC and ABCG families in order to resist chemotherapeutic treatments (Gottesman and Ling, 2006). Similarly, parasitic nematodes appear to preferentially use the upregulation of ABC transporters to gain resistance to anthelmintics (Kaschny et al., 2015). While it is clear that some species have a greater propensity to use transporters as a resistance mechanisms, it is unclear whether this difference is caused by the difference in drugs used to control each one or whether it reflects differences in the genetic arsenal contained within a species. Furthermore, the types of transporters upregulated varies depending on the species. Fungi tend to use members of the major facilitator superfamily to efflux toxins while animals rely much more heavily on ABC transporters. Bacteria employ a variety of multi drug efflux pumps including MATE transporters in addition to ABCs and MFSs (Higgins, 1992).

It is unclear whether variation in ABC transporters incurs any fitness costs. As was discussed with receptor and P450 modifications, fitness costs determine how likely a resistance mechanism is to evolve in a field setting. The *D. melanogaster* transporter *Mdr49*, for example, was shown to be a transporter of endogenous germline attractants, although a clean deletion of the gene suggests that removing it has no noticeable fitness defect (Ricardo and Leehman 2009, Denecke unpublished). *MRP4* has also been implicated in combating oxidative stress response and ageing by comparing *MRP* hypomorphs to a wild type line after significant backcrossing (Huang et al., 2014). Most famously several members of the ABCG family transport pigment into the eye and are responsible for some of the most well known *D. melanogaster* phenotypes such as *white* (Ewart & Howells 1998). Although existing happily in laboratory settings, the *white* gene does carry a severe fitness cost because the pigment deficiency inhibits eyesight (Cosens and Briscoe, 1972). However, the fitness costs of field derived resistance alleles have not been analysed, and no transporter that has been linked to drug resistance has clearly showed any fitness effects. In the current work, no fitness cost was observed when *Mdr65* was deleted from the genome.

**5.2.4.2) Substrate Specificity**

As was true for DMEs, the ability of transporter proteins to confer resistance to a given insecticide depends on the ability of a specific transporter to bind that insecticide. In chapter IV, the capacity of *Mdr65* to confer resistance was found to depend on the which insecticide was used. There was no discernible pattern based on the chemical structure of the several insecticides tested (Section 4.3.2). The extent to which the specificity observed is conserved in *Mdr65* orthologues and paralogues is unknown. P-gp mediated transport is the primary mechanism for ivermectin resistance in nematodes (Kaschny et. al 2015), yet data from Chapter IV suggests that *D. melanogaster Mdr65*
does not have a high affinity for ivermectin (Fig 4.6). These data suggest that differences exist between the biochemistry of P-gp orthologues from widely divergent taxa. If similar differences are observed between mammalian and *D. melanogaster* P-gp, it may not be possible to use the structural model of mammalian P-gp to create models of insect orthologues to predict their capacity to confer insecticide resistance. Substrate specificity may be more highly conserved within species although there is too little data to draw any firm conclusions. The *Mdr65* parologue *Mdr49* was capable of conferring resistance to DDT, in concordance with the findings of the current work (Fig 4.6; Seong et. al 2016). Furthermore, a genotype with a clean knockout of *Mdr49* showed differences in resistance to spinosad and nitenpyram, which overlap with observations for *Mdr65* (Denecke unpublished). Future studies looking at knockouts or overexpression lines of closely related ABC transporters, will help answer this question. The controlled expression of ABC transporter genes from other species in *D. melanogaster* may permit their substrate specificity to be examined.

The broad substrate specificity of some or all ABC transporters may underpin their ability to confer resistance in field populations. A similar advantage can be assumed for polyspecific DMEs, which can convey resistance to many insecticides. This is especially relevant in recent times due to the proliferation of resistance management strategies involving the rotation of different insecticide classes and allows refuge areas to select against fitness reducing resistance alleles. DMEs and ABCs, which confer polyspecific resistance at relatively low fitness costs, may be preferred over target site resistance which confer resistance to specific modes of action and often impose fitness costs. This may occur in spite of the relatively low levels of resistance conferred by metabolic resistance compared to target site mediated resistance.

