

Understanding insecticide detoxification in the leafcutter bee, *Megachile rotundata*

Submitted by Angela Hayward to the University of Exeter
as a thesis for the degree of
Doctor of Philosophy in Biological Sciences
In September 2021

This thesis is available for Library use on the understanding that it is
copyright material and that no quotation from the thesis may be
published without proper acknowledgement.

I certify that all material in this thesis which is not my own work has been
identified and that any material that has been submitted and approved for the
award of a degree by this or any other University has been acknowledged.

(Signature)

Abstract

Recent research on three managed bee pollinators, namely the Western honeybee (*Apis mellifera*), buff-tailed bumblebee (*Bombus terrestris*) and red mason bee (*Osmia bicornis*), has demonstrated that cytochrome P450 enzymes belonging to the CYP9Q and CYP9BU lineages provide protection to certain insecticides from three different mode of action classes. The alfalfa leafcutter bee (*Megachile rotundata*) is the world's most economically important solitary bee species. Nonetheless, it is unclear whether this species has a *CYP9Q/BU* ortholog that could afford similar levels of protection against insecticides, to those seen in other managed bee pollinators.

To address this question, the *M. rotundata* CYPome was curated and examined using phylogenetic and syntenic analyses. These investigations revealed that this species lacks a *CYP9Q/BU* ortholog or closely related P450 enzyme. Topical insecticide bioassays using *M. rotundata* determined that the species exhibits high sensitivity to all the compounds known to be detoxified by the *CYP9Q/BU* lineage. For example, *M. rotundata* is >2,500-fold more sensitive to thiacloprid than *A. mellifera*. Functional studies, using *M. rotundata* native microsomes, revealed no significant level of metabolism of any of the insecticides known to be detoxified by the *CYP9Q/BU* lineages. Radioligand competition assays, on head membrane preparations, showed no significant difference in binding affinity of neonicotinoid and butanolide insecticides at the nicotinic acetylcholine receptor. Taken together these findings indicate that the lack of a *CYP9Q/BU* ortholog, or closely related P450, in *M. rotundata* correlates with an inability of microsomal P450s to metabolise certain insecticides *in vitro* and a high sensitivity to these compounds *in vivo*.

To understand how wide spread the lack of a *CYP9Q/BU* ortholog might be, genomic and transcriptomic CYP9 sequences from 75 bee species were examined, using phylogenetic analyses. Five of the six bee families included in the phylogeny had genes that were closely related to, or shared a recent ancestor with *CYP9Q/BU* lineage (~97%). The Megachilidae showed a lower prevalence of *CYP9Q/BU* orthologs (50%), most notably in the two species of *Megachile*, where, based on phylogenetic and syntenic analyses, *CYP9DM*

genes have evolved in place of the *CYP9Q/BU* lineage. Sequencing of the transcriptomes of three UK and one Canadian *Megachile* species further substantiated that this entire genus (~1500 species) may have *CYP9DM* rather than *CYP9Q/BU* genes. Functional expression of this P450 suggested that, in contrast to several Megachilidae *CYP9BU*-like enzymes, *CYP9DM* P450s lack the capacity to bind neonicotinoids.

From the results generated in this thesis it is clear that the use of other managed bee species as a proxy for *M. rotundata* in ecotoxicological testing is unreliable. This has important implications for regulatory risk assessments. The data also illustrate the utility of using phylogenetic analyses, with targeted functional studies, as a tool to predict the level of sensitivity to insecticides of a bee species.

Table of contents

Abstract	2
List of tables	9
List of figures	12
Acknowledgements	21
List of Abbreviations	23
Chapter one: General introduction	26
1.1 Food security	26
1.1.1 <i>The essential role of pollinators in food security</i>	26
1.2. Megachile rotundata	27
1.2.1 <i>Description</i>	27
1.2.2 <i>Distribution and habitat</i>	28
1.2.3 <i>Life cycle</i>	28
1.2.4 <i>Nest recognition</i>	31
1.2.5 <i>Sex determination</i>	31
1.2.6 <i>Sex ratio</i>	32
1.2.7 <i>Diet</i>	32
1.2.8 <i>Agricultural importance</i>	33
1.3 Agricultural losses due to pest species	36
1.4 Chemical control of agricultural insect pests	37
1.4.1 <i>Insecticides targeting the GABA-gated chloride channel</i>	39
1.4.2 <i>Insecticides targeting acetylcholinesterases (AChEs)</i>	39
1.4.3 <i>Insecticides targeting voltage-gated sodium channels</i>	40
1.4.4 <i>Insecticides targeting nicotinic acetylcholine receptor (nAChR)</i>	41
1.4.5 <i>Insecticides targeting ryanodine receptor (RyR)</i>	45
1.5 Plant secondary metabolites	46
1.5.1 <i>Alkaloids</i>	46
1.6 Mechanisms involved in insecticide resistance and tolerance	47
1.6.1 <i>Metabolic mechanisms (pharmacokinetic mechanisms)</i>	48
1.6.2 <i>Target site modification (toxicodynamic mechanisms)</i>	57
1.6.3 <i>Physiological adaptation</i>	57
1.6.4 <i>Behavioural resistance</i>	58
1.7 Insecticide detoxification in bees	59
1.7.1 <i>Functional importance of CCEs in bees</i>	59
1.7.2 <i>Functional importance of GSTs in bees</i>	60
1.7.3 <i>Functional importance of P450s in bees</i>	60
1.8 Concerns and risks to pollinator health	65
1.8.1 <i>Habitat loss, degradation and loss of resource diversity</i>	65
1.8.2 <i>Agricultural intensification</i>	66
1.8.3 <i>Use of insecticides</i>	67
1.8.4 <i>Climate change</i>	70
1.8.5 <i>Pests and pathogens</i>	73

1.8.6 Risks pertinent to <i>M. rotundata</i>	74
1.9 Aims and objectives of this PhD	75
1.9.1 Chapter 3	76
1.9.2 Chapter 4	76
1.9.3 Chapter 5	76
1.9.4 Chapter 6	76
Chapter two: General materials and methods	77
2.1 Materials	77
2.1.1 Kits	77
2.1.2 Solutions and buffers	77
2.2 Molecular methods	77
2.2.1 Centrifugation	77
2.2.2 RNA extraction	78
2.2.3 First strand cDNA synthesis	78
2.2.4 DNA extraction	79
2.2.5 Polymerase chain reaction (PCR)	79
2.2.6 Agarose gel electrophoresis	81
2.2.7 PCR purification	81
2.2.8 PCR Cloning	82
2.3 Biochemical methods	84
2.3.1 Bradfords protein assay	84
2.4 Care and maintenance of insect cell lines	84
2.4.1 Initiation of insect cell lines from frozen stocks	84
2.4.2 Passaging cells	85
2.4.3 Determining cell density and viability	86
Chapter three: Curation of the CYPome of <i>M. rotundata</i> and comparison to other bee species	87
3.1 Introduction	87
3.1.1 Bee evolution	87
3.1.2 Comparative genomics	88
3.1.3 Molecular phylogenetic analysis	90
3.1.4 Chapter aims and underpinning questions	91
3.2 Methods	92
3.2.1 Curation of the CYPome of <i>M. rotundata</i>	92
3.2.2 Comparison of the CYPome of <i>M. rotundata</i> with other managed bee pollinators	93
3.2.3 Conserved synteny analysis	93
3.2.4 Investigating the structural homology of the CYP9 subfamily of P450s	94
3.2.5 Comparison of the structural homology of CYP9Q, CYP9BU and CYP9DM proteins	95
3.2.6 Distribution of the CYP9 subfamily across 12 available bee genomes	95
3.3 Results	96
3.3.1 Curation of the CYPome of <i>M. rotundata</i>	96
3.3.2 Comparison of the CYPome of <i>M. rotundata</i> with other managed pollinators	99
3.3.2.1 The CYP3 clan	102
3.3.3 Conserved synteny analysis	104
3.3.4 The structural homology of the CYP9 subfamily	110

3.3.5 Comparison of the structural homology of CYP9Q, CYP9BU and CYP9DM proteins	113
3.3.6 Distribution of the CYP9 subfamily across 12 available bee genomes	116
3.4 Discussion	119
3.4.1 The CYPome in <i>M. rotundata</i>	119
3.4.2 The CYP9 subfamily audit.....	122
Chapter four: Acute contact toxicity bioassays of select insecticides against <i>M. rotundata</i>	127
4.1 Introduction	127
4.1.1 Toxicity tests in <i>M. rotundata</i> to date.....	128
4.1.2 Acute contact toxicity test.....	129
4.1.3 Chapter aims and underpinning questions.....	130
4.2 Methods	130
4.2.1 Care and maintenance of bees	130
4.2.2 Insecticides.....	133
4.2.3 Preparation of bees.....	134
4.2.4 Range finding test	134
4.2.5 Test conditions.....	135
4.2.6 Analysis.....	136
4.2.7 Comparison with other managed bee pollinator species.....	136
4.3 Results.....	137
4.3.1 Range finding tests	137
4.3.2 Acute contact toxicity bioassay of <i>M. rotundata</i>	138
4.3.3 Comparison of <i>M. rotundata</i> tolerance to insecticides with three other managed bee pollinator species.....	141
4.4 Discussion	145
Chapter five: Functional characterisation of <i>M. rotundata</i> P450s	151
5.1 Introduction	151
5.1.1 Metabolism assays.....	152
5.1.2 Microsomes.....	153
5.1.3 Heterologous expression of <i>M. rotundata</i> P450s in insect cell lines.....	154
5.1.4 Assays with analysis using tandem LC-MS/MS.....	156
5.1.5 Model substrate profiling	156
5.1.6 Binding affinity of insecticides targeting the nicotinic acetylcholine receptor (nAChR)	157
5.1.7 Chapter aims and underpinning questions	157
5.2 Methods	158
5.2.1 Preparation of <i>M. rotundata</i> microsomes.....	158
5.2.2 Functional expression of <i>M. rotundata</i> P450s in insect cell lines.....	159
5.2.3 Metabolism assays with analysis using tandem LC-MS/MS.....	168
5.2.4 Model substrate profiling	169
5.3 Results.....	173
5.3.1 Screening male and female <i>M. rotundata</i> microsomes for metabolic activity	173
5.3.2 Expression of recombinant CYP9 P450s in an insect cell line	174
5.3.3 Metabolism assays with analysis using tandem LC-MS/MS.....	176
5.3.4 Model substrate profiling	182

5.3.5 Binding affinity of insecticides targeting the nicotinic acetylcholine receptor (nAChR)	187
5.4 Discussion	188
5.4.1 Preparation of functional microsomes from <i>M. rotundata</i>	188
5.4.2 Metabolic capability of native microsomes with analysis using LC-MS/MS	189
5.4.3 Model substrate profiling	196
Chapter six: Evolutionary analyses of the CYP9 subfamily of P450s across the Hymenoptera, with focus on the Anthophila	199
6.1 Introduction	199
6.1.1 Evolution of the Hymenoptera	199
6.1.2 P450 evolution	202
6.1.3 RNA sequencing and analysis	203
6.1.4 Chapter aims and underpinning questions	203
6.2 Methods	205
6.2.1 Surveying the available CYP9 subfamily sequences of bee species	205
6.2.2 Selection of candidate Megachilidae CYP9 sequences for functional expression	207
6.2.3 Model substrate metabolism assays	207
6.2.4 Surveying CYP9 subfamily sequences in representative Hymenoptera species	208
6.2.5 Conserved synteny analysis across the Hymenoptera	209
6.2.6 Transcriptomics from four species of Megachile genus bees	209
6.3 Results	213
6.3.1 The phylogeny of the CYP9 subfamily across bee families	213
6.3.2 Selection of candidate Megachilidae CYP9 sequences for functional expression	221
6.3.4 Fluorescence-based model substrate assays	223
6.3.5 The phylogeny of the CYP9 subfamily across the Hymenoptera	229
6.3.6 Conserved synteny analysis of the membralin /alpha catulin-associated CYP9 cluster across Hymenoptera	237
6.3.7 Transcriptomics results from four species of Megachile bees	240
6.4 Discussion	247
6.4.1 The phylogeny of the CYP9 subfamily across bee families	247
6.4.2 Fluorescence-based model substrate assays	248
6.4.3 Phylogenetic and syntenic analyses of the CYP9 subfamily across the Hymenoptera	250
6.4.4 Transcriptomics results from four species of Megachile bees	251
Chapter seven: General discussion	253
7.1 Understanding the molecular determinants of insecticide sensitivity in <i>M. rotundata</i>	254
7.2 Phylogenetic and syntenic analyses of the CYP9 subfamily of P450s	257
7.3 Metabolic profiling of recombinant Megachilidae CYP9BU-like enzymes	262
7.4 The Megachile CYP9DM enzymes	264
7.5 Implications	268
7.6 Applications	269
7.7 Future work	271
Appendices	274

Appendix chapter two	274
Appendix chapter three.....	276
Appendix chapter four.....	286
Appendix chapter five	290
Appendix chapter six.....	298
<i>Bibliography</i>	310

List of tables

Table 1.1: Insecticides recommended for insect management on crops pollinated by <i>M. rotundata</i>	34
Table 1.2: <i>M. rotundata</i> pollination services	35
Table 1.3: Evidence for P450 involvement in xenobiotic metabolism in bee pollinators	62
Table 2.1: Cycling conditions used for PCR using Taq polymerase (times for DreamTaq Green PCR Master Mix (2X) (Thermo Scientific) and MyTaq TM Red Mix (Bioline Meridian Bioscience) shown in black and red respectively)	80
Table 2.2: Cycling conditions used for PCR using Phusion [®] high-fidelity DNA polymerase (New England BioLabs Inc.)	81
Table 2.3: Insect cell line frozen stocks	85
Table 3.1: Description of the CYPome of <i>M. rotundata</i>	97
Table 3.2: Macro-synteny of the CYP9 and CYP6AS clusters between the managed bee pollinators. [CYP6AS data for <i>O. bicornis</i> not included as <10% conserved pairs found across 12 scaffolds]	106
Table 3.3: Genetic neighbourhood of <i>CYP9DN1</i> in two Megachilidae and one Halictidae bee species. Flanking genes found across all three species shaded dark grey and those found in two species light grey.	107
Table 3.4: Percent identity with <i>A. mellifera</i> protein sequences for the primary flanking genes from the syntenic block that includes the CYP9 cluster.	109
Table 4.1: EPA classification categories of pesticide toxicity in bees based on LD ₅₀ values [493].	130
Table 4.2: Scoring criteria for acute contact toxicity test using <i>M. rotundata</i> (based on [484].	136
Table 4.3: Concentration of insecticides used in range finding tests	137
Table 4.4: Concentration range of insecticides used in <i>M. rotundata</i> acute contact bioassay	138
Table 4.5: Response of <i>M. rotundata</i> to topically applied neonicotinoid insecticides at 48 h	138
Table 4.6: Response of <i>M. rotundata</i> to topically applied pyrethroid insecticides at 48 h.	139
Table 4.7: Response of <i>M. rotundata</i> to topically applied organophosphate insecticides at 48 h	140
Table 4.8: Difference in LD ₅₀ values between the four managed bee pollinators of select insecticides.	144
Table 5.1: Compounds used in metabolism assays	169
Table 5.2: The chemical structure, molecular weight and excitation/emission wavelengths (nm) of P450 fluorescent model substrates used in this study ...	170
Table 5.3: Expression of recombinant P450s using Gateway [®] cloning technology and transfection in insect cell lines.	175
Table 5.4: Welch's <i>t</i> -test results for metabolism of synthetic insecticides by native microsomal preparations of <i>M. rotundata</i> as measured by LC-MS/MS (1 h incubation at 30°C ±NADPH).	178
Table 5.5: Results for univariate analysis of variance using general linear model for metabolism of synthetic insecticides by native microsomal preparations of <i>M.</i>	

<i>rotundata</i> as measured by LC-MS/MS (1 h incubation at 30°C ±NADPH). Tests of between-subject effects with <i>post-hoc</i> multiple comparisons of depletion using a Bonferroni adjustment.....	179
Table 5.6: Welch's <i>t</i> -test results of percent depletion of parent compound of alkaloid allelochemicals by native microsomal preparations of <i>M. rotundata</i> as measured by LC-MS/MS (1 h incubation at 30°C ±NADPH).....	181
Table 5.7: Results for univariate analysis of variance using general linear model for metabolism of alkaloid allelochemicals by native microsomal preparations of <i>M. rotundata</i> as measured by LC-MS/MS (1 h incubation at 30°C ±NADPH). Tests of between-subject effects with <i>post-hoc</i> multiple comparisons of depletion using a Bonferroni adjustment.	181
Table 5.8: Model substrate standard curves	183
Table 5.9: Reduction of HC production (%), after incubation (1 h) with the neonicotinoid insecticides thiacloprid (TCP) and imidacloprid (IMI).	186
Table 6.1: List of 75 species of bee used in phylogenetic analyses, with brief notes on family, subfamily, life history and ecology	214
Table 6.2: Number of sequences of each CYP9 lineage from 75 species of bee across six families (all data taken from NCBI genome and transcriptome Shotgun Assembly (TSA) databases).	215
Table 6.3: Partial sequences, found during BLASTn searches of the NCBI databases, for bee species without a CYP9P representative in the Bayesian inferred phylogeny [445] of the CYP9 cluster (figure 6.7).	217
Table 6.4: Partial sequences, found during BLASTn searches of the NCBI databases, for bee species without a CYP9Q/BU/DL representative in the Bayesian inferred phylogeny [445] of the CYP9 cluster (figure 6.7).....	220
Table 6.5: Expression of recombinant P450s using Gateway® cloning technology and transfection of insect cell lines.	221
Table 6.6: Reduction of HC production (%) by Megachilidae recombinant CYP9s, after incubation (1 h), with the neonicotinoid insecticides thiacloprid (TCP) and imidacloprid (IMI).	225
Table 6.7: Kinetics data for MC/EC metabolism (resulting in HC) by Megachilidae CYP9s co-incubated with thiacloprid or imidacloprid	227
Table 6.8: List of species of wasp, ant and sawfly species included in phylogenetic analyses.	229
Table 6.9: Numbers of wasp and ant CYP9 sequences that share an ancestral node with the CYP9Q/BU/DL clade in bees.	235
Table 6.10: Scaffolds from the genomes of wasp, ant and sawfly species containing syntenic blocks of genes also found in bees.....	238
Table 6.11: Percent identity with <i>A. rosae</i> protein sequences for the primary flanking genes from the syntenic block that includes the CYP9 cluster across Hymenopteran species. Data generated by alignment of protein sequences in Geneious using MUSCLE (version 3.5, default settings) [451].	240
Table 6.12: Inventory of the lineages of the CYP9 subfamily of <i>A. mellifera</i> , <i>B. terrestris</i> , <i>O. bicornis</i> , <i>D. novaeangliae</i> and five species of <i>Megachile</i> bees. .	241

Table 6.13: Percentage identity and number of non-synonymous changes between the CYP9R sequences from the RNAseq data for *M. centuncularis*, *M. lapponica*, *M. leachella* and *M. willughbiella*.243

List of figures

Figure 1.1: Adult <i>M. rotundata</i> (a): female; (b): male. Image (a) credit Pitts-Singer, T. US department of agriculture. Public domain.	28
Figure 1.2: Diagram of the life cycle of <i>M. rotundata</i> showing the univoltine and bivoltine cycles.	31
Figure 1.3: Total weight of pesticide used globally shown per annum (1990-2018) in tonnes. [Data taken from: Food and Agriculture Organisation of the United Nations [114] (FAOSTAT accessed 16/09/20). Pesticides include: insecticides, herbicides, fungicides, bactericides, rodenticides, disinfectants, mineral oils and plant growth regulators].....	37
Figure 1.4: 2D chemical structure of two organophosphate insecticides (a): Chlorpyrifos (CPS) [139], (b): Coumaphos (CMP) [140].	40
Figure 1.5: 2D chemical structure of three pyrethroid insecticides (a): Deltamethrin (DMT) [147], (b): <i>tau</i> -Fluvalinate (τ -FLV) [148], (c): <i>alpha</i> -Cypermethrin (α -CP) [149].	41
Figure 1.6: 2D chemical structure of Nithiazine [159], the negatively charged nitro-group, also present in the <i>N</i> -nitroguanidine neonicotinoid insecticides, is outlined with a red dashed line.	43
Figure 1.7: 2D chemical structure of neonicotinoid insecticides. (a): Thiacloprid [160], (b): Acetamiprid [161], (c): Imidacloprid [162], (d): Clothianidin [163]. The negatively charged cyano-groups are outlined with a solid red line and negatively charged nitro-groups with a dashed red line.	43
Figure 1.8: 2D chemical structure of the butanolide insecticide Flupyradifurone (FPF) [172].	44
Figure 1.9: 2D chemical structure of two diamide insecticides. (a): Flubendiamide [181] and (b): Chlorantraniliprole [182].	46
Figure 1.10: 2D chemical structure of alkaloid allelochemicals (a): Nicotine (NCT) [189], (b): Atropine (APN) [190], (c): Cytisine (CTS) [191], (d): Anabasine (ABS) [192].	47
Figure 1.11: The main routes of 'xenobiotic compound-insect' interactions. Interactions shown: (i) avoidance, (a behavioural response - genetic or learned); (ii) penetration; (iii) three-phase metabolic process (iv) excretion (both metabolite and unaltered substrate) (v) sequestration. Yellow squares annotated with an X indicate a xenobiotic compound; target sites (shown in purple) annotated with a W indicate 'wild-type' receptors and those marked M indicate 'mutant-type' receptors with the potential to confer resistance.	49
Figure 1.12: Scheme of accepted P450 nomenclature. Amended from [218]...	51
Figure 1.13: Generalised primary structure of a microsomal CYP, showing conserved and variable regions, with highly conserved amino acids in bold [218, 243, 244].	53
Figure 1.14: Tertiary structure of <i>Homo sapiens</i> CYP4A3 (PDB: 4D6Z [254]) (a): Secondary structures displayed; α -helices coloured green and β -sheets cyan. The conserved cysteine is coloured yellow and marked with a C (b): Conserved motifs (M) are coloured red and substrate recognition sites (SRSs) blue. The	

crystal structure CYP4A3 (PDB: 4D6Z [254]) is rendered with the surface depicted in mesh. [Figure created using UCSF Chimera version 1.10.1]	55
Figure 1.15: Global land use for food production. Amended from [315].....	67
Figure 1.16: Total incidents of bee poisonings attributed to pesticide each year (1988-2019) in the UK. Data represented by a yellow circle are for the UK and are taken from [325, 326]; data represented by a blue square are for England and Scotland only and are taken from [327, 328, 330]. Graph drawn in GraphPad Prism V 8.1.0.	68
Figure 1.17: Insecticide used in tonnes per annum in the United States of America, the European Union and the United Kingdom (1990-2018). [Data taken from: Food and Agriculture Organisation of the United Nations [114] (FAOSTAT accessed 16/09/20); graph drawn in GraphPad Prism V 8.1.0.	70
Figure 1.18: The potential routes of exposure to insecticides to <i>M. rotundata</i> (solid lines represent primary routes and dashed lines secondary routes). Amended from [396].	75
Figure 3.1 (a): Evolutionary history of bees showing the topology of relationships between families. Genera shown: <i>Apis</i> [401], <i>Bombus</i> [402], <i>Melipona</i> [403], <i>Eufreisea</i> [404], <i>Habropoda</i> [405], <i>Megachile</i> [406], <i>Osmia</i> [407], <i>Heriades</i> [408], <i>Andrena</i> [409], <i>Melitta</i> [410], <i>Dioxys</i> [411], <i>Chelostoma</i> [412], <i>Hylaeus</i> [413], <i>Colletes</i> [414], <i>Ctenocolletes</i> [415], <i>Stenotritus</i> [416], <i>Dufourea</i> [417], <i>Nomia</i> [418], <i>Lasioglossum</i> [419], <i>Panurgus</i> [420], <i>Camptopoeum</i> [421], <i>Macropis</i> [422], <i>Dasyopoda</i> [423] (b): Bar chart showing approximate number of species for each bee family. Number of species per family taken from [424].	88
Figure 3.2: Examples of synteny blocks (SBs) showing conserved synteny (**) and conserved linkage (*) with possible scenarios that lead to breaks in SBs. Genes are denoted with numbers, and species by a super-script letter (^a / ^b). Orthologous genes are linked by a dashed line. The minimum number of orthologs that constitute an SB is shown here as three.	89
Figure 3.3: Phylogeny of <i>M. rotundata</i> CYPome, rooted on camphor hydroxylase (P450cam; <i>P. putida</i>). Phylogeny estimated using PhyML Maximum likelihood algorithm [453], with branch support of 200 bootstraps shown as %. Blue shading denotes a pseudogene; ancestral node of the CYP bloom in the CYP6AS subfamily marked with a yellow circle and the CYP9 subfamily with a red circle.	98
Figure 3.4: Distribution of the CYP clans in <i>M. rotundata</i> compared with three insect species (<i>A. mellifera</i> [211]; <i>D. melanogaster</i> [222] & <i>A. gambiae</i> [212])	99
Figure 3.5: Heat map of the distance matrix (percentage identity) of the CYPomes across four species of managed bee pollinators (<i>A. mellifera</i> ; <i>B. terrestris</i> ; <i>O. bicornis</i> and <i>M. rotundata</i>). Matrix generated using Bayesian inference estimation [Chain length: 1,100,000; Subsampling frequency: 200; Burn-in length: 100,000; Heated chains: 4; Heated chain temperature: 0.2] ..	100
Figure 3.6: Bayesian inference phylogeny of the CYPomes of four managed bee pollinators (<i>A. mellifera</i> ; <i>B. terrestris</i> ; <i>O. bicornis</i> and <i>M. rotundata</i>), using substitution model LG+G [454]. Sequence names coloured by species: <i>A. mellifera</i> - red; <i>B. terrestris</i> - green; <i>O. bicornis</i> – light blue and <i>M. rotundata</i> –	

dark blue. Posterior probability of nodes shown as a % probability. Tree rooted on camphor hydroxylase (P450cam; *P. putida*). All protein sequences accessed from NCBI protein database..... 101

Figure 3.7: Distribution of CYP9 subfamily across four species of managed bee pollinators. 103

Figure 3.8: Structure of the transcribed region of *CYP9DN1* in two Megachilidae species. Exons are shown in red and introns in blue. 103

Figure 3.9: A map of macro-synteny between *A. mellifera* DH4 linkage group LG14 (Amel_HAV3.1) in comparison to five scaffolds from the genome of *M. rotundata* (scaffolds: 1303, 0244, 0120, 0464 and 0030). *A. mellifera* CYP9 members are shown in red in the list of genes. Large blocks of conserved genomic content and gene order can be seen, although there is also evidence of inversions and translocations. The region 500Kbp upstream and downstream of the CYP9 cluster is mapped to three scaffolds in *M. rotundata* (scaffolds: 1303, 0244 and 0120). Evolutionary divergence between Apidae and Megachilidae ~104-125mya. 105

Figure 3.10: (a) Phylogeny of CYP9 amino acid sequences from six bee species (*A. mellifera*; *B. terrestris*; *O. bicornis*, *M. rotundata*, *D. novaeangliae* and *C. gigas*). Phylogeny estimated using PhyML Maximum likelihood algorithm [453] and substitution model LG+G [454], with branch support of 50 bootstraps, shown as %, rooted on *Nasonia vitripennis* CYP9AG4. Scale bar represents 40 substitutions per 100 residues. Branch lengths, where shown, are in italics. Each monophyletic CYP9 lineage is denoted by colour and the ancestral node is marked with a circle. Sequence name coloured by family. (b) Syntenic relationship at the CYP9 loci in six bee species across four families (schematic representation only, not to scale). CYP9 genes are coloured by lineage. Arrows denote reading frame. 108

Figure 3.11: Tertiary structure of *H. sapiens* CYP4A3 (PDB: 4D6Z [254]), showing conserved protein sequence in the CYP9 subfamily in four managed bee pollinators (*A. mellifera*; *B. terrestris*; *O. bicornis* and *M. rotundata*). Distal and proximal faces of the molecule are included in the figures. (a) Conserved motifs (M) are coloured red and substrate recognition sites (SRSs) blue. The crystal structure CYP4A3 (PDB: 4D6Z [254]) is depicted with a solid surface. (b) Conservation of CYP9 genes across four managed bee pollinators. Level of conservation is depicted in a magenta to cyan (conserved to variable) gradient. (c) Multiple sequence alignment (MSA) of CYP9 genes from four managed bee pollinators. Conserved motifs are enclosed in red and SRSs in blue. The conserved cysteine is enclosed in yellow. Figure x(a) and (b) created using UCSF Chimera version 1.10.1. 111

Figure 3.12: Conservation of CYP9 protein lineages across four managed bee pollinators (*A. mellifera*; *B. terrestris*; *O. bicornis* and *M. rotundata*). Level of conservation is depicted in a magenta to cyan (conserved to variable) gradient on crystal structure *H. sapiens* CYP4A3 (PDB: 4D6Z [254]). Percent identity shown is to *A. mellifera* sequences (*CYP9R1*, *CYP9P1* and *CYP9Q3*). Figure created using UCSF Chimera version 1.10.1..... 112

Figure 3.13: Multiple sequence alignment of *A. mellifera* and *B. terrestris* CYP9Qs with *O. bicornis* CYP9BUs and *M. rotundata* CYP9DMs [aligned in Geneious version 10.2.3 (Biomatters) using MUSCLE [451] (version 3.5, default settings). The sequences are coloured black to white according to their similarity. Conserved motifs (M) and substrate recognition sites (SRS) are shaded red and blue respectively and represented by annotations below chain A of the crystal structure CYP4A3 (PDB: 4D6Z [254]). Secondary structures are annotated above chain A 4D6Z; dark cyan cylinders and dashed black boxes represent α -helices [with reference to the crystallographic structures of PDB: 4D6Z; 1TQN (CYP3A4: *H. sapiens* [250, 254]) and P450cam [248]]. Amino-acid substitutions and *M. rotundata* specific gaps are highlighted in orange..... 114

Figure 3.14: (a) Ribbon diagram of *H. sapiens* CYP3A4 (PDB: 4D6Z [254]), showing the secondary elements. β -sheets are coloured purple and coils are grey. Helices are coloured pale blue apart from helix I and helix L which are bright blue and cyan respectively. The highly conserved cysteine is coloured yellow (b) The P450 active site/heme-binding pocket, created by helix L and the central region of helix I. The heme group is depicted in light brown with a dark orange sphere marking the central iron (Fe) atom. The highly conserved cysteine is coloured yellow, other conserved residues are coloured red. (c) The P450 active site/heme-binding pocket with amino acid substitutions in CYP9DMs (*M. rotundata*) compared to CYP9Q/BUs coloured green. Figure created using UCSF Chimera version 1.10.1..... 115

Figure 3.15: Phylogeny of CYP9 amino acid sequences from 12 bee species: *A. mellifera*; *A. cerana*; *A. dorsata*; *A. florea*; *B. terrestris*; *B. impatiens*; *D. novaeangliae*; *E. mexicana*; *H. laboriosa*; *M. quadrifasciata*; *M. rotundata* and *O. bicornis*. Phylogeny estimated using PhyML Maximum likelihood algorithm [453] and substitution model LG+G [454]. The branches show relative time. . 118

Figure 4.1: *M. rotundata* cocoons for incubation (groups of approximately 60 cocoons per box)..... 131

Figure 4.2: Incubator set up showing the sequential groups of cocoons (coloured by week)..... 132

Figure 4.3: Female *M. rotundata* in holding cages (with enrichment and feeders) 133

Figure 4.4: Experimental set up for female *M. rotundata* bioassay 134

Figure 4.5: Experimental set up for range-finding test using male *M. rotundata* 135

Figure 4.6: 48 h acute contact toxicity dose-response curves for one nitro-group imidacloprid (IMI) and two *N*-cyano-group neonicotinoid insecticides thiacloprid (THC) and acetamiprid (ACE) against *M. rotundata*. Slope values for the steepest part of the dose-response curve are shown for each compound..... 139

Figure 4.7: 48 h acute contact toxicity dose-response curves for two pyrethroid insecticides, deltamethrin (DMT) and *tau*-fluvalinate (τ -FLV) against *M. rotundata*. Slope values for the steepest part of the dose-response curve are shown for each compound. 140

Figure 4.8: 48 h acute contact toxicity dose-response curves for two organophosphate insecticides coumaphos (CMP) and chlorpyrifos (CPS)

against <i>M. rotundata</i> . Slope values for the steepest part of the dose-response curve are shown for each compound.	141
Figure 4.9: LD ₅₀ values for <i>M. rotundata</i> (48 h) after exposure to topically applied thiacloprid (TCP), imidacloprid (IMI) and acetamiprid (ACE), compared to data for three other managed bee pollinators (<i>A. mellifera</i> , <i>B. terrestris</i> and <i>O. bicornis</i>); error bars indicate 95% CIs. Sensitivity thresholds are depicted according to the EPA toxicity ratings [493] (neonicotinoid data for <i>A. mellifera</i> , <i>B. terrestris</i> and <i>O. bicornis</i> taken from [137, 155, 168, 226, 446, 463])......	142
Figure 4.10: LD ₅₀ values for <i>M. rotundata</i> (48 h) after exposure to topically applied <i>tau</i> -fluvalinate (τ -FLV) and deltamethrin (DMT), compared to data for three other managed bee pollinators (<i>A. mellifera</i> , <i>B. terrestris</i> and <i>O. bicornis</i>); error bars indicate 95% CIs. Sensitivity thresholds are depicted according to the EPA toxicity ratings [493] (pyrethroid data for <i>A. mellifera</i> , <i>B. terrestris</i> and <i>O. bicornis</i> taken from [137, 145, 282, 463])......	143
Figure 4.11: LD ₅₀ values for <i>M. rotundata</i> (48 h) after exposure to topically applied coumaphos (CMP) and chlorpyrifos (CPS), compared to data for three other managed bee pollinators (<i>A. mellifera</i> , <i>B. terrestris</i> and <i>O. bicornis</i>); error bars indicate 95% CIs. Sensitivity thresholds are depicted according to the EPA toxicity ratings [493] (organophosphate data for <i>A. mellifera</i> , <i>B. terrestris</i> and <i>O. bicornis</i> taken from [29, 137, 138, 463])......	144
Figure 5.1: Selection of <i>M. rotundata</i> P450s for heterologous expression.....	155
Figure 5.2: Enzyme activity in microsomal preparations using male, female and dissected female heads and thoraxes of <i>M. rotundata</i> . Activity is given in relative fluorescence units (RFU) based on screening against MC and EFC. Data are mean values \pm SD (n=3). Analysis performed using a Welch's t-test (two-tailed; degrees of freedom = 2.044) with significant differences indicated by *** $p < 0.0005$	174
Figure 5.3: Metabolism of the neonicotinoid insecticides thiacloprid (TCP), acetamiprid (ACE), imidacloprid (IMI), clothianidin (CTN) and the butenolide insecticide flupyradifurone (FPF) by native microsomal preparations of <i>M. rotundata</i> as measured by LC-MS/MS (1 h incubation at 30°C \pm NADPH). The error bars indicate 95% CIs.	176
Figure 5.4: Metabolism of the pyrethroid insecticides <i>tau</i> -fluvalinate (τ -FLV), deltamethrin (DMT) and <i>alpha</i> -cypermethrin (α -CMT) by native microsomal preparations of <i>M. rotundata</i> as measured by LC-MS/MS (1 h incubation at 30°C \pm NADPH). The error bars indicate 95% CIs.	177
Figure 5.5: Metabolism of the organophosphate insecticides coumaphos (CMP), Chlorpyrifos (CPS) and the diamide insecticide flubendiamide (FBD) by native microsomal preparations of <i>M. rotundata</i> as measured by LC-MS/MS (1 h incubation at 30°C \pm NADPH). The error bars indicate 95% CIs. Analysis performed using a Welch's t-test (two-tailed) with significant differences indicated by * $p < 0.05$	178
Figure 5.6: Metabolism of select alkaloid allelochemicals by native microsomal preparations of <i>M. rotundata</i> as measured by LC-MS/MS (1 h incubation at 30°C \pm NADPH). The error bars indicate 95% CIs. Analysis performed using a	

Welch's t-test (two-tailed) with significant differences in percent depletion of parent compound indicated by **p<0.01 and ***p<0.001. 180

Figure 5.7: CYP9R1 enzyme activity (fluorescence) against the model substrate EFC, incubated at 25, 30, 35 and 40°C, over 180 minutes. Data points are mean values (n=8). 182

Figure 5.8: Metabolic activity of male *M. rotundata* microsomes against coumarin- and resorufin-derived fluorescent model substrates. Data points are mean values ±SD (n=3). Analysis performed using a Welch's t-test (two-tailed; degrees of freedom = 2.241) with significant differences indicated by ***p<0.0002 and ****p<0.0001. 184

Figure 5.9: (a) Metabolic activity of recombinant CYP9 P450s against coumarin-derived fluorescent model substrates. Data points are mean values ±SD (n=3) CYP9 P450s are coloured by CYP9 lineage. (b) Schematic of the phylogenetic relationship between the CYP9 lineages. (c) Heatmap showing the metabolic activity of recombinant CYP9 P450s against coumarin-derived fluorescent model substrates. 185

Figure 5.10: Insecticide mediated inhibition of 7-hydroxycoumarin (HC) formation by recombinantly expressed CYP9 enzymes incubated with different concentrations of thiacloprid (TCP) and imidacloprid (IMI). Data are mean values ± S.D. (n=4)..... 187

Figure 5.11: Binding affinity (IC₅₀ values) of selected insecticides to nAChR head membrane preparations of *A. mellifera* and *M. rotundata*. Data for *A. mellifera* taken from [446]. Figure taken from Hayward *et al.*, 2019..... 188

Figure 6.1: The evolutionary history of the Hymenoptera, showing phylogenetic relationships, branches not to scale. Estimates of divergence times associated with key evolutionary events are indicated at their respective nodes [585, 587]. Triangular branches indicative of extant species not drawn to scale. NCBI genome assembly numbers as of December 2020. Photos used: *Athalia rosae* [590], *Cephus cincta* [591], *Cotesia vestalis* [592], *Nasonia* sp. [593], *Microplitis demolitor* [594], *Chrysis* sp. [595], *Polistes rothneyi* [596], ant [597]. 200

Figure 6.2: (a) and (b) Male willowherb leafcutter bee (*M. lapponica* Thomson, 1872), photos show the long wings which have a smoky subhyaline appearance. (c) Male *M. rotundata* photo shows the shorter wings with less prominent veining. 210

Figure 6.3: (a) Female patchwork leafcutter bee (*M. centuncularis* Linnaeus, 1758) [622]. (b) Female silvery leafcutter bee (*M. leachella* Curtis, 1828) [623]. (c) Female Willughby's leafcutter bee (*M. willughbiella* Kirby, 1802) [624]. 211

Figure 6.4: The number of contigs and the sequence length for each *Megachile* species dataset. 212

Figure 6.5: Schematic showing the synteny of the CYP9 cluster in bees across four families. 213

Figure 6.6: (a) Bayesian inference phylogeny [445] of the CYP9 subfamily across 75 species of bee, using substitution model LG+G [454]. Posterior probability of nodes shown as a % probability. Tree rooted on *N. vitripennis* CYP9AG4. Sequences are coloured by bee family and annotated with an abbreviated form of the family name. Single sequence branches are labelled

with a circle, coloured by family. All nucleotide sequences accessed from NCBI databases. (b) Schematic of the phylogenetic relationship between bee families.	216
Figure 6.7: Bayesian inference phylogeny [445] of the CYP9 cluster across 75 species of bee, using substitution model LG+G [454]. Posterior probability of nodes shown as a % probability. Tree rooted on <i>N. vitripennis</i> CYP9AG4. Sequences are coloured by bee family and annotated with an abbreviated form of the family name. Single sequence branches are labelled with a circle, coloured by family. All nucleotide sequences accessed from NCBI databases. (b) Schematic of the phylogenetic relationship between bee families.	218
Figure 6.8: (a) Bayesian inference phylogeny [445] of the CYP9Q/BU/DL clade across 75 species of bee, using substitution model LG+G [454]. Posterior probability of nodes shown as a % probability. Tree rooted on <i>N. vitripennis</i> CYP9AG4. Sequences are coloured by bee family and annotated with an abbreviated form of species name (see table 6.1). All nucleotide sequences accessed from NCBI databases. (b) Schematic of the phylogenetic relationship between bee families. (c) Annotations indicating Megachilidae CYP9BU-like sequences selected for functional expression.	222
Figure 6.9: Model substrate profile of Megachilidae CYP9BU-like expressed recombinant proteins against coumarin-derived fluorescent model substrates.	224
Figure 6.10: (a) Insecticide mediated inhibition of 7-hydroxycoumarin (HC) formation by recombinantly expressed Megachilidae CYP9BU-like enzymes incubated with different concentrations of thiacloprid (TCP) and imidacloprid (IMI). Data are mean values \pm S.D. (n=4). (b) Schematic of the phylogenetic relationship between the CYP9BU genes.	228
Figure 6.11: (a) PhyML maximum likelihood phylogeny [453] (substitution model JTT+G [627] with branch support of 50 bootstraps), of the CYP9DN lineage, across the Order: Hymenoptera. Tree rooted on <i>M. domestica</i> CYP4ae1. Sequences are coloured by superfamily (see table 6.8), apart from bees which are coloured by family. Nodes that mark the divergence of superfamilies are denoted by coloured circles. All nucleotide sequences accessed from NCBI databases. (b) Schematic of the phylogenetic relationship of the Hymenoptera. Photos used: <i>Athalia rosae</i> [590], <i>Cotesia vestalis</i> [592], <i>Nasonia</i> sp. [593], <i>Microplitis demolitor</i> [594], <i>Chrysis</i> sp. [595], <i>Polistes rothneyi</i> [596], ant [597]. (c) Schematic of the phylogenetic relationship between bee families.	231
Figure 6.12: (a) Bayesian inference [445] phylogeny of the CYP9 cluster across Hymenoptera species, using substitution model JTT (G+I) [627]. Posterior probability of nodes shown as a % probability. Tree rooted on <i>M. domestica</i> CYP4ae1. Sequences are coloured by superfamily, apart from bees which are coloured by family. All nucleotide sequences accessed from NCBI databases. Nodes that mark the divergence of CYP9 lineages are denoted by coloured circles. (b) Schematic of the phylogenetic relationship of the Hymenoptera. Photos used: <i>A. rosae</i> [590], <i>C. vestalis</i> [592], <i>Nasonia</i> sp. [593], <i>M. demolitor</i> [594], <i>Chrysis</i> sp. [595], <i>P. rothneyi</i> [596], ant [597]. (c) Schematic of the phylogenetic relationship between bee families.	233

Figure 6.13: (a) PhyML maximum likelihood phylogeny [453] (substitution model JTT+G [627]) of the CYP9Q/BU/DL lineage, across the Order: Hymenoptera. Tree rooted on *M. domestica* CYP4ae1. Sequences are coloured by superfamily (see table 6.8), apart from bees which are coloured by family. Ancestral node of bees coloured with a red circle. All nucleotide sequences accessed from NCBI databases. (b) Schematic of the phylogenetic relationship between bee families. (c) Schematic of the phylogenetic relationship of the Hymenoptera. Photos used: *A. rosae* [590], *C. vestalis* [592], *Nasonia* sp. [593], *M. demolitor* [594], *Chrysis* sp. [595], *P. rothneyi* [596], ant [597]..... 236

Figure 6.14: Conserved tandem arrangements of the membralin/ alpha catulin-associated CYP9 loci across the Hymenoptera. Photos used: *A. mellifera*, [401], *M. rotundata* [406], *C. costatus* [628], *C. floridanus* [629], *P. dominula* [630], *V. mandarinia* [631], *N. vitripennis* [632], *T. pretiosum* [633], *Athalia rosae* [590], *Neodiprion lecontei* [634], *Cephus cinctus* [591]. 239

Figure 6.15: PhyML maximum likelihood phylogeny [453] (using substitution model LG+G [454], branch support of 50 bootstraps) of the CYP9 subfamily sequences from *A. mellifera*, *B. terrestris*, *O. bicornis*, *D. novaeangliae*, and five species of *Megachile* bee. Tree rooted on *N. vitripennis* CYP9AG4. Apidae and Halictidae sequences coloured by family: red and dark cyan respectively. Megachilidae bees coloured by genus, *Osmia* and *Megachile*: light blue and dark blue respectively. Branch lengths (substitutions per site) shown for ancestral nodes of each CYP9 lineage. 242

Figure 6.16: Heat map of the percent identity of *CYP9DM* sequences from five species *Megachile*. 244

Figure 6.17: Multiple sequence alignment of: *A. mellifera*, *B. terrestris*, *O. bicornis* and *D. novaeangliae* CYP9Q/BU/DLs with *Megachile* species CYP9DM sequences. Aligned in Geneious version 10.2.3 (Biomatters) using MUSCLE [451] (version 3.5, default settings). The sequences are coloured black to white according to their similarity. Conserved motifs (M) and substrate recognition sites (SRS) are shaded red and blue respectively and represented by annotations below *A. mellifera* CYP9Q3. Secondary structures are annotated: dark cyan cylinders and dashed black boxes represent α -helices (with reference to the crystallographic structures of PDB: 4D6Z; 1TQN CYP3A4: *H. sapiens* [250, 254] and P450cam [248]). Amino-acid substitutions and *Megachile* specific gaps are highlighted in orange. 245

Figure 7.1: Representation of the phylogenetic relationship in the Megachilidae family, showing the position of the genera represented in the phylogenetic analyses in chapter six [640]. Branches are coloured by currently recognised tribes [72]. Ancestral nodes leading to species with CYP9 enzymes capable of binding neonicotinoid insecticides are marked with a red circle. 259

Figure 7.2: Ribbon diagram of conserved motifs and active site/heme-binding pocket of *H. sapiens* CYP93A4 (PDB: 4D6Z [254]). (a) Conserved motifs (M1-5) coloured red to show similarity, with respect to *A. mellifera*, *B. terrestris*, *O. bicornis* CYP9Q/BU enzymes, conserved cysteine (C) shown in yellow and the heme- group in green. (b) Conserved motifs (M1-5) coloured red and blue to indicate similarity or difference between CYP9Q/BU and CYP9DM enzymes

from five species *Megachile*. Figure created using UCSF Chimera version
1.10.1.266
Figure 7.3: Suggested framework or 'tool-kit' to predict bee sensitivity to
insecticides. Boxes (1) - (3) show the pipe-line of the comparative genomics,
heterologous expression of P450 enzymes and metabolic profiling that make up
the 'tool-kit'.270

Acknowledgements

First and foremost, I would like to say a massive thank you to my supervisor Chris Bass, for his patience, knowledge, invaluable advice and encouragement. Thank you for always having the time to talk and for allowing me to explore my ideas throughout this PhD study. I would also like to thank my second supervisor, Jeremy Field for all his support, particularly for his expertise in collecting and identifying specimens. I am also very grateful to my colleagues at Bayer, especially Ralf Nauen and Julian Haas, for the many insightful discussions, support and excellent collaboration. This PhD was funded by BBSRC and Bayer CropScience and I am grateful to both bodies for the opportunity and support.

A huge thank you goes to Kat Beadle who helped me settle in during my first year and who mentored me, patiently giving me advice on anything from lab basics, to bee care and handling – I am indebted and very grateful Kat. I'd also like to thank Emma Randall for her support with everything in my first couple of years. Special thanks go to Bartek Troczka for his generosity with his time, patience, knowledge and doughnuts over the years – I have learned so much and am very grateful, Bartek. To Mark Mallott, a huge thank you for the advice and support, but also for your unfailing kindness and sense of humour. I'd also like to mention Ben Hunt for his bioinformatics expertise. Lastly, to Ellie Bushnell a thank you for the friendship and for being my writing buddy.

I would like to recognise the entire Bass lab and the crew at the SERSF PhD office, past and present, who were always there to provide support, advice and humour. You all deserve a mention, so here goes – Adam, Amy, Ana, Andy, Charlie, Emma, James, Jo, Kumar, Nasser, Vicky, Zoltan and anyone else I may have forgotten! You all made it a friendly and inclusive place to work.

I would like to say thank you to Professor Mick Bailey – for all the help and support, but mostly for believing in me before I believed in myself. A huge thank you must go to all my friends in Somerset, especially Kathy for her proof reading and punctuation skills, but also Shannon, Debs, Sean, Pete, Tim and Theo – I couldn't have made it without you.

Most importantly, I am grateful to my family, especially my children Daniel and Hope for their unconditional and unequivocal support. They encouraged me to take this opportunity, giving me the courage to move to Cornwall by myself and so, I would like to dedicate this thesis to them. Thank you for always listening to me and for believing I could do this.

Lastly, I have to thank Dotty the cat for keeping me sane during a pandemic, two lock-downs and the months of writing up.

List of Abbreviations

aa	Amino acid
ACE	Acetamidiprid
ACh	Acetylcholine
BFC	7-Benzoyloxy-4-(trifluoromethyl)-coumarin
BI	Bayesian inference
BLAST	Basic local alignment search tool
BSA	Bovine serum albumin
BOMFC	7-(benzyloxymethoxy)-4-trifluoromethylcoumarin
BOMR	(benzyloxymethoxy)resorufin
bp	Base pairs
C	Celsius
CCE	Carboxyl/cholinesterase
cDNA	Complementary deoxyribonucleic acid
CI (95%)	Confidence interval (95%)
CMP	Coumaphos
α-CMT	<i>alpha</i> -Cypermethrin
CPR	Cytochrome P450 reductase
CPS	Chlorpyrifos
CTN	Clothianidin
CYP	Cytochrome P450
CYPome	Cytochrome P450 complement
DMSO	Dimethyl sulfoxide
DMT	Deltamethrin
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
DTT	Dithiothreitol
EC	7-ethoxy-coumarin
EDTA	Ethylenediaminetetraacetic acid
EFC	7-ethoxy-4-trifluoro-methylcoumarin
EFSA	European food safety authority
EPA	U.S. environmental protection agency
ER	Ethoxy-resorufin
FBD	Flubendiamide

FBS	Fetal bovine serum
τ-FLV	<i>tau</i> -Fluvalinate
FPF	Flupyradifurone
g	Gram
GC content	Guanine-cytosine content
g-force	Relative centrifugation force
GST	Glutathione-S-transferase
HC	7-Hydroxycoumarin
HFC	7-hydroxy-4-(trifluoromethyl)-coumarin
HR	7-Hydroxy-resorufin
IC₅₀	Half maximal (50%) inhibitory concentration
IMI	Imidaclopid
IPM	Integrated pest management
IRAC	Insecticide resistance action committee
Kb	Kilo base pairs
LB	Lysogeny (Luria) broth
LC-MS/MS	Liquid chromatography-mass spectrometry (with two mass spectrometry detectors)
LD₅₀	Dose required to kill 50% of individuals in a given population
MC	7-methoxy-coumarin
MEGA	Molecular Evolutionary Genetics Analysis
MFC	7-methoxy-4-(trifluoromethyl)coumarin
mg	Milligram
mM	Millimole
MoA	Mode of action
MOBFC	7-(4-methoxybenzyloxy)-4-trifluoromethylcoumarin
MR	Methoxy-resorufin
mRNA	Messenger ribonucleic acid
MUSCLE	Multiple Sequence Comparison by Log-Expectation
nAChR	Nicotinic acetylcholine receptor
NADPH	Nicotinamide adenine dinucleotide phosphate
NCBI	National centre for biotechnology information
ng	Nanogram

OECD	Organisation for Economic Co-operation and Development
OOMR	(octyloxymethoxy)resorufin
OP	Organophosphate
ORF	Open reading frame
P450	Cytochrome P450
PBO	Piperonyl butoxide
PCR	Polymerase chain reaction
PHYML	Phylogenetic maximum likelihood
pmol	Picomole
POLO	Probit or Logit analysis
ppb	Parts per billion
ppm	Parts per million
PR	Pentoxo-resorufin
RFU	Relative fluorescence units
RNA	Ribonucleic acid
rpm	Revolutions per minute
SE	Standard error
SOC	Super optimal broth with catabolite repression
TAE	Tris-acetate-EDTA
TCP	Thiacloprid
T_m	Primer melting temperature
μg	Microgram
μM	Micromole
V	Voltage

Chapter one: General introduction

1.1 Food security

The food and agriculture organization of the United Nations (FAO) defines food security as: *'a situation that exists when all people, at all times, have physical, social and economic access to sufficient, safe and nutritious food that meets their dietary needs and food preferences for an active and healthy life'* [1]. Over the last 50 years the Green Revolution [2] has enabled global food product to double [3, 4]. This increase in food production allowed for explosive growth in the global human population [5, 6]. In 1950 the human population was 2.5 billion, but it is projected that by the middle of this century that figure will have reached 9-10 billion [5-7]. In the last three decades, the global prevalence of undernourishment has dropped from 18.6% to 8.9% [8, 9]. However, approximately 60 million more people have become affected by hunger since 2014, with 690 million people classified as undernourished in 2019 [9]. Projected figures suggest the COVID-19 pandemic may add a further 83-132 million people to the figures for 2020 [9].

1.1.1 The essential role of pollinators in food security

Animal pollinators play a crucial ecological role and are vital to ensure not only human food security, but also the stability of natural ecosystems [10, 11]. Crops, such as fruit, vegetables, seeds and nuts, depend on animal pollinators and account for 35% of global food production [10, 12]. The economic value of the 105 most common crops that rely on insect pollination is put at US\$800 billion per annum [13]. Many crops that are vital sources of micronutrients essential to human health, such as folate and vitamin A, are specifically dependent on insect pollination [14]. Micronutrient deficiencies in human populations could therefore be related to a disruption in crop pollination [14, 15]. The vast majority of these insect pollinators are non-managed, wild taxa and include over 20,000 species of bee [10, 12].

There is growing concern about the reported decline of insect pollinator species and the resulting potential impacts on biodiversity levels and future food security [16-20]. Pollinator declines have been reported in all continents, except Antarctica [21]. As an example, 9.2% of wild bee populations in Europe have

been given the IUCN Red List status of ‘threatened’, and there is also good evidence for a global decline in species of bumblebee (*Bombus*) [22-24]. Changing land-use, agricultural intensification and use of insecticides are considered by many to be the drivers of these declines [22, 25-27]. The Western honeybee (*Apis mellifera* Linnaeus) is one of the best studied insect species [28]. The possible impact of insecticides on the health of honeybee colonies has been the subject of many scientific publications [29-31] and much media interest [32-34]. There is, however, uncertainty over whether the data gathered by monitoring *A. mellifera* can be applied to wild bee populations [28, 35-38]. Loss of insect biodiversity is occurring within the wider context of climate change and both of these factors are expected to exacerbate food insecurity, particularly in the developing world, by disrupting food availability, food accessibility and food utilisation [39-42].

Any increase in the risk of crop loss or failure has the potential to threaten global food security [43]. Increases in food prices due to lack of supply have been linked to social unrest and ‘food riots’ [44, 45]. To combat the joint risks of food insecurity and social discord, farmers need the necessary tools to protect their crops, secure a harvest and safeguard their livelihoods [12, 46]. However, the diversity, abundance and health of insect pollinators must be considered in the production of any pesticide [12, 47].

1.2. *Megachile rotundata*

1.2.1 Description

M. rotundata is among the smallest leafcutting bees with females having an average body weight of 35 mg [48]. They are considerably smaller than *A. mellifera* (hatching adult worker bee weight 100 mg [49]). *M. rotundata* is sexually dimorphic with females being 1.2 - 1.3 times larger than males (females 8 – 9 mm; males 7 – 8 mm long [50]). Both sexes are dark grey in colour, but females have black eyes and white setae (see figure 1.1(a)), whereas males have blueish eyes, yellow/cream spots on the abdomen and yellow setae on the face (see figure 1.1(b)) [48, 50, 51]. Females are longer lived than males: 30 - 60 days compared to 15 - 35 days [48]. *M. rotundata* is a solitary species but prefers to nest gregariously in above ground cavities, such as hollow twigs, crevices in trees or holes in buildings [52]. They also readily

accept man-made nesting holes making it an ideal species for commercial pollination services [52, 53].



(a)

(b)

Figure 1.1: Adult *M. rotundata* (a): female; (b): male. Image (a) credit Pitts-Singer, T. US department of agriculture. Public domain.

1.2.2 Distribution and habitat

M. rotundata is native to south-western Asia and south-eastern Europe [54]. The species was introduced to North America accidentally one or more times in the 1930's [52]. The subsequent use of the species in commercial pollination has led to it being introduced across the Americas, Australia and New Zealand, where it is still currently used as a managed pollinator [52, 55, 56]. *M. rotundata* is not managed for pollination services in its native range [57]. However, European alfalfa production only accounts for around one-fifth of world production whereas North America produces two-thirds of world alfalfa seed [53, 57]. The yield of European alfalfa seed is comparable to that found in the Americas which implies native, feral *M. rotundata* provide adequate pollination [53, 57].

1.2.3 Life cycle

1.2.3.1 Emergence and mating

M. rotundata is a summer-flying bee; the adults naturally emerge in the spring or summer months as temperatures increase and allow for the completion of development [53, 55]. Like most bees *M. rotundata* is protandrous (males emerging before females [48]). Individual emergence is regulated by the

development stage of their nest mates in the tubular nest, with emerging bees waiting for the siblings ahead of them to leave before exiting themselves [58]. This allows for the exodus from nests to be orderly and diminishes the risk that individuals will chew through their siblings as they emerge [58]. Female *M. rotundata* are monogamous, mating once almost immediately after they emerge, thereafter actively resisting further attempts at copulation [48, 59, 60]. All females will have sperm in their spermatheca after a week [61]. Males can inseminate multiple females and they patrol both nest sites and flowers looking for receptive females [59]. While the eggs are developing the female will select a nesting site and begin to construct and provision the first cell within 24 - 48 hours [48, 62].

1.2.3.2 Nest construction

M. rotundata constructs tubular-shaped nests in cavities, such as cracks or hollows in trees, plant stems and reeds [53]. They prefer cavities that are about the same size as their bodies (4.5 – 6 mm diameter [63]). As their common name implies, female leafcutting bees use their mandibles to cut leaf pieces from nearby plants to line the nesting tunnel they have selected [48]. Walls, partitions and plugs are constructed and glued together using a salivary secretion produced by chewing the edges of the plant tissue to a sticky pulp [48, 53]. *M. rotundata* prefers soft, pliable leaves and will readily use alfalfa and canola making them useful for commercial pollination [48, 56]. The nesting material of *M. rotundata* forms a hydrophobic layer and may show anti-microbial action, thus possibly serving a similar protective function to the lining secreted by Dufour's gland in other bee species [64, 65]. Each cell requires 14 - 15 leaf pieces and under ideal foraging conditions is constructed in approximately 2.5 hours [48, 66].

Once the first cell is lined the female begins to provision it with nectar and pollen [48, 53]. *M. rotundata* has a short flight range and when the female is collecting provisions she forages within a few hundred feet of the nest site [67]. In early foraging trips the female carries a high ratio of pollen to nectar (80:20), but the proportion of nectar increases with subsequent trips until the final trip where only nectar is brought back [66-68]. She then backs into the tunnel and lays a single egg on the provision mass [66]. The cell is then capped with

circular leaf pieces [53]. Once the first cell is complete a second cell is initiated in front of it, and this continues in a linear fashion until the tunnel is almost full [66, 67]. Depending on the length of the cavity, 8 - 12 cells are constructed in a tunnel [67]. When the last cell is finished, a plug made of 10 - 50 circular pieces of leaf cemented together is constructed that seals the entrance [67]. The female will then select and construct another nest tunnel [53]. Under ideal conditions females can lay two eggs a day and complete 57 cells with eggs in their lifetime (Maeta & Kitamura, 2005 cited in [56]). When food resources are scarce fewer cells are made [69].

1.2.3.3 Larval development

M. rotundata are holometabolous and have 5 larval instars [62]. The life cycle of *M. rotundata* is shown in figure 1.2. During development through the 3rd and 4th instar stages, the larvae continue to feed but do not defecate [56, 62]. The fifth and final instar is mobile and will consume the remaining provision and defecate in preparation to spin a cocoon and enter diapause [62]. As with other species of long-tongued bees, the cocoon is made from a silk-like material produced by the salivary glands [56, 62]. *M. rotundata* shows facultative bivoltinism and, with the correct cues, up to half of the offspring will be non-diapausing, or second generation bees (see figure 1.2), emerging in the summer of their natal year and needing as little as 5 days to pupate [58, 62]. Most bivoltinism will be seen in the first nests of the season [70]. The remaining offspring will overwinter as diapausing prepupae [56, 62]. The toughness of the cocoon and the cell walls, made up of leaf material, allow the prepupae to be safely harvested commercially using a cell remover and mechanised tumbler [48, 56, 71].

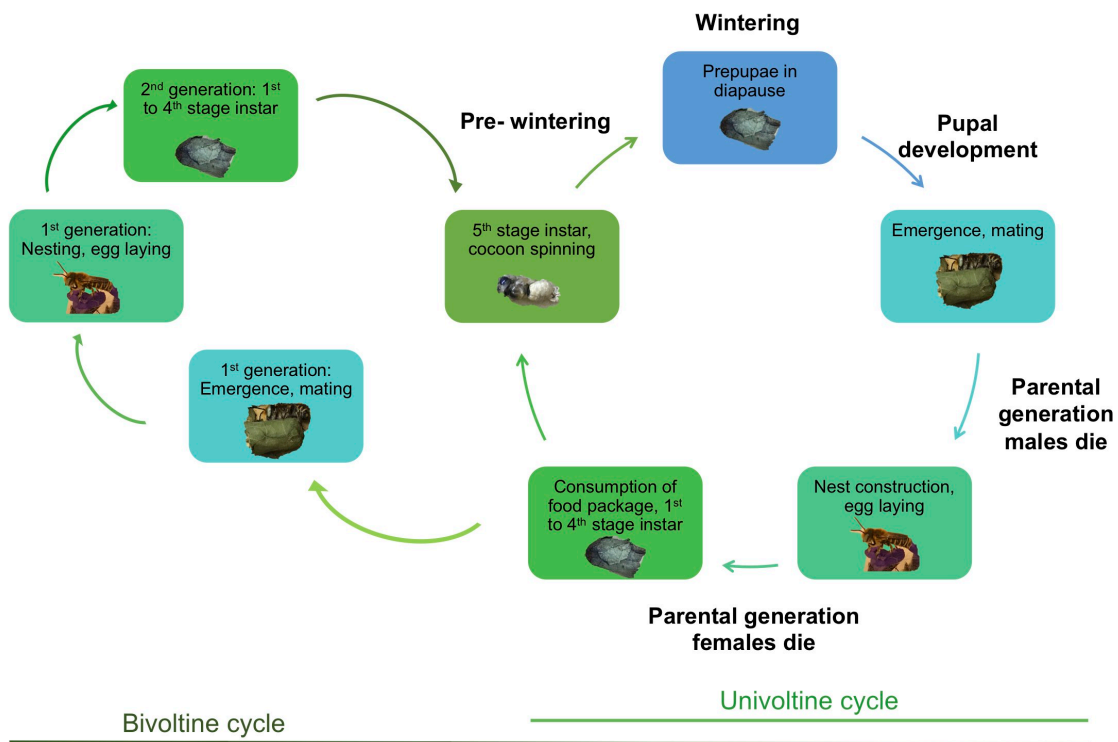


Figure 1.2: Diagram of the life cycle of *M. rotundata* showing the univoltine and bivoltine cycles.

1.2.4 Nest recognition

Like many other solitary bees species *M. rotundata* is gregarious and can nest in highly dense aggregations [72]. This necessitates that each female be able to recognise her own nest among those of her conspecifics [73, 74]. Female bees are adept at short-range orientation to their nests using both visual and olfactory cues [73-75]. On the way in and out of a nest female *M. rotundata* wipe the tunnel with their abdomens leaving olfactory cues that if removed cause confusion when the bee returns [75].

1.2.5 Sex determination

In common with most Hymenoptera, *M. rotundata* are haplodiploid [76]. Fertilized eggs get two sets of chromosomes (16 from each parent) and develop into diploid females. Unfertilized eggs only contain 16 maternal chromosomes and develop into haploid males [76]. The female controls fertilization during oviposition. She pauses her abdominal contractions briefly, presumably to allow the egg to receive sperm from the spermatheca, when depositing female eggs, but not with male eggs [60]. Female *M. rotundata* eggs are almost always laid in the inner cells of a tunnel and are provisioned with a greater supply of food [68].

Males are usually laid in the cells at the outer end of the tunnel and, on average, receive 17% less provision [68]. It may be that the depth of tunnel is associated with the stimulus to fertilize the egg [60]. Like many other species of Hymenoptera a small proportion of *M. rotundata* males will be diploid (2 - 7%) [77]. It is generally accepted, that with some exceptions, such as the solitary vespid wasp *Euodynerus foraminatus*, most diploid males in the Hymenoptera are sterile or, should they mate successfully, will produce sterile triploid progeny [78]. These individuals will have been provisioned at the same level as females and will attain a similar adult size and weight and so present a high cost to their parents [77, 78].

1.2.6 Sex ratio

Non-diapausing, second generation bees have a different sex ratio to spring-emergent generations, with more females emerging in the summer (64.5%) than in the spring (38.3%; [79]. Commercially, the optimal M:F ratio is considered to be 2:1, which is similar to the level found naturally in bees emerging from diapause in the spring [79, 80]. However, nesting conditions affect the sex ratio of adults and this is not always achieved [80]. The sex ratio of the population is affected by nesting tunnel length and diameter, which can create a bias towards males [63]. Tunnels with diameters of 4 mm can produce a skew to males of 11:1; whereas nests with diameters of 5.5 mm had ratios of 3:1 and those with 6 mm diameters obtained the commercially optimal level [63]. The optimal length of tunnel is 5 - 7.5 cm [63]. If the ratio is skewed towards males from the optimal level, sexual harassment of females has a negative effect on fecundity, until the point where male-male aggression minimises the effect (above 4:1) [81]. Availability of floral resource does not appear to affect sex ratio but when limited it does decrease overall fecundity and size of female offspring [69].

1.2.7 Diet

M. rotundata is a polylectic species, with individuals collecting pollen from taxonomically diverse plants, with a wide variation in flower structures [55, 82]. However, female *M. rotundata* have a strong preference for certain legumes, particularly those of the genus *Melilotus* and *Medicago* (which includes alfalfa [55]). When nest sites are placed for commercial pollination of an alfalfa crop

over 90% of the pollen loads of female *M. rotundata* will be alfalfa [82]. The female prefers to forage within a 50 m radius of the nest site she has selected and this will, in part, determine the plants that she visits [82]. However, when unmanaged the female can forage for a distance of up to 1.6 km [53]. Ogilvie and Forrest [83] define floral resource phenology as the floral species available to a bee population through time. *M. rotundata* has a relatively short flight-season and this, combined with a limited foraging range, means the species has restricted access to floral resources both spatially and temporally [53, 55]. This in turn may have affected the evolution of bee-plant interactions in the species [83].

1.2.8 Agricultural importance

1.2.8.1 Alfalfa production

As its name implies the alfalfa leafcutting bee is widely used in alfalfa production [57]. Alfalfa is the world's most important forage crop [52]. It is grown to produce hay that is fed to livestock and to produce seed for cultivation [52]. Alfalfa flowers have an explosive pollination system (characteristic of all *Medicago* species) in which the stamens and pistil, held under tension by the keel, flip upwards, releasing pollen when a bee trips the mechanism [48, 84]. Honey bees do not pollinate alfalfa effectively because they probe the flowers from the side, avoiding the tripping mechanism (22% visited flowers tripped); whereas, *M. rotundata* forces the keel apart to access nectar tripping up to 78% of visited flowers [48, 84]. Use of *M. rotundata* as a pollinator increases the potential yield of seed in Canada from 50 kg/ha to 1000 kg/ha [57].

Although there are important differences between the cultivation of alfalfa for the production of hay versus seed, the two crops share many insect pests, such as the alfalfa weevil (*Hypera postica*) and *Lygus* bugs [85]. In order to harvest seed, the alfalfa plant needs to mature beyond the vegetative plant stage at which hay is cut [85]. Management of a crop while it is in bloom necessitates the need for pest control to be balanced against pollinator health [85, 86]. The insecticides recommended for use in insect management where *M. rotundata* is used as a pollinator are shown in table 1.1. In general, alfalfa seed growers rely heavily on the chemical control of pests [85, 87]. The typical management program involves a pre-bloom 'clean-up spray' using a mixture of

organophosphate and pyrethroid insecticides (often dimethoate, chlorpyrifos, bifenthrin and/or *lambda*-cyhalothrin) [85, 87]. However, the lower-risk products novaluron, indoxcarb and flonicamid (see table 1.1) have now been registered for use on alfalfa and are beginning to be used as alternatives in the pre-bloom clean-up spray [87]. These pre-bloom treatments are applied prior to the placement of *M. rotundata* into the fields [85, 87]. After this point, any insecticide treatments involve short-residual or ‘low bee-risk’ compounds and are carried out during the evening, or at night, after the bees have stopped foraging [85, 87, 88]. In general, three to five pest control treatments are applied, per growing season, to a crop of alfalfa for seed production [85]. Flonicamid is often applied in evening application during bloom for the management of *Lygus* bugs and more recently sulfoxaflor has also been recommended as a reduced-risk insecticide for evening use during flowering [87]. Although the more general advice to producers is to read the pesticide labels thoroughly, with particular attention to the specific precautionary and advisory statements found in the Environmental Hazards section as this should contain the most up-to-date advice [88].

Table 1.1: Insecticides recommended for insect management on crops pollinated by *M. rotundata*

Insecticide	MoA	MoA group	Crop
<i>alpha</i> -Cypermethrin	Sodium channel modulator (Pyrethroid)	1	Alfalfa
<i>beta</i> -Cyfluthrin	Sodium channel modulator (Pyrethroid)	1	Alfalfa
Bifenthrin	Sodium channel modulator (Pyrethroid)	1	Alfalfa
Carbaryl	Acetylcholinesterase (AChE) inhibitor (Carbamate)	1	Alfalfa
Chlorantraniliprole	Ryanodine receptor modulator (Diamide)	1	Alfalfa
Chlorpyrifos	Acetylcholinesterase (AChE) inhibitor (Organophosphate)	1	Alfalfa
Clothianidin	Nicotinic acetylcholine receptor (nAChR) competitive modulator (Neonicotinoid)	1	Canola
Cyfluthrin	Sodium channel modulator (Pyrethroid)	1	Alfalfa
Deltamethrin	Sodium channel modulator (Pyrethroid)	1	Canola/ blueberries
Dimethoate	Acetylcholinesterase (AChE) inhibitor (Organophosphate)	1	Alfalfa
Flonicamid	Modulators of chlordontal organs	1	Alfalfa
Flubendiamide	Ryanodine receptor modulator (Diamide)	1	Blueberries
Flupyradifurone	Nicotinic acetylcholine receptor (nAChR) competitive modulator (Butenolide)	1	Alfalfa
<i>gamma</i> -Cyhalothrin	Sodium channel modulator (Pyrethroid)	1	Alfalfa
Imidacloprid	Nicotinic acetylcholine receptor (nAChR) competitive modulator (Neonicotinoid)	1	Canola
Indoxacarb	Voltage-dependent sodium channel blocker (Oxadiazine)	1	Alfalfa
<i>lambda</i> -Cyhalothrin	Sodium channel modulator (Pyrethroid)	1	Alfalfa
Methomyl	Acetylcholinesterase (AChE) inhibitor (Carbamate)	1	Alfalfa
Novaluron	Inhibitor of chitin biosynthesis affecting CHS1 (Benzoylueas)	2	Alfalfa/ Canola
Permethrin	Sodium channel modulator (Pyrethroid)	1	Alfalfa
Phosmet	Acetylcholinesterase (AChE) inhibitor (Organophosphate)	1	Canola/ blueberries
Spinetoram	Nicotinic acetylcholine receptor (nAChR) allosteric modulator - Site 1	1	Blueberries
Spinosad	Nicotinic acetylcholine receptor (nAChR) allosteric modulator - Site 1	1	Blueberries
Sulfoxaflor	Nicotinic acetylcholine receptor (nAChR) competitive modulator	1	Alfalfa
<i>zeta</i> -Cypermethrin	Sodium channel modulator (Pyrethroid)	1	Alfalfa

Targeted physiology MoA group 1: Nerve and muscle; MoA group 2: Growth and development. References: Alfalfa [89], Canola [90], Blueberries [91].

Despite the wide-spread use of *M. rotundata* as a managed pollinator, the guidance for reporting incidences of bee poisoning remains weighted towards *A. mellifera*, with references to hives, pollen, wax and honey in the official guidance on how to deal with a possible bee poisoning event [92-95]. However, it is clear that the signs of *M. rotundata* poisoning are well-known and distinctive [88, 96, 97]. A lack of female nesting activity and inordinate numbers of dead males on the ground in front of nesting boards are cited as the main indicators [88, 96, 97]. Unlike with *A. mellifera*, however, it is unclear from the publicly available records how common these type of incidences are [94].

1.2.8.2 Other crop plants and applications

M. rotundata is polylectic and their application as commercial pollinators is not limited to alfalfa [55]. They are efficient pollinators of many different types of flowering plant (see table 1.2). They have equal foraging rates to honey bees and, with certain flowers, increased pollen collection [98]. Their limited flight range makes them ideal pollinators for the production of hybrid seed and for use in enclosed spaces [99, 100].

Table 1.2: *M. rotundata* pollination services

Crop plant	Species	Comments
Canola (oilseed rape) [98]	<i>Brassica napus</i>	Equal foraging to <i>A. mellifera</i> ; increased pollen collection
Canola (oilseed rape) [100]	<i>Brassica napus</i>	Production of hybrid seed in tents
Wild carrot [99]	<i>Daucus carota</i>	Production of hybrid seed in tents
Annual clover [101]	<i>Trifolium</i> spp.	Production of forage crop
Perennial clover [102]	<i>Trifolium</i> spp.	Production of forage crop
Lowbush blueberries [103]	<i>Vaccinium angustifolium</i>	80% fruit set (<i>A. mellifera</i> - 59%)
Purple prairie-clover [104]	<i>Dalea purpurea</i>	Restoration of wild habitats
Prairie legumes [105]	<i>Dalea</i> spp.	Restoration of wild habitats

1.2.8.3 Commercial management

M. rotundata has been managed commercially in North America since the early 1960's [52]. The most common management system is the 'loose cell system', in which bee cells are removed from nesting boards and separated [106]. There have been many articles written over the decades that outline good management of *M. rotundata* for pollination services [48, 52, 106]. Advice on

best practices in regard to more current issues, such as biosecurity or pesticide use, is still regularly published [88, 107].

1.2.8.4 Economic importance

In terms of economic value, the pollination services provided by *M. rotundata* are only exceeded by those of the honey bee [55, 56]. The total value of alfalfa pollination by *M. rotundata* in the US, considering both hay and seed, is between \$5-7 billion per annum [56, 108]. Due to its' use as a commercial pollinator, the species is considered to be responsible for two-thirds of the world production of alfalfa seed [56].

1.3 Agricultural losses due to pest species

Since the origins of agriculture, farmers and gardeners have had to cope with pest species that impact the yield of crops grown for human use or consumption. Pest organisms can be pathogens (viruses, bacteria and fungi), animals (rodents, birds, insects and other invertebrates) or other plants [109]. To avoid such losses, pest species have been controlled manually (eg. removal of larvae by hand), biologically (eg. encouragement of predator species) and chemically (application of pesticides) [109]. Crop losses due to pests can be divided into pre-harvest and post-harvest losses [109, 110]. Pre-harvest loss estimates vary between crops and countries, but are generally put at between 25% and 50%, with insects accounting for at least 15% of that loss [109, 110]. Post-harvest losses are estimated at a further 10% and although cereal staples are lost in this manner, fruit, vegetables and root crops are more susceptible [110, 111]. A total of 1,663 species of insect are known to infest post-harvest storage of crops [112]. Total global crop loss, pre- and post-harvest, due to insect consumption, is estimated to account for food capable of feeding 1 billion people [113]. The global annual usage (tonnes) of pesticides per annum has increased year on year (see figure 1.3), but crop losses due to pests have not decreased significantly over the last 40 years [113].

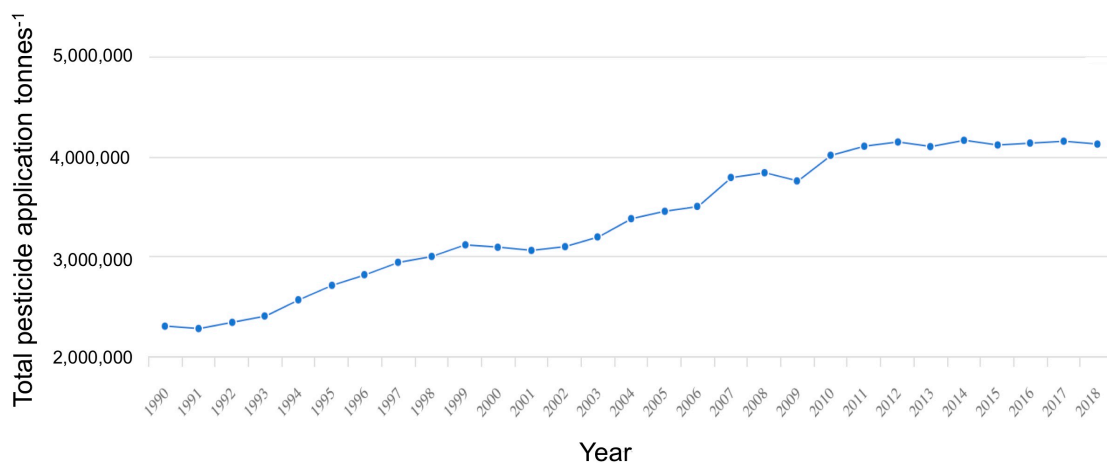


Figure 1.3: Total weight of pesticide used globally shown per annum (1990-2018) in tonnes. [Data taken from: Food and Agriculture Organisation of the United Nations [114] (FAOSTAT accessed 16/09/20). Pesticides include: insecticides, herbicides, fungicides, bactericides, rodenticides, disinfectants, mineral oils and plant growth regulators]

These data include compounds used to control animal (both vertebrate and invertebrate), plant and pathogenic pests. However, usage is not uniform across the planet, with Asia, the Americas, Europe, Africa and Oceania accounting for approximately 52%, 32%, 12%, 2% and 1.7% respectively of the total usage in 2017. The type of pesticide applied also varies by region. For example, in North America and Western Europe unwanted plant species are routinely controlled using herbicides, whereas in other areas cheaper farm labour allows for weeding by hand [115].

1.4 Chemical control of agricultural insect pests

The first use of inorganic compounds, such as sulphur, as insecticides is recorded 3500-4500 years ago by Sumerian and Chinese societies and insecticidal compounds have been applied to crops ever since [109]. Up until the 1800s control was mostly left to these inorganic compounds, certain botanicals, such as nicotine (recorded use from 1690), and the manual removal of insect pests [116]. Around 1820 pyrethrum was extracted from *Chrysanthemum cinerariaefolium* (Asteraceae) flowers and rotenone was isolated from the roots of plants in the genus *Derris* [117, 118]. Pyrethrum remained the most important botanical insecticide for almost two hundred years [116]. Synthetic organics developed in the 1940s to 1970s, were cheaper, more

readily available and more stable; as such, they largely replaced the inorganic and botanical insecticides [116]. New chemical classes of neurotoxicants such as organochlorines (eg. DDT), organophosphates (eg. parathion), methylcarbamates (eg. carbaryl) and pyrethroids (eg. deltamethrin) were introduced during these decades [109, 116]. The persistence in the environment of the early synthetic organics, adverse effects on ecosystems and non-target organisms, public health concerns and the development of insecticide resistance led to stronger regulatory assessment in the industry [116]. Several compounds were withdrawn from the market or had their use restricted [109]. In response to this there was a move to create biodegradable compounds and pesticides with higher potency and enhanced specificity [116]. With the use of each chemical class the evolution of resistant strains of pest species limits their effectiveness [119]. The potential impact of insecticide resistance on human health and future food security drives the development of new insecticide compounds [120]. In the period since WWII at least one new major chemical class of insecticide has been developed each decade [121]. In the 1990s the development of neonicotinoids began to replace systemic organophosphates and methylcarbamates for use against sucking insect pests [116, 118]. By 2010 their use had grown rapidly and they accounted for >25% global insecticide market value [122]. The insecticide resistance action committee (IRAC) classifies insecticides by their mode of action (MoA) [123, 124].

Although there are 32 known MoA groups [124], there are five main MoA groups:

1. Nerve and muscle targets
2. Growth and development targets
3. Respiration targets
4. Midgut targets
5. Unknown or uncertain MoA

With the exception of novaluron, all the insecticides recommended for use on alfalfa [89] belong to the first MoA group: those that target nerve and muscle (see table 1.1).

1.4.1 Insecticides targeting the GABA-gated chloride channel

1.4.1.1 Cyclodiene organochlorines

Organochloride (OC) is a generic term for pesticides that contain chlorine [125]. OCs target the gamma-amino butyric acid (GABA) regulated chloride channel (or GABA-gated chloride receptor) in neurons, reducing neuronal inhibition and causing hyperexcitation, convulsions and death [126]. OCs were the first class of synthetic organic insecticides to be developed, the first and most famous of which is DDT [127]. Many compounds from this class have been removed from the market due to persistence in the environment [127].

1.4.1.2 Phenylpyrazoles (Fiproles)

Phenylpyrazoles are broad-spectrum insecticides characterised by a central pyrazole ring [127]. Fipronil, the first compound in this group, was registered for use in the US in 1996 [128]. These insecticides target the GABA-gated chloride channel and have a high binding affinity for insect receptor complexes [129]. They are used against chewing and sucking insects [130]. Fipronil is classified as highly toxic to *A. mellifera* and bumblebees (*Bombus* spp.) [131-133].

1.4.2 Insecticides targeting acetylcholinesterases (AChEs)

This group of insecticides is toxic to insects, and potentially mammals, as it targets acetylcholinesterase (AChE) enzymes [134, 135]. Acetylcholine (ACh) is found in vertebrates and invertebrates and is involved in the transmission of nerve impulses at synaptic and neuromuscular junctions. It is the chief excitatory neurotransmitter in the insect central nervous system (CNS) [124] and is released from vesicles in a nerve ending in response to a nerve impulse causing stimulation of the nerve or muscle [136]. AChEs effect the rapid hydrolysis of the neurotransmitter ACh into the inactive products choline and acetic acid [136]. The enzyme acts in a regulatory manner by reducing the concentration of ACh at the junction and hence the level of stimulation. The inhibition of AChEs by this group of insecticides leads to hyperexcitation, rapid twitching, paralysis and death [124].