### 5.2.4.3) Tissue Specificity

The precise sub-cellular expression pattern of drug transporters is likely very important in determining their capacity to confer resistance. Upregulation of a transporter facing, for example, into the CNS, would increase the amount of insecticide reaching its target and thus decrease resistance. In chapter IV, attention was restricted to *Mdr65*, which is expressed almost exclusively in the BBB in larvae and also in the malpighian tubules in adults (Fig 4.3), and likely plays an excretory role in both these tissues (Mayer et. al 2009, Fig 4.9). However, other members of the ABCB family in *Drosophila*, such as *Mdr49* and *Mdr50* are expressed in the malpighian tubules and midguts, while *CG10226* shows similar expression patterns to *Mdr65*. Studying these genes may reveal insights to how insecticides and other toxins are excreted and gain entry to the hemolymph respectively. Wu et al. (1991) found that an *Mdr49* hypomorph showed increased
susceptibility to colchicine, and Seong et. al (2016) transgenically overexpressed Mdr49, which conferred DDT resistance. No functional data exists on Mdr50 or CG10226. While this research is still in its infancy a limited subset of unpublished data hints at a potentially interesting story for these paralogues. A genotype carrying a clean deletion of Mdr49 show susceptibility to nitenpyram but resistance to spinosad (Denecke unpublished). While the magnitude of such differences were far smaller than those observed for Mdr65 the fact that the differences were in opposite directions hints at a more nuanced role for transporters expressed in metabolic tissue. As Mdr49 is expressed in both midgut and malpighian tubule, it may very well be that the transporter is simultaneously importing and exporting insecticides depending on the tissue in question. Differences among insecticides in terms of overall resistance could be explained by the particular kinetics of the insecticide. The far larger spinosad may rely on active uptake from the midgut far more than a smaller nitenpyram, which could feasibly diffuse into the body and would be far more influenced by transport in the tubules. These proposals are nothing more than speculation, but the studies required to examine the movement of insecticides and metabolites around the insect body and the impact on resistance, can easily be achieved using the genetic toolkit in D. melanogaster.

5.2.4.4) The Integration of Metabolism and Transport

Finally, one of the biggest mysteries regarding insecticide metabolism is the interaction if any between the processes of insecticide modification and insecticide transport. In Chapter III, it was demonstrated that increased expression of P450s underpinned differences in the amount of metabolites excreted, but that these differences were not found in the larval bodies (Fig 3.9; Table 3.5). However, when Cyp6g1 was removed from RAL_517, there were less metabolites in the body and media (Fig 3.13). This discrepancy can be attributed to variation in transportation mechanisms present among DGRP genotypes, compared to the lack of such variation in the controlled background of the Cyp6g1 knockout. Most notably is the case of RAL_509, which had high levels of metabolism, but was hyper-susceptible due to its inability to efflux metabolised imidacloprid.

Whether ABC transporters (active transport) influence this difference is unclear. Mdr65 did not appear to have a high affinity for imidacloprid based on data from Chapter IV, but it is possible that it or other ABCs would have a higher affinity for metabolites of imidacloprid or other insecticides and could then act in a synergistic way with P450 mediated metabolism. There is precedent for this elsewhere as the mammalian ABC protein MRP was shown to have a higher affinity for drug metabolites produced by glutathione-S transferase (Keppler et al., 1999). However, there is a difference here in that glutathione-s-transferase metabolites typically have a glutathione molecule attached, which may the common factor that allows MRP to recognize them.
Alternatively, the metabolites could travel passively out of the body and bypass active transport.
Indeed, more hydrophilic compounds are generally more easily excreted from the mammalian body.
No SNP within any ABCB transporter appeared to be associated with imidacloprid response, even
when rare variants not included in the initial GWAS were considered. However, it is possible that
such variation has gone unnoticed in the DGRP if improperly annotated.

5.2.5) Non-Paradigmatic Genes

Thus far, insecticide target receptors, P450s, and ABC transporters have been the focus of
the discussion about insecticide biology. These categories represent the mechanisms, which likely
have the largest effect on resistance or carry the least fitness cost relative to the resistance they
provide. In other words, they are the most likely mechanisms for resistance to evolve in field
settings. However, they do not tell the whole story. The GWAS and TWAS candidate genes from
Chapter II demonstrate that a field population such as the DGRP has a gradient of resistance levels
and that this gradient can be explained by many genetic variants each of which contribute only
slightly to the overall phenotype. However, these genes were ignored in favour of the lower hanging
fruit, extending our understanding of *Cyp6g1*, already proven to have a role in imidacloprid
metabolism. This is a common phenomenon in the academic scientific community pressured by
time constraints and the inability to publish negative results. It is somewhat ironic that GWAS that
are initiated with the goal of finding unknown genes, can end with a focus on the known. However
in this case, the previously unidentified genes implicated, such as *Sickie* and *Tweek*, are difficult to
study because they appear to be lethal when knocked out completely and hypomorphs show severe
fitness defects (Abe et. al 2014, Verstreken et. al 2009). Further this type of variation is not what
was observed in the DGRP, but if the observed variation were regenerated in a defined genetic
background it may not produce a detectable resistance phenotype because of a small effect size or
the phenotype observed in the DGRP was background specific. Validating the observations for
genes such as *Sickie* and *Tweek* and gaining a mechanistic understanding of their role in
imidacloprid resistance proved to be a bridge too far. Nonetheless, a case for widespread, subtle
modifications in CNS structure and function playing a role in the response to imidacloprid was
established by the weight of numbers of CNS candidates identified (Table 3.3).

By way of contrast to *Sickie* and *Tweek*, the complete removal of genes such as
*Cyp6g1* or *Mdr65* showed no noticeable fitness effects. This is likely due to where these gene
products lie in the Drosophila regulatory hierarchy. While *Sickie* and *Tweek* are involved in
processes (axon growth, synaptic transmission) that affect so many other processes downstream,
*Cyp6g1* and *Mdr65* are not known to be involved in any essential endogenous processes (although
this has not been studied in detail). If the role of a gene is restricted to responding to xenobiotics, then these genes become very easy to study because they are endpoints, regulated by upstream processes, but only acting in a very circumscribed way. However, limiting the scope of studies in this way has severely limited the discussion of the phenotype of insecticide resistance to the description of a small subset of genes which only tell a fraction of the total story. Because the CNS genes were not pursued functionally, firm conclusions could not be made about whether particular variant was involved in imidacloprid resistance, and if so, what it was doing. To overcome the difficulties of strengthening the case for genes exemplified by Sickie and Tweek, GWAS on other populations may be a useful first step. The finding of the same candidates in independent studies would strengthen the case for their involvement in the imidacloprid response.

What can be obtained from the data presented is an estimation of how complex the biology of insecticide resistance is. Far from being influenced by only a handful of receptors, P450s and transporters, there are a myriad of other factors that make a substantial contribution to the level of resistance of a genotype. These factors are not classified as resistance genes either because they contribute in very small ways to the overall phenotype, or because the alleles of the genes necessary for resistance carry large fitness costs. Either way, their lack of ability to confer a detectable difference in resistance levels in a field setting has led to their being overlooked by researchers in favour of lower hanging fruit. In applied settings this makes sense due to the economic relevance of the variation that shifts resistance levels above thresholds of insecticide concentration applied. But to really understand the genetic basis of insecticide resistance and its potential to evolve it must be understood that, like all phenotypes, it is influenced by variation in many genes. It is enmeshed in an almost infinitely complex genetic and physiological network; subtle variations in individual genes will contribute very slightly to the overall resistance phenotype, but summed or synergized the contribution can be substantial.

5.3) The Genetic Concepts Emphasized by This Work

5.3.1) Introduction

Understanding the biology of insecticide resistance is of critical importance, but the more abstract genetic concepts touched on in this work also merit further discussion. A transition can be made from addressing the question how do insects interact with insecticides to the broader question How is the genetic basis of traits, such as insecticide resistance, described. When determining the genetic basis for traits, the goal is to first identify which loci in the genome contribute to the overall
trait and to then attempt to quantify the contribution of each. Broadly, two strategies for accomplishing this will be discussed. The first involves associating pre-existing genetic and phenotypic variation in order to implicate genes in a phenotype. The other employs transgenic techniques to manipulate individual candidates to see how these manipulations influence the overall phenotype.