1.4.2.1 Organophosphates (OPs)

Organophosphate (OP) is a generic term for pesticides that contain carbon (hence-*organo*), hydrogen and phosphorus [125]. They are derived from one of

the phosphorous acids and work by inhibiting AChEs [127]. OP's are generally more toxic to vertebrates than other classes of insecticides, however, they are not chemically stable and therefore do not persist in the environment [127]. Due to the relatively high toxicity of OPs the use of the class was reappraised during the 1990s and many were withdrawn from the market [127]. Two OPs were used in this PhD: coumaphos and chlorpyrifos. Their chemical structures are shown in figure 1.4. Chlorpyrifos is classified as highly toxic to *A. mellifera* and the buff-tailed bumblebee (*Bombus terrestris*) [29, 137], whereas coumaphos is classified as practically non-toxic to both species and is regularly used as an in-hive acaricide to protect against the varroa mite (*Varroa destructor*) [137, 138].

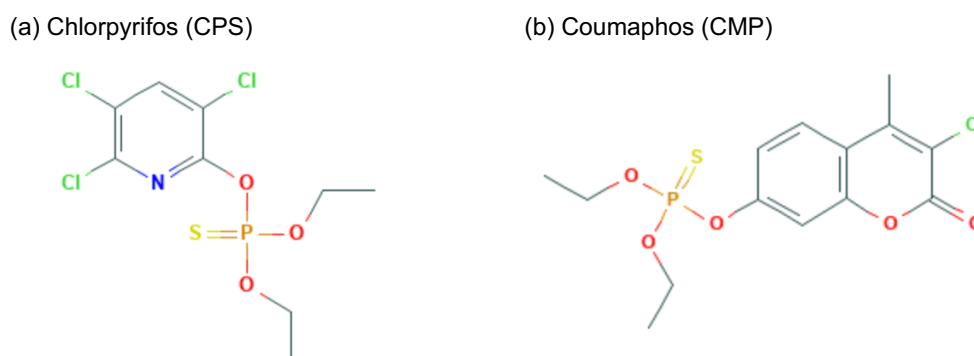


Figure 1.4: 2D chemical structure of two organophosphate insecticides (a): Chlorpyrifos (CPS) [139], (b): Coumaphos (CMP) [140].

1.4.2.2 Carbamates

Carbamates are derived from carbamic acid [127]. Like OPs they inhibit AChEs and have similar physical effects on their target insect [127]. Carbamates tend to have a low toxicity to mammals [127]. Carbaryl and methomyl are registered for use in alfalfa production [89] (see table 1.2). Both these insecticides are classified as highly toxic to *A. mellifera* [141], but are moderately toxic to *B. terrestris* [142].

1.4.3 Insecticides targeting voltage-gated sodium channels

Members of this group of insecticides are considered as axonic poisons. They target the sodium channel in the neuronal membrane and keep it open [124, 127]. This allows sodium ions to continually enter the axon causing

hyperexcitation and repetitive discharges [124, 127]. This eventually leads to paralysis and death [127].

1.4.3.1 Pyrethroids

Pyrethroids are synthetic analogs of pyrethrum, the naturally occurring compound found in *C. cinerariaefolium* [117, 143]. Extracting the natural compound from the dried flowers of the plant is costly and the resulting pyrethrum is unstable in sunlight [117]. Pyrethrum extract contains six insecticidal esters (pyrethrins). Pyrethroids are derived from these natural esters, but are more stable, although they do not generally persist in the environment [127, 143]. Historically, pyrethroids have been grouped into two sub-classes: type I and type II, based on their structure and toxicity [144]. Three type II pyrethroids were used in this PhD: deltamethrin, *tau*-fluvalinate and *alpha*-cypermethrin. Their chemical structures are shown in figure 1.5. Both deltamethrin and cypermethrin are classified as highly toxic to *A. mellifera* [145, 146], and the former is also highly toxic to *B. terrestris* [137]. In contrast, *tau*-fluvalinate is classified as moderately toxic to *A. mellifera* [138] and practically non-toxic to *B. terrestris* [137]. As a result, *tau*-Fluvalinate has been used as an in-hive treatment against *V. destructor* [138].

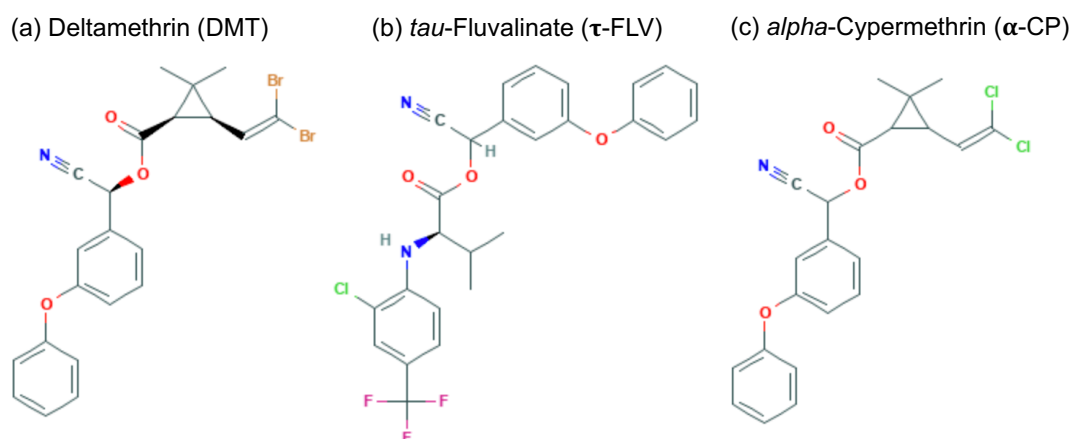


Figure 1.5: 2D chemical structure of three pyrethroid insecticides (a): Deltamethrin (DMT) [147], (b): *tau*-Fluvalinate (τ -FLV) [148], (c): *alpha*-Cypermethrin (α -CP) [149].

1.4.4 Insecticides targeting nicotinic acetylcholine receptor (nAChR)

Members of this group of insecticides are agonists for nicotinic acetylcholine receptors (nAChR) [150, 151]. nAChR respond to the neurotransmitter ACh.

Agents that disrupt the functioning of the nAChR are good candidates for potent insecticides [119]. AChEs cannot breakdown these agonists leading to a hyperexcited state [152].

1.4.4.1 Neonicotinoids

Neonicotinoids are classified as nAChR competitive modulators and as such they bind to the receptor causing hyperexcitation, paralysis and death [124]. Neonicotinoids are water soluble and so they can be used as systemic insecticides. Their use as seed dressings means they are taken up into the tissues of the plants they are protecting, thus reducing the dosage used and need for repeated applications [153]. Neonicotinoids are modelled on nithiazine (see figure 1.6), a chemical developed at Purdue University in the late 1960s [154]. They are more photostable than nithiazine, making them highly efficient for use in agriculture [118]. Neonicotinoids are classified chemically into three subclasses; *N*-nitroguanidines (eg. imidacloprid), nitromethylenes (eg. nitenpyram) and *N*-cyanoamidines (eg. thiacloprid; see figure 1.7) [155, 156]. Due to differences in binding site interactions and detoxification, neonicotinoids are selectively more toxic to insects than mammals [118]. The key element of this selectivity is the negatively charged tip of the nitro or cyano group which interacts with positively charged amino acid residues (eg. lysine, arginine or histidine) present in the insect, but absent in mammalian nAChRs [157]. This specificity for insect nAChRs has led to the global use of this class of insecticides and they now account for approximately one quarter of all insecticide use [158].

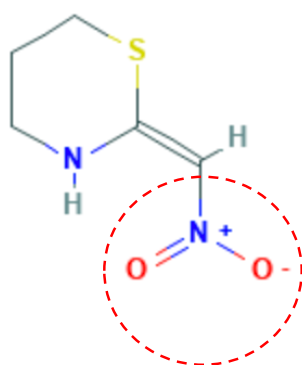
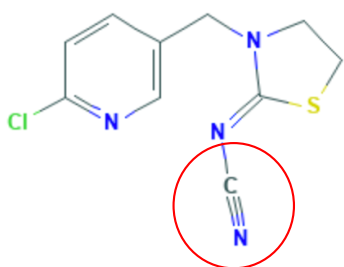
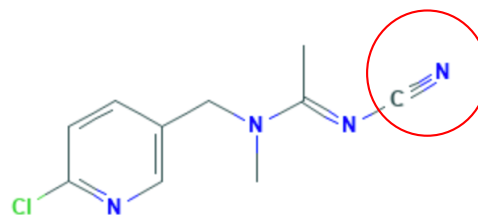


Figure 1.6: 2D chemical structure of Nithiazine [159], the negatively charged nitro-group, also present in the *N*-nitroguanidine neonicotinoid insecticides, is outlined with a red dashed line.

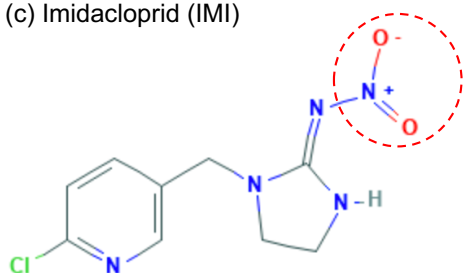
(a) Thiacloprid (THC)



(b) Acetamiprid (ACE)



(c) Imidacloprid (IMI)



(d) Clothianidin (CLO)

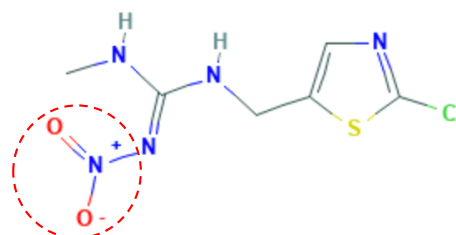


Figure 1.7: 2D chemical structure of neonicotinoid insecticides. (a): Thiacloprid [160], (b): Acetamiprid [161], (c): Imidacloprid [162], (d): Clothianidin [163]. The negatively charged cyano-groups are outlined with a solid red line and negatively charged nitro-groups with a dashed red line.

There is a marked difference in the toxicity to *A. mellifera* and *B. terrestris* between the subclasses [137, 155]. Toxicity data are usually presented in the form of an LD₅₀ (where LD stands for 'Lethal Dose'), the measure of the median single dose or concentration of the test substance (active ingredient – a.i.) that

causes death in 50% of a population, usually within 48 hours [164-166]. The magnitude of difference between acute contact toxicity LD₅₀ (48 h) values for thiacloprid and imidacloprid being 479-816-fold in *A. mellifera* and >250-fold in *B. terrestris* [137, 155]. The lower toxicity of the *N*-cyanoamidine neonicotinoids has been attributed to their more rapid metabolism [155, 167, 168] and the higher toxicity of *N*-nitroguanidine metabolites such as imidacloprid-olefin [169]. Four neonicotinoids were used in this PhD, two examples of a *N*-nitroguanidine pharmacophore: imidacloprid and clothianidin; and two examples of a *N*-cyanoamidine pharmacophore: thiacloprid and acetamiprid. Their chemical structures are shown in figure 1.7. Imidacloprid and clothianidin are classified as highly toxic to both *A. mellifera* and *B. terrestris* [137, 155, 170]. Whereas, acetamiprid is classified as moderately toxic to *A. mellifera* [155] and practically non-toxic to *B. terrestris* [137] and thiacloprid as practically non-toxic to both species [137, 155].

1.4.4.2 Butenolides

Butenolides are classified as nAChR competitive modulators and as such they bind to the receptor causing hyperexcitation, paralysis and death [124]. Although they share the same MoA as neonicotinoids, butenolides are chemically distinct and are therefore classified separately [171]. Flupyradifurone is the first representative of this new class of insecticide [171]. The chemical structure of flupyradifurone is shown in figure 1.8. Flupyradifurone is classified as practically non-toxic to both *A. mellifera* and *B. terrestris* with an acute contact LD₅₀ >100µg bee⁻¹ to both species [171].

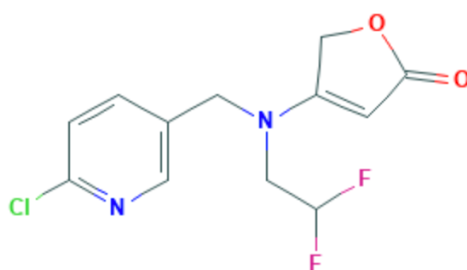


Figure 1.8: 2D chemical structure of the butenolide insecticide Flupyradifurone (FPF) [172].

1.4.4.3 nAChR allosteric modulators – site 1 (Spinosyns)

This group of insecticides acts by disrupting the post-synaptic binding of acetylcholine to nAChR's [124]. Spinosyns are one of the newer classes of insecticides and they are derived from *Saccharopolyspora spinosa*, a soil actinomycete [151]. Spinosad and spinetoram are registered for use for lowbush blueberries (*Vaccinium angustifolium*) crops [91] (see table 1.2). Both insecticides are classified as highly toxic to *A. mellifera* [145, 173].

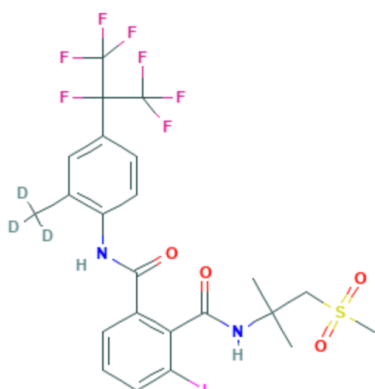
1.4.5 Insecticides targeting ryanodine receptor (RyR)

The ryanodine receptor (RyR) is named after the plant alkaloid ryanodine, a natural insecticide from *Ryania speciosa*, which binds to RyRs with high affinity [174, 175]. The RyRs is a non-voltage-gated calcium channel [175]. RyRs mediate the release of intracellular stores of Ca^{2+} leading to the activation of muscles [124, 174, 175]. As such they are an excellent target for insect control as over excitation will lead to contraction and paralysis [124, 175]. Ryanodine itself has limited agricultural utility due to unacceptable levels of toxicity to mammals [176].

1.4.5.1 Diamides

In the 1990s RyR modulators with high selectivity towards insect over mammalian receptors (>500-fold [176]) were developed and diamide insecticides were introduced to the market [177]. Chlorantraniliprole is classified as highly toxic to *A. mellifera*, although it is considered to pose low risk to other beneficial insect species [175, 178]. Flubendiamide is practically non-toxic to *A. mellifera* (>200 μ g ai/bee) and many other beneficial insect species, but has possible adverse effects on aquatic invertebrates [178-180]. The chemical structure of flubendiamide and chlorantraniliprole are shown in figure 1.9.

(a) Flubendiamide



(b) Chlorantraniliprole

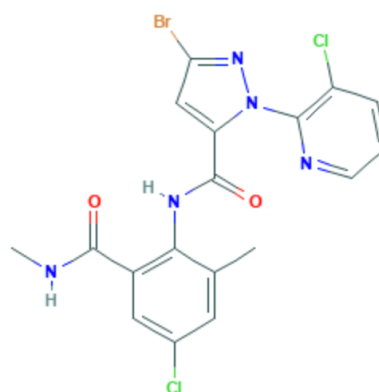


Figure 1.9: 2D chemical structure of two diamide insecticides. (a): Flubendiamide [181] and (b): Chlorantraniliprole [182].

1.5 Plant secondary metabolites

Plants form the base of most food chains, serving as the ultimate food source for heterotrophic organisms. To protect themselves against microorganisms and herbivores, plants produce a diverse array of chemicals known as secondary metabolites, or allelochemicals [183, 184]. Plant defence chemicals generally occur as mixtures of structurally similar compounds, in contrast to synthetic insecticides which are normally single compounds applied to control a pest insect species [185]. To date, over 30,000 secondary metabolites have been reported [186]. Secondary metabolites are divided into three groups: nitrogen-containing alkaloids, hydroxylated aromatic ring-containing phenolic compounds and isoprene unit-containing terpenoids [187]. This PhD only included the study of certain alkaloids.

1.5.1 Alkaloids

Some 300 plant families produce alkaloids and over 10,000 compounds have been isolated [187]. This group of secondary metabolites produces pronounced physiological effects in animals and can often be very toxic to vertebrates (eg. strychnine and curarine) [186]. The presence of alkaloids is not restricted to the vegetative parts of the plant, but can also be found in pollen and nectar [184, 188]. As such, bees regularly encounter these plant defensive chemicals [184]. Four alkaloids were used in this PhD: nicotine, atropine, cytosine and anabasine. Their chemical structures are shown in figure 1.10.

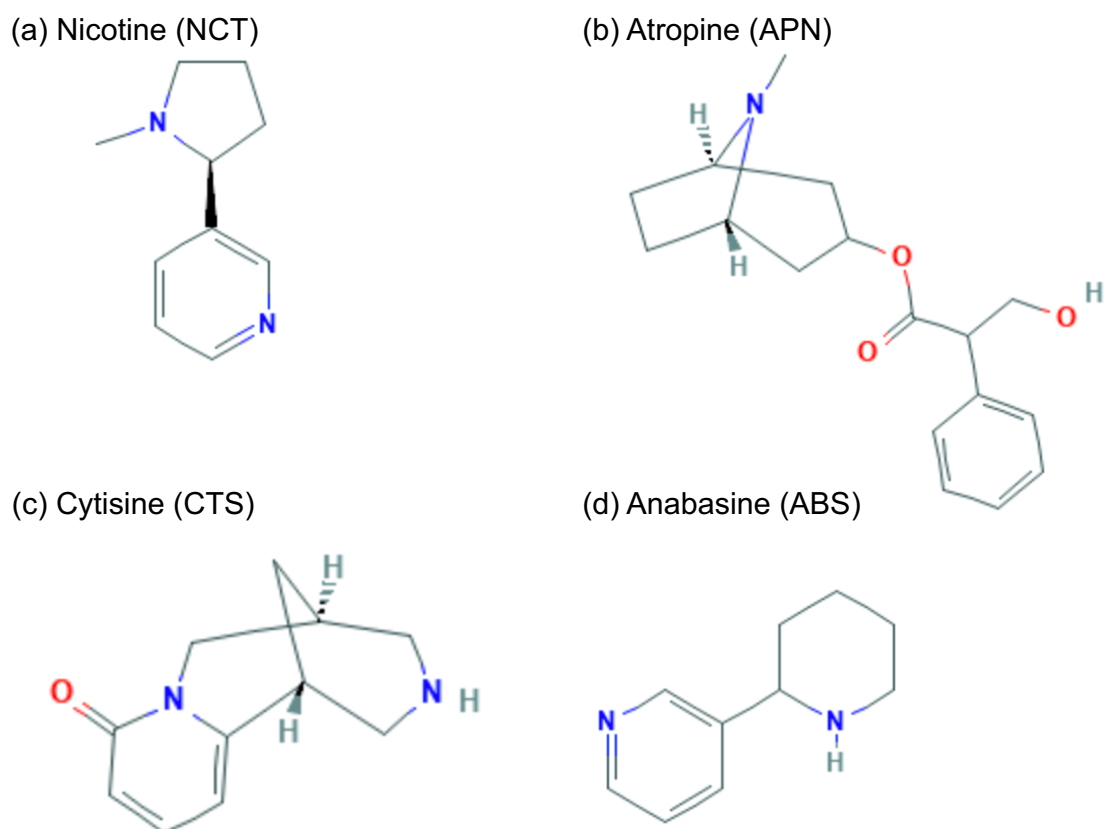


Figure 1.10: 2D chemical structure of alkaloid allelochemicals (a): Nicotine (NCT) [189], (b): Atropine (APN) [190], (c): Cytisine (CTS) [191], (d): Anabasine (ABS) [192].

Nicotine targets the nAChR and is found in the Solanaceae family (nightshades) of plants that includes the tobacco plant (*Nicotiana tabacum*) [117]. Atropine targets the muscarinic acetylcholine receptor (mAChR) and is found in the Solanaceae family, including *Atropa belladonna* (deadly nightshade), *Datura stramonium* (jimson weed) and *Mandragora officinarum* (mandrake) [193]. Cytisine is considered a partial agonist of the nAChR and is found in the Leguminosae family, including *Laburnum anagyroides* [194]. Anabasine targets the nAChR and is found in *Nicotiana glauca* (tree tobacco) [195]. All four compounds are considered highly toxic insecticides [186].

1.6 Mechanisms involved in insecticide resistance and tolerance

Insecticide resistance refers to the heritable ability in an organism to withstand exposure to a dose of toxicant that would kill the majority of a susceptible population [196]. The first report on insecticide resistance concerned the use of sulphur-lime to control *Quadraspidiotus perniciosus* (San Jose scale) [197]. It

was published in 1914 and, with the emergence of synthetic insecticides and their increasing use, reported incidences of resistance have risen ever since [123]. The application of chemicals to control an insect species effectively places those populations under selective pressure. Different genotypes within a population may have differing tolerance of an insecticide [120, 198]. The selective advantage of more tolerant individuals will lead to the need to increase the doses required to control future populations, resulting almost inevitably to a resistant population [116]. Incidences of overuse and misuse have also contributed to selection of resistance, and over 600 arthropod species now show some level of pesticide resistance [116, 199]. Of these 178 are species that negatively affect human health, and 265 species are agricultural pests [200]. The mechanisms for insecticide resistance can be broadly classified into four categories: metabolic (enhanced insecticide detoxification, excretion or sequestration); target site modification (decreased sensitivity/binding); physiological (eg. increased cuticular thickness) and behavioural (eg. avoidance of contact with insecticide) [120, 198, 201].

In contrast to insecticide resistance, insecticide tolerance is less easy to define, but here it is taken to refer to the constitutional ability of an organism to cope with synthetic xenobiotics due to an inherent metabolic pathway [202, 203]. Herbivorous insects have evolved mechanisms to cope with the toxicological challenge presented by the plant secondary metabolites they encounter in their diet [202, 203]. The metabolic pathways involved in adaptations to allelochemicals, which allow for insecticide tolerance, are the same as those implicated in metabolic resistance to insecticides [204]. Innate enzymatic pathways in *A. mellifera* are implicated in the tolerance to certain synthetic insecticides found in the species [202].

1.6.1 Metabolic mechanisms (pharmacokinetic mechanisms)

Insects defend themselves against xenobiotic and toxic compounds through a three-phase process that converts the xenobiotic into less-toxic, water-soluble metabolites that can be excreted [205-207] (see figure 1.11). In phase I the substrate undergoes enzymatic reduction, oxidation or hydroxylation; phase II involves conjugation of phase I products, and phase III consists of the transportation of the phase II conjugates for excretion [206, 208] (see figure

1.11). Insects rely on a taxon-specific suite of detoxification enzymes to metabolise the plant secondary metabolites they encounter in their environment [209, 210]. Three superfamilies of enzymes are involved in detoxification of xenobiotics compounds in insects: cytochrome P450s (P450s); carboxyl/cholinesterases (CCEs) and glutathione-S-transferases (GSTs) [120, 208, 211]. One common feature of all these superfamilies of enzymes in insects is the number of duplicated genes [212]. A second, more poorly understood metabolic defence deployed by certain herbivorous insects is the sequestration of xenobiotics for safe storage or to provide a defensive advantage against predators [198, 213-215]. Sequestration has been linked to resistance in many insect species [216].

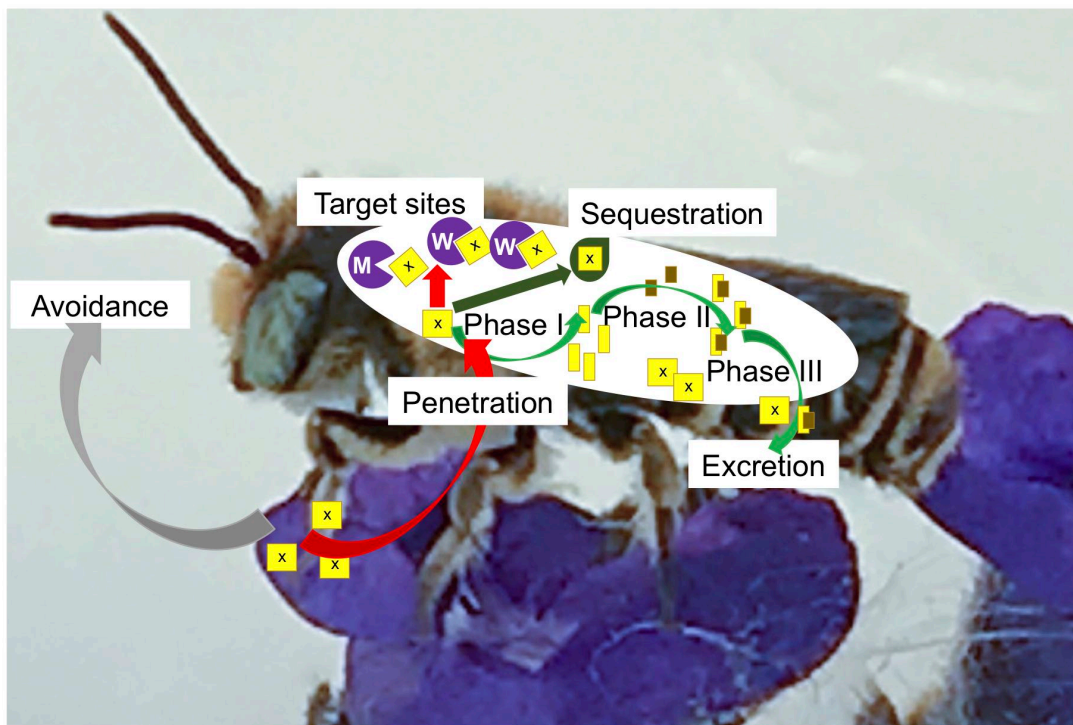


Figure 1.11: The main routes of 'xenobiotic compound-insect' interactions. Interactions shown: (i) avoidance, (a behavioural response - genetic or learned); (ii) penetration; (iii) three-phase metabolic process (iv) excretion (both metabolite and unaltered substrate) (v) sequestration. Yellow squares annotated with an **X** indicate a xenobiotic compound; target sites (shown in purple) annotated with a **W** indicate 'wild-type' receptors and those marked **M** indicate 'mutant-type' receptors with the potential to confer resistance.

1.6.1.1 Cytochrome P450s (phase I)

Cytochrome P450s (CYPs) form one of the largest gene families [217, 218]. The evolutionary origin of the ancestor of all P450 enzymes lies in prokaryotes and CYP genes are found in almost all living organisms [219]. CYP genes are broadly considered as 'environmental response' genes [217]. They code for P450 enzymes, heme-thiolate proteins, best known for their functions as monooxygenases, that play a role in the biosynthesis and metabolic detoxification of endogenous substrates, such as pheromones, hormones and lipids, as well as targeting xenobiotic compounds [217, 218]. The absorbance peak near 450 nm of the Fe^{II}-CO complex, formed when a reduced P450 enzyme binds CO gives the superfamily its name [220]. The genetic diversity of the P450 superfamily, together with its' substrate promiscuity and catalytic versatility means it is the only metabolic system that has been implicated in resistance to all classes of insecticides [204]. P450s are linked to resistance through coding sequence mutations, upregulation of the genes and overexpression [204].

The total number of CYP genes in a genome, the CYPome, can vary radically between species and in insects it ranges from 36 in the body louse (*Pediculus humanus*) [221], 85 in the fruit fly (*Drosophila melanogaster*) [222], up to 170 in the southern house mosquito (*Culex quinquefasciatus*) [223]. Across CYPomes it is common to find one or more subfamilies whose members have proliferated. This is known as a CYP 'bloom'. These blooms represent recent, independent expansion of the subfamily involved [218]. The size of CYPome cannot be used as a good proxy for functional repertoire. The body louse with its' seemingly impoverished CYPome has become resistant to many classes of insecticide [221]. Indeed, the human CYPome contains only 57 CYP genes [224]. However, two members of this suite (*CYP3A4* and *CYP2D6*) are potent metabolisers, being responsible for the biotransformation of over 50% of clinically used drugs [225]. The CYPome of *A. mellifera* contains only 47 genes and that of *B. terrestris* 49 genes [226]. However, in these bee species members of the *CYP9Q* lineage are capable of metabolizing insecticides belonging to three different classes and certain P450s from this family may function as potent metabolisers [168].

A process of extensive gene duplication and neofunctionalization has driven the diversity of the P450 superfamily, allowing for novel functions to evolve in response to new environmental stresses, whilst ancestral metabolic competency is retained [217, 219]. To avoid confusion with the naming of this diverse superfamily of enzymes, rules of nomenclature were introduced in 1987 [227] and have been revised and updated as necessary [228, 229]. All P450s are named with CYP as a prefix (see figure 1.12). An Arabic numeral then designates the family (members >40% identical), a capital letter(s) designates the subfamily (members >55% identical) and finally an Arabic numeral designated the individual gene [229] (see figure 1.12).

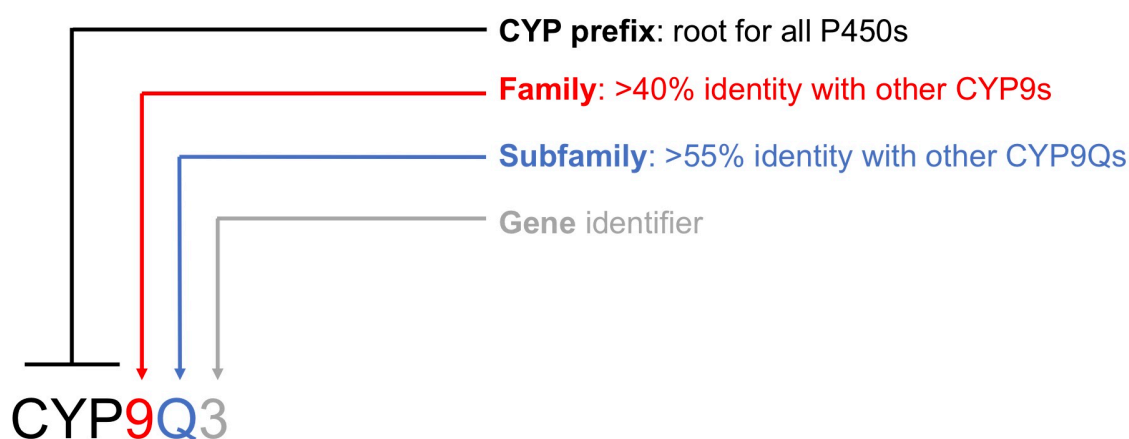


Figure 1.12: Scheme of accepted P450 nomenclature. Amended from [218].

Allelic variants of a gene (>96% identical) are named by subscripts v1, v2 [217]. Allelic variants of CYP genes appear to be common and the regions of the genome coding for them can be highly variable [218, 230]. As such a gene that has recently been converted to a pseudogene in one population may still be active in another [218]. Allelic variants of *CYP2D6*, the human potent metaboliser, have different metabolic abilities, producing an enzyme whose ability is classified as absent, poor, intermediate, efficient or ultra-rapid [231]. Copy number variation (CNV) is also commonly found within the CYPomes of insects and CNV has been implicated in insecticide resistance in *D. melanogaster* [232]. It can be difficult to distinguish allelic variants from recently duplicated genes as the latter can differ by as little as a single amino acid, however, genotyping may be useful in predicting possible interpopulation

differences in metabolism [212, 233]. The higher order grouping into CYP clans was introduced in 1998 [234] and as such insect CYP genes are distributed into four clans: CYP2, CYP3, CYP4 and mitochondrial CYP clans [235] .

Mitochondrial CYP clan

This clan of P450s appears to only be found in animals [235]. In vertebrates all members of this clan are involved in endogenous metabolism [235]. Some insects however, have two types of mitochondrial CYP genes: one group of highly conserved sequences that code for enzymes involved in essential physiological functions and a taxon-specific group of genes that are rapidly evolving and may be involved in the metabolism of xenobiotics and insecticide resistance [235]. In *A. mellifera* it is likely that members of this clan are involved in ecdysteroid biosynthesis [211].

CYP2 clan

The CYP2 clan is an ancient group of P450s and a CYP2 member was probably the ancestor of the mitochondrial CYP clan [236]. P450 enzymes from the CYP2 clan are mostly involved in essential physiological functions [235]. For example, in *A. mellifera* it is likely that the CYP2 gene CYP15A1 is involved in juvenile hormone biosynthesis [211].

CYP3 clan

Genes from the CYP3 clan are numerous and diverse in insects and are often found in large clusters [235, 237]. There is a significant amount of evidence that members of the CYP3 clan are linked to the metabolism of xenobiotics [238] [239]. The CYP3 clan contains the CYP6 and CYP9 families [218]. Members of both these families contain members with known xenobiotic detoxification capacity [168, 238]. In *A. mellifera* there is clear evidence of recent and rapid duplication in the CYP6 and CYP9 families and members of both are organised in clusters on linkage groups in which they appear (linkage groups 13 and 14) [211].

CYP4 clan

The CYP4 clan is expanded and hugely diversified in insects [235]. Some CYP4 genes are involved in essential physiological functions, such as

pheromone metabolism, but others are linked to the detoxification of xenobiotics [235]. Members of this clan have also been implicated in insecticide resistance to pyrethroids [240], DDT [241], OPs and carbamates [242]. The CYP4 clan in *A. mellifera* has few members, whose functions are not known, but they are likely to be involved in ecdysteroid biosynthesis and lipid metabolism [211].

1.6.1.1.1 Diversity and conservation of CYP gene sequences

The sequence identity of P450s is often extremely low (<20%) and across the group there are only three residues considered to be invariant in the protein sequence [243]. However, there are five conserved motifs within the sequence (see figure 1.13). These conserved elements surround the core of the protein that contains the heme group that is responsible for electron transfer and oxygen activation [218, 219]. This heme group forms the 450 nm absorption peak from which P450s get their name [219, 243]. The more variable regions, which include six substrate recognition sites (SRSs; see figure 1.13), are responsible for recognising and binding substrates such as insecticides [219, 243, 244]. SRS4 contains conserved elements that are involved in the reaction mechanism (the oxygen-binding motif in the I-helix), as well as residues associated with substrate recognition, so is more highly conserved than the other SRSs [243].

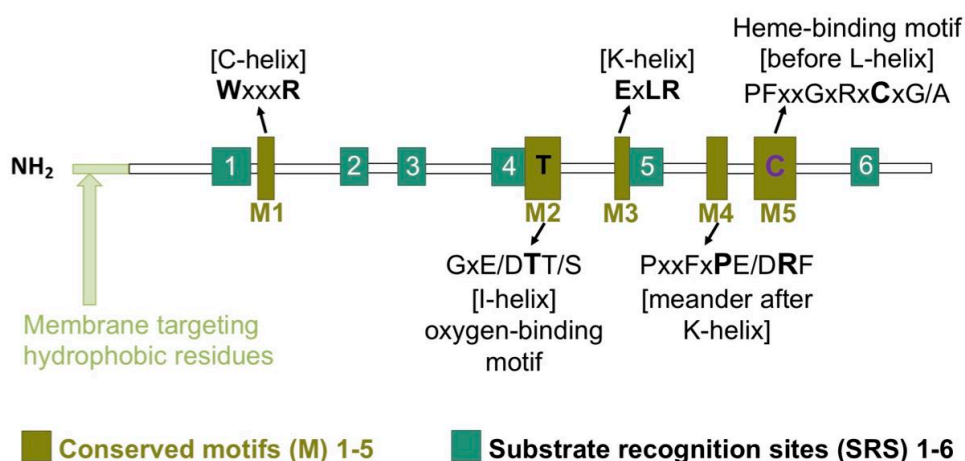


Figure 1.13: Generalised primary structure of a microsomal CYP, showing conserved and variable regions, with highly conserved amino acids in bold [218, 243, 244].

The conserved motifs found in sequences throughout the P450 superfamily enable the correct identification and alignment of otherwise very diverse sequences [243, 245]. However, recently several insect P450s from the CYP3 clan have been identified that lack the conserved cysteine in the heme-binding motif and the WxxxR motif in the C-helix [245]. Other insect P450s lack the conserved threonine in the oxygen-binding motif in the I-helix [245-247] and so caution should be exercised in identifying sequences and it may be that some P450s previously categorised as pseudogenes may actually be functional genes [218].

1.6.1.1.2 P450 tertiary structure

In general, all P450s are composed of 4 β -sheets (presence of β 5 sheet is variable) and 13 α -helices [243, 248] (see figure 1.14(a)). Despite the tremendous diversity in sequence, structural determination of bacterial and eukaryotic P450s, using crystallisation and X-ray diffraction shows that there is a conserved structural fold in the tertiary structure that allows for a common topography [219, 248-250]. The basic 3-dimensional structure of a P450 is that of a triangular prism, with a 'helix-rich' domain and a 'helix-poor' domain [248, 249] (see figure 1.14(a)). The five conserved motifs and the six SRSs all appear in a 'central band' around the tertiary structure of the P450 protein, with the majority of the SRSs occurring in the proximal face (helix-poor domain; see figure 1.14(b)). This 'central band' of the P450 includes the active site/heme pocket and substrate access channel, found on the distal face [219, 248]. These sites are involved, not only in substrate recognition and binding, but also with orientation of substrate within the pocket [219]. In the investigation of potential effects of the diversity in amino acid sequence of P450s on function, structural homology modelling has been a useful tool [251, 252]. One of the biggest limitations of structural homology is the small number of crystal structures from eukaryotic species available [251, 253]. To date the crystal structure of an insect P450 has not been resolved successfully.

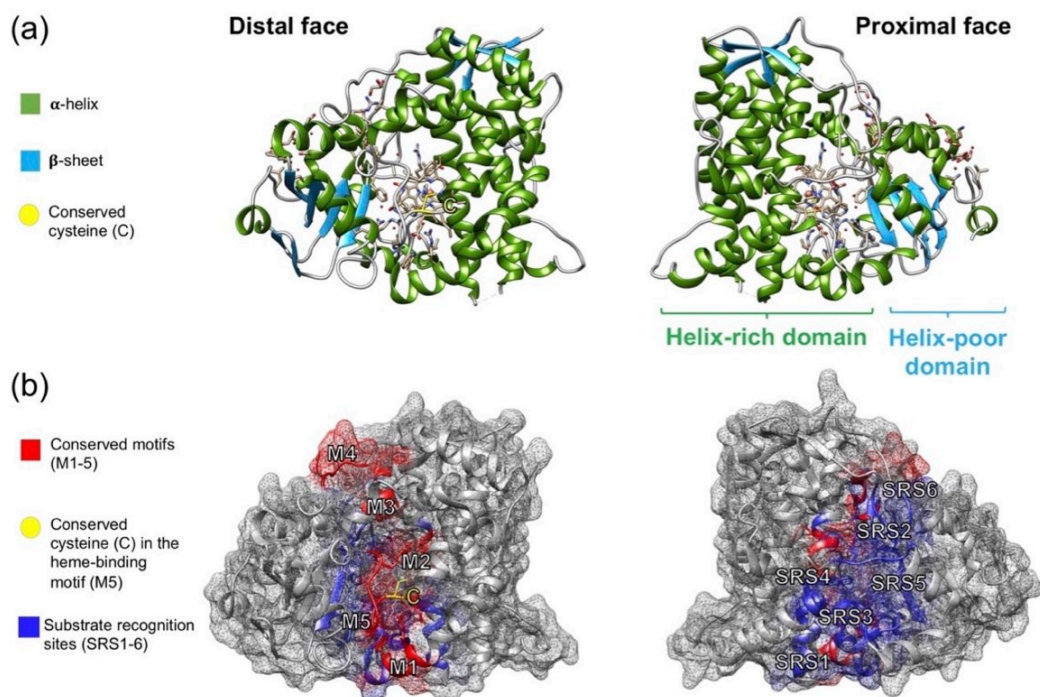


Figure 1.14: Tertiary structure of *Homo sapiens* CYP4A3 (PDB: 4D6Z [254]) (a): Secondary structures displayed; α -helices coloured green and β -sheets cyan. The conserved cysteine is coloured yellow and marked with a C (b): Conserved motifs (M) are coloured red and substrate recognition sites (SRSs) blue. The crystal structure CYP4A3 (PDB: 4D6Z [254]) is rendered with the surface depicted in mesh. [Figure created using UCSF Chimera version 1.10.1]

1.6.1.2 Carboxyl/cholinesterases (CCEs; phase I detoxification)

Most of the widely used insecticides, including many pyrethroids and all OPs and carbamates, are esters [255]. In almost all cases hydrolysis of the ester group into its corresponding acid and alcohol groups by an esterase, results in less toxic metabolites that can be excreted [120, 255]. In a few cases, such as the voltage-dependent sodium channel blocker, indoxacarb, the insecticide depends on this hydrolysis for its' toxicity [255]. In insects the main group of esterases of interest hydrolyse carboxylic acid esters and so are termed carboxylesterases [255]. The closely related cholinesterases, such as AChE, target choline-based esters and are also of interest in the context of insecticide resistance [255]. CCEs have been documented as mediating resistance to OPs, carbamates and pyrethroids through gene amplification, coding sequence mutations and upregulation [204]. The overexpression of certain CCEs have

even been implicated in the potential resistance to the neonicotinoid imidacloprid in the tarnished plant bug (*Lygus lineolaris*) [256]. Esterases frequently mediate the sequestration of insecticides and examples of this process are found across many species, including aphids and mosquitos [216].

Unlike the nomenclature associated with P450s, the rules for classification of CCEs have not been clarified, however they are divided into 13 clades across three major classes [211, 212]. The number of CCE genes varies between insect species. *D. melanogaster* has 35 sequences and the mosquito *Anopheles gambiae* has 51, most of which appear in gene clusters [212]. *A. mellifera* only has 24 CCE sequences, across the three classes, but with sequences from only 10 of the 13 clades [211]. Most CCE genes in *A. mellifera* are dispersed across 11 chromosomes, with only nine genes found in four small clusters [211]. This implies there is little evidence for recent gene duplication and neofunctionalization [211].

1.6.1.3 Glutathione-S-transferases (phase II detoxification)

Glutathione-S-transferases (GSTs) are known to mediate resistance to OPs, organochlorines and pyrethroids by gene amplification and overexpression [204]. They may also be implicated in resistance to certain neonicotinoids [257]. There are two main groups of GSTs in insects: microsomal and cytosolic [258]. Microsomal GSTs have not been implicated in insecticide metabolism [258]. Classification and nomenclature of GSTs has evolved from a numerical system based on the order of their elution to a system based on phylogenetic analyses with mammalian sequences, dividing genes into six classes designated by the Greek letters: delta, epsilon, sigma, theta, zeta and omega [258]. The number of GST genes varies between insect species. *D. melanogaster* has 39 sequences and *A. gambiae* has 31, with most of them falling into the insect-specific delta and epsilon clades [212]. *A. mellifera* however, only has 10 GST genes [211]. It only has one delta clade member and does not have any epsilon clade genes [211].

1.6.1.4 Excretion of conjugates by efflux pumps (phase III detoxification)

Cells and organelles are all described by lipid membranes and to survive they need to rely on transporters to traffic diverse compounds across their

membranes [259]. Specific proteins associated with the membrane mediate transmembrane transport [259]. Among these transport proteins the ubiquitous ATP-binding cassette (ABC) protein family is one of the largest groups [260]. ABC transporters utilise the energy from ATP-hydrolysis to transport a wide array of hydrophobic substrates, including the metabolites produced by phase I and II detoxification of drugs and insecticides [260]. Overexpression of ABC proteins and increased level of efflux have been implicated in insecticide resistance [261-264]. There is evidence that insecticides from three MoA classes are substrates of ABC proteins in *A. mellifera* [265]. It appears that these transport proteins may be involved in protecting the honeybee from certain pesticides [265].

1.6.2 Target site modification (toxicodynamic mechanisms)

An insecticide is classified by its MoA, which is dependent on the target site it can bind to [124]. Changes in the receptor molecule can reduce the ability of the insecticide to bind, or create an altered response to the compound, either way potentially reducing the toxic effects on the insect [216, 255]. Target site modification is the primary cause of resistance to many insecticides [255]. In most cases the modifications take the form of point mutations that lead to changes in the amino acids of the receptor molecule that are critical to the recognition of an insecticide [255]. These mutations can be lethal, but can also lead to resistance [255]. Established target-site mutations have been described for 13 different classes of MoA [266]. Target site mutations linked to resistance/tolerance have not as yet been described in *A. mellifera* [211].

1.6.3 Physiological adaptation

Target-site modification and metabolic resistance have been the most extensively studied insecticide resistance mechanisms. However, other physiological resistance mechanisms exist, such as reduced penetration through the cuticle and increased excretion [120, 267].

1.6.3.1 Decreased rates of penetration

To reach its target site, an externally applied insecticide must first penetrate the major barrier of an insect's cuticle [267]. The effect of reduced penetration on the toxicity of an insecticide is small, generally in the range of ~2-3-fold [255].

However, it can act as an intensifier of resistance when it appears with target-site modification, or metabolic resistance in the same species and has been reported with DDT, OPs, carbamates, pyrethroids and neonicotinoids [255, 264, 268, 269]. Penetration resistance mechanisms involve two main factors: the thickness or the composition of cuticle [267]. In *A. mellifera* there have been reported race-based differences in the rates of penetration of insecticides [270].

1.6.3.2 Increased excretion – removal of unaltered substrate by efflux pumps (Phase 0 detoxification)

In addition to their role as phase III transmembrane transporter proteins, it has been established that certain ABC transporters are involved in the efficient and direct transport of unaltered insecticide [260, 264]. This process moves the insecticide away from internal sites to the cuticle, thereby decreasing internal accumulation (see figure 1.11) [255]. The increase in efflux, along with removal of the compound by metabolism or sequestration, will cause a reduction in the level of toxicant reaching its target site and so prevent the exertion of its effect [260]. Increased efflux of insecticide has been linked to resistance [264].

1.6.4 Behavioural resistance

Behavioural resistance is generally taken to refer to avoidance behaviours that allow an insect to survive in what would otherwise be a toxic environment [271]. Behaviours include the cessation of feeding and movement away from the area that has been treated [271]. There is an expectation that insects will exhibit a preference for feeding on untreated plants over those treated with pesticide. However, choice studies with *A. mellifera* and *B. terrestris* resulted in counterintuitive findings that individuals prefer nectar containing certain neonicotinoid insecticides, fungicides and herbicides [272, 273]. Yet, in a study of *B. terrestris*, neonicotinoid insecticides appeared to reduce free-flying foraging behaviours, and in *A. mellifera* the application of certain fungicides on cranberry (*Vaccinium macrocarpon*) increased bee foraging on non-crop plants [274, 275]. Behavioural resistance is often harder to investigate than other mechanisms as it can require populations to be monitored in the field and unequivocal examples are scarce [276-278].

1.7 Insecticide detoxification in bees

In common with other insects, bees have evolved sophisticated detoxification systems that can either metabolize endogenous substrates and xenobiotics into non-toxic compounds, or produce metabolites that can be eliminated rapidly (see figure 1.11) [120]. These detoxification pathways are vital in determining the level of sensitivity of bees to insecticides [206]. To date, most studies on the detoxification pathways of bees have focused on members of the Apidea, the largest bee family which accounts for 30% of all species [206]. This family contains *A. mellifera* and *B. terrestris*, both of which are economically important managed pollinators.

A. mellifera has significantly fewer protein coding genes annotated in its genome than other insect species, with a marked reduction in the three superfamilies of enzymes are involved in detoxification of xenobiotics compounds (CCEs, GSTs and P450s) [211]. For example, the CYPome of *A. mellifera* contains only 46 genes [211]. It has been hypothesized that, because of this lack of detoxification, *A. mellifera* may be more sensitive to insecticides than other insect species enzymes [211, 279]. However, a thorough review of data for 62 insecticides across six classes revealed that, whilst *A. mellifera* is sensitive to specific compounds, it does not exhibit a general lack of tolerance in comparison to other insects [29].

1.7.1 Functional importance of CCEs in bees

Although *A. mellifera* only has 24 CCE genes, it has representatives of all three classes, although only 8 are in the class most associated with xenobiotic detoxification [211]. In *D. melanogaster* and *A. gambiae* many CCE genes are found in clusters, but in *A. mellifera* they are dispersed across 11 chromosomes, with only nine genes found in four small clusters [211]. This implies there is little evidence for recent gene duplication and neofunctionalization [211].

CCE genes have been shown to be upregulated in *A. mellifera* in response to exposure to coumaphos [280, 281]. Using enzymatic biomarker analysis CCE activity has also been induced by exposure to fipronil, spinosad and deltamethrin, insecticides from 3 different classes [145]. Synergism studies

using the CCE inhibitor S,S,S-tributylphosphorotrithioate (DEF) have shown that DEF increases the toxicity of certain pyrethroid and OP insecticides to *A. mellifera* [282, 283]. This implies the involvement of CCEs in their metabolism [282, 283]. Further, certain P450 enzymes from the CYP9 subfamily in *A. mellifera* have been shown to hydroxylate *tau*-fluvalinate to a metabolite suitable for cleavage by CCEs [284].

1.7.2 Functional importance of GSTs in bees

A. mellifera only has about a third of the complement of GSTs found in other insects such *D. melanogaster* [211]. This reduction in numbers is particularly evident in the delta and epsilon classes, known to be involved in insecticide metabolism in other insects [285]. With only one delta class member and no epsilon class genes, it might be expected that the species would show a high level of sensitivity to certain insecticides [211]. Although studies using the GST inhibitor diethyl maleate (DEM) showed a low level of synergism for the pyrethroid lambda-cyhalothrin, the absence of interaction between DEM and the pyrethroids cyfluthrin, tau-fluvalinate, the OP coumaphos and the acaricides fenpyroximate, amitraz and thymol suggests that GSTs do not play a primary role in the tolerance of many insecticides in *A. mellifera* [282, 283]. However, GST genes have been shown to be upregulated in *A. mellifera* in response to exposure to coumaphos [280, 281].

1.7.3 Functional importance of P450s in bees

In *A. mellifera* the P450 superfamily appears to play the primary role in the metabolic detoxification of naturally occurring and synthetic xenobiotics [155, 282, 286]. Genome wide expression studies have identified P450s genes that are upregulated by exposure to xenobiotics [280, 287]. Enzyme assays and synergism bioassay studies using the P450 inhibitor piperonyl butoxide (PBO) against various insecticides have yielded results that indicate a strong detoxification role for P450s (see table 1.3) [155, 282, 286]. *A. mellifera* utilizes P450s in the detoxification of insecticides across three classes, including certain neonicotinoids [155], pyrethroids and OPs [283] as well as the phenoxy pyrazole acaricide fenpyroximate [283] and the mycotoxin aflatoxin B1 [288].

Heterologous expression of individual P450s, using baculovirus in lepidopteran

cells is a widely used method to study metabolic activity and substrate selectivity [218]. Recent such work on *A. mellifera* and *B. terrestris* has shown that cytochrome P450 enzymes belonging to the CYP9Q lineage are capable of metabolizing insecticides belonging to three different classes (see table 1.3) [168, 284]. Members of the CYP9Q lineage have also been implicated in the differential toxicity of *N*-nitroguanidine and *N*-cyanoamidine neonicotinoid insecticides [168]. To date, much less work has been conducted on detoxification pathways in non-Apidae bees, however, publications on managed solitary bee species are now emerging, for example recent work on the red mason bee (*Osmia bicornis*) [226], a member of the Megachilidae family. Like *A. mellifera* and *B. terrestris* this species exhibits marked differences in sensitivity to *N*-nitroguanidine and *N*-cyanoamidine neonicotinoids [226]. Although *O. bicornis* lacks CYP9Q genes, it has two closely related genes, CYP9BU1 and CYP9BU2 that detoxify *N*-cyanoamidine neonicotinoids [226]. It may be therefore that certain P450s from the CYP9 subfamily function as potent metabolisers in a similar way to human CYP3A4 and CYP2D6 [225].

Many of the P450s implicated in the detoxification of synthetic insecticides in *A. mellifera*, such as the CYP9Qs, are expressed constitutively at low levels [284]. It therefore appears likely their ability to metabolise synthetic insecticides is not the result of intensive pesticide selection, but rather the consequence of a similarity to natural substrates of the P450s [284]. For example, *A. mellifera* pollinates flowers from the Asteraceae and Solanaceae families that produce the allelochemicals pyrethrum and nicotine respectively [117] and as such may have evolved mechanisms to detoxify these naturally encountered compounds [284].

Worldwide, *M. rotundata* is the most economically important managed solitary bee pollinator [56, 289]. However, it is, as yet, unclear whether the differential sensitivity to neonicotinoids and the tolerance of the acaricides, tau-fluvalinate and coumaphos, also extends to this species. Without this data the use of *A. mellifera* and *B. terrestris* as proxies for this species in ecotoxicological studies may not prove reliable.

Table 1.3: Evidence for P450 involvement in xenobiotic metabolism in bee pollinators

Compound	MoA	Result	Method
<i>Neonicotinoid</i>			
Acetamiprid [155]	nAChR agonist	Increase in toxicity when bees treated with a P450 inhibitor	PBO synergism
Acetamiprid [247]	nAChR agonist	Insecticide metabolism in specific expressed P450s (CYP9 subfamily)	Baculovirus expression of CYP9s and metabolic assays
Acetamiprid [137]	nAChR agonist	Increase in toxicity when bees treated with a P450 inhibitor	PBO synergism
Imidacloprid [287]	nAChR agonist	Overexpression of P450s (CYP4, CYP6 and CYP9 clades)	Larval dietary exposure & MicroRNA sequencing
Thiacloprid [155]	nAChR agonist	Increase in toxicity when bees treated with a P450 inhibitor	PBO synergism
Thiacloprid [168]	nAChR agonist	Increase in toxicity when bees treated with a P450 inhibitor; insecticide metabolism in specific expressed P450s (CYP9 subfamily)	PBO synergism; Baculovirus expression of CYP9s and metabolic assays
Thiacloprid [226]	nAChR agonist	Increase in toxicity when bees treated with a P450 inhibitor; insecticide metabolism in specific expressed P450s (CYP9 subfamily)	PBO synergism; Baculovirus expression of CYP9s and metabolic assays
Thiacloprid [247]	nAChR agonist	Insecticide metabolism in specific expressed P450s (CYP9 clade)	Baculovirus expression of CYP9s and metabolic assays
Thiacloprid [137]	nAChR agonist	Increase in toxicity when bees treated with a P450 inhibitor	PBO synergism
<i>Pyrethroid</i>			

Bifenthrin [284]	Sodium channel modulators	Upregulated <i>CYP9Q</i> P450s	Adult dietary exposure & PCR technique
λ -Cyhalothrin [290]	Sodium channel modulators	Mid-guts incubated with a P450 inhibitor and insecticide followed by chromatography to determine metabolite production	Prochloraz synergism
<i>Tau</i> -fluvalinate [284]	Sodium channel modulators	Insecticide metabolism in specific expressed P450s (<i>CYP3</i> clade). Upregulated <i>CYP9Q</i> P450s	Baculovirus expression of <i>CYP6s</i> and <i>CYP9s</i> and metabolic assays. Adult dietary exposure & PCR technique
<i>Tau</i> -fluvalinate [283]	Sodium channel modulators	Increase in toxicity when bees treated with a P450 inhibitor	PBO synergism
<i>Tau</i> -fluvalinate [280]	Sodium channel modulators	Upregulated P450s from the <i>CYP3</i> clade	Adult dietary exposure & microarray analysis
<i>Tau</i> -fluvalinate [137]	Sodium channel modulators	Increase in toxicity when bees treated with a P450 inhibitor	PBO synergism
<i>Organophosphate</i>			
Coumaphos [284]	AChE inhibitors	Insecticide metabolism in specific expressed P450s (<i>CYP3</i> clade)	Baculovirus expression of <i>CYP6s</i> and <i>CYP9s</i> and metabolic assays
Coumaphos [291]	AChE inhibitors	Upregulated P450s from the <i>CYP2</i> clade	In hive acaricide treatment. Quantitative RT-PCR technique
Coumaphos [283]	AChE inhibitors	Increase in toxicity when bees treated with a P450 inhibitor	PBO synergism

Coumaphos [280]	AChE inhibitors	Upregulated P450s from the CYP2 and CYP3 clades	Adult dietary exposure & microarray analysis
<i>Pyrazole</i>			
Fenpyroximate (Phenoxy pyrazole) [283]	Mitochondrial complex I electron transport inhibitors	Increase in toxicity when bees treated with a P450 inhibitor	PBO synergism
<i>Mycotoxin</i>			
Aflatoxin B1 [288]	Carcinogen	Increase in toxicity when bees treated with a P450 inhibitor	PBO synergism
<i>Plant secondary metabolites</i>			
Nicotine (Alkaloid) [281]	nAChR agonist	Upregulated P450s from the CYP3 clade	Adult dietary exposure & metabolomic and proteomic profiling
Quercetin (Flavaonoid) [292]		Insecticide metabolism in specific expressed P450s (CYP6A subfamily)	Baculovirus expression of CYP6As and metabolic assays
Quercetin (Flavaonoid) [284]		Insecticide metabolism in specific expressed P450s (CYP3 clade)	Baculovirus expression of CYP6s and CYP9s and metabolic assays
<i>Honey constituent</i>			
<i>p</i> -coumaric acid [293]		Upregulated P450s from the CYP3 clade	Adult dietary exposure & RNAseq analysis

1.8 Concerns and risks to pollinator health

The importance of invertebrate declines has been highlighted by the fate of pollinators [16, 294-296] but it is clear that as a group, invertebrates provide many ecosystem services other than pollination, including, filtration of water, formation of soil and biological control [297]. Shocking reports of massive declines (>78%) in insect biomass over the last few decades raise questions over the impact to the diversity and health of pollinator species, in particular wild bees [19, 298-300]. Animal pollinators are implicated in the increase in yield of 70-84% of crop plants [301, 302]; their loss may threaten ecosystem resilience, local and global food security [15, 303]. With well documented declines in *A. mellifera* numbers due to *V. destructor*, the need to understand the risks to other bee species is pressing [17, 295, 304]. The responsibility for the loss of insect biodiversity has been attributed to habitat loss, agricultural intensification, use of insecticides, climate change and diseases [4, 36, 296, 300, 304, 305].

1.8.1 Habitat loss, degradation and loss of resource diversity

Urbanisation, changes in global land-use, loss of ecosystem diversity and habitat fragmentation have been linked to declining insect species [19, 305]. Angiosperms (flowering plants) are the most diverse group of land plants and the key adaptation linked to their success is animal-mediated pollination [306, 307]. The adaptive radiation of bees occurs in the Mid to Late Cretaceous, at the same time as the expansion in eudicot angiosperms [308]. As such, robust plant-pollinator networks have evolved between many angiosperms and bee species [305]. Although these networks incorporate species redundancy and behavioural flexibility, some ecological models predict sudden crashes in plant diversity with continued pollinator declines [309, 310]. Of all of the drivers associated with the decline in diversity and abundance of bee pollinators, habitat loss appears to be the most crucial [17, 304, 311]. Changes in land-use, such as forest clearance, urbanisation and the expansion of the road network lead to habitat loss and degradation [4, 312, 313]. This impacts bee pollinators through the loss of floral resources, nesting sites and nesting materials [4, 304, 312]. Mitigation of some of these effects is possible, for example by promoting flowering lawns or the removal of road-verge cuttings [312-314]. Although an important driver of land-use change, urban and built up land only covers 1% of the habitable land surface (see figure 1.15) [315].

1.8.2 Agricultural intensification

The largest change in global land-use has come from the intensification of agriculture [4, 305]. Land set aside for agricultural use, both as cropland and pasture, has become the largest terrestrial biome on the planet, covering 40-50% of all habitable land (see figure 1.15) [4, 315, 316]. Modern agricultural practices have been successful in increasing crop yields, however, they have also degraded and fragmented native habitats [4]. Mechanisation of tillage, the creation of monocultures and the use of chemical fertilizers, herbicides and insecticides are all implicated in habitat degradation through losses in diversity of soil biota, plant and insect species [40, 317, 318]. Biodiversity is also diminished by the reduction in land area covered by riparian buffer strips, hedgerows, thickets and woodland that could act as wildlife corridors [317]. Forestry practices also affect the balance of arthropod communities [19]. Like other arthropods, the level of sensitivity of bees to land use change and habitat fragmentation will depend on life history traits and physiological tolerance. A study of 613 bee taxa found that species nesting above ground showed a nine-fold greater reduction in abundance in response to agricultural intensification and habitat fragmentation, compared to ground-nesting species [319]. However, ground-nesting bees were more negatively affected by tillage [319]. In general, agricultural intensification is expected to be particularly disadvantageous to solitary bees because they tend to have more specific habitat requirements, including the need for nesting sites and materials as well as the close proximity of floral resources [320]. Unlike social species, solitary bees also lack any buffer against losses. Most solitary bees prefer to forage within 300m of their nesting site and there can be not only a decrease in offspring numbers when they are forced to extend their foraging distance, but also a change to the sex ratio of the offspring produced [321, 322]. In the face of an ever-increasing human population, the global demand for food can only increase, with the risk of rising food prices (predicted to be ~30-70% by 2050) and resulting heterogeneous food insecurity [7]. The challenge is to grow, harvest and store enough food in a resilient and sustainable manner, whilst also managing the threat of agricultural losses to pest species [5, 7, 109, 316]. Diversifying landscapes by the protection of natural habitat, increasing riparian buffer strips and planting wildflower areas, may increase wild bee diversity and abundance, thereby increasing the resilience of the agricultural ecosystem [4, 21, 320].

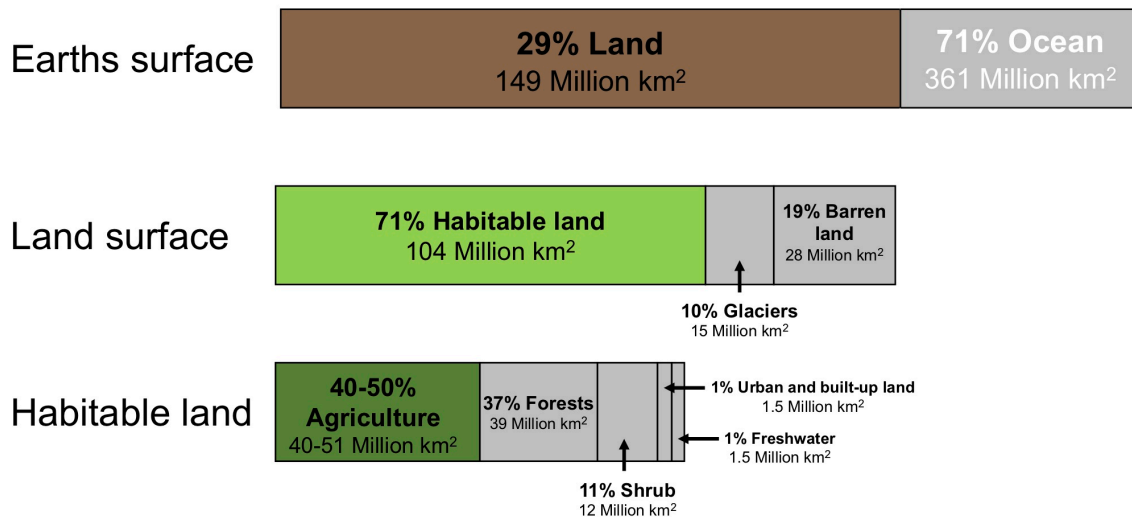


Figure 1.15: Global land use for food production. Amended from [315].

1.8.3 Use of insecticides

Intensive crop management often relies on the chemical control of pest insect species. Landscape and regional surveys have shown that the species richness of wild bees is lower where risk of exposure to insecticide is high [323]. A study of 613 bee taxa reported that social bee species showed a greater reduction in abundance (70%) compared to solitary species (29%) in response to pesticide use [319]. Incidences of inadvertent poisoning of bees and other non-target insects have been documented since the 1870's in North America [324]. In the UK, potential bee poisoning incidences are investigated under the auspices of the Wildlife Incident Investigation Scheme (WIIS), which allows for post-registration surveillance of pesticides [325]. Pesticide poisonings are evaluated from field information or residue analysis and are then assigned to one of four categories:

1. Approved use – the pesticide was used according to its approval.
2. Misuse – careless, accidental or wilful failure to adhere to the conditions of the approval of the pesticide.
3. Abuse – deliberate, illegal use of a pesticide to poison, or attempt to poison bees.
4. Unspecified – the source of the pesticide is uncertain [325, 326].

It is clear from figure 1.16 that the number of incidents of poisoning attributed to pesticides have been declining since the 1980s. To mark this trend more easily

it is helpful to divide the data into decades. In the decade 1988-1998 there were, on average, 27 incidences of poisoning attributed to pesticides per year, although it is clear from figure 1.16 that the number of cases was declining dramatically during the second half of the 1990s [325]. In the next decade (1999-2009) the average rate of cases fell to approximately 7 incidences per year, and in the decade (2010-2019) average cases continued to fall slightly to 6.5 per year [326-328]. This decline in bee poisoning events has also been seen in other developed countries, such as Germany [329]. These declines are attributed to changes in agricultural practices, more stringent regulations and improved product labelling [326]. The insecticides implicated in poisoning incidences up until the mid 2000s were predominantly organophosphates and carbamates, although several pyrethroids and two organochlorides were also identified [325, 326, 330]. From the mid 2000s to date, carbamates are still the predominant insecticides implicated in bee poisoning, however, neonicotinoid and pyrethroid insecticides are reported at similar levels to organophosphates for this period [328]. It should be noted that the data shown are almost exclusively for *A. mellifera*, with infrequent reports of bumblebee (*Bombus* spp) poisonings.

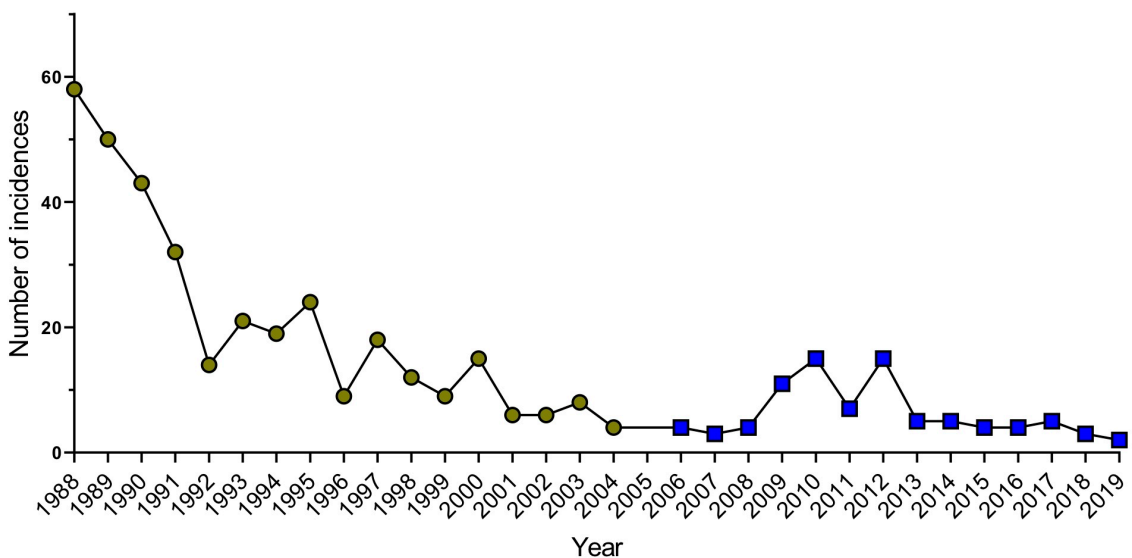


Figure 1.16: Total incidents of bee poisonings attributed to pesticide each year (1988-2019) in the UK. Data represented by a yellow circle are for the UK and are taken from [325, 326]; data represented by a blue square are for England and Scotland only and are taken from [327, 328, 330]. Graph drawn in GraphPad Prism V 8.1.0.

Although there has generally been a focus on the exposure of *A. mellifera* to pesticides and the levels of pesticide residues found in honeybee products, such as honey and wax, the level of pesticides found in pollen and nectar from plants near managed bees has also been investigated [331-334]. Insecticide residues have been found in wild flowers close to agricultural crops which poses an added risk not only to *A. mellifera*, but to all bee species foraging in the area [335, 336]. Due to the essential pollination services provided by wild solitary bees, the potential risk from insecticide residues to non-*Apis* species is of great concern [158].

Neonicotinoids, out of all the classes of insecticides, have received significant attention due to concerns over their effects on non-target insect species and their persistence in the environment [158, 337]. In May 2013 the European Commission brought in severe restrictions on the use of three *N*-nitroguanidine neonicotinoids: clothianidin, imidacloprid and thiamethoxam due to concern over safety to bees (Regulation (EU) No 485/2013) [338-340]. The restrictions were reinforced in May 2018, to encompass a total ban of the outdoor use of the three compounds (Regulation (EU) Nos 2018/783, 2018/784, 2018/785). The ban on these compounds has resulted in many farmers relying on alternative insecticides, including those requiring soil or foliar application, in particular pyrethroids [341]. The most commonly used of these was cypermethrin, a pyrethroid that is highly toxic to *A. mellifera* and *B. terrestris*: acute contact toxicity LD₅₀ value ~ 0.124-0.26 µg ai/bee and 0.17 µg ai/bee respectively [342-346]. There is therefore a risk that the insecticides being used to replace *N*-nitroguanidine neonicotinoids could be similarly toxic to bee species [346, 347].

However, in the decades since the 1990s, in most of the developed world the use of insecticides, measured as total weight of chemicals applied (tonnage) has generally been stable or has decreased (see figure 1.17). During the last few decades there has been a move towards expanding pest control science to incorporate the principles of applied ecology [348]. Integrated pest management (IPM) combines biological, physical and cultural tools with judicious use of chemical agents in the control of pest species [305, 349]. IPM encourages the

growth of healthy crops with the minimum amount of disruption to agricultural ecosystems [349].

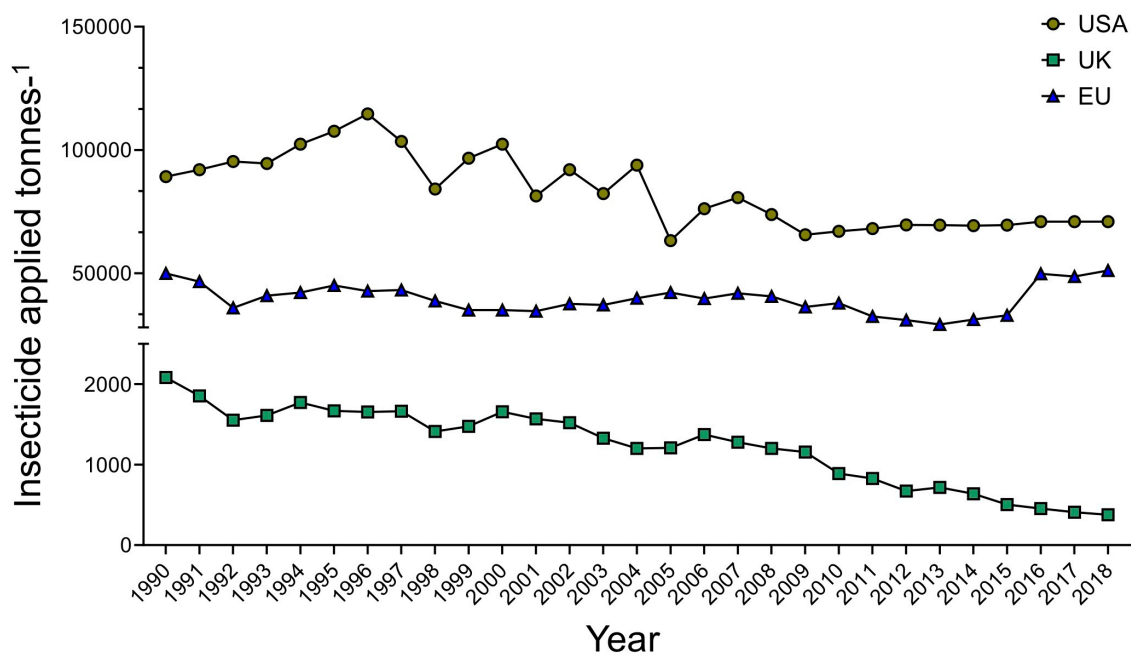


Figure 1.17: Insecticide used in tonnes per annum in the United States of America, the European Union and the United Kingdom (1990-2018). [Data taken from: Food and Agriculture Organisation of the United Nations [114] (FAOSTAT accessed 16/09/20); graph drawn in GraphPad Prism V 8.1.0.

1.8.4 Climate change

Climate change models predict strong upheavals that will impact natural systems with changes to rainfall patterns, melting ice and snow, retreating ice-caps, warmer temperatures, desertification and a greater frequency of extreme weather events [42, 350]. These effects will be spatially heterogeneous [351]. This spatial heterogeneity will be associated with a global increase in the number and intensity of extreme weather events, such as hurricanes, storms, tsunamis, wild-fires and floods [352]. These climate-events, rather than climate-trends are less predictable and therefore hold the potential for massive ecological impacts [353]. In general, bees do not forage in heavy rain or strong winds ($>5 \text{ m}^{-1}\text{s}^{-1}$), although some larger species such as bumblebees can be considered as ‘all-weather’ feeders and would therefore be more affected by the increased levels of precipitation and turbulent wind-flows associated with extreme weather events [354, 355]. The nature of the weather event will, in part, determine the impact on a species; for example, ground-nesting bees are

predicted to be particularly vulnerable to flooding events [354]. In a study of 216 tropical bee species found in the Carajás National Forest in Brazil, 95% of species were predicted to suffer population decrease due to climate change, using models based on bee distribution data from 1981-2020 [356]. In line with other studies the 'winners' were found to be wide-range habitat polylectic generalists, and the 'losers', habitat specialists and those species with parasitic, univoltine and oligolectic traits [319, 356, 357]. Losses of solitary species are expected to be more prevalent, particularly in above-ground cavity-nesters such as members of the Megachilidae family [319].

Studies indicate that rising environmental temperature under climate change, is likely to be the factor that has the greatest direct impact on insect populations [358]. Globally, temperature is predicted to reach 1.5°C above pre-industrial levels between 2030 and 2052 [42]. Insects are generally considered as ectotherms and as such, their physiology can be severely affected by the environmental temperature [359-361]. In many insects however, the ability to thermoregulate has evolved [362-364]. To initiate powered flight, insects need a relatively high body temperature (>30°C) and so many species including bees, moths, butterflies and beetles have evolved pre-flight thermogenesis via shivering and substrate cycling [362, 363, 365, 366]. Conversely, to reduce their body temperature below the environmental temperature, insects use ventilatory mechanisms, both external through flapping wings or moving legs, and internal, via the tracheal system [362, 367]. However, in general, warmer conditions are expected to accelerate the metabolic rate of insects, which in turn will increase their need to consume food, something that is expected to increase crop losses to agricultural pest species [43]. Rising temperatures and drought will likely result in a decreased production of flowers and nectar, which in turn would reduce, not only crop yields, but also the overwintering survival of pollinator populations [354]. Insects can mitigate the adverse impacts of climate change by acclimation, phenological changes, behavioural plasticity, shifting their geographical ranges and evolutionary adaptation [368].

Acclimation and range-shifting

Insect species found in high latitudes are expected to have broad thermal tolerances and so may cope with increases in temperature [368]. Whereas even

small degrees of warming will be associated with insect population declines in the tropics, as the species found there are already living close to their optimal temperature [368]. In general, warming has caused range expansion towards the poles and higher elevations in animal species [369-373]. Land use change and habitat fragmentation will impact how possible it is for insect species to shift their ranges in order to offset warming [368].

Phenotypic adaptation

Phenotypic adaptation is the response of a population to environmental change that occurs without genetic mutation and is often referred to as plastic change [374]. Environmental changes, particularly warming temperatures, affect the behaviour of insect species, such as changes in seasonal timings (phenology) and life history traits [360]. Phenotypic adaptations include changes to development time, diapause, voltinism, colouration and body size [375, 376].

In general, the trend in the phenology of insects is an advancement of spring events and delay of autumn events [373, 376-379]. The trend for earlier emergence is particularly marked in bee species that are active in the spring and those that overwinter as adults, such as the European orchard bee (*Osmia cornuta*) [380, 381]. Overall the foraging season of bees is predicted to increase in duration with increasing temperatures [381].

The primary concern with changes in phenology is the potential for phenological mismatch or trophic asynchrony between interacting species, such as plant-pollinator networks [377]. Most of the data for climate related impacts on bees are for generalist, polylectic species that are able to collect pollen from diverse plant species. Oligolectic bee species, that are more dependent on particular plants and are therefore more susceptible to phenological mismatch, have rarely been included in studies [377]. In general, however, it appears that specialist species, of both plants and pollinators, will be more at risk from phenological mismatch than generalists [382].

Evolutionary adaptation

Adaptive evolution occurs when genetic change alters the phenotype [374]. Evolutionary responses to environmental change can include changes in body

size, thermal sensitivity, seasonality traits and colouration [376, 383, 384]. *A. mellifera* has an almost global occurrence, in highly diverse climatic conditions, indicating it has great social resilience and perhaps indicates a potential for significant adaptive capacity [355]. Conversely, the Asian species of honeybee, such as the Eastern honeybee (*A. cerana*), the giant honeybee (*A. dorsata*) and the dwarf honeybee (*A. florea*) have not expanded their ranges outside of Asia, which may indicate they will be less able to adapt to climate change [355]. However, overall there is a paucity of evidence of the impact of evolutionary responses to mitigate the effects of climate change in the context of insects as a whole [368]. The fossil record of the Pleistocene glacial and interglacial periods indicates that in response to the changes in climate, species tended to shift their distribution, rather than remaining in their home ranges and evolving new forms [375].

1.8.5 Pests and pathogens

All bees have a wide range of parasitoids, parasites and pathogens (including fungi, bacteria, protozoans and viruses) found in association with them [26]. The role of these pests and pathogens in bee mortality and colony-collapse disorder in *A. mellifera* has received much attention [202, 385]. It has become evident that many bee diseases and parasites have been allowed to spread outside of their natural geographical ranges and hosts by the commercial trade in honeybee (*Apis* spp.) and bumblebee (*Bombus* spp.) colonies [26, 386, 387]. This type of pathogen spillover has been well-studied in vertebrates, particularly humans, where it is often associated with increased virulence and mortality in the naïve host population [388, 389].

V. destructor was originally described as a mite parasitising *A. cerana* in Asia, but since its host shift to *A. mellifera*, it is now found globally, with the exception of Australia, and is a known vector of many lethal honeybee viruses, such as Deformed wing virus (DWV) [385, 387]. A similar host shift from *A. cerana* to *A. mellifera* has occurred with the fungal microsporidian *Nosema ceranae*, which has now spread from Asia to Europe and the Americas [26]. Studies have highlighted the shift of several viruses, microsporidia and parasites associated with *Apis* spp. to other host species from at least three bee families (Megachilidae, Andrenidae and Apidae) [386, 387, 390, 391]. The most

common pathogen infecting *M. rotundata* is the fungus *Ascosphaera aggregata* Skou that causes the devastating larval disease, chalkbrood, [56, 392]. Although *A. aggregata* is related to the *A. apis*, the fungus that causes chalkbrood in *A. mellifera*, there appears to be no spillover or cross infection [56]. However, there is an undoubted potential for disease spillover from managed *A. mellifera* and *B. terrestris* colonies to populations of wild bees [391]. The extent of the negative impacts of this remain largely unknown [387].

1.8.6 Risks pertinent to *M. rotundata*

The biology and life cycle of *M. rotundata* create a different set of risks from insecticide exposure than those faced by *A. mellifera* (see figure 1.18). There are five main areas of difference:

1. *M. rotundata* is among the smallest leafcutter bees, with females having an average body weight of 35 mg [48]. They are considerably smaller than *A. mellifera* (hatching adult worker bee weight 100 mg [49]). If exposed to insecticide spray, *M. rotundata* will receive a higher dose of insecticide, in relative terms, due to its smaller size and higher surface-area-to-volume ratio [393-395].
2. *M. rotundata* has a significantly smaller flight range than that of *A. mellifera*. When foraging in a field that has been exposed to insecticide, there is less chance that uncontaminated provisions from other fields would be added to, and therefore dilute, contaminated pollen and nectar before an egg is sealed in a nest cell [396].
3. *M. rotundata* larvae feed directly on a provision ball made up solely of pollen and nectar whereas the brood of *A. mellifera* are fed on secretions from nurse bees [396, 397]. For the first 3-4 days *A. mellifera* larvae consume a caste-specific jelly produced from the hypopharyngeal glands of adult worker bees which contains no pollen [398]. Subsequently they are fed a modified form of the jelly which contains pollen [398]. It may be that nursing the larvae in this manner protects them against xenobiotics [397].
4. Female *M. rotundata* not only forage for pollen and nectar, they also cut and collect leaf pieces to build their nests. These leaf pieces are a further source of exposure to both adults and developing larvae [396].
5. If exposed to a fatal dose of insecticide, all reproductive capability for the individual *M. rotundata* is finished, whereas the loss of a single *A. mellifera*

worker may be of little or no consequence, to the reproductive capability of her colony [396].