5.3.2) Implication/Measurement by Association

5.3.2.1) Variations in Assays and Analysis can Implicate Different Genes

Focusing on the association approach first, the situation must first be imagined where every gene in the genome (~15,000 in the case of D. melanogaster) may have the capacity to influence a given trait. The task is to then narrow that number down as much as possible, identifying genuine candidate genes. This is done first and foremost by associating QTLs or transcripts with phenotypic variation as was accomplished in the GWAS and TWAS performed in Chapter III. Each candidate is described as having a probability (p-value) of influencing a phenotype, and the magnitude of their contribution can be estimated by modelling the effect of having a particular allele or amount of transcript. However, the determination of which QTLs and transcripts count as “significantly associated” is ill defined. Once a list of candidates is chosen, a further ambiguity exists as to which candidates to pursue functionally (Section 5.2.5.1). Furthermore, the choice of bioassay and method of analysis (Section 1.5) can have profound effects in determining which loci are implicated. Increasing the number of assays and analyses will accordingly increase the number of true positives but also the number of false positives. The upshot of this is that association studies should be seen as starting points for downstream analysis rather than definitive proof that a particular gene does or does not influence a trait.

5.3.2.2) The Pros and Cons of Association

Despite being a common technique to implicate genes in insecticide resistance, the utility of association studies is sometimes overstated. Association does not test a gene’s involvement in a phenotype, but rather its association with a phenotype. Factors such as linkage disequilibrium and transcriptional co-regulation of transcripts can confound seemingly significant results. Correlation does not equal causation, and the failure to understand this can lead to false positives. An association detected in the current study between RMR value at 25ppm and copy number variation at the Cyp12a4/Cyp12a5 locus proved to be spurious as transgenic expression yielded no significant difference in phenotype (Denecke unpublished). False negatives can also be present as not every QTL or transcript in the genome has sufficient genetic variation to be significantly associated with
any phenotype Indeed, genes that have vital functions and a role in the process under investigation are less likely to be detected in GWAS because of a lack of variation. As a result the usefulness of techniques such as GWA has been questioned by some (Visscher et al., 2012). Although most (including the author) do not doubt the usefulness of association studies, the claims made be such studies must be circumscribed and their significant findings validated by functional genetics. Exaggeration of significance and effect sizes notwithstanding, association is a valuable tool to describe the genetic basis of a trait, and the ability to examine the genetic basis of a trait in an unbiased way is indispensable. Psychological studies have found that conformation bias is an issue in the scientific community and the ability to remove any a priori assumptions with a forward genetic screen is a powerful way to implicate unknown candidate genes and to confirm previously implicated candidates.

5.3.3) Implication/Measurement by Functional Genetics

5.3.3.1) Overexpression

One of the most direct ways to functionally implicate a gene in a given phenotype is to overexpress it transgenically. In D. melanogaster, this makes use of the GAL4-UAS system. Even though this is often employed to mimic naturally occurring situations, this need not always be the case. Genotypes with AA or BA Cyp6g1 haplotypes show increased expression of the gene in the digestive tissues of the midgut Malpighian tubules and the fat body, and it was shown that this conferred insecticide resistance (Chung et al. 2007). However, nothing was known about how the upregulation of such a gene in other tissues could affect insecticide resistance. Such a case is easy to imagine because different enhancers usually drive the expression of genes in different tissues. In chapter II, an alternative evolutionary scenario was simulated in which expression of Cyp6g1 was enriched only in the CNS (Fig 2.8). This scenario is not unprecedented as an individual P450 overexpressed in the CNS was found to underpin field derived resistance to deltamethrin in Tribolium castaneum (Zhu et. al 2010). However, while overexpression permits the simulation of evolutionary scenarios it is limited by the types of genes it can test. P450s are particularly good models because resistance is usually mediated by overexpression. Furthermore, overexpression does not measure the contribution of a gene in a field population and should be seen as a measure of what the gene product can do, not what it is doing in a given organism.

5.3.3.2) Background

To measure the contribution of an identified gene to a phenotype, the gene must be subtracted from the genome either by knockdown or knockout. The various pros and cons of such
methods have previously been discussed (Section 1.7.3), and comparing case and control remains the only conclusive way to measure an allele’s contribution to a phenotype. What must be emphasized though is the fact that this measurement can only take place within a defined genetic background that includes the particular allele of the gene being subtracted. In chapter III, Cyp6g1 was removed from 3 different backgrounds and only the RAL_517-Cyp6g1KO showed any phenotypic difference from its control (Fig 3.11, 3.13). This is almost certainly due to the high levels of expression found only in the RAL_517 background, but it stands as a lesson for how important the genetic background is. Another example can be found with the Drosophila P-gp orthologue Mdr49, of which only one allele was capable of conferring resistance to DDT (Seong et. al 2016). Different versions of the same gene often have distinct properties and deleting them may have very different effects. Furthermore, even when the same allele is considered, gene products do not operate in vacuums and epistasis must be considered when characterizing the role of a gene in a phenotype.

Slightly more serious is the common practice of failing to control for genetic background when possible or ignoring differences in background as unrelated to the phenotype. For instance, in one of the original studies implicating the Mdr65 gene in drug transport, a heterogeneous background was used (Mayer et. al 2009). While this is not a problem with alleles of extremely large effect, more subtle differences can be masked by genetic differences apart from the loci in question. For example, removal of the same nAChR subunits from multiple backgrounds often finds a degree of background dependent effects (Luong unpublished). Until recently, these practices were unavoidable due to the logistical limitations of techniques such as “ends out” or TALENS. However, with recent developments in CRISPR-Cas9 precisely controlling for background is only dependent on checking for off target effects, and properly balancing crosses when transferring chromosomes. The standard of precision in genetic manipulation has been raised, and it is important for new studies to strive to minimize any undefined differences between case and control. Genetic background issues have been much harder to resolve in pest systems, meaning that it is impossible to attribute the fold difference in resistance between resistant and susceptible strains to a single gene. As CRISPR-Cas9 becomes increasingly available in pest systems, this issue can be addressed.

5.3.3.3) Assessing Resistance versus Probing Biology

The findings from chapters III and IV are similar in some ways. The former used a forward genetic screen and the latter used a literature search to find a gene that was thought to contribute to insecticide resistance. The gene was then removed from defined genetic background and the resistance of the knockout was compared to that of the control. However, it is important to
distinguish between the resistance caused by Cyp6g1 and the biological role played by Mdr65. The former does not play any apparent role in insecticide biology unless it is overexpressed, at which point the contribution is significant. Conversely, removing Mdr65 from a susceptible laboratory strain affected the resistance significantly. This would suggest that Mdr65 is playing a role in insecticide biology in general, while this would only be true for a subset of Cyp6g1 genotypes. True, Mdr65 has not been assessed in multiple backgrounds as Cyp6g1 was, but there is no indication that there is any derived Mdr65 allele present in Wxac.

This example highlights the difference between the biology and resistance of insecticides and has implications of how the future of the study of insecticide resistance will likely proceed. Until recently, the description of the resistance has been limited to what natural selection has provided, and, as a result, only a few genes comprising a few resistance mechanisms have been described in detail. This description is incomplete and many factors underlying resistance are unknown partly due to the masking effect of these more easily identified genes and mechanisms. Moving forward, I recommend that the insecticide resistance community begin to move towards a more proactive approach, systematically manipulating candidate genes that may not have been previously found in field populations. It is only through broadening our understanding of insecticide biology that more rational insecticide design and better pest management strategies can be achieved.