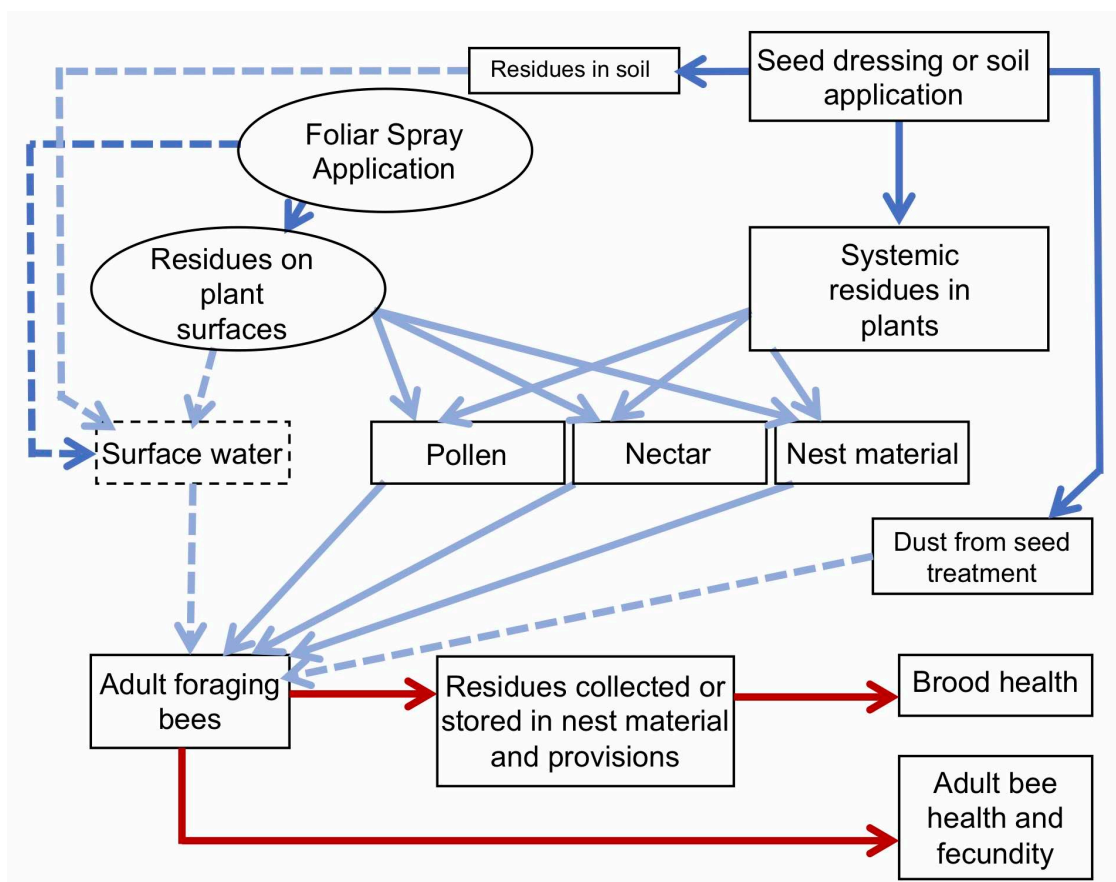


Figure 1.18: The potential routes of exposure to insecticides to *M. rotundata* (solid lines represent primary routes and dashed lines secondary routes). Amended from [396].

1.9 Aims and objectives of this PhD

The overarching aim of this PhD was to understand the molecular and biochemical basis of the sensitivity/tolerance of the economically important solitary bee pollinator *M. rotundata* to select insecticides. Specifically, this thesis addresses (1) the repertoire of P450 enzymes present; (2) the *in vivo* sensitivity/tolerance to select insecticides; (3) the functional characterisation and metabolic profiling of key candidate P450s in this species. It also assesses the evidence for the evolution of a CYP9 P450 profile similar to that found in *M. rotundata*, within the *Megachile* genus and Megachilidae family, but then more widely across the Hymenopteran order. This thesis will contribute towards the development of a bee ‘tool-kit’ which can facilitate future screening of lead

compounds developed for use as insecticides. The key scientific questions underpinning each experimental chapter are briefly outlined below.

1.9.1 Chapter 3

This chapter addressed the question of whether the genome of *M. rotundata* encodes P450 enzymes belonging, or closely related to, the CYP9Q and CYP9BU lineages that have been shown to be able to metabolise insecticides from three classes in other bee species (see section 1.7.3).

1.9.2 Chapter 4

This chapter addressed the important question of whether the intrinsic differential sensitivity to certain insecticides observed in *A. mellifera*, *B. terrestris* and *O. bicornis* extends to *M. rotundata* (see section 1.7.3).

1.9.3 Chapter 5

This chapter addressed the question of the functional significance of the CYP9 repertoire reported in Chapter 3 in relation to the sensitivity/tolerance to insecticides observed in Chapter 4.

1.9.4 Chapter 6

This chapter examined the question of how often the array of CYP9 P450s described in Chapter 3 is found in other species of bees.

Chapter two: General materials and methods

2.1 Materials

2.1.1 Kits

CloneJET PCR Cloning Kit (Thermo Scientific)

E.Z.N.A.® Insect DNA Kit (OMEGA bio-tek)

GeneJET Plasmid Miniprep Kit (Thermo Scientific)

Isolate II RNA Mini Kit (Bioline. Meridian Life Science)

Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Scientific)

Monarch® DNA Gel Extraction Kit (New England BioLabs Inc.)

Monarch® PCR and DNA Cleanup Kit (5 µg) (New England BioLabs Inc.)

QIAquick® Gel Extraction Kit (QIAGEN)

QIAquick® PCR Purification Kit (QIAGEN)

Qubit® dsDNA BR Assay Kit (Molecular Probes Life Technologies)

Qubit® RNA BR Assay Kit (Molecular Probes Life Technologies)

Qubit® RNA IQ Assay Kit (Molecular Probes Life Technologies)

SuperScript™ III First-Strand Synthesis SuperMix (Invitrogen)

2.1.2 Solutions and buffers

0.1M Na/K-phosphate buffer, pH 7.6 (potassium-phosphate buffer; see appendix table 2.1)

Homogenisation buffer, pH 7.4 (see appendix table 2.2)

Buffer R, pH 7.6 (see appendix table 2.3)

Phosphate buffered saline (PBS – Sigma)

Tween 20 (Acros Organics)

Formaldehyde (Sigma-Aldrich)

Stop solution (see appendix table 2.4)

50% (w/v) sucrose (Tate and Lyle granulated white sugar)

LB agar plates and LB broth (see appendix table 2.5)

2.2 Molecular methods

2.2.1 Centrifugation

Small volume (<2 ml) centrifugation steps were carried out in model 5424 (Eppendorf) and 5418 (Eppendorf) microcentrifuges, or a model 5430R

(Eppendorf) centrifuge. Centrifugation of larger volumes (>2 ml) was carried out either in a model 5430R (Eppendorf) or model 5810R (Eppendorf) centrifuge.

2.2.2 RNA extraction

Bees were snap frozen in liquid nitrogen and stored at -80°C prior to use in RNA extraction. Specimens were weighed before being ground to a fine powder using a pre-cooled, RNase-free ceramic mortar and pestle in combination with liquid nitrogen. RNA was extracted using a Bioline Isolate II RNA Mini Kit (Bioline Reagents) following the manufacturers recommended protocol. The amount of lysis buffer added to tissue from larger specimens (>30 mg) was doubled and the sample was split across two columns. RNA was eluted in 30 µl RNase-free water and stored at -80°C.

The quality of RNA was assessed from the ratios of the absorbance maxima (A_{260}/A_{280} and A_{260}/A_{230}) measured, using 1 µl of sample, on a NanoDrop One (Thermo Scientific) or NanoDrop 2000 spectrophotometer (Thermo Scientific). RNA quantity and quality was further assessed by analysis of samples in a Qubit 4 fluorometer (Invitrogen by Thermo Fisher Scientific). 1 µl of sample was used with the Qubit® RNA BR Assay kit to assess RNA quantity and a further 1 µl with the Qubit® RNA IQ Assay kit to assess the RNA quality and integrity following the manufacturers protocols. Samples were accepted as pure with ratios: $A_{260}/A_{280} \sim 2.0$ and $A_{260}/A_{230} > 1.9$.

RNA integrity was visualised on an agarose gel. 0.5 µl of RNA sample was incubated with 4 µl 2X RNA loading dye (Thermo Scientific) and 3.5 µl nuclease-free water at 70°C for 10 minutes. Samples were run on a 1% TAE agarose gel (see section 2.2.5), containing 1 µl RedSafe nucleic acid staining solution (20,000 X) (iNtRON biotechnology) per 20 ml agarose, at 75V/200 mA for 45-60 minutes.

2.2.3 First strand cDNA synthesis

Maxima H Minus First Strand cDNA Synthesis Kit (Thermo Scientific) or SuperScript™ III First-Strand Synthesis SuperMix (Invitrogen) was used to synthesise complementary DNA (cDNA) from extracted RNA for use in PCR reactions following the manufacturer's protocols. 1-2 µg of cDNA was

synthesised in each 20 µl reaction. cDNA was stored at -80°C or -20°C for use in the short-term.

2.2.4 DNA extraction

Bees were snap frozen in liquid nitrogen and stored at -80°C prior to use in DNA extraction. Specimens were ground to a fine powder using a pre-cooled, DNase-free ceramic mortar and pestle in combination with liquid nitrogen. DNA was extracted using the E.Z.N.A.® Insect DNA Kit (OMEGA bio-tek) following the manufacturers recommended protocol. The incubation time used with the Proteinase K solution was 2-4 hours. For specimens yielding >50 mg tissue, the quantities of CTL Buffer and Proteinase K solution were doubled and the sample split across two columns. DNA was eluted in 50 µl elution buffer and stored at -20°C.

The quality of DNA was assessed from the ratios of the absorbance maxima (A_{260}/A_{280} and A_{260}/A_{230}) measured, using 1 µl of sample, on a NanoDrop One (Thermo Scientific) or NanoDrop 2000 spectrophotometer (Thermo Scientific). Samples were accepted as pure with ratios: $A_{260}/A_{280} \sim 1.8$ and $A_{260}/A_{230} > 1.9$. DNA quantity was further assessed by analysis of samples in a Qubit 4 fluorometer (Invitrogen by Thermo Fisher Scientific). 1 µl of sample was used with the Qubit® dsDNA BR Assay Kit to assess DNA quantity.

2.2.5 Polymerase chain reaction (PCR)

2.2.5.1 Primer design

Primers for use in this PhD were designed using the following protocol:

- a) The genetic sequence of interest was visualised in Geneious version 10.2.3 (Biomatters) and potential sites for primers identified
- b) Potential primer sequence design was checked using Primer3web version_4.0.0
https://primer3plus.com/primer3web/primer3web_input.htm
- c) Sequences that met the following criteria were selected: length – 18-26 bp; salt adjusted primer melting temperature T_m (°C) – 58-63°C; GC content (%) – 40-60%

- d) Primers were checked for secondary structures and self-complimentarity using the online Thermo Fisher Scientific Multiple Primer Analyzer tool

<https://www.thermofisher.com/uk/en/home/brands/thermo-scientific/molecular-biology/molecular-biology-learning-center/molecular-biology-resource-library/thermo-scientific-web-tools/multiple-primer-analyzer.html>

Primer stocks were synthesised by Sigma-Aldrich (Poole, UK) at a concentration of 100 µM. A working stock for use in PCR, at 10 µM, was made using nuclease-free water. All primers were stored at -20°C. See appendix table 2.6 for primer sequences used in this PhD.

2.2.5.2 PCR

All PCR reactions were carried out in 0.2 ml thin-walled PCR tubes (Star Lab) and run using a thermal cycler (T100™ Thermal Cycler, Bio-Rad). The annealing temperatures were determined using the online New England BioLabs Inc. T_m calculator (<https://tmcalsculator.neb.com/#!/main>). A control replicate, containing nuclease-free water in place of template DNA, was performed alongside each PCR reaction.

2.2.5.2.1 TA end PCR product

To produce a sticky-end PCR product, where a deoxyadenosine triphosphate (dA) is added to the 3'-end of the amplified DNA, DreamTaq Green PCR Master Mix (2X) (Thermo Scientific) and MyTaq™ Red Mix (Bioline Meridian Bioscience) were used. Reactions with both Taq polymerase mixes were performed following the manufacturer's recommended protocols and thermal cycling conditions as outlined in table 2.1, using 50-100 ng of template DNA.

Table 2.1: Cycling conditions used for PCR using Taq polymerase (times for DreamTaq Green PCR Master Mix (2X) (Thermo Scientific) and MyTaq™ Red Mix (Bioline Meridian Bioscience) shown in black and red respectively)

Step	Temperature °C	Time (min:sec)	Number of cycles
Initial denaturation	95	3:00 (1:00)	1
Denaturation	95	0:30 (0:15)	35
Annealing	60	0:30 (0:15)	
Extension	72	1:30 (1:30)	
Final extension	72	5:00 (5:00)	1

2.2.5.2.2 Blunt end PCR product

To produce a blunt-end PCR product Phusion® high-fidelity DNA polymerase (New England BioLabs Inc.) was used. Reactions were performed following the manufacturer's recommended protocol and thermal cycling conditions as outlined in table 2.2, using 50-100 ng of template DNA.

Table 2.2: Cycling conditions used for PCR using Phusion® high-fidelity DNA polymerase (New England BioLabs Inc.)

Step	Temperature °C	Time (min:sec)	Number of cycles
Initial denaturation	98	0:30	1
Denaturation	98	0:10	35
Annealing	62	0:30	
Extension	72	1:30	
Final extension	72	10:00	1

2.2.6 Agarose gel electrophoresis

PCR products, colony PCR products and recombinant baculovirus DNA were visualised by gel electrophoresis. Gels were prepared using genetic analysis grade agarose (Fisher BioReagents) and Tris-acetate-EDTA (TAE) buffer (Thermo Scientific) containing 1 µl RedSafe nucleic acid staining solution (20,000 X) (iNtRON biotechnology) per 20 ml agarose. Where necessary 1 µl of 6X loading dye (Thermo Scientific or New England BioLabs Inc.) was added to 5 µl product. A 1 kb Plus DNA ladder (Thermo Scientific) was run alongside samples to allow for the estimation of the size of products. In general products were run at 75V/200 mA for 45-60 minutes. Products were visualised on a 302 nm UV trans-illuminator (Molecular Imager® Gel Doc™ XR+ Imaging System, Bio-Rad).

2.2.7 PCR purification

2.2.7.1. Column purification

Column purification of PCR DNA product was performed to remove enzymes, dNTPs, primers and short-failed PCR products (<300 bp). Monarch® PCR and DNA Cleanup Kit (5 µg) (New England BioLabs Inc.) or QIAquick® PCR Purification Kit (QIAGEN) kits were used and purification performed following the manufacturer's recommended protocols. The quality of DNA was assessed as described in section 2.2.4. The purified product was stored at -20°C.

2.2.7.2 Gel extraction and clean up

Where it was necessary to isolate and purify DNA fragment based on size, gel extraction and clean up was performed. PCR products were run on a 1.5% TAE agarose gel (see section 2.2.6), at 75V/200 mA for 45-60 minutes. Bands of the desired size were excised from the gel with a scalpel, using a blue light safe transilluminator (Dark Reader DR-46-B, Clare Chemical Research). Monarch® DNA Gel Extraction Kit (New England BioLabs Inc.) or QIAquick® Gel Extraction Kit (QIAGEN) kits were used and purification performed following the manufacturer's recommended protocols. The quality of DNA was assessed as described in section 2.2.4. The purified product was stored at -20°C.

2.2.8 PCR Cloning

2.2.8.1 Ligation into plasmid vector using T4 ligase

DNA fragments (~150 ng per reaction) were ligated into a pJET1.2/blunt cloning vector using the CloneJET PCR Cloning Kit (Thermo Scientific) and following the manufacturer's recommended protocol. Reactions were carried out in 0.2 ml thin-walled PCR tubes (Star Lab) and the ligation mixture was incubated at room temperature for 30 minutes.

2.2.8.2 PCR cloning of Gateway® entry clones

PCR product flanked with bacteriophage λ *attB* sites were inserted into pDONR™221 donor vector, following the protocols set out in the Gateway® Technology with Clonase® II user guide (Invitrogen). See sections 5.2.2.2.1 and 5.2.2.2.2 for detailed description of methods.

2.2.8.3 Transformation using chemically competent *E. coli* cells

Library Efficiency® DH5 α ™ (Invitrogen) or Max Efficiency® DH5 α ™ (Invitrogen) chemically competent *E. coli* cells were used as host cells for transformation.

1.5 μ l recombinant pJET1.2/blunt vector was used with 35 μ l Max Efficiency® DH5 α ™ (Invitrogen) competent *E. coli* cells, following the manufacturer's recommended protocol. Transformed cells were spread onto LB agar plates (see appendix table 2.5), containing ampicillin (100 μ g/ml) and incubated overnight at 37°C.

1 μ l recombinant pDONRTM221 vector was used with 100 μ l Library Efficiency[®] DH5 α TM (Invitrogen) competent *E. coli* cells, following the manufacturer's recommended protocol. Transformed cells were spread onto LB agar plates containing kanamycin (50 μ g/ml) and incubated overnight at 37 °C.

2.2.8.4 Colony PCR

The success of the recombination reaction was verified via colony PCR, using primers specific to the vector. pJET 1.2 (10 μ M) and M13 (10 μ M) primers were used with pJET1.2/blunt cloning vector and pDONRTM221 vector respectively. DreamTaq Green PCR Master Mix (2X) (Thermo Scientific) and MyTaqTM Red Mix (Bioline Meridian Bioscience) were used as described in section 2.2.5.2.1 and table 2.1, but with *E. coli* cells from the transformation LB agar plates used as DNA template. A fresh LB agar plate, containing the appropriate antibiotic, was seeded with bacteria from individual colonies taken from the transformation plates, using a sterile 10 μ l pipette tip (Star Labs), which was then placed into a 0.2 ml thin-walled PCR tubes (Star Lab) containing the PCR mastermix and primers. The seeded LB agar plates were incubated overnight at 37°C. Between 6-12 colonies were screened for each DNA insert. To confirm the presence of the insert, 5 μ l of PCR product was run on a 1% TAE agarose gel (see section 2.2.6).

2.2.8.5 Plasmid DNA minipreps

The GeneJET Plasmid Miniprep Kit (Thermo Scientific) was used to extract plasmid DNA from selected recombinant *E. coli* colonies, following the manufacturers recommended protocol. Using the TAE agarose gel of the colony PCR as a reference, individual colonies were selected and using a sterile 10 μ l pipette tip (Star Labs) 5 ml of LB broth (see appendix table 2.5), supplemented with the appropriate antibiotic was inoculated. The culture was incubated overnight for approximately 16 hours at 37°C while shaking at 225 rpm. Cells were harvested and pelleted by centrifugation at 3000 rpm for 15 minutes. DNA was eluted in 30 μ l nuclease-free water, pre-heated to 65-70°C. The quality of the plasmid DNA was assessed as described in section 2.2.4. The purified product was stored at -20°C.

2.2.8.6 Evaluation of plasmid DNA

The sequence of the inserted DNA fragment was determined by Eurofins Genomics using automated Sanger sequencing. Premixed samples (DNA and primer) were submitted to Eurofins Genomics using their TubeSeq Service in accordance with their submission criteria.

<https://eurofinsgenomics.eu/en/custom-dna-sequencing/eurofins-services/tubeseq-service/> Insert DNA sequence was verified by visualisation in Geneious version 10.2.3 (Biomatters).

2.3 Biochemical methods

2.3.1 Bradford's protein assay

The concentration of protein preparations, both native microsomes and expressed P450s, were determined according to Bradford (1976), using Bradford reagent (Sigma-Aldrich). A serial dilution of the protein standard bovine serum albumin (BSA) (Sigma-Aldrich) and the protein preparations was carried out in 25 µl buffer R, pH 7.6 (10, 5, 2.5, 1.25, 0.625, 0.3125 mg/ml and 1:1, 1:2, 1:4, 1:8, 1:16 and 1:32 respectively). 250 µl Bradford reagent was added to 5 µl of each dilution and plates were left for 15-20 minutes at room temperature. Technical replicates were performed in triplicate. All absorbance readings were taken in clear, flat-bottomed 96-well plates (CytoOne), at 595 nm using a SpectraMax Plus 384 microplate reader (Molecular Devices).

2.4 Care and maintenance of insect cell lines

2.4.1 Initiation of insect cell lines from frozen stocks

The handling of all insect cell lines was performed in a sterile laminar flow hood (Astec Microflow Peroxide Class II) using aseptic techniques. Stocks of Sf9 and High Five™ cells, suspended in freezing medium (see table 2.3) and sealed in 1.5ml cryovials were stored in a cryogenic dewar containing liquid nitrogen. A frozen cryogenic vial was thawed rapidly in a 37°C water bath until almost completely defrosted. The cell suspension was transferred immediately, dropwise, into a T-25 treated tissue flask (CytoOne) containing 4 ml of the appropriate medium at room temperature. The flask was rocked gently to spread the cells evenly and then placed in a 27°C incubator (Sanyo MIR 553) for 30-45 mins to allow the cells to attach to the flask surface. Visual inspection using a microscope (Leitz DM IL) was undertaken to confirm attachment. The

media was aspirated to remove debris, dead cells and the DMSO found in the freezing medium and replaced with 5 ml of fresh media at room temperature. The flask was then returned to the 27°C incubator (Sanyo MIR 553) for 24 hours or until the cells reached ~90% confluency. Cells were passaged and expanded once this stage of confluency was reached.

Table 2.3: Insect cell line frozen stocks

Cell Line	Freezing Medium	Density (cells/ml)	Media type
Sf9	60% Graces Insect Medium 30% FBS 10% DMSO	1×10^7	Sf-900™ II SFM (Gibco™ – Thermo Fisher)
High Five™	42.5% conditioned Express Five™ SFM 42.5% fresh Express Five™ SFM 10% DMSO 5% FBS	3×10^6	Express Five™ SFM (Gibco™ – Thermo Fisher)

2.4.2 Passaging cells

Sf9 cells

Cells were passaged and expanded when in mid-log phase, typically once ~90% confluency was reached. Cells were dislodged from the monolayer by tapping the sides of the flask. Visual inspection using a microscope (Leitz DM IL) was undertaken to confirm cells had dislodged. Once ~50% of the cells had dislodged the media containing the cells was aspirated and transferred to a fresh T-25 treated tissue flask and the cells were spread evenly by gently rocking the flask. 5 ml fresh media at room temperature was added to the initial flask. Both flasks were placed in a 27°C incubator (Sanyo MIR 553). Sf9 cell cultures were passaged approximately every 2-3 days and High Five™ cell cultures every 2 days. To create a suspension culture all the cells were dislodged from the monolayer. Once ~100% of the cells were dislodged, the cell suspension was aspirated and transferred to a sterile 100 ml Erlenmeyer flask containing 5 ml of fresh media which was placed on a rocker plate (110 rpm; Thermo Scientific) in the 27°C incubator (Sanyo MIR 553). Cells in suspension were grown to a density of $\sim 2 \times 10^6 - 3 \times 10^6$ viable cells/ml (>90% viability), at which point they were still in log phase. Sf9 cells were then passaged at a density of $\sim 8 \times 10^5$ cells/ml and High Five™ cells at a density of $\sim 5 \times 10^5$ cells/ml in 25 ml media on a rocker plate (110 rpm; Thermo Scientific) in the 27°C incubator (Sanyo MIR 553).

High Five™ cells

In general, the protocol described above for Sf9 cells was followed, however, High Five™ cells aggregate to form large clumps and so the suspension culture was initially supplemented with heparin (Sigma) at 10 U/ml of culture. High Five™ cells were grown for several passages before being weaned from heparin. The Express Five™ SFM (Thermo Fisher) media needed to be supplemented with L-glutamine (Thermo Fisher) at a final concentration of 16 mM prior to use.

2.4.3 Determining cell density and viability

Prior to passaging, the density and viability of the suspension culture was assessed using 0.4% trypan blue solution (Gibco™). A 1:5 dilution of the suspension cell culture was made with 0.4% trypan blue solution. 10 µl of this was added to a haemocytometer slide (Neubauer chamber, Marienfeld) and placed under a microscope (Leitz DM IL). Cells that had taken up the trypan blue were considered to be non-viable. Cell density was calculated using the following formula:

$$\text{Cell density } \left(\frac{\text{cells}}{\text{ml}} \right) = \text{Number of viable cells} \times \text{dilution factor} \times 10,000$$

Cell density and viability were calculated at least twice and the average figures were calculated.

Chapter three: Curation of the CYPome of *M. rotundata* and comparison to other bee species

3.1 Introduction

3.1.1 Bee evolution

Bees (Order: Hymenoptera; Superfamily: Apoidea; Clade: Anthophila) are thought to have evolved from a carnivorous wasp ancestor in the early to mid-Cretaceous, approximately 100-149 mya [64, 72, 399]. The transition from a carnivorous diet to pollinivory is a significant driver of the rapid diversification and proliferation in bee lineages [399, 400]. Bees are monophyletic and consist of seven families that broadly fall into two groups: the long-tongued (L-T) bees and the short-tongued (S-T) bees (see figure 3.1(a)) [72, 399]. The Melittidae family is sister to all other bee families (see figure 3.1(a)), and so it is likely that the ancestral 'proto-bee' was melittid-like [399]. L-T bees is a monophyletic group that contains two major bee families, the Apidae and the Megachilidae, which together account for ~49% of all bee species (see figure 3.1 (a) and (b)). The Megachilidae family, which contains *M. rotundata*, diverged from the common ancestor of all L-T bees 104-125 mya [64]. The long timescale from the last common ancestor of the L-T bees and the rapid expansion of bee lineages at that timepoint, allows for the possibility that diverse genomes have evolved in the Apidae and Megachilidae families [399]. Bees are very important wild and managed agricultural pollinators [36]. This chapter will primarily focus on three managed pollinator species: *A. mellifera*, *B. terrestris* and *O. bicornis*, in addition to *M. rotundata*.

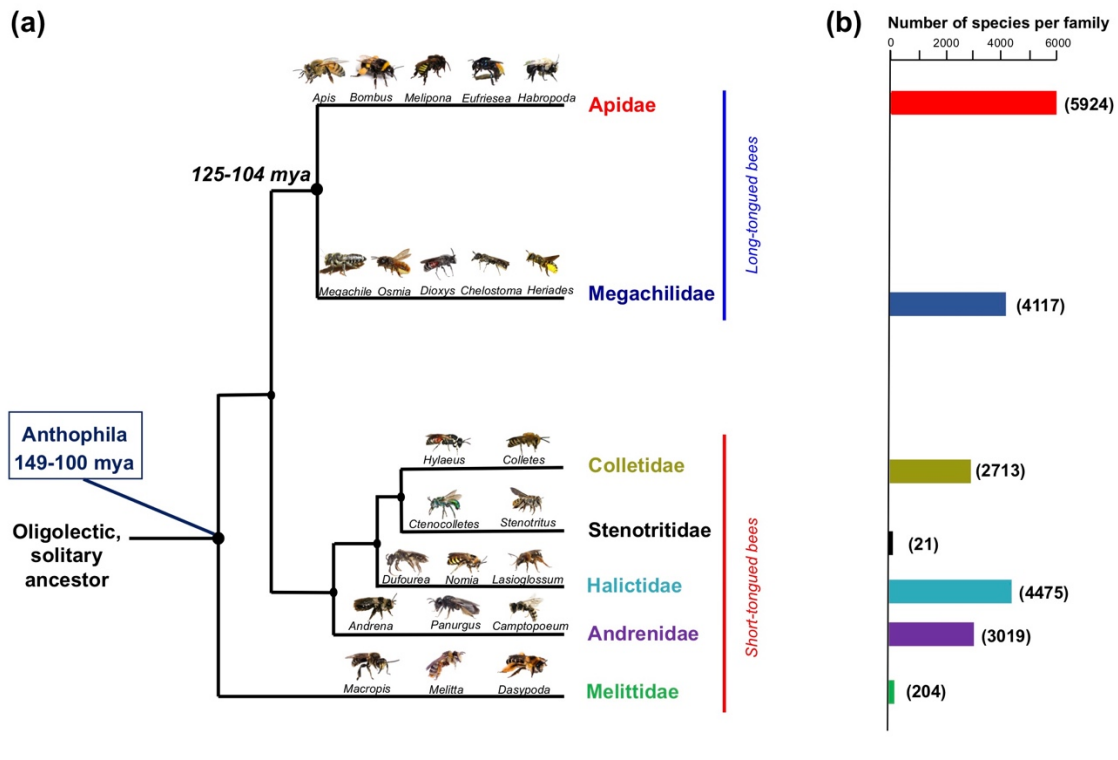


Figure 3.1 (a): Evolutionary history of bees showing the topology of relationships between families. Genera shown: *Apis* [401], *Bombus* [402], *Melipona* [403], *Eufreisea* [404], *Habropoda* [405], *Megachile* [406], *Osmia* [407], *Heriades* [408], *Andrena* [409], *Melitta* [410], *Dioxys* [411], *Chelostoma* [412], *Hylaeus* [413], *Colletes* [414], *Ctenocolletes* [415], *Stenotritus* [416], *Dufourea* [417], *Nomia* [418], *Lasioglossum* [419], *Panurgus* [420], *Camptopoeum* [421], *Macropis* [422], *Dasypoda* [423] (b): Bar chart showing approximate number of species for each bee family. Number of species per family taken from [424].

3.1.2 Comparative genomics

Comparative genomics is a discipline that reveals the evolutionary relationship between species. It relies on both sequencing technology and computational power in order to compare two or more genomes with accuracy [425-427]. Evolutionary processes such as sequence mutations, gene duplication, gene loss and chromosomal rearrangements can often make marked differences in genomes, even between those of closely related species [427-429]. After speciation, independent evolution allows for diversifying variations to accumulate, which can result in the genomes containing mosaics of chromosomal regions [430]. One of the primary tools of comparative genomics

is multiple sequence alignment (MSA) however, different MSA tools do not always give identical results [431]. An additional strategy that can be used alongside MSAs is the identification of genomic regions showing conserved synteny (collinearity), or conserved linkage [427, 432]. Synteny is the presence of two or more orthologous genes on the same scaffold/chromosome and can be thought of as evidence for residual ancestral genetic arrangement [433, 434]. Conserved synteny is defined as two or more homologous genes that are syntenic in two or more species, irrespective of gene order [432, 435]. When genes show conserved synteny and gene order is maintained, then it is defined as conserved linkage [432, 435]. The term synteny block (SB) is used to describe synteny, conserved synteny and conserved linkage (see figure 3.2) [427]. SBs have been used as a proxy for the degree of genome rearrangement [435, 436].

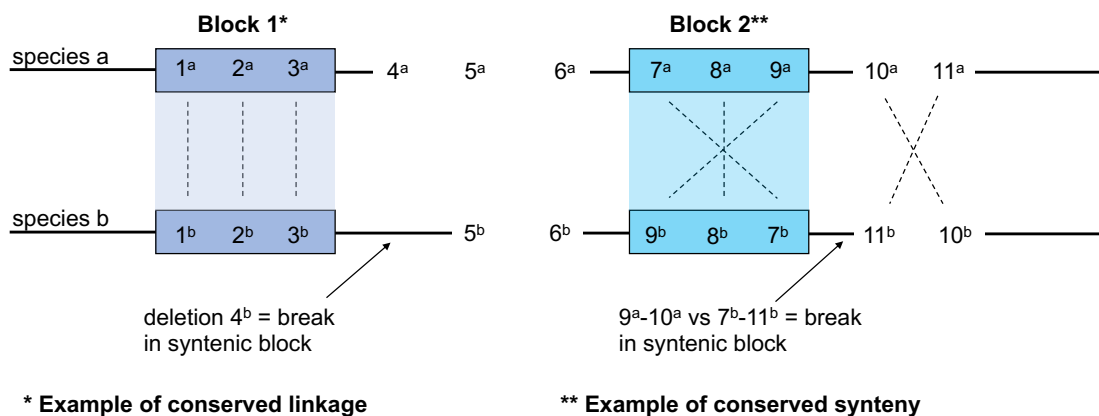


Figure 3.2: Examples of synteny blocks (SBs) showing conserved synteny (**) and conserved linkage (*) with possible scenarios that lead to breaks in SBs. Genes are denoted with numbers, and species by a super-script letter (^a / ^b). Orthologous genes are linked by a dashed line. The minimum number of orthologs that constitute an SB is shown here as three.

Synteny can also be thought of in terms of macro-synteny and micro-synteny [437], where macro-synteny refers to the conservation of chromosomal content and micro-synteny to smaller regions of conserved gene neighbourhoods [437].

3.1.3 Molecular phylogenetic analysis

A phylogenetic tree is a non-reticulated graph that is generally depicted as a bifurcating diagram [438, 439]. The external nodes or tips of the tree are referred to as leaves and represent the data inputted into the estimation. The internal nodes of the tree represent hypothetical ancestral data [440]. A 'gene' tree infers the evolutionary relationship of DNA, RNA or protein sequences, with the internal nodes representing hypothetical ancestral sequences [441]. Branches connect the nodes of the tree and their lengths represent the amount of change (usually substitutions per site) between an ancestor and its descendent [440]. The method used to infer the phylogenetic relationship between data is either distance-based or character-based [440]. Distance-based methods compute the pairwise distance between all the sequences included in an MSA. This results in a distance matrix which is then used to infer a tree [440].

Character-based methods compare all the sequences in an MSA, considering each column or site (character) in the alignment and estimate many trees from the data [438]. A score is then produced for each tree to determine the most probable or likely tree [440]. To be able to describe the MSA data and produce a reliable tree-score, the algorithm needs to hold some prior assumptions about the process of nucleotide, or amino acid substitution at different sites [438, 440]. These assumptions are provided by the use of a substitution model [440, 442]. Distance-based methods have the advantage of computational efficiency and, as such, are very useful for large data sets. They are most suited to data with low levels of sequence divergence [440]. Given the low sequence identity found across P450 sequences (see section 1.6.1.1.1), only character-based approaches were used to infer phylogenetic trees from the data in this PhD. Two methods were employed, both requiring the use of a substitution model: maximum likelihood (ML) and Bayesian inference (BI).

3.1.3.1 Maximum likelihood (ML)

ML calculates the log-likelihood value as the tree-score and measures the fit of the tree to the data (MSA) [440, 443]. In other words, it seeks a single tree that maximises the likelihood of observing the data given the tree and substitution model [440, 443]. In order to provide a measure of confidence at each node,

non-parametric bootstrapping is commonly used with ML tree estimation [443]. Bootstrapping generates a number of pseudo data sets by resampling sites at random from the MSA data. Sets of trees are then generated and compared, allowing for proportions of trees that include the same clades to be presented as a bootstrap support value [443]. Historically ML was considered computationally slow and costly, however recent advances, such as the PhyML algorithm [444] have solved many of these issues [440]. The PhyML algorithm [444] with non-parametric bootstrapping was used to infer ML phylogeny in this chapter.

3.1.3.2 Bayesian inference (BI)

BI [445] calculates the posterior probability for each node position as the tree-score [440, 443] and seeks the best set of trees from which to infer a consensus tree [438, 443]. In other words, it seeks the most probable tree given the data (MSA) and the substitution model. A major breakthrough in the algorithms used to infer phylogeny was the application of the Metropolis-Coupled Monte Carlo Markov Chain (MCMCMC or MC³) to BI [443]. With MC³ a set of independent searches (chains), normally four, are run simultaneously [443]. The first of these searches is referred to as the 'cold chain' and the others as 'hot chains' [438]. The chains occasionally swap roles and one of the hot chains is assigned as the cold chain. In this way the chains are able to exchange information, thereby increasing the chance of finding the optimal set of trees from which to infer a consensus tree [443]. BI as a methodology, is considered able to yield results that directly answer a biological question in terms that are easy to interpret [440].

3.1.4 Chapter aims and underpinning questions

A. mellifera and *B. terrestris* only have about half the number of P450s found in *D. melanogaster* (see section 1.7), but it is not clear whether this pattern is ubiquitous across all bees. To date there is a lack of data concerning the repertoire of P450 genes in solitary bee species. This chapter aims to curate the CYPome of *M. rotundata* and to audit the genes that fall into the four CYP clans found across insect species. This chapter will add to the understanding of how wide-spread the depauperate numbers of P450s are within bee species, specifically within a solitary, non-Apidae species.

As discussed in sections 1.4.4.1 and 1.7.3 members of the *CYP9Q* and *CYP9BU* lineages are involved in the differential sensitivity to insecticides from three MoA classes in *A. mellifera*, *B. terrestris* and *O. bicornis* [137, 155, 168, 226, 446]. This chapter will specifically focus on an audit of the CYP9 subfamily in *M. rotundata* to establish whether it contains candidate *CYP9Q/BU*-like, or closely related genes, that could be predicted to provide intrinsic tolerance to certain synthetic insecticides.

The following key questions will be addressed:

Firstly, is the presence of insecticide-degrading P450 enzymes ubiquitous across all bees? Secondly, can phylogeny be used as a reliable tool in the prediction of the function of P450 genes in the case of bees? That is, can a prediction of differential sensitivity to the different classes of neonicotinoid insecticides be made based on the presence of a gene that is closely related to the *CYP9Q/BU* lineage? Equally, could the lack of such a gene be used to predict an increased sensitivity to insecticides from three MoA classes?

The results from sections 3.3.2 and 3.3.6 of this chapter are published in Hayward *et al.*, 2019 *Nature Ecology and Evolution* 3(11):1521-1524 [447].

3.2 Methods

3.2.1 Curation of the *CYPome* of *M. rotundata*

Sequences encoding *M. rotundata* P450s were identified and assembled using three separate interrogations of the National Centre for Biotechnology Information (NCBI) non-redundant protein sequences database (nr) (<https://www.ncbi.nlm.nih.gov/>) (K. Singh, 2021 personal communication [448]). G1: text searches of existing annotated *M. rotundata* P450s; G2: OrthoFinder-v1.1.8 [226, 449] was used to define groups of orthologous genes (orthogroups) between the *M. rotundata* and *A. mellifera* proteomes and G3: an iterative Protein Basic Local Alignment Search Tool (BLASTp) (blastp:2.5.0+) [450] search of the *M. rotundata* proteome using annotated *A. mellifera* P450s as query sequences. The resulting 3 fasta files were amalgamated in Geneious version 10.2.3 (Biomatters) and all duplicates were removed. The remaining sequences were aligned using MUSCLE (version 3.5, default settings)[451]. All

sequences with pairwise distances of 50% or above were examined manually and the shortest isoforms of any duplicates were removed. The remaining sequences were used as the query in a BLASTp [450] search of *A. mellifera* in the NCBI protein database. All sequences that did not yield a P450 as a hit were removed. The resulting set of *M. rotundata* sequences were sent to David R. Nelson to be classified using the recognised P450 nomenclature [228, 229]. The final set of *M. rotundata* P450s was aligned with the outgroup P450cam, the camphor hydroxylase from *Pseudomonas putida* [211]; >gi|117297|sp|P00183.2| Cytochrome CYP-cam; CPXA_PSEPU) in Geneious version 10.2.3 (Biomatters) using MUSCLE (version 3.5, default settings) [451]. P450cam was used as the crystal structure has been determined [248]. This alignment was used to generate phylogenetic trees using Maximum likelihood algorithm [452, 453] and Bayesian inference [445] algorithms [Substitution model: LG+G [454]; Chain length: 1,100,000; Subsampling frequency: 200; Burn-in length: 100,000; Heated chains: 4; Heated chain temperature: 0.2].

3.2.2 Comparison of the CYPome of *M. rotundata* with other managed bee pollinators

P450 sequences for *A. mellifera*, *B. terrestris* and *O. bicornis* were obtained from the NCBI protein database. *M. rotundata* P450 sequences were aligned with those from three other managed bee pollinators and the outgroup P450cam in Geneious version 10.2.3 (Biomatters) using MUSCLE (version 3.5, default settings) [451]. MEGAX [455] was used to find the best-fit model of amino acid substitution using a Maximum Likelihood fit of 56 different models. Parameters including substitution model, proportion of invariable sites and rate variation were calculated (see appendix table 3.1). The substitution model with the lowest Bayesian Information Criterion (BIC) score was selected for use in phylogeny estimation. The alignment was used to generate phylogenetic trees using Maximum likelihood and Bayesian inference [445] algorithms as described in the previous section.

3.2.3 Conserved synteny analysis

The scaffolds containing CYP9 genes from each managed pollinator species were downloaded from the NCBI nucleotide database and imported into Geneious version 10.2.3 (Biomatters). *A. mellifera* (DH4 linkage group LG14,

Amel_HAv3.1 WGS) was used as the reference and the regions upstream and downstream of all CYP9 genes were examined and all flanking genes noted. Genes in a ~500Kbp region flanking the CYP9 genes in *A. mellifera* were used to identify potential orthologous regions in the genomes of the other bee species. This section of *A. mellifera* linkage group LG14 was used as a query sequence in a BLASTn search through the genomes of other managed bee pollinators to look for evidence of macro-synteny and micro-synteny. For a region to be considered as showing micro-synteny the minimum requirement was the conservation of two neighbouring homologs with no more than five unrelated genes in the intervening DNA. Genes identified in a micro-syntenic block in all four species were aligned using MUSCLE (version 3.5, default settings) [451] to discover their percentage identity and enable an assessment of their homology. To better understand the general level of synteny found between bee genomes, the region ~500Kbp upstream and downstream of a second CYP cluster (the CYP6AS lineage) was examined. In *A. mellifera* the CYP6AS cluster contains 15 CYP6 genes and is located on the DH4 linkage group LG13. This region was used as described above to determine syntenic conservation in the other managed bee pollinator species.

To investigate whether micro-synteny of the CYP9 cluster is conserved between S-T and L-T bees (see figure 3.1 (a)), scaffolds containing the CYP9 cluster from two further bee genomes were added to the analysis: the Halictidae, *Dufourea novaeangliae* and the Colletidae, *Colletes gigas*. BLASTn searches using query sequences from *A. mellifera* (*CYP9P1*, *CYP9R1* and *CYP9Q3*) and *M. rotundata* (*CYP9DM1* and *CYP9DN1*) were used to locate the CYP9 subfamily, and *D. novaeangliae* gene sequences to locate the CYP9 flanking genes in the unannotated *C. gigas* genome.

3.2.4 Investigating the structural homology of the CYP9 subfamily of P450s

To select the most appropriate crystal structure *A. mellifera* *CYP9Q3* was used as a query sequence in a BLASTp [450] search of the Protein Data Bank (RCSB: PDB) gated for 'eukaryotes (taxid:2759)'. The top hit crystal structure (*Homo sapiens* *CYP3A4* PDB: 4D6Z [254]) was used in all structural homology (e-value: 2e-66). To understand the secondary and tertiary structure of *M.*

rotundata CYP9s, the protein sequences were aligned with CYP9 sequences from *A. mellifera*, *B. terrestris* and *O. bicornis*; and chain A of PDB: 4D6Z in Geneious version 10.2.3 (Biomatters) using MUSCLE [451] (version 3.5, default settings). With reference to the resolved structure of *P. putida* P450cam [248], the positions of the 5 conserved motifs and 6 SRSs were annotated on the alignment in Geneious. This was then used to manually overlay the positions onto the crystal structure, PDB: 4D6Z, using UCSF Chimera version 1.10.1. Well-defined and manually checked MSAs have been used effectively in this manner to model P450s [251].

3.2.5 Comparison of the structural homology of CYP9Q, CYP9BU and CYP9DM proteins

To identify candidate genes in *M. rotundata* with homology to the CYP9Q/BU lineage both phylogeny and syntenic analysis were used. The protein sequences of the candidate homolog genes were aligned with the CYP9Qs, CYP9BUs and chain A of PDB: 4D6Z [254] in Geneious version 10.2.3 (Biomatters) using MUSCLE [451] (version 3.5, default settings). With reference to the resolved structure of *P. putida* P450cam [248] PDB: 4D6Z and PDB: 1TQN [250], the secondary structure (α -helices and β -sheets) and the positions of the 5 conserved motifs and 6 SRSs were annotated on the alignment in Geneious. This was then used to manually overlay the positions onto the crystal structure, PDB: 4D6Z, using UCSF Chimera version 1.10.1.

3.2.6 Distribution of the CYP9 subfamily across 12 available bee genomes

All other available genomes of bee species were searched and sequences encoding CYP9 subfamily P450 enzymes were identified and assembled as described in section 3.2.1. (8 bee species: *Apis cerana*, *Apis dorsata*, *Apis florea*, *Bombus impatiens*, *D. novaeangliae*, *Eufriesea mexicana*, *Habropoda laboriosa* and *Melipona quadrifasciata*; data accessed from NCBI protein database 2017 & 2018). The resulting sequences were aligned to those of the managed bee pollinator species and the outgroup P450cam in Geneious version 10.2.3 (Biomatters) using MUSCLE [451] (version 3.5, default settings). Parameters including proportion of invariable sites and gamma rate were optimised and the Le & Gascuel amino acid substitution matrix [454] with gamma rates (LG+G) was selected based on the Bayesian information criterion

(see appendix table 3.2). Phylogeny was estimated using a Maximum likelihood algorithm in MEGAX [455] that estimated relative time rather than substitutions per site [456].

3.3 Results

3.3.1 Curation of the CYPome of *M. rotundata*

Forty-nine full length CYPs were identified from the interrogation of the database and classified using the recognised nomenclature. A full list of CYP names, accession numbers and size are provided in table 3.1, and the phylogenetic relation between the genes in the CYPome are shown in figure 3.3. One putative pseudogene, *CYP6BE1P*, appears in the list, shown highlighted in blue in figure 3.3. The majority (over 65%) of *M. rotundata* CYPs are in the CYP3 clan and overall, the breakdown of the CYP clans is remarkably similar to that of *A. mellifera* (see figure 3.4). There is clear evidence of a CYP bloom in the subfamily of *CYP6AS* genes which has 13 members; the ancestral node is highlighted in yellow in figure 3.3. The 9 CYP9 subfamily genes form a monophyletic clade within the CYP3 clan, the ancestral node is highlighted in red (see figure 3.3).

Both *A. mellifera* and *M. rotundata* show a depauperate CYP4 clan when compared to *D. melanogaster* and *A. gambiae* (see figure 3.4). The CYP4 clan is highly diverse in many other insect species and is known to be involved in xenobiotic metabolism and chemical communication, as well as being involved in endogenous functions such as ecdysteroid synthesis and lipid metabolism [211]. CYP4 member genes have also been implicated in insecticide metabolism and resistance, such as resistance to pyrethroids in *Blattella germanica* [457] and DDT in *Anopheles* mosquitos [241].

Table 3.1: Description of the CYPome of *M. rotundata*

P450 Clan	Name	Protein accession number	Nucleotide accession number	Protein length
CYP2	CYP15A1	XP_012138027.1	XM_012282637.1	503
	CYP18A1	XP_012142657.1	XM_012287267.1	538
	CYP303A1	XP_012153072.1	XM_012297682.1	564
	CYP305D1	XP_003702175.1	XM_003702127.2	485
	CYP307B1	XP_012139366.1	XM_012283976.1	507
	CYP343A1	XP_012138467.1	XM_012283077.1	515
	CYP369A1	XP_003703633.1	XM_003703585.2	505
CYP3	CYP6AQ52	XP_012136078.1	XM_012280688.1	515
	CYP6AQ53	XP_012136079.1	XM_012280689.1	515
	CYP6AQ54	XP_003701310.1	XM_003701262.2	515
	CYP6AS108	XP_012146067.1	XM_012290677.1	370
	CYP6AS109	XP_012150537.1	XM_012295147.1	502
	CYP6AS110	XP_012136792.1	XM_012281402.1	499
	CYP6AS111	XP_012143066.1	XM_012287676.1	501
	CYP6AS112	XP_003704416.1	XM_003704368.2	510
	CYP6AS113	XP_012143067.1	XM_012287677.1	501
	CYP6AS114	XP_012143073.1	XM_012287683.1	499
	CYP6AS115	XP_012143072.1	XM_012287682.1	501
	CYP6AS116	XP_003704412.1	XM_003704364.2	512
	CYP6AS117	XP_012143062.1	XM_012287672.1	502
	CYP6AS118	XP_003704278.1	XM_003704230.2	499
	CYP6AS119	XP_012140372.1	XM_012284982.1	501
	CYP6AS120	XP_003704414.1	XM_003704366.2	499
	CYP6BC1	XP_003703200.1	XM_003703152.2	517
	CYP6BD1	XP_003699913.1	XM_003699865.2	503
	CYP6BE1	XP_012153915.1	XM_012298525.1	511
	CYP6BE1P	XP_012153916.1	XM_012298526.1	440
	CYP9DN1	XP_003703411.1	XM_003703363.2	522
	CYP9P2	XP_012145771.1	XM_012290381.1	515
	CYP9P22	XP_012145774.1	XM_012290384.1	506
	CYP9P23	XP_012145773.1	XM_012290383.1	512
	CYP9R1	XP_003705491.1	XM_003705443.2	516
	CYP9R58	XP_003705489.1	XM_003705441.2	516
	CYP9R59	XP_012145777.1	XM_012290387.1	518
	CYP9DM1	XP_003705488.1	XM_003705440.2	498
	CYP9DM2	XP_003705490.2	XM_003705442.2	531
	CYP336A33	XP_003702293.1	XM_003702245.2	501
	CYP336A34	XP_003702219.1	XM_003702171.2	501
	CYP336M1	XP_012138246.1	XM_012282856.1	412
	CYP4	CYP4G11	XP_012145465.1	XM_012290075.1
CYP4G202		XP_003700755.1	XM_003700707.2	561
CYP4AA1		XP_012137854.1	XM_012282464.1	515
CYP4AV1		XP_003699552.2	XM_003699504.2	514
MITOCHONDRIAL	CYP301A1	XP_003703366.2	XM_003703318.2	530
	CYP301B1	XP_003703365.1	XM_003703317.2	492
	CYP302A1	XP_012135500.1	XM_012280110.1	519
	CYP314A1	XP_012150013.1	XM_012294623.1	525
	CYP315A1	XP_003704053.1	XM_003704005.2	533
CYP334A1	XP_003700083.1	XM_003700035.2	582	

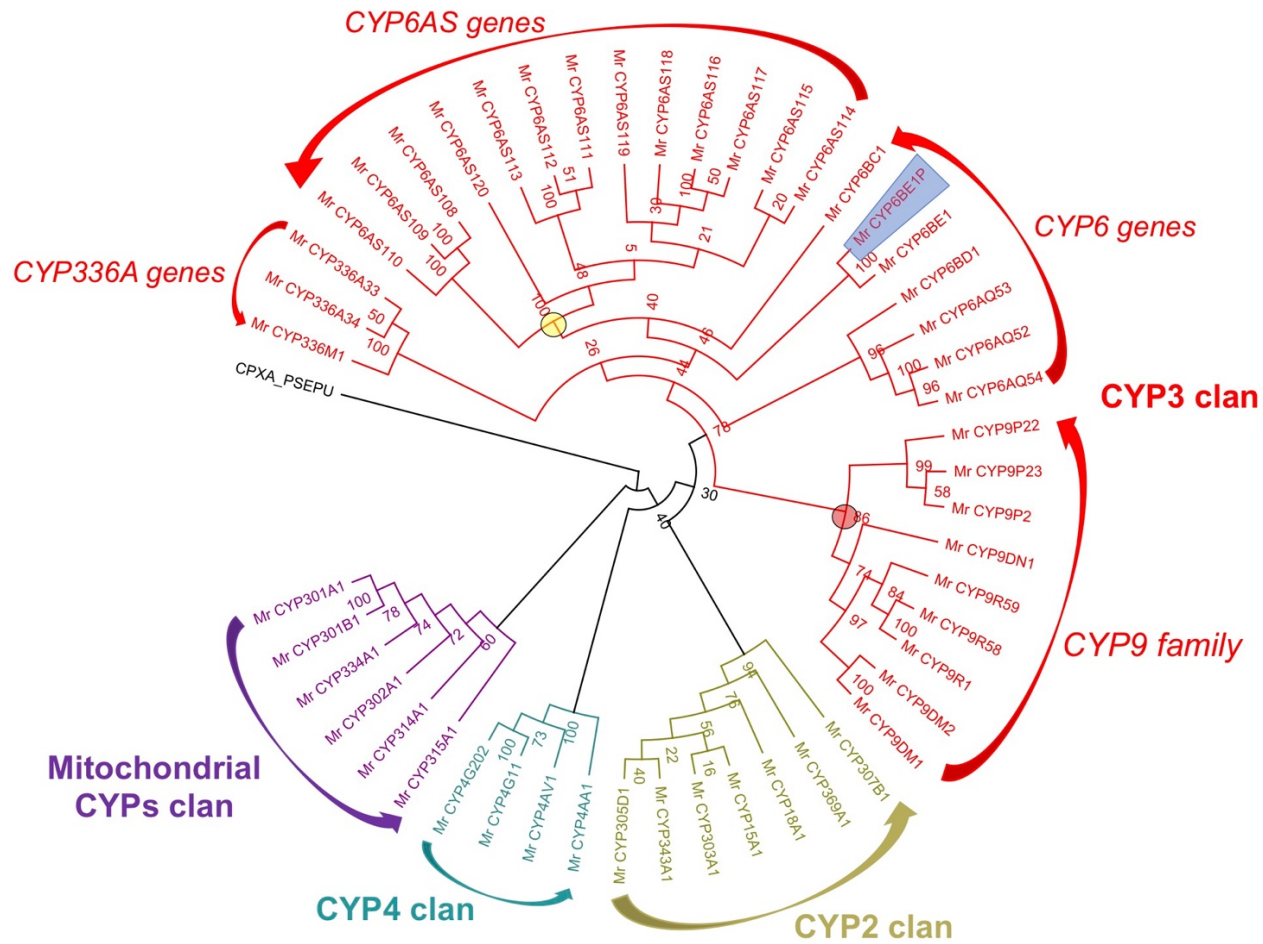


Figure 3.3: Phylogeny of *M. rotundata* CYPome, rooted on camphor hydroxylase (P450cam; *P. putida*). Phylogeny estimated using PhyML Maximum likelihood algorithm [453], with branch support of 200 bootstraps shown as %. Blue shading denotes a pseudogene; ancestral node of the CYP bloom in the *CYP6AS* subfamily marked with a yellow circle and the *CYP9* subfamily with a red circle.

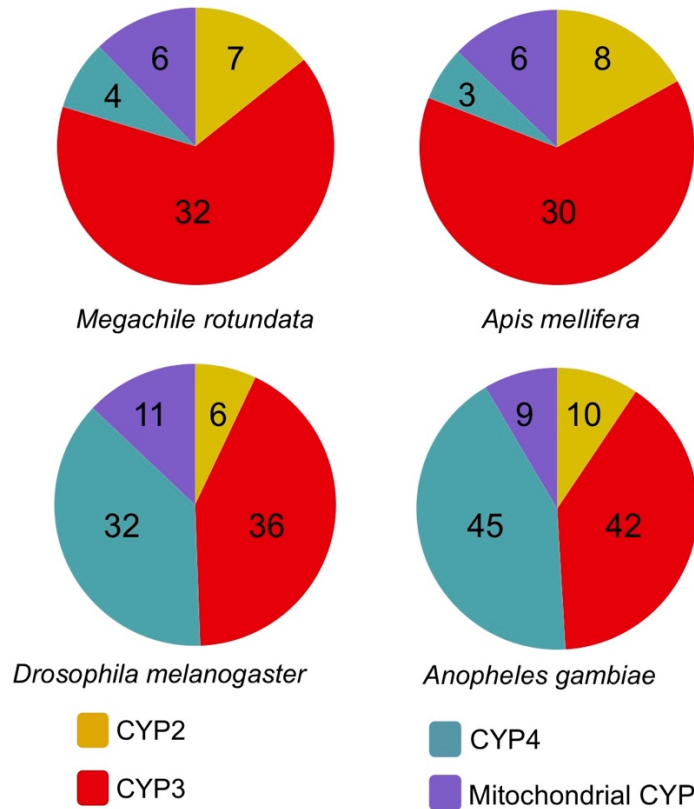


Figure 3.4: Distribution of the CYP clans in *M. rotundata* compared with three insect species (*A. mellifera* [211]; *D. melanogaster* [222] & *A. gambiae* [212])

3.3.2 Comparison of the CYPome of *M. rotundata* with other managed pollinators

The overall number of functional P450s in the CYPomes of the four species is very similar (*A. mellifera*: 46; *B. terrestris*: 49; *O. bicornis*: 52 and *M. rotundata*: 48; see appendix table 3.3). As can be seen from the heat map of the Bayesian inference distance matrix (see figure 3.5), the sequences form distinct groups, which correspond to the four clans of CYP genes found across all insects. The boundaries of the CYP clans are distinct, as are the sub-divisions by gene family and sub-family found within the CYP3 clan (see figure 3.5). The members of the mitochondrial CYP clan and those of the CYP2 and CYP4 clans are almost identical across all four managed bee pollinators (see figure 3.5 and appendix table 3.3). The mitochondrial clan is the only one that has 1:1 orthology across all members (see appendix table 3.3). The major differences between the CYPomes are found in the CYP3 clan. In the

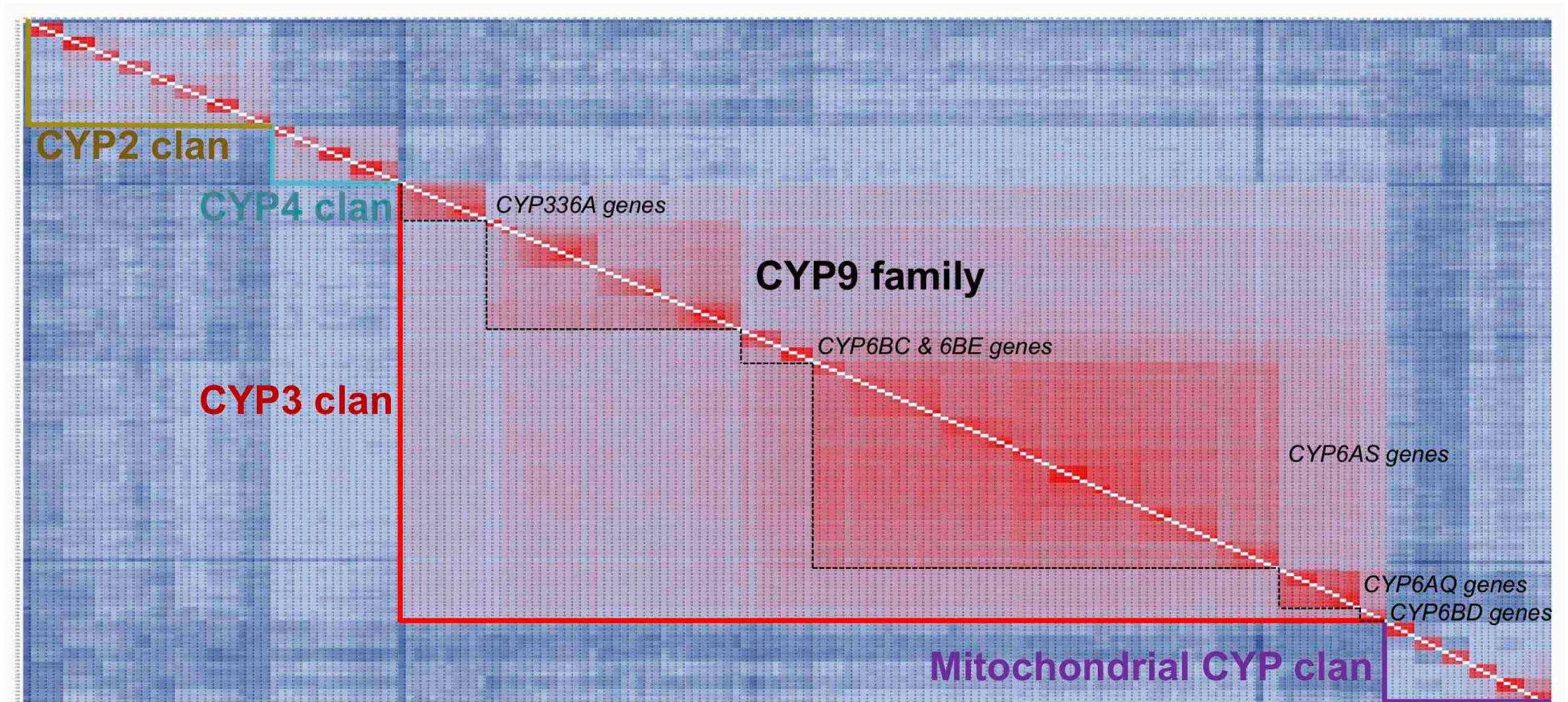


Figure 3.5: Heat map of the distance matrix (percentage identity) of the CYPomes across four species of managed bee pollinators (*A. mellifera*; *B. terrestris*; *O. bicornis* and *M. rotundata*). Matrix generated using Bayesian inference estimation [Chain length: 1,100,000; Subsampling frequency: 200; Burn-in length: 100,000; Heated chains: 4; Heated chain temperature: 0.2]

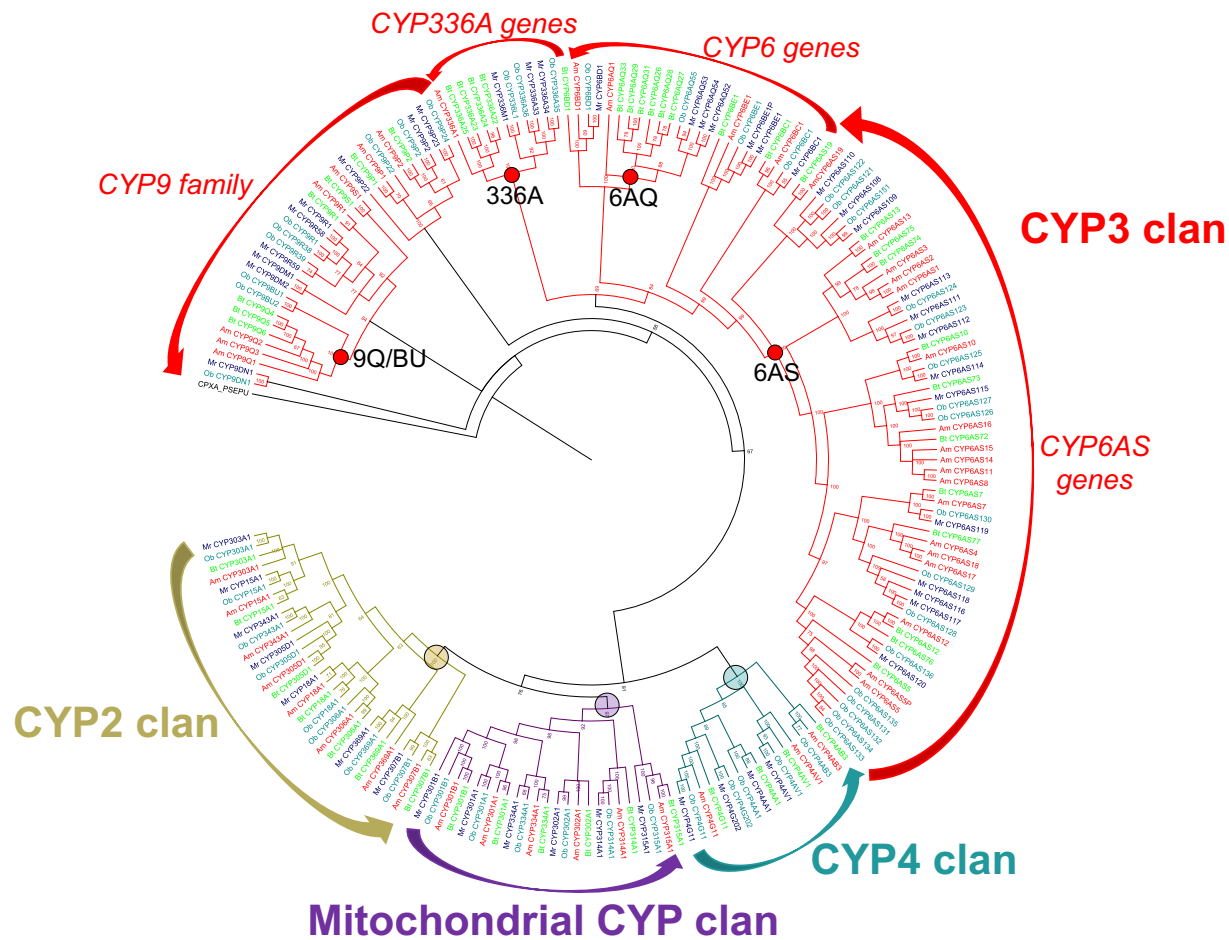


Figure 3.6: Bayesian inference phylogeny of the CYPomes of four managed bee pollinators (*A. mellifera*; *B. terrestris*; *O. bicornis* and *M. rotundata*), using substitution model LG+G [454]. Sequence names coloured by species: *A. mellifera* - red; *B. terrestris* - green; *O. bicornis* – light blue and *M. rotundata* – dark blue. Posterior probability of nodes shown as a % probability. Tree rooted on camphor hydroxylase (P450cam; *P. putida*). All protein sequences accessed from NCBI protein database.

phylogeny of the CYPomes (figure 3.6) there is clear 'clumping' of genes in the CYP3 clan due to the differential expansion of P450 lineages in the four species.

3.3.2.1 The CYP3 clan

CYP6 family

There is clear evidence of an expansion (or bloom) within the subfamily of CYP6AS genes in all four species (see figure 3.6 and appendix table 3.3). *A. mellifera* has 17 CYP6AS genes; *O. bicornis*: 17; *M. rotundata*: 13 and *B. terrestris*: 12. This gene bloom accounts for ~59%, 38%, 52% and 42% of the entire CYP3 clan respectively and, given the position of the ancestral node, (marked in red in fig 3.6), it is clear that the origin of the expansion of this gene subfamily predates the divergence of the species. *B. terrestris* and *M. rotundata* show an expanded repertoire of CYP6AQ genes (6 and 3 members respectively) whereas *A. mellifera* and *O. bicornis* have only a single CYP6AQ (see appendix table 3.3). However, the duplication of CYP6AQ genes appears to have occurred after the divergence of the species (see node marked in red in figure 3.6). The remaining CYP6 genes all have 1:1 orthologs (see figure 3.6 and appendix table 3.3).

CYP336A family

A. mellifera has one CYP336A gene, whereas the other three species show an expanded CYP336A subfamily (*B. terrestris*: 4 members; *M. rotundata* and *O. bicornis*: 3 members) (see figure 3.6 and appendix table 3.3).

CYP9 subfamily

The Apidae species have 7 CYP9 genes, whereas the Megachilidae have 9 (see figure 3.7). Both Apidae species have one CYP9S1 gene and three members of the CYP9Q lineage, all of which are absent in *M. rotundata* and *O. bicornis*. However, *O. bicornis* has two CYP9BU genes which share a common ancestor with the CYP9Q lineage (see figure 3.6). Both Megachilidae species show an expanded repertoire of CYP9R lineage genes (3 members each) (see figure 3.7 and appendix table 3.3). *M. rotundata* is the only species to have CYP9DM genes.

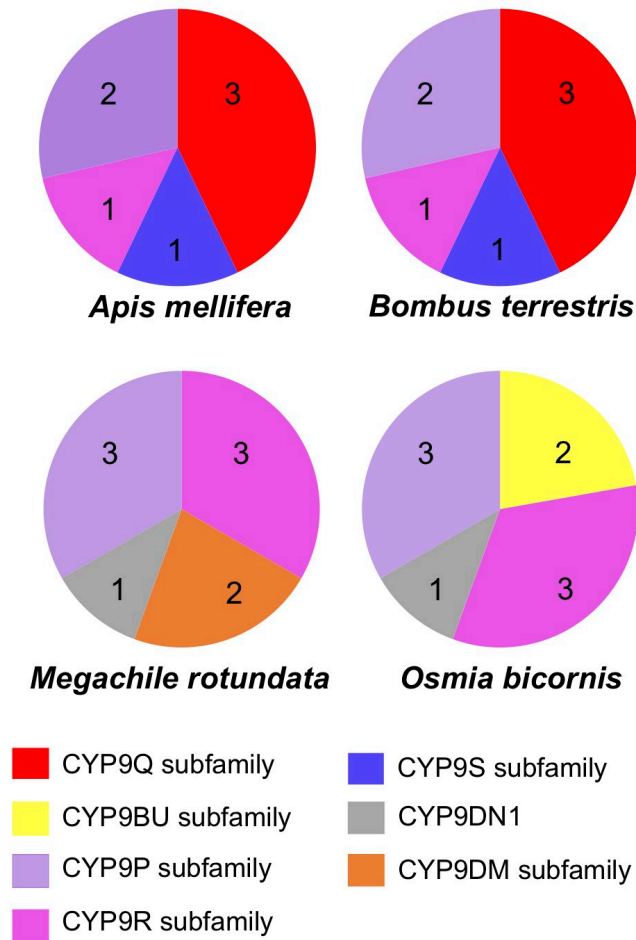


Figure 3.7: Distribution of CYP9 subfamily across four species of managed bee pollinators.

CYP9DN1 is only found in the Megachilidae species; no homolog was found in the Apidae bees (see figure 3.7). Unlike the other CYP9s which are single-exon (intronless) genes, *CYP9DN1* contains one intron (see figure 3.8).

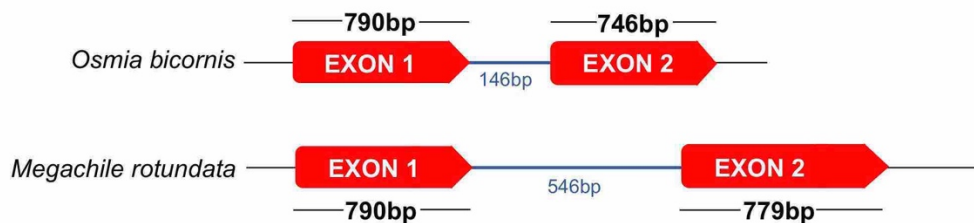


Figure 3.8: Structure of the transcribed region of *CYP9DN1* in two Megachilidae species. Exons are shown in red and introns in blue.

3.3.3 Conserved synteny analysis

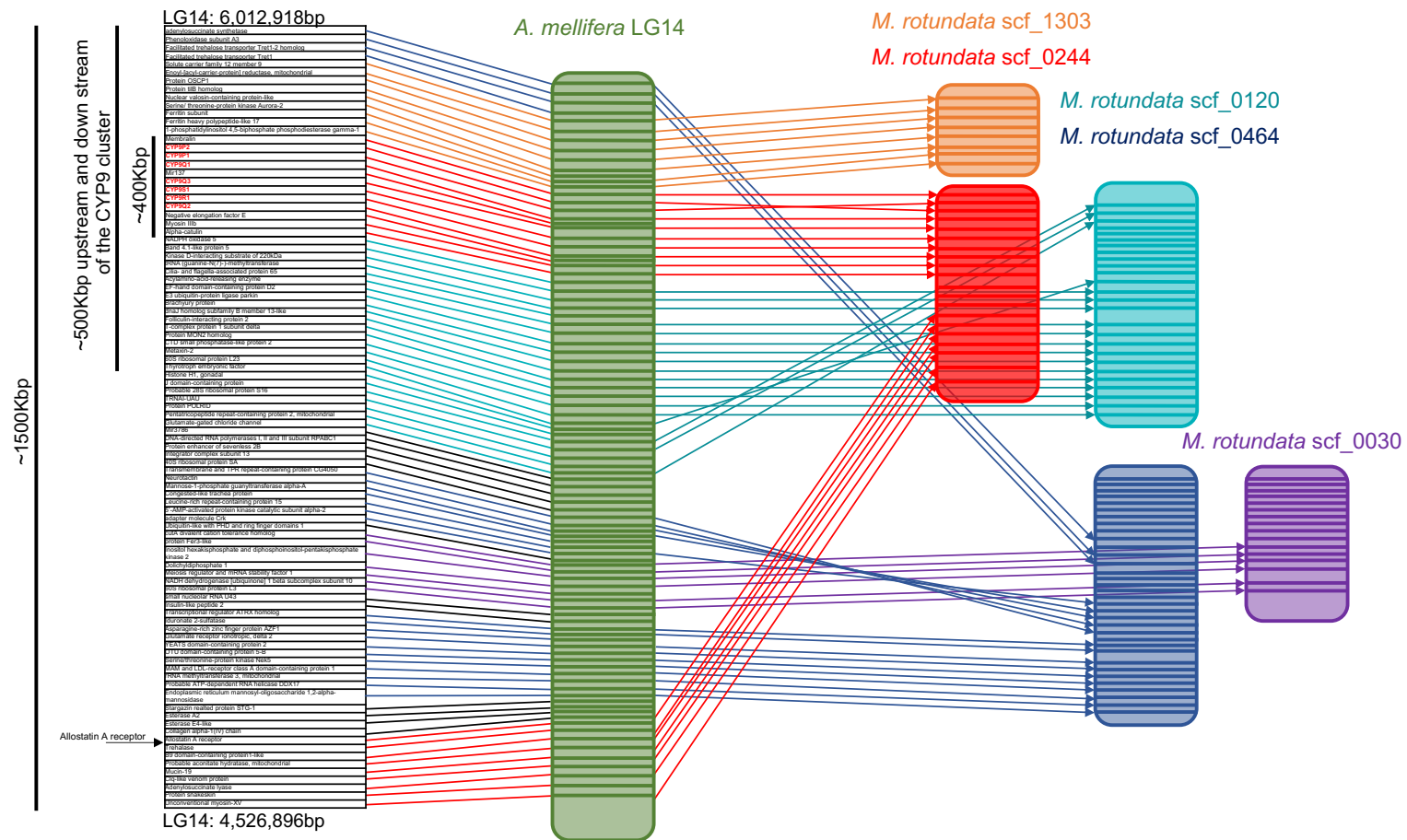
3.3.3.1 Macro-synteny of the CYP9 subfamily

In *A. mellifera* all members of the CYP9 subfamily appear in a single cluster on one scaffold (DH4 linkage group LG14, Amel_HAv3.1) (see figure 3.9 and appendix table 3.4). In the other species, one CYP9 gene appears on a separate scaffold. *CYP9S1* in *B. terrestris* is annotated on a short scaffold (Bter_1.0 GroupUn997; 5678bp), rather than in the CYP9 cluster. *CYP9DN1*, which is only found in the Megachilidae species, is also located on a separate scaffold (Obicornis)_v3 Scaffold 00191 and MROT_1.0_scf_0129).

There appears to be good macro-synteny of the main CYP9 cluster between *A. mellifera* and *B. terrestris*, although the genes in the latter are located across two scaffolds (LG B01, Bter_1.0 WGS and LG B14, Bter_1.0 WGS) (see appendix table 3.4). There is good conservation of genomic content, with 68.4% (26/38) genes present in the region 500Kbp upstream and downstream of the CYP9 cluster in *A. mellifera*, being found on the two *Bombus* scaffolds (see table 3.2 and appendix table 3.4). The two Megachilidae species also show evidence of regions of good macro-synteny. 68.4% (26/38) and 55.3% (21/38) of the genes present in the CYP9 cluster region in *A. mellifera* are found in *M. rotundata* and *O. bicornis* respectively (see table 3.2 and appendix table 3.4). However, the same region of *A. mellifera* LG14 is represented across four separate scaffolds in *M. rotundata* (MROT_1.0: scf_0464; scf_1303; scf_0244 and scf_0120) and three in *O. bicornis* (Obicornis_v3: scf00020; scf00060 and scf00090; see figure 3.9 and appendix table 3.4). There is clear evidence of at least one incidence of synteny breakage and genomic rearrangement common to both Megachilidae species. This synteny break results in *membralin* becoming flanked, after the insertion of two genes, by *allostatin A receptor* (see figure 3.9 and appendix table 3.4).

3.3.3.2 Macro-synteny of the CYP6AS cluster

The region 500Kbp upstream and downstream of the *CYP6AS* cluster in *A. mellifera* (DH4 linkage group LG13, Amel_HAv3.1), used to identify potential orthologous regions in the other species, is mapped to a single scaffold in *B. terrestris* (LG B13, Bter_1.0 WGS) and two scaffolds in *M. rotundata*



(MROT_1.0_scf_0147, MROT_1.0_scf_0128) (see table 3.2 and appendix table 3.5). The *CYP6AS* genes in *O. bicornis* are located across 20 scaffolds (see appendix table 3.5). Of these, 11 contain only one *CYP6AS* sequence and are referred to as orphans in appendix table 3.5. Only two *O. bicornis* scaffolds (Scaffold00161 and Scaffold00374) have any flanking genes in common with the regions in the other bee genomes (see appendix table 3.5).

Table 3.2: Macro-synteny of the CYP9 and CYP6AS clusters between the managed bee pollinators. [CYP6AS data for *O. bicornis* not included as <10% conserved pairs found across 12 scaffolds]

Species	CYP cluster [genomic position]	Conserved pairs	Total number non-CYP genes	Proportion conserved (%)
<i>B. terrestris</i>	CYP6AS [LG B13]	22	26	84.6
	CYP9 [LG B01, LG B14]	26	38	68.4
<i>M. rotundata</i>	CYP6AS [MROT_1.0_scf_0147, MROT_1.0_scf_0128]	16	26	61.5
	CYP9 [MROT_1.0_scf_1303, MROT_1.0_scf_0244, MROT_1.0_scf_0120]	26	38	68.4
<i>O. bicornis</i>	CYP9 [Obicornis_v3_scf00020, Obicornis_v3_scf00060, Obicornis_v3_scf00090]	21	38	55.3

3.3.3.3 Macro-synteny of the *CYP9DN1*-containing scaffold

The *CYP9DN1*-containing scaffold in *O. bicornis* (Scaffold 00191) is only 229,147 bp long and *CYP9DN1* is located 24,955-22,927 (complement) from the beginning of the scaffold. The lack of sequence further than 22,927 bp downstream of *CYP9DN1* in *O. bicornis*, and the fact that this gene only appears in the two Megachilidae species, makes any analysis of synteny unreliable. Although there is evidence of micro-synteny with the conservation of four genes in the genetic neighbourhood of *CYP9DN1* in the Megachilidae (see table 3.3). Due to the size of the scaffold in *O. bicornis*, the *CYP9DN1*-containing scaffold in the genome of *D. novaeangliae* (Halictidae) was also examined for evidence of conserved synteny. 57.14% of the immediate neighbouring genes found upstream and downstream in *M. rotundata* were also

present in *D. novaeangliae* (see Table 3.3), potentially implying a good level of syntenic conservation across bee families.

Table 3.3: Genetic neighbourhood of *CYP9DN1* in two Megachilidae and one Halictidae bee species. Flanking genes found across all three species shaded dark grey and those found in two species light grey.

<i>M. rotundata</i>	<i>O. bicornis</i>	<i>D. novaeangliae</i>
<i>MOB kinase activator-like 2</i> – LOC100876017	N/A	<i>Palmitoyltransferase Hip 14</i> – LOC108571567
<i>Palmitoyltransferase ZDHHC17</i> – LOC100875905	N/A	<i>MOB kinase activator-like 2</i> – LOC108571518
<i>Dynein heavy chain 1, axonemal</i> – LOC100874893	<i>Dynein heavy chain 1, axonemal</i> – LOC114879560 & LOC114879567	<i>Dynein heavy chain 1, axonemal</i> – LOC108571486
<i>CYP9DN1</i> – LOC100875680	<i>CYP9DN1</i> – LOC114879550	<i>CYP9DN1</i> – LOC108571625
<i>Alpha-2C adrenergic receptor-like</i> – LOC105662486	<i>Thyrotropin-releasing hormone receptor-like</i> – LOC114879552	<i>Serine tRNA ligase-like</i> LOC108571561
<i>Cytochrome b-c1 complex subunit 8</i> – LOC105662487	<i>Cytochrome b-c1 complex subunit 8</i> – LOC114879554	<i>AT-rich interactive domain-containing protein 2</i> – LOC108571634
	<i>Serine-tRNA synthetase-like protein Slimp</i> – LOC114879551	
<i>Zinc finger protein 593 homolog</i> – LOC100876834	<i>Zinc finger protein 593 homolog</i> – LOC114879553	
<i>AT-rich interactive domain-containing protein 2</i> – LOC100876726	<i>AT-rich interactive domain-containing protein 2</i> – LOC114879556	

3.3.3.4 Micro-synteny of the CYP9 cluster

To investigate the evolutionary relationship of the CYP9 cluster across four bee families, a phylogeny of the protein sequences was estimated using a maximum likelihood algorithm (see figure 3.10(a)). To allow for branch support data to be shown clearly, not all branch length data for the CYP9 phylogeny is displayed. However, the range of branch lengths is: 0.02 -1.1. The longest branch length in the phylogeny leads to the *M. rotundata*-specific *CYP9DM* lineage and is estimated at 1.1 substitutions per site (see figure 3.10(a)). In comparison the branch lengths leading to the *CYP9Q* lineage and the *O. bicornis*-specific *CYP9BU* lineage are 0.16 and 0.17 substitutions per site respectively (see figure 3.10(a)).

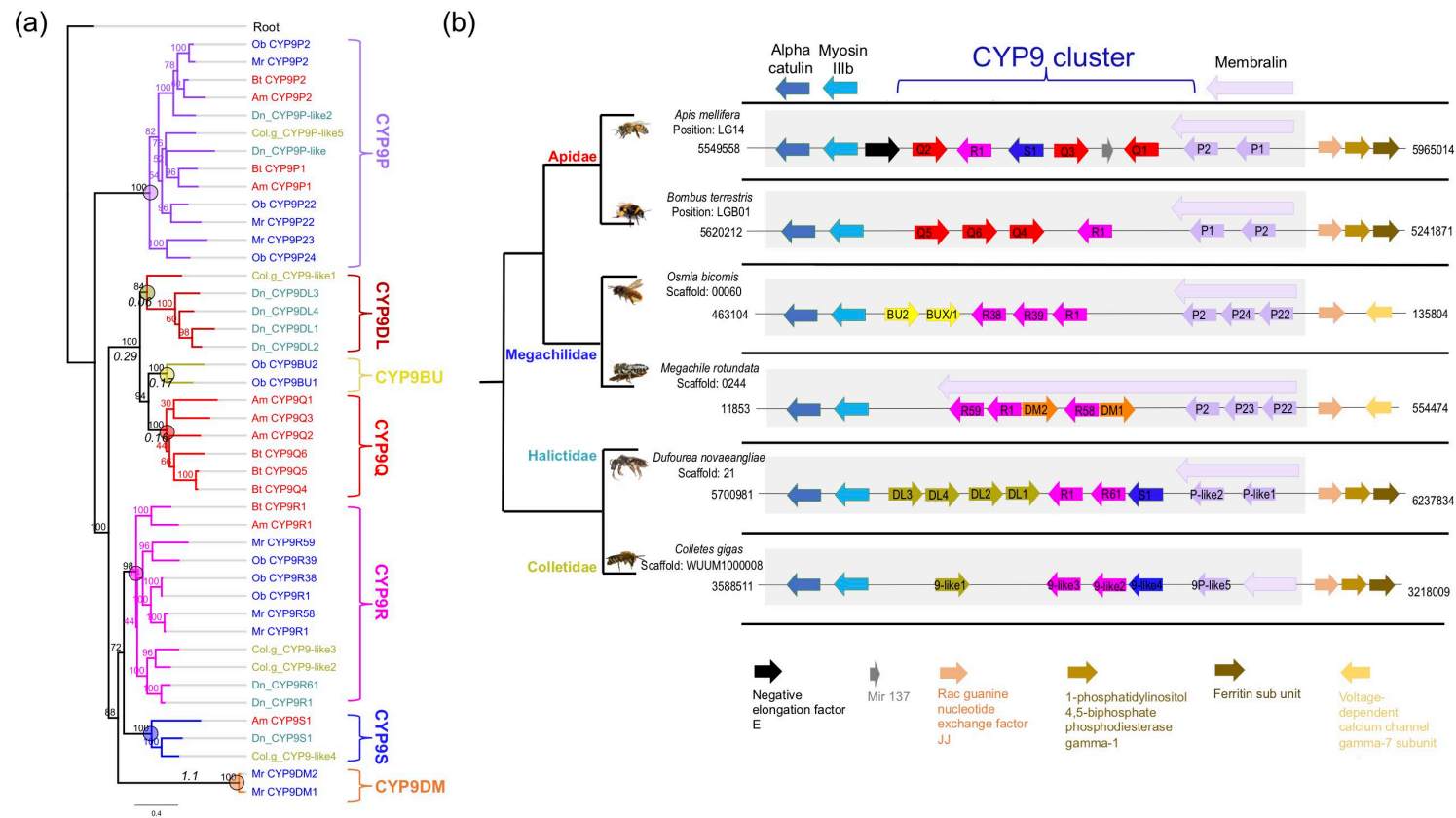


Figure 3.10: (a) Phylogeny of CYP9 amino acid sequences from six bee species (*A. mellifera*; *B. terrestris*; *O. bicornis*, *M. rotundata*, *D. novaeangliae* and *C. gigas*). Phylogeny estimated using PhyML Maximum likelihood algorithm [453] and substitution model LG+G [454], with branch support of 50 bootstraps, shown as %, rooted on *Nasonia vitripennis* CYP9AG4. Scale bar represents 40 substitutions per 100 residues. Branch lengths, where shown, are in italics. Each monophyletic CYP9 lineage is denoted by colour and the ancestral node is marked with a circle. Sequence name coloured by family. (b) Syntenic relationship at the CYP9 loci in six bee species across four families (schematic representation only, not to scale). CYP9 genes are coloured by lineage. Arrows denote reading frame.

The region immediately flanking the CYP9 cluster in *A. mellifera*, spanning *ferritin sub unit* to *alpha catulin*, was examined for micro-synteny (see figure 3.10(b)). The shortest of the primary CYP9 cluster-containing scaffolds isolated was in *O. bicornis* (Scaffold000060: 850,338bp) and the longest in *B. terrestris* (LG B01: 17,153,651bp). Members of the *CYP9P* lineage are consistently associated with *membralin* and are found in the upstream portion of the CYP9 cluster (see figure 3.10(b)). *Myosin IIIb* and *alpha catulin* are found downstream of members of the *CYP9R* lineage in all four species (see figure 3.10(b)). The degree of similarity of these three flanking-genes across six bee species is shown in table 3.4. The degree of identity increases if Megachilidae species are compared solely within family (*membralin*: 80.397%; *myosin IIIb*: 87.888% and *alpha catulin*: 98.343%).

Table 3.4: Percent identity with *A. mellifera* protein sequences for the primary flanking genes from the syntenic block that includes the CYP9 cluster.

	% identity to <i>A. mellifera</i> sequence				
	<i>Membralin</i>	<i>Myosin IIIb</i>	<i>Alpha catulin</i>	<i>CYP9P2</i>	<i>CYP9R1</i>
<i>B. terrestris</i>	85.331	85.283	97.176	77.543	66.792
<i>O. bicornis</i>	72.727	81.813	92.000	71.098	60.038
<i>M. rotundata</i>	73.217	78.765	92.588	68.208	60.342
<i>D. novaeangliae</i>	70.508	78.045	76.793	67.823	60.755
<i>C. gigas</i>	66.762	79.300	92.471	58.949	60.264

The first gene located upstream from *membralin* is annotated in the NCBI database as: *uncharacterised protein* LOC551741 in *A. mellifera*; *putative protein tag-52* in *B. terrestris* and *FYVE, RhoGEF and PH domain containing protein 4* in *O. bicornis*. However, when used as query sequences in a BLASTn search of the *M. rotundata* genome, all three of these sequences return reliable hits for *rac guanine nucleotide exchange factor JJ* (E value *A. mellifera*: $1e^{-111}$; *B. terrestris*: 0.0 and *O. bicornis*: 0.0). The synteny depicted in figure 3.10(b) therefore shows *rac guanine nucleotide exchange factor JJ* in this position across these bee species. The two genes found further upstream from *rac guanine nucleotide exchange factor JJ* show good synteny in the Apidae, Halictidae and Colletidae species, with those species showing 1-*phosphatidylinositol 4,5-biphosphate phosphodiesterase gamma 1* and *ferritin sub unit* in identical order and direction of reading frame (see figure 3.10(b)). Both of these genes appear on different scaffolds in the Megachilidae species

(see appendix table 3.4). The Megachilidae species show the insertion of *voltage-dependent calcium channel gamma-7 subunit* and one other gene, after *rac guanine nucleotide exchange factor JJ*, at the point of the synteny breakage and genomic rearrangement, before the sequence joins to the *allostatin A receptor* gene (see figure 3.10(b) and appendix table 3.4).

3.3.4 The structural homology of the CYP9 subfamily

The MSA of CYP9 sequences shown in figure 3.11(c) was used to overlay the crystal structure of *H. sapiens CYP4A3* (PDB: 4D6Z [254]) using UCSF Chimera version 1.10.1. The protein was then rendered by conservation and coloured using a magenta-cyan gradient (conserved-variable) (see figure 3.11(b)). The central part of the distal face contains most of the conserved motifs and is the site of the heme-binding motif (M5; conserved cysteine marked in yellow) (see figure 3.11(a)). When overlaid with the MSA (see figure 3.11(b)), the distal face of the molecule shows the CYP9 sequences to be extremely well conserved across the species. In figure 3.11(b) the heme-binding motif is surrounded almost exclusively by highly conserved (magenta) residues. In contrast, the proximal face, that houses the main SRSs (see figure 3.11(b)) has more residues that are in the white-cyan range of the colour-gradient, implying that these parts of the primary sequences are much less well conserved across the CYP9 subfamily.

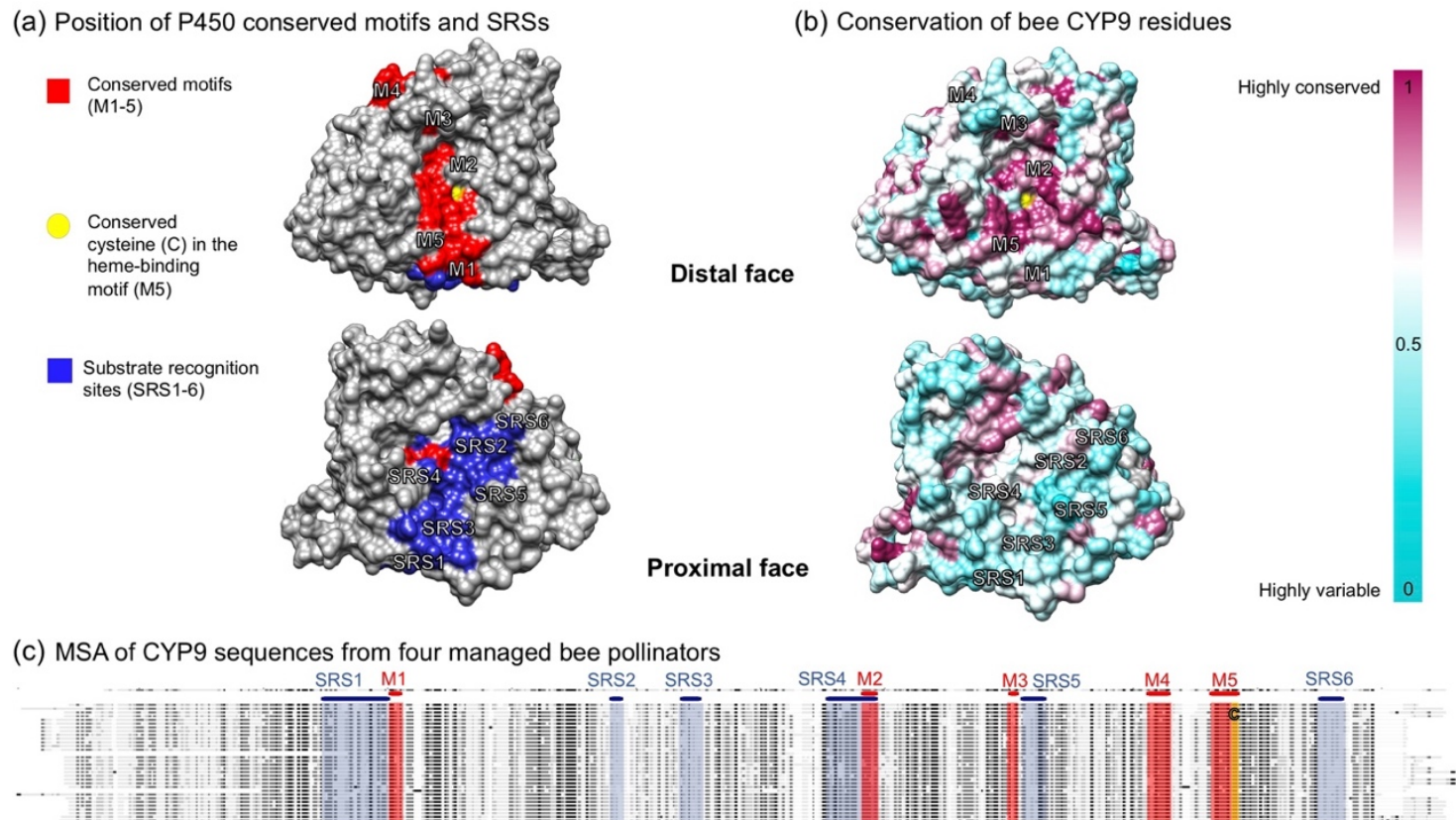


Figure 3.11: Tertiary structure of *H. sapiens* CYP4A3 (PDB: 4D6Z [254]), showing conserved protein sequence in the CYP9 subfamily in four managed bee pollinators (*A. mellifera*; *B. terrestris*; *O. bicornis* and *M. rotundata*). Distal and proximal faces of the molecule are included in the figures. (a) Conserved motifs (M) are coloured red and substrate recognition sites (SRSs) blue. The crystal structure CYP4A3 (PDB: 4D6Z [254]) is depicted with a solid surface. (b) Conservation of CYP9 genes across four managed bee pollinators. Level of conservation is depicted in a magenta to cyan (conserved to variable) gradient. (c) Multiple sequence alignment (MSA) of CYP9 genes from four managed bee pollinators. Conserved motifs are enclosed in red and SRSs in blue. The conserved cysteine is enclosed in yellow. Figure x(a) and (b) created using UCSF Chimera version 1.10.1.

3.3.4.1 Structural conservation of the CYP9 lineages

To investigate the conservation of the different CYP9 lineages, separate MSAs were used to overlay the crystal structure of *H. sapiens* CYP4A3 (PDB: 4D6Z [254]) using UCSF Chimera version 1.10.1 (see figure 3.12). Of the three CYP9 lineages found in the main CYP9 cluster, CYP9R is the most highly conserved between species (54-67% identity) (see figure 3.12). In both the CYP9R and CYP9P lineages, the SRS-containing proximal face shows regions of high conservation (magenta colour). The least well conserved of the CYP9 lineages are the CYP9Q/BU/DM sequences. Even without the inclusion of the CYP9DM genes, the percent identity ranges from 47-58%. The differences are even greater when only the conserved motifs and SRSs are aligned, with the CYP9P, CYP9R and CYP9Q/BU/DM lineages having 57-82%, 58-75% and 29-60% identity respectively (see appendix table 3.6).

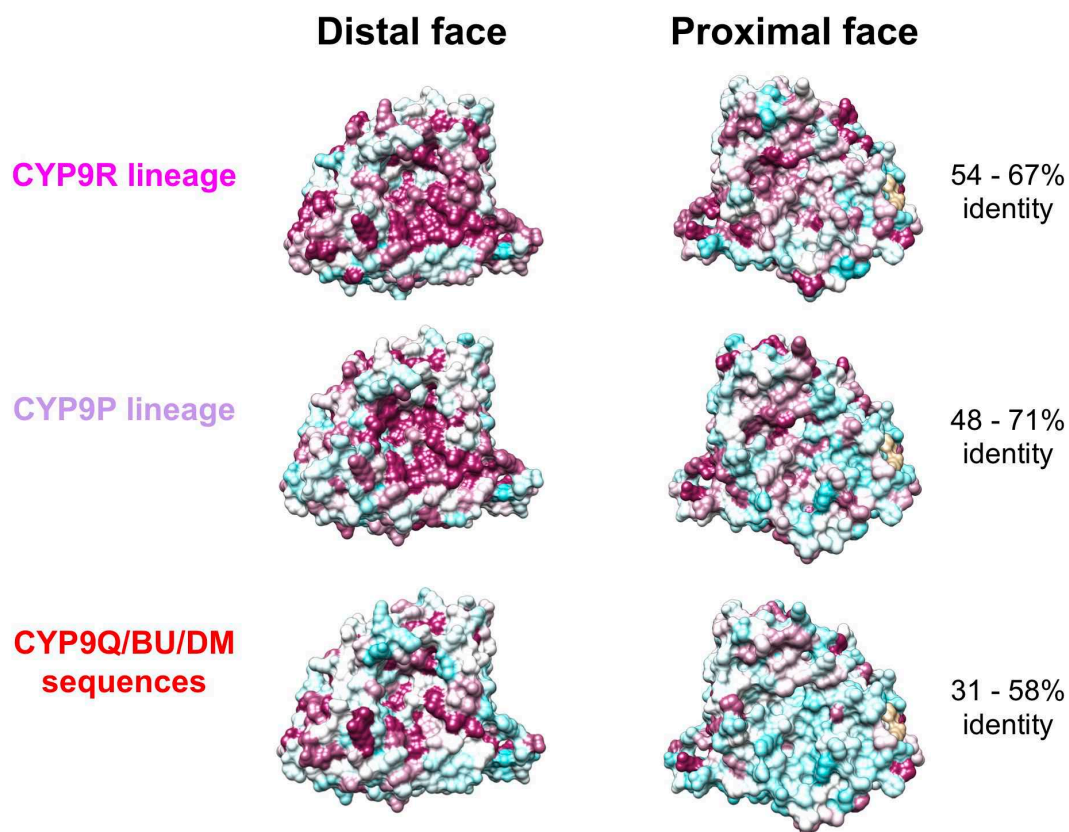


Figure 3.12: Conservation of CYP9 protein lineages across four managed bee pollinators (*A. mellifera*; *B. terrestris*; *O. bicornis* and *M. rotundata*). Level of conservation is depicted in a magenta to cyan (conserved to variable) gradient on crystal structure *H. sapiens* CYP4A3 (PDB: 4D6Z [254]). Percent identity shown is to *A. mellifera* sequences (*CYP9R1*, *CYP9P1* and *CYP9Q3*). Figure created using UCSF Chimera version 1.10.1.

3.3.5 Comparison of the structural homology of CYP9Q, CYP9BU and CYP9DM proteins

The CYP9 cluster in *M. rotundata* shows *CYP9DM1* and *CYP9DM2* in association with *CYP9R* genes, upstream of *myosin IIIb* (see figure 3.10(b)). The organisation of this genomic block is similar to that of the region containing *CYP9Q2* and *CYP9Q3* in *A. mellifera* (see figure 3.10(b)). Both position and reading frame direction appear to be close enough between the two species to warrant the *CYP9DM* sequences being investigated as potential *CYP9Q*, or *CYP9BU* homologs. However, the MSA of the *CYP9Q*s, *CYP9BU*s and *CYP9DM* amino acid sequences highlights some clear variances in the conserved motifs of the *M. rotundata* proteins, shown highlighted in orange in figure 3.13.

There are major substitutions within the structurally critical oxygen-binding motif (M2: GxE/DTT/S) (see figure 1.13). Both *CYP9DM*s have identical residues, NSAST, versus the consensus for the other bees, GFDTV (residues 332-336), highlighted in orange in figure 3.13. From this it appears that the oxygen-binding motif in both *CYP9DM*s is divergent, not only from other *CYP9* subfamily genes in bees, but also from P450 sequences more generally (for generalised primary structure see figure 1.13). The *CYP9DM* oxygen-binding motif has substitutions in four out of five amino acid residues when compared pairwise across the MSA. It does however, share the polar-polar, threonine-serine substitution (position 335 of the MSA) found in *CYP9BU1* (*O. bicornis*); *CYP9Q3* (*A. mellifera*) and *CYP9Q6* (*B. terrestris*) (see figure 3.13).

The conserved hydrophobic glycine (G) residue at the start of the oxygen-binding motif is replaced with a polar asparagine (N) in both *CYP9DM* sequences (position 332 of the MSA). An initial G substitution also occurs in *CYP9Q1* (G to serine (S): hydrophobic to polar) and *CYP9BU2* (G to alanine (A): both hydrophobic). In both *CYP9DM* sequences the charged acid residue (aspartic/glutamic acid; E/D), position 334 of the MSA, has been replaced with a hydrophobic A (see figure 3.13). This substitution appears to be unique to the *CYP9DM*s.

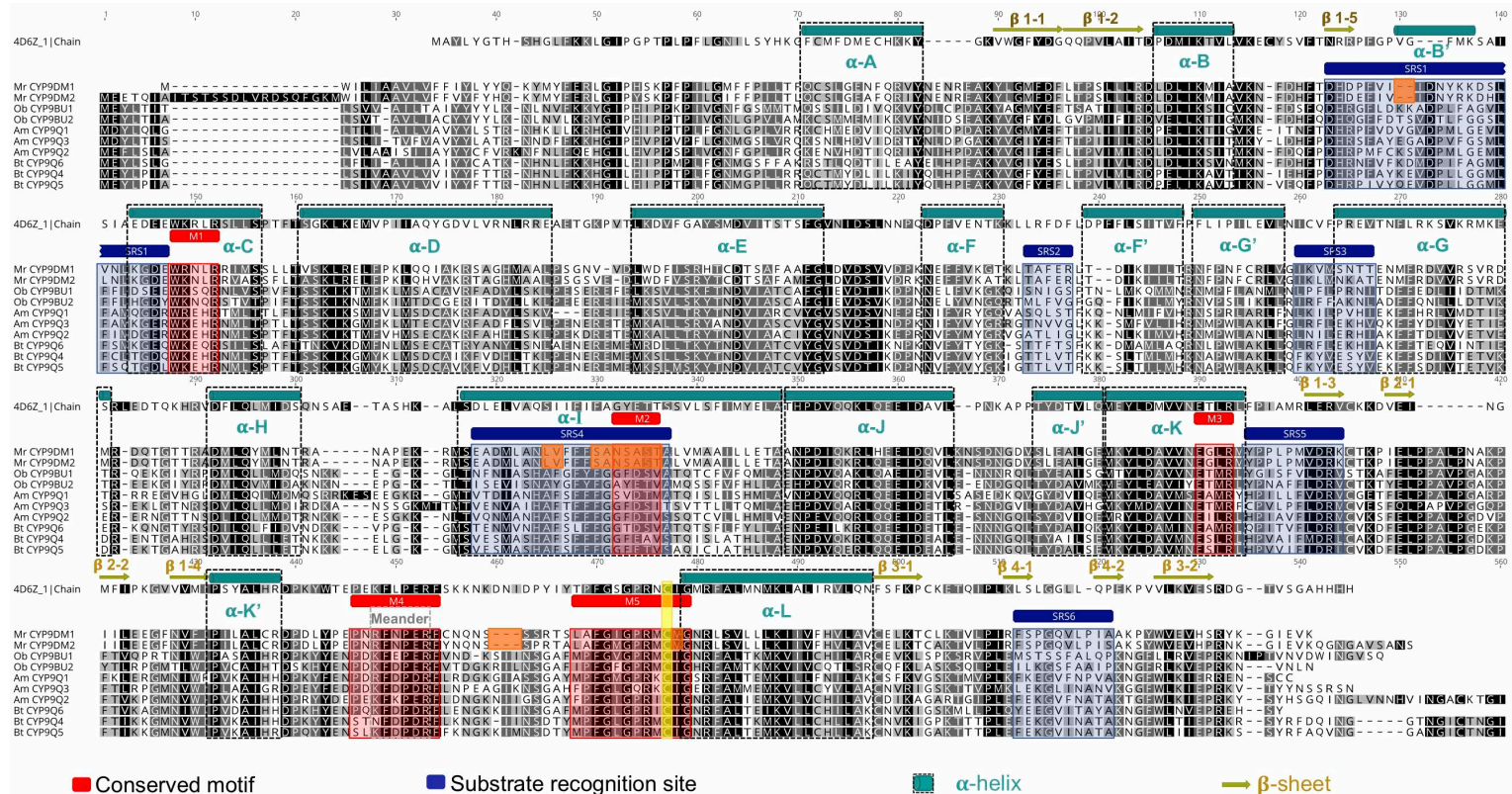


Figure 3.13: Multiple sequence alignment of *A. mellifera* and *B. terrestris* CYP9Qs with *O. bicornis* CYP9BUs and *M. rotundata* CYP9DMs [aligned in Geneious version 10.2.3 (Biomatters) using MUSCLE [451] (version 3.5, default settings). The sequences are coloured black to white according to their similarity. Conserved motifs (M) and substrate recognition sites (SRS) are shaded red and blue respectively and represented by annotations below chain A of the crystal structure *CYP4A3* (PDB: 4D6Z [254]). Secondary structures are annotated above chain A 4D6Z; dark cyan cylinders and dashed black boxes represent α-helices [with reference to the crystallographic structures of PDB: 4D6Z; 1TQN (*CYP3A4*: *H. sapiens* [250, 254]) and P450cam [248)]. Amino-acid substitutions and *M. rotundata* specific gaps are highlighted in orange.

There is also an isoleucine (I) to methionine (M) substitution (both hydrophobic residues) in the heme-binding motif (M5) (see figure 1.13) of both CYP9DMs (position 477 of the MSA) which changes the sequence immediately after the highly conserved cysteine, with CIG becoming CMG (see figure 3.13).

In figure 3.14(c) the structural positions of the substitutions in the CYP9DM M2 and M5 conserved motifs are modelled on the active site segment of *H. sapiens* CYP4A3 (PDB: 4D6Z [254]). The oxygen-binding motif (central part of helix I), the heme-binding motif and helix L are fundamental to the correct formation of the active site/heme binding pocket, the functionally important part of the protein (see figure 3.14(a) and (b)) [248, 249].

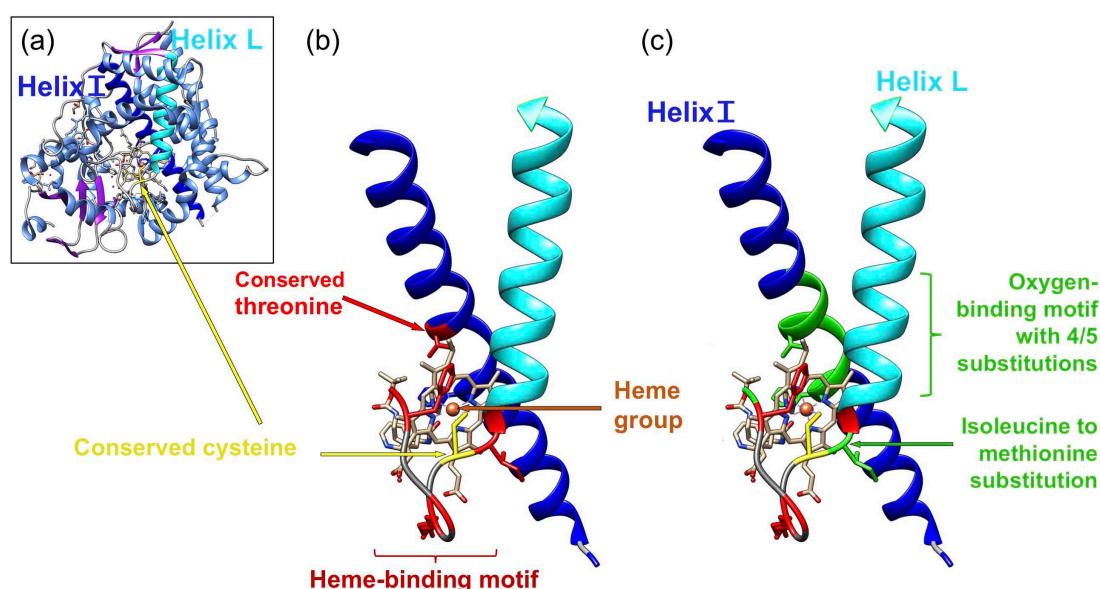


Figure 3.14: (a) Ribbon diagram of *H. sapiens* CYP3A4 (PDB: 4D6Z [254]), showing the secondary elements. β -sheets are coloured purple and coils are grey. Helices are coloured pale blue apart from helix I and helix L which are bright blue and cyan respectively. The highly conserved cysteine is coloured yellow (b) The P450 active site/heme-binding pocket, created by helix L and the central region of helix I. The heme group is depicted in light brown with a dark orange sphere marking the central iron (Fe) atom. The highly conserved cysteine is coloured yellow, other conserved residues are coloured red. (c) The P450 active site/heme-binding pocket with amino acid substitutions in CYP9DMs (*M. rotundata*) compared to CYP9Q/BUs coloured green. Figure created using UCSF Chimera version 1.10.1.

The substitutions found in both CYP9DM sequences are marked in green in figure 3.14(c) and can be seen to surround the heme group. Any major disruption to this part of a P450 molecule could potentially alter not only the substrate specificity, but also the correct folding of the protein, which is critical to its function.

3.3.6 Distribution of the CYP9 subfamily across 12 available bee genomes

Of the 12 available bee genomes (accessed from the NCBI database 2017-2018), 9 were from the Apidae family, 2 from the Megachilidae and 1 from the Halictidae (the genome for the Colletidae *C. gigas* was not released until 2020 and is not included in this phylogeny). The relationship between the CYP9 genes across the 12 species is in general agreement with the known phylogeny of bees (see figure 3.1(a)). However, the distribution of genes into the CYP9 lineages across the 12 species of bee studied is not even (see appendix table 3.7).

3.3.6.1 CYP9Q-like lineage

The 9 Apidae species all have CYP9Q genes, with the most basal of the species, *H. laboriosa*, having only CYP9Q9. *M. quadrifasciata* and *E. mexicana* appear to have had a single duplication event, whereas the *Apis* and *Bombus* species all have three CYP9Q genes. *O. bicornis* has 2 CYP9BU genes and *D. novaeangliae* 4 CYP9DL genes that share a recent common ancestor with the Apidae CYP9Q lineage, ancestral node highlighted in red in figure 3.15. The Maximum likelihood phylogeny in figure 3.15 estimates the relative time for divergence from *D. novaeangliae* CYP9DL and *O. bicornis* CYP9BU genes to CYP9Q genes, as 0.72 and 0.32 respectively. The only species that has no CYP9Q or closely related gene is *M. rotundata* (see figure 3.15 and appendix table 3.7).

3.3.6.2 CYP9R-like lineage

All species have at least one CYP9R gene, with *M. quadrifasciata* and *E. mexicana* showing recent duplication events, indicated by short branch lengths (see figure 3.15). All the CYP9R sequences share a single ancestral node, marked in pink, and duplication events are found within species rather than between (see figure 3.15).

3.3.6.3 *CYP9S*-like lineage

7 of the Apidae species and the Halictidae *D. novaeangliae* have one *CYP9S* gene. *D. novaeangliae CYP9S1* is basal to the lineage. Neither of the Megachilidae species have a *CYP9S* gene (see figure 3.15).

3.3.6.4 *CYP9P*-like lineage

9 of the 12 species have *CYP9P* genes annotated their genomes. However, when the scaffold containing the CYP9 cluster is examined manually using Geneious version 10.2.3 (Biomatters), the three species that appear to lack *CYP9Ps*: *A. florea*, *D. novaeangliae* and *H. laboriosa*, all have at least two candidate sequences that clade with the other *CYP9Ps* (see appendix figure 3.1). It appears therefore, that all 12 species have *CYP9P* genes.

The *CYP9P* lineage shows incidences of less recent duplication events than those seen in the *CYP9R* lineage, indicated by longer branch lengths and distinct separate clades of sequences. There are four distinct *CYP9P* lineages: 9P1, 9P2, 9P22 and 9P23-25, the ancestral nodes of these clades are numbered and marked in purple in figure 3.15. The *CYP9P* lineage appears to be the ancestral form of CYP9 gene.

3.3.6.5 *CYP9DM* lineage

Only *M. rotundata* has *CYP9DM* genes, shown in orange in figure 3.15. These genes are ancestral to the *CYP9Q*, *CYP9S* and *CYP9R* lineages.

3.3.6.6 *CYP9DN*-like lineage

Both Megachilidae and 3 of the Apidae species have a single *CYP9DN* gene.

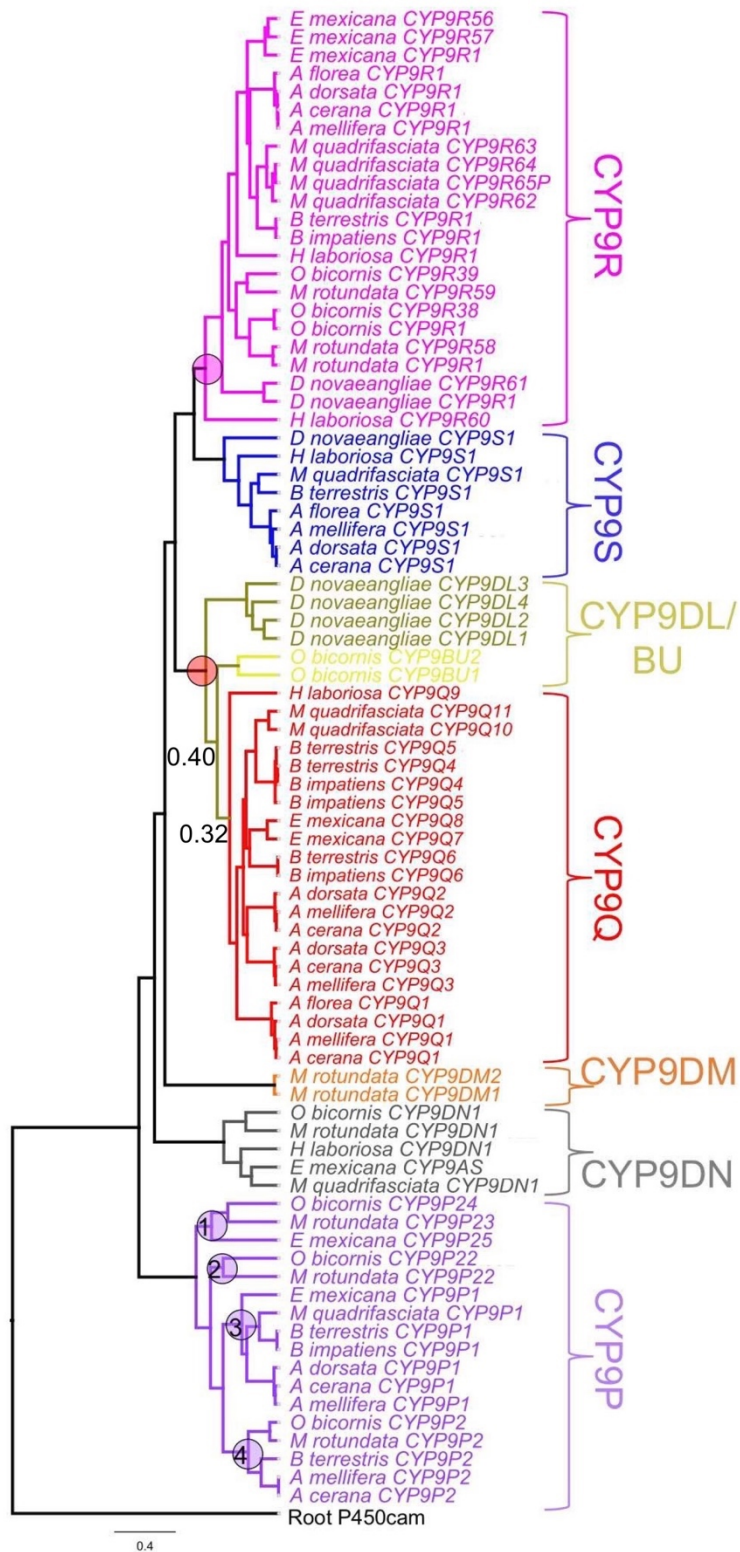


Figure 3.15: Phylogeny of CYP9 amino acid sequences from 12 bee species: *A. mellifera*; *A. cerana*; *A. dorsata*; *A. florea*; *B. terrestris*; *B. impatiens*; *D. novaeangliae*; *E. mexicana*; *H. laboriosa*; *M. quadrifasciata*; *M. rotundata* and *O. bicornis*. Phylogeny estimated using PhyML Maximum likelihood algorithm [453] and substitution model LG+G [454]. The branches show relative time.

3.4 Discussion

3.4.1 The CYPome in *M. rotundata*

The discovery of the substantially reduced number of detoxification genes in the *A. mellifera* genome [458], led to the question of whether this is something common to all bee species. The data from this chapter shows the paucity of P450 genes found in *A. mellifera* (46 sequences) is also reflected in *M. rotundata* (49 sequences). In agreement with other research, such as the 2018 study by Johnson *et al.* [459], this reduced P450 repertoire extends across the other 10 species of bee included in this chapter, none of which have a CYPome containing more than 53 sequences. A contracted CYPome, therefore appears to be the norm across bee families. The 12 species examined here have diverse and contrasting life histories, which range from solitary, through primitively social, to obligatory eusocial and yet all show the reduced suite of P450s. The initial suggestion that *A. mellifera* might offset its' reduced repertoire of P450s, with complex social behaviours and colony-level homeostasis, therefore seems less likely to be the case [29, 211, 460].

When compared to other insects such as *D. melanogaster* or *A. gambiae*, *M. rotundata* has an impoverished CYP4 clan, analogous to that seen in *A. mellifera* [211]. Both bee species have 12-14% of the number of CYP4 sequences found in other insect genomes [206]. CYP4 clan members are numerous, highly diverse and have been implicated in both endogenous and exogenous functions in other insect species [211, 241, 457]. Although definitive functions have not been determined for the CYP4 P450s in *A. mellifera*, they have all been implicated in endogenous roles such as lipid metabolism and ecdysteroid synthesis [211]. The *M. rotundata* CYP4 sequences group broadly into four clades with the sequences from the other managed bee pollinators, see figure 3.6. It is reasonable to assume that the sequences are therefore likely to be orthologs, with similar endogenous functions.

The CYP2 and mitochondrial CYP clans in *M. rotundata* also contain broadly similar numbers of P450s to those of *A. mellifera*. Once again, the sequences from the managed bee pollinators group into distinct clades, see figure 3.6. Several members of these P450 clans in *A. mellifera* have 1:1 orthologs in *D.*

melanogaster, with functions relating to hormone biosynthesis [211]. It is reasonable to assume that the orthologous genes in *M. rotundata* will code for enzymes with similar functions.

In common with *A. mellifera*, the most populous CYP clan in *M. rotundata* is the CYP3 clan, see table 3.1 and figure 3.6. This clan is associated with xenobiotic detoxification in many insect species [206]. The clustering of *CYP6AS* and *CYP9* sequences in the genomes of *A. mellifera*, *B. terrestris*, *O. bicornis*, *D. novaeangliae* and *M. rotundata* is clear evidence of relatively recent duplication events within these subfamilies across bee species, see table 3.2 and figure 3.10.

All the major duplications in the CYP9 subfamily appear to have arisen before the divergence of the bee families, with all species having members of the *CYP9P* and *CYP9R* lineages, see figure 3.15. With the exception of *CYP9DN1*, which is found on a separate scaffold, the CYP9 sequences are all intronless genes. Intronless genes are suggestive of an ancestral retrotransposition event into the germ line [211, 461]. The genome of *A. mellifera* shows a relative lack of transposable elements (TEs) when compared to other insect genomes, with an overall TE content of 4.25% compared to a median value of 24.4% for other insect species [458, 462]. However, the retrotransposition event for the CYP9 cluster would have occurred before the divergence of the bee families, and there is ample evidence of degraded copies and remnants of TEs in the genome of *A. mellifera* [458].

When only the four managed bee species were considered, it appeared that *CYP9DN1* was specific to the Megachilidae, see figure 3.6. However, when the 12 bee genomes are included in a phylogeny, three members of the Apidae family, *H. laboriosa*, *E. mexicana* and *M. quadrifasciata*, have a *CYP9DN1* gene, see figure 3.15. Apart from the obvious caveat that four genomes are too few to draw reliable conclusions from, the presence of a *CYP9DN1* gene in some members of the Apidae family begs the question of why there is an absence in the *Apis* and *Bombus* genera. It would appear there has been a loss of this CYP9 gene in these genera.

The other notable CYP bloom found in *M. rotundata*, and the other managed bee pollinators, is in the *CYP6AS* subfamily. There appear to have been eight *CYP6AS* duplications that occurred before the divergence of the Apidae and Megachilidae families, see figure 3.6. In six of these clades additional, more recent, species-specific duplications have occurred. *CYP6AS* genes have been implicated in the metabolism of certain flavonol phytochemicals, such as quercetin, that are found in nectar, pollen and resin, [286, 292, 459]. The importance of flavanols as a class of phytochemical that impacts bees in general, cannot be underestimated. Flavanols are universally present in both nectar and pollen [459]. Leafcutter bees, such as *M. rotundata*, also construct a nest lined with cut angiosperm leaves and in this process are exposed to a further source of flavanols [459]. The larvae of leafcutter bees may also be exposed to flavanols present in cut foliage, as the compounds could leach into the mass provisions in the nest cell [459].

Two other P450 lineages in the CYP3 clan, *CYP6AQ* and *CYP336A*, show expansion in *M. rotundata* when compared to *A. mellifera*, see figure 3.6. Both these lineages have three members in *M. rotundata* compared to a single sequence in *A. mellifera*. However, there has been expansion of the *CYP336A* lineage in *O. bicornis* and of both lineages in *B. terrestris*. *CYP6AQ55*, *CYP336A35* and *CYP336A36* have been implicated in the metabolism of the phytochemicals: nicotine, atropine and hyoscyne respectively in *O. bicornis* [463].

There are distinct differences between the CYPomes of bee species, particularly within the CYP3 clan. The recent, species-specific duplications seen in the CYP3 of the managed bee pollinators may reflect a subtle difference in exposure to specific phytochemicals, caused by the disparity in diet and life history. Although *M. rotundata* and *O. bicornis* are polylectic, they are undoubtedly temporally and spatially constrained in the diversity and availability of floral resource, when compared to the annual eusocial *B. terrestris* or the perennial eusocial *A. mellifera*. The temporal constraint of a short, annual life-cycle of the two solitary bees, may have resulted in the evolution of a P450 repertoire tailored to plant species available during the flight season of each individual species.

In conclusion it appears that the reduced inventory of P450s, compared to other insect species such as *D. melanogaster* or *A. gambiae*, first described in the *A. mellifera* genome, is ubiquitous across three bee families. It is therefore unlikely that whether the species is organised socially, or is solitary, explains the depauperate CYPome. One consequence of the selection pressure from plant biosynthesis of allelochemicals on phytophagous insects is undoubtedly that sequential tandem gene duplication events, followed by neofunctionalization of enzymes, have occurred in insect P450s, such as those seen in the CYP9 and CYP6AS subfamilies [217, 464, 465]. However, a second well-established feature of P450s is that, in certain cases, such as *H. sapiens* CYP3A4 and CYP2D6, a single enzyme metabolises multiple, structurally diverse substrates [225, 465]. The evolution of such promiscuous enzymes could negate the drive to increase the number of P450 genes. In the CYP9Q and CYP9BU lineages *A. mellifera*, *B. terrestris* and *O. bicornis* have evolved enzymes capable of metabolising multiple substrates [137, 168, 226].

3.4.2 The CYP9 subfamily audit

There are remarkable similarities in the genomic structure of the CYP9 cluster of genes across bee species, but there are also key differences (see figure 3.10(a) and (b)). Most notably, are the differences with the insecticide-degrading CYP9 P450 enzymes (i.e. the CYP9Q and CYP9BU lineages). It is clear from figure 3.15, that of the 12 species included, only *M. rotundata* lacks a gene that clades with CYP9Q and CYP9BU lineages. Indeed, the Halictidae *D. novaeangliae* has four CYP9DL genes that share a common ancestor with the CYP9Q lineage [447]. That *M. rotundata* lacks a CYP9Q/BU ortholog shows that these enzymes pre-adapted to metabolise certain insecticides are not present across all bee species.

The obvious question that arises from this is: Can we use the lack of a CYP9Q/BU-like ortholog to predict an increased sensitivity to certain insecticides? Without reliable toxicological data and/or study of the molecular basis of detoxification in this species, caution needs to be exercised. It is possible that *M. rotundata* has evolved other P450s, or indeed, non-P450 detoxification enzymes that are capable of metabolising a similar set of

insecticides to the *CYP9Q/BU*-like sequences. From the data currently available on the biotransformation systems across bee species it is not clear that phylogeny of the CYP3 clan can be described in terms of functional clades.

When syntenic analysis is included alongside phylogeny, it is evident that the main CYP9 cluster sits within a syntenic block that is conserved across four bee families, see figure 3.10(b). The phylogeny (figure 3.10(a)) puts *M. rotundata* *CYP9DM* genes distant from *CYP9Q/BU/DL* sequences, with a branch length of 1.1 substitutions per 1.0 residues. This indicates the potential for a large number of substitutions and consequential sequence change. However, *CYP9DM1* and *CYP9DM2* occupy a similar position and reading frame direction to the *CYP9Q/BU/DL* sequences within the genomic landscape. In their 2006 study, Zdobnov & Bork reported a linear correlation between protein sequence identity and gene order, which predicts that ancestral gene order is lost when protein identity is lower than 50% [430]. It may be that the synteny observed around the CYP9 cluster in bees is a result of the flanking genes being so highly conserved, see table 3.4, rather than a product of the P450s themselves. Sequence identity in P450 proteins is notoriously low, sometimes as low as ~20% [243] and so, it may well be that, for gene order to be maintained around a CYP cluster, highly-conserved flanking genes are essential. However, conservation of synteny has been observed with genes below the threshold of 50% identity in Lepidoptera and so there may be constraints, other than protein sequence identity, that prevent complete scrambling of gene order [466].

There is good evidence for at least one genomic rearrangement, that occurred after the divergence of Apidae and Megachilidae families (timescale 104-125 mya), which results in *membralin* becoming flanked, by *allostatin A receptor*, see figure 3.9 and appendix table 3.4. There is a high level of gene order conservation, in the bee families, in the region 500 kb upstream and downstream of the CYP9 cluster, with 55-68% of the Apidae genes having orthologs present in the Megachilidae, see table 3.2. A similarly high level of synteny also appears to be present in the region around the main CYP6AS cluster across *A. mellifera*, *B. terrestris* and *M. rotundata* (see table 3.2). These levels are in keeping with the higher end of the spectrum of synteny found across insect orders [430]. Genome assemblies are imperfect and can be

fragmented, something that is known to lead to errors in synteny analyses. For example, it is often the case that the length of synteny blocks is underestimated due to fragmented genomes. The high level of synteny shown here, might therefore, reflect constrained genomic shuffling in the 104-125 million years since the divergence of the Apidae and Megachilidae families. Studies of the recombination landscape of bee genomes has revealed that although *A. mellifera* and *B. terrestris* have very high recombination rates when compared to *M. rotundata*, the recombination landscape of the three species is very similar [467-469]. In comparison to the rest of the genome, recombination rates in the coding regions of all three species are significantly reduced [467-469]. The observation of high levels of synteny and the implication of constrained genomic shuffling is in keeping with the idea of slow genome evolution in bees.

The tertiary protein structures of the CYP9R and CYP9P lineages are relatively well conserved (48-71%), see figure 3.12. There is greater identity of these lineages when only the conserved structural elements and substrate recognition are considered (58-82%), see appendix table 3.6. This degree of identity is due to the levels of conservation seen in the distal face of the molecule that houses the conserved motifs in both these lineages, figure 3.12. In the CYP9Rs the proximal face, which houses the SRSs, appears more poorly conserved, figure 3.12. This implies that the SRSs of the *CYP9R* lineage, across the managed bee species, are diverse, something that presumably allows for a range of substrate specificity. The phylogeny of the CYP9 subfamily indicates that the *CYP9P* lineage is likely to be sister to the other lineages, see figure 3.15. It may be therefore, that these enzymes have functions that are conserved across bee species. The later lineages, that presumably arose from original *CYP9P* duplication events, may have evolved more species-specific functions. Of all the CYP9 lineages the *CYP9Q/BU/DMs* have the lowest level of conservation, see figure 3.12. The conserved motifs and SRSs for CYP9DM1 and CYP9DM2 show the lowest similarity (>40% identity) to the other members of the *CYP9Q/BU* lineage, with the greatest similarity being with *A. mellifera* CYP9Q1 (37.2%), see appendix table 3.6.

The conserved structural core of P450 enzymes is comprised of a four-helix bundle, see section 1.6.1.1.2 and figure 1.14 [249]. When CYP9DM1 and

CYP9DM2 are modelled on *H. sapiens* CYP3A4 (PDB: 4D6Z [254]) and the portion of the protein which houses the active site is examined, it can be seen that helices I and L, which surround the heme molecule, contain six amino acid substitutions, see figure 3.14(b) and 3.14(c). Helix L forms part of the heme-binding region providing the highly conserved G from the 'CIG' end of conserved motif five (M5), see figure 3.13. Helix I contains the oxygen-binding motif (M2) and SRS4, see figure 3.13, and is therefore involved in both substrate binding and the reaction mechanism [249]. The sequence of helix I in CYP9DM enzymes is divergent, not only from other bee CYP9s, but also from the more generally accepted structure of P450s, as it lacks the conserved G and charged acid residue (E/D) of the M2 motif. This structural difference in combination with the fact that the *CYP9DMs* are phylogenetically distinct raises the question of whether the tertiary folding and/or function of these enzymes may also deviate in some manner.

In conclusion, of the 12 species examined, only *M. rotundata* lacks a *CYP9Q/BU/DL* ortholog, or indeed any sequence that shares a common ancestor with the clade. Using syntenic analysis of the CYP9 cluster the most likely candidates for *CYP9Q/BU/DL* orthologs are the *CYP9DM* sequences. However, these sequences appear to be structurally divergent from other CYP9 enzymes. This has potential implications for the ability of *M. rotundata* to metabolise certain insecticides. The sensitivity or tolerance of *M. rotundata* to insecticides from three MoA classes is explored using *in vivo* acute contact bioassays and *in vitro* functional analyses in chapters four and five of this PhD. These data will help elucidate whether phylogeny has the potential to be used to predict insecticide tolerance or sensitivity.

Given that the CYP9Q/BU-like enzymes, known to provide intrinsic tolerance to certain synthetic insecticides, are not universally present in bees, there is a question of how wide-spread a lack of *CYP9Q/BU*-like orthologs might be. Although there is a dearth of bee genomes from non-Apidae families, the fact that *O. bicornis* has *CYP9BU* genes tells us that their absence is not common to all Megachilidae species. However, there is a need to investigate CYP9 cluster sequences from more bee species, particularly those from non-Apidae families.

It would also be highly informative to have data from other species of the *Megachile* genus. These issues are addressed in chapter six of this PhD.

Chapter four: Acute contact toxicity bioassays of select insecticides against *M. rotundata*

4.1 Introduction

Insecticides are widely used in agriculture to control pests, in order to eliminate yield losses and maintain the quality of the crop (see section 1.8.3). For an insecticide to be registered as safe to use, the full environmental risk must be evaluated [470]. Water, air and soil contamination must be considered, as well as the potential toxic effects to non-target species, such as pollinators or entomophagous insects [470]. To this end, toxicity testing of non-target species needs to be carried out. The testing process consists of 3-tiers: acute toxicity tests; semi-field condition tests and field tests [396, 470, 471]. Currently, pesticide registration in the US and EU only requires acute contact and oral toxicity tests on a single bee species, *A. mellifera* [472-475]. These first-tier tests using *A. mellifera* are designed to flag up compounds that are likely to be of minimal risk to bees and so are suitable to go forward for further safety testing [472]. The need to develop standardised test protocols for bumble bees and solitary species has been recognised since the European Food Safety Authority (EFSA) published its' guidance document on the risk assessment of pesticides on bees in 2013 [472]. However, registration regulations have not changed at the time of writing.

It has been suggested that, with the use of an assessment (multiplication) factor, *A. mellifera* can be used as a surrogate in ecotoxicology tests for other bees, including solitary species [476-478]. The concept of extrapolating data, from one species to another, based on body weight, is widely used in toxicology, particularly with vertebrate species in the sphere of human medicine [479, 480]. The biological basis for extrapolating experimental data from one animal species to another, is built on an underlying assumption that they are physiological and biochemical similar [479]. There is, however, acknowledgment that singular differences may occur in one species that are not found in others [479]. In the paradigm of bee toxicology, there is some agreement that applying an assessment or bridging factor of 10, to *A. mellifera* LD₅₀ endpoints, is sufficient to protect 95% of other bee species [476-478]. As such, an assessment figure of 10 is included in the European Chemical

Agency's 2020 scoping document, which sets out preliminary considerations for guidance on the risk of biocides to bees [475].

With the reports of insect population declines [16, 20, 298, 481] it has become more essential to determine the inherent toxicity of insecticides to other pollinator species, particularly to other managed bee pollinators such as *M. rotundata*, rather than to rely solely on an assessment factor. The Organisation for Economic Cooperation and Development (OECD) has developed standard guidelines for assessing the toxicity of insecticides to *A. mellifera* [482, 483] and *B. terrestris/ impatiens* [484, 485]. These guidelines describe standard methodologies to assess the toxicity of chemicals to *A. mellifera* and *Bombus* species via the two primary routes of exposure: contact (from droplets of a spray) and oral (consumption of contaminated pollen or nectar). However, no standardized methodologies for assessing the toxicity of chemicals to solitary bees have been published by the OECD as yet. Nevertheless, toxicological studies of certain species are beginning to be published [226, 486], and *Osmia* species have been suggested as potential model species for solitary bees [472]. Unlike *A. mellifera*, solitary bees do not share food via trophallaxis and therefore may need to be housed individually, in a similar fashion to *Bombus* species, for oral toxicity tests [485]. Work on *O. bicornis* has shown that there may be issues with solitary bees not feeding well under test conditions, which compromises the reliability and robustness of the results [463]. Given the limited seasonal availability of *M. rotundata*, only acute contact toxicity tests were carried out in this PhD. The methods outlined below are principally based on OECD Test No. 214 [482] and OECD Test No. 246 [484], with reference to the International Commission for Plant Pollinator Relationships (ICPPR) Solitary bee, Acute Contact Toxicity Test protocol [487].

4.1.1 Toxicity tests in *M. rotundata* to date

Given the economic and agricultural importance of *M. rotundata*, there have been studies of insecticide toxicity published on the species since the 1960s [90, 91, 488-491]. However, the results of these studies report LD₅₀ values in µg a.i./bee, µg a.i./g bee, mg a.i./l and % v/w, making direct comparison between them difficult. Furthermore, different methodologies were used in the collection of the data, making straightforward comparison with the data from other bee

species extremely problematic [29, 477]. In 1994, Helson *et al.* reported that *M. rotundata* was less susceptible than *A. mellifera*, but more susceptible than *B. terrestris*, to six insecticides from three different MoA classes [490]. Whereas, in 2003, Devillers *et al.* reported that *M. rotundata* was more susceptible to 32 insecticides (across six MoA classes) than either *A. mellifera* or *B. terrestris*, but less susceptible than the alkali bee (*Nomia melanderi*) [394]. However, the size of bee may be important when reporting LD₅₀s in µg a.i./bee. In general, the order of sensitivity to insecticides reported in the literature is *B. terrestris* < *A. mellifera* < *M. rotundata*, but in certain cases the discrepancy in body size between the species might explain the data [394]. Many of the compounds that have been tested against *M. rotundata* in these studies have since been withdrawn from the market. There is therefore a need for toxicity data for this species for insecticides currently registered for agricultural use.

4.1.2 Acute contact toxicity test

Acute contact toxicity tests reflect exposure to an insecticide via droplets of a spray or contaminated nest material. They are first tier tests that determine the intrinsic toxicity of an insecticide after topical application of a single dose at a range of concentrations, within a maximum time period (usually 96h). The tests performed should follow standard guidelines [482, 484]. LD₅₀ is usually expressed in mass of a.i. (µg, ng) per individual or mass of a.i (µg, ng) per mg body weight [166, 492]. Laboratory testing of the acute toxicity of a compound has two main advantages. Firstly, the LD₅₀ is a measurement of an absolute level of toxicity and is not linked to a recommended dose for field application [164]. Secondly, the units in which an LD₅₀ is expressed means that these measures are potentially comparable to each other, more so if the methodology is standardised [164, 166]. LD₅₀ values are also used to calculate the toxicity exposure ratio (TER), hazard quotients (HQ) and risk quotients (RQ) [166, 470]. These three measures (TER, HQ and RQ) are used to determine the risk from an insecticide to a bee species in the field [166, 470]. The U.S. environmental protection agency (EPA) classifies the risk from pesticides to bees into broad categories based on the measurement of their LD₅₀s (see table 4.1) [493]. However, it is important that the final assessment of risk from any substance be established by integrating the available evidence from all tiers of testing [396].

Table 4.1: EPA classification categories of pesticide toxicity in bees based on LD₅₀ values [493].

Classification category	LD₅₀ (µg ai/bee)
Category I – Highly toxic	< 2.0
Category II – Moderately toxic	2.0 - 10.9
Category III – Practically non-toxic	> 11

4.1.3 Chapter aims and underpinning questions

This chapter aims to provide acute contact toxicity data for *M. rotundata* using methods based on the OECD guidelines for *A. mellifera* and *B. terrestris/impatiens* [482, 484]. This methodology will produce robust data to inform on insecticide toxicity in a solitary, non-Apidae bee and allow for comparison with other species (*A. mellifera*, *B. terrestris* and *O. bicornis*) tested using similar methodology [137, 168, 226].

As mentioned in sections 1.4.4.1 and 1.7.3, *A. mellifera*, *B. terrestris* and *O. bicornis* exhibit a marked difference in sensitivity to *N*-nitroguanidine and *N*-cyanoamidine neonicotinoid insecticides [137, 155, 168, 226, 446]. It is not clear whether this differential sensitivity will extend to *M. rotundata* and this chapter aims to answer this important question.

In chapter three the curation of the *M. rotundata* CYPome highlighted that this species lacks an ortholog of the insecticide-degrading CYP9 P450 enzymes from either the *CYP9Q* or *CYP9BU* lineages [137, 155, 168, 226, 446]. This chapter aims to provide key results needed to test the veracity of using phylogenetics as part of a ‘tool-kit’ to predict function in this species. This approach may then inform the screening of future insecticides.

The results from sections 4.3.2.1 and 4.3.3.1 of this chapter are published in Hayward *et al.*, 2019 *Nature Ecology and Evolution* 3(11):1521-1524 [447].

4.2 Methods

4.2.1 Care and maintenance of bees

For this PhD *M. rotundata* cocoons were obtained from Bayer (AG, Crop Science Division, Leverkusen, Germany). Approximately 5000 cocoons were

received in 2018 and 2019. On arrival the cocoons were stored, in groups of approximately 100, in aerated plastic containers (125 x 85 x 65 mm) at 4°C in complete darkness. These conditions ensured that prepupae remained in diapause. To enable the sequential emergence of experimental groups of bees, approximately 180 cocoons were placed into 3 plastic containers (125 x 85 x 65 mm; see figure 4.1) and were warmed for 24 h under the following conditions: temperature: 24°C, humidity: 55%. The cocoons were then incubated for 28 days under the following conditions: temperature: 30°C, humidity: 55%, L/D cycle 24 h dark. Enough cocoons were sequentially incubated to provide two experimental groups per week (see figure 4.2). Males emerged day 18-22; females emerged day 21-28 (for full incubation and emergence schedules for the species see [48]).



Figure 4.1: *M. rotundata* cocoons for incubation (groups of approximately 60 cocoons per box)

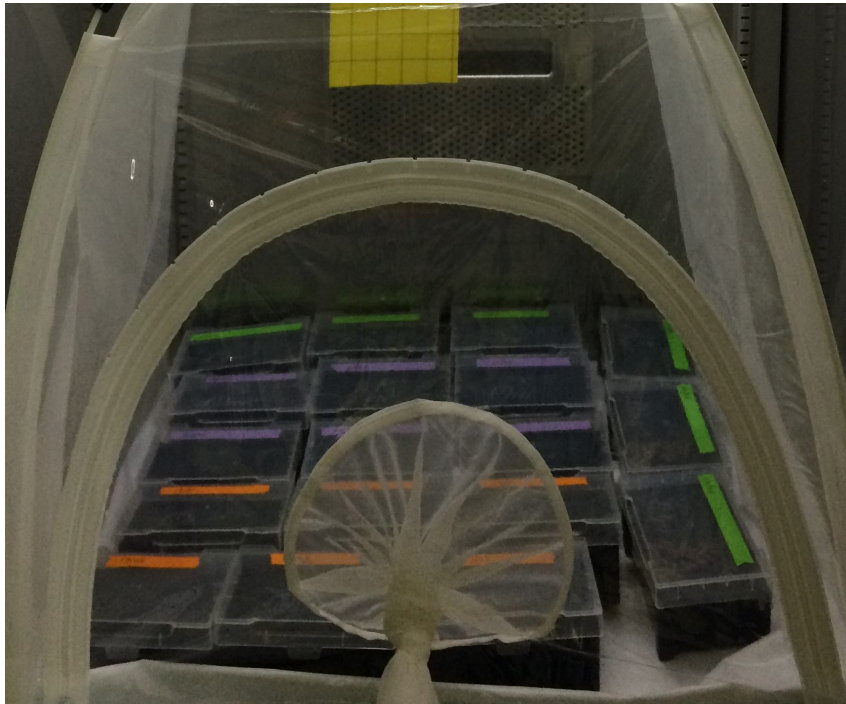


Figure 4.2: Incubator set up showing the sequential groups of cocoons (coloured by week)

Once emerged, bees were separated by sex based on the colour of their eyes and setae on their heads (males: blue eyes, yellow setae; females: black eyes and white setae [48]). Bees were placed in well-ventilated plastic holding cages (120 mm diameter x 100 mm high; see figure 4.3) in groups of 10-15 individuals. Sucrose solution in purified water with a final concentration of 500 g/l (50% w/v) was used as food and provided *ad libitum* at all times, using soaked cotton wool in a feeder (35 mm diameter 10 mm high). Feeders were placed on the ground of the cages. Other solitary species have been documented preferring the opportunity to play/have hiding places [487] and so holding cages were provided with enrichment in the form of paper (see figure 4.3). During holding and testing disturbance to the bees was minimised.

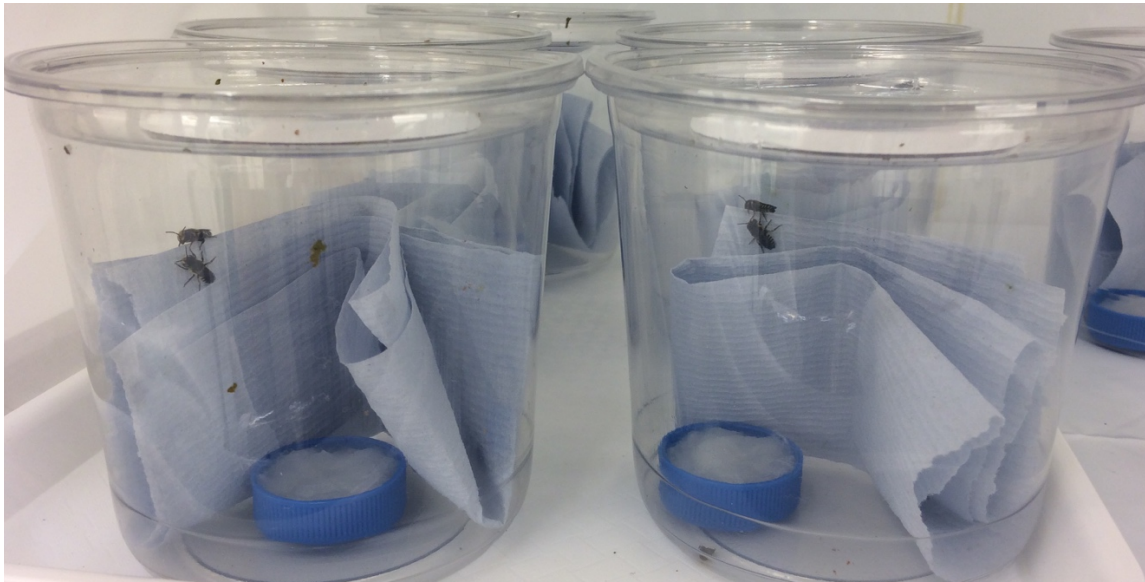


Figure 4.3: Female *M. rotundata* in holding cages (with enrichment and feeders)

4.2.2 Insecticides

To determine whether the marked difference in sensitivity to the different classes of neonicotinoid insecticides found in other bee species extends to *M. rotundata*, two *N*-cyanoamidine (thiacloprid and acetamiprid) and one *N*-nitroguanidine neonicotinoid (imidacloprid) were selected as exemplars of their class [137, 155, 168, 226, 446]. The P450 enzymes that metabolise *N*-cyanoamidine neonicotinoids in *A. mellifera* and *B. terrestris* have also been shown to metabolise the pyrethroid *tau*-fluvalinate and the organophosphate coumaphos [168]. Both these insecticides were selected for testing along with a second exemplar of their class known to be highly toxic to other bee species. For this, the pyrethroid deltamethrin and the organophosphate chlorpyrifos were selected.

Technical grade insecticides were obtained from Sigma. A stock solution of 100 µg/ µl (100,000 ppm) for each compound was prepared by diluting the insecticide in 100% acetone. All stock solutions were made in glass vials and stored at 4°C in the dark. Further dilutions were made from the stock solutions. An aliquot of the lowest and highest concentration of each insecticide used was frozen at -20 for analytical determination of the concentration if required. Pyrethroids are known to bind strongly to plastics and so all experimental cages were discarded at the end of bioassays using *tau*-fluvalinate and deltamethrin, to prevent a build-up of the compounds or contamination of other tests. Range

finding tests were carried out using male or female bees, to determine the experimental concentrations of each test compound.

4.2.3 Preparation of bees

Holding cages were selected randomly and the bees were anaesthetized in groups of 10 using CO₂. Moribund bees were rejected and replaced by healthy individuals before starting the test. Anaesthetised bees were individually treated by topical application with a hand-held micro-applicator (Hamilton repeating dispenser PB600-1). A volume of 1 µl of test substance solution was applied to the dorsal side of the thorax of each bee (between the neck and wing base). No enrichment was provided in the experimental test cages (see figure 4.4). A minimum of three replicate test groups, each of ten female bees (~24-48 h old), were dosed with each insecticide concentration. One control group, each of ten female bees, dosed with 1 µl 100% acetone was included in each bioassay. Bees that did not recover within 30 minutes of treatment were excluded from the experiment.



Figure 4.4: Experimental set up for female *M. rotundata* bioassay

4.2.4 Range finding test

Range finding tests were carried out to determine the experimental concentrations of each test compound. In 2019, range finding bioassays were performed using male *M. rotundata* to allow more females to be included in the experimental tests [478]. The considerable size difference between male and

female *M. rotundata* [48] was taken into consideration when deciding on the experimental concentrations. To prevent males huddling together in tight balls and enable accurate scoring, cages containing males for range-finding tests were provided with filter paper as enrichment (see fig 4.5).



Figure 4.5: Experimental set up for range-finding test using male *M. rotundata*

4.2.5 Test conditions

The bees were held in test cages at a temperature of $24 \pm 2^{\circ}\text{C}$, relative humidity of 55% and light/dark cycle of 8/16 h throughout the test. The duration of the test was 96 h. Mortality was recorded 4 h after dosing and thereafter at 24 h, 48 h, 72 h and 96 h. Bees were assessed and scored according to the criteria set out in table 4.2. All abnormal behavioural effects observed during the testing period were recorded. Each test cage was filmed at each time point as a reference.

Table 4.2: Scoring criteria for acute contact toxicity test using *M. rotundata* (based on [484]).

Scoring criteria post exposure			
Unaffected	Affected	Moribund	Dead
<p>Individual:</p> <ul style="list-style-type: none"> • Can stand in a coordinated manner • Can walk in a coordinated manner • Can fly in a coordinated manner and lands well • Is feeding <p>Scored as 'No effect' (NE)</p>	<p>Individual:</p> <ul style="list-style-type: none"> • Cannot fly in a coordinated manner, briefly takes off but quickly falls down or hits the cage walls, lands poorly • Is crawling and uncoordinated • May recover <p>Scored as 'Spasm' (S)</p> <ul style="list-style-type: none"> • Is motionless as if frozen, but will move if nudged with forceps <p>Scored as 'Frozen' (F)</p>	<p>Individual:</p> <ul style="list-style-type: none"> • Is motionless and unresponsive to stimuli, but upright <p>Scored as 'Immobile/Paralysis (IP)</p> <ul style="list-style-type: none"> • Lies on its back, moving legs and antennae, but unable to move in a coordinated manner • Cannot walk in a coordinated manner, does not right itself if pushed over gently with forceps • Unlikely to recover <p>Scored as 'Effectively dead" (ED)</p>	<p>Individual:</p> <ul style="list-style-type: none"> • Is immobile, does not move when touched with tip of forceps • Only twitches slightly when touched with tip of forceps <p>Scored as 'Dead' (D)</p>

4.2.6 Analysis

Bioassays with control mortality >10% were excluded from analysis. The relationship between concentration and mortality was determined using probit analysis [494]. LD₅₀ values and their respective 95% confidence interval values were calculated in Le Ora Software PoloPus (Version 1.0) and IBM SPSS (Version 26). Correction for control mortality was made using Abbott's correction [495]. A non-linear regression was selected to generate a dose response curve in GraphPad Prism (Version 8.1.0 (325)) at each observation time (24 h, 48 h, 72 h and 96 h) and the slope (\pm SE) at the steepest part of the curve was calculated.

4.2.7 Comparison with other managed bee pollinator species

The standardised protocol for acute contact toxicity testing in bees [482, 484] allows for data generated under similar conditions to be compared across species. Three species of managed bee pollinator, namely *A. mellifera*, *B.*

terrestris and *O. bicornis* were selected for data comparison. These species were specifically chosen because the majority of the acute contact toxicity test data needed could be taken from publications from within either the Bass Group (University of Exeter), or Bayer AG, Crop Science Division (Leverkusen, Germany), where tests were performed under similar conditions and using comparable protocols [137, 168, 226].

4.3 Results

4.3.1 Range finding tests

Range finding tests were carried out using the concentrations shown in table 4.3.

Table 4.3: Concentration of insecticides used in range finding tests

Compound	Concentration $\mu\text{g}/\mu\text{l}$ ($\mu\text{g}/\text{bee}$)
Neonicotinoids	
Thiacloprid	50, 20, 1, 0.1, 0.01, 0.001, 0.0001, 0.00001, 0.000001
Imidacloprid	0.001, 0.0001, 0.00001, 0.000001, 0.0000001, 0.00000001
Acetamiprid	0.5, 0.2, 0.08, 0.05, 0.032, 0.02, 0.0128, 0.008, 0.0032, 0.00128
Pyrethroids	
<i>tau</i> -fluvalinate	0.2, 0.1, 0.05, 0.025, 0.0125
Deltamethrin	0.04, 0.02, 0.01, 0.004, 0.002, 0.001, 0.0005, 0.0004, 0.00025, 0.0002, 0.0001, 0.00005, 0.000025
Organophosphates	
Coumaphos	100, 50, 25, 12.5, 6.25, 2.5, 1, 0.4, 0.16
Chlorpyrifos	0.4, 0.2, 0.1, 0.05, 0.025, 0.0125, 0.00625, 0.003125, 0.0015625

Based on the results from these tests the final ranges of concentrations used in the bioassays are shown in table 4.4.

Table 4.4: Concentration range of insecticides used in *M. rotundata* acute contact bioassay

Compound	Concentration $\mu\text{g}/\mu\text{l}$ ($\mu\text{g}/\text{bee}$)
Neonicotinoids	
Thiacloprid	0.05, 0.02, 0.008, 0.0032, 0.00128
Imidacloprid	0.01, 0.004, 0.0016, 0.00064, 0.000256
Acetamiprid	0.5, 0.2, 0.08, 0.032, 0.0128
Pyrethroids	
<i>tau</i> -fluvalinate	0.2, 0.1, 0.05, 0.025, 0.0125
Deltamethrin	0.02, 0.008, 0.0032, 0.00128, 0.000512
Organophosphates	
Coumaphos	2.0, 1.0, 0.5, 0.25, 0.125
Chlorpyrifos	0.05, 0.025, 0.0125, 0.00625, 0.003125

4.3.2 Acute contact toxicity bioassay of *M. rotundata*

4.3.2.1 Neonicotinoid insecticides

As can be seen in figure 4.6 and table 4.5 all three neonicotinoid insecticides tested are highly toxic to *M. rotundata* ($\text{LD}_{50} < 2 \mu\text{g}/\text{bee}$; see table 4.1). The acute contact LD_{50} (48 h) of imidacloprid ($0.001 \mu\text{g}/\text{bee}$) > thiacloprid ($0.013 \mu\text{g}/\text{bee}$) > acetamiprid ($0.179 \mu\text{g}/\text{bee}$), with an overall 179-fold difference between the highest and lowest values. The data and dose-response curves at each observation point are shown in appendix table 4.1 and appendix figure 4.1.

Table 4.5: Response of *M. rotundata* to topically applied neonicotinoid insecticides at 48 h.

Compound tested	LD_{50} 48h ($\mu\text{g ai}/\text{bee}$)	Lower CI 95%	Upper CI 95%	Hill slope	SE \pm
Thiacloprid	0.013	0.012	0.017	3.997	0.585
Imidacloprid	0.001	0.001	0.002	915.807*	206.532
Acetamiprid	0.179	0.115	0.351	1.523	0.219

95% confidence intervals calculated using a heterogeneity factor (the chi-square value divided by the df) as the p value for the Pearson goodness-of fit chi-square test was significant ($p < 0.150$). *result calculated without converting doses to logarithms.

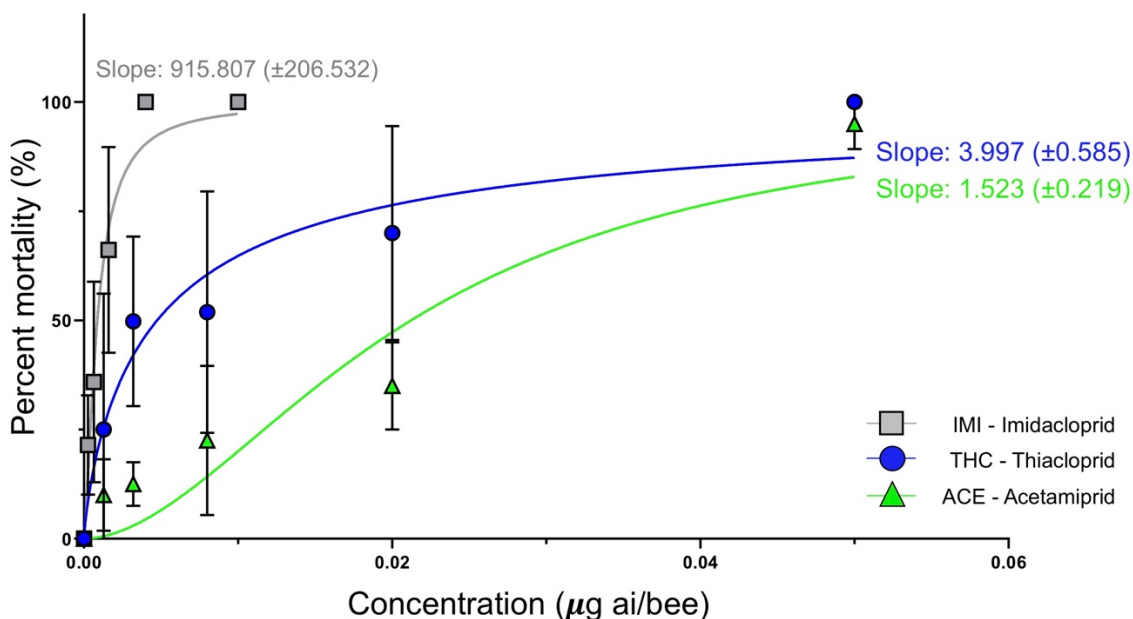


Figure 4.6: 48 h acute contact toxicity dose-response curves for one nitro-group imidacloprid (IMI) and two *N*-cyano-group neonicotinoid insecticides thiacloprid (THC) and acetamiprid (ACE) against *M. rotundata*. Slope values for the steepest part of the dose-response curve are shown for each compound.

4.3.2.2 Pyrethroid insecticides

As can be seen in figure 4.7 and table 4.6 both pyrethroid insecticides tested are highly toxic to *M. rotundata* ($LD_{50} < 2 \mu\text{g}/\text{bee}$; see table 4.1). The acute contact LD_{50} (48 h) of deltamethrin ($0.004 \mu\text{g}/\text{bee}$) $>$ *tau*-fluvalinate ($0.061 \mu\text{g}/\text{bee}$), with an overall 15.25-fold difference between the two compounds. The data and dose-response curves at each observation point are shown in appendix table 4.2 appendix figure 4.2.

Table 4.6: Response of *M. rotundata* to topically applied pyrethroid insecticides at 48 h.

Compound tested	LD_{50} 48h ($\mu\text{g ai}/\text{bee}$)	Lower CI 95%	Upper CI 95%	Hill slope	SE \pm
<i>tau</i> -fluvalinate	0.061	0.044	0.092	2.208	0.339
Deltamethrin	0.004	0.003	0.008	359.620*	2.9930

95% confidence intervals calculated using a heterogeneity factor (the chi-square value divided by the df) as the p value for the Pearson goodness-of fit chi-square test was significant ($p < 0.150$). *results calculated without converting doses to logarithms.

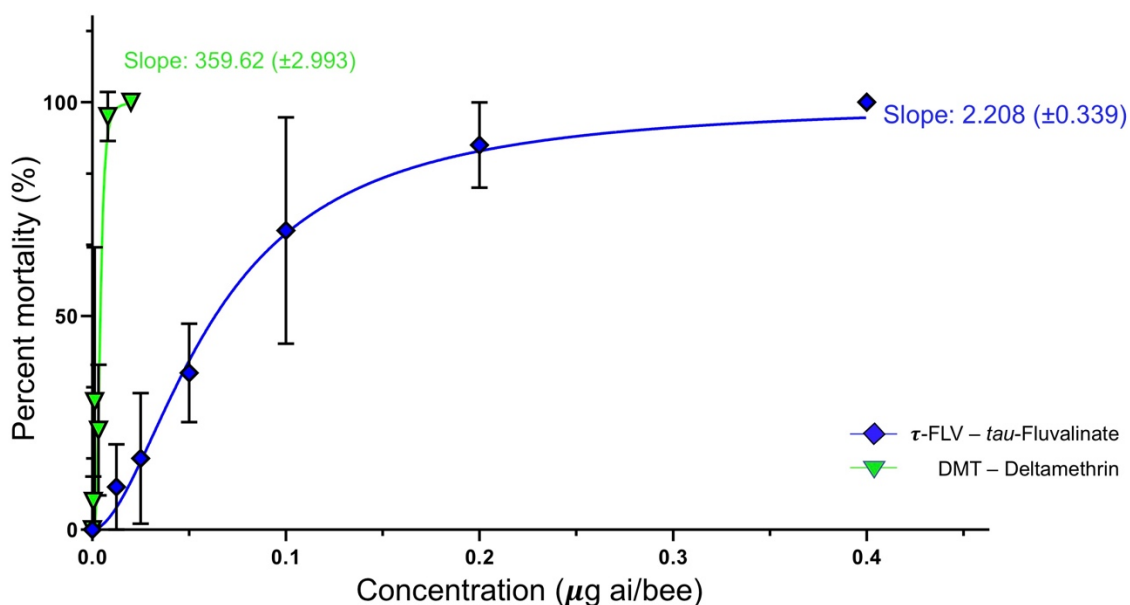


Figure 4.7: 48 h acute contact toxicity dose-response curves for two pyrethroid insecticides, deltamethrin (DMT) and *tau*-fluvalinate (τ -FLV) against *M. rotundata*. Slope values for the steepest part of the dose-response curve are shown for each compound.

4.3.2.3 Organophosphate insecticides

As seen in figure 4.8 and table 4.7 both organophosphate insecticides tested are highly toxic to *M. rotundata* ($LD_{50} < 2 \mu\text{g}/\text{bee}$; see table 4.1). The acute contact LD_{50} (48 h) of chlorpyrifos ($0.017 \mu\text{g}/\text{bee}$) > coumaphos ($0.557 \mu\text{g}/\text{bee}$), with an overall 32.76-fold difference between the two compounds. The data and dose-response curves at each observation point are shown in appendix table 4.3 and appendix figure 4.3.

Table 4.7: Response of *M. rotundata* to topically applied organophosphate insecticides at 48 h.

Compound tested	LD_{50} 48h ($\mu\text{g ai}/\text{bee}$)	Lower CI 95%	Upper CI 95%	Hill slope	SE \pm
Coumaphos	0.557	0.433	0.720	2.700	0.298
Chlorpyrifos	0.017	0.015	0.020	10.277	1.274

[95% confidence intervals calculated using a heterogeneity factor (the chi-square value divided by the df) as the p value for the Pearson goodness-of fit chi-square test was significant ($p < 0.150$)]

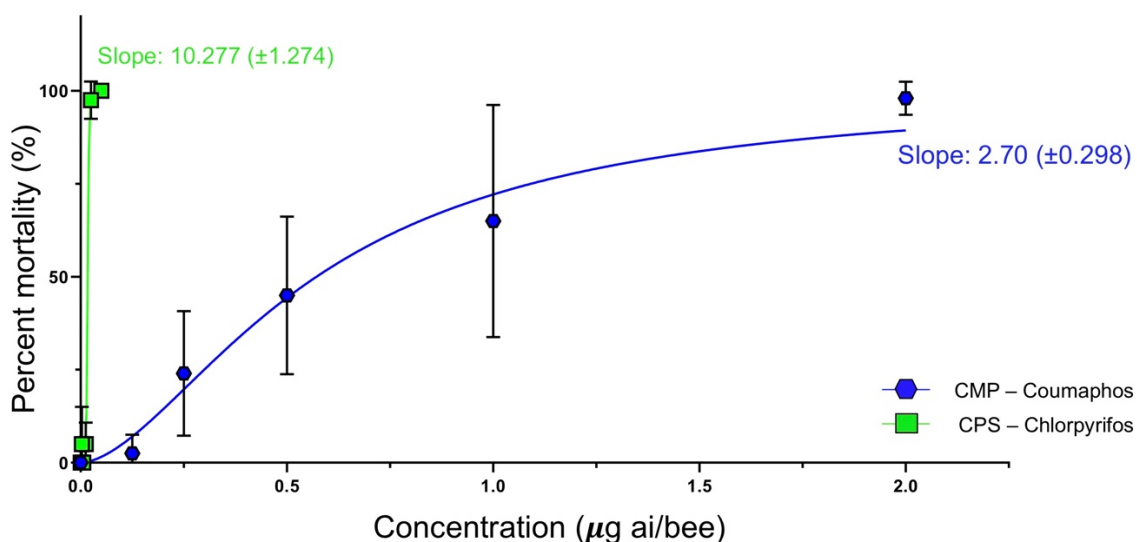


Figure 4.8: 48 h acute contact toxicity dose-response curves for two organophosphate insecticides coumaphos (CMP) and chlorpyrifos (CPS) against *M. rotundata*. Slope values for the steepest part of the dose-response curve are shown for each compound.

All the compounds tested across the three classes of insecticide were highly toxic when applied topically to *M. rotundata*. The species is most sensitive to the neonicotinoid imidacloprid (acute contact LD₅₀ (48 h) 0.001 µg a.i./bee) and least sensitive to the organophosphate coumaphos (acute contact LD₅₀ (48 h) 0.557 µg a.i./bee). The acute contact LD₅₀ (48 h) of imidacloprid < deltamethrin < thiacloprid < chlorpyrifos < tau-fluvalinate < acetamiprid < coumaphos, with an overall 557-fold difference across all the compounds.

4.3.3 Comparison of *M. rotundata* tolerance to insecticides with three other managed bee pollinator species

4.3.3.1 Neonicotinoid insecticides

The magnitude of difference between acute contact toxicity LD₅₀ (48 h) values for thiacloprid and imidacloprid is 479-816-fold in *A. mellifera*, >250-fold in *B. terrestris* and >2000-fold in *O. bicornis* (see figure 4.9) [137, 155, 226]. There is no such differential sensitivity in *M. rotundata*, where only a 13-fold difference between acute contact toxicity LD₅₀ (48 h) values for thiacloprid and imidacloprid was observed. *M. rotundata* is >2500-fold more sensitive than *A.*

meliifera and >6,500-fold more sensitive than *B. terrestris* or *O. bicornis* to thiacloprid (see figure 4.9) [168, 463]. *M. rotundata* also lacks the significant differential sensitivity between imidacloprid and acetamiprid found in the other bee species (see table 4.5 and figure 4.9).