5.6) Concluding Remarks

Despite recent advances in non-insecticide pest control techniques (Sinkins and Gould, 2006), chemical insecticides remain the best weapon in the arsenal of society to ensure food security and public health. It is therefore crucial that a better understanding of resistance to insecticides is achieved. Although a prodigious amount of scientific literature has been devoted to trying to understand the genetics insecticide resistance, surprisingly little is known about which QTLs or transcripts are involved in resistance and how much they contribute to the resistance phenotype. This is concerning for two reasons. First, it makes it harder to manage resistance. Secondly, it means that there is little knowledge of the potential evolutionary pathways that may be taken by pests in the future, leaving us open to evolutionary blind sides from genes that were never previously on our radar. Deep probing of insecticide biology in model organisms remains one of the best ways to remedy this deficiency.

While this work has dealt primarily with the genetic basis for insecticide resistance, and should first and foremost be seen in this light, it also has a larger pharmacological purpose. Insecticides, like all drugs, act on and are acted upon by their host organisms in many different
ways. Variation in these interactions has the capacity to change how the host organism responds and the effectiveness of the treatment. New bioassays and the genetic techniques described in this thesis help us understand these interactions and the path that an insecticide takes from the time that it first enters the body until the time it is excreted or until the insect dies. In using insecticides, the goal is to eliminate the target organism, but the same kinetics apply to other pharmacological applications, such as the use of drugs in medicine.

In the broadest of all possible senses, this work, like so many before it, relates to evolution. Insecticide resistance allows for a unique opportunity to study the physiological and genetic mechanisms underpinning a recent, strong selective pressure. Furthermore, advances in biological and computational technology enables us to begin to predict how evolution will proceed, rather than looking back and trying to figure out what happened.
Bibliography


doi:10.1073/pnas.0502901102


144


## Appendix 1
### Primer Sequences

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence</th>
<th>Set</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mdr65KO_R</td>
<td>AGG TTC ACG ATG ACT GGT GG</td>
<td>Mdr65KO</td>
</tr>
<tr>
<td>Mdr65KO_F</td>
<td>ACC CGA ACG GTG CGA TAT ATG</td>
<td>Mdr65KO</td>
</tr>
<tr>
<td>Cyp6g1KO_F</td>
<td>CGAGATAAGATAATGCTGAG</td>
<td>Cyp6g1KO</td>
</tr>
<tr>
<td>Cyp6g1KO_R</td>
<td>ATTGCCCCCATCAACTAATCAC</td>
<td>Cyp6g1KO</td>
</tr>
<tr>
<td>Mdr65_PCFD4_F</td>
<td>ATATATAGGAAAAGATATCCCGGTT GAACTTCGATGGGAAGTCAGTG CGCTTTTTAGAGCTAGAAATAC AAG</td>
<td>Mdr65_PCFD4</td>
</tr>
<tr>
<td>Mdr65_PCFD4_R</td>
<td>ATATATAGGAAAAGATATCCCGGTT GAACTTCGATGGGAAGTCAGTG CGCTTTTTAGAGCTAGAAATAC AAG</td>
<td>Mdr65_PCFD4</td>
</tr>
<tr>
<td>Cyp6g1_PCFD4_F</td>
<td>TATATAGGAAAAGATATCCCGGTT GAACTTCGATGGGAAGTCAGTG CGCTTTTTAGAGCTAGAAATAGC AAG</td>
<td>Cyp6g1_PCFD4</td>
</tr>
<tr>
<td>Cyp6g1_PCFD4_R</td>
<td>ATTTTTACCTTGGCTATTTCCTAGCTC TAAAACGCATAGTGAGGCTTCG ACGACGTTAATGGAAATAGG TC</td>
<td>Cyp6g1_PCFD4</td>
</tr>
</tbody>
</table>
Author/s:
Denecke, Shane

Title:
Probing insecticide biology using Drosophila melanogaster

Date:
2017

Persistent Link:
http://hdl.handle.net/11343/191214

File Description:
Probing Insecticide Biology Using Drosophila melanogaster

Terms and Conditions:
Terms and Conditions: Copyright in works deposited in Minerva Access is retained by the copyright owner. The work may not be altered without permission from the copyright owner. Readers may only download, print and save electronic copies of whole works for their own personal non-commercial use. Any use that exceeds these limits requires permission from the copyright owner. Attribution is essential when quoting or paraphrasing from these works.