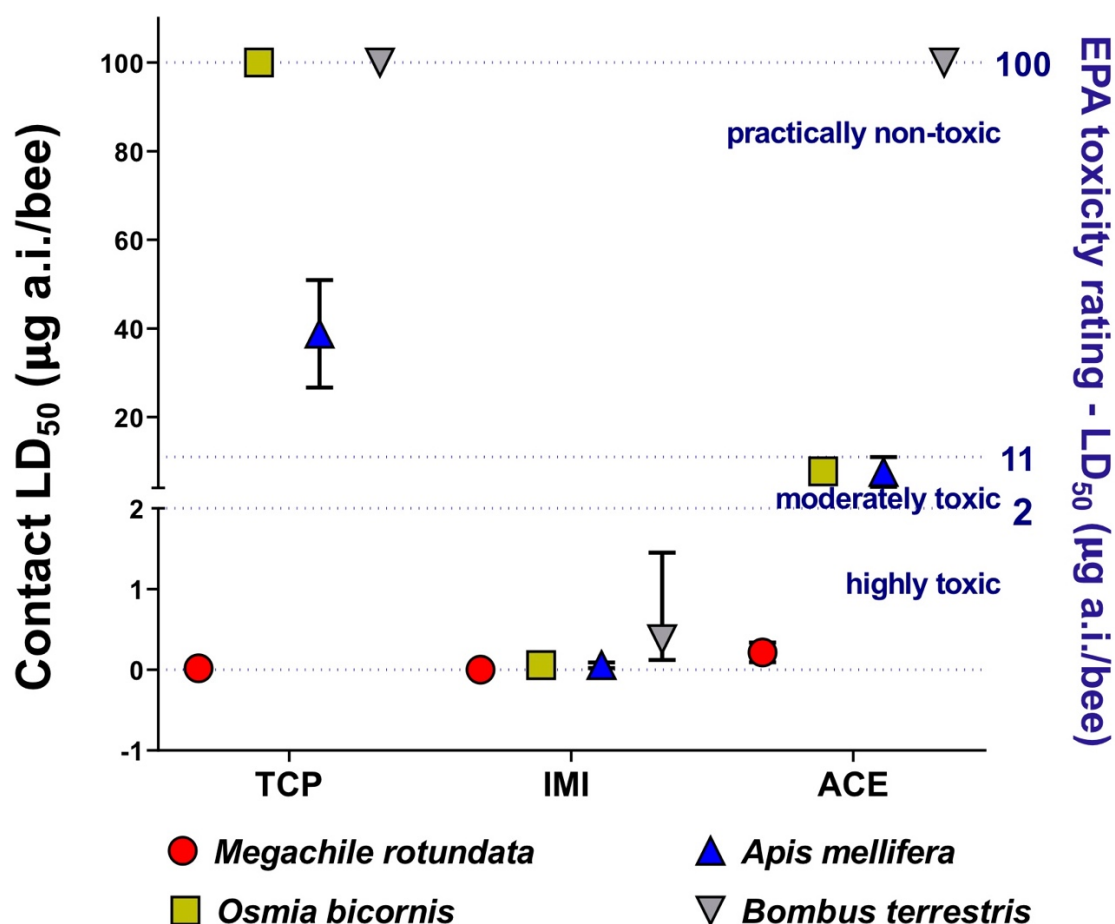


Figure 4.9: LD₅₀ values for *M. rotundata* (48 h) after exposure to topically applied thiacloprid (TCP), imidacloprid (IMI) and acetamiprid (ACE), compared to data for three other managed bee pollinators (*A. mellifera*, *B. terrestris* and *O. bicornis*); error bars indicate 95% CIs. Sensitivity thresholds are depicted according to the EPA toxicity ratings [493] (neonicotinoid data for *A. mellifera*, *B. terrestris* and *O. bicornis* taken from [137, 155, 168, 226, 446, 463]).

4.3.3.2 Pyrethroid insecticides

The magnitude of difference between acute contact toxicity LD₅₀ (48 h) values for tau-fluvalinate and deltamethrin is >150-fold in *A. mellifera*, >17-fold in *B. terrestris* and >15-fold in *O. bicornis* (see figure 4.10) [137, 145, 282, 463]. As

such, the 15.25-fold difference between the acute contact toxicity LD₅₀ (48 h) values for *M. rotundata* is similar to that shown by *B. terrestris* and *O. bicornis*. However, *tau*-fluvalinate is moderately toxic to *O. bicornis* (LD₅₀ (48 h) 3.811 µg a.i./bee) [463] and practically non-toxic to *B. terrestris* (LD₅₀ (48 h) 18.71 µg a.i./bee) [137], but is highly toxic to *M. rotundata* (LD₅₀ (48 h) 0.061 µg a.i./bee; see figure 4.10).

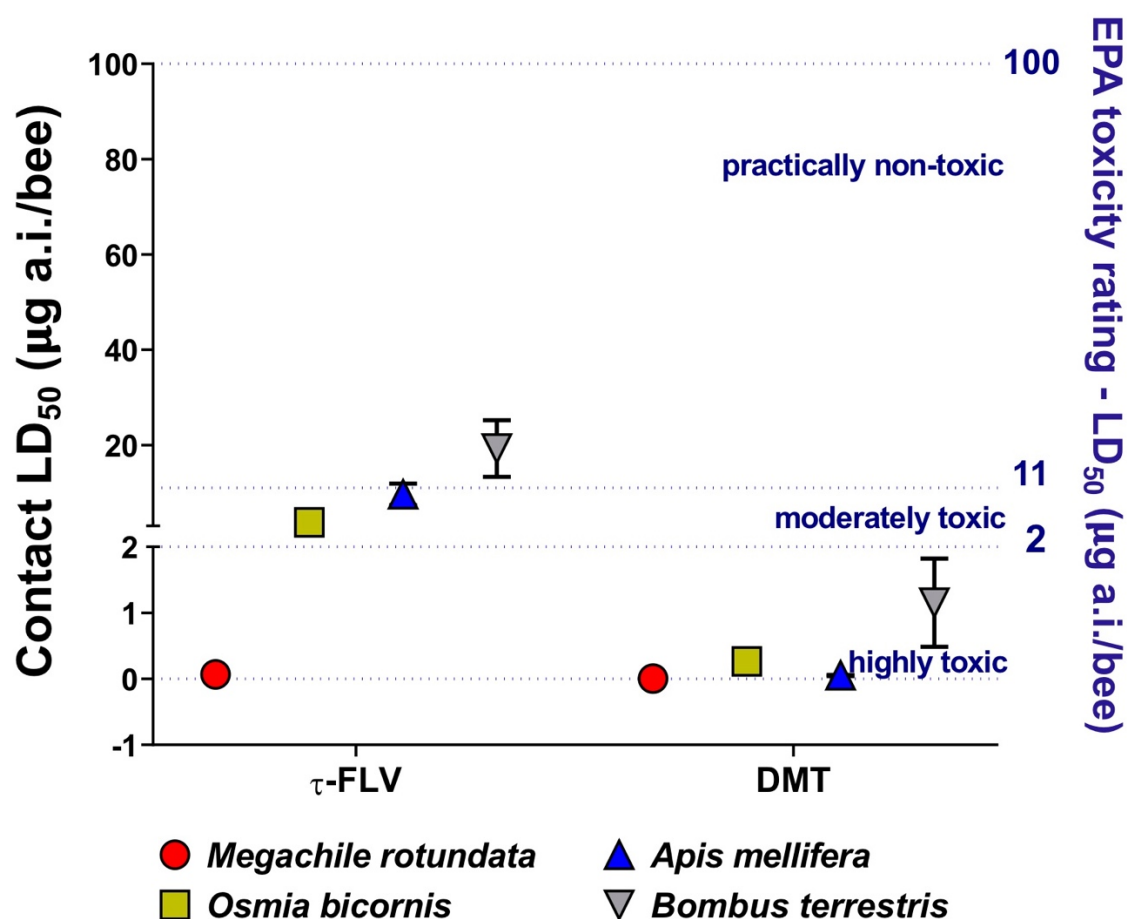


Figure 4.10: LD₅₀ values for *M. rotundata* (48 h) after exposure to topically applied *tau*-fluvalinate (τ-FLV) and deltamethrin (DMT), compared to data for three other managed bee pollinators (*A. mellifera*, *B. terrestris* and *O. bicornis*); error bars indicate 95% CIs. Sensitivity thresholds are depicted according to the EPA toxicity ratings [493] (pyrethroid data for *A. mellifera*, *B. terrestris* and *O. bicornis* taken from [137, 145, 282, 463]).

4.3.3.3 Organophosphate insecticides

The magnitude of difference between acute contact toxicity LD₅₀ (48 h) values for coumaphos and chlorpyrifos is >250-fold in *A. mellifera*, >150-fold in *B.*

terrestris and >1000-fold in *O. bicornis*, compared to the ~32-fold difference found in *M. rotundata* (see figure 4.11) [29, 137, 138, 463]. Coumaphos is also considered practically non-toxic to the other bee species, with an LD₅₀ (48 h) of >100 µg a.i./bee in *B. terrestris* and *O. bicornis* (see figure 4.11) [137, 463].

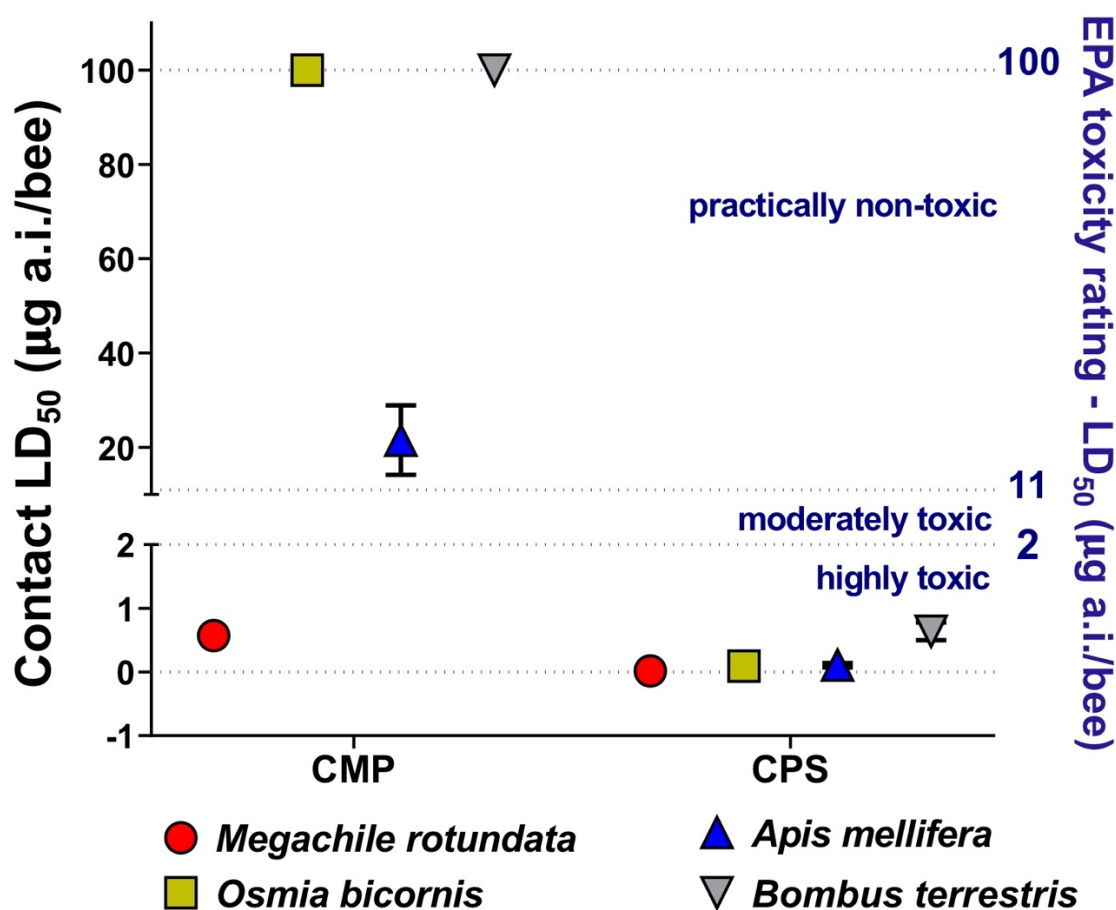


Figure 4.11: LD₅₀ values for *M. rotundata* (48 h) after exposure to topically applied coumaphos (CMP) and chlorpyrifos (CPS), compared to data for three other managed bee pollinators (*A. mellifera*, *B. terrestris* and *O. bicornis*); error bars indicate 95% CIs. Sensitivity thresholds are depicted according to the EPA toxicity ratings [493] (organophosphate data for *A. mellifera*, *B. terrestris* and *O. bicornis* taken from [29, 137, 138, 463]).

Table 4.8: Difference in LD₅₀ values between the four managed bee pollinators of select insecticides.

Insecticide	Fold difference in sensitivity compared to <i>M. rotundata</i>		
	<i>A. mellifera</i>	<i>B. terrestris</i>	<i>O. bicornis</i>
Thiacloprid	2500	6500	6500
<i>tau</i> -Fluvalinate	132	306	62
Coumaphos	35	180	180

4.4 Discussion

Despite the fact that *M. rotundata* is a managed pollinator of great economic importance, there is a lack of reliable data concerning the intrinsic tolerance/sensitivity to insecticides in the species, including those compounds licenced for use on the crops it pollinates. The aim of this chapter was to provide toxicological data that could be compared to that from other managed bee pollinators. To that end, the acute contact LD₅₀ data presented here, clearly indicate that *M. rotundata* exhibits a low intrinsic tolerance for insecticides across three distinct MoA classes: neonicotinoids, pyrethroids and organophosphates. All seven insecticides tested in this PhD are highly toxic to the species, with LD₅₀ values of <2 µg a.i./ bee (see section 4.3.2). Importantly, this includes four insecticides, that based on tests with *A. mellifera*, are considered 'bee-friendly' compounds. Indeed, coumaphos and *tau*-fluvalinate are routinely used as in-hive miticidal treatments for *V. destructor* [496, 497]. Thiocloprid and coumaphos are rated practically non-toxic; LD₅₀ >11 µg a.i./ bee), and acetamiprid and *tau*-fluvalinate as moderately toxic (LD₅₀ 2.0-10.9 µg a.i./ bee) to *A. mellifera*. Furthermore, the marked difference in sensitivity to *N*-nitroguanidine and *N*-cyanoamidine neonicotinoid insecticides observed in the other managed bee pollinators (250-2000-fold) is not seen in *M. rotundata* (see section 4.3.3.1). This lack of differential sensitivity in *M. rotundata* extends to the other two MoA classes tested in this PhD (see sections 4.3.3.2 and 4.3.3.3). Further corroborating my findings, results from collaborators at Bayer (AG, Crop Science Division, Leverkusen, Germany) show that a higher level of sensitivity extends to the butenolide insecticide, flupyradifurone (FPF) [447]. This compound is rated as highly toxic to *M. rotundata* (LD₅₀ (72 h) of 0.092 µg a.i./ bee) [447], but as practically non-toxic to the other managed bee pollinators [447, 463, 498].

The disparity between *M. rotundata* and the other managed bee pollinators (see table 4.8), is too great to be explained by body size alone. This is particularly true in the case of thiocloprid, where the difference between the LD₅₀ values in *M. rotundata* is >2500-fold for *A. mellifera* and >6500-fold for *B. terrestris* and *O. bicornis*, whereas the differential in body weight is 4-fold in *A. mellifera* and

O. bicornis (assuming an average bodyweight ratio of 35:130 mg) and 17-fold in *B. terrestris* (assuming an average bodyweight ratio of 35:600 mg).

Currently, the insecticide risk assessment for bees only requires acute contact and acute oral LD₅₀ values for *A. mellifera* to be established [472, 499]. Given the LD₅₀ values generated here, even without further information regarding the mechanisms that underpin them, the risk to *M. rotundata*, of relying on *A. mellifera* as a proxy in toxicological testing, is obvious. The assessment or bridging factor of 10, suggested in previous research, [476-478] would not reduce the risk to *M. rotundata* to a safe level. For example, the LD₅₀ for thiacloprid is 37.8 µg a.i./bee in *A. mellifera* [168], so an assessment factor would indicate that ~4 µg a.i./bee would be acceptable in *M. rotundata*. The LD₅₀ found here, for thiacloprid in *M. rotundata*, is 0.013 µg a.i./bee a figure that is considerably lower (>250-fold).

As mentioned in section 4.1, extrapolating data between vertebrate species, based on body weight, is widely used in toxicology [479]. The underlying principle of allometric scaling, with the use of an assessment factor, is based on the $\frac{3}{4}$ power scaling relationship between body mass and metabolic rate observed in animals [480]. However, the metabolism of drugs and xenobiotics is not always found to be comparable between species, and so, it might be prudent to exercise caution when applying this form of scaling [479]. For example, it was originally thought that the $\frac{3}{4}$ power scaling extended to plants [480], but more recently, this has been shown not to be the case [500]. Likewise, there is some argument over whether the $\frac{3}{4}$ power scaling relationship holds across insect species [501-503]. Furthermore, it is possible that differences may occur in the internal environment of individual species [479]. Where that difference occurs in the CYPome of a species, such as the lack of a *CYP9Q/BU* ortholog in *M. rotundata*, there is potential for a corresponding change in substrate specificity.

The fact that all seven insecticides tested here are highly toxic to *M. rotundata* raises the question: What is the root cause of the comparatively high level of insecticide sensitivity in this species? In *A. mellifera*, *B. terrestris* and *O. bicornis*, the *CYP9Q* and *CYP9BU* lineages of P450s have been demonstrated

to be involved in the metabolism of *N*-cyanoamidine neonicotinoids, *tau*-fluvalinate and coumaphos [137, 168, 226]. These CYP9s have also been implicated in the detoxification of flupyradifurone in *A. mellifera* and *O. bicornis* [463, 498]. The results presented in the previous chapter, clearly indicate that *M. rotundata* lacks a *CYP9Q/BU* ortholog. The best candidate genes in the species, the *CYP9DMs*, are phylogenetically and structurally distinct from sequences in the insecticide-degrading *CYP9Q/BU* lineages. This lack of a *CYP9Q/BU* ortholog alone raises some doubt over the capacity of the species to metabolise the insecticides linked with that enzyme lineage (i.e. thiacloprid, acetamiprid, *tau*-fluvalinate, coumaphos and flupyradifurone) [137, 168, 226, 498].

However, although it is tempting to assume that the lack of a *CYP9Q/BU* ortholog is the explanation for the high level of sensitivity in the species, there are other plausible reasons that cannot be discounted at this point. There could be some form of toxicodynamic mechanism or physiological adaptation (see sections 1.6.2 and 1.6.3), for example, an increased affinity of a compound at its receptor site, or an increased rate of penetration through the cuticle.

In the other bee pollinators, evidence for the involvement of P450s in the metabolism of individual insecticides was provided from the results of synergism bioassay studies, using the P450 inhibitor PBO [137, 168, 226]. Unfortunately, due to the high level of sensitivity shown by *M. rotundata* to the insecticides used in this PhD, bioassays using a synergist would not yield useful data. The role for P450 enzymes in detoxification has been investigated, in other bee species, by assessing the capacity of native microsomes (a source of total P450 enzymes extracted from the endoplasmic reticulum) to metabolise insecticides *in vitro*. Heterologous expression of candidate P450s, identified using phylogenetic analyses, has allowed for the functional characterisation of specific enzymes in bees. For example, *CYP9Q3* in *A. mellifera* and *CYP9BU1* in *O. bicornis* [168, 226]. Both of these techniques were used to investigate the metabolic profile of P450s in *M. rotundata* (see chapter 5).

To study the rates of cuticular penetration, radiolabelled ligands of neonicotinoid and butenolide insecticides ($[^{14}\text{C}]$ imidacloprid; $[^{14}\text{C}]$ thiacloprid; $[^{14}\text{C}]$

acetamiprid and [¹⁴C] flupyradifurone) have been used successfully in other bee species [226, 463]. The protocol calls for individual bees to be dosed with 5000 ppm (equivalent to 5 µg a.i./ bee) of radiolabelled insecticide. The radiolabelled insecticide is then washed off experimental groups of five bees at 0, 2, 4 and 24 h [226, 463]. The high level of sensitivity to neonicotinoids and flupyradifurone, observed in *M. rotundata*, mean a dose of these compounds at 5000 ppm, even over 24 h, would be expected to be lethal in many individual bees. Altering the dose, to one more appropriate for *M. rotundata*, runs the risk that the resultant level of radiolabelled insecticide is too weak for reliable measurement. This effectively rules this protocol out for use in *M. rotundata*.

However, when it comes to toxicodynamic mechanisms, in particular the study of binding affinity at the nicotinic acetylcholine receptor site, there is an established *in vitro* protocol. This methodology measures the displacement of tritiated imidacloprid ([³H] imidacloprid), by unlabelled neonicotinoid or butenolide insecticide, at the nicotinic acetylcholine receptor using head membrane preparations [168, 226, 446]. This protocol is used to establish an inhibitory concentration (IC₅₀) value and, as it is *in vitro* rather than *in vivo*, could be applied to investigate the binding affinity of these insecticides in *M. rotundata*.

Acute oral bioassays were not undertaken in this PhD due to difficulties with solitary bees feeding under test conditions (see section 4.1). Further work on establishing a standard protocol for assessing acute LD₅₀ (oral) in solitary bees, similar to OECD tests No 213 [483] and No 247 [485] for *A. mellifera* and *B. terrestris*, is needed. Furthermore, this chapter has only established toxicological results for adult bees and has not generated data on the sensitivity of *M. rotundata* larvae. There is undoubtedly a potential risk from chronic exposure to certain insecticides in the provision mass to the larvae of this species. There is an established OECD protocol for testing *A. mellifera* larvae (OECD Test No. 237) [504], but no such methodology exists for solitary bees as yet.

It must also be remembered that acute toxicity tests form the basis of the first of three tiers in process to register an insecticide safe for use (see section 4.1). As

such, they do not relate to a realistic field-use level or to the residual concentrations found in field margins, concentrations that are often reported in parts-per-billion (ppb) [146, 158, 331, 505]. In their 2010 paper, Mullin *et al.*, convert standard LD₅₀ values reported in µg/bee to ppb using the formula $1000 \div \text{average bee weight g}$ [331]. For *A. mellifera* this amounts to a multiple of 10,000 [331], whereas, for *M. rotundata* this multiple would be in the region of 28,000 (assuming an average body weight of 0.035 g). Applying this calculation to the LD₅₀ values generated in this chapter gives us lethal contact concentrations that range from 28 ppb for imidacloprid to 15,596 ppb for coumaphos.

In their 2016 study, David *et al.*, report levels of residues from five different neonicotinoids in pollen in the U.K. These range from 3.6-19 ppb in oil seed rape, and from 0.13-2.8 in wildflower pollen, where ppb = ng/g wet weight sample [505]. Of the five neonicotinoids, thiacloprid was recorded at the highest residual level in the oil seed rape crop (19 ppb). Using the conversion formula above, that value corresponds to 0.0019 µg/bee for *A. mellifera* and 0.0007 µg/bee for *M. rotundata*. In both species this level of thiacloprid would be below their acute toxicity LD₅₀ value and it seems unlikely there is a high risk of direct mortality [158].

There has been a great deal of interest and research into the chronic exposure to insecticides of *A. mellifera* and *B. terrestris* [158, 294, 506-511]. There is strong evidence for the sub-lethal effects of chronic exposure in these species. Exposure has been linked to a decreased immune response and an increase in the susceptibility to other stressors, such as the *V. destructor* or pathogens [512, 513]. Chronic exposure is also thought to impair fertility, development, memory and learning in both *A. mellifera* and *B. terrestris* [506, 510, 511, 514]. A sub-lethal injury to cognitive ability in any bee species has the potential to impact negatively on foraging and homing ability [158, 506]. Similar sub-lethal effects in solitary bees, particularly a decreased ability to forage or find a nesting site, could have a very deleterious impact on their survival.

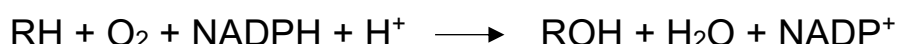
In conclusion, this data establishes that *M. rotundata* is very sensitive to insecticides across three MoA classes. The underlying mechanism of this

sensitivity is not certain, although the lack of a *CYP9Q/BU* ortholog is of prime interest and this was further investigated in chapter five. Overall, the most important take home message from these data is, that the use of *A. mellifera* as a proxy for *M. rotundata* in ecotoxicology risk assessments is unreliable.

Chapter five: Functional characterisation of *M. rotundata* P450s

5.1 Introduction

P450 enzymes are generally thought of as monooxygenases, although they also perform other catalytic functions, e.g. reductases and desaturases [515, 516]. The simple stoichiometry that describes the monooxygenase reaction is shown below:



where R is the substrate. To obtain the electrons required for catalysis from nicotinamide adenine dinucleotide phosphate (NADPH), P450s rely on a redox-partner, cytochrome P450 reductase (CPR) and/or, in some cases, cytochrome *b*₅ [515, 517]. In insects P450s have been shown to play a role in the detoxification of plant secondary metabolites and synthetic insecticides [168, 209], and have been implicated in insecticide resistance [204, 218]. Indeed, the ability of herbivorous insects to metabolise a wide range of phytochemicals is considered a good example of the co-evolutionary theory proposed by Erlich and Raven [518, 519]. However, there is a paucity of data on the substrate specificity of most insect P450s, making predictions about their metabolic abilities difficult [520, 521].

There are two main modes of transcription of genes coding for detoxification enzymes [204, 522]. The first is constitutive expression, where genes are continuously transcribed, and the second is the induction of gene expression following substrate exposure [202, 206]. The induction of specific gene transcription is thought to be associated with lower overall metabolic costs, when compared to the constitutive expression of enzymes [518]. Insect P450s are known to be induced by phytochemicals in several species. For example, *CYP6B1* and *CYP6B3* are induced in the black swallowtail butterfly (*Papilio polyxenes*) in response to exposure to xanthotoxin, a phytotoxin found in the plants they use as a food source [523, 524]. In *A. mellifera* pollen extracts, such as the flavonoid quercetin and *p*-coumaric acid, have been shown to upregulate the transcription of *CYP6AS* and *CYP9Q* genes [284, 286]. This form of regulation allows insects to directly respond to a change in their environment by generating an appropriate detoxification defence [218, 518].

However, only a third of the *D. melanogaster* CYPome was found to be inducible by xenobiotics [525]. In comparison to the strong induction response of P450 transcription levels to naturally occurring xenobiotics in *D. melanogaster*, the response to synthetic insecticides is often more minimal, and is absent for several compounds such as the neonicotinoid nitenpyram and the organophosphate diazinon [521, 525]. Thus, it appears likely that, in insects, constitutive expression of P450s is the key transcriptional regulatory strategy when it comes to detoxifying synthetic insecticides [204, 521]. Indeed, the constitutive overexpression of P450s belonging to the CYP3, CYP4 and mitochondrial clans has frequently been implicated in insecticide resistance [204]. Relative to the levels of overexpression linked to resistance, genes from the *CYP9Q* lineage are constitutively expressed at low levels in *A. mellifera* [284]. However, they are also transcriptionally induced in response to exposure to certain insecticides and honey extracts [284]. It is likely that the orchestration of the *CYP9Q* activity is a complex interaction of transcription factors, coregulators and signalling pathways similar to those reported for other P450s and, as such, is not yet fully understood [526]. Nevertheless, the level of their constitutive expression suggests that the ability of these P450s to metabolise insecticides from three MOA classes does not result from insecticide selection pressure, but is more likely due to coincidental structural similarity to natural substrates encountered by *A. mellifera* [284].

5.1.1 Metabolism assays

Measuring the activity of P450s is challenging because the CYPome contains a large number of enzymes, each catalysing a specific range of substrates [218]. Individual P450s can also have a broad range of substrates and some may be particularly promiscuous in terms of their substrate profile [168, 225]. Enzyme-catalysed reactions are often described in terms of Michaelis-Menten kinetics, where reaction rate (formation of product) is shown to vary in relation to increasing substrate concentration [527, 528]. A characteristic hyperbolic curve is obtained from this type of reaction by plotting level of product formed against substrate concentration. From this it is possible to determine the Michaelis constant K_M , which describes the substrate concentration required for an enzyme to function at half its maximum rate (V_{max}) [527]. The K_M value reflects

the specificity and binding affinity of an enzyme, a small value indicating strong substrate binding and higher specificity [529].

There are several established methods for investigating enzyme metabolism, including kinetic metabolism assays, liquid chromatography-mass spectroscopy (LC-MS) and fluorometric assays using model substrates [168, 463]. Tandem LC-MS/MS (using two mass spectrometry detectors) and fluorometric assays were used in this chapter to functionally characterise *M. rotundata* native microsomes and recombinant P450 proteins.

5.1.2 Microsomes

Native microsomes (microsomal P450 preparations) are fragments of endoplasmic reticulum and attached ribosomes and, as such, contain an undetermined number and quantity of P450 and non-P450 enzymes [218, 530]. Microsomal preparations have been used in the *in vitro* study of xenobiotics, in vertebrates and insects since the 1960s [531-533]. Metabolism assays using native microsomes, extracted from insect homogenates, is a useful first step in understanding the overall functional characteristics of the CYPome of a species [218]. It also captures instances where P450s act together in concert to boost detoxification, something that is not measured by studying recombinant P450s individually [534]. However, any measurement of microsomal metabolism of substance 'x' into substance 'y', must always be interpreted as the sum of all P450s capable of that reaction and of any non-P450 enzymes that may be involved [218].

There are well established protocols for the preparation of insect microsomes using homogenisation and differential centrifugation [530]. However, native microsomes isolated from *A. mellifera* tissue have resulted in a preparation that exhibits limited P450 activity due to the presence of an inhibitory protein found in the mid-gut [535, 536]. More recently, it has been discovered that after the removal of the venom gland sting complex *A. mellifera* native microsomes become highly active [534]. Bees are part of the infraorder Aculeata, a monophyletic group, that have a characteristic feature; a sting [537]. The sting is a modified ovipositor, used exclusively for defence, and as such is only present in females [538]. The presence of phospholipase A₂, a powerful P450

inhibitor, in the venom sac of *A. mellifera* appears to account for the isolation of non-functional microsomes from intact workers [534]. However, microsomes prepared from both female and male *O. bicornis* show good P450 activity, and so it is not clear that the presence of an inhibitory protein is ubiquitous in all bee species [226]. Prior to their use in metabolism assays, the activity of both male and female *M. rotundata* native microsomes were screened using model substrates. A high level of sensitivity to insecticides was observed in *M. rotundata* in acute contact bioassays performed in this PhD (see section 4.3.2). Native microsomes are enriched subcellular extracts that contain multiple P450 and non-P450 enzymes, as such, metabolism assays involving *M. rotundata* microsomes will allow an understanding of whether this sensitivity is correlated to a lack of enzymatic ability.

5.1.3 Heterologous expression of *M. rotundata* P450s in insect cell lines

A number of different methods have been employed to express P450 proteins, including prokaryotic (e.g. *E. coli*) and eukaryotic (e.g. yeast and insect cells) systems. This chapter details the functional expression of *M. rotundata* P450s using Gateway® cloning technology (Invitrogen) and BaculoDirect™ baculovirus expression system (Invitrogen). The use of Gateway® cloning technology enables the transfer of DNA sequences derived from purified PCR products into multiple vector systems, including baculovirus [539]. The transfection of recombinant baculovirus vector into insect cell lines, has been established as a successful protocol in the expression of insect P450s, including those from bee species [137, 168, 226, 540]. This system yields a very good level of protein expression because of the correct eukaryotic modes of post-transcription and post-translation in the cell lines [541]. To maximise the yield of secreted P450s, the presence of the redox-partner NADPH cytochrome P450 reductase (CPR) is essential. Thus, co-infection of cell lines with baculovirus containing P450 sequences and baculovirus containing a CPR sequence is necessary. There is a single, well conserved CPR gene in insects [218]. In this project, co-infection of insect cell lines was achieved using a baculovirus containing house fly (*Musca domestica*) CPR, due to its previous success in the expression of insect P450s [540].

In eukaryotes all cytochrome P450s are evolved from a common ancestral CYP51 sequence [236]. Sequence identities between P450s can be low (10-30%), however, the topology and tertiary structure is well conserved, particularly the heme-binding core structure [542]. This gives P450s their characteristic spectral absorption peak (450 nm). A CO-difference spectral assay can determine the concentration of a recombinant P450 and also to check its' integrity, as loss of enzyme activity has been associated with the loss of a peak at 450 nm and gain of one at 420 nm [220].

5.1.3.1 Selection of candidate *M. rotundata* P450s for heterologous expression

The substrate profile of most P450 enzymes in the *A. mellifera* CYPome have not, as yet, been fully characterised [206]. However, the *CYP9Q/BU* lineage of P450s, known to be involved in the detoxification of synthetic insecticides, has been well studied in *A. mellifera*, *B. terrestris* and *O. bicornis* [137, 168, 463]. To add to this knowledge base, the focus in the selection of *M. rotundata* P450s in this PhD was the CYP9 subfamily. Phylogenetic and syntenic analyses (see sections 3.3.3 and 3.3.3) were used to inform the selection (see figure 5.1 (a) and (b)).

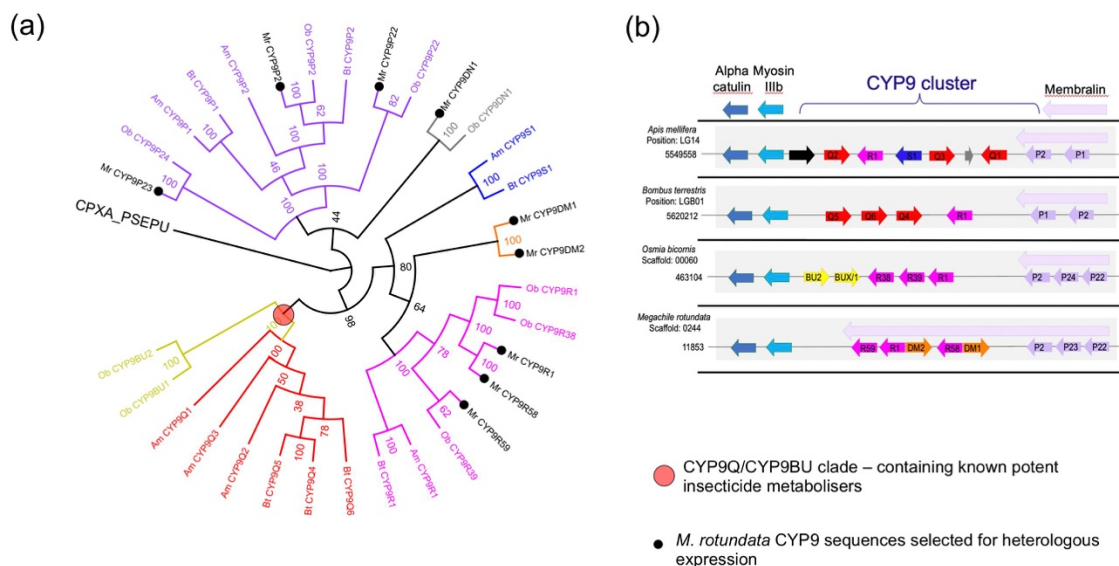


Figure 5.1: Selection of *M. rotundata* P450s for heterologous expression.

(a) Phylogeny of the CYP9 subfamily in four species managed bee pollinators. Phylogeny estimated using PhyML Maximum likelihood algorithm [453] using substitution model LG+G [454], with branch support of 50 bootstraps shown as %. (b) Synteny of the CYP9 cluster in four species of managed bee pollinators.

M. rotundata does not have a *CYP9Q/BU* ortholog (see figure 5.1 (a)), although there is good synteny of the CYP9 cluster across the managed bee pollinators (see figure 5.1 (b)). *CYP9DM1* and *CYP9DM2* in *M. rotundata* are present in the CYP9 cluster in a similar position to that seen in the *CYP9Q/BU* genes of other species (see figure 5.1 (b)). All nine members of the *M. rotundata* CYP9 subfamily were selected for heterologous expression, with *CYP9DM1* and *CYP9DM2* marked as of particular interest.

5.1.4 Assays with analysis using tandem LC-MS/MS

Tandem LC-MS/MS combines the separating power of liquid chromatography with the detection power of mass spectrometry. Mass spectrometry allows for the separation and identification of charged molecules [543]. LC-MS/MS has been used to separate and identify the hydroxylated metabolites of synthetic insecticides [168, 544]. It allows for identification and quantification of both parent compound, and known hydroxylated metabolites, after standardised incubation of native microsomes or recombinant P450 proteins with natural or synthetic xenobiotic compounds [168]. All the insecticides used in the acute contact bioassays in this project, were also used in metabolism assays with LC-MS/MS analysis. However, further insecticides and naturally occurring plant allelochemicals (alkaloids) were also used to gather a more complete functional profile of the CYPome.

5.1.5 Model substrate profiling

Model substrate profiling of native microsomes or recombinant P450s relies on the metabolism of a low fluorescent substrate to a high fluorescent product that can be measured in a spectrofluorometer [545-547]. Coumarin- and resorufin-derived compounds are the most widely used fluorophores for preparing fluorogenic model substrates [546, 547]. Fluorometric assays have been used successfully in the functional characterisation of insect microsomes and recombinant P450s, including bee species [168, 247, 548]. The standard protocol for fluorometric assays using model substrates to profile metabolic ability is based on that described by Boulenc *et al.*, in 1992 [546].

More recently, a fluorescence-based high-throughput assay using recombinant P450s has been developed to predict the interaction of model substrates with pesticidal compounds, such as neonicotinoid insecticides [549]. A modified protocol, based on the methodology published by Haas and Nauen in 2021 [549] is used in this study to investigate neonicotinoid mediated inhibition of model substrate metabolism by a *M. rotundata* recombinant CYP9DM enzyme.

5.1.6 Binding affinity of insecticides targeting the nicotinic acetylcholine receptor (nAChR)

As discussed in section 4.4, the high level of sensitivity to neonicotinoid and butenolide insecticides, observed in *M. rotundata*, could be due to the toxicodynamic mechanism of increased affinity at the nAChR site, rather than a lack of metabolic activity. Binding affinity at the nAChR is established by measuring the displacement of tritiated imidacloprid ($[^3\text{H}]$ imidacloprid) by unlabelled imidacloprid, thiacloprid or flupyradifurone, using head membrane preparations [168, 171, 550]. This work was done by colleagues at Bayer (AG, Crop Science, Leverkusen, Germany) in 2019 [447]. As part of my CASE placement with in March 2020, I was due to repeat the work, so I could present the binding affinity study fully in this thesis. However, due to the pandemic I was unable to travel to Germany. I will therefore include the results from the 2019 work in brief in this chapter.

5.1.7 Chapter aims and underpinning questions

In comparison to many species of pest insects, the isolation of functional microsomes from *A. mellifera* is not straightforward, due to the presence of a powerful P450 inhibitory factor in the venom sac. To ensure native microsomes extracted from *M. rotundata* were functional, prior to their use in metabolism assays, the presence/absence of a P450 inhibitory factor in the venom of the species was determined.

The main aim of this chapter, however, is to determine whether there is a correlation between the increased sensitivity to insecticides, observed *in vivo* in *M. rotundata* in chapter four, and a lack of *in vitro* metabolic activity. To this end the catalytic activity of native *M. rotundata* microsomes and functionally expressed candidate P450s is assessed, using tandem LC-MS/MS and model

substrate fluorescence assays. This methodology will allow the intrinsic metabolism of the insecticides used in the acute contact bioassays to be assessed. Radioligand binding studies carried out on *M. rotundata* head membrane preparations in 2019 by colleagues at Bayer will also inform on binding affinity at the nAChR [447].

As mentioned in section 1.5, foraging bees are exposed to a diverse range of plant defence chemicals [186]. To establish a fuller metabolic profile of *M. rotundata*, the ability of native microsomes to metabolise naturally occurring plant allelochemicals was investigated with four alkaloids (see section 1.5), using tandem LC-MS/MS analyses.

In chapter three, the curation of the *M. rotundata* CYPome highlighted the lack of a *CYP9Q/9BU* ortholog. In this chapter, through the first functional characterisation of P450s in *M. rotundata*, the metabolic profile of enzymes from the CYP9 subfamily in this species is examined using fluorometric assays.

The data from this chapter will be used to assess the veracity of using *in vitro* studies, with expressed P450s and/or native microsomes, alongside phylogenetics, as part of a pipe-line in a ‘tool-kit’ for future screening of insecticides.

The results from sections 5.3.3.1, 5.3.3.4 and 5.3.5 of this chapter are published in Hayward *et al.*, 2019 *Nature Ecology and Evolution* 3(11):1521-1524.

5.2 Methods

5.2.1 Preparation of *M. rotundata* microsomes

Microsomes were prepared following a standard protocol of homogenisation and differential centrifugation [530]. Native microsomes were prepared for both male and female *M. rotundata*. Approximately 60 snap frozen adult bees were homogenised, over ice, in homogenisation buffer (see appendix table 2.2) using a mortar and pestle. Cells were further broken down using a tissue grinder (Polytron PT 1600E) and a sonifier (Branson digital sonifier 250) set at 50% amplitude, pulsing on for two seconds, off for two seconds, for one minute. The

homogenate was centrifuged at 4000 g for 10 minutes at 4°C to pellet the exoskeleton (Eppendorf centrifuge 5810R). The supernatant was removed and centrifuged at 15,700 g for 15 minutes (Beckman Coulter Optima XPN-100 Ultracentrifuge, Rotor type Ti 32) to sediment larger subcellular particles [530]. The supernatant was removed and centrifuged at 100,000 g for 1 hour (Beckman Coulter Optima XPN-100 Ultracentrifuge, Rotor type Ti 32). This supernatant was discarded and the pellet resuspended in buffer R (see appendix table 2.3) using a glass homogeniser. The protein concentration was determined using a Bradford's protein assay, as described in section 2.3.1 [551]. Microsome preparations were stored at -80°C.

5.2.1.1 Screening male and female *M. rotundata* microsomes for activity against model substrates

Intact male and female *M. rotundata* were used to isolate native microsomes as described in section 5.2.1 above. As the venom gland sting complex is found in the abdomen of female bees, microsomes from dissected heads and thoraxes of females were also isolated [552]. Dissections of frozen female bees were performed on dry ice immediately prior to homogenisation and use in native microsome extraction. All three microsomal preparations were screened against 7-methoxy coumarin (MC) and 7-ethoxy-4-(trifluoromethyl)-coumarin (EFC) following the protocol outlined in section 5.2.4.3 below. Three replicates were carried out for each data point.

5.2.2 Functional expression of *M. rotundata* P450s in insect cell lines

Gateway® cloning technology (Invitrogen) was used to create *M. rotundata* recombinant P450s. A schematic of the Gateway® BP and LR recombination reactions are shown in appendix figures 5.1. BP and LR reactions were performed in a sterile laminar flow hood (Astec Microflow Peroxide Class II).

5.2.2.1 Synthesis of Candidate Gateway® P450 Entry Clones

Due to a high degree of similarity with other CYP9 subfamily genes, two of the nine genes from the CYP9 subfamily in *M. rotundata* were not suitable as candidates for the creation of Gateway® Entry Clones from PCR product. These two genes, *CYP9R58* and *CYP9DM1*, were synthesised *in vitro* by GeneArt (Life Technologies). The sequences were codon optimised for *S.*

frugiperda and delivered in pDONR221 plasmids (see appendix figure 5.2) with flanking Gateway® *attL1* (vector-N₇₅ CCA ACT TTG TAC AAA AAA GCA GGC TTC-insert) and *attL2* sites (insert-TGG GTC GAA AGA ACA TGT TTC AAC C N₇₅- vector) and kanamycin resistance. On arrival the quantity and quality of the entry clone was assessed using a NanoDrop One (Thermo Scientific) or NanoDrop 2000 spectrophotometer (Thermo Scientific). The clones were then stored at -20°C until needed.

5.2.2.2 Creating Gateway® Entry Clones from PCR product

5.2.2.2.1 PCR amplification of Gateway® plasmid inserts using the 2-step 12*attB* site PCR protocol (Invitrogen)

Template specific primers containing 12 nucleotides of the *attB* sites were designed for the remaining CYP9 genes as follows:

12*attB*1: 5'-AA AAA GCA GGC TNN-(template specific sequence)-3'

12*attB*2: 5'-A GAA AGC TGG GTN-(template specific sequence)-3'

For the exact sequences used see appendix table 2.6.

Universal *attB* adapter primer sequences were:

***attB*1 adapter primer:**5'-GGGG ACA AGT TTG TAC AAA AAA GCA GGC T-3'

***attB*2 adapter primer:**5'-GGGG AC CAC TTT GTA CAA GAA AGC TGG GT-3'

The following protocol was used, based on the 2-step 12*attB* site PCR protocol (Invitrogen), to amplify *attB*-flanked DNA fragments. Blunt-ended product was required and so Phusion® high-fidelity DNA polymerase (New England BioLabs Inc.) was used in PCR. Step 1 PCR reactions (20 µl) contained 40 ng template DNA, 10 µM template specific 12*attB*1 forward and reverse primers, 0.4 µl 10mM dNTPs, 4 µl Phusion® HF buffer and 0.2 µl Phusion® DNA polymerase. PCR reaction temperature cycling conditions were 98°C for 2 minutes, followed by 10 cycles of 98°C for 15 seconds (denaturation), 64°C for 30 seconds (annealing) and 72°C for 2 minutes (extension).

Step 2 PCR reactions (20 µl) contained 4.0 µl PCR product Step 1, 10 µM universal *attB*1 adapter primer forward and reverse primers, 0.4 µl 10mM dNTPs, 4 µl Phusion® HF buffer and 0.2 µl Phusion® DNA polymerase. PCR reaction temperature cycling conditions were: phase 1, 98°C for 2 minutes,

followed by 5 cycles of 98°C for 15 seconds (denaturation), 60°C for 30 seconds (annealing) and 72°C for 2 minutes (extension); phase 2, 30 cycles of 98°C for 15 seconds (denaturation), 64°C for 30 seconds (annealing), 72°C for 2 minutes (extension) and a final extension of 72°C for 10 minutes. To remove any *attB* primers or *attB* primer-dimers, the PCR product was purified using gel extraction (see section 2.2.6.2). The PCR product was run on a 1.5% TAE agarose gel at 65 V for 1 hour and 20 minutes (see section 2.2.5).

5.2.2.2.2 The BP recombination reaction

BP reactions (10 µl) contained 2 µl of BP Clonase® II enzyme mix, 50 fmol *attB*-PCR product and 50 fmol pDONR™221 donor vector and were performed according to the Gateway® Clonase II protocol (Invitrogen, clonase II 2012). The resulting BP reaction was then used to transform chemically competent *E. coli* cells.

5.2.2.2.3 Transformation of chemically competent cells

1 µl of each BP reaction was used to transform Library Efficiency® DH5α™ competent cells (Invitrogen) and the culture was plated out on LB agar plates containing kanamycin, as outlined in section 2.2.8.3.

5.2.2.2.4 Colony PCR to verify Gateway® plasmid inserts

PCR was used to confirm the presence of plasmid inserts as described in section 2.2.8.4. For each gene of interest 6 colonies were used in the colony PCR and to inoculate a fresh LB agar plate containing kanamycin. Plated colonies were incubated overnight at 37°C. PCR cycling condition were as described in section 2.2.8.4. The PCR product was visualised on a 1% TAE agarose gel to confirm the presence of the plasmid insert as described in section 2.2.6.

5.2.2.2.5 Column purification of Gateway® BP Entry Clones

For each gene of interest, the two colonies that showed the clearest bands, when visualised on a 1% TAE agarose gel, were purified as described in section 2.2.8.5. The BP entry clones were stored at -20°C until needed.

5.2.2.3 Creating Gateway® Expression Clones from Entry Clones using the BaculoDirect™ Baculovirus Expression System (Invitrogen)

5.2.2.3.1 The LR recombination reaction

LR reactions (10 µl) contained 10 µl of BaculoDirect™ C-Term Linear DNA (Invitrogen), 100-300 ng BP entry clone and 4 µl of LR Clonase® II enzyme mix. Reactions were performed according to the BaculoDirect™ baculovirus expression system protocol (Invitrogen BaculoDirect 2012). A map of BaculoDirect™ C-Term linear DNA is shown in appendix figure 5.3.

5.2.2.3.2 Screening LR recombinant baculovirus DNA using PCR

The LR reaction was analysed by PCR using DreamTaq Green PCR Master Mix (2X) (Thermo Scientific) as outlined in section 2.2.5.2.1. A 2 µl aliquot of the LR reaction was diluted 200-fold and 2 µl of this was used in a PCR. PCR reactions (25 µl) contained 2 µl diluted LR reaction, 10 µM stocks of Polyhedrin forward primer and V5 reverse primers and 12.5 µl DreamTaq mastermix. PCR reaction temperature cycling conditions were 95°C for 5 minutes, followed by 25 cycles of 94°C for 45 seconds (denaturation), 52°C for 1 minute (annealing) and 72°C for 90 seconds (extension), and then a final extension at 72°C for 10 minutes. The products (~1700-1800 bp) were visualised on a 1% TAE agarose gel run at 70 V for 1 hour and 20 minutes as described in section 2.2.6. The LR reaction product was then directly transfected into the Sf9 insect cell line, as described in section 5.2.2.5.1.

5.2.2.4 Expression of recombinant P450s in insect cell lines

Insect cell lines that support efficient replication of baculovirus have been derived from the Fall army worm (*Spodoptera frugiperda*) and the cabbage looper moth (*Trichoplusia ni*) [553]. Two commercially available insect cell lines were used in this project: *S. frugiperda* Sf9s and *T. ni* High Five™ cell lines (Invitrogen). Cell lines were maintained as described in section 2.4.

5.2.2.4.1 Transfection of the LR recombination reaction into Sf9 insect cells

Transfection of Sf9 cells was performed following the protocol outlined in the BaculoDirect™ baculovirus expression system manual [554]. In each well of a treated six-well plate (Falcon), 8×10^5 Sf9 cells ($1.5\text{--}2.5 \times 10^6$ cells/ml) at >95% viability were seeded in 2 ml Unsupplemented Grace's Insect Medium (Thermo

Fisher). The cells were evenly distributed and allowed to attach to the plate for 15-30 minutes at room temperature. Two wells of Sf9 cells were transfected with each LR recombination reaction. The following solutions were prepared in 1.5 ml microcentrifuge tubes for each transfection sample: Transfection mixture A - 16 μ l Cellfectin® II reagent and 200 μ l Grace's Insect Medium (Gibco); Transfection mixture B - 18 μ l LR recombination reaction and 200 μ l Grace's Insect Medium. Once the cells were attached, transfection mixtures A and B were combined and mixed gently by tapping the tube. The resulting solutions were incubated at room temperature for 25-35 minutes. After incubation the transfection mixture was divided equally into the two prepared wells. The plates were then incubated at 27°C for five hours. After this time the transfection mixture was removed and replaced with 2 ml of fresh Sf-900™ II SFM media supplemented with 100 μ M ganciclovir (Thermo Fisher) (a nucleoside analog that negatively selects against non-recombinant baculovirus) to each well. The plates were placed in a sealed plastic bag containing a moist paper towel to prevent evaporation and incubated at 27°C for 72 hours, or until signs of viral infection could be seen.

5.2.2.4.2 Isolation of P1 viral stock

Plates were inspected daily after 48 hours using a microscope (Leitz DM IL). The inclusion of two control wells per plate, containing cells that were not transfected, provided a clear comparison between healthy cells and those showing signs of viral infection, such as an increase in size or detachment from the plate (lysis). Once transfected cells showed clear signs of infection (>72 h), the medium from each well was aspirated into sterile 15 ml tubes (Falcon). Both wells of each LR reaction were combined (~4 ml medium). The tubes were centrifuged at 3500 rpm for 5 minutes to remove cell debris. The supernatant was transferred into fresh sterile 15ml tubes (Falcon) and 5% foetal bovine serum (FBS) (Sigma) was added. P1 viral stocks were protected from light and stored at 4°C.

5.2.2.4.3 Preparation and isolation of the P2 viral stock

In each well of a treated six-well plate (Falcon), 8×10^5 Sf9 cells (1.5 - 2.5×10^6 cells/ml) at >95% viability were seeded in 2 ml fresh Sf-900™ II SFM media (Thermo Fisher) supplemented with 100 μ M ganciclovir (Thermo Fisher). The

cells were evenly distributed and allowed to attach to the plate for 15-30 minutes at room temperature. Two wells of Sf9 cells were seeded for each P1 viral stock. Once the cells were attached, 200 µl of the P1 viral stock was added, dropwise into each well. The plates were placed in a sealed plastic bag containing a moist paper towel to prevent evaporation and incubated at 27°C for 72 hours, or until signs of viral infection could be seen. Once cells showed clear signs of infection, the P2 viral stock was harvested and stored as described in 5.2.2.5.2.

5.2.2.4.4 Quantification of viral titre by enzyme-linked immunosorbent assay (ELISA)

Assay set up and infection

Viral infection at the correct multiplicity of infection (MOI) is essential to achieving optimal protein yields with baculovirus expression systems and therefore quantification of the viral titre of P2 and P3 viral stocks was performed. The following protocol was adapted from BacPAK™ Baculovirus Rapid Titration Kit, Clontech [555]. One row (12 wells) of a tissue culture treated 96-well plate (CytoOne) was seeded with Sf9 cells in log phase (6.5×10^4 cells/well or $3-4 \times 10^5$ cells/ml) at >95% viability for each viral stock to be quantified. The plate was then incubated for 30-60 mins at 27°C to allow the cells to attach. In a separate 96-well plate serial dilutions of the viral stocks to be assessed were prepared, to give final dilutions of 10^{-3} , 10^{-4} , 10^{-5} and 10^{-6} of viral stock in Sf-900™ II SFM media. The plate was checked under a microscope (Leitz DM IL) to ensure the cells had attached. The media was aspirated from the wells, taking care not to scrape the cells from the bottom of the plate. 25 µl aliquots of the viral dilutions were added to the appropriate well, see appendix figure 5.4 for plate layout. Media was not removed from the negative control wells (C). The plate was then incubated at 27°C on a shaking platform (110 rpm; Thermo Scientific) for 50 minutes. The media was aspirated from the wells and 50 µl of fresh Sf-900™ II SFM media was added to each well. The plate was placed in a sealed plastic bag containing a moist paper towel to prevent evaporation and incubated at 27°C for 43-47 hours.

Virus detection

The reagents shown in appendix table 5.1 were prepared, volumes shown are sufficient for one 96-well plate. 100 µl of media was first removed from the negative control wells. Then 150 µl of freshly prepared ice-cold formyl buffered acetone was added to the wells and the plate was incubated at room temperature for 10 minutes. The plate was then emptied and tapped onto paper towel and washed by adding 200 µl PBS + 0.05% Tween 20 and shaking at room temperature for five minutes. This was repeated three times, emptying the plate between washes. 50 µl diluted normal goat serum was added to each well and the plate was incubated on the shaker for 5 minutes at room temperature. The diluted goat serum was emptied out and the plate tapped out well on paper towel. 25 µl of diluted mouse gp64 antibody was added to each well and the plate was incubated on the shaker at 37°C for 25 minutes. The plate was emptied and washed twice by adding 200 µl PBS + 0.05% Tween 20 and shaking as before. 50 µl of goat anti-mouse antibody/ HRP conjugate was added to the wells and the plate was incubated on the shaker at 37°C for 25 minutes. The plate was emptied and washed three times by adding 200 µl PBS + 0.05% Tween 20 and shaking as before. 50 µl of TrueBlue™ Peroxidase Substrate (KPL) was added to the wells and the plate was incubated at room temperature for 2-3 hours, although a preliminary (less accurate) estimate of viral titre can be made after as little as 10 minutes after adding the substrate.

Determination of viral titre

The number of stained foci of infection in the wells (see appendix figure 5.5) were counted using a microscope (Leitz DM IL). The highest dilution wells containing a reasonable number of foci (~5-25) were used. Each discrete cluster of stained cells was counted as one focus. The average number of foci per well at a dilution was used to calculate the viral titre with the following equation:

Viral titre (pfu/ml) = average no of foci x dilution factor x normalisation factor
(Normalisation factor = 40 when viral infection is done using 25 µl).

5.2.2.4.5 Preparation and isolation of the P3 viral stock

Sf9 cells were seeded at a density of 2×10^6 cells/ ml in 50 ml Sf-900™ II SFM media in a sterile 250 ml Erlenmeyer flask (Corning). The amount of P2 stock

needed to infect this culture was calculated using a MOI of 0.1-0.2 and the following equation:

$$P2 \text{ inoculum required (ml)} = \frac{MOI \text{ (pfu/cell)} \text{ (cell density} \times \text{volume)}}{Viral \text{ titre (pfu/ml)}}$$

The inoculum was added to the culture and the flask incubated at 27°C on a rocking platform (110 rpm, Thermo Scientific). After 48 hours the viability of the cell culture was assessed as described in section 5.2.2.4.4. When the viability was <10% but >0% the P3 viral stock was harvested (generally up to 7 days post infection). The culture from each flask was transferred into sterile 50 ml tubes (Falcon) and centrifuged at 3500 rpm for 10 minutes to remove cell debris. The supernatant was transferred to fresh, sterile 50 ml tubes and 5% FBS added. P3 stock was protected from light and stored at 4°C. The viral titre of the P3 stock was quantified as described in section 5.2.2.5.4.

5.2.2.4.6 Expression of recombinant P450 proteins

Culture for protein expression

High Five™ cells were seeded at a density of 2x10⁶ cells/ ml in 80 ml Express Five™ SFM media (Thermo Fisher) in a sterile 250 ml Erlenmeyer flask (Corning). P450s are heme-containing proteins and to functionally express them successfully a source of ferrous iron and a heme-precursor are both needed. 10mM Iron (III) citrate (Aldrich) and 10mM 5-aminolevulinic acid (Sigma) were therefore added to the culture at a final concentration of 0.1 mM. For optimal enzymatic activity the presence of the redox-partner cytochrome P450 reductase (CPR) is required. To provide this the High Five™ culture was co-infected with baculovirus P3 viral stock produced from recombinant P450s and baculovirus P3 viral stock produced with recombinant CPR. In this instance house fly (*M. domestica*) CPR was used. House fly P2 viral stocks were kindly supplied by Christoph Zimmer (Syngenta). The amount of recombinant P450 P3 stock and CPR P3 stock needed to infect this culture were calculated using a MOI of 3.0 and 0.5 respectively, using the equation in section 5.2.2.5.5. The infected cells were incubated at 27°C on a shaking platform (110 rpm; Thermo Scientific) for 24 hours. A second dose of 10 mM Iron (III) citrate (Aldrich) and 10 mM 5-aminolevulinic acid (Sigma) to a final concentration of 0.1 mM was added and the culture was incubated at 27°C on a rocker plate (110 rpm;

Thermo Scientific) for a further 36 hours. The cells were then harvested using differential centrifugation.

Harvesting the protein

50 ml of the culture was transferred into a 50 ml tube (Falcon) and centrifuged at 1000 rpm at 15°C for 5 minutes. The supernatant was discarded and the remaining 30 ml of culture was added to the 50 ml tube. The tube was centrifuged at 1000 rpm at 15°C for 10 minutes and the supernatant discarded. The cell pellet was then resuspended in 30 ml of homogenisation buffer (see appendix table 2.2). The cells were broken up using a sonifier (Branson digital sonifier 250) set at 50% amplitude, pulsing on for two seconds, off for two seconds, for one minute. The resulting solution was centrifuged at 2700 rpm at 4°C for 10 minutes. The supernatant was transferred to fresh tubes and centrifuged at 100,000 g (~30,000 rpm; Beckman Coulter Optima XPN-100 Ultracentrifuge, Rotor type Ti 32) at 4°C for one hour. The supernatant was discarded and the remaining pellet resuspended in 1-4 ml buffer R (see appendix table 2.3) using a glass homogeniser (VWR).

5.2.2.4.7 CO difference spectroscopy

The CO-difference spectral assay protocol performed for each recombinant P450 was adapted from Guengerich *et al.*, 2009 [556]. A double beam Specord® 200 PLUS spectrophotometer (Analytik Jena) was used to measure the absorbance (nm) of the recombinant P450 samples. In a 15 ml tube (Falcon) 100 µl of recombinant protein (+0.05% BSA) was added to 1.9 ml of potassium phosphate buffer (see appendix table 2.1). A pinch of sodium dithionite, Na₂S₂O₄ (Sigma) was added and the tube was vortexed thoroughly. The sample was divided equally into the measurement (M) and reference (R) cuvettes (Hellma Analytics) and a baseline reference measurement was taken according to the manufacturers protocol. The M cuvette was removed and carbon monoxide (CO) was bubbled through the sample at a rate of 2 bubbles per second for 30 seconds. The M cuvette was replaced in the spectrophotometer and the absorbance measurement was repeated. The concentration of the recombinant P450 can be calculated using the difference between the absorbance at 450 nm and 490 nm and the extinction coefficient $\epsilon = 91 \text{ mM}^{-1} \text{ cm}^{-1}$ [220] using the formula below:

$$P450 \text{ concentration (nmol/ml)} = \frac{(\Delta A_{450} - \Delta A_{490})}{0.091} \times 20$$

The total protein concentration of the sample was determined using a Bradford's assay as described in section 2.3.1. The recombinant P450s were dispensed in 200-500 µl aliquots into cryogenic vials, snap frozen in liquid nitrogen and stored at -80°C.

5.2.3 Metabolism assays with analysis using tandem LC-MS/MS

Native microsomes (160 µg/well) or recombinant P450s (5 pmol/well) in 0.1 M potassium phosphate buffer (see appendix table 2.1) were incubated for 1 hour (with shaking) with a range of substrates, at a concentration of 10µM, in a total assay volume of 200 µl, at 30±1°C, in the presence or absence of a NADPH regeneration system. Three replicates were performed for each data point. Samples incubated without NADPH served as controls. The reactions were terminated by the addition of ice-cold acetonitrile (to 80% final concentration). Samples were incubated at 4°C overnight and then centrifuged at 3,000 g for 10 minutes to pellet any precipitation of protein. LC-MS/MS was performed as previously described by Zhu *et al.*, 2010 [557], using an Acquity UPLC (Waters) coupled to an API 4000 mass spectrometer (Sciex) and an Infinity II UHPLC (Agilent Technologies; reverse phase mode) coupled to a QTRAP 6500 mass spectrometer (Sciex) employing electrospray ionization. The recovery rates of parent compounds incubated without NADPH were normally close to 100%. LC-MS/MS was carried out in house by Bayer (AG, Crop Science, Leverkusen, Germany). Calculations were carried out using Microsoft Excel (2013) and GraphPad Prism (Version 7.03). A Welch's t-test [558] was used to analyse the difference between incubation with and without NADPH for each compound. Substrates tested are shown in table 5.1.

Table 5.1: Compounds used in metabolism assays

Insecticide class	Name	Source
NEONICOTINOID	Imidacloprid	Sigma Aldrich
	Thiacloprid	Sigma Aldrich
	Acetamiprid	Sigma Aldrich
	Clothianidin	Sigma Aldrich
BUTENOLIDE	Flupyradifurone	Sigma Aldrich
PYRETHROID	Deltamethrin	Sigma Aldrich
	<i>tau</i> -Fluvalinate	Sigma Aldrich
	<i>alpha</i> -Cypermethrin	Sigma Aldrich
ORGANOPHOSPHATE	Coumaphos	Sigma Aldrich
	Chlorpyrifos	Sigma Aldrich
DIAMIDE	Flubendiamide	Sigma Aldrich
Allelochemical (alkaloid)	Associated plant species	Source
Nicotine	Solanaceae family e.g. Tobacco plant (<i>Nicotiana tabacum</i>)	
Anabasine	Tree tobacco (<i>Nicotiana glauca</i>)	Sigma Aldrich
Atropine	Solanaceae family e.g. Deadly nightshade (<i>Atropa belladonna</i>)	Sigma Aldrich
Cytisine	Fabaceae family e.g. Laburnum (<i>Laburnum anagyroides</i>)	Sigma Aldrich

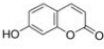
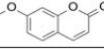
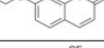
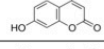
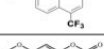
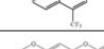


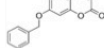
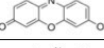
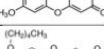
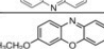
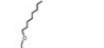
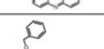
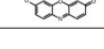
5.2.4 Model substrate profiling

The model substrates used are shown in table 5.2. All assays using model substrates were performed in a black flat-bottomed 96-well plate (4-tititude). Data were recorded using a SpectraMax M2 Multi-mode plate reader (Molecular Devices).

5.2.4.1 Temperature optimisation

The effect of temperature on enzyme activity is well established and is important to consider here since the way P450s respond to temperature may be fundamental to their biological role (Peterson et al., 2007; Daniel & Danson, 2013). Temperature optimisation of metabolism assay conditions was performed for *A. mellifera* and *O. bicornis* and the optimal temperature was found to be 30°C, however, for *B. terrestris* the temperature was lower ~25°C (Manjon et al., 2018). Temperature optimisation of the assay conditions to determine the optimal temperature for the activity of *M. rotundata* P450s was therefore performed as follows. The assay was performed in a 100 µl reaction containing 2 pmol of a randomly selected expressed recombinant P450, 1 mM NADPH (Sigma) and 50 mM of EFC (Sigma). Samples were incubated at 25,

Table 5.2: The chemical structure, molecular weight and excitation/emission wavelengths (nm) of P450 fluorescent model substrates used in this study

Chemical root	Substrate	Substrate abbreviation	Reference	Structure	Chem. Formula	MW g/mol	Excitation (nm)	Emission (nm)	Supplier	Conc in assay
coumarin	7-Hydroxycoumarin	HC			C ₉ H ₆ O ₃	162.14			Sigma	50mM (50μM in test)
coumarin	7-methoxy-coumarin	MC	HC		C ₁₀ H ₈ O ₃	176.17	390	465	Sigma	50mM (50μM in test)
coumarin	7-ethoxy-coumarin	EC	HC		C ₁₁ H ₁₀ O ₃	190.195	390	465	Sigma	50mM (50μM in test)
coumarin (trifluoromethyl)	7-hydroxy-4-(trifluoromethyl)-coumarin	HFC			C ₁₀ H ₅ F ₃ O ₃	230.14			Sigma	50mM (50μM in test)
coumarin (trifluoromethyl)	7-ethoxy-4-(trifluoromethyl)coumarin	EFC	HFC		C ₁₂ H ₉ F ₃ O ₃	258.19	410	510	Sigma	50mM (50μM in test)
coumarin (trifluoromethyl)	7-methoxy-4-(trifluoromethyl)coumarin	MFC	HFC		C ₁₁ H ₇ F ₃ O ₃	244.17	410	535	Sigma	50mM (50μM in test)
coumarin (trifluoromethyl)	7-benzyloxy-4-(trifluoromethyl)-coumarin	BFC	HFC		C ₁₇ H ₁₁ F ₃ O ₃	320.26	410	535	Sigma	50mM (50μM in test)
coumarin (trifluoromethyl)	7-(4-methoxybenzyloxy)-4-trifluoromethylcoumarin	MOBFC	HFC		C ₁₈ H ₁₃ F ₃ O ₄	350.29	405	510	Invitrogen	50mM (50μM in test)
coumarin (trifluoromethyl)	7-(benzyloxymethoxy)-4-trifluoromethylcoumarin	BOMFC	HFC				405	510	Bayer (custom synthesized by Enamine Ltd. (Kiev, Ukraine))	50mM (50μM in test)
resorufin	7-Hydroxy-resorufin	HR			C ₁₂ H ₇ NO ₃	213.19			Sigma	1mM (1μM in test)
resorufin	Methoxy-resorufin	MR	HR		C ₁₃ H ₉ NO ₃	227.26	535	590	Sigma	1mM (1μM in test)
resorufin	Pentoxymethoxy-resorufin	PR	HR		C ₁₇ H ₁₇ NO ₃	283.33	535	590	Sigma	1mM (1μM in test)
resorufin	Ethoxy-resorufin	ER	HR		C ₁₄ H ₁₁ NO ₃	241.2	535	590	Sigma	1mM (1μM in test)
resorufin	(octyloxymethoxy)resorufin	OOMR	HR		C ₂₁ H ₂₉ NO ₄	355.43	520	590	Invitrogen	1mM (1μM in test)
resorufin	(benzyloxymethoxy)resorufin	BOMR	HR		C ₂₀ H ₁₉ NO ₄	333.34	520	590	Invitrogen	1mM (1μM in test)

30, 35 and 40°C for three hours. Four replicates were performed for each data point and the assay was repeated twice for all temperatures. P450 incubated without NADPH and wells containing only potassium phosphate buffer (see appendix table 2.1) served as controls. Data were recorded every minute at excitation/emission wavelength suitable for EFC (410/510 nm, see table 5.2). Average control measurements were subtracted from average fluorescent measurement at each data point and the results graphed in Microsoft® Excel (Version 16.35).

5.2.4.2 Standard curves for model substrates

Standard curves were created using the following standards: HC for coumarin, HFC for trifluoromethyl coumarin and HR for resorufin model substrates, to allow prediction of the concentration of fluorescent product formed in kinetic assays with microsomes or recombinant P450s (see table 5.2). Each standard was diluted to a range of concentrations (0, 5, 10, 15, 20, 30, 50, 60, 80 and 100 pmol) using potassium phosphate buffer (see appendix table 2.1). 100 µl of each concentration of HFC and HR, and 200 µl of each concentration of HC, were added to each well, with four replicates performed for each data point. The plate was incubated and read at the optimal temperature. The fluorescence was recorded at excitation/emission wavelengths suitable for each model substrate (see table 5.2). The 'TRENDLINE' function in Microsoft® Excel (Version 16.35) was used to graph the standard curves and to calculate the R² values.

5.2.4.3. Model substrate metabolism assays

Metabolism assays were performed in a 100 µl reaction in each well, containing 50 µg microsomal protein or 2 pmol of recombinant P450, 1 mM NADPH (Sigma) and either a coumarin-based (50 mM) or resorufin-based (1 mM) substrate. Three replicates were carried out for each data point. Samples were mixed and incubated at the optimal temperature. Microsomal protein or P450s incubated without NADPH, and wells incubated without protein containing only potassium phosphate buffer (see appendix table 2.1) served as controls. *O. bicornis* microsomes acted as a further comparative control for microsomal assays. Data were recorded every minute, with mixing between each measurement, for 60 minutes at the appropriate excitation/emission wavelengths (nm) for each model substrate (see table 5.2). An endpoint reading

was also taken. MC and EC have a similar emission wavelength (465 nm) to NADPH (460 nm). Reactions containing these model substrates were terminated (after incubation for 60 minutes) prior to measurement by the addition of 100 μ l of a stop solution (see appendix table 2.4), making a 200 μ l reaction in total. The plates were incubated at the optimal temperature for a further 50 minutes and the data recorded as before at the correct excitation/emission wavelengths (390/465 nm). An endpoint reading was also taken. Calculations were carried out using Microsoft® Excel (Version 16.35). Control measurements were subtracted from substrate measurements and the 'TREND' function was used to calculate the y intercept using the appropriate standard curve. For each model substrate used three replicates for all the P450s screened were run within one plate. The resulting data were analysed using GraphPad Prism (Version 7.03).

5.2.4.4 Fluorescence-based assay to assess insecticide mediated inhibition

Of the two *M. rotundata* CYP9DM protein sequences, CYP9DM2 was found to be more active against coumarin-based model substrates, and was therefore selected for use with thiacloprid and imidacloprid, in an insecticide mediated inhibition assay. *O. bicornis* CYP9BU1, a P450 known to metabolise thiacloprid, was also selected as a comparison. Viral stocks of *CYP9BU1* were kindly provided by Kat Beadle. The method outlined below is amended from protocols published by Haas and Nauen, 2021 [549] and Haas *et al.*, 2021 [550]. For each recombinant CYP9, a coumarin-based model substrate (see table 5.9) which showed good activity was selected as a suitable fluorescent probe substrates for use. Inhibition assays were performed in a total reaction of 50 μ l per well. Four replicates were carried out for each data point. Each replicate consisted of 25 μ l of suitable fluorescent probe substrate (final concentration range 0.5-200 μ M), 1 mM NADPH (Sigma) and 25 μ l recombinant P450, plus a competing insecticide. Final inhibitor insecticide concentrations of 10 μ M and 100 μ M were used. Each recombinant P450 was pre-incubated with the inhibitor at 20°C for 10 minutes, before the addition of the 25 μ l of substrate (with or without 1 mM NADPH). Recombinant P450 protein incubated without NADPH, and wells incubated without protein, containing only potassium phosphate buffer (see appendix table 2.1) served as controls. Reactions were incubated at 25°C for 60 minutes at the appropriate excitation/emission

wavelengths for the chosen model substrate, see table 5.2. Reactions were terminated (after incubation for 60 minutes) by the addition of 50 μl of a stop solution (see appendix table 2.4) making a 100 μl reaction in total. The plates were incubated at 25°C, for a further 50 minutes after which an endpoint reading was taken. Calculations were carried out using Microsoft® Excel (Version 16.35). Control measurements were subtracted from substrate measurements and the 'TREND' function was used to calculate the y intercept using the appropriate standard curve. The data were analysed for evidence of inhibition by non-linear regression, where possible assuming Michaelis-Menten kinetics, using GraphPad Prism (Version 7.03).

5.3 Results

5.3.1 Screening male and female *M. rotundata* microsomes for metabolic activity

Microsomes isolated from intact male and female *M. rotundata* show a dramatic difference in functional activity when assayed against two coumarin-derived model substrates ($p < 0.0005$) (see figure 5.2). In fact, microsomes isolated from intact females show no activity against MC (average RFU/ min^{-1} mg protein $^{-1}$ = 1.85) and virtually none against EFC (average RFU/ min^{-1} mg protein $^{-1}$ = 7.98) compared to those from male microsomes (average RFU/ min^{-1} mg protein $^{-1}$ against MC = 804.16; average RFU/ min^{-1} mg protein $^{-1}$ against EFC = 274.45). Microsomes isolated from female heads and thoraxes show significantly more activity to both coumarins compared to microsomes isolated from intact females ($p < 0.0005$) (see figure 5.2). It therefore appears that, in common with *A. mellifera*, the venom sac of *M. rotundata* contains a P450 inhibitor. The solution to this problem, in *A. mellifera*, is to dissect the venom gland sting complex out from the worker bees before homogenisation (Zaworra, 2019). However, unlike *A. mellifera*, where most individuals are female workers, the sex ratio, male: female, of *M. rotundata* is 2:1 [79, 80]. The standard protocol for the acute contact bioassays of insecticides [482] calls for the use of female bees. This leaves a surplus of male bees that can be snap frozen and used to prepare microsomes, for use in metabolism assays, as needed. As such, only microsomes extracted from male bees were used in this study.

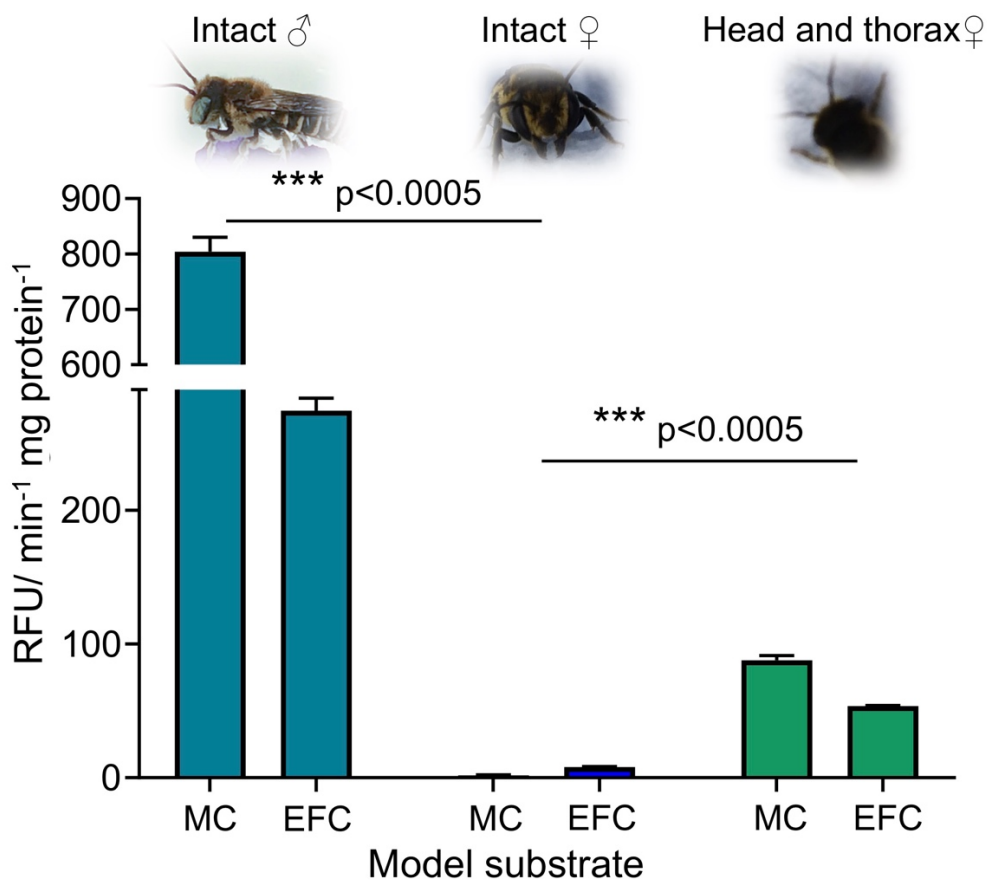


Figure 5.2: Enzyme activity in microsomal preparations using male, female and dissected female heads and thoraxes of *M. rotundata*. Activity is given in relative fluorescence units (RFU) based on screening against MC and EFC. Data are mean values \pm SD (n=3). Analysis performed using a Welch's t-test (two-tailed; degrees of freedom = 2.044) with significant differences indicated by *** $p < 0.0005$.

5.3.2 Expression of recombinant CYP9 P450s in an insect cell line

There are nine CYP9 family members in *M. rotundata*, of these two were synthesised by GeneArt (Life Technologies) using pDONR221 vector plasmids as described in section 5.2.2.1. PCR amplification of Gateway® plasmid inserts using the 2-step 12attB site PCR protocol (Invitrogen) was used to create BP entry clones for the remaining seven genes. Six reactions were successful producing PCR product flanked with attB sites (see appendix figure 5.6). The

protocol was not successful at amplifying *CYP9R59*. Gel extraction was used to purify the 12*attB* site PCR product as described in section 2.2.7.2 (see table 5.3). Four recombinant CYP9s produced P450 proteins that were correctly folded and produced a characteristic spectral absorption peak (450 nm) (see table 5.3 and appendix figure 5.7). Where possible, the concentration of recombinant P450 protein was calculated as described in section 5.2.2.5.7. Both CYP9DM2 and CYP9P2 produced a peak at 420 nm rather than 450 nm (see appendix figure 5.7). The global protein content of all the recombinant P450s was calculated using a Bradfords assay as described in section 2.3.1 [551] (see table 5.3).

Table 5.3: Expression of recombinant P450s using Gateway® cloning technology and transfection in insect cell lines.

Name	Size in bp	Entry clone method	<i>attB</i> -PCR product ng/μl	Entry clone concentration ng/μl	Amount of entry clone used in LR reaction μl	Viral titer P2 stock	Volume to transfect P3 stock μl (MOI = 0.1)	Viral titer P3 stock	Volume to transfect for expression ml (MOI = 3)	450nm peak	Amount of P450 nMol/ml	Amount of protein (mg/ml)
CYP9R1	1612	BP recombination reaction	73.2	208 255.9	1	1.42E+07	530	2.11E+08	1.71	Y	10.35	36.65
CYP9P2	1609	BP recombination reaction	63.0	100.2 142.2	1	2.96E+07	250	1.48E+08	2.43	N	N/A	33.18
CYP9P22	1582	BP recombination reaction	163.6	No colonies	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
CYP9P23	1600	BP recombination reaction	31.0	304.6 313.2	1	1.84E+07	410	6.87E+07	5.24	Y	4.40	31.45
CYP9DM2	1657	BP recombination reaction	32.1	325.9 325.1	1	1.75E+07	430	3.73E+08	0.97	N	N/A	25.87
CYP9DN1	1630	BP recombination reaction	88.1	No colonies	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
CYP9R58	1543	GeneArt synthesis	N/A	48.6	2	7.33E+07	100	4.00E+07	9.0	Y	8.37	42.33
CYP9DM1	1597	GeneArt synthesis	N/A	33.3	2	1.52E+07	490	5.87E+07	6.1	Y	0.66	24.69

5.3.3 Metabolism assays with analysis using tandem LC-MS/MS

5.3.3.1 Neonicotinoid and butenolide insecticides

Analysis of parent compound depletion by LC-MS/MS after incubation of *M. rotundata* microsomes with *N*-cyanoamidine and *N*-nitroguanidine neonicotinoids and the butenolide flupyradifurone showed no significant metabolism of any of the insecticides (see figure 5.3 and table 5.4). No significant difference in metabolism between *N*-cyanoamidine and *N*-nitroguanidine compounds was observed (see table 5.5).

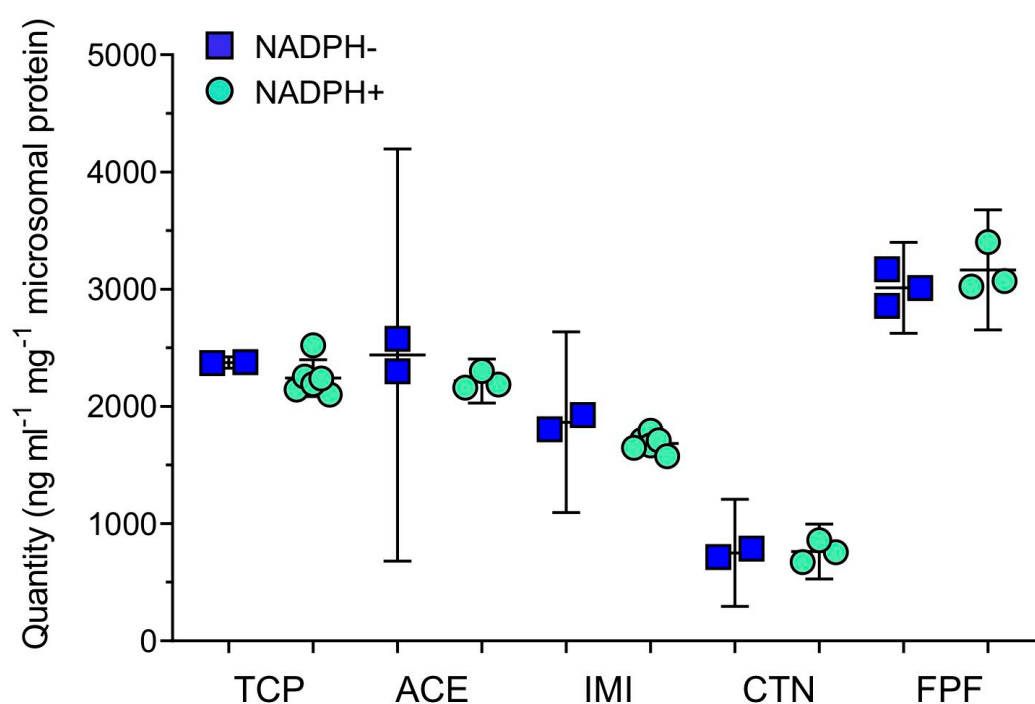


Figure 5.3: Metabolism of the neonicotinoid insecticides thiacloprid (TCP), acetamiprid (ACE), imidacloprid (IMI), clothianidin (CTN) and the butenolide insecticide flupyradifurone (FPF) by native microsomal preparations of *M. rotundata* as measured by LC-MS/MS (1 h incubation at 30°C \pm NADPH). The error bars indicate 95% CIs.

5.3.3.2 Pyrethroid insecticides

Analysis of parent compound depletion by LC-MS/MS after incubation of *M. rotundata* microsomes with select pyrethroid insecticides showed no significant metabolism of any of the insecticides (see figure 5.4 and table 5.4). No significant difference in metabolism between the pyrethroid compounds was observed (see table 5.5).

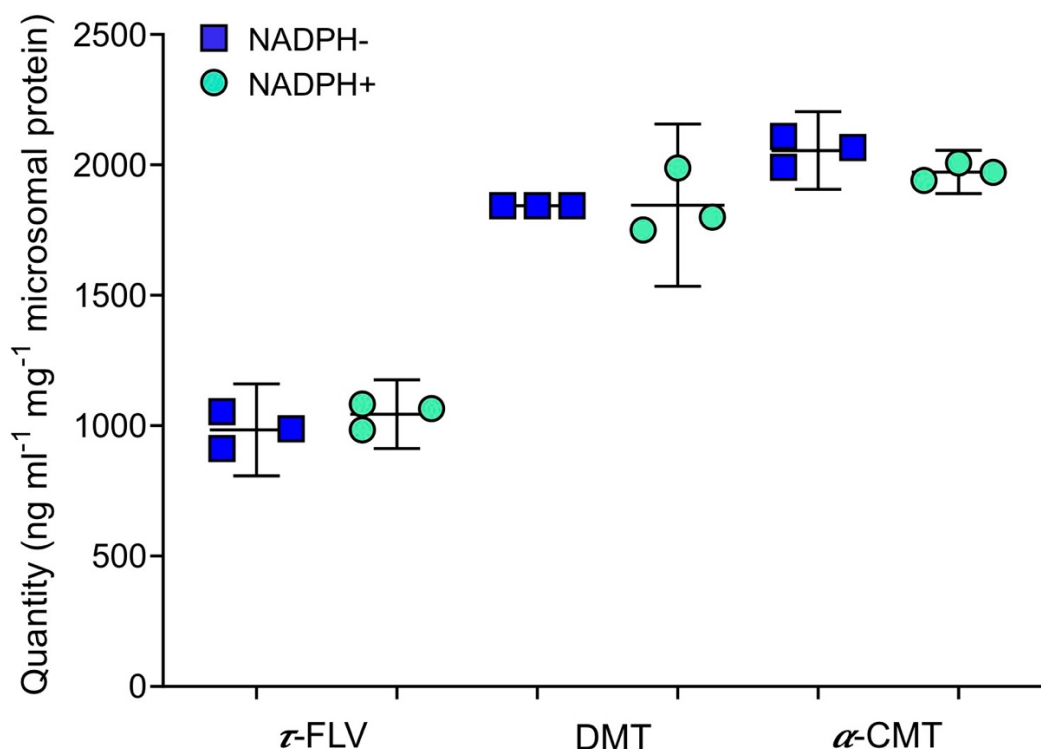


Figure 5.4: Metabolism of the pyrethroid insecticides *tau*-fluvalinate (τ -FLV), deltamethrin (DMT) and *alpha*-cypermethrin (α -CMT) by native microsomal preparations of *M. rotundata* as measured by LC-MS/MS (1 h incubation at 30°C \pm NADPH). The error bars indicate 95% CIs.

5.3.3.3 Organophosphate and diamide insecticides

Incubation of *M. rotundata* microsomes with two organophosphate insecticides and analysis of parent compound depletion by LC-MS/MS showed that both were significantly metabolised ($p < 0.05$) (see figure 5.5 and table 5.4).

Approximately 13% of coumaphos and 14% of chlorpyrifos was metabolised in one hour. However, after a correction for multiple testing was applied there was

no significant difference in metabolism between the organophosphate compounds (see table 5.5).

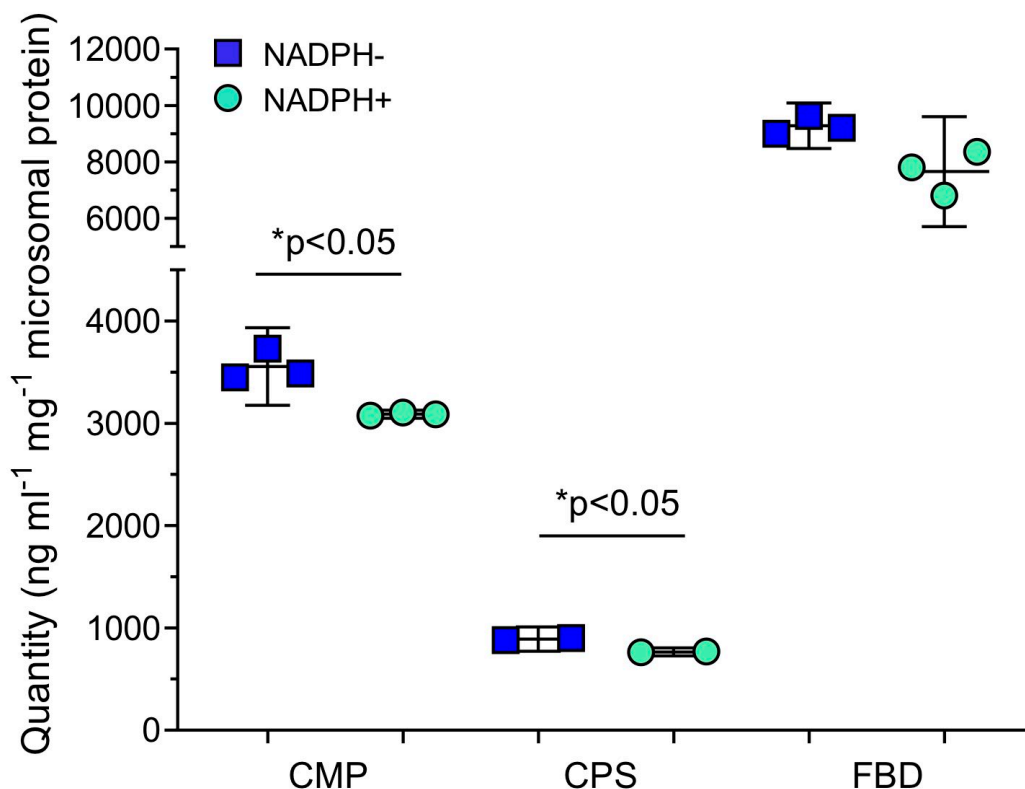


Figure 5.5: Metabolism of the organophosphate insecticides coumaphos (CMP), Chlorpyrifos (CPS) and the diamide insecticide flubendiamide (FBD) by native microsomal preparations of *M. rotundata* as measured by LC-MS/MS (1 h incubation at 30°C ±NADPH). The error bars indicate 95% CIs. Analysis performed using a Welch's t-test (two-tailed) with significant differences indicated by *p<0.05.

Table 5.4: Welch's t-test results for metabolism of synthetic insecticides by native microsomal preparations of *M. rotundata* as measured by LC-MS/MS (1 h incubation at 30°C ±NADPH).

Insecticide	Welch's t-test (two-tailed)		
	p value	Sig p<0.05	df
Thiacloprid	0.0804	N	5.038
Acetamiprid	0.3382	N	1.201
Imidacloprid	0.1529	N	1.528
Clothianidin	0.866	N	2.988
Flupyradifurone	0.3684	N	3.723

Tau-fluvalinate	0.3123	N	3.703
Deltamethrin	0.9796	N	2.000
<i>alpha</i> -cypermethrin	0.1244	N	3.140
Coumaphos	0.0327	Y*	2.042
Chlorpyrifos	0.0304	Y*	1.22
Flubendiamide	0.0537	N	2.666

Table 5.5: Results for univariate analysis of variance using general linear model for metabolism of synthetic insecticides by native microsomal preparations of *M. rotundata* as measured by LC-MS/MS (1 h incubation at 30°C ±NADPH). Tests of between-subject effects with *post-hoc* multiple comparisons of depletion using a Bonferroni adjustment.

	GLM between-subjects test		Post-hoc pairwise comparison (Bonferroni adjustment)
	p value	F	Sig. p<0.005 ($\alpha/10$)
Neonicotinoid			
Thiacloprid	0.021	3.60	N
Acetamiprid			
Imidacloprid			
Clothianidin			
Flupyradifurone			
Pyrethroid			Sig. p<0.017 ($\alpha/3$)
Tau-fluvalinate	0.641	0.461	N
Deltamethrin			
Alpha-cypermethrin			
Organophosphates			Sig. p<0.017 ($\alpha/3$)
Coumaphos	0.711	0.366	N
Chlorpyrifos			
Flubendiamide			

Due to the lack of metabolism of synthetic insecticides shown by *M. rotundata* native microsomes, assays to *examine* Michaelis-Menten kinetics were not performed.

5.3.3.4 Alkaloid allelochemicals

Incubation of *M. rotundata* microsomes with four naturally occurring alkaloid allelochemicals and analysis of parent compound depletion by LC-MS/MS showed that two of them; nicotine and cytosine, were significantly and rapidly metabolised ($p<0.01$ and $p<0.001$ respectively) (see figure 5.6 and table 5.6).

There were differences in the depletion of parent compound between the alkaloid allelochemicals (see table 5.7). The depletion of cytisine and nicotine was significantly greater than that of anabasine or atropine ($p < 0.008$, see table 5.7). Approximately 40% of nicotine and 55% of cytisine was metabolised in one hour, compared to only 17% of anabasine. There was no evidence that atropine was metabolised.

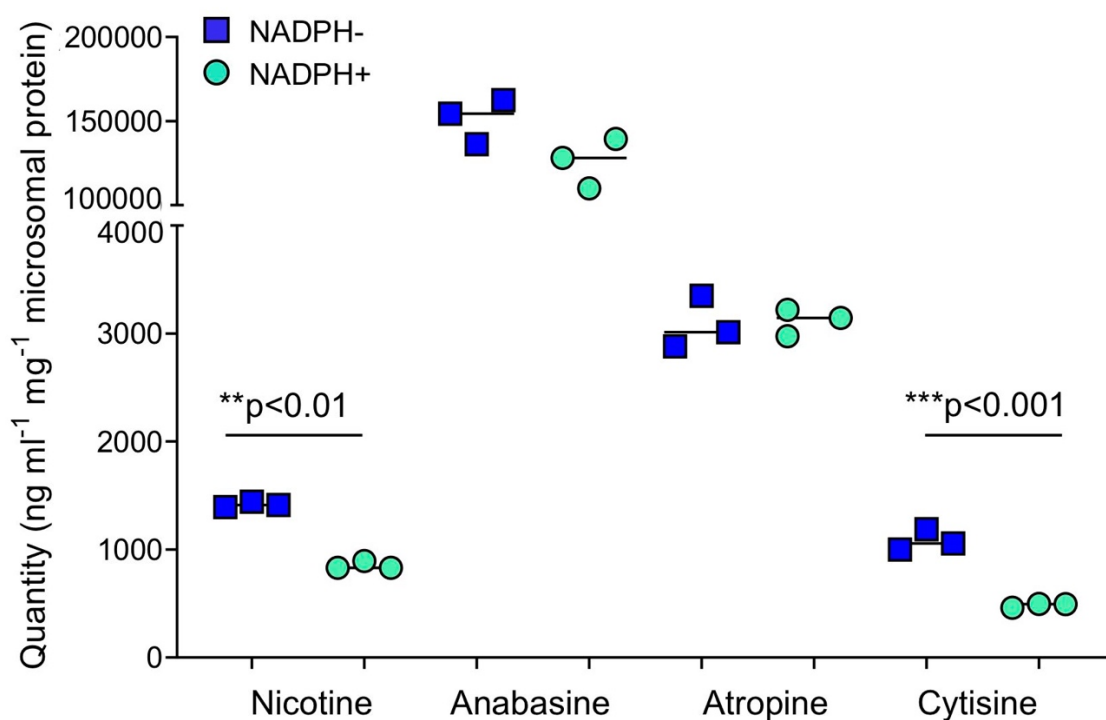


Figure 5.6: Metabolism of select alkaloid allelochemicals by native microsomal preparations of *M. rotundata* as measured by LC-MS/MS (1 h incubation at 30°C \pm NADPH). The error bars indicate 95% CIs. Analysis performed using a Welch's t-test (two-tailed) with significant differences in percent depletion of parent compound indicated by ** $p < 0.01$ and *** $p < 0.001$.

Table 5.6: Welch's *t*-test results of percent depletion of parent compound of alkaloid allelochemicals by native microsomal preparations of *M. rotundata* as measured by LC-MS/MS (1 h incubation at 30°C ±NADPH).

Alkaloid	Welch's <i>t</i> -test (two-tailed)			% depletion of parent compound
	p value	Sig p<0.05	df	
Nicotine	0.0014	Y**	2	40
Anabasine	0.0985	N	2	17
Atropine	0.7071	N	2	0
Cytisine	0.0004	Y***	2	55

Table 5.7: Results for univariate analysis of variance using general linear model for metabolism of alkaloid allelochemicals by native microsomal preparations of *M. rotundata* as measured by LC-MS/MS (1 h incubation at 30°C ±NADPH). Tests of between-subject effects with *post-hoc* multiple comparisons of depletion using a Bonferroni adjustment.

Alkaloid	GLM between-subjects test		Sig p<0.05
	p value	F	
Nicotine	<0.001	52.805	Y
Anabasine			
Atropine			
Cytisine			
		Post-hoc pairwise comparison (Bonferroni adjustment)	
Alkaloid 1	Alkaloid 2	Sig p<0.008 ($\alpha/6$)	Sig p<0.008 ($\alpha/6$)
Nicotine	Anabasine	0.008*	Y
	Atropine	<0.001*	Y
	Cytisine	0.084	N
Anabasine	Atropine	0.034	N
	Cytisine	<0.001*	Y
Atropine	Cytisine	<0.001*	Y

5.3.4 Model substrate profiling

5.3.4.1 Temperature optimisation

As shown in figure 5.7, incubation of CYP9R1 with EFC at a range of temperatures (25-40°C) revealed that the optimal temperature for P450 activity was ~25°C. The enzyme activity was markedly lower at the higher temperatures (35° and 40°C). This may be due to the P450 undergoing irreversible thermal inactivation at higher temperatures and changing to a denatured state. At 25°C the reaction rate in the first 60 minutes is still in the linear range and so this was the incubation time period used in all assays.

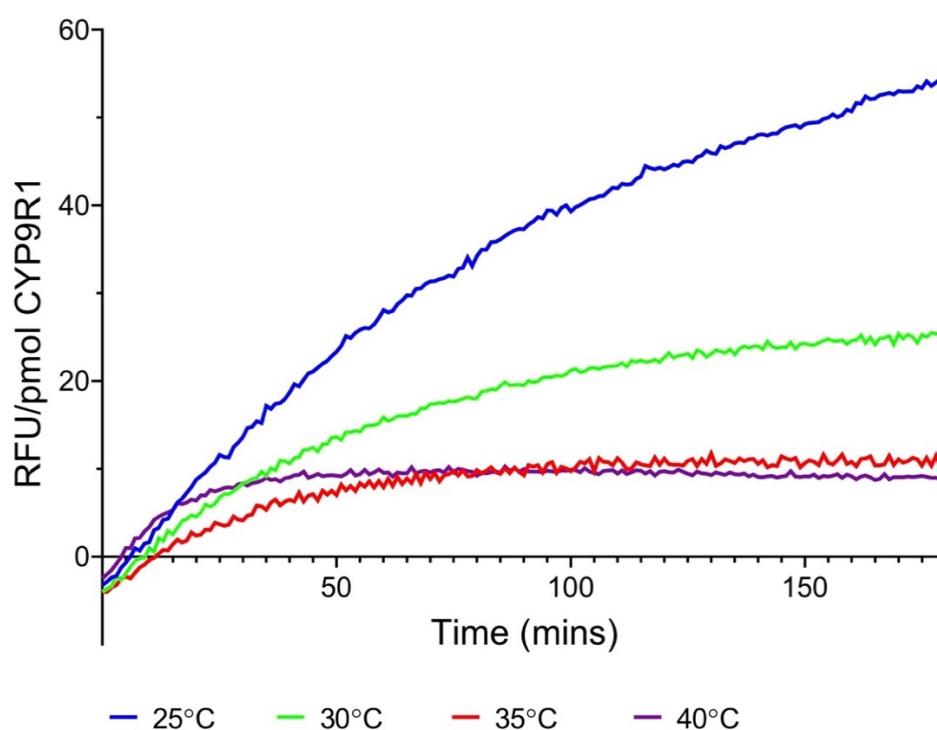


Figure 5.7: CYP9R1 enzyme activity (fluorescence) against the model substrate EFC, incubated at 25, 30, 35 and 40°C, over 180 minutes. Data points are mean values (n=8).

5.3.4.2 Standard curves for model substrates

Due to the range in excitation/emission wavelengths six standard curves were needed (see table 5.8). The standard curve graphs all produced good R^2 values (0.9989-0.9997). The standard curves used can be found in appendix figures 5.8 and 5.9.

Table 5.8: Model substrate standard curves

Standard	Excitation/Emission fluorescence wavelength (nm)	Model substrate	R ²
HC	390/465	EC MC	0.9989
HFC	410/510	EFC	0.9997
HFC	410/535	MFC BFC	0.9995
HFC	405/510	MOBFC	0.9997
HR	535/590	MR ER PR	0.9994
HR	520/590	OOMR BOMR	0.9992

5.3.4.3 Model substrate metabolism assays

5.3.4.3.1 Assays against *M. rotundata* microsomes

M. rotundata microsomes show a preference for coumarin-derived rather than resorufin-derived model substrates (see table 5.2; see figure 5.8). In fact, they showed no metabolic activity against any of the resorufin-derived model substrates used. Against coumarin-derived model substrates there was broad pattern of specificity, with the highest activity shown against the smaller coumarin-based substrates MC and EC (MW: 176.17 and 190.195 g/mol respectively). The highest specific activity was for the smallest molecule, MC. There were significantly lower rates of metabolism against the bulkier (trifluoromethyl)-coumarin substrates (MW: 244.17- 350.29 g/mol; $p < 0.0002$).

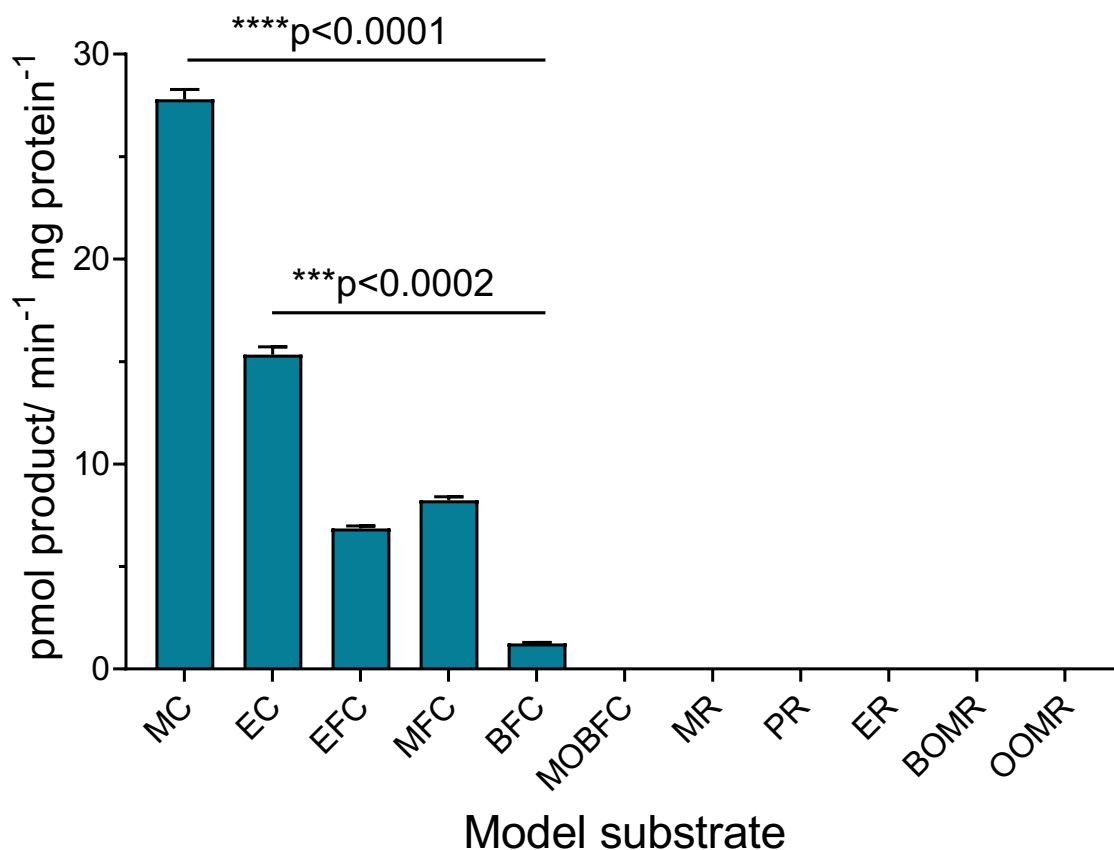


Figure 5.8: Metabolic activity of male *M. rotundata* microsomes against coumarin- and resorufin-derived fluorescent model substrates. Data points are mean values \pm SD (n=3). Analysis performed using a Welch's t-test (two-tailed; degrees of freedom = 2.241) with significant differences indicated by ***p<0.0002 and ****p<0.0001.

5.3.4.3.2 Assays against recombinant CYP9 family P450s

All six recombinant CYP9 P450s were screened in kinetic assays using fluorescent model substrates (see table 5.2). Activity was found in all six, even the two proteins that produced a spectral absorption peak of 420 nm rather than 450 nm (CYP9P2 and CYP9DM2) (see section 5.3.2). The preference for coumarin-derived model substrates seen in the assay using microsomes was also found in the assays using recombinant CYP9 P450s (see figure 5.9 (a) and (c)), and likewise there was no metabolic activity against any of the resorufin-derived model substrates used. The two P450s from the *CYP9R* lineage were the most promiscuous of the tested P450s, each showing metabolic activity against four coumarin-derived substrates (see figure 5.9 (a)

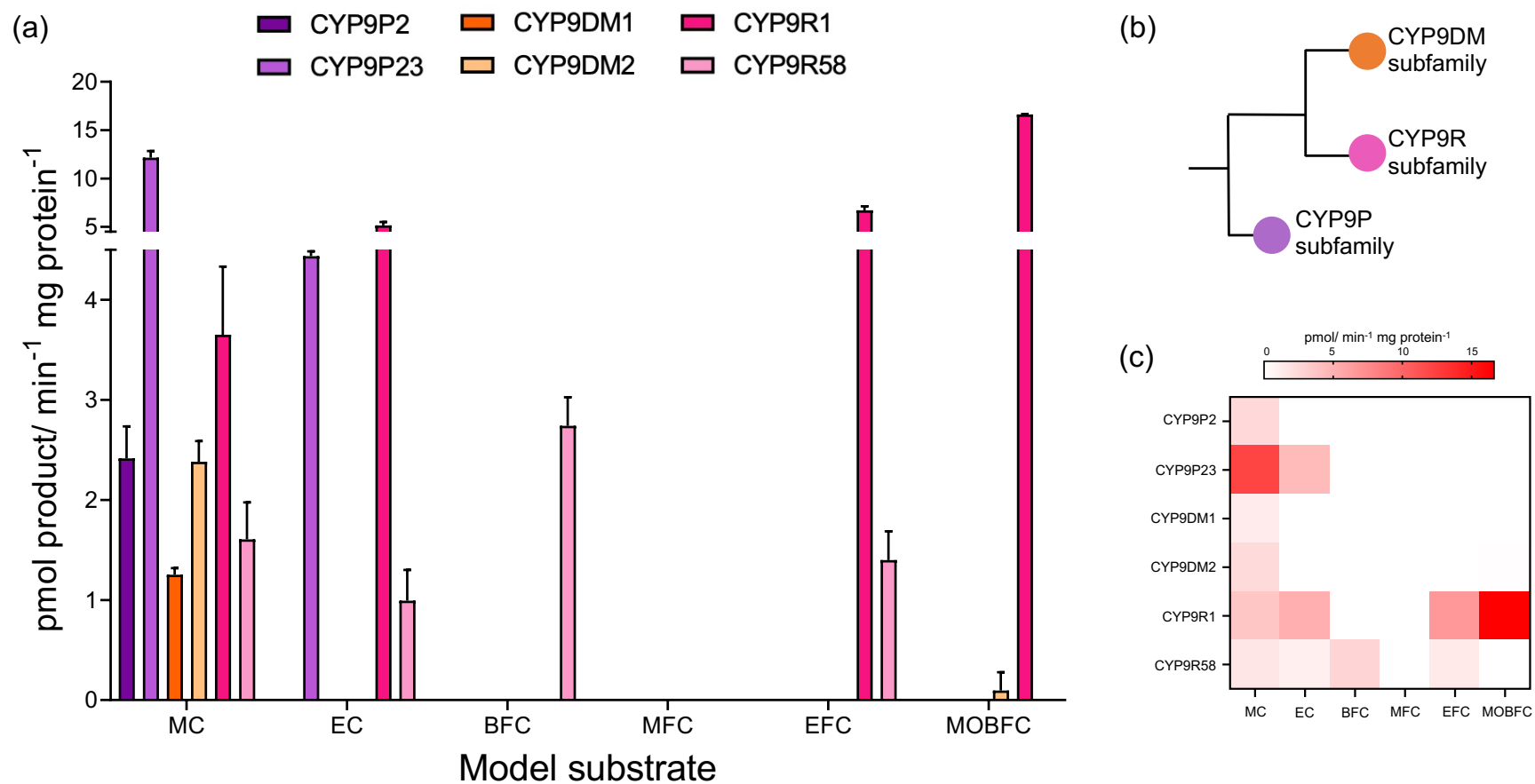


Figure 5.9: (a) Metabolic activity of recombinant CYP9 P450s against coumarin-derived fluorescent model substrates. Data points are mean values \pm SD (n=3) CYP9 P450s are coloured by CYP9 lineage. (b) Schematic of the phylogenetic relationship between the CYP9 lineages. (c) Heatmap showing the metabolic activity of recombinant CYP9 P450s against coumarin-derived fluorescent model substrates.

and (c)). With the exception of very low production of fluorescent product from incubation of CYP9DM2 with MOBFC ($\sim 0.1 \pm 0.15$ pmol product/ min⁻¹ mg protein⁻¹) the P450s from the *CYP9R* lineage were the only enzymes that showed metabolism against any of the larger (trifluoromethyl)-coumarin substrates, EFC, BFC and MOBFC (MW: 258.19, 320.26 and 350.29 g/mol respectively). Both CYP9DM enzymes, which appear to be unique to *M. rotundata*, showed activity against the smallest coumarin MC (MW: 176.17 g/mol). Of the *CYP9P* lineage, CYP9P23 showed metabolic activity against both smaller coumarin-based substrates, whereas CYP9P2 was only active against the smallest model substrate MC. In contrast to the results from the assay using microsomes, none of the expressed CYP9 P450s enzymes showed activity against MFC.

5.3.4.3.3 Fluorescence-based assay to assess insecticide mediated inhibition

Based on model substrate profiling, MC was selected for use against both *O. bicornis* CYP9BU1 and *M. rotundata* CYP9DM2 sequences in insecticide mediated inhibition assays. Thiacloprid had the strongest inhibitory effect on MC metabolism by CYP9BU1 ($p < 0.01$; see table 5.9 and figure 5.10). Indeed, incubation with differing concentrations of both thiacloprid and imidacloprid interfered with HC formation by CYP9BU1 (see table 5.9 and figure 5.10). Conversely, very low levels of insecticide mediated inhibition of HC production were observed in CYP9DM2 with both insecticides (see table 5.9 and figure 5.10).

Table 5.9: Reduction of HC production (%), after incubation (1 h) with the neonicotinoid insecticides thiacloprid (TCP) and imidacloprid (IMI).

CYP9	% Reduction of HC production			
	TCP	p value (Welch's t-test)	IMI	p value (Welch's t-test)
<i>O. bicornis</i> CYP9BU1	32.7	0.0014**	14.1	0.0387*
<i>M. rotundata</i> CYP9DM2	9.2	0.6145	5.5	0.2167

Depletion values compare levels of HC production at 0 μ M insecticide inhibitor to those at 100 μ M. Values used are for 200 μ M fluorescent probe concentration (MC).

None of the reactions followed standard Michaelis-Menten kinetics (see figure 5.10), as evidenced by the lack of a hyperbolic plot. Co-incubation of MC and imidacloprid with both recombinant CYP9s appear to show sigmoidal kinetics, whereas, the plots generated by co-incubation with thiacloprid appear to show biphasic kinetics [528, 559]. V_{max} or K_M values were calculated for the sigmoidal plots and are shown in appendix table 5.2, although they should be considered as indicative only as the reaction did not follow Michaelis-Menten kinetics.

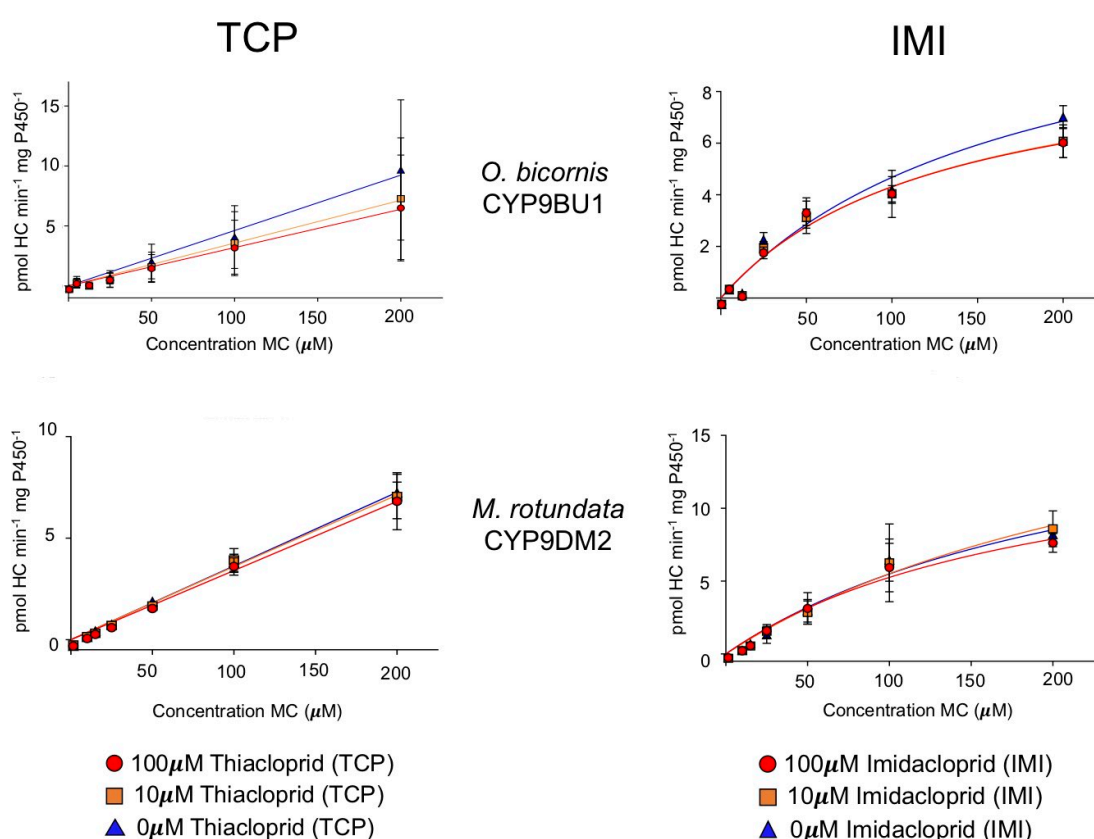


Figure 5.10: Insecticide mediated inhibition of 7-hydroxycoumarin (HC) formation by recombinantly expressed CYP9 enzymes incubated with different concentrations of thiacloprid (TCP) and imidacloprid (IMI). Data are mean values \pm S.D. (n=4).

5.3.5 Binding affinity of insecticides targeting the nicotinic acetylcholine receptor (nAChR)

Brief results for the binding affinity are presented here (see section 5.1.6). All three insecticides used in this study (thiacloprid, imidacloprid and flupyradifurone) were found to bind reversibly with nanomolar affinity (see figure 5.11). All three compounds generated similar IC_{50} values (the concentration of

unlabelled insecticide that displaced 50% of the radiolabelled [^3H]imidacloprid) to those reported for *A. mellifera*. No significant difference was observed in the specific binding of the insecticides at the receptor.

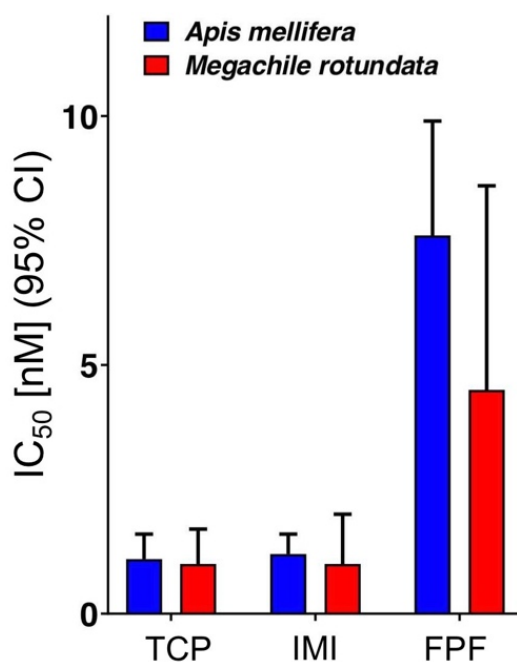


Figure 5.11: Binding affinity (IC_{50} values) of selected insecticides to nAChR head membrane preparations of *A. mellifera* and *M. rotundata*. Data for *A. mellifera* taken from [446]. Figure taken from Hayward *et al.*, 2019.

5.4 Discussion

5.4.1 Preparation of functional microsomes from *M. rotundata*

The presence of a powerful P450 inhibitory factor in the venom sac of *A. mellifera*, which is released in the process of extracting native microsomes, was confirmed by Zaworra and Nauen in 2019 [534]. To investigate whether a similar inhibitory factor is present in the venom sac of *M. rotundata*, native microsomes were extracted from both female (stinger complex present) and male (no stinger complex) bees. A significant difference in the activity against coumarin-based fluorogenic model substrates was observed between the sexes (see figure 5.2). Microsomes extracted from dissected female heads and thoraces also showed significantly more activity to coumarins than those from intact females. This suggests that, in common with *A. mellifera*, *M. rotundata* microsomes extracted from intact female bees would not be functional, due to the presence of a P450 inhibitory factor. This contrasts to work in *O. bicornis*,

where no significant difference in the activity of microsomes prepared from male and female bees was observed [463]. It seems therefore, that the presence of a P450 inhibitory factor in the venom sac is not ubiquitous across bee species. Furthermore, the absence of an inhibitory factor in *O. bicornis* indicates that its presence cannot be predicted by family. Further work, like that done in *A. mellifera* [534], is needed to isolate and characterise the endogenous inhibitory factor, in order to determine whether phospholipase A₂ is also present in *M. rotundata* venom. However, in solitary bee species, where the sex ratio allows, it would be prudent to use males in the extraction of native microsomes.

5.4.2 Metabolic capability of native microsomes with analysis using LC-MS/MS

5.4.2.1 Synthetic insecticides

In contrast to other managed bee pollinators, *M. rotundata* shows a clear *in vivo* sensitivity to synthetic insecticides (see section 4.3.2). Of particular note is the lack of a differential sensitivity to *N*-cyanoamidine and *N*-nitroguanidine neonicotinoid insecticides. In other bee species, P450 enzymes from the *CYP9Q/BU* lineages have been shown to provide protection to insecticides across three MoA classes, including *N*-cyanoamidine neonicotinoids. However, *M. rotundata* lacks a *CYP9Q/BU* ortholog (see section 3.3.6). This raises the question: Does *M. rotundata* lack enzymes capable of metabolising the compounds known to be detoxified by the *CYP9Q/BU* lineage? To this end, the metabolic ability of *M. rotundata* native microsomes was investigated with the insecticides known to be targeted by *CYP9Q/BU* enzymes: thiacloprid, acetamiprid, flupyradifurone, tau-fluvalinate and coumaphos. As with the acute contact bioassays, a challenge to the metabolic ability of *M. rotundata* microsomes was also posed by an exemplar of each MoA class known to be highly toxic to bees (imidacloprid, deltamethrin and chlorpyrifos) (see section 4.2.2).

LC-MS/MS analyses showed that, of the compounds known to be metabolised by the *CYP9Q/BU* lineage, only coumaphos exhibits any significant reduction in parent compound (see figure 5.5). Although it should be noted that the low number of replicates, required for LC-MS/MS analyses, generated a small sample size, which by default, reduced the statistical power of the t-test used. In

the acute contact bioassay tests against *M. rotundata* in this project, coumaphos was the least toxic, with an LD₅₀ (48 h) of 0.557 µg a.i./bee (see table 4.7). However, that value is still well below the EPA 2 µg a.i./bee threshold for classification as a highly toxic substance (see table 4.1) [493].

Nevertheless, it is interesting that there appears to be a level of metabolism for both organophosphates. Substrate specificity in P450s is often broad, where one enzyme is capable of binding multiple compounds (such as CYP9Q3 in *A. mellifera* [168] and CYP3A4 in *H. sapiens* [225]). However, some P450s have overlapping specificities, where different enzymes are capable of binding the same substrate [560]. Indeed, the broad and overlapping nature of substrate interactions in P450 enzymes precludes their naming by the reactions they catalyse [560]. In *O. bicornis*, the most significant depletion of coumaphos was found to be due to CYP336A36 rather than CYP9BU1 or CYP9BU2 [463]. Native microsomes contain multiple P450 and non-P450 enzymes and so metabolism of a compound cannot be attributed to an individual enzyme. It is possible that, in common with *O. bicornis*, one of the CYP336A enzymes found in *M. rotundata* (or an entirely different enzyme) has some metabolic ability against coumaphos. Likewise, metabolism of chlorpyrifos was seen by CYP9R1 in *O. bicornis*, although this did not prevent the insecticide from being rated highly toxic in acute contact bioassays [463].

There was no metabolism observed of any of the neonicotinoid, butenolide or pyrethroid insecticides by *M. rotundata* microsomes. None of the insecticides known to be metabolised by the CYP9Q/BU lineages (i.e. thiacloprid, acetamiprid, tau-fluvalinate and flupyradifurone) showed any significant depletion in parent compound. Quantifying the depletion of parent compound, as described above, is a simple method that yields an estimate of intrinsic metabolism [527]. For a more comprehensive understanding of insecticide metabolism, measurement of metabolite production is required. To allow this to happen, the metabolites for the insecticides in question must have been identified and synthetic standards prepared [527]. Standards are available for the neonicotinoids, flupyradifurone and deltamethrin, but not for tau-fluvalinate or the organophosphates. Where measurement was possible (i.e. thiacloprid, imidacloprid, acetamiprid, flupyradifurone and deltamethrin) there was some production of hydroxylated (OH) metabolites, which indicates a low-level of

metabolism (see appendix table 5.3). However, the compounds known to be highly toxic, such as imidacloprid and deltamethrin, also show a level of metabolite production. As such the LC-MS/MS data need to be interpreted holistically, taking parent compound depletion and production of metabolites into account. Metabolite production alone does not indicate successful detoxification of an insecticide to a safe level *in vivo*. It should also be considered that metabolites of insecticides cannot be assumed to be less toxic than their parent compounds. For example, whereas 5-hydroxy-thiacloprid is classified as practically non-toxic to *A. mellifera* [168], 5-hydroxy-imidacloprid is highly toxic (0.159 µg a.i./bee) and olefin more toxic than imidacloprid itself (0.036 vs 0.041 µg a.i./bee) [446].

One explanation for the *in vivo* lack of tolerance to insecticides observed in *M. rotundata* is an increased affinity of these compounds at the target receptor. Neonicotinoid and butenolide insecticides target the nAChR and, as such, the binding affinity of thiacloprid, imidacloprid and flupyradifurone was investigated by colleagues at Bayer (AG, Crop Science, Leverkusen, Germany). As can be seen in section 5.3.5, *M. rotundata* head membrane preparations generated similar IC₅₀ values to those reported for *A. mellifera*. The lack of tolerance to *N*-cyanoamidine neonicotinoid insecticides and flupyradifurone, in *M. rotundata*, cannot therefore be explained by an enhanced affinity of these compounds at the receptor site.

From the LC-MS/MS and radioligand binding study data combined, the most likely explanation for the high sensitivity observed in *M. rotundata*, is an absence of enzymes capable of metabolising the insecticides used. Specifically, the lack of a *CYP9Q/BU* ortholog is correlated with an inability to metabolise the insecticides targeted by those lineages in other bees (i.e. *N*-cyanoamidine neonicotinoids, flupyradifurone, *tau*-fluvalinate and coumaphos). The absence of metabolism by native microsomes indicates that, not only does *M. rotundata* lack a *CYP9Q/BU* ortholog, it also has not evolved alternative P450s, or non-P450 detoxification enzymes, that are capable of performing similar functions. It also rules out metabolism by *M. rotundata* P450s acting in concert to metabolise these insecticides [534].

In addition to the seven insecticides used in the acute contact bioassays, the metabolism of the diamide, flubendiamide, by *M. rotundata* native microsomes was also investigated using LC-MS/MS. There was a reduction in the parent compound found in flubendiamide, although the depletion was at a fairly low level (<18%). One repetition of a range-finding test using two diamide insecticides (flubendiamide and chlorantraniliprole) against male *M. rotundata*, was performed in July 2019, with a view to informing bioassay concentrations for use in 2020. Unfortunately, due to the pandemic, no acute contact bioassay work was performed in 2020. However, the initial results of the range-finding tests indicate that it is unlikely that flubendiamide and chlorantraniliprole would be classified as highly toxic to *M. rotundata* (see appendix table 5.4).

In drug-drug interaction (DDI) studies in human medicine, assays using liver microsomes are considered to be the *in vitro* 'gold-standard' test [561]. Human liver microsomes are employed in the DDI studies required for drug regulatory purposes, as they are considered to closely mimic the *in vivo* situation [561]. The initial aim of this part of the PhD was to look for a correlation between enzymatic function and *in vivo* sensitivity to insecticides in *M. rotundata*. The data generated by the insecticide metabolism assays using native microsomes was sufficient to establish that P450 enzymes present in the microsomes of *M. rotundata* are inefficient metabolisers of most of the insecticides tested in this chapter. The individual expressed recombinant CYP9 P450s were therefore not screened for insecticide metabolism using LC-MS/MS. However, ideally, for absolute clarity and certainty about the metabolic abilities of the CYP9 subfamily in *M. rotundata*, this work should be undertaken at some point in the future.

There are two main questions that arise from these data. Firstly, is the current regime of toxicology regulation and risk assessment sufficient to protect *M. rotundata*, particularly with reference to its use as a managed pollinator of agricultural crops? Secondly, is *M. rotundata* the only species of bee that lacks a *CYP9Q/BU* ortholog? The first of these questions is perhaps the easier one to answer. The data generated in chapters four and five of this PhD strongly indicate that the use of *A. mellifera*, *B. terrestris* or *O. bicornis* as surrogates for *M. rotundata* in toxicological risk assessment, across four MoA classes, is

unreliable. This has immediate and significant implications for the safe use of insecticides on the crops pollinated by *M. rotundata*. However, it also has the potential for wider significance, which leads us to the second question raised: How wide-spread is the lack of this lineage of P450s? The lack is obviously not family-wide as *CYP9Q* and *CYP9BU* genes are present in the Apidae and Megachilidae, while the Halictidae *D. novaeangliae* has four closely related *CYP9DL* genes (see section 3.3.6). Nonetheless, even should this loss only prove to be genus-wide, the *Megachile* genus contains ~1,500 species (Wedmann, 2009), and alone accounts for approximately 7.5% of all bee species. The lack of published bee genomes means the search for CYP9 subfamily P450s should be widened to include an interrogation of the NCBI transcriptomics database, which holds information from over 150 assemblies, covering six of the seven bee families. Further transcriptomics sequencing of other *Megachile* species would also add data on the genus. These approaches are addressed in chapter six of this PhD.

5.4.2.2. Plant allelochemicals

Plants produce a wide variety of allelochemicals that act as part of a defence system against herbivorous insects [184, 562]. Alkaloids are one of the most important groups of allelochemical, containing an enormous variety of compounds, including many substances known to be highly toxic such as nicotine, morphine and strychnine [183].

In contrast to the lack of metabolism observed with synthetic insecticides, *M. rotundata* native microsomes show good depletion of the naturally occurring plant allelochemicals nicotine and cytosine (see figure 5.6). Significant depletion of parent compound was observed in nicotine and cytosine with 40% ($p < 0.01^{**}$) and 55% ($p < 0.001^{***}$) reduction respectively. There was also a low-level of parent compound reduction observed in the nicotine-related, anabasine (~17%). However, no metabolism was observed in the muscarinic acetylcholine receptor agonist, atropine (see figure 5.6). A synthetic standard for the nicotine metabolite, cotinine, is available and so cotinine production was also measured using LC-MS/MS. Cotinine was produced ($\sim 32 \text{ ng/ml}^{-1} \text{ mg protein}^{-1}$), further supporting the metabolism of nicotine by *M. rotundata* microsomes (see appendix table 5.3).

Nicotine is a pyridine alkaloid found in the Solanaceae family of plants, which includes tobacco (*Nicotiana tabacum*) and certain food crops from the nightshade family, such as tomatoes and potatoes [117]. It targets the nAChR, giving it a similar mode of action to neonicotinoid insecticides, and it has a long history of use as an insecticide [117, 152]. In human medicine, research into smoking cessation has determined that CYP2A6 is the most important enzyme involved in nicotine metabolism [563, 564]. In vertebrates 70-80% of nicotine is metabolised in the liver and excreted with its metabolites in the urine [565]. Few insect species are known to tolerate nicotine in their diet, and those that can are most often pest species, such as *Myzus persicae* (peach-potato aphid) [566], or *Bemisia tabaci* (sweet potato whitefly) [567]. CYP3 clade P450 enzymes have been implicated in the metabolism of nicotine by these pest species; CYP6CY3 in *M. persicae* [566], and CYP6CM1 in *B. tabaci* [567].

Interestingly, *A. mellifera* appears to have the ability to metabolise nicotine and subsequently has a level of tolerance to the compound in its diet [562, 568]. Under experimental feeding conditions, using sucrose solution containing nicotine, the level of nicotine found in honey was 90% lower than that added in the artificial diet [568]. Although the enzymes responsible for the metabolism of have not yet been identified, constitutively expressed P450 enzymes are thought to be the most likely candidates [562]. In *O. bicornis* LC-MS/MS analyses showed ~90% reduction in nicotine concentration after incubation with both native microsomes and with functionally expressed CYP6AQ55 [463]. It may be that, at least in *O. bicornis*, the CYP6AQ lineage of P450s provides a level of tolerance to dietary nicotine. *O. bicornis* and *A. mellifera* have a single CYP6AQ gene (see section 3.3.2.1), which contrasts with the expanded repertoire of CYP6AQ genes found in *M. rotundata* (three members) and *B. terrestris* (six members) (see figure 3.6). Heterologous expression of *M. rotundata* CYP6AQ enzymes would be needed to discover whether this lineage is responsible for the metabolism of nicotine observed in LC-MS/MS analyses with native microsomes.

Anabasine is found in the tree tobacco plant (*Nicotiana glauca*) which is a close relative of *N. tabacum*. Like nicotine, anabasine targets the nAChR and it has

been used as an insecticide [569] and a biomarker in smoking cessation studies in human medicine [570]. In humans, the P450 that is active against nicotine, CYP2A6, has also been linked with anabasine metabolism, using inhibition assays with nicotine and a coumarin-based fluorogenic model substrate [571]. In spite of the structural similarity of nicotine and anabasine the amount of nicotine depleted by *M. rotundata* microsomes is more than double that of anabasine (~40% versus ~17%). A similar difference was observed between nicotine and anabasine depletion by *O. bicornis* microsomes (~95% versus ~40%) [463]. However, CYP6AQ55 did not show any metabolic ability against anabasine in *O. bicornis* [463]. It seems that in contrast to what is found in humans, a separate P450, or non-P450 enzyme is responsible for the metabolism of anabasine in *O. bicornis* and this, therefore, may also be the case in *M. rotundata*.

Of the four alkaloids used, cytisine showed the most significant depletion by *M. rotundata* microsomes. Cytisine is a quinolizidine alkaloid found in the Fabaceae or Leguminosae family [194]. This family of plants includes legumes, such as lupins (*Lupinus* species) and birds-foot trefoil (*Lotus cornicularis*) known to attract bees of the *Megachile* genus [572-574]. Cytisine is considered to be a partial agonist of the nAChR [575]. In human medicine it is widely used in the treatment of nicotine addiction and is licenced for this purpose in parts of Europe and Asia [576]. However, there is limited data available on the pharmacokinetics of cytisine in vertebrates and insects [576, 577]. Animal and human studies have shown that cytisine undergoes only minimal metabolism, with 90-95% of the dose being excreted unchanged in the urine [576, 578]. The results here imply that this is not the case in *M. rotundata*, as there is significant depletion (55%) in parent compound by native microsomes, indicating metabolism. As a genus, *Megachile* bees are known to have a preference for plants from the Fabaceae family [573, 574]. Specifically, *M. rotundata* is known to have a strong preference for birds-foot trefoil (*L. cornicularis*) and crown vetch (*Coronilla varia*) [572]. Although neither of these legumes is linked with a high cytisine content, they do produce other allelochemicals known to be toxic, such as linamarin, lotaustralin and coronillin [579, 580].

Together, the results from the LC-MS/MS analyses of alkaloid metabolism assays indicate two things. Firstly, they give comfort that the protocols used in this section are correct, as one explanation for the lack of metabolism seen for synthetic insecticides could be the use of unsuitable methodologies. Secondly, and more importantly, they suggest that *M. rotundata* has evolved species-appropriate detoxification pathways (involving P450 and non-P450 enzymes) to metabolise the diverse array of xenobiotics naturally encountered in its diet and ecology (i.e. plant allelochemicals in nectar, pollen and leaf materials). In contrast to other bee species, however, these existing detoxification pathways are not recruited to protect *M. rotundata* against insecticides [168].

5.4.3 Model substrate profiling

The metabolic activity of *M. rotundata* native microsomes and functionally expressed CYP9 P450s was explored using a range of fluorescent coumarin- and resorufin-based model substrates. The CYP9 P450s screened were: CYP9P2, CYP9P23, CYP9R1, CYP9R58, CYP9DM1 and CYP9DM2, and as such, included representatives from the three lineages found in the CYP9 cluster (see section 3.3.3). No functional activity was recorded against any of the resorufin-based substrates by either the microsomes or any of the expressed CYP9s (see figures 5.8 and 5.9). This is different to the metabolic profile observed in *O. bicornis*, where low-level activity by native microsomes was recorded against several resorufin-based model substrates [463].

All six *M. rotundata* CYP9s showed activity against 7-methoxy coumarin (MC), and three were also active against 7-ethoxy coumarin (EC), the two smallest coumarin-based model substrates (see figure 5.9 (c)). The CYP9R enzymes exhibited the broadest specificity, each showing activity against four coumarin-based substrates (66.67%). Conversely, CYP9P2 and CYP9DM1 had the narrowest specificity and only showed activity against MC. These results are indicative of two things. Firstly, it seems that the active binding pocket in many *M. rotundata* CYP9s is capable of binding and metabolising multiple substrates. Secondly, there is overlapping substrate specificity between the different CYP9s. Indeed, all six functionally expressed CYP9s are able to bind and metabolise MC.

There is a paucity of published material which include data on model substrate profiling of P450s in bee species. However, due to the importance of the *CYP9Q/BU* lineages there are data for these enzymes in *A. mellifera*, *B. terrestris* and *O. bicornis*. All three CYP9Q enzymes in *A. mellifera* exhibit broad substrate specificity, showing activity to five out of six coumarin-based model substrates tested, with CYP9Q3 showing activity against all six substrates [168]. Both *O. bicornis* CYP9BU enzymes showed a similar pattern of specificity to that seen in *A. mellifera*, with activity against five of the six model substrates [463]. Conversely, the two *B. terrestris* CYP9Q enzymes screened showed narrower specificity. CYP9Q4 was only active against two, and CYP9Q5 against three, of the six model substrates [168]. As a final comparison, we have the phylogenetically and structurally distinct CYP9DM enzymes in *M. rotundata*. CYP9DM1 is only active against one and CYP9DM2 against two of the six model substrates. There is a remarkable difference in breadth of specificity reported between the two Apidae species, which makes forming any conclusions across the four managed bee pollinators difficult. It may be that the protocol needs to have stricter controls, such as the type and sensitivity of plate reader, before useful comparison between studies can be made.

The recent publication of a protocol for a fluorescence-based assay to assess insecticide mediated inhibition using expressed recombinant P450s from *A. mellifera* (see Haas and Nauen, 2021), opens the real possibility of developing high-throughput *in vitro* tests to screen new pesticidal chemical and to predict interactions between chemotypes in bee species [549, 550]. In the insecticide mediated inhibition assays performed using *O. bicornis* CYP9BU1 and *M. rotundata* CYP9DM2, both P450s exhibited non-Michaelis-Menten kinetics (sigmoidal and biphasic plots) (see figure 5.10). It is well established that many P450 mediated reactions follow non-Michaelis-Menten (atypical) kinetics [528, 561, 581]. Generally, this is thought to be due to the ability of the P450 to simultaneously bind more than one substrate molecule [561, 581]. The binding of multiple ligands produces changes in the kinetic parameters. For example, autoactivation, an effect that is known to produce sigmoidal kinetics plots [528, 561, 581]. The broad substrate specificity of certain P450s, such as *H. sapiens* CYP3A4, is thought to be facilitated by large active sites and well-defined binding pockets that can accommodate two or more substrate molecules [582].

The *CYP9Q/BU* lineage is known to contain P450s with broad substrate specificity, capable of metabolising insecticides from four MoA classes. The atypical sigmoidal and biphasic plots generated by co-incubation of 7-methoxycoumarin and insecticide, with recombinant CYP9BU1 and CYP9DM2, may be explained by both substrates being bound at the same time.

The insecticide mediated inhibition assays were run over a timescale of one hour, and so, the rate of inhibitor turnover (i.e. the extent to which the insecticide is metabolised) should be considered [561]. A significant reduction in the parent compound might impact the endpoint readings taken. It should also be noted that any metabolites produced could also act as inhibitors, perhaps showing stronger effects than the parent compound [561]. Further work is needed to be able to account for what impacts these type of effects have when comparing results across enzymes and species.

However, although the kinetics exhibited in these insecticide mediated inhibition assays are atypical, there is a lack of HC depletion in those run against CYP9DM2 when compared to CYP9BU1 (see table 5.9). This implies that CYP9DM2 does not bind or metabolise either of the tested neonicotinoids effectively, something that supports the LC-MS/MS data generated from incubations of *M. rotundata* native microsomes (see section 5.3.3.1), and the *in vivo* acute toxicity bioassay results (see section 4.3 2.1). For *O. bicornis* these results indicate that CYP9BU1 binds and metabolises thiacloprid more strongly than imidacloprid, a finding that is in keeping with LC-MS/MS data from other studies [226, 463].

In conclusion, there is a correlation between the lack of a *CYP9Q/BU* ortholog and an inability to metabolise certain insecticides in *M. rotundata*. This raises a real possibility of using phylogeny to predict function, thereby allowing the creation of a bee ‘tool-kit’ that can be applied to pesticide screening in order to inform product regulators. This could be envisioned as a step-by-step process, or pipe-line, involving phylogeny, structural modelling, MSAs and functional assays. This type of approach could form the basis of a framework that applies a toxicogenomics approach to assess the likelihood of sensitivity, or tolerance, to insecticides in a bee species.

Chapter six: Evolutionary analyses of the CYP9 subfamily of P450s across the Hymenoptera, with focus on the Anthophila

6.1 Introduction

6.1.1 Evolution of the Hymenoptera

The Hymenoptera are one of the most successful and diverse orders of insects with more than 153,000 described and perhaps up to one million undescribed extant species [583-586]. The earliest Hymenoptera fossils date to the Triassic period, but age estimates from molecular data put the origin of the order in the late Carboniferous (~311 million years ago (mya)) [584, 587]. The order underwent an explosive radiation during the late Jurassic and Cretaceous periods [583]. Most of the extant families had appeared by the mid-late Cretaceous, see figure 6.1 [583]. Extant Hymenoptera fill a wide variety of important ecological niches in most terrestrial ecosystems, with species functioning as predators, pollinators, decomposers and parasites [584, 587]. Each species will have evolved the complex suite of physiological and behavioural adaptations needed to fill its ecological role. The order is divided broadly into the 'broad-waisted' Symphyta (sawflies) and the 'wasp-waisted' Apocrita [583, 585]. Sawflies are mainly phytophagous, so it is thought the ancestor of the Hymenoptera was likely to be phytophagous, feeding on pollen, shoots and leaves [583].

Parasitism appears to have evolved only once (approximately 247 mya, see node marked in red in figure 6.1), appearing in the common ancestor of parasitic wood-wasps (Orussoidea) and the Apocrita [583, 585]. The evolution of the wasp-waist, a constriction between the first and second abdominal segments, allowed for a dramatic increase in the manoeuvrability of the rear end of the wasp, especially the ovipositor [588]. This major innovation undoubtedly contributed to the massive success of the Parasitica, with the Ichneumonidea and Chalcidoidea together accounting for well over 60,000 extant species (see figure 6.1) [587, 588]. The early radiation of the Parasitica also coincides with the diversification of their host lineages [585, 589].

The most likely ancestor of the Aculeata was an ectoparasitoid wasp that attacked and paralysed a concealed host before laying an egg in, or on, the

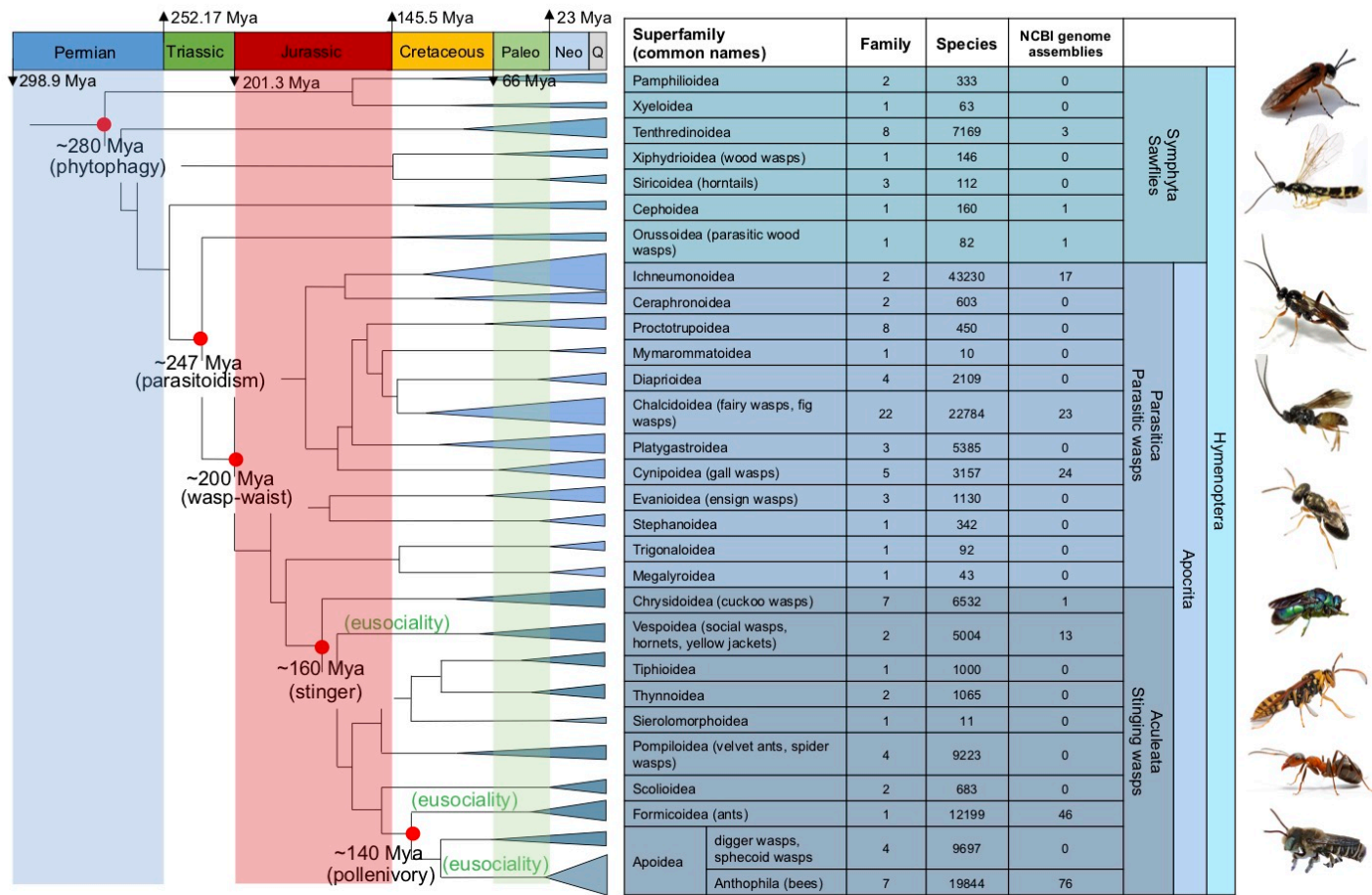


Figure 6.1: The evolutionary history of the Hymenoptera, showing phylogenetic relationships, branches not to scale. Estimates of divergence times associated with key evolutionary events are indicated at their respective nodes [585, 587]. Triangular branches indicative of extant species not drawn to scale. NCBI genome assembly numbers as of December 2020. Photos used: *Athalia rosae* [590], *Cephus cincta* [591], *Cotesia vestalis* [592], *Nasonia* sp. [593], *Microplitis demolitor* [594], *Chrysis* sp. [595], *Polistes rothneyi* [596], ant [597].

body [598]. In certain lineages, the Vespidae, Formicoidea and Apoidea, the behaviours associated with parasitoidism evolved and species adopted a more active predatory way of life [583, 598]. These changes coincided with other life-history changes such as nest building, prey transportation and increasing levels of parental care [598]. Eusocial behaviours also arose on multiple occasions within these lineages [598-600].

A further novel dietary specialism, the consumption of pollen (pollenivory) rather than insect prey, arose in the pollen wasps (Vespidae: Masarinae) and bees (Anthophila) (see figure 6.1) [72, 601, 602]. Bees rely almost exclusively on floral resources with nectar (carbohydrates), pollen (protein) and oils providing their principle dietary needs [603]. However, there is evidence that some species supplement this with foods from other sources, for example certain stingless bees (Apidae: Meliponinae) visit fruits, collect aphid secretions and plant exudates [604]. Three species of neotropical stingless bees from the *Trigona* genus have completely abandoned floral resources and rely on obligate necrophagy, feeding on carrion [604-606]. One of these necrophagic species, *Trigona hygokea*, has also been documented preying on live larvae in wasp nests [606]. This behaviour represents the culmination of the evolutionary process, from the initial switch from basal Hymenopteran phytophagy, to carnivory in the Parasitica, followed by the reversal to phytophagy in the pollinivorous Anthophila, and finally a re-reversal back to carnivory [604, 606]. This evolutionary transition to predation reflects the malleability of Hymenopteran traits over evolutionary time, to fill ecological niches, rather than plasticity of behaviour [606].

Genomic resources for the Hymenopteran order are patchy, and there is not even coverage, see figure 6.1. Many superfamilies have not had any assembled genomes released, whereas others, such the Apoidea (containing bees) and the Formicoidea (ants) are comparatively well represented.

The ancestral method of sex determination in the Hymenoptera is haplodiploidy, or arrhenotokous parthenogenesis [607, 608]. In this system, males develop from unfertilized eggs and females develop from fertilized eggs. Whilst this system allows for parental control over offspring sex ratio and benefits the

mother, in that her sons will always pass her genes to their offspring, haploid individuals are expected to be less viable [607, 609]. Diploid organisms can have sizeable numbers of deleterious recessive alleles in their genomes and survive, whereas the presence of these in a haploid individual might leave it unlikely to survive [607, 609]. Over time however, this could allow selection to eliminate deleterious mutations and fix beneficial mutations, thereby offsetting the fitness costs to the species [609]. This process is likely to lead to the build-up of beneficial alleles, and is something that could become marked in species that have lower recombination rates in the coding regions of their genomes, something that is found in certain bee species [467-469, 609].

In general, Hymenoptera are seen as beneficial to humans, and apart from some sawflies and seed-feeding chalcids there are few examples of agricultural pests among the Hymenoptera [583, 610]. In fact, due to the range of host species, many parasitoid wasp species are used in the biological control of arthropod pests in agriculture and forestry [611]. However, globally there are more than 10 million stings to humans every year from the Aculeata, and approximately 3% of adults and 0.4-0.8% of children experience a life-threatening reaction to the venom [586].

6.1.2 P450 evolution

The evolution of the P450 superfamily appears to be highly dynamic by nature. One gauge of the evolutionary pattern is the number of genes that make up the CYPome of a species [461, 612]. However, as discussed in chapter three of this PhD, this is not the only measure that should be used, and diversity of P450s by clan, family, subfamily and lineage should also be taken into account when looking for insights into CYPome evolution [218]. Comparative genomics has shown that there is extensive macro-synteny across vertebrate genomes [612, 613]. There have been several studies that have investigated conserved synteny of various gene families in insects. For example, the *metallothionein*, *trehalase* and the *Osiris* gene families [614-616]. The P450 superfamily is considered as a model for gene family evolution, and the sequential tandem duplications that lead to clusters and blooms of CYP genes, make them ideal candidates for syntenic comparison between genomes of closely related and more distant species [461, 466]. Synteny analysis of P450 clusters in

Lepidopteran species reported a high number of syntenic breaks and evidence of frequent rearrangements [466]. There is an absence of similar syntenic analysis of P450 clusters in Hymenopteran species.

6.1.3 RNA sequencing and analysis

Molecular biology studies the flow of information from genes, stored in DNA, via transcription into RNA, to translation into proteins. The transcriptome consists of all RNA transcripts, coding and non-coding, that are expressed in a cell or tissue. Next generation sequencing (NGS) technology, such as Illumina, enable RNA analysis by synthesising complementary DNA (cDNA) from a mRNA template and attaching adaptors to one or both ends of these cDNA fragments [617]. The double-stranded cDNA molecules are then sequenced. In the case of Illumina sequencing this is achieved using fluorescently labelled dNTPs, with each of the four bases: adenine (A), cytosine (C), guanine (G) and thymine (T) having a unique colour. This process is commonly referred to as RNA-seq (RNA sequencing). RNA-seq has low error rates (<1%) and can be used for *de novo* transcriptome assembly [617, 618].

6.1.4 Chapter aims and underpinning questions

In *A. mellifera*, *B. terrestris* and *O. bicornis*, members of the CYP9Q and CYP9BU lineages are involved in the metabolism of insecticides from three MOA classes (see sections 1.4.4.1 and 1.7.3) [137, 155, 168, 226, 446]. *M. rotundata* does not have a CYP9Q/BU ortholog, and this correlates with a high sensitivity to certain insecticides (see sections 4.3.2 and 5.3.3). Given this correlation it is imperative we establish how wide-spread a lack of CYP9Q/BU-like orthologs might be across bee species. This chapter will initially focus on delineating the genes that make up the body of CYP9 subfamily across bee families, and using comparative genomics to determine whether there is a predictive pattern.

O. bicornis and *M. rotundata* are both Megachilidae species, so it is clear that the lack of a CYP9Q/BU ortholog is not family-wide. However, there is a lack of information on the metabolic ability of the CYP9 genes in other Megachilidae species. Using phylogenetic analyses, this chapter aims to select candidate CYP9BU-like genes from as many Megachilidae species as possible, with a

view to functionally expressing them. This will allow the exploration of the metabolic profile of these enzymes. Specifically, fluorescence-based insecticide mediated inhibition assays are used to infer evidence of differential metabolism of *N*-cyanoamidine and *N*-nitroguanidine neonicotinoids, such as that observed in *O. bicornis* CYP9BU1.

To help reliably predict the presence of *CYP9Q/BU-like* sequences across all bee families, understanding the ancestral origins of the lineage is key. To achieve this, P450 sequences from species of other Hymenopteran superfamilies were used in phylogenetic and syntenic analyses, to look for evidence of conservation in the CYP9 subfamily. This enables the evolution of the *CYP9Q/BU* lineage to be examined, to determine whether it is bee-specific, or whether other Hymenopteran species have genes that share a common ancestor. This chapter will also look for evidence of whether the genomic landscape of the CYP9 cluster found in bees is a conserved artefact that predates the divergence of the Apoidea.

There is a lack of genomic and transcriptomic data from other species from the *Megachile* genus. There is currently only data on one other species, *M. willughbiella*, held in the NCBI databases (data accessed July 2020). To fully answer the question of how wide-spread the lack of a *CYP9Q/BU* ortholog is, the obvious starting point is other *Megachile* species. To that end, this chapter describes the curation of the CYP9 subfamily, from the transcriptomic assemblies of *M. willughbiella* and three further *Megachile* species. This will determine the presence/absence of either a *CYP9Q/BU* or *CYP9DM* ortholog in these species.

The following key questions will be addressed:

Firstly, how wide-spread is the presence of insecticide-degrading P450 enzymes across bee families and genera?

Secondly, can comparative genomics, in combination with P450 synthesis and targeted *in vitro* functional analyses, be used as the basis for a predictive tool that can be applied to create a robust insecticide risk assessment framework for bees?

Thirdly, given that the CYP9 cluster shows a high degree of conserved synteny across Apidae, Megachilidae, Halictidae and Colletidae species (see section 3.3.3.4), can we trace the ancestry of this key subfamily of detoxification genes back through the Hymenoptera Order?

6.2 Methods

6.2.1 Surveying the available CYP9 subfamily sequences of bee species

6.2.1.1 Searching the NCBI genomic database

The NCBI genome database was searched for published genomes of 'Apoidea (bees)'. The nucleotide sequences for *A. mellifera* CYP9Q3, CYP9P1, CYP9R1 and *M. rotundata* CYP9DN1 were used as the query sequences in a BLASTn search through the assembly of the genome of each bee species to find CYP9 homologs. All resulting hit tables were downloaded. Scaffolds containing CYP9 P450s were downloaded as a GenBank (full) file and imported into Geneious version 10.2.3 (Biomatters). Unannotated CYP9 sequences were found using the 'find in document' tool and the BLASTn alignment results. All CYP9 sequences were annotated in the Geneious file. A folder of CYP9 sequences for each bee species was created and each nucleotide sequence was translated and manually inspected for the presence of conserved motifs four and five (M4 & M5). Partial sequences and those that contained stop codons were removed. Unless already named, CYP9 sequences were named numerically for each species (eg. *Lasioglossum albipes*: La_CYP9-like1, La_CYP9-like2).

6.2.1.2 Searching the NCBI transcriptomic database

To discover which species of bee were represented in the NCBI Transcriptome Shotgun Assembly (TSA) database, the nucleotide sequence for *A. mellifera* CYP9Q3 was used as a query sequence in a BLASTn search, gated by bees (taxid:34735), through the TSA database, see appendix table 6.1. The hit table was downloaded and used as a reference of bee species for use in defining the limits of subsequent BLASTn searches. The nucleotide sequences for *A. mellifera* CYP9Q3, CYP9P1, CYP9R1 and *M. rotundata* CYP9DN1 were used as the query sequences in a BLASTn search through the TSA database, limited by each bee species in the hit table from the initial BLASTn search (see

appendix table 6.1), to find CYP9 homologs. All resulting hit tables were downloaded. Scaffolds containing CYP9s were downloaded as a GenBank (full) file and imported into Geneious version 10.2.3 (Biomatters). Unannotated CYP9 sequences were found using the 'find in document' tool and the BLASTn alignment results. All CYP9 sequences were annotated in the Geneious file. A folder of CYP9 sequences for each bee species was created and each nucleotide sequence was translated and manually inspected for the presence of conserved motifs four and five (M4 & M5). Partial sequences and those that contained stop codons were removed. CYP9 sequences were named numerically for each species as described above.

6.2.1.3 Estimating the phylogeny of the CYP9 subfamily across bee families

The resulting CYP9 sequences from the genome and transcriptome databases were translated into protein sequences in Geneious (version 10.2.3 (Biomatters)). All partial sequences were removed, and the remaining sequences were aligned with the outgroup sequence CYP9AG4 from the parasitoid Chalcidoidea wasp *Nasonia vitripennis*, in Geneious using MUSCLE [451] (version 3.5, default settings). MEGAX [455] was used to find the best-fit model of amino acid substitution, using a maximum likelihood fit of 56 different models as described in section 3.2.2, see appendix table 6.2. The alignment was used to generate phylogenetic trees using Maximum likelihood [453] (Substitution model: LG+G [454]) and Bayesian inference [445] algorithms (Substitution model: LG+G [454]; Chain length: 1,100,000; Subsampling frequency: 200; Burn-in length: 100,000; Heated chains: 4; Heated chain temperature: 0.2). For all species, partial sequences were translated and aligned with the CYP9 sequences of *A. mellifera*, *B. terrestris*, *O. bicornis*, *M. rotundata* and *D. novaeangliae* using MUSCLE. These MSAs were used to generate phylogenetic trees using a maximum likelihood algorithm [453] (Substitution model: LG+G [454]), in order to infer which CYP9 lineage best described each partial sequence.

Sequences that appeared in the *CYP9DN* subfamily were removed in order to focus on the main CYP9 cluster and the remaining sequences were realigned in Geneious using MUSCLE [451] (version 3.5, default settings). This second MSA was used to generate phylogenetic trees using Maximum likelihood [453]

(Substitution model: LG+G [454]) and Bayesian inference [445] algorithms (Substitution model: LG+G [454]; Chain length: 1,100,000; Subsampling frequency: 200; Burn-in length: 100,000; Heated chains: 4; Heated chain temperature: 0.2).

6.2.2 Selection of candidate Megachilidae CYP9 sequences for functional expression

In *O. bicornis* the CYP9BU lineage is involved in the metabolism insecticides from three MOA classes [137, 155, 168, 226, 446] (see sections 1.4.4.1 and 1.7.3) and so this clade was selected and examined. Sequences that appeared in a clade with *O. bicornis* CYP9BU1 and CYP9BU2, were chosen for functional expression. These potential CYP9BU orthologs were synthesised *in vitro* by Twist Bioscience. The sequences were delivered in pTwistENTR plasmids with flanking Gateway® *attL1* (vector-N₇₅ CCA ACT TTG TAC AAA AAA GCA GGC TTC-insert) and *attL2* sites (insert-TGG GTC GAA AGA ACA TGT TTC AAC C N₇₅- vector) and kanamycin resistance (see appendix figure 6.1 for plasmid map). On arrival, the quantity and quality of the entry clones were assessed using a NanoDrop One (Thermo Scientific) or NanoDrop 2000 spectrophotometer (Thermo Scientific). The entry clones were then stored at -20°C until needed. Expression clones were created from the entry clones, through Gateway® cloning technology using BaculoDirect™ C-Term linear DNA, as described in chapter five, section 5.2.2.3. Recombinant CYP9 proteins were then expressed using insect cell lines as described in section 5.2.2.4 to 5.2.2.5.7. Viral stocks of *O. bicornis* CYP9BU1 and CYP9BU2 were kindly provided by Kat Beadle and recombinant proteins were expressed from these for use in model substrate kinetic assays.

6.2.3 Model substrate metabolism assays

All assays using model substrates were performed in a black flat-bottomed 96-well plate (4-titude). Data were recorded using a SpectraMax M2 Multi-mode plate reader (Molecular Devices).

6.2.3.1 Model substrate profiling

The Megachilidae recombinant expressed CYP9 proteins were screened against six coumarin-based model substrates (MC, EC, EFC, MFC, BFC and

MOBFC, see table 5.9) as described in section 5.2.4.3. Standard curves for HC for coumarin and HFC for trifluoromethyl coumarin were created as outlined in section 5.2.4.2.

6.2.3.2 Fluorescence-based assay to assess insecticide mediated inhibition

The recombinant expressed Megachilidae CYP9 proteins were used in fluorescence-based assay to assess insecticide mediated inhibition using the methodology outlined in section 5.2.4.4.

6.2.4 Surveying CYP9 subfamily sequences in representative Hymenoptera species

6.2.4.1 Searching genomic data of ants, wasp and sawfly species

The NCBI genome database was searched for published genomes of Hymenoptera (hymenopterans). The genomes of three Symphyta species were available for inclusion in BLASTn searches (NCBI genome database, accessed March 2020): one Cephoidea and two Tenthredinoidea. Five species from the Formicoidea (ants), Vespoidea (social wasps), Cynipoidea (gall wasps), Chalcidoidea (chalcid wasps) and Ichneumonoidea (parasitoid wasps), whose genomes were annotated (NCBI Release Annotation), were selected as candidate species for inclusion in BLASTn searches and phylogenetic analyses. No Chrysididae (cuckoo wasp) genomes had NCBI Release Annotation, and only one candidate species (*Goniozus legneri*) was discovered for inclusion. The nucleotide sequences for *A. mellifera* CYP9Q3, CYP9P1, CYP9R1 and *M. rotundata* CYP9DN1 were used as the query sequences in a BLASTn search through the assembly of the genome of each Hymenopteran species to find CYP9 homologs. The resulting sequences were treated as described in section 6.2.1.1.

6.2.4.2 Estimating the phylogeny of the CYP9 subfamily across the Hymenoptera

The resulting Hymenopteran CYP9 sequences were translated into protein sequences in Geneious version 10.2.3 (Biomatters). The sequences were aligned with the CYP9s from the bee species (see sections 6.2.1.1 and 6.2.1.2) using *M. domestica* CYP4ae1 as an outgroup, in Geneious using MUSCLE [451] (version 3.5, default settings). MEGAX [455] was used to find the best-fit

model of amino acid substitution using a Maximum Likelihood fit of 56 different models as described in section 3.2.2, see appendix table 6.4. Phylogeny was estimated of the entire CYP9 subfamily. The CYP9DN sequences were removed and a phylogeny of the CYP9 cluster sequences was estimated.

6.2.5 Conserved synteny analysis across the Hymenoptera

The scaffolds containing CYP9 genes from the genomes of Hymenopteran species with NCBI Release Annotation were downloaded from the NCBI nucleotide database and imported into Geneious version 10.2.3 (Biomatters). Using the micro-synteny analysis from section 3.3.3.1 as a reference, the regions upstream and downstream of all CYP9 genes were examined and all flanking genes noted, to identify potential orthologous regions in the genomes. For the CYP9 region to be considered as showing micro-synteny the minimum requirement was the conservation of two neighbouring homologs with no more than five unrelated genes in the intervening DNA.

Protein identified in a micro-synteny block in all species were aligned using MUSCLE [451] (version 3.5, default settings) to discover their percentage identity and enable an assessment of their homology.

6.2.6 Transcriptomics from four species of *Megachile* genus bees

6.2.6.1 Identification of Canadian native *Megachile* species

In August 2018 (13.08.18), a male bee of a native Canadian species emerged alongside the *M. rotundata* being incubated for use in the acute contact bioassays. The species was identified by Professor Jeremy Field (Professor of Evolutionary Biology, University of Exeter) using the Canadian Journal of Arthropod Identification No 18, available online [619] (https://cjai.biologicalsurvey.ca/srpg_18/C39/C39.html). The literature on the commercial management of *M. rotundata* suggests that *M. relativa* Cresson is the most likely native Canadian species to emerge during incubation [103, 620]. However, in this instance, the vein patterning and prominence of the hypostomal tubercle identified the species as the Willowherb leafcutter bee, *M. lapponica* (see figure 6.2). The *M. lapponica* specimen was snap frozen in liquid N₂ and stored at -80°C.



(a) (b) (c)
 Figure 6.2: (a) and (b) Male willowherb leafcutter bee (*M. lapponica* Thomson, 1872), photos show the long wings which have a smoky subhyaline appearance. (c) Male *M. rotundata* photo shows the shorter wings with less prominent veining.

6.2.6.2 Collection and identification of U.K. native *Megachile* species

Trips to collect specimens of U.K. native *Megachile* species were organised with Professor Jeremy Field (Professor of Evolutionary Biology, University of Exeter) in July 2019. There are seven extant species of the *Megachile* genus native to the U.K. [621], three of which were collected for use (see figure 6.3). The main collection site was Gyllyngdune Gardens, Falmouth Cornwall (50° 8' 48.69" N, 5° 3' 51.20" W) where female *M. centuncularis* and *M. willughbiella* were captured. Females of the coastal species *M. leachella* were collected at Loe Bar, Porthleven, Cornwall, (50° 4' 9.87" N, 5° 17' 39.83" W). For collection, we walked slowly, examining plants for bees, and captured specimens using nets. Specimens were identified by Professor Field, snap frozen in liquid N₂ and stored at -80°C.



Figure 6.3: (a) Female patchwork leafcutter bee (*M. centuncularis* Linnaeus, 1758) [622]. (b) Female silvery leafcutter bee (*M. leachella* Curtis, 1828) [623]. (c) Female Willughby's leafcutter bee (*M. willughbiella* Kirby, 1802) [624].

6.2.6.3 Transcriptome assembly

RNA was extracted from the single Canadian species and the three U.K. species of *Megachile* bee as described in section 2.2.2. The quality and quantity of the RNA was assessed using Qubit® RNA BR and Qubit® RNA IQ Assay kits, and RNA integrity was visualised on an agarose gel, as described in section 2.2.2. Transcriptome libraries (250-300 bp insert cDNA) were generated by Novogene UK (Cambridge, CB4 0FW, UK; <https://en.novogene.com/>) from 10,000 ng total RNA, and run using high-throughput sequencing (Illumina HiSeq) with 150 bp paired-end chemistry, with 30 Gb raw data obtained per sample. Quality control was run by Novogene using Agilent 2100 and q-PCR; the results are shown in appendix table 6.3 and appendix figure 6.2 . All the RNA samples yielded a high number and percentage of clean reads (2.14×10^8 – 2.99×10^8 ; 96.9 – 98.77%). The qualified libraries were pooled according to concentration and expected data volume and then run into an Illumina sequencer. Resulting data was supplied in fastq file format.

For each dataset the reads were quality trimmed, *in silico* normalised and assembled using Trinity version 2.8.4 [625]. Trimming was performed using Trimmomatic [626] as implemented in Trinity with default settings, and strand-specific assembly was specified using the parameter `-SS_lib_type RF`. The resulting contigs were imported, in fasta file format, into Geneious. The number of contigs and the sequence length for each *Megachile* species is shown in figure 6.4.

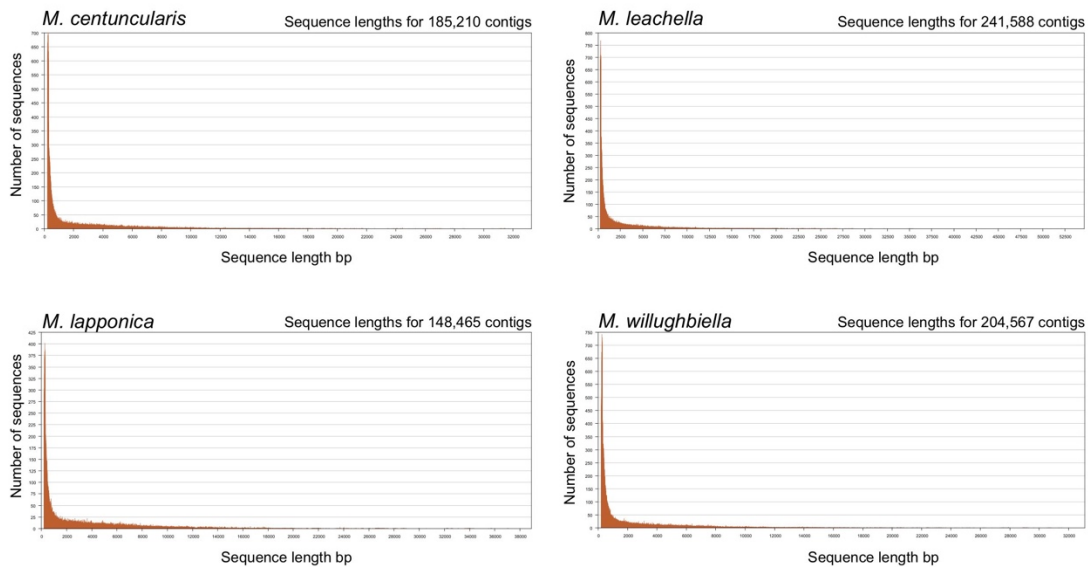


Figure 6.4: The number of contigs and the sequence length for each *Megachile* species dataset.

6.2.6.4 Manual curation to extract CYP9 sequences

The trimmed reads were imported into Geneious version 10.2.3 (Biomatters) and a BLAST database was created from each dataset. The nucleotide sequences for *A. mellifera* *CYP9Q3*, *CYP9P1*, *CYP9R1*, *M. rotundata* *CYP9DN1* and *CYP9DM1* were used as the query sequences in a discontinuous BLAST search through the database for each species. Using the resulting hit table and 'query centric view' tab to avoid duplicates, the top hits were selected and examined manually for the presence of a P450 sequence. Where the contigs were long enough they were examined for the presence of multiple P450 genes. A folder of CYP9 sequences for each bee species was created and each nucleotide sequence was translated and manually inspected for the presence of conserved motifs four and five (M4 & M5). Partial sequences, duplicates and those that contained stop codons were removed. The resulting protein sequences were aligned with CYP9s from *A. mellifera*, *B. terrestris*, *M. rotundata*, *O. bicornis*, *D. novaeangliae* CYP9s and *N. vitripennis* CYP9AG4 using MUSCLE [451] (version 3.5, default settings). Phylogeny was estimated using the PhyML maximum likelihood algorithm [453] (substitution model LG+G [454], with branch support of 50 bootstraps), to identify CYP9 lineages. The *Megachile* CYP9 sequences were then named by species,

lineage and number (e.g. *M. leachella* CYP9P2-like, *M. leachella* CYP9R-cluster1_1).

6.3 Results

6.3.1 The phylogeny of the CYP9 subfamily across bee families

BLASTn searches of the NCBI genome and transcriptome shotgun assembly (TSA) databases were performed in May and June 2020. In all 75 species of bee that had a least one full length CYP9 sequence were discovered – see table 6.1 for a full list of species. Sequences were found for six of the seven bee families. The smallest family, the Stenotritidae, which only accounts for 21 species, was found to be data deficient (as of June 2020). There was a skew in the numbers of species in favour of the Apidae, with over half of the species found from this family (41/75). The Megachilidae, Halictidae, Andrenidae yielded twelve (16%), ten (13.3%) and six (8%) species respectively; and both the Colletidae and Melittidae three (4%). In total 569 sequences were discovered, of which 103 were excluded because they were only partial, leaving 466 full length sequences for translation and inclusion in an MSA and phylogeny.

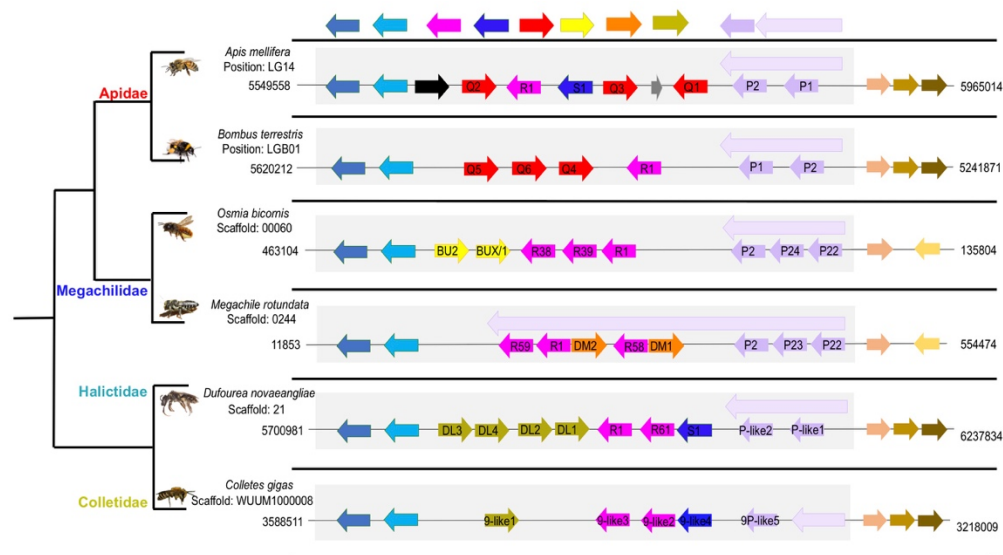


Figure 6.5: Schematic showing the synteny of the CYP9 cluster in bees across four families.

Table 6.1: List of 75 species of bee used in phylogenetic analyses, with brief notes on family, subfamily, life history and ecology

Species	Abbreviation used in tree	Family	Subfamily	No CYP9Q/BU/DL-like genes	Diet	Nesting	Social/ solitary	Notes	Database
<i>Ammobates syriacus</i>	As	Apidae	Nomadinae	1	Nectar	None	Solitary/ Cleptoparasite	Clepto parasite - Cuckoo bee	TSA
<i>Andrena cineraria</i>	AnCi	Andrenidae	Andreninae	N/F	Polylectic	Below ground	Solitary	Ash-mining bee	TSA
<i>Andrena fulva</i>	Anf	Andrenidae	Andreninae	N/F	Polylectic	Below ground	Solitary	Tawny mining bee	TSA
<i>Andrena haemorrhoa</i>	Ah	Andrenidae	Andreninae	1	Polylectic	Below ground	Solitary	Orange-tailed mining bee	TSA
<i>Andrena vaga</i>	Av	Andrenidae	Andreninae	1	Oligolectic - prefers Salix (Willows)	Below ground	Solitary	Grey-backed mining bee	TSA
<i>Anthidium manicatum</i>	Anma	Megachilidae	Megachilinae	N/F	Polylectic	Above ground, cavity nesting	Solitary	European wool carder bee	TSA
<i>Apis cerana</i>	Ac	Apidae	Apinae	3	Polylectic/ trophallaxis	Colony/ Hive/ Above ground	Social	Asiatic honeybee	Genome
<i>Apis dorsata</i>	Ad	Apidae	Apinae	3	Polylectic/ trophallaxis	Colony/ Hive - single comb/ Above ground	Social	Giant honeybee	Genome
<i>Apis florea</i>	Af	Apidae	Apinae	3	Polylectic/ trophallaxis	Colony/ Hive - single comb/ Above ground	Social	Dwarf honeybee	Genome
<i>Apis mellifera</i>	Am	Apidae	Apinae	3	Polylectic/ trophallaxis	Colony/ Hive/ Above ground	Social	European honeybee	Genome
<i>Bombus bifarius</i>	Bb	Apidae	Apinae	3	Polylectic	Colony	Social	Two-form bumblebee	Genome
<i>Bombus campestris</i>	Bc	Apidae	Apinae	3	Nectar	None	Solitary/ Cleptoparasite	Clepto parasite - Cuckoo bee	Genome
<i>Bombus cryptarum</i>	Bc	Apidae	Apinae	3	Polylectic	Colony/ Below ground	Social	Cryptic white-tailed bumblebee	TSA
<i>Bombus impatiens</i>	Bi	Apidae	Apinae	3	Polylectic	Colony/ Below ground	Social	Common Eastern bumblebee	Genome
<i>Bombus pascuorum</i>	Bp	Apidae	Apinae	3	Polylectic	Colony/ Above or below ground	Social	Common carder bee	Genome
<i>Bombus pyrosoma</i>	Bp	Apidae	Apinae	N/F	Polylectic	Colony	Social	Himalayan bumblebee	Genome
<i>Bombus rufestris</i>	Bf	Apidae	Apinae	3	Nectar	None	Solitary/ Cleptoparasite	Clepto parasite - Red-tailed cuckoo bumblebee	TSA
<i>Bombus terrestris</i>	Bt	Apidae	Apinae	3	Polylectic	Colony/ Below ground	Social	Buff-tailed bumblebee	Genome
<i>Bombus vancouverensis</i>	Bv	Apidae	Apinae	3	Polylectic	Colony/ Below ground/ Surface	Social	Vancouver island bumblebee	Genome
<i>Bombus vosnesenkii</i>	Bvo	Apidae	Apinae	3	Polylectic	Colony/ Below ground	Social	Yellow-faced bumblebee	Genome
<i>Campilopoeum sacrum</i>	Cas	Andrenidae	Panurginae	3	Oligolectic on Asteraceae	Below ground	Solitary	Mining bee - Oligolectic on Asteraceae	TSA
<i>Ceratina australensis</i>	Ca	Apidae	Xylocopinae	2	Polylectic	Above ground/ Dead wood/ Stems	Socially polymorphic/ Solitary/ Primitively social	Small carpenter bee	Genome
<i>Ceratina calcarata</i>	Cc	Apidae	Xylocopinae	4	Polylectic	Above ground/ Dead wood/ Stems	Solitary/ subsocial	Carpenter bee	Genome
<i>Ceratina chalybea</i>	Cch	Apidae	Xylocopinae	2	Polylectic	Above ground/ Dead wood/ Stems	Solitary/ nest guarding	Small carpenter bee	TSA
<i>Chelostoma florisomme</i>	Cf	Megachilidae	Megachilinae	1	Oligolectic - only <i>Ranunculus</i> spp	Above ground, cavity nesting	Solitary	Large scissor bee	TSA
<i>Coelioxys conoidea</i>	Coco	Megachilidae	Megachilinae	N/F	Nectar	None	Solitary/ Cleptoparasite	Clepto parasite - Large sharp-tailed bee - cuckoo bee of <i>Megachile maritima</i>	TSA
<i>Colletes curvicaulis</i>	Colc	Colletidae	Colletinae	1	Oligolectic - mostly Salix spp	Below ground	Solitary	Spring mining bee	TSA
<i>Colletes gigas</i>	Colg	Colletidae	Colletinae	1	Oligolectic	Below ground	Solitary	Plasterer bee	Genome
<i>Dasydota hirtipes</i>	Dh	Melittidae	Dasydopodinae	1	Oligolectic - mostly Asteraceae spp	Below ground	Solitary	Pantalon bee or hairy legged mining bee	TSA
<i>Dioxys cincta</i>	Dc	Megachilidae	Megachilinae	1	Nectar	None	Solitary/ Cleptoparasite	Clepto parasite - Cuckoo bee of <i>Megachile</i> spp.	TSA
<i>Dufourea dentiventris</i>	Dd	Halictidae	Rophitinae	2	Oligolectic - <i>Campanula</i> spp	Below ground	Solitary	European sweat bee	TSA
<i>Dufourea novaeangliae</i>	Dn	Halictidae	Rophitinae	4	Only gather pollen from Fickel weed (<i>Pontederia cordata</i>). Nectar from others sources too.	Below ground	Solitary	Sweat bee	Genome
<i>Epeolus variegatus</i>	Ev	Apidae	Nomadinae	2	Nectar	None	Solitary/ Cleptoparasite	Clepto parasite - Black-thighed Epeolus - cuckoo bee of <i>Colletes</i> spp.	TSA
<i>Eucera nigrescens</i>	En	Apidae	Apinae	1	Oligolectic - vetches	Below ground	Solitary	Early long-horned bee	TSA
<i>Eucera plumigera</i>	Ep	Apidae	Apinae	1	Unknown	Below ground	Solitary	Long-horned bee	TSA
<i>Eucera synaca</i>	Es	Apidae	Apinae	1	Unknown	Below ground	Solitary	Long-horned bee	TSA
<i>Eufriesea mexicana</i>	Em	Apidae	Apinae	2	Polylectic	Above ground, cavity nesting	Facultatively simple social	Orchid bee	Genome
<i>Euglossa cordata</i>	Ec	Apidae	Apinae	1	Polylectic	Above ground, cavity nesting	Primitively eusocial	Orchid bee	TSA
<i>Euglossa dilemma</i>	Ed	Apidae	Apinae	2	Polylectic	Above ground, cavity nesting/ propolis resin construction	Solitary	Orchid bee	Genome
<i>Eulaema nigrita</i>	Eun	Apidae	Apinae	N/F	Polylectic	Above ground, cavity nesting	Solitary/ communal	Orchid bee	TSA
<i>Frieseomelitta varia</i>	Fv	Apidae	Apinae	5	Polylectic/ trophallaxis/ trophic egg laying	Colony/ Above ground	Social	Yellow marmalade bee. Sterile worker caste.	Genome
<i>Habropoda laboriosa</i>	Hl	Apidae	Apinae	1	Polylectic - preference for <i>Vaccinium</i> spp pollen	Below ground	Solitary	Southeastern blueberry bee	Genome
<i>Halictus quadricinctus</i>	Hq	Halictidae	Halictinae	2	Polylectic	Below ground	Solitary/ communal	Giant furrow bee - ground nesting	TSA
<i>Haeriades truncorum</i>	Ht	Megachilidae	Megachilinae	2	Oligolectic - yellow flowered Asteraceae	Above ground/ Cavity nesting/ <i>Rubus</i> stems - resin used to line nest	Solitary	Large-headed resin bee	TSA
<i>Heterotrigona itama</i>	Hi	Apidae	Apinae	1	Polylectic	Colony/ Above ground/ cavity nesting	Social	Malaysian stingless bee	TSA
<i>Hylaeus variegatus</i>	Hv	Colletidae	Colletinae	1	Polylectic	Above ground/ cavity nesting/ Below ground	Solitary	Romanian bee - nests in the ground	TSA
<i>Lasoglossum albipes</i>	La	Halictidae	Halictinae	2	Polylectic	Below ground	Facultatively simple social	Bloomed furrow bee - sweat bee	Genome
<i>Lasoglossum xanthopus</i>	Lx	Halictidae	Halictinae	1	Polylectic	Below ground	Solitary	Orange-footed furrow bee - sweat bee	TSA
<i>Lepidotrigona ventralis</i>	Lv	Apidae	Apinae	2	Polylectic/ trophic egg laying	Colony/ Above ground/ tree cavity	Social	Stingless bee	Genome
<i>Lithurgus chrysurus</i>	Lc	Megachilidae	Lithurginae	N/F	Oligolectic - Centaurea spp	Above ground, wood boring	Solitary	Mediterranean wood-boring bee	TSA
<i>Macropis fulvipes</i>	Mf	Melittidae	Melittinae	1	Oligolectic - <i>Lysimachia</i> (Primulaceae)	Below ground	Solitary	Oil collecting bees that dig a nest	TSA
<i>Megachile rotundata</i>	Mr	Megachilidae	Megachilinae	0	Polylectic	Above ground - stems, reeds - leaf material used to line nest	Solitary/ gregarious	Alfalfa leaf-cutter bee	Genome
<i>Megachile willughbiella</i>	Mw	Megachilidae	Megachilinae	0	Polylectic	Below ground, soil - or wood - leaf material used to line nest	Solitary	Willughby's leaf-cutter bee	TSA
<i>Megalopta genalis</i>	Mg	Halictidae	Halictinae	4	Polylectic/ trophallaxis	Above ground in dead wood, vnes	Facultatively social/ solitary	Sweat bee - nocturnal species	Genome
<i>Melipona quadricincta</i>	Mq	Apidae	Apinae	2	Polylectic	Colony/ Above ground/ mud construction in trees	Social	Stingless bee	Genome
<i>Melitta haemorrhoidalis</i>	Mh	Melittidae	Melittinae	1	Oligolectic - bellflowers (<i>Campanula</i> spp.)	Below ground	Solitary	Bellflower blunthorn bee - specialist on chalk and limestone - burrows a nest	TSA
<i>Nomada lathburiana</i>	Nl	Apidae	Nomadinae	1	Nectar	None	Solitary/ Cleptoparasite	Clepto parasite - Lathbury's nomad bee - cuckoo bee of <i>Andrena cineraria</i>	TSA
<i>Nomia diversipes</i>	Nd	Halictidae	Nominae	1	Polylectic	Below ground - soil	Solitary	European sweat bee	TSA
<i>Nomia melanderi</i>	Nm	Halictidae	Nominae	1	Preference for alfalfa	Below ground - soil	Solitary/ gregarious	Alkali bee	Genome
<i>Osmia bicornis</i>	Ob	Megachilidae	Megachilinae	2	Polylectic	Above ground - cavity - mud used to line nest	Solitary	Red mason bee	Genome
<i>Osmia comuta</i>	Oc	Megachilidae	Megachilinae	2	Polylectic	Above ground - cavity - mud used to line nest	Solitary	European orchard bee	TSA
<i>Osmia lignaria</i>	Ol	Megachilidae	Megachilinae	2	Polylectic	Above ground - cavity/ reeds - mud used to line nest	Solitary	Blue orchard bee	Genome
<i>Panurgus dentipes</i>	Pd	Andrenidae	Panurginae	1	Restricted to Asteraceae	Below ground	Solitary	Mining bee	TSA
<i>Sphexodes albilabris</i>	Sa	Halictidae	Halictinae	N/F	Nectar	None	Solitary/ Cleptoparasite	Clepto parasite - Cuckoo bee of <i>Colletes curvicaulis</i>	TSA
<i>Stelis punctulatisssima</i>	Sp	Megachilidae	Megachilinae	N/F	Nectar	None	Solitary/ Cleptoparasite	Banded dark bee	TSA
<i>Systropha curvicornis</i>	Sc	Halictidae	Rophitinae	1	Oligolectic - <i>Convolvulus</i> species	Below ground	Solitary	Old world sweat bee	TSA
<i>Tetragonula carbonaria</i>	Tc	Apidae	Apinae	2	Polylectic	Above ground, cavity nesting - resin used to line colony entrance	Social	Stingless bee - Sugarbag bee	Genome
<i>Tetragonula clypearis</i>	Tcl	Apidae	Apinae	1	Polylectic	Above ground, cavity nesting	Social	Stingless bee - Sugarbag bee	Genome
<i>Tetragonula davenporti</i>	Td	Apidae	Apinae	2	Polylectic	Above ground, cavity nesting	Social	Stingless bee - Sugarbag bee	Genome
<i>Tetragonula hockingsi</i>	Th	Apidae	Apinae	3	Polylectic	Above ground, cavity nesting	Social	Stingless bee - Sugarbag bee	Genome
<i>Tetragonula mellipes</i>	Tm	Apidae	Apinae	4	Polylectic	Above ground, cavity nesting	Social	Stingless bee	Genome
<i>Tetralonia malvae</i>	Tmal	Apidae	Apinae	1	Oligolectic - <i>Malva arborea</i> tree mallow	Below ground - soil	Solitary	Long-horned bee	TSA
<i>Tetraloniella nigripes</i>	Tn	Apidae	Apinae	1	Polylectic	Below ground - soil	Solitary	Long-horned bee	TSA
<i>Thyrsus orbatu</i>	To	Apidae	Apinae	1	Nectar	None	Solitary/ Cleptoparasite	Clepto parasite - Cuckoo bee of <i>Amegilla</i> spp.	TSA
<i>Xylocopa violacea</i>	Xv	Apidae	Xylocopinae	1	Polylectic	Above ground - dead wood	Solitary	Violet carpenter bee	TSA

Based on the genomic data available, the CYP9 subfamily in bees, with the exception of *CYP9DN1*, is found in a cluster flanked by the *membralin*, *alpha catulin* and *myosin IIIb* genes (see figure 6.5). Unfortunately, even though CYP9 sequences were found for the Melittidae and Andrenidae families, the data was taken from the TSA database and, as such, there is no information on the genomic position of those sequences. However, the results below are described in terms of the full CYP9 subfamily (i.e. including *CYP9DN1*) and the CYP9 cluster. The phylogeny for the entire CYP9 subfamily is shown in figure 6.6, and the numbers of species found containing full-length sequences of each CYP9 lineage in table 6.2.

Table 6.2: Number of sequences of each CYP9 lineage from 75 species of bee across six families (all data taken from NCBI genome and transcriptome Shotgun Assembly (TSA) databases).

Family (number of species)	CYP9DN1	CYP9P	CYP9R	CYP9Q/BU/DL	CYP9DM
Apidae (41)	10	35	38	39	0
Megachilidae (12)	12	12	12	6	2
Halictidae (10)	8	9	10	9	0
Colletidae (3)	3	3	3	3	0
Melittidae (3)	3	1	3	3	0
Andrenidae (6)	5	4	6	4	0

6.3.1.1 The *CYP9DN1* lineage

It is clear that in common with *A. mellifera* and *B. terrestris* (see section 3.3.2), not all Apidae species appear to have *CYP9DN1* sequences, see figure 6.6 and table 6.2. In fact, only around a quarter to a third of the Apidae species found have a *CYP9DN1* gene, see table 6.2. Of the ten Apidae species with a *CYP9DN1* nine were from the Apinae subfamily and one from the Xylocopinae. With the exception of one Andrenidae and two Halictidae species, a *CYP9DN1* gene is present in all other bees, see table 6.2. The exceptions are *Panurgus dentipes* (Andrenidae), *Lasioglossum xanthopus* (Halictidae) and *Dufourea dentiventris* (Halictidae). However, the BLASTn searches for these species yielded: five (498-1167 bp), three (378-1482 bp) and seven (504-1464 bp) partial sequences respectively, so one, or more of these fragments could indicate the presence of the gene. It appears possible, therefore, that *CYP9DN1* is ubiquitous in non-Apidae bees.

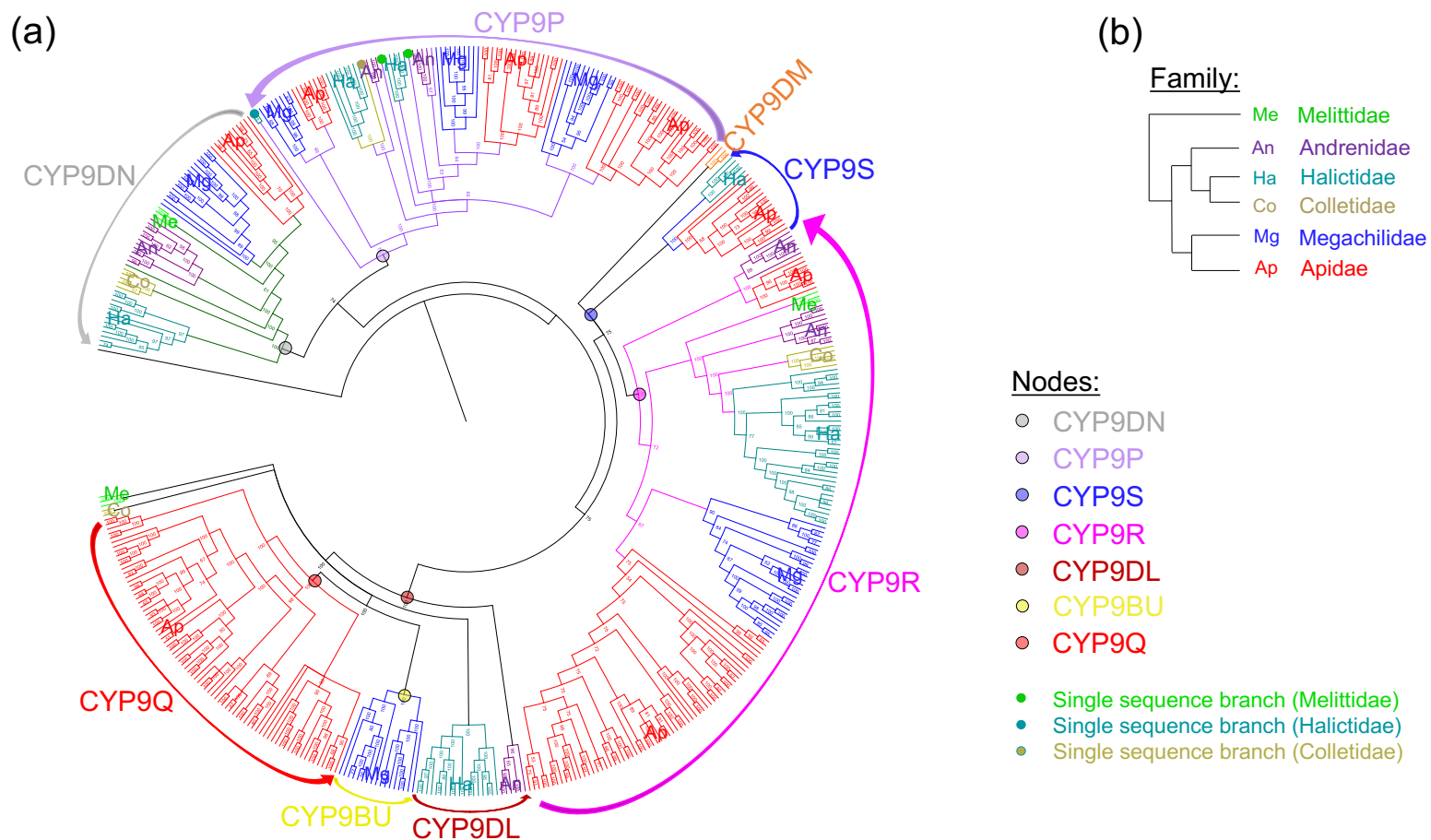


Figure 6.6: (a) Bayesian inference phylogeny [445] of the CYP9 subfamily across 75 species of bee, using substitution model LG+G [454]. Posterior probability of nodes shown as a % probability. Tree rooted on *N. vitripennis* CYP9AG4. Sequences are coloured by bee family and annotated with an abbreviated form of the family name. Single sequence branches are labelled with a circle, coloured by family. All nucleotide sequences accessed from NCBI databases. (b) Schematic of the phylogenetic relationship between bee families.

6.3.1.2 The *CYP9P* lineage

The phylogeny of the CYP9 cluster is shown in figure 6.7. The long branch leading to the ancestral node of the clade indicates that the *CYP9P* lineage is sister to the other CYP9 cluster lineages (see figure 6.7). When the partial sequences that were found during the in initial BLASTn searches were included in the analysis (see table 6.3), only three of the 75 species of bee showed no clear evidence for the presence of a *CYP9P* gene. These were *Dasypoda hirtipes* (Melittidae), *Panurgus dentipes* (Halictidae) and *Lepidotrigona ventralis* (Apidae). Only three CYP9 sequences and one pseudogene were discovered for *D. hirtipes*. The pseudogene has one stop codon (at position 907-909 of the nucleotide sequence), however, when it is run in a phylogeny estimate it clades with the *CYP9P* lineage from other bee species. Overall, it appears that the majority of bee species have *CYP9P* genes. It also seems that the duplication of an ancestral *CYP9P* sequence occurred before the divergence of the families, as evidenced by three distinct clades in the lineage (see figure 6.7).

Table 6.3: Partial sequences, found during BLASTn searches of the NCBI databases, for bee species without a *CYP9P* representative in the Bayesian inferred phylogeny [445] of the CYP9 cluster (figure 6.7).

Species (Family)	No. of partial sequences	Range of length of partial sequences (bp)	No. of partial sequences that clade with CYP9Ps
<i>Melitta haemorrhoidalis</i> (Melittidae)	2	1011-1257	1
<i>Dasypoda hirtipes</i> (Melittidae)	0	N/A	0
<i>Panurgus dentipes</i> (Andrenidae)	5	498-1167	0
<i>Andrena vaga</i> (Andrenidae)	2	874-984	1
<i>Nomia diversipes</i> (Halictidae)	7	504-1464	2
<i>Euglossa cordata</i> (Apidae)	4	429-1440	1
<i>Epeolus variegatus</i> (Apidae)	3	1476-1530	2
<i>Eucera nigrescens</i> (Apidae)	2	805-925	1
<i>Heterotrigona itama</i> (Apidae)	6	421-771	1
<i>Lepidotrigona ventralis</i> (Apidae)	3	801-1407	0
<i>Thyreus orbatus</i> (Apidae)	3	657-1493	3

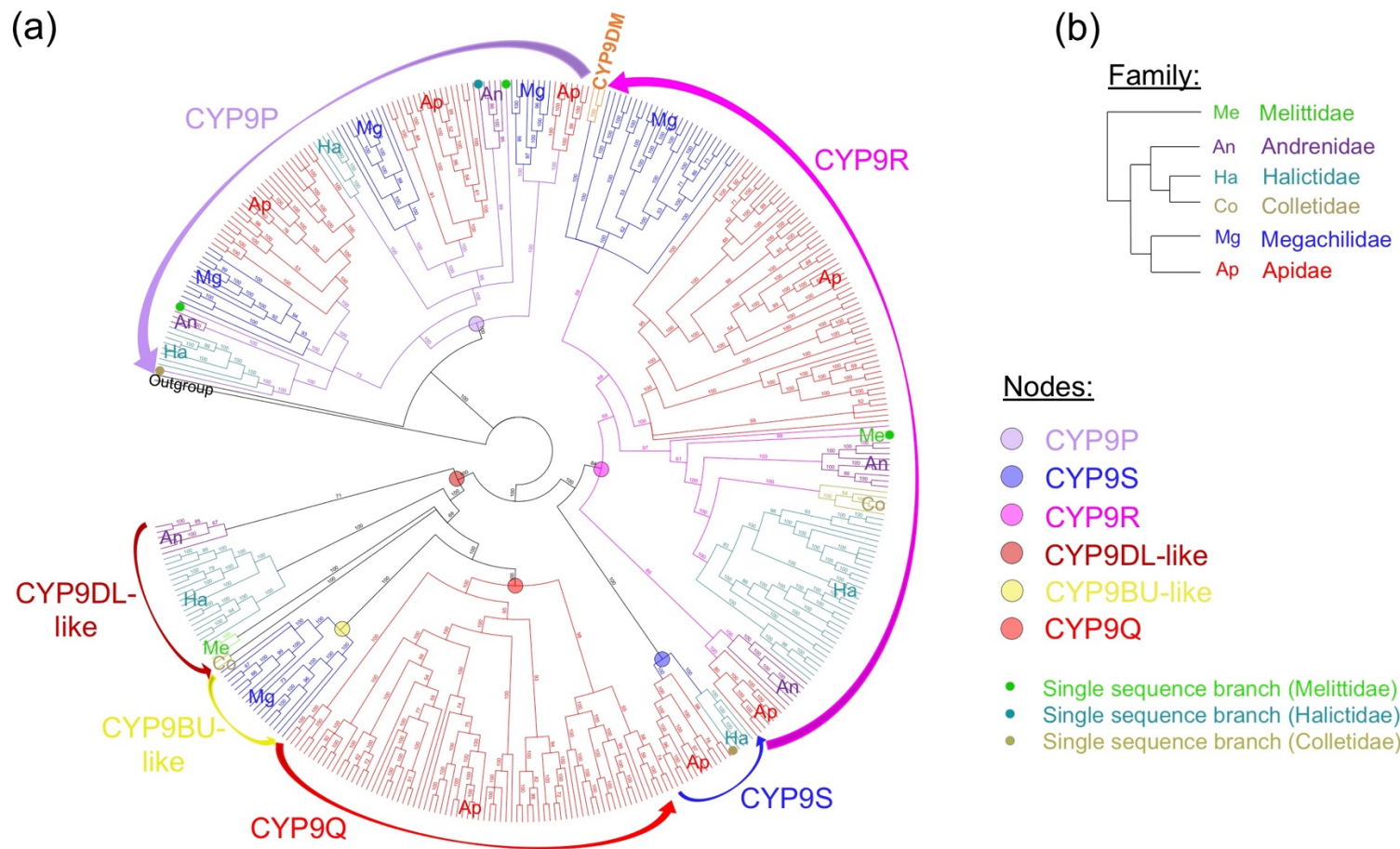


Figure 6.7: Bayesian inference phylogeny [445] of the CYP9 cluster across 75 species of bee, using substitution model LG+G [454]. Posterior probability of nodes shown as a % probability. Tree rooted on *N. vitripennis* CYP9AG4. Sequences are coloured by bee family and annotated with an abbreviated form of the family name. Single sequence branches are labelled with a circle, coloured by family. All nucleotide sequences accessed from NCBI databases. (b) Schematic of the phylogenetic relationship between bee families.

6.3.1.3 The *CYP9R* lineage

The vast majority of bees have a *CYP9R* lineage member (72/75 species). The three species that appear to lack a *CYP9R* gene are all from the Apidae family: *Ammobates syriacus*, *Eulaema nigrita* and *Tetragonula carbonaria*. *T. carbonaria* has a partial sequence (977 bp) that clades as a *CYP9R*. Only three sequences were found for *A. syriacus*, and *E. nigrita* has three partial sequences that were excluded. The apparent lack of a *CYP9R* gene in these species may well be an artifact in these cases. Although it should be noted that *A. syriacus* is a cleptoparasitic species (see table 6.1), and therefore does not collect pollen for its offspring, potentially reducing its dietary exposure to allelochemicals. A reduction in the detoxification gene repertoire therefore should not be ruled out for this species.

6.3.1.4 The *CYP9S* lineage

Only three families appear to have *CYP9S* genes: five Halictidae species, twelve Apidae species and one Colletidae (see figure 6.7). No Melittidae, Megachilidae or Andrenidae species *CYP9S* sequences were found. The *CYP9S* lineage is basal to the larger, more diverse *CYP9R* clade. It may be, therefore, that only certain species have retained this more ancestral form of P450. The Apidae and Halictidae species with *CYP9S* genes are from two subfamilies, in both cases: the Apinae and the Xylocopinae and the Halictinae and the Rophitinae respectively. *Colletes gigas*, the only Colletidae species with a published genome, also has a *CYP9S* gene. It may be that the other two Colletidae species do have *CYP9S*, but as with the other TSA data, due to the exclusion of partial sequences they are absent from the phylogeny in this case.

6.3.1.5 The *CYP9Q/BU/DL* lineage

Of the 41 Apidae species, 39 have at least one *CYP9Q-like* sequence that appears in the phylogeny shown in figures 6.7 and 6.8(a). The two species that appear to lack an *CYP9Q* ortholog have partial sequences that clade with the lineage when they are included in phylogeny estimation (see table 6.4). It seems, therefore safe to say that all Apidae species have at least one *CYP9Q-like* sequence. Likewise, it is reasonable to assume that the lack of a *CYP9Q/BU/DL* ortholog in the two species *Andrena* species is due to an artifact in the data. There are two other *Andrena* species that have *CYP9Q/BU/DL*

sequences (see figure 6.8(a)), and *A. cineraria* has a partial sequence that clades with those when it is included (see table 6.4). Only one Halictidae species, *Sphecodes albilabris*, has no *CYP9DL* ortholog and the partial sequence for this species clades with the *CYP9P* lineage. Like *A. syriacus*, mentioned in section 6.3.1.3, *S. albilabris* is a cleptoparasitic species (see table 6.1) and so may possess a reduced detoxification gene repertoire.

Table 6.4: Partial sequences, found during BLASTn searches of the NCBI databases, for bee species without a *CYP9Q/BU/DL* representative in the Bayesian inferred phylogeny [445] of the CYP9 cluster (figure 6.7).

Species (Family)	No. of partial sequences	Range of length of partial sequences (bp)	No. of partial sequences that clade with CYP9Q/BU/DLs
<i>Sphecodes albilabris</i> (Halictidae)	1	1024	0
<i>Andrena cineraria</i> (Andrenidae)	4	1215-1554	1
<i>Andrena fulva</i> (Andrenidae)	3	1092-1348	0
<i>Bombus pyrosoma</i> (Apidae)	1	1077	1
<i>Eulaema nigrita</i> (Apidae)	3	345-1047	3
<i>Anthidium manicatum</i> (Megachilidae)	0	N/A	0
<i>Coelioxys conoidea</i> (Megachilidae)	1	1172	0
<i>Lithurgus chrysurus</i> (Megachilidae)	1	1224	0
<i>Stelis punctulatissima</i> (Megachilidae)	3	917-1511	0
<i>M. willughbiella</i> (Megachilidae)	1	1127	0
<i>M. rotundata</i> (Megachilidae)	0	N/A	0

Only half of the Megachilidae species (6/12) have a *CYP9BU* ortholog (see table 6.2 and figure 6.8(a)). Both of the *Megachile* species have *CYP9DM* sequences that appear in a basal position in the *CYP9R* clade (see figure 6.7). Of the remaining four species (*A. manicatum*, *C. conoidea*, *L. chrysurus* and *S. punctulatissima*) two are cleptoparasitic species (see table 6.1).

6.3.2 Selection of candidate Megachilidae CYP9 sequences for functional expression

The six Megachilidae species that have at least one *CYP9BU*-like sequences are: *O. bicornis*, *Osmia lignaria*, *Osmia cornuta*, *Chelostoma florissomne*, *Dioxys cincta* and *Heriades truncorum* (see figure 6.8 (a)). Three species have sequences that clade with *CYP9BU1* (basal node marked with a green circle in figure 6.8 (a)) and five species have sequences that clade with *CYP9BU2* (basal node marked with a blue circle in figure 6.8 (a)). Three of the species are from the *Osmia* genus. In order to avoid including too many closely related species, the decision was taken to include only one *Osmia* species (further to *O. bicornis*). *O. lignaria* has a published genome and so was chosen for inclusion. In total five *CYP9BU*-like sequences were chosen for functional characterisation, two *CYP9BU1*-like: *O. lignaria* CYP9-like and *H. truncorum* CYP9-like1 and three *CYP9BU2*-like: *H. truncorum* CYP9-like2, *D. cincta* CYP9-like and *C. florissomne* CYP9-like (see figures 6.8 (a) and (c)).

Gateway® expressions clones were successfully created for all five *CYP9BU*-like sequences. Recombinant CYP9 proteins were then expressed using insect cell lines and a CO-difference spectral assay was performed as described in section 5.2.2.4.7, and the global protein content was calculated using a Bradfords assay as described in section 2.3.1 [551] (see table 6.5 and appendix figure 6.3). Three of the five *CYP9BU*-like sequences produced a peak at 450 nm in the CO-difference spectral assay, but the two *H. truncorum* sequences produced a peak at 420 nm (see table 6.5 and appendix figure 6.3).

Table 6.5: Expression of recombinant P450s using Gateway® cloning technology and transfection of insect cell lines.

Name	Size in bp	Entry clone concentration ng/μl	Amount of entry clone used in LR reaction μl	Viral titer P2 stock	Volume to transfect P3 stock μl (MOI = 0.1)	Viral titer P3 stock	Volume to transfect for expression ml (MOI = 3 -10)	450nm peak	Amount of P450 nMol/ml	Amount of protein (mg/ml)
Cf_1	1566	896.0	1	6.07E+07	125	7.07E+08	0.51	Y	1.5604	36.35
Dc_1	1524	313	1	7.20E+07	105	5.60E+08	0.65	Y	1.5824	32.86
Ht_1	1581	1017	1	9.80E+07	80	6.80E+08	0.55	N	N/A	29.68
Ht_2	1554	213	1	6.33E+07	120	6.13E+08	0.60	N	N/A	19.27
Ol_1	1566	392	1	4.73E+07	160	5.60E+08	2.50	Y	5.6484	36.61
BU1	1554	N/A	N/A	6.67E+07	115	5.47E+08	0.70	Y	4.5495	31.96
BU2	1530	N/A	N/A	N/A	N/A	1.40E+08	2.60	Y	1.2308	36.42

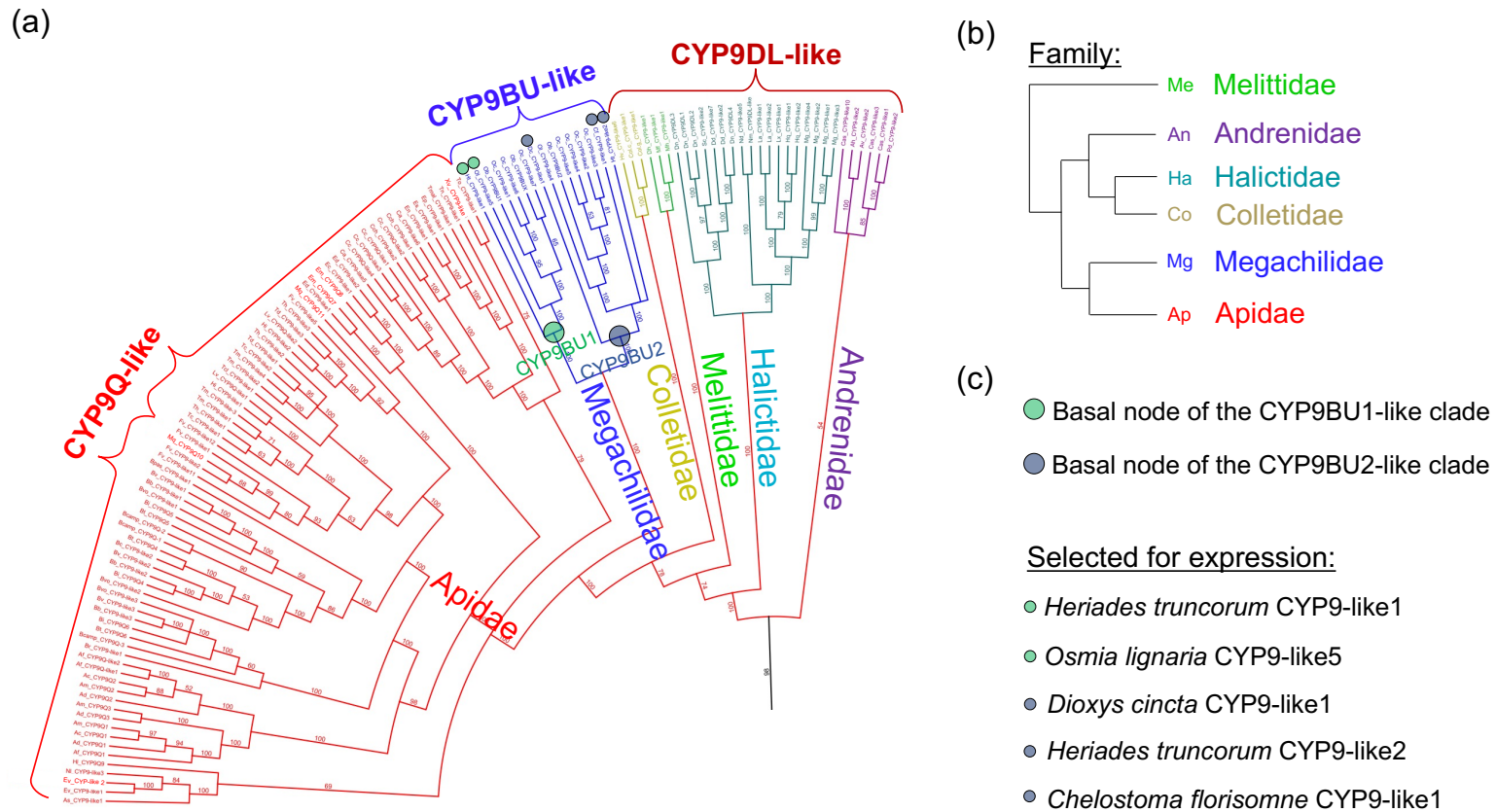


Figure 6.8: (a) Bayesian inference phylogeny [445] of the CYP9Q/BU/DL clade across 75 species of bee, using substitution model LG+G [454]. Posterior probability of nodes shown as a % probability. Tree rooted on *N. vitripennis* CYP9AG4. Sequences are coloured by bee family and annotated with an abbreviated form of species name (see table 6.1). All nucleotide sequences accessed from NCBI databases. (b) Schematic of the phylogenetic relationship between bee families. (c) Annotations indicating Megachilidae CYP9BU-like sequences selected for functional expression.

6.3.4 Fluorescence-based model substrate assays

6.3.4.1 Model substrate profiling

All five recombinant Megachilidae CYP9BU-like sequences, and *O. bicornis* CYP9BU1 and CYP9BU2, were screened in kinetic assays using coumarin-derived fluorescent model substrates (see table 5.2). Activity was found in four out of the five P450s expressed. The two *H. truncorum* P450s, produced a spectral absorption peak of 420 nm rather than 450 nm (see appendix figure 6.3), and only *H. truncorum* CYP9-like1 showed metabolic activity. *O. bicornis* CYP9BU1 and CYP9BU2 both produced a spectral absorption peak of 450 nm (see appendix figure 6.3) and were metabolically active.

The six Megachilidae P450s screened were all found to be active against EC, and five showed activity against MC, the two smallest coumarins (MW: 190.195 and 176.17 g/mol respectively). *O. bicornis* CYP9BU2 and *D. cincta* CYP9-like were the most promiscuous of the tested P450s, each being active against five of the six coumarin-derived model substrates. *H. truncorum* CYP9-like1 was the only other P450 that showed activity against the larger (trifluoromethyl)-coumarin substrates, with metabolism of BFC (MW: 320.26 g/mol). None of the expressed CYP9 P450s enzymes showed activity against MFC.

The model substrate profiling was split into two groups, corresponding to the CYP9BU1 and CYP9BU2 clades (see figure 6.8 (a)) to look for a pattern in the observed activity. In general, the CYP9s that appear in the CYP9BU2 clade (see figure 6.9 (d) and (e)) were active against a wider spectrum of coumarin-derived model substrates than those more closely related to CYP9BU1 (see figure 6.9 (a) and (c)). Two pairings of CYP9s were active against an identical panel of the substrates: *O. bicornis* CYP9BU1 with *C. florisomne* CYP9BU-like; and *O. bicornis* CYP9BU2 with *D. cincta* CYP9BU-like (see figure 6.9 (a) and (d)). However, *C. florisomne* CYP9BU-like clades with *O. bicornis* CYP9BU2 in the phylogeny and so it appears that there is no recognisable model substrate signature that defines either CYP9BU1-like or CYP9BU2-like sequences.

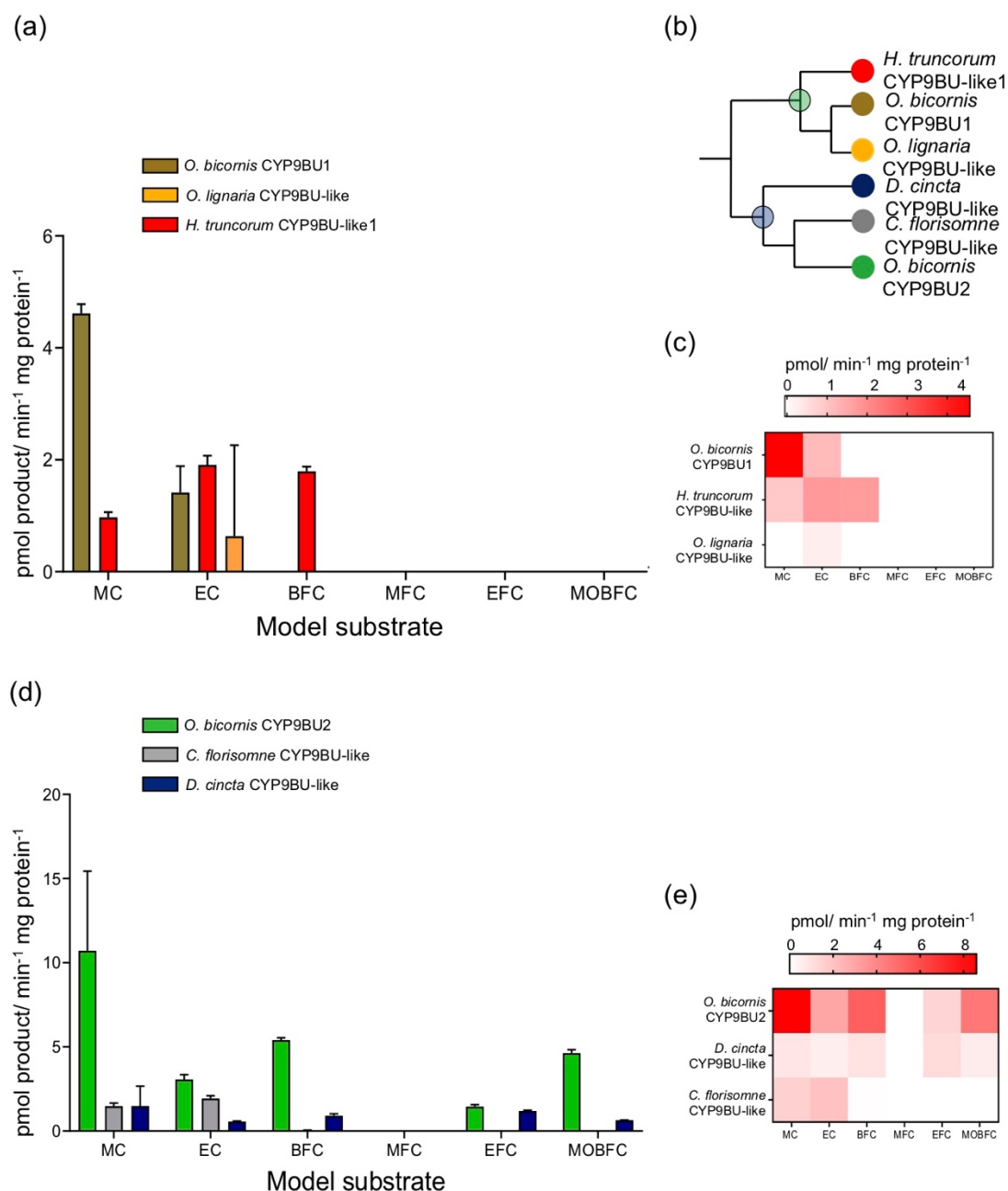


Figure 6.9: Model substrate profile of Megachilidae CYP9BU-like expressed recombinant proteins against coumarin-derived fluorescent model substrates. (a) Metabolic profile of proteins in the CYP9BU1 clade. CYP9 P450s are coloured by species. (b) Schematic of the phylogenetic relationship between the CYP9BU genes. (c) Heatmap showing the metabolic activity of recombinant CYP9BU1 clade against coumarin-derived fluorescent model substrates. (d) Metabolic profile of proteins in the CYP9BU2 clade. CYP9 P450s are coloured by species. (e) Heatmap showing the metabolic activity of recombinant CYP9BU2 clade against coumarin-derived fluorescent model substrates. All data points are mean values \pm SD (n=3).

6.3.4.2 Fluorescence-based assay to assess insecticide mediated inhibition

Based on model substrate profiling, EC or MC was selected for use against the Megachilidae CYP9BU-like sequences in insecticide mediated inhibition assays (see figure 6.10). MC was used against *D. cincta* CYP9-like, *O. bicornis* CYP9BU1 and CYP9BU2; EC against *O. lignaria* CYP9-like, *C. florisomne* CYP9-like and *H. truncorum* CYP9-like1.

Three of the recombinant CYP9s, *O. lignaria* CYP9-like, *C. florisomne* CYP9-like and *O. bicornis* CYP9BU1 showed a significant reduction in HC production after incubation with both thiacloprid and imidacloprid (see table 6.6 and figure 6.10). The inhibitory effect of both neonicotinoids on the metabolism of EC in *C. florisomne* was particularly strong with a significant reduction in HC production ($p < 0.001$) between 0 μM insecticide and 100 μM (i.e. comparing red and blue lines on the graphs in figure 6.9). *H. truncorum* CYP9-like1 displayed unusual activity, whereby incubation with thiacloprid had a significant inhibitory effect on HC production ($p < 0.05$), but the presence of imidacloprid had a significant facilitatory effect ($p < 0.0001$) (see table 6.6 and figure 6.10).

Table 6.6: Reduction of HC production (%) by Megachilidae recombinant CYP9s, after incubation (1 h), with the neonicotinoid insecticides thiacloprid (TCP) and imidacloprid (IMI).

CYP9	% Reduction of HC production			
	TCP	p value (Welch's t-test)	IMI	p value (Welch's t-test)
<i>O. bicornis</i> CYP9BU1	32.7	0.0014**	14.1	0.0387*
<i>O. lignaria</i> CYP9-like	74.2	0.0347*	100	0.0012**
<i>H. truncorum</i> CYP9-like1	47.1	0.0256*	-80.0	0.0002***
<i>O. bicornis</i> CYP9BU2	5.3	0.8796	10.3	0.6981
<i>C. florisomne</i> CYP9-like	64.2	0.0048**	71.3	0.0007***
<i>D. cincta</i> CYP9-like	28.8	0.1702	30.0	0.0302*

Depletion values compare levels of HC production at 0 μM insecticide inhibitor to those at 100 μM . Values used are for 200 μM fluorescent probe concentration (MC/EC). Depletion values for *O. bicornis* CYP9BU2 are averaged values at 200 μM and 100 μM fluorescent probe concentration.

D. cincta CYP9-like showed a significant inhibition in HC production in the presence of imidacloprid ($p < 0.05$). Although incubation of *D. cincta* CYP9-like with thiacloprid did not produce a significant reduction in HC, there was a 28.8%

decrease in the level of product between 0 μM and 100 μM insecticide concentrations. *O. bicornis* CYP9BU2 was the only protein that exhibited no significant inhibition of HC production by either neonicotinoid insecticide (see table 6.6).

Most of the kinetic activity observed in the Megachilidae CYP9s are atypical (see figure 6.10). Four out of the six P450s exhibit non-Michaelis-Menten kinetics, as seen by the lack of a hyperbolic plot. Co-incubation with thiacloprid results in biphasic plots for most of the Megachilidae CYP9s (4/6) and so no V_{max} or K_{M} values could be calculated. Co-incubation with thiacloprid in the case of *O. lignaria* CYP9-like and *C. florisomne* CYP9-like exhibited sigmoidal plots and so an attempt to calculate V_{max} or K_{M} values was made (see table 6.7). However, the values for V_{max} or K_{M} at 100 μM (and their 95% confidence intervals) were unstable or reported as infinite.

Co-incubation with imidacloprid resulted in biphasic plots for two of the Megachilidae CYP9s, *O. bicornis* CYP9BU2 and *D. cincta* CYP9-like. A hyperbolic plot (Michaelis-Menten kinetics) was exhibited by *O. lignaria* CYP9-like without imidacloprid (0 μM), but a biphasic plot at 100 μM . No V_{max} or K_{M} values could be calculated for the biphasic plots. Co-incubation with imidacloprid resulted in sigmoidal plots for two Megachilidae CYP9s: *O. bicornis* CYP9BU1 and *H. truncorum* CYP9-like1 (see figure 6.10). V_{max} or K_{M} values were calculated for the kinetics of both sigmoidal plots although these data should be considered as indicative only as the reaction did not follow Michaelis-Menten kinetics (see table 6.7). The result for *O. bicornis* CYP9BU1 indicates a decrease in both K_{M} and V_{max} ($K_{\text{M}100 \mu\text{M}} < K_{\text{M}0 \mu\text{M}}$ and $V_{\text{max}100 \mu\text{M}} < V_{\text{max}0 \mu\text{M}}$), which may be indicative of a heterotropic interaction resulting in non-competitive or partial inhibition of HC production by the presence of imidacloprid. In the case of *H. truncorum* CYP9-like1, the results suggest that the presence of imidacloprid increases the binding affinity for EC molecules ($K_{\text{M}100 \mu\text{M}} < K_{\text{M}0 \mu\text{M}}$ or decrease in K_{M}) and increases the rate of HC formation ($V_{\text{max}100 \mu\text{M}} > V_{\text{max}0 \mu\text{M}}$ or increase in V_{max}) (see table 6.7) [581]. This appears to be an example of heteroactivation rather than a form of competition or inhibition. Co-incubation with imidacloprid and *C. florisomne* CYP9-like exhibits Michaelis-Menten kinetics (see figure 6.9). In this instance

the fluorometric assay revealed competitive inhibition, demonstrated by a dramatic increase in K_M value ($K_M100 \mu\text{M} > K_M0 \mu\text{M}$) with no significant change in V_{max} (see table 6.7) [549].

Table 6.7: Kinetics data for MC/EC metabolism (resulting in HC) by Megachilidae CYP9s co-incubated with thiacloprid or imidacloprid

	Conc substrate (μM)	K_M (μM)	95% CI	V_{max} ($\text{pmol}/\text{min}^{-1} \text{mg protein}^{-1}$)	95% CI	Adjusted R^2
<i>O. bicornis</i> CYP9BU1 + IMI	100	144.9	87.64-271.3	10.40	8.011-15.11	0.9255
	0	200.4	123.0-378.4	13.82	10.47-20.97	0.9493
<i>O. lignaria</i> + TCP	100	N/A	N/A	N/A	N/A	N/A
	0	41.19	24.65-69.85	1.157	0.9561-1.427	0.8557
<i>O. lignaria</i> + IMI	100	N/A	N/A	N/A	N/A	N/A
	0	15.97	8.157-29.71	0.9241	0.7626-1.120	0.7425
<i>H. truncorum</i> + IMI	100	138.8	106.8-185.3	20.60	17.94-24.23	0.9645
	0	196.2	140.1-292.5	13.20	10.87-17.00	0.9591
<i>C. florisomne</i> + TCP	100	N/A	N/A	N/A	N/A	N/A
	0	31.92	22.34-45.43	5.267	4.673-5.973	0.9104
<i>C. florisomne</i> + IMI	100	504.5	97.05-infinity	4.395	1.501-infinity	0.7151
	0	19.67	14.24-26.75	4.724	4.296-5.199	0.9132

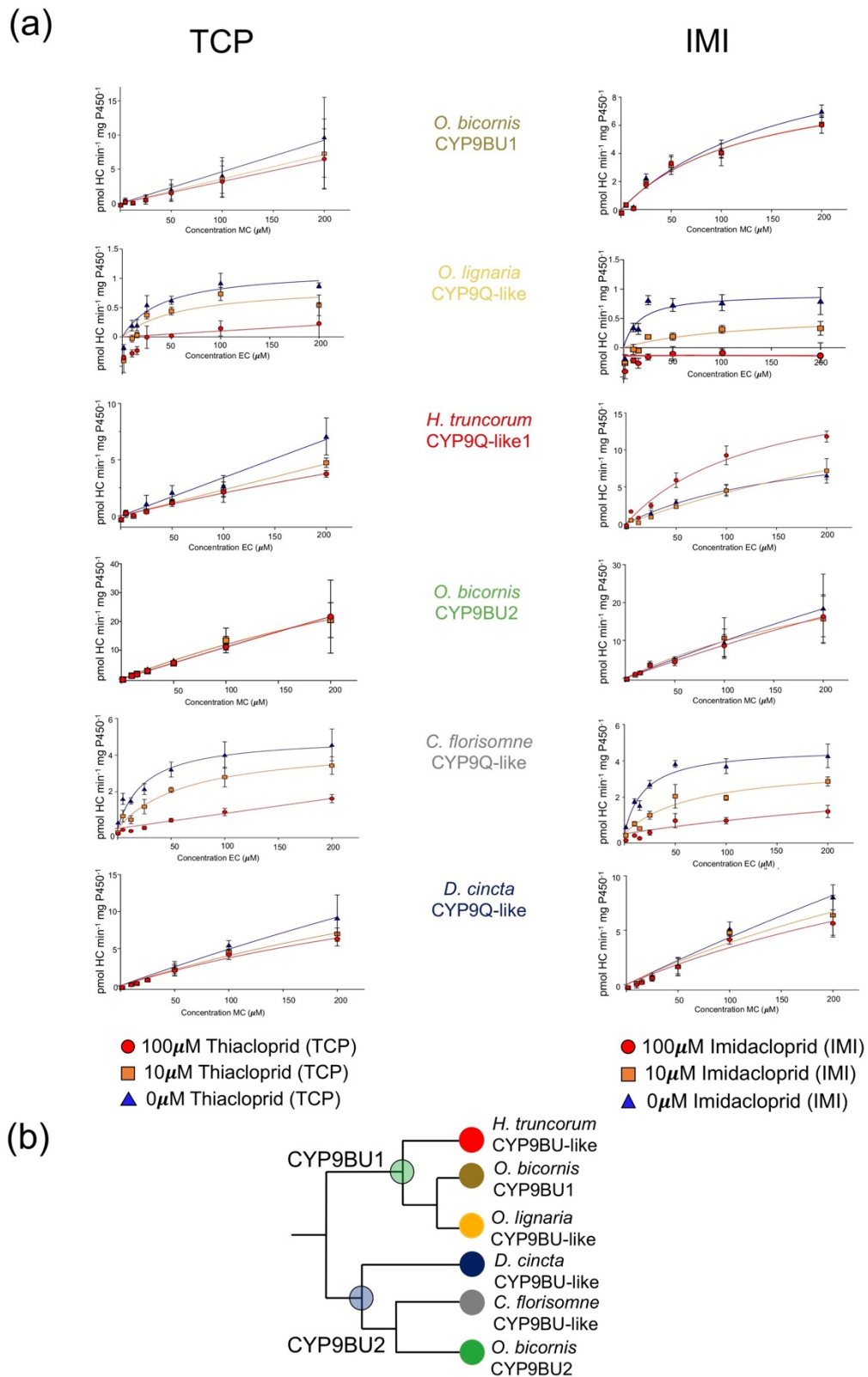


Figure 6.10: (a) Insecticide mediated inhibition of 7-hydroxycoumarin (HC) formation by recombinantly expressed Megachilidae CYP9BU-like enzymes incubated with different concentrations of thiacloprid (TCP) and imidacloprid (IMI). Data are mean values \pm S.D. (n=4). (b) Schematic of the phylogenetic relationship between the CYP9BU genes.

6.3.5 The phylogeny of the CYP9 subfamily across the Hymenoptera

BLASTn searches of the NCBI genome database were performed in June and July 2020. The list of species chosen for inclusion in the phylogenetic analyses is shown in table 6.8. In total 436 sequences were discovered, of which 96 were excluded because they were only partial, which left 340 full length sequences for translation which, with the 466 sequences found in bee species (see section 6.3.1), made a total of 806 sequences for inclusion in an MSA and phylogenetic analyses.

Table 6.8: List of species of wasp, ant and sawfly species included in phylogenetic analyses.

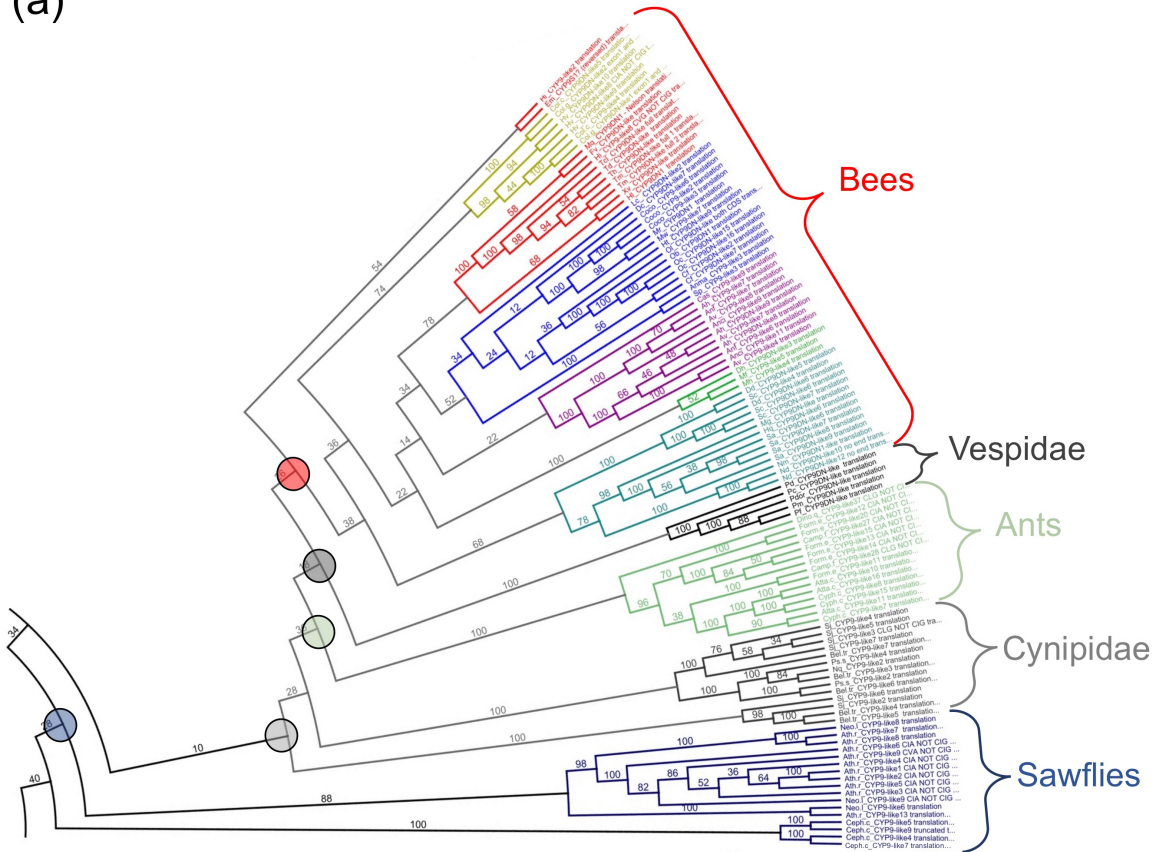
Species	Abbreviation used	Family or subfamily	Genome assembly release (NCBI)
Wasps			
<i>Belonocnema treatae</i>	Bel.tr	Cynipoidea	B_treatae_v1
<i>Leptopilina heterotama</i>	Lep.h	Cynipoidea	ASM1001604v1
<i>Neuroterus quercusbaccarum</i>	Nq	Cynipoidea	Neuqba
<i>Pseudoneuroterus saliens</i>	Ps.s	Cynipoidea	Neusal
<i>Synergus japonica</i>	Sj	Cynipoidea	Synjap
<i>Copidosoma floridanum</i>	Cop.f	Chalcidoidea	Cflo_2.0
<i>Ceratosolen solmsi marchali</i>	Csm	Chalcidoidea	CerSol_1.0
<i>Nasonia giraulti</i>	Ng	Chalcidoidea	Ngir_1.0
<i>Nasonia vitripennis</i>	Nv	Chalcidoidea	Nvit_psr_1.1
<i>Trichogramma pretiosum</i>	Trich.p	Chalcidoidea	Tpre_2.0
<i>Polistes canadensis</i>	Pc	Vespoidea	ASM131383v1
<i>Polistes dominula</i>	Pd	Vespoidea	Pdom r1.2
<i>Polistes dorsalis</i>	Pdor	Vespoidea	CU_Pdor_10
<i>Polistes fuscatus</i>	Pf	Vespoidea	CU_Pfus_HIC
<i>Polistes metricus</i>	Pm	Vespoidea	CU_Pmet_PB
<i>Goniozus legneri</i>	Gl	Chrysoidea	ASM305509v1
<i>Cotesia vestalis</i>	Cv	Ichneumonoidea	ASM95615v1
<i>Diachasma alloeum</i>	Da	Ichneumonoidea	Dall2.0
<i>Fopius arisanus</i>	Fa	Ichneumonoidea	ASM80636v1
<i>Macrocentrus cingulum</i>	Mac.c	Ichneumonoidea	MCINOVS1.0
<i>Microplitis demolitor</i>	Md	Ichneumonoidea	Mdem2
Ants			
<i>Atta cephalotes</i>	Atta.c	Myrmicinae	Attacep1.0
<i>Camponotus floridanus</i>	Camp.f	Formicinae	Cflo_v7.5
<i>Cyphomyrmex costatus</i>	Cyph.c	Myrmicinae	Ccos1.0
<i>Dinoponera quadriceps</i>	Dino.q	Ponerinae	ASM131382v1
<i>Formica exsecta</i>	Form.e	Formicinae	ASM365146v1
Sawflies			
<i>Athalia rosae</i>	Ath.r	Tenthredinidae	Aros_2.0
<i>Cephus cinctus</i>	Ceph.c	Cephidae	Ccin1
<i>Neodiprion lecontei</i>	Neo.l	Diprionidae	Nlec1.0

The Jones-Taylor-Thornton (JTT) with gamma rate (1.358) and invariant sites (0.01) was selected for use in estimating phylogeny of the Hymenopteran CYP9 subfamily protein sequences (see appendix table 6.4).

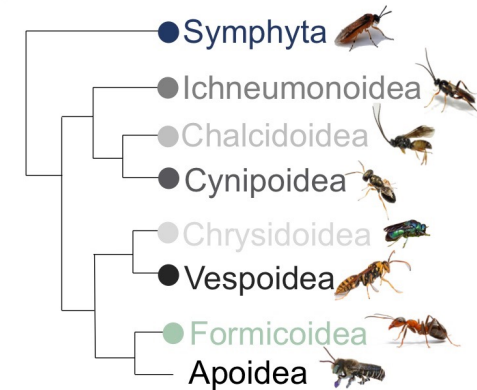
6.3.5.1 The *CYP9DN* lineage

The *CYP9DN1* gene, in bees, is found on a separate scaffold to the CYP9 cluster. A close-up of the *CYP9DN* clade of the PhyML phylogeny indicates that all three sawfly species (see table 6.8) have duplicated the *CYP9DN* gene, although the duplication look to be species-specific rather than shared ancestral duplications (see figure 6.11 (a)). Two of the wasp superfamilies have *CYP9DN* sequences. All five species of Cynipoidea (see table 6.8) have at least one *CYP9DN* gene; *N. quercusbaccarum* has a single sequence whereas *S. japonica* has six. It appears as if one ancestral duplication of CYP9DN may have occurred in this superfamily as *B. treatae* and *S. japonica* have genes that form two clades with the lineage (see figure 6.11 (a)). The five Vespiodea species all have a single *CYP9DN* sequence. All five species of ant (see table 6.8) have at least one *CYP9DN* gene. *D. quadriceps* has a single *CYP9DN* gene that is sister to the other ant sequences. *A. cephalotes* and *C. costatus* each have three *CYP9DN* sequences that clade together, indicating these duplications are likely to be species-specific, whereas the two Formicinae species (see table 6.8) have genes that appear in the two clades of ant sequences (see figure 6.11 (a)). *C. floridanus* has a single gene in each clade, whereas *F. exsecta* has five sequences in one of the clades (see figure 6.11 (a)). So, it appears that there may have been an ancestral duplication of the *CYP9DN* gene that occurred after the divergence of the Formicinae and Myrmicinae from the Ponerinae. No Chalcidoidea, Ichneumonoidea, or Chrysoidea species had *CYP9DN* sequences.

(a)



(b) Hymenoptera:



(c) Bees coloured by family:

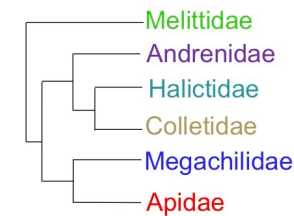


Figure 6.11: (a) PhyML maximum likelihood phylogeny [453] (substitution model JTT+G [627] with branch support of 50 bootstraps), of the CYP9DN lineage, across the Order: Hymenoptera. Tree rooted on *M. domestica* CYP4ae1. Sequences are coloured by superfamily (see table 6.8), apart from bees which are coloured by family. Nodes that mark the divergence of superfamilies are denoted by coloured circles. All nucleotide sequences accessed from NCBI databases. (b) Schematic of the phylogenetic relationship of the Hymenoptera. Photos used: *Athalia rosae* [590], *Cotesia vestalis* [592], *Nasonia* sp. [593], *Microplitis demolitor* [594], *Chrysis* sp. [595], *Polistes rothneyi* [596], ant [597]. (c) Schematic of the phylogenetic relationship between bee families.

6.3.5.2 The *CYP9P* lineage

Both the PhyML [453] phylogeny (see appendix figure 6.4) and the Bayesian inference [445] phylogeny of the CYP9 cluster (see figure 6.12 (a)) place the *CYP9P* lineage as sister to the other CYP9 lineages. All the wasp (Cynipoidea, Chalcidoidea, Vespoidea, Chrysoidea and Ichneumonoidea), ant and sawfly species have at least one *CYP9P* sequence. There are two separate clades of sawfly CYP9Ps which indicates a probable ancestral duplication. All five species of the Vespoidea and Ichneumonoidea only have a single *CYP9P* sequence. However, there is evidence of duplication within the other superfamilies of wasps. For example, *G. legneri*, the only Chrysoidea species, has eight *CYP9P* sequences (see figure 6.12 (a)). It also appears that species-specific duplication has occurred in the Cynipoidea. Although *S. japonica*, *N. quercusbaccarum* and *P. saliens* only have a single *CYP9P* sequence, *B. treatae* has two and *L. heterotama* has eight. In the Chalcidoidea only *C. solmsi marchali* has a single *CYP9P* sequence, with the other species having at least two. *C. floridanum* and *T. pretiosum* have *CYP9Ps* that appear in two clades, indicating possible ancestral duplication in this superfamily. In the ants the two Myrmicinae species have a single *CYP9P*. The Formicinae show lineage-specific duplication, with *C. floridanus* having two and *F. exsecta* five sequences. However, *D. quadriceps* has the largest contingent of *CYP9P* genes of all the species included in the phylogeny, with 15 sequences.

6.3.5.3 The *CYP9R* lineage

It is not clear from looking at both the PhyML and Bayesian inference phylogenies where the ancestral node for the *CYP9R* lineage should be placed. The PhyML phylogeny of the entire CYP9 subfamily (see appendix figure 6.4) shows the *CYP9R* sharing an ancestral sequence with ants, Chrysoidea, Vespoidea and Ichneumonoidea, whereas the Bayesian inference phylogeny of the CYP9 cluster shows the *CYP9R* and *CYP9S* lineages as bee-specific (see figure 6.12 (a)). To help resolve this, the *CYP9R*, *CYP9S* and *CYP9Q/BU/DL* protein sequences in bees were aligned with the Hymenopteran sequences that were related to them in (i.e. no *CYP9P* sequences included) and a further PhyML phylogeny was estimated, see appendix figure 6.5. All the wasp and ant

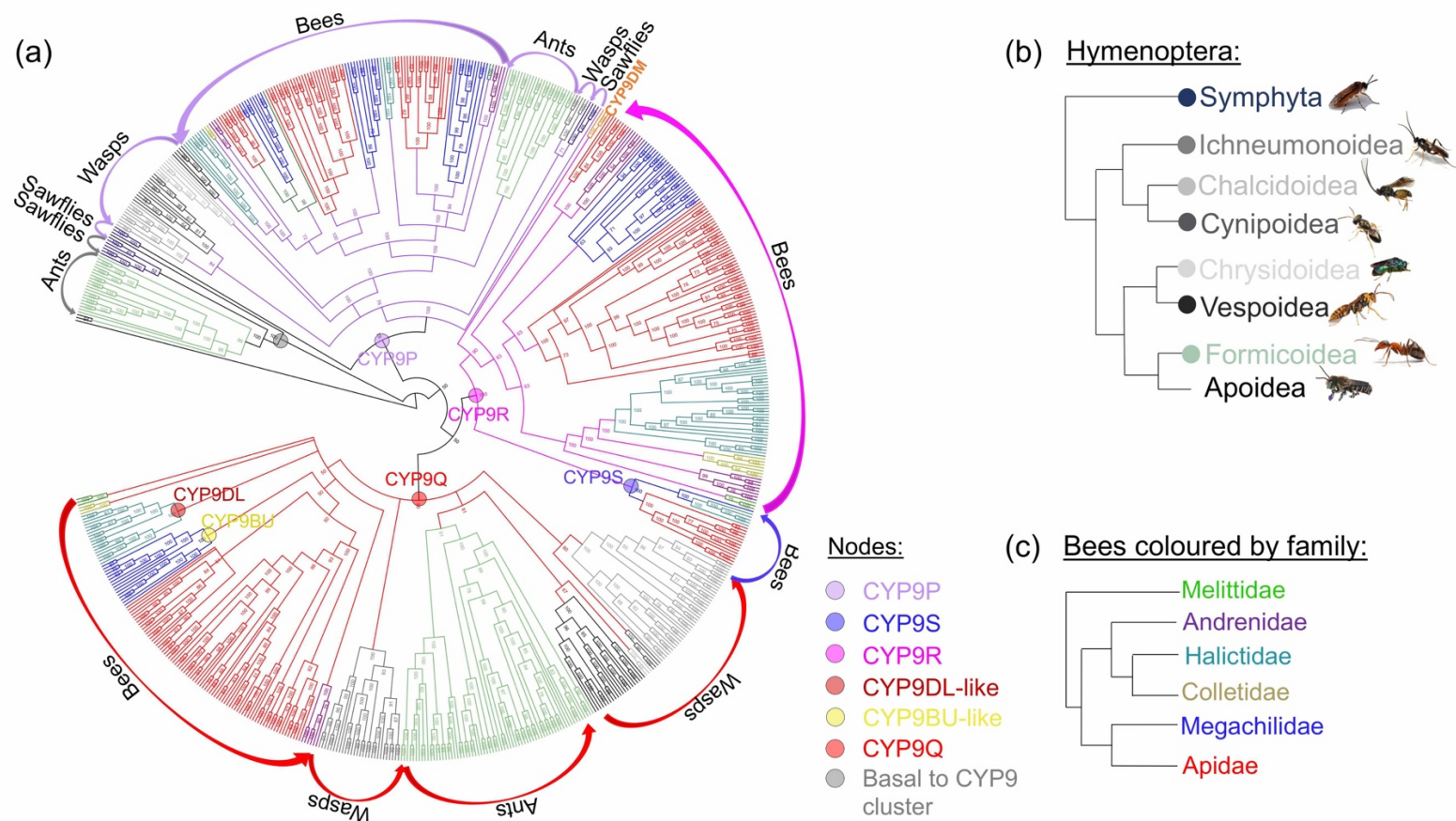


Figure 6.12: (a) Bayesian inference [445] phylogeny of the CYP9 cluster across Hymenoptera species, using substitution model JTT (G+I) [627]. Posterior probability of nodes shown as a % probability. Tree rooted on *M. domestica* CYP4ae1. Sequences are coloured by superfamily, apart from bees which are coloured by family. All nucleotide sequences accessed from NCBI databases. Nodes that mark the divergence of CYP9 lineages are denoted by coloured circles. (b) Schematic of the phylogenetic relationship of the Hymenoptera. Photos used: *A. rosae* [590], *C. vestalis* [592], *Nasonia* sp. [593], *M. demolitor* [594], *Chrysis* sp. [595], *P. rothneyi* [596], ant [597]. (c) Schematic of the phylogenetic relationship between bee families.

sequences in this phylogeny shared an ancestral node with the *CYP9Q/BU/DL* lineage, which implies that the *CYP9R* and *CYP9S* lineages are bee-specific. This phylogeny also shows that, in bees, the *CYP9Q/BU/DL* clade is ancestral to the *CYP9R* lineage.

6.3.5.4 The *CYP9Q/BU/DL* lineage

Species from each of the five superfamilies of wasp, included in the phylogeny, have sequences that share an ancestral node with the *CYP9Q/BU/DL* clade in bees (see table 6.9 and figure 6.13 (a)). Apart from the Cynipoidea, in which only two out of the five species have sequences that appear in this clade, all other wasp species have at least one gene that shares an ancestral node with the *CYP9Q/BU/DL* clade (see table 6.9 and figure 6.13 (a)). The sequences of the two Cynipoidea species appear together in a single clade, and so it is likely that the duplication events that have occurred are species-specific. The sequences of all the Vespoidea species appear in four distinct clades, although it should be noted that they are all from the same genus, *Polistes*. There is also good evidence for at least two discreet clades of CYP9 genes in this lineage in the Ichneumonoidea and Chalcidoidea. The formation of well-defined clades, each of which includes sequences from multiple species, indicates ancestral duplication before the radiation of the species.

In the ant species, this gene lineage appears to have undergone a CYP bloom (see table 6.9 and figure 6.13 (a)). This is particularly obvious in *C. floridanus* which has 23 sequences. Some of the duplication events appear to precede the divergence of the species, as evidenced by at least three distinct clades (see table 6.9 and figure 6.13 (a)). In this instance, the numbers of CYP9 sequences per species does not appear to be predicted by subfamily (see table 6.9).

Apart from a single *M. willughbiella* sequence, no species included in any of the phylogenetic analyses has a CYP9 gene that clades with *M. rotundata* *CYP9DM1* and *CYP9DM2*. The three *CYP9DM* sequences discovered appear to be ancestral to the *CYP9R* lineage rather than to be part of the *CYP9Q/BU/DL* lineage. All the wasp and ant CYP9 sequences shown in figure 6.13 (a), therefore, appear to be more closely related to the bee *CYP9Q/BU/DL* lineage than the *Megachile* *CYP9DM* genes are.

Table 6.9: Numbers of wasp and ant CYP9 sequences that share an ancestral node with the CYP9Q/BU/DL clade in bees.

Species	Family or subfamily	Number of CYP9 sequences	Number of CYP9Q/BU/DL clades
Wasps			
<i>Cotesia vestalis</i>	Ichneumonoidea	3	1
<i>Diachasma alloenum</i>	Ichneumonoidea	8	3
<i>Fopius arisanus</i>	Ichneumonoidea	6	3
<i>Macrocentrus cingulum</i>	Ichneumonoidea	5	2
<i>Microplitis demolitor</i>	Ichneumonoidea	5	2
<i>Goniozus legneri</i>	Chrysoidea	3	1
<i>Polistes canadensis</i>	Vespoidea	4	4
<i>Polistes dominula</i>	Vespoidea	4	4
<i>Polistes dorsalis</i>	Vespoidea	4	4
<i>Polistes fuscatus</i>	Vespoidea	5	4
<i>Polistes metricus</i>	Vespoidea	5	4
<i>Copidosoma floridanum</i>	Chalcidoidea	9	2
<i>Ceratosolen solmsi marchali</i>	Chalcidoidea	1	1
<i>Nasonia giraulti</i>	Chalcidoidea	7	2
<i>Nasonia vitripennis</i>	Chalcidoidea	14	2
<i>Trichogramma pretiosum</i>	Chalcidoidea	9	2
<i>Belonocnema treatae</i>	Cynipoidea	2	1
<i>Synergus japonica</i>	Cynipoidea	2	1
Ants			
<i>Atta cephalotes</i>	Myrmicinae	10	3
<i>Camponotus floridanus</i>	Formicinae	23	4
<i>Cyphomyrmex costatus</i>	Myrmicinae	12	4
<i>Dinoponera quadriceps</i>	Ponerinae	17	3
<i>Formica exsecta</i>	Formicinae	6	4

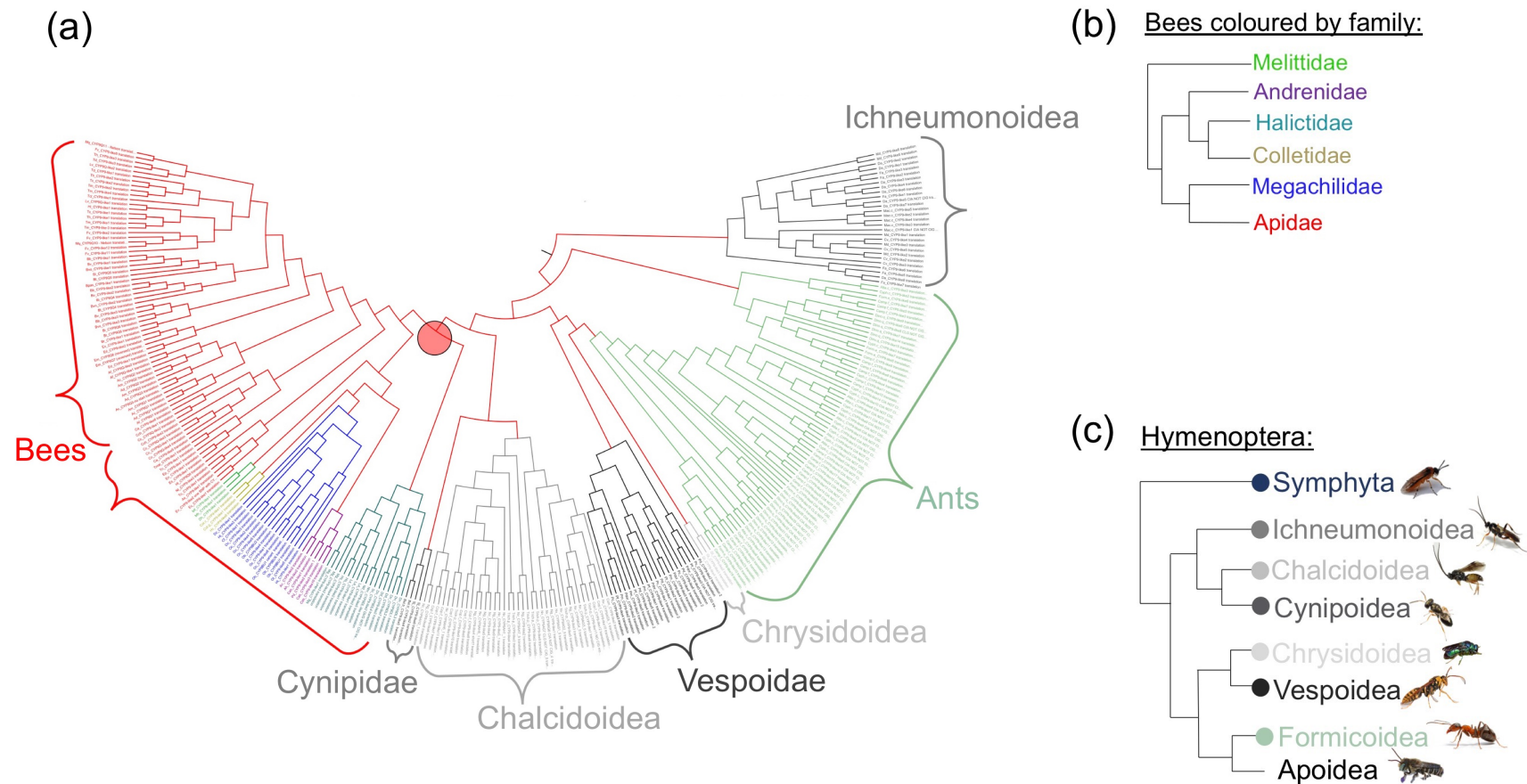


Figure 6.13: (a) PhyML maximum likelihood phylogeny [453] (substitution model JTT+G [627]) of the CYP9Q/BU/DL lineage, across the Order: Hymenoptera. Tree rooted on *M. domestica* CYP4ae1. Sequences are coloured by superfamily (see table 6.8), apart from bees which are coloured by family. Ancestral node of bees coloured with a red circle. All nucleotide sequences accessed from NCBI databases. (b) Schematic of the phylogenetic relationship between bee families. (c) Schematic of the phylogenetic relationship of the Hymenoptera. Photos used: *A. rosae* [590], *C. vestalis* [592], *Nasonia* sp. [593], *M. demolitor* [594], *Chrysis* sp. [595], *P. rothneyi* [596], ant [597].

6.3.6 Conserved synteny analysis of the *membralin* /*alpha catulin*-associated CYP9 cluster across Hymenoptera

The CYP9 scaffolds for each species of wasp, ant and sawfly species included in the phylogenetic analyses were investigated for evidence of a shared conserved synteny with the CYP9 cluster, identified in bees in section 3.3.3.4 and illustrated in figure 6.5. All five species of ant showed good conservation of the syntenic block containing the CYP9 cluster (see table 6.10 and figure 6.14). The evolutionary distance between ants and bees is ~140 mya, but both relative gene orientation and gene order have been conserved. As seen in figure 6.13 (a), there is evidence for a CYP bloom in the CYPs in ant species, and the syntenic analysis here, confirms that these duplications have taken place within the CYP9 cluster syntenic block (see table 6.10 and figure 6.14).

The genome of the only Chrysidoidea species included in the phylogeny, *G. legneri*, is not annotated and as such synteny analysis was not undertaken for this species. None of the Ichneumonidae or Cynipoidea species had any CYP9s associated with *membralin*, *alpha catulin* or *myosin IIIb* (i.e. the genes that flank the CYP9 cluster in bees). In fact, where these genes were annotated in the Ichneumonidae and Cynipoidea, they appeared on separate scaffolds. The Ichneumonidae sequences appear in a clade that is basal to the rest of the phylogeny (see figure 6.13 (a)). The Cynipoidea sequences group as a sister clade to the Chalcidoidea.

The two remaining superfamilies of wasps included in the phylogenetic analyses, the Chalcidoidea and the Vespoidea, both showed good conservation of the syntenic block containing the CYP9 cluster (see table 6.10). CYP9 genes are consistently found associated with *membralin* in both superfamilies, with *myosin IIIb* and *alpha catulin* found downstream, as is the case across the bee families (see table 6.10 and figure 6.14). The genome of *Vespa mandarinia* became available in the NCBI database in August 2020, as such the species was included in the syntenic analyses to add a non-*Polistes* species, although it does not appear in the phylogeny. Interestingly, in the all Vespoidea species that had annotated genes, the direction of the reading frame and the gene order of *alpha catulin* and *myosin IIIb* were reversed compared to bees and chalcids (see figure 6.14). The last common ancestor of Vespoidea wasps and bees

existed ~150 mya, but for the Chalcidoidea the evolutionary distance is probably over 200 mya (see figure 6.1).

Table 6.10: Scaffolds from the genomes of wasp, ant and sawfly species containing syntenic blocks of genes also found in bees.

Species	NCBI scaffold accession number	<i>Alpha catulin</i>	<i>Myosin IIIb</i>	Number of CYP9s	<i>Membralin</i>
Chalcidoidea					
<i>C. floridanum</i>	NW_019379459			1	✓
<i>C. solmsi marchali</i>	NW_011948421	✓		1	✓
<i>N. giraulti</i>	Not annotated				
<i>N. vitripennis</i>	Chr5 NC_045761	✓	✓	6	✓
<i>T. pretiosum</i>	NW_019641630	✓	✓	2	✓
Vespoidea					
<i>P. canadensis</i>	NW_014569587	✓	✓	4	✓
<i>P. dominula</i>	NW_015149002	✓	✓	7	✓
<i>P. dorsalis</i>	Not annotated				
<i>P. fuscatus</i>	Not annotated				
<i>P. metricus</i>	Not annotated				
<i>Vespa mandarinia</i>	NW_023395841	✓	✓	10	✓
Ants					
<i>A. cephalotes</i>	NW_012130066	✓	✓	4	✓
<i>C. floridanus</i>	NW_020229325	✓	✓	26	✓
<i>C. costatus</i>	NW_017280322	✓	✓	6	✓
<i>D. quadriceps</i>	NW_014555795	✓	✓	10	✓
<i>F. exsecta</i>	NW_021849649	✓	✓	10	✓
Sawflies					
<i>A. rosae</i>	NW_020311375	✓	✓	1	
	NW_020311970			1	✓
<i>C. cinctus</i>	NW_014333138	✓	✓	1	
	NW_014332979			1	✓
	NW_014333322			4	✓
<i>N. lecontei</i>	NW_015385094	✓	✓	1	
	NW_015384993			1	✓

The three species of sawfly included in the phylogenetic analysis all have at least one CYP9 gene associated with *membralin*. *Cephus cincta* has two *membralin* sequences (*membralin*: LOC107265992 and *membralin*-like LOC107270115), one of which is associated with four CYP9 sequences (see

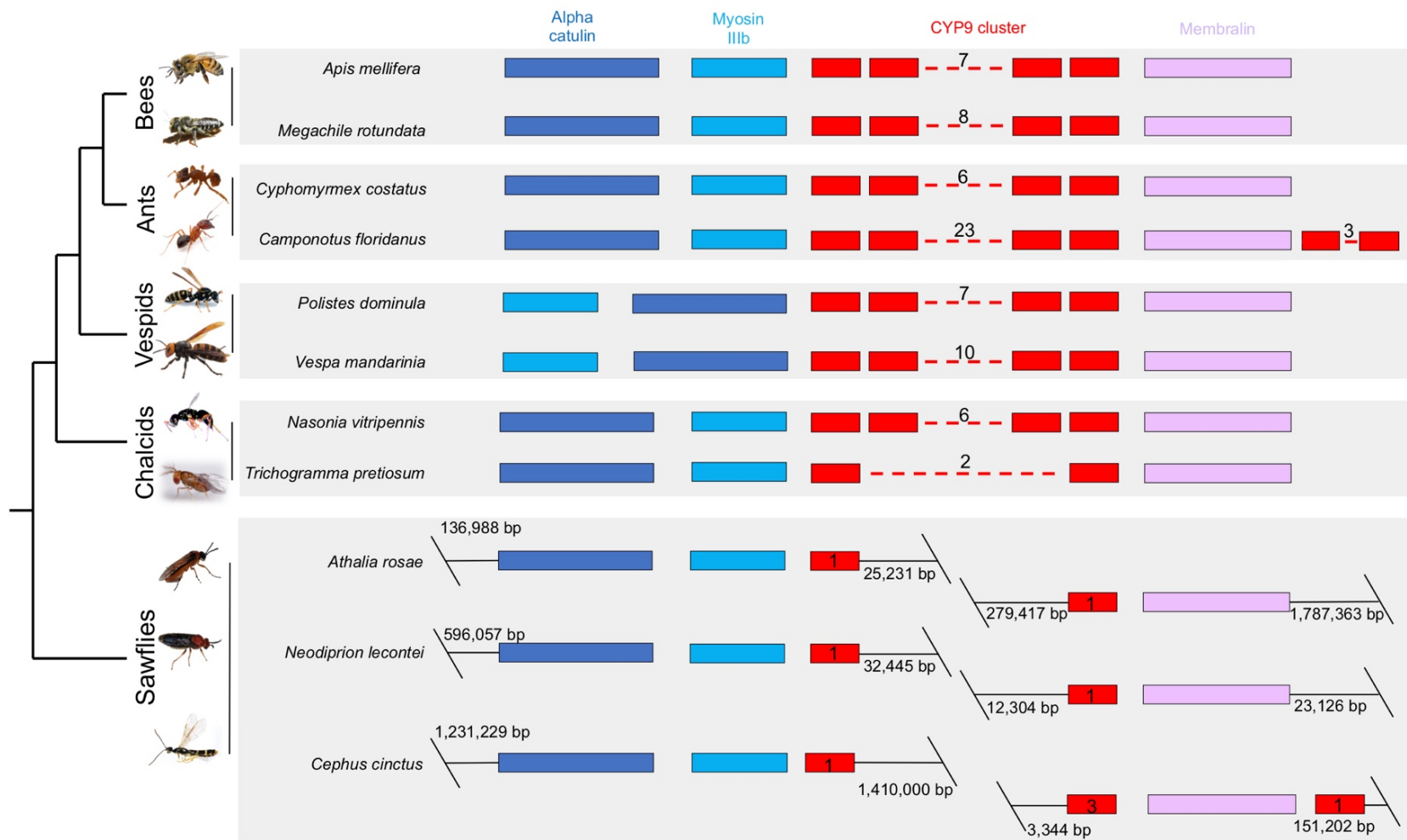


Figure 6.14: Conserved tandem arrangements of the membrainin/ alpha catulin-associated CYP9 loci across the Hymenoptera. Photos used: *A. mellifera*, [401], *M. rotundata* [406], *C. costatus* [628], *C. floridanus* [629], *P. dominula* [630], *V. mandarinia* [631], *N. vitripennis* [632], *T. pretiosum* [633], *Athalia rosae* [590], *Neodiprion lecontei* [634], *Cephus cinctus* [591].

table 6.10). *Alpha catulin* and *myosin IIIb* appear together on a different scaffold in all three species and, in each case, are associated with a single CYP9 sequence (see table 6.10 and figure 6.14). For all the species included in the phylogeny, shown in figure 6.12, where genomic data is available, the CYP9 sequence found associated with *membralin* clades with the *CYP9P* lineage.

The degree of similarity (% identity) of these three flanking-genes across the Hymenopteran species included in figure 6.14, is shown in table 6.11. It appears that *alpha catulin* is extremely well conserved across the order, ranging from 74% (chalcid to sawfly) to 99% (sawfly to sawfly) identity (see table 6.11).

Table 6.11: Percent identity with *A. rosae* protein sequences for the primary flanking genes from the syntenic block that includes the CYP9 cluster across Hymenopteran species. Data generated by alignment of protein sequences in Geneious using MUSCLE (version 3.5, default settings) [451].

Species	% identity to <i>A. rosae</i> sequence		
	<i>Membralin</i>	<i>Myosin IIIb</i>	<i>Alpha catulin</i>
Bees			
<i>A. mellifera</i>	56.812	65.932	82.005
<i>M. rotundata</i>	53.474	69.964	84.830
Ants			
<i>C. costatus</i>	59.796	65.782	84.878
<i>C. floridanus</i>	59.490	69.798	85.158
Vespoidea			
<i>P. dominula</i>	54.254	67.606	79.254
<i>V. mandarinia</i>	59.852	67.995	82.491
Chalcidoidea			
<i>N. vitripennis</i>	55.297	67.364	74.543
<i>T. pretiosum</i>	53.797	63.529	83.016
Sawflies			
<i>N. lecontei</i>	72.306	88.332	99.081
<i>C. cinctus</i>	65.894 33.113*	73.983	85.680

* This value is for *C. cinctus membralin*-like which is a partial sequence – only 313 amino acid residues long.

6.3.7 Transcriptomics results from four species of *Megachile* bees

The repertoire of CYP9 sequences identified in each of the *Megachile* bees is shown in table 6.12. The CYP9 subfamily in these species is broadly similar to that observed in *M. rotundata* (see figure 6.15 and table 6.12).

6.3.7.1 The *CYP9DN* lineage

All four species have a *CYP9DN* sequence (see figure 6.15 and table 6.12). This lineage has the second longest ancestral branch length (0.49 substitutions per site) which implies a large degree of evolutionary distance between *CYP9DN* sequences and the other *CYP9* lineages.

Table 6.12: Inventory of the lineages of the *CYP9* subfamily of *A. mellifera*, *B. terrestris*, *O. bicornis*, *D. novaeangliae* and five species of *Megachile* bees.

Species	<i>CYP9DN</i>	<i>CYP9DM</i>	<i>CYP9P</i>	<i>CYP9Q/BU/DL</i>	<i>CYP9R</i>	<i>CYP9S</i>
<i>A. mellifera</i>	0	0	2	3	1	1
<i>B. terrestris</i>	0	0	2	3	1	0
<i>D. novaeangliae</i>	0	0	2	4	2	1
<i>O. bicornis</i>	1	0	3	2	3	0
<i>M. rotundata</i>	1	2	3	0	3	0
<i>M. centuncularis</i>	1	1	3	0	11	0
<i>M. lapponica</i>	1	1	3	0	6	0
<i>M. leachella</i>	1	3	3	0	7	0
<i>M. willughbiella</i>	1	4	4	0	3	0

6.3.7.2 The *CYP9P* lineage

Three of the *Megachile* species have three *CYP9P* sequences, but *M. willughbiella* has four (see figure 6.15 and table 6.12). The *CYP9P* sequences assemble into three clades: *CYP9P2*, *CYP9P22* and *CYP9P23* (see figure 6.15). The three distinct *CYP9P* clades, support the earlier finding that there was a duplication event in the *CYP9P* lineage before the divergence of the bee families (see section 6.3.1.2). The two *M. willughbiella* nucleotide sequences that clade into the *CYP9P23* cluster are 99.74% identical. There are 4 single nucleotide polymorphisms (SNP) between the two sequences, three of which are synonymous and one is non-synonymous. The non-synonymous SNP creates an amino acid change at position 487 of the protein: serine to threonine (S – T).

6.3.7.3 The *CYP9R* lineage

M. rotundata and *O. bicornis* have three *CYP9R* genes, although these group into two main clades: *CYP9R1/38/58* and *CYP9R39/59* (see figure 6.15). This duplication occurred before the divergence of the two species. However, *A. mellifera* and *B. terrestris* have only a single *CYP9R* gene, and although *D. novaeangliae* has two, they form a distinct clade. This implies that at least some

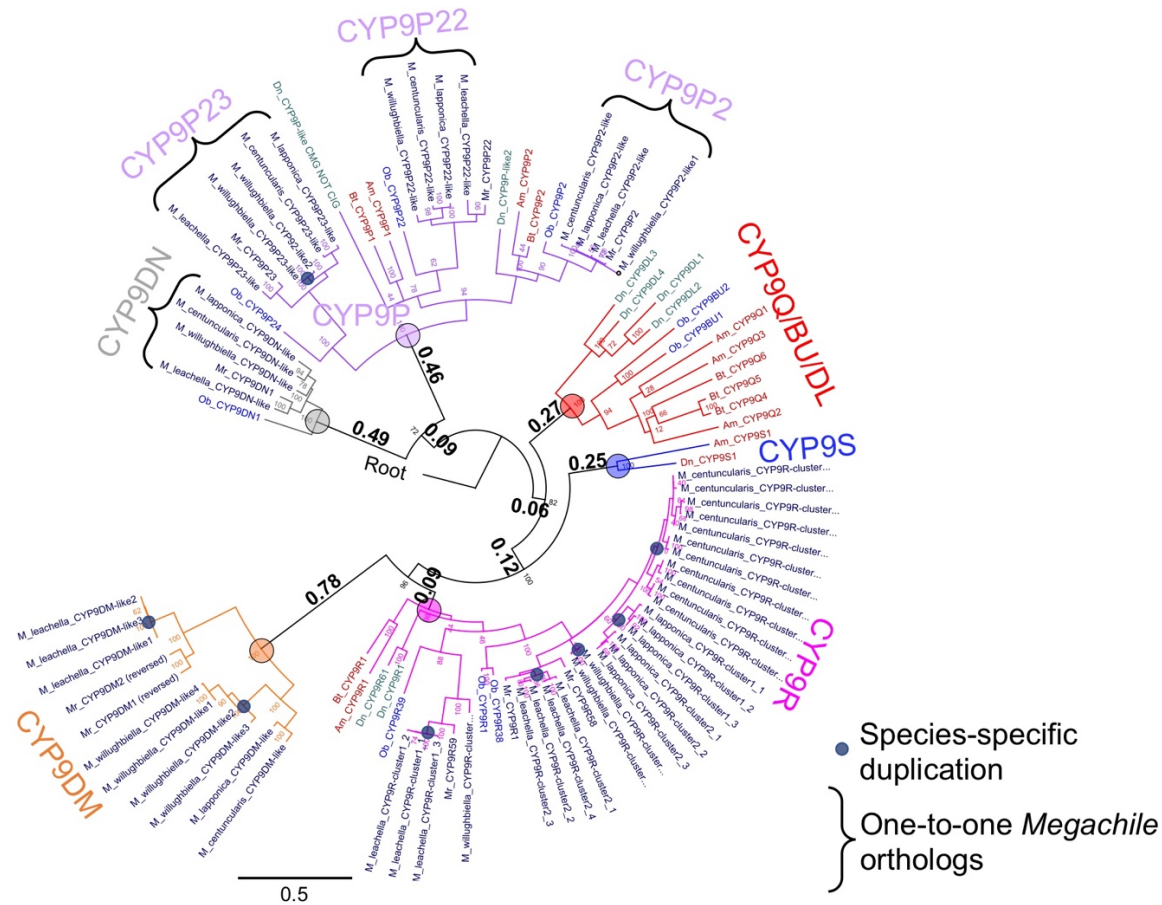


Figure 6.15: PhyML maximum likelihood phylogeny [453] (using substitution model LG+G [454], branch support of 50 bootstraps) of the CYP9 subfamily sequences from *A. mellifera*, *B. terrestris*, *O. bicornis*, *D. novaeangliae*, and five species of *Megachile* bee. Tree rooted on *N. vitripennis* CYP9AG4. Apidae and Halictidae sequences coloured by family: red and dark cyan respectively. Megachilidae bees coloured by genus, *Osmia* and *Megachile*: light blue and dark blue respectively. Branch lengths (substitutions per site) shown for ancestral nodes of each CYP9 lineage.

of the duplications in this lineage are species-specific (i.e. arising after they diverge). The short branch lengths in the *CYP9R* lineage also indicate close sequence identity and recent divergence (see figure 6.15).

Although the manual curation of the RNAseq contigs shows high numbers of *CYP9R* genes in the four *Megachile* species, it is likely that this is an anomaly. When these sequences are aligned, it is clear that they cluster into distinct clades (see appendix figure 6.6). *M. lapponica* and *M. willughbiella* have six and three sequences respectively, that fall into two clades (see table 6.13). *M. centuncularis* and *M. leachella* have eleven and seven sequences respectively, that fall into three clades (see figure 6.15 and table 6.13).

Table 6.13: Percentage identity and number of non-synonymous changes between the *CYP9R* sequences from the RNAseq data for *M. centuncularis*, *M. lapponica*, *M. leachella* and *M. willughbiella*.

Species	Clade 1 % identity (number of non-synonymous changes)	Clade 2 % identity (number of non- synonymous changes)	Clade 3 % identity (number of non- synonymous changes)
<i>M. centuncularis</i>	95.349 – 99.031 (2-16)	94.380 – 98.643 (4-19)	98.643 (7)
<i>M. lapponica</i>	96.899 – 99.031 (2-5)	97.868 – 99.031 (1-3)	–
<i>M. leachella</i>	99.197 – 99.799 (1-4)	99.225 (1)	99.225 (1)
<i>M. willughbiella</i>	100 (only one sequence)	99.419 (3)	–

It is likely that the high number of *CYP9Rs* generated by the BLAST searches of the RNA-seq contigs in three of the species represents an artefact of *de novo* transcriptome assembly, and at least, some of these sequences represent different isoforms of the same gene. Further work, such as PCR verification, would likely resolve the numbers of *CYP9R* sequences in the *M. centuncularis* and *M. leachella* down [635]. However, for the purposes of determining the repertoire of the *CYP9* subfamily all the sequences were included in the analyses, and it is clear that the four species have at least two *CYP9R* sequences in their *CYPomes*.

6.3.7.4 The *CYP9S* lineage

None of the *Megachile* species has a *CYP9S* sequence (see table 6.12 and figure 6.15). Although this clade is ancestral to the *CYP9R* lineage, it is also quite distinct with an ancestral branch length of 0.25 substitutions per site (see figure 6.15).

6.3.7.5 The *CYP9Q/BU/DL* lineage

None of the *Megachile* species has a *CYP9Q/BU/DL* sequence (see table 6.12 and figure 6.15).

6.3.7.6 The *CYP9DM* lineage

The five *Megachile* species (*M. rotundata* and the four species collected for RNAseq) all have at least one *CYP9DM* sequence (see table 6.12 and figure 6.15). The ancestral branch length of this lineage is the longest at 0.78 substitutions per site, which supports the results from section 3.3.3.4 where a very long branch length for the *CYP9DMs* (1.1 substitutions per site) was also seen in a phylogeny across four bee families. The *CYP9DM* sequences group into two main clades by species, rather than sequence (see figures 6.15 and 6.16). *M. leachella* and *M. rotundata* clade together with percent identity ranging from 81.729 to 99.398%. *M. willughbiella*, *M. centuncularis* and *M. lapponica* group into the second clade with percent identity ranging from 75.449 to 99.402%.

	M_leachella CYP9DM3	M_leachella CYP9DM1	M_leachella CYP9DM2	Mr_CYP9DM1	Mr_CYP9DM2	M_willughbiella CYP9DM3	M_willughbiella CYP9DM2	M_willughbiella CYP9DM1	M_willughbiella CYP9DM4	M_centuncularis CYP9DM	M_lapponica CYP9DM
M_leachella_CYP9DM3		98.795	99.398	81.928	82.932	59.642	57.455	56.461	56.859	60.558	61.753
M_leachella_CYP9DM1	98.795		99.398	81.928	82.932	59.642	57.853	56.859	57.256	60.359	61.554
M_leachella_CYP9DM2	99.398	99.398		81.727	82.731	59.443	57.654	56.66	57.058	60.359	61.554
Mr_CYP9DM1	81.928	81.928	81.727		91.165	59.642	58.847	59.245	59.642	62.749	63.745
Mr_CYP9DM2	82.932	82.932	82.731	91.165		59.642	59.245	58.847	59.245	61.554	61.952
M_willughbiella_CYP9DM3	59.642	59.642	59.443	59.642	59.642		94.024	88.446	89.044	75.551	76.2
M_willughbiella_CYP9DM2	57.455	57.853	57.654	58.847	59.245	94.024		94.422	93.825	75.449	76.295
M_willughbiella_CYP9DM1	56.461	56.859	56.66	59.245	58.847	88.446	94.422		99.402	75.449	76.295
M_willughbiella_CYP9DM4	56.859	57.256	57.058	59.642	59.245	89.044	93.825	99.402		75.449	75.896
M_centuncularis_CYP9DM	60.558	60.359	60.359	62.749	61.554	75.551	75.449	75.449	75.449		93.186
M_lapponica_CYP9DM	61.753	61.554	61.554	63.745	61.952	76.2	76.295	76.295	75.896	93.186	

Figure 6.16: Heat map of the percent identity of *CYP9DM* sequences from five species *Megachile*.

In section 3.3.5 the structural homology of *CYP9Q*, *CYP9BU* and *CYP9DM* protein sequences was examined. A major substitution in the oxygen-binding motif (M2) unique to *M. rotundata* *CYP9DM* proteins was uncovered. In total there are six amino acid substitutions in helix I and helix L surrounding the

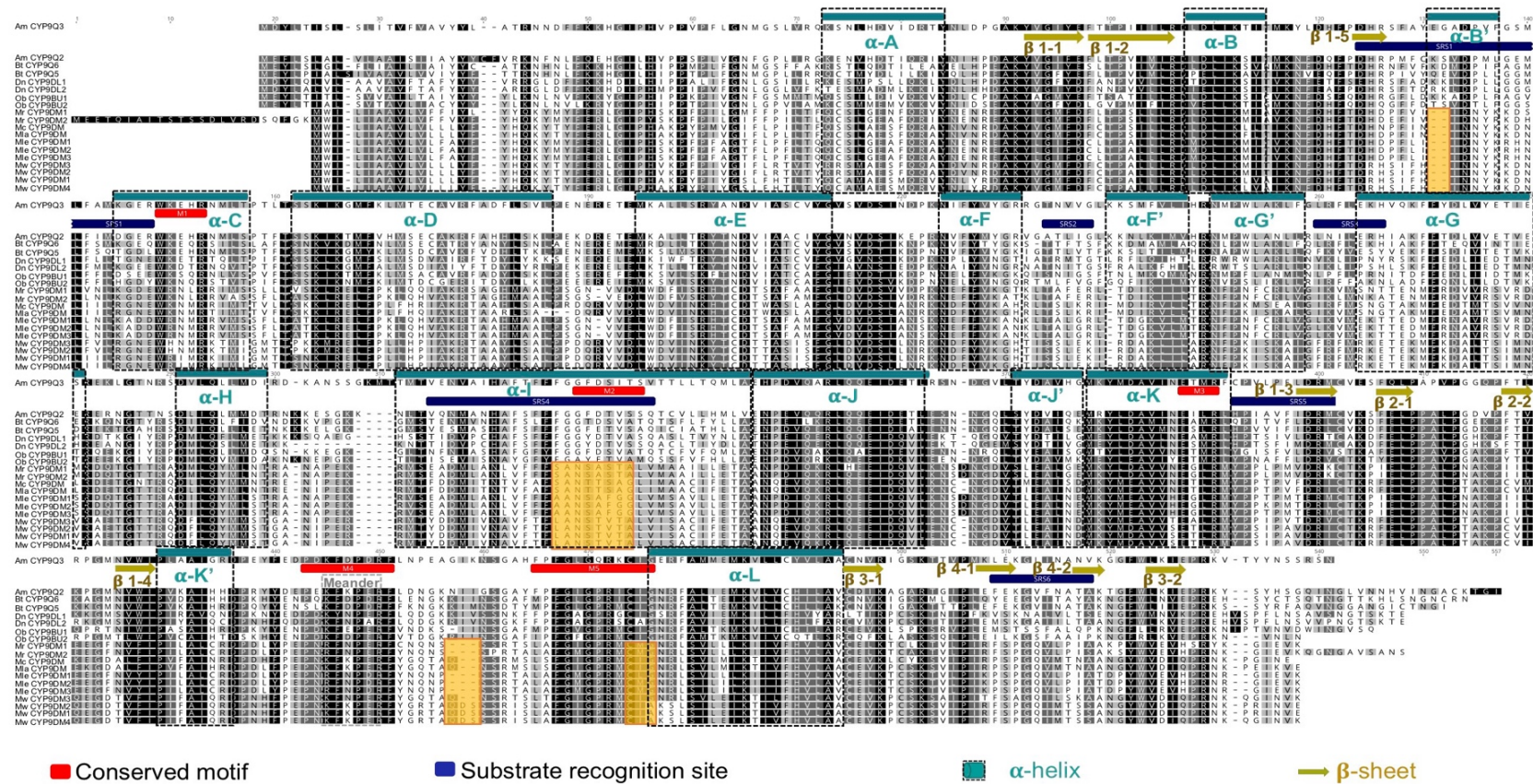


Figure 6.17: Multiple sequence alignment of: *A. mellifera*, *B. terrestris*, *O. bicornis* and *D. novaeangliae* CYP9Q/BU/DLs with *Megachile* species CYP9DM sequences. Aligned in Geneious version 10.2.3 (Biomatters) using MUSCLE [451] (version 3.5, default settings). The sequences are coloured black to white according to their similarity. Conserved motifs (M) and substrate recognition sites (SRS) are shaded red and blue respectively and represented by annotations below *A. mellifera* CYP9Q3. Secondary structures are annotated: dark cyan cylinders and dashed black boxes represent α -helices (with reference to the crystallographic structures of PDB: 4D6Z; 1TQN CYP3A4; *H. sapiens* [250, 254] and P450cam [248]). Amino-acid substitutions and *Megachile* specific gaps are highlighted in orange.

heme molecule in both *M. rotundata* CYP9DMs. To determine whether this structural change is also present in sequences from other *Megachile* species, all the CYP9DMs were aligned with *A. mellifera* CYP9Q2 and CYP9Q3; *B. terrestris* CYP9Q5 and CYP9Q6; *O. bicornis* CYP9BU1 and CYP9BU2 and *D. novaeangliae* CYP9DL1 and CYP9DL2 using MUSCLE [451]. The resulting MSA was annotated to show the secondary structures, conserved motifs (M1-5) and substrate recognition sites (SRS), and is shown in figure 6.17. All the CYP9DMs have substitutions in the M2 region (motif: G-x-E/D-T-T/S). The M2 motif in the *Megachile* species is: N-S/T-A/T-**S/F/V**-T/A/G, versus the CYP9Q/BU/DL sequences: G-F/T/Y-D/E-**T/S**-V (residues 329-334), highlighted in orange in figure 6.17. The two major substitutions are, firstly, the replacement of the conserved hydrophobic glycine (G) residue at the start of the oxygen-binding motif with a polar asparagine (N) and secondly, the loss of the charged acid residue (aspartic/glutamic acid; E/D), position 331 of the MSA, which has been replaced with a hydrophobic alanine (A). None of the *Megachile* CYP9DMs have the conserved threonine (**T**) at position 332 of the MSA, but a substitution here is not unique to the *Megachile* species as there is a polar-polar, threonine-serine substitution in CYP9BU1, CYP9Q3 and CYP9Q6 (see figure 6.17).

There is also an isoleucine (I) to methionine (M) substitution (both hydrophobic residues) in the heme-binding motif (M5) of *M. leachella* CYP9DM2 and CYP9DM3 and both *M. rotundata* CYP9DMs (position 474- 476 of the MSA), which changes the sequence immediately after the highly conserved cysteine; CIG becoming CMG (see figure 6.17). All four *M. willughbiella* CYP9DMs have a G to A substitution (both hydrophobic) at position 476 of the MSA which changes this part of the conserved motif from CIG to CIA (see figure 6.17). However, from the manual curation of the CYP9 sequences from the NCBI TSA database (see section 6.2.1.2), substitutions in this part of the M5 conserved motif (i.e. CIG - CIA- CMG) are not uncommon in bee species.

All the CYP9DM sequences have a three-residue deletion in SRS1 (position 131-132 of the MSA, shown highlighted in orange). *M. rotundata* and *M. leachella* CYP9DMs have a second three-residue deletion between M4 and M5 (position 457- 459 of the MSA, highlighted in orange). *M. centuncularis*

CYP9DM, *M. lapponica* CYP9DM and *M. willughbiella* CYP9DM3 have a two-residue deletion at the same position.

The conserved structural core of P450 enzymes, houses the active site of the enzymes and is made up of a four-helix bundle, which includes helix I and helix L (see section 1.6.1.1.2 and figure 1.14). This portion of the enzyme is involved in both binding the substrate and the reaction mechanism [249]. It is clear that this portion of the CYP9DM sequences for all five *Megachile* species is divergent from other CYP9s. Indeed, a manual check of the alignment used to create the phylogeny of the Hymenoptera, shown in figure 6.11, confirmed that only the three CYP9DM sequences lacked the charged acid residue (E/D) in M2.

6.4 Discussion

6.4.1 The phylogeny of the CYP9 subfamily across bee families

The substitution of a *CYP9Q/BU/DL* ortholog for a *CYP9DM* gene correlates to a high degree of sensitivity to insecticides in *M. rotundata*. Therefore, the crucial question of how wide-spread a lack of *CYP9Q/BU*-like orthologs might be across bee species needs to be addressed. The phylogenies estimated from the CYP9 subfamily and the CYP9 cluster proteins of 75 species, across six of the seven families of bee, clearly illustrate that there is not a lack of *CYP9Q/BU/DL* orthologs in general. There is an obvious exception of some Megachilidae species, in particular the two species of *Megachile* (see section 6.3.1). However, the Melittidae are considered to be a sister group to all other bee families (see figure 3.1) and so the presence of a *CYP9Q/BU/DL* ortholog in all three species (*D. hirtipes*, *M. fulvipes* and *M. haemorrhoidalis*) is indicative that the presence of this lineage of genes predates the divergence of bees from one another. It therefore appears, that the absence of a *CYP9Q/BU/DL* ortholog amounts to a loss of this gene lineage.

When the CYP9 cluster in *M. rotundata* is compared to that of other bees there is good syntenic conservation, with *CYP9DM1* and *CYP9DM2* genes holding the same genomic position and reading frame direction as the *CYP9Q/BU/DL* genes from in other bee species (see section 3.3.3.4). Nevertheless, these two lineages have diverged enough at the DNA sequence and protein levels, to

have become evolutionary distinct from each other. There is a skew in the available data towards the Apidae, and therefore a comparative lack of information for the other families. As more RNA-seq data for Megachilidae species becomes available, particularly those belonging to the subfamily Megachilinae and Megachilini tribe it might be possible to determine whether there is a more widespread pattern of *CYP9Q/BU/DL* loss at these phylogenetic levels. The most recent phylogenetic analyses of the Megachilini tribe place the cleptoparasitic genera *Coelioxys* and *Radoszkowskiana* as a sister group to *Megachile* [636]. Interestingly, the TSA data available for *Coelioxys conoidea* indicated that this species may not have either *CYP9Q/BU/DL* or *CYP9DM* like sequences. *C. conoidea* looks to have expanded its *CYP9DN1* lineage, having three sequences rather than one, the remaining three sequences discovered for the species grouped with the *CYP9P* and *CYP9R* lineages.

On the whole, while it appears that the lack of a *CYP9Q/BU/DL* ortholog is not commonly found among bees, certain members of the Megachilidae family have lost the lineage. In particular it appears that some, or all, of the *Megachile* genus, which accounts for ~1500 species, may have *CYP9DM* rather than *CYP9Q/BU/DL* genes, with the corresponding potential for low tolerance to insecticides.

6.4.2 Fluorescence-based model substrate assays

There are orthologs of the insecticide-degrading *O. bicornis* *CYP9BU* enzymes in half of the Megachilidae species included in the phylogeny (see section 6.3.1). Sequences from four species were recombinantly expressed using Gateway® cloning technology and insect cell lines. The resulting enzymes were used in fluorescence-based insecticide mediated inhibition assays using *N*-cyanoamidine and *N*-nitroguanidine neonicotinoids and coumarin-based model substrates (see section 6.3.4.2). Unlike LC-MS/MS these fluorescence-based assays are not a definitive measure of the metabolic ability of a P450, or an assessment of metabolite production. Nonetheless, if there is inhibition of fluorescent product (HC) formation, it does infer that the insecticide is able to bind to the catalytic site of the recombinant protein [549, 550].

The results for the recombinant Megachilidae CYP9s substrate profiling indicate that there is no recognisable model substrate signature that defines either CYP9BU1-like or CYP9BU2-like sequences (see figure 6.8). Likewise, the inhibition studies also reveal a lack of a clear pattern of activity that correlates to the position the of the sequence in the phylogeny in figure 6.7 (a) (i.e. whether the enzyme claded with *O. bicornis* CYP9BU1 or CYP9BU2) (see figure 6.9 (a)). For example, *O. bicornis* CYP9BU1, *H. trunctorum* CYP9-like1 and *C. florisomne* CYP9-like show the most significant reduction in HC production when co-incubated with thiacloprid. However, although *H. trunctorum* CYP9-like1 clades as 'CYP9BU1-like' in the phylogeny, *C. florisomne* CYP9-like clades as 'CYP9BU2-like' (see figure 6.7 (a)). *O. bicornis* CYP9BU2 showed no significant reduction in HC production after co-incubation with either neonicotinoid, whereas the other two 'CYP9BU2-like' sequences both showed inhibition of product with exposure to imidacloprid (see table 6.6 and figure 6.9 (a)). The production of HC was significantly inhibited by both neonicotinoids after co-incubation with three recombinant Megachilidae CYP9s (*O. bicornis* CYP9BU1, *O. lignaria* CYP9-like and *C. florisomne* CYP9-like) (see table 6.6).

H. trunctorum CYP9-like1 shows heteroactivation when co-incubated with EC and imidacloprid. This type of cooperative effect is found with other P450 enzymes, such as *H. sapiens* CYP3A4, and also in *H. sapiens* liver microsomes [581, 637]. It is often explained in terms of the existence of multiple binding sites, whereby the two substrates involved in the heteroactivation are bound to separate sites on the enzyme [637]. Although heteroactivation of EC metabolism by imidacloprid by *H. trunctorum* CYP9-like implies that the insecticide is bound to the enzyme, it is possible that this binding occurs at an effector site rather than the catalytic site. This would imply that there is no metabolism of imidacloprid, but rather that its presence acts only to increase the metabolism of the EC.

Overall, it is clear from the model substrate profiling that all the recombinant Megachilidae CYP9s show metabolic activity. There is also a strong inference that most of the enzymes are capable of binding at least one neonicotinoid insecticide. It also appears that four of them may bind both *N*-cyanoamidine and

N-nitroguanidine neonicotinoids, although in the case of *H. truncorum* CYP9-like1 it is not clear that binding imidacloprid would result in its metabolism.

The next step would be to repeat the fluorescence-based insecticide mediated inhibition assays using the pyrethroid *tau*-fluvalinate and the organophosphate coumaphos. Both these insecticides are known to be metabolised by the CYP9Q and CYP9BU lineages. To further strengthen the use of fluorescence-based assays as a predictive tool, LC-MS/MS analyses of insecticide incubations with the neonicotinoid insecticides for the recombinant Megachilidae CYP9s from this PhD should be carried out. This would give definitive data on both metabolism of the parent compound and the production of hydroxylated metabolites. However, this is an initial step towards designing the pipeline for a predictive tool-kit that uses comparative genomics, in combination with P450 synthesis and targeted *in vitro* functional analyses, to assess the risk posed by insecticides to bees.

6.4.3 Phylogenetic and syntenic analyses of the CYP9 subfamily across the Hymenoptera

From the phylogenetic and syntenic analyses of the Hymenopteran CYP9 sequences it is clear that this subfamily of P450 enzymes is ancient. Most Hymenopteran species have at least one *CYP9P* sequence that is always associated with the *membralin* gene. In the Chalcidoidea: Apocrita the CYP9 cluster has appeared, where the P450 genes are flanked by *membralin* on one side and *alpha catulin* and *myosin IIIb* on the other (see figure 6.13). This conserved synteny is also found in the Vespoidea and the Formicoidea (see figure 6.13). There is no evidence that this genomic structure is present in the other species-rich Apocrita superfamily, the Ichneumonioidea, although the CYP9 sequences for these wasps fall into two main clades, one clade appears with the *CYP9Ps* and the other with the *CYP9Q/BU/DLs*. The conserved synteny of the CYP9 cluster found in the Chalcidoidea is also not evident in the Cynipoidea or Chrysidoidea, however, both of these wasp superfamilies appear to have CYP9s that clade as *CYP9Ps* and *CYP9Q/BU/DLs* (see figure 6.11 (a)). It does seem that in several of the Apocrita superfamilies the region containing the CYP9 cluster has been conserved, over a long timescale. This syntenic conservation may well be due to selection pressure and could therefore be

functionally relevant. It also appears that there was an ancestral genome rearrangement which brought the genomic region containing the *membralin* gene together with the region containing *alpha catulin* and *myosin IIIb*. Both these regions had CYP9 sequences associated with them. This rearrangement must either have occurred at some point before the divergence of the last common ancestor of the Chalcidoidea and the Vespoidea, (~ 240 mya), or it occurred twice. It seems less likely to assume an identical genomic rearrangement occurred twice during Hymenopteran evolution. If the rearrangement occurred at the point where the Parasitica arose as a group, it does not appear to have been conserved in the Ichneumoniodea or Cynipoidea. However overall, it does appear that the CYP9 cluster of P450s is an ancient, well-conserved and remarkably syntenic gene subfamily that most likely arose in the Jurassic period due to a genome rearrangement.

6.4.4 Transcriptomics results from four species of *Megachile* bees

The quality control results on the RNA extracted from each specimen of *Megachile* bee (*M. centuncularis*, *M. lapponica*, *M. leachella* and *M. willughbiella*), both the in-house Qubit® RNA assays (>99% large RNA) and the Agilent 2100 and q-PCR run by Novogene, indicated that the RNA was of good quality and size (see appendix table 6.3). The quality of the RNA allowed good sequencing with a low sequence error rate (see appendix figure 6.7). The RNA-seq data, therefore was able to reliably resolve the inventory of the CYP9 subfamily in the four species. The data indicate that the CYP9 inventory in all four *Megachile* species is broadly similar to that found in *M. rotundata* (see figure 6.14).

Species of *Megachile* are found on all continents, with the exception of Antarctica [638, 639]. Of the five *Megachile* species used in this PhD, four are native to Europe (three to the UK, one to Eastern Eurasia) and one to North America (Canada). Although some species of *Megachile*, that nest in wood or stems, are thought to have dispersed by the transportation of their nests on floating islands of vegetation, and in certain cases by passage of nests in timber on cargo ships, the global distribution of the genus implies that most species have evolved *in situ* [72, 638]. The *Megachile* genus is subdivided into 53 subgenera [72]. *M. rotundata* and *M. leachella* fall into the *Eutricharaea*

Thomson subgenus, *M. centuncularis* and *M. lapponica* into the *Megachile* Latreille subgenus and *M. willughbiella* into the *Xanthosarus* Robertson subgenus [72]. The *Eutricharaea* Thomson and *Xanthosarus* Robertson subgenera are more closely related to each other than they are to the *Megachile* Latreille subgenus [640]. It is clear, however, that the evolutionary separation of the *CYP9DM* genes from the *CYP9Q/BU/DL* lineage pre-dates the divergence of these subgenera. It seems likely that this phenomena could be genus wide, or, at the very least, be present in the majority of *Megachile* subgenera. This raises obvious concerns for the safe use of insecticides, not only for the pollination services provided by *M. rotundata*, but also for the health of wild *Megachile* species. As well as being one of the largest genera of bees, *Megachile* is also among the most widely dispersed [638, 639]. The role that each individual species of *Megachile* plays in the maintenance of their ecosystem is mostly undocumented. The impact of the genus would generally be reported as part of the amalgam termed 'wild-bees'. Calculations of the pollination service value of wild-bees to the US economy put the figure at over \$3 billion per annum [108, 641]. The proportion of this that can be attributed to the genus *Megachile* is unknown. However undoubtedly, wild *Megachile* bees perform an important ecosystem service by pollinating both native species of plant, cultivated plants and agricultural crops [108, 642]. There is an obvious need to ensure adequate safeguarding of this genus of wild pollinators.

In conclusion, the loss of a *CYP9Q/BU/DL* ortholog is not wide-spread across the bee families, but appears to be a phenomena that is confined to certain members the Megachilidae family. Furthermore, when present, the ortholog may not always have a broad substrate specificity. It seems likely that the majority, if not the entirety, of the *Megachile* genus have evolved *CYP9DM* genes and do not have sequences that clade with the *CYP9Q/BU/DL* lineage. It may be that other Megachilidae species such as *C. conoidea* may also have lost the *CYP9Q/BU/DL* lineage. This loss becomes even more curious when the evolutionary ancestry of the CYP9 cluster through the Order Hymenoptera is considered. From the phylogenetic relationship and syntenic order observed in CYP9 subfamily of the Chalcidoidea, Vespoidea and Formicoidea, it is clear that there has been substantial selection pressure to conserve this group of genes.

Chapter seven: General discussion

As discussed in the introduction, the recent reports of dramatic declines in insect numbers and diversity, and the potential impact of this on pollination services, make understanding non-target effects of pesticides on beneficial organisms, such as bees, critical [16, 17, 36, 643, 644]. To accurately inform the development of new, pest-specific compounds, that have low toxicity to non-target insects, a working knowledge of the metabolic pathways and detoxification mechanisms of beneficial insects is required. It has already been established that P450s from the *CYP9Q/BU* lineage are key determinants of insecticide sensitivity in *A. mellifera*, *B. terrestris* and *O. bicornis*. This PhD examined the possibility of whether these enzymes, capable of metabolising insecticides, are ubiquitous across bees, by initially studying the solitary leafcutter bee, *M. rotundata*. *In silico* analyses were then extended more widely across bee species, and to the Hymenoptera, to determine whether members of the CYP9 subfamily are highly conserved; something that could indicate strong selection pressure and have functional relevance.

In toto this work informs the development of a bee ‘toolkit’ or framework, such as that shown in figure 7.3. The application of a toxicogenomics approach to predict the level of sensitivity of a species of bee to insecticidal compounds, could facilitate future screening of lead compounds and underpin the current conventional toxicity trials. However, the data generated here also allow for the exploration of the evolution of CYP9 subfamily of genes in bees, and, more widely, their ancestry in the order Hymenoptera. This approach, whilst based in evolutionary biology, could also be applied to informing the assessment of insecticidal products. Tracing the evolution of key lineages of detoxification enzymes through a diverse insect order, such as the Hymenoptera, could elucidate where potential tolerances or sensitivities are likely to arise, thereby informing where future toxicity trials might be needed. For example, the many non-bee pollinators, or key insect groups such as the Chacidoidea fig wasps.

7.1 Understanding the molecular determinants of insecticide sensitivity in *M. rotundata*

Does M. rotundata have a CYP9Q/BU ortholog?

Are CYP9Q/BU orthologs ubiquitous in bees?

What are the implications in terms of insecticide sensitivity?

Can we predict insecticide sensitivity from phylogeny?

What are the wider implications?

The initial curation of the *M. rotundata* CYPome in chapter three showed that in common with other bee species, such as *A. mellifera* and *B. terrestris*, the species has a comparable number of P450s, split with similar weight across the four CYP clans found in insects (see section 3.3.1). However, phylogenetic analyses with 12 bee species, across three families (Apidae, Megachilidae and Halictidae) highlighted that, despite the overall similarity in CYPome content, there were key differences in the CYP9 subfamily of P450s. In particular, *M. rotundata* is the only species not to have a sequence that was part of the CYP9Q/BU/DL lineage (see section 3.3.6).

Syntenic analyses show that *M. rotundata* CYP9DM genes occupy a similar position and reading frame direction to the CYP9Q/BU/DL sequences within the genomic landscape (see section 3.3.3.4), but the phylogenetic analyses suggest they are evolutionarily distinct (branch length of 1.1 substitutions per 1.0 residues, see figure 3.10 a). Long branch lengths can imply structural differences and this is confirmed by modelling the tertiary structure of the protein. In both CYP9DMs there are key mutations in the active site formed by helices I (oxygen-binding) and L (heme-binding). In particular, there is a CYP9DM-specific loss of the charged acid residue (E/D) of the oxygen-binding motif (M2) not found in any other bee CYP9s.

Although these analyses answer the initial question of whether CYP9Q-like enzymes, that are able to metabolise certain synthetic insecticides, are ubiquitous across all bee species (they are not); they also raise further questions. Firstly, considering *in silico* work in isolation, is it possible to accurately predict that these structurally divergent CYP9 enzymes would also be functionally deviant? Secondly, do these data inform whether *M. rotundata* is

likely to tolerate the insecticides known to be metabolised by CYP9Q/BU enzymes *in vivo*? The answer is that we cannot accurately answer either these questions without further investigation. It would be premature to answer these questions on *in silico* work alone, as there is a lack information concerning the metabolic capabilities of the species. To provide these necessary data, three approaches were taken.

Firstly, in chapter four, acute contact toxicity bioassays using *M. rotundata* were performed, focused around the insecticides known to be metabolised by CYP9Q and CYP9BU enzymes.

Secondly, in chapter five, *M. rotundata* native microsomes were used in LC-MS/MS analyses against the insecticides known to be metabolised by CYP9Q and CYP9BU enzymes. In the case of the neonicotinoid insecticides, head membrane preparations were also used to assess binding affinity at the receptor site.

Lastly, in chapter five, key *M. rotundata* CYP9 enzymes were expressed, using heterologous expression, and their functional characterisation assessed using fluorometric assays.

The acute contact bioassays revealed that the three 'bee-friendly' insecticides are all highly toxic to *M. rotundata* ($LD_{50} < 2 \mu\text{g a.i./bee}$) [493]. The difference in toxicity level between *M. rotundata* and other managed pollinator species is particularly significant in the case of the neonicotinoid thiacloprid. *M. rotundata* is >2500-fold more sensitive than *A. mellifera* and >6,500-fold more sensitive than *B. terrestris* or *O. bicornis* to thiacloprid (see section 4.3.3) [168, 226]. The fold-difference in toxicity between *M. rotundata* and the other managed bee pollinators is less marked with *tau*-fluvalinate and coumaphos (ranging from 35-fold to >300-fold), but is still clearly greater than would be expected from difference in body size alone (~4 -17-fold).

Native microsomes from male *M. rotundata* were used in LC-MS/MS analyses due to the presence of a powerful P450 inhibitor in homogenates extracted from females (see section 5.3.1). No significant metabolism of the insecticides

thiacloprid, acetamiprid, flupyradifurone or *tau*-fluvalinate was observed (see section 5.3.3). The binding affinity of thiacloprid, imidacloprid and flupyradifurone was investigated (see section 5.3.5), and no difference was observed in the binding of these compounds at the receptor site between *M. rotundata* and *A. mellifera*. The high toxicity rating of these insecticides in *M. rotundata* is therefore not attributable to increased binding affinity.

The same protocol for LC-MS/MS analyses was used with *M. rotundata* native microsomes against four naturally occurring plant allelochemicals (see section 5.3.3.4). These data suggest that whilst *M. rotundata* has an appropriate repertoire of detoxification enzymes to metabolise xenobiotics naturally encountered in its diet and ecology, it cannot recruit these proteins to protect itself against synthetic insecticides.

The metabolic profile of *M. rotundata* CYP9s was examined using heterologous expression (see section 5.3.4). For six out of the nine *M. rotundata* CYP9s a functional enzyme was expressed. All six metabolised coumarin-derived fluorescent model substrates, including both CYP9DM proteins. Fluorescence-based assay, using the neonicotinoid insecticides imidacloprid and thiacloprid as inhibitory compounds, indicated that unlike *O. bicornis* CYP9BU1, *M. rotundata* CYP9DM2 does not bind or metabolise either compound.

The overall conclusion is that there is a correlation between the lack of a CYP9Q/BU ortholog and an inability to metabolise certain insecticides in *M. rotundata*. This answers the question of what implications the loss of a CYP9Q/BU ortholog might have: This loss is correlated with a high level of sensitivity to certain insecticides.

These data strongly indicate that the use of *A. mellifera* as a proxy for *M. rotundata* in ecotoxicology risk assessments is unreliable. One route to address this is to add more species into the initial tier of testing, and both *B. terrestris* and *O. bicornis* have been suggested as good candidates for this [472, 645-647]. Indeed, the OECD has published standardised protocols for testing for *B. terrestris* [484, 485]. However, in order to reduce the uncertainty level around the current tier one tests, any new species added would need to present a

different toxicology profile to *A. mellifera*. More specifically, it would need to be a more sensitive species [645]. Neither *B. terrestris* or *O. bicornis* consistently exhibit higher levels of sensitivity to insecticides across different MoA classes, and as such, they may not be suitable for use as additional proxy species [137, 168, 226, 463, 645]. However, it seems that *M. rotundata* might meet the criteria required from a new surrogate species for tier one testing of insecticides [447].

These data also raise the real possibility of using phylogeny, in combination with targeted *in vitro* functional analyses, to predict function. As it stands, the current risk assessment for pesticide registration, relies on toxicity data for *A. mellifera* as a proxy for the ~20,000 other species of bee. There is, however, a growing body of publications that highlight the impacts of specific life-history traits, such as nesting behaviour or social structure, on pesticide exposure and sensitivity [395, 646-649]. As such, the pesticide risk assessment process is under review by regulatory agencies in Europe and the North America [472, 650]. Due to the number of bees species and their diverse ecology, there are obvious knowledge gaps which could hinder the development of a more inclusive risk assessment [646-649]. The data in this thesis adds to the body of work that advocates the application of molecular methods currently used in medicine, such as RNA-seq, gene expression assays and fluorescence-based inhibitory assays, to the toxicological risk assessment of beneficial insects [550, 561, 643, 651]. Exploiting the genomic information of a species, with targeted function analyses, in conjunction with knowledge of life history traits would allow a toxicogenomic approach to be applied to the current risk assessment strategies [549, 550, 643].

7.2 Phylogenetic and syntenic analyses of the CYP9 subfamily of P450s

How wide-spread across bee families is the loss of a CYP9Q/BU/DL ortholog?

Do all Megachile lack a CYP9Q/BU ortholog?

Can we trace the ancestry of the CYP9Q/BU ortholog?

The fact that *M. rotundata* does not have CYP9Q/BU-like enzymes, something that is correlated with a highly sensitivity to insecticides, raises the key question: How wide-spread is this loss across bee species? The managed bee pollinators

included in toxicity comparisons in chapter four, are split across two families of bees (the Apidae and Megachilidae), and so the disparity in sensitivity observed in the species cannot be explained in terms of a family-wide trait. All species of Megachilidae are solitary, and so this difference in sensitivity is also not best described in terms of social structure and behaviour. There are ~20,000 described species of bee, across seven families, with diverse ecology, physiology and life history [72]. There is, therefore, clearly potential for a range of different responses to insecticide exposure [477, 645].

In chapter six, phylogenetic analyses of CYP9 sequences from 75 species, across six of the seven bee families were examined. Species from all six families were found to have CYP9 sequences that could be categorised as belonging to the *CYP9DN*, *CYP9P*, *CYP9R* or *CYP9Q/BU/DL* lineages, see table 6.2. The vast majority of the 75 species of bee have at least one full, or partial sequence that clades with the *CYP9Q/BU/DL* lineage (67/75). The Melittidae are considered to be a basal, sister clade to the other bee families (see figure 3.1), and so the presence of a *CYP9Q/BU/DL* ortholog in all three Melittidae species (representative of two of the three subfamilies, see table 6.1) indicates that the presence of this lineage pre-dates the divergence of bees from one another. However, only half of the Megachilidae species (6/12) have a definite *CYP9Q/BU/DL* ortholog. *Anthidium manicatum*, *Coelioxys conoidea*, *Lithurgus chrysurus*, *Stelis punctulatissima* and the two *Megachile* bees (*M. rotundata* and *M. willughbiella*) did not have a definite *CYP9Q/BU/DL* sequence. Although five of these species belong to the subfamily Megachilinae, (*L. chrysurus* alone belongs to the Lithurginae), they can be differentiated from each other at the level of tribe [72] (see figure 7.1). The six species that lack a *CYP9Q/BU/DL* sequence fall into three tribes: the Lithurgini, Megachilini and Anthidini (see figure 7.1). There is a skew in the Megachilidae sequences towards Osmiini tribe (5/12), meaning that information from these genera account for almost half of the Megachilidae data, but it should be noted that all the species from this tribe have at least one *CYP9Q/BU/DL* sequence.

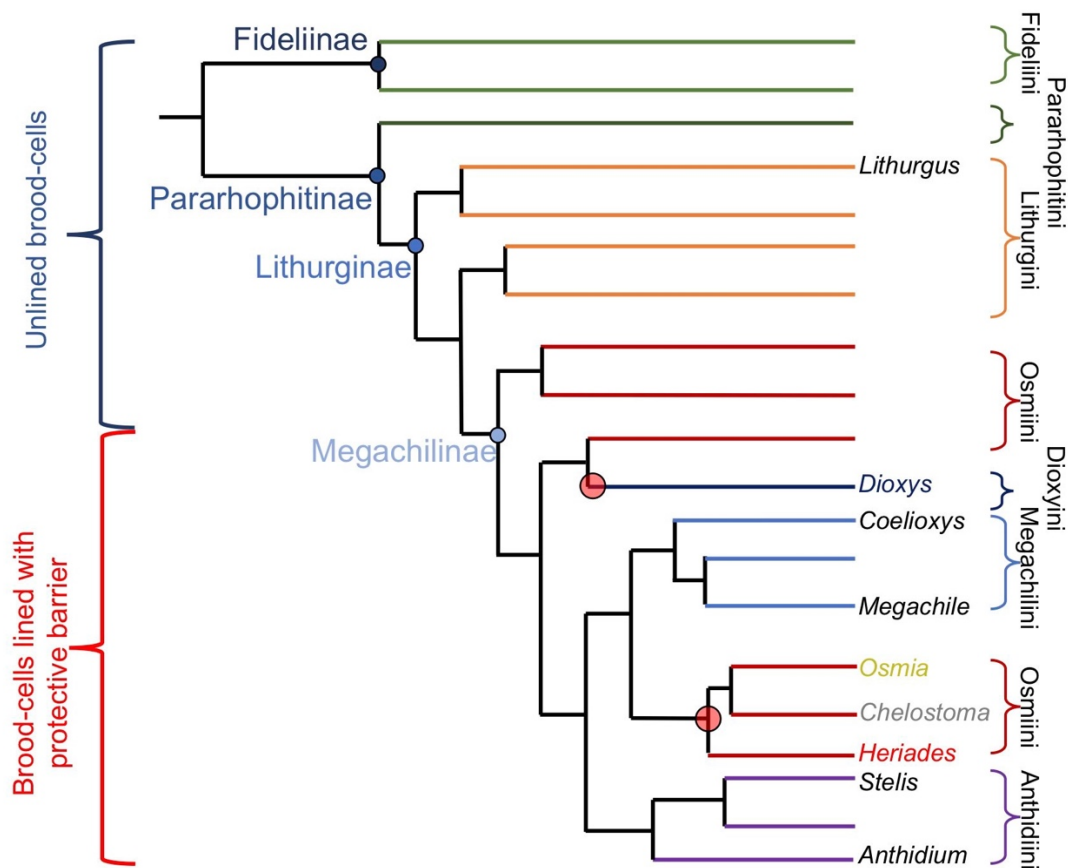


Figure 7.1: Representation of the phylogenetic relationship in the Megachilidae family, showing the position of the genera represented in the phylogenetic analyses in chapter six [640]. Branches are coloured by currently recognised tribes [72]. Ancestral nodes leading to species with CYP9 enzymes capable of binding neonicotinoid insecticides are marked with a red circle.

Megachile is one of the largest genera of bees, accounting for ~1500 species, and leafcutter bees are found on every continent, except Antarctica [64, 638, 639]. Subsequently, there is an obvious need to establish whether the loss of a *CYP9Q/BU/DL* ortholog is genus-wide. RNA was extracted from three UK native and one Canadian *Megachile* species. These four species represent three of the 53 *Megachile* subgenera: *Eutricharaea*, *Megachile* and *Xanthosarus*. The RNA-seq results were unambiguous: None of the *Megachile* species had a *CYP9Q/BU/DL* ortholog; all four had at least one *CYP9DM* gene.

Overall, these data indicate that the loss of the *CYP9Q/BU/DL* lineage is not wide-spread across bees. In five out of the six families of bee included in the phylogenetic analyses, there was an almost ubiquitous presence of at least one

CYP9Q/BU/DL ortholog (see section 6.3.1.5). However, there has been a loss within the Megachilidae. Most notably, it appears likely that the entire *Megachile* genus, and perhaps the Megachilini tribe, have lost the *CYP9Q/BU/DL* lineage. In the case of the *Megachile* species, the *CYP9Q/BU/DL* lineage has been replaced by the structurally divergent *CYP9DM* sequences.

When the phylogenetic analyses are extended to include Hymenopteran species (see figure 6.1 and section 6.3.5), it becomes apparent that CYP9 P450 enzymes belong to an ancient subfamily that is present in ants and wasps, as well as bees, and whose origins can be traced back to sawflies (see figure 6.11). All the Hymenoptera species included in the analyses have at least one *CYP9P* gene (see section 6.3.5.2). From the phylogenetic analyses it is evident that the *CYP9P* lineage is the ancestral form of CYP9 P450, and so the cluster of genes was most likely formed by the tandem duplication of an ancestral *CYP9P* gene.

Even more extraordinary, the results of the syntenic analyses indicate that the conserved syntenic CYP9 block, initially observed in bees (see section 3.3.3.4), is also widely found in wasps (Vespoidea and Chalcidoidea) and ants (see section 6.3.6). This common genomic landscape sees the CYP9 cluster flanked by *membralin* on one side and *alpha catulin* and *myosin IIIb* on the other. Genes from the *CYP9P* lineage are found in association with *membralin* in bees, ants, Vespoidea, Chalcidoidea and all three sawflies (*Athalia rosae*, *Cephus cinctus* and *Neodiprion lecontei*). The three flanking genes: *membralin*, *alpha catulin* and *myosin IIIb* are found on two separate scaffolds in the sawfly species. In addition to the *CYP9P* gene which is found in association with *membralin*, there is a another CYP9 sequence found in association with *alpha catulin* and *myosin IIIb* on the second scaffold in sawflies (see figure 6.14). It appears that at some point, before the divergence of the Chalcidoidea, there was a genome rearrangement which brought the region containing *membralin* together with the region containing *alpha catulin* and *myosin IIIb*. It may be therefore, that the *CYP9P* and *CYP9Q/BU/DL* lineages evolved from different ancestral genes: the first derived from the scaffold containing *membralin* and the second from that containing *alpha catulin* and *myosin IIIb*.

There is no evidence of a *CYP9P-membralin* association in the Ichneumonoidea or Cynipoidea superfamilies of wasp, and it appears that the syntenic CYP9 block has not been conserved in these groups. The Ichneumonoidea are mostly parasitoids, as are many of the Cynipoidea, however the latter family also contains the phytophagous gall-wasps [652, 653]. The adults of many species of parasitic wasps feed on nectar and so, in common with bees, would be expected to encounter plant allelochemicals in their diet. Despite the lack of conserved synteny in certain wasp superfamilies, the origin of the *CYP9P-membralin* association appears to predate the divergence of the Hymenoptera from the other Endopterygota, as it is found in all three sawfly species. This association has therefore been conserved across diverse groups for well over 200 million years.

All the sawfly CYP9 sequences are intronless and, with the exception of two Chalcidoidea species *Nasonia vitripennis* (2/9 sequences) and *Copidosoma floridanum* (1/3 sequences), this gene structure is conserved in the CYP9 genes throughout the Hymenoptera. In the two Chalcidoidea species, the intron-containing CYP9s were associated with *alpha catulin* and not *membralin*, however, all three clade with the *CYP9P* lineage. One of the mechanisms thought to be responsible for generating intronless genes is retrotransposition [654]. It is not uncommon for intronless genes to be duplicated, giving rise to families of clustered intronless genes, such as the *H. sapiens* type I interferon gene cluster [655]. Intronless genes have the obvious advantages of faster protein production and lower energy cost and so this type gene structure is highly expedient in an environmental response gene, such as a P450 [217, 656]. It is clear from the intronless structure of the sawfly *CYP9P* gene, and its association with '*membralin*', that the ancestral retrotransposition event that led to the formation of the CYP9 cluster must have occurred before the divergence of the Hymenoptera, with expansion of the subfamily occurring later. It appears there must have been selection pressure to conserve the CYP9 syntenic block in diverse superfamilies across the Hymenoptera. As such, there is good evidence that the CYP9 subfamily represents a group of highly conserved, Hymenoptera-specific, environmental response genes.

7.3 Metabolic profiling of recombinant Megachilidae CYP9BU-like enzymes

Do the Megachilidae CYP9BU-like enzymes have activity against synthetic insecticides?

Can we reliably predict function from phylogeny?

The phylogeny of 75 species of bee (see section 6.3.2) was used to select candidate *CYP9BU*-orthologs from the Megachilidae species. Five sequences were chosen, two that were similar to *O. bicornis CYP9BU1*, and three that were similar to *O. bicornis CYP9BU2*. Four of the recombinant expressed enzymes were metabolically active (see section 6.3.4): two *CYP9BU1*-like (*Osmia lignaria* and *Heriades truncorum*) and two *CYP9BU2*-like (*Dioxys cincta* and *Chelostoma florissomne*). In addition to these proteins, the availability of *O. bicornis CYP9BU1* and *CYP9BU2* viral stocks allowed for a comparison to these enzymes [226, 463]. *M. rotundata CYP9DM2* protein was also screened using the same protocols (see section 5.3.4.3.3). In total therefore seven Megachilidae recombinant CYP9 enzymes from six species were investigated for metabolic profile using fluorescence-based model substrate screening and insecticide mediated inhibition assays (see section 6.3.4). The four Megachilidae species are split across two tribes: the Osmiini (three species) and Dioxyini (one species).

Only *O. bicornis CYP9BU2* and *M. rotundata CYP9DM2* enzymes exhibited no significant inhibition of HC production by either insecticide, implying that neither protein binds or metabolises the tested neonicotinoids effectively. In the case of *O. bicornis*, this is supported by the higher K_M values obtained for *CYP9BU2* in comparison to *CYP9BU1* (TCP: $K_M = 47.5 \mu\text{M}$ versus $5.5\mu\text{M}$; IMI: $K_M = 171.5 \mu\text{M}$ versus $50.9\mu\text{M}$) [463]. For *CYP9DM2* the lack of inhibition of HC production supports the LC-MS/MS data generated using *M. rotundata* native microsomes (see section 5.3.3.1). There is a strong inference that all the other Megachilidae recombinant enzymes are capable of binding at least one neonicotinoid insecticide and that four of them may bind both *N*-cyanoamidine and *N*-nitroguanidine neonicotinoids.

There was no discernible pattern of activity in the Megachilidae recombinant CYP9 enzymes that could be related to their phylogenetic relationship to either CYP9BU1-like or CYP9BU2-like. The results of the insecticide mediated inhibition assays indicate that the majority of these enzymes exhibit atypical (non-Michaelis-Menten) kinetics, something that is not uncommon in P450s [528, 561, 581]. It is also in keeping with the results found for the substrate binding kinetics of *A. mellifera* CYP9Q2 [549]. There are examples of both non-competitive and competitive inhibition, but also one example of heteroactivation in the co-incubation of *H. truncorum* CYP9-like1 with EC and imidacloprid (see section 6.3.4.2). Heteroactivation is generally explained in terms of multiple binding sites [637, 657, 658]. Although it is probably more accurate to describe a highly potent P450 as having a sufficiently large active site, that can be accessed by substrates bound to different regions of the enzyme [561, 659]. The highly potent metaboliser *H. sapiens* CYP3A4 is known to be capable of binding three substrates simultaneously [561]. The hypothesis that, due to the ability to metabolise multiple diverse insecticides, enzymes from the CYP9Q/BU lineage could be considered as highly potent, promiscuous enzymes, akin to *H. sapiens* CYP3A4, is consistent with finding heterotropic enhancement effects within the ranks of their orthologs.

The fact that out of the 12 species of Megachilidae included in the phylogenetic analyses of 75 bee species, only six had *CYP9BU* orthologs, meant that there were very few sequences available to select from in choosing candidates for expression. Of the six available species, three were from the *Osmia* genus, one being *O. bicornis*. The four candidate species, whilst the most complete selection available, are not truly representative of the broader Megachilidae family (see figure 7.1). Only one Megachilidae species with CYP9 genes that could be described as 'non-*CYP9Q/BU*-like' (i.e. *M. rotundata* CYP9DM1 and CYP9DM2), was available to be included in the metabolic profiling and insecticide mediated inhibition assays.

The Megachilidae is a large, cosmopolitan family that contains over 4000 species of bee (see figure 3.1) [660]. As such, there is an obvious paucity of data for this family, something that makes drawing any conclusions problematic and potentially unreliable. That being said, with the exception of *O. bicornis*

CYP9BU2, it appears that all the CYP9BU-like enzymes are capable of binding at least one neonicotinoid insecticide. Conversely, *M. rotundata* CYP9DM enzymes appear not to bind either thiacloprid or imidacloprid, something that supports what was observed in LC-MS/MS analyses using native microsomes.

For the data from the fluorescence-based inhibition assays to be considered reliable, the results would need to be verified, using insecticide assays and LC-MS/MS analyses [550]. Once a good correlation between data generated from these two protocols has been established, the need to rely on the more costly LC-MS/MS analyses will be reduced. However, these data illustrate the possibility of predicting function from phylogeny, at least in as much as informing decisions on which species should be prioritised for further assessment, such as acute contact toxicity tests.

7.4 The Megachile CYP9DM enzymes

What are the key differences between a CYP9Q/BU/DL and a CYP9DM protein?

Do CYP9DMs have activity against synthetic insecticides?

What life-history or ecological trait depends on the presence of a CYP9DM enzyme?

The phylogenetic analyses throughout this PhD have been consistent in identifying *CYP9DM* sequences as being distinct from *CYP9Q/BU* sequences. Long branch lengths in phylogenetic trees were produced for *M. rotundata* *CYP9DM1* and *CYP9DM2* (see sections 3.3.2 and 3.3.6), and then again for the *CYP9DM* sequences from the other four *Megachile* species (see sections 3.3.3.4 and 6.3.7). It does therefore appear that the *Megachile* genus has evolved a separate lineage of CYP9 enzymes.

The nucleotide sequences of *M. rotundata* *CYP9DM1* and *CYP9DM2* are 85.8% identical to each other and the proteins are 91.2% identical (44 amino acid substitutions). Moreover, both enzymes are functional, in that they can metabolise coumarin-based fluorescent model substrates (see section 5.3.4). Indeed, the two CYP9DM enzymes appear to have slightly different substrate specificities. Both are able to metabolise the coumarin MC, but CYP9DM2 is

also active against the larger (trifluoromethyl)- coumarin MOBFC. This implies there has been neofunctionalization within this gene lineage, something that could lead to the evolution of discrete biological functions.

All the *CYP9DM* sequences from the five species share a modified active site/heme-binding pocket (see section 6.3.7.6). In figure 7.2 (a) and (b) the key differences between the conserved motifs of the *CYP9Q/BU* lineage and the *CYP9DM* sequences are represented. Motifs M1, M3 and M4 are well conserved across the *CYP9Q/BU* and *CYP9DM* lineages (see figure 7.2 (b)). The real differences between these *CYP9* lineages appear in the oxygen-binding motif, M2 (motif: G-x-E/D-T-T/S) (see section 6.3.7.6). This motif is a vital part of any P450 enzyme, as it is involved in both binding a substrate (SRS4) and its metabolism (M2) [218, 249]. The loss of both the conserved glycine (G) and the charged acid residue (aspartic/glutamic acid; E/D) is unusual. Indeed, it is not found in any other Hymenopteran *CYP9* sequences included in the phylogenetic analyses. In fact, the loss of the conserved G and E/D residues is not seen in any other bee P450s, apart from certain the members of the *CYP2* clan (*CYP15A1*, *CYP305D1*, *CYP307B1* and *CYP369A1*). The *CYP2* clan is thought to be the ancestral form of P450s in insects [218, 236]. It is clear therefore that, whilst unusual, the substitution of these highly conserved residues does not necessarily prevent the correct folding of the protein.

The substitution of the anionic amino acid (E/D) for a non-charged residue, either alanine (A; non-polar, hydrophobic) or threonine (T; polar) at the centre of the oxygen-binding motif of the *CYP9DM* enzymes will alter the binding of that portion of the active site. A change in the charge of a binding site can have a dramatic effect on its affinity for different substrates. For example, the substitution of a cationic subsite in the insect nicotinic acetylcholine receptor site (nAChR) for an anionic site in the mammalian nAChR allows for the selectivity of neonicotinoids to insects [118, 157]. Work on the potent metaboliser, *H. sapiens* *CYP3A4*, has shown the importance of key amino acid residues in the SRSs, including Glu308 and Thr309 from the SRS4/M2 motif

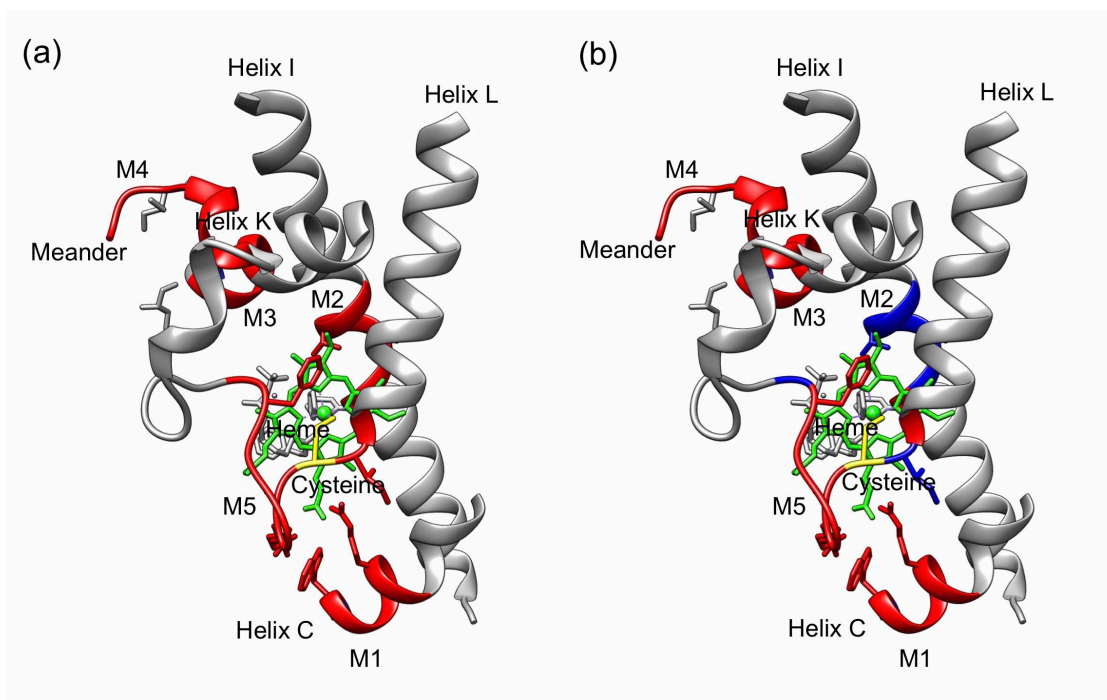


Figure 7.2: Ribbon diagram of conserved motifs and active site/heme-binding pocket of *H. sapiens* CYP93A4 (PDB: 4D6Z [254]). (a) Conserved motifs (M1-5) coloured red to show similarity, with respect to *A. mellifera*, *B. terrestris*, *O. bicornis* CYP9Q/BU enzymes, conserved cysteine (C) shown in yellow and the heme- group in green. (b) Conserved motifs (M1-5) coloured red and blue to indicate similarity or difference between CYP9Q/BU and CYP9DM enzymes from five species *Megachile*. Figure created using UCSF Chimera version 1.10.1.

[659]. As such, the loss of the anionic residue in the CYP9DM enzymes (equivalent to Glu308), might indicate that these enzymes bind a different group of substrates those of the CYP9Q/BU lineage.

The phylogenetic analyses of the 104 Hymenopteran species did not reveal any CYP9 sequences, apart from those of *Megachile* bees, that were categorised as being CYP9DM-like (see section 6.3.5). This is indicative that this lineage is specific to certain bee species, and perhaps even *Megachile*-specific. This raises the question of what selection pressure was involved in changing a gene lineage, that codes for potent metabolising enzymes capable of binding multiple substrates?

All solitary bees that build nests (i.e. not brood parasites) need to construct brood cells. These cells house the developing larvae and the provision mass, which contains pollen, nectar and floral oils. To prevent the provision mass from becoming spoiled by fungal or microbial infection, the vast majority of bees line their brood cells with a protective waterproof, microbe-resistant barrier [64, 641]. In most bee families (Andrenidae, Apidae, Colletidae and Halictidae) this is achieved using secretions from the Dufour's gland, which is located close to the venom-sac [64, 641]. These secretions create a hydrophobic, microbe-resistant layer that coats the inside of the brood cell. In contrast, the majority of species in the Megachilidae family use a diverse selection of foreign material to line their nests, although the Fideliinae and Pararhophitinae subfamilies construct unlined nests in the soil and the Lithurginae excavate unlined nests in wood or stems (see figure 7.1) [64, 641]. The remaining Megachilidae species use mud, resin, fibre, pebbles, petals, leaf masticate or leaf pieces and sometimes a mix of these materials, to construct waterproof and microbe-resistant linings for their brood cells [64, 641].

In the case of species of the *Megachile* genus (leafcutter bees), the brood cells are lined with discs of leaves or petals. The use of plant-derived material for this purpose has allowed the genus to diversify, colonise temperate regions and to become species-abundant [64, 641]. By default, however, collecting material by cutting leaves, actively damages the plant from which it is taken. It could be therefore, that there is some form of an induced defence response in the plant in reaction to the tissue damage caused. Plants have evolved complex morphological and biochemical defences to minimise damage from herbivorous insects [661, 662]. For example, the plant *Arabidopsis thaliana*, produces allelochemicals (glucosinolates and anthocyanins) in response to leaf vibrations caused by insect chewing [663]. Whilst leafcutter bees are not ingesting the leaf discs they cut, it is possible that they come into contact with plant allelochemicals. One possible dynamic for what has driven the pressure to evolve a unique clique of detoxification enzymes in *Megachile* bees, is perhaps that in becoming bees that cut leaves, they are faced with a different array of plant defence chemicals. This would be an application of Occam's razor: Leafcutter bees need distinct detoxification enzymes because they cut leaves. However, more genomic and transcriptomic data is needed before reliable

evidence of a definite ecological pattern or evolutionary signature can be elucidated for the *CYP9DM* lineage.

7.5 Implications

M. rotundata is the most intensively managed solitary bee species, used widely in agriculture for pollination services [56]. The most obvious implication of the work described in this thesis therefore, relates to the safety of insecticides used on the crops this species pollinates. Specifically, my findings strongly suggest that insecticides: thiacloprid, flupyradifurone, *tau*-fluvalinate and coumaphos should not be used where *M. rotundata* can come into direct contact with them. The product type, timing and method of application of insecticides all need to be considered for crops where this bee is the choice of managed pollinator.

As discussed in section 7.1, it is clear that the current use of *A. mellifera* as a proxy for *M. rotundata* in ecotoxicological testing is unreliable. Given the LD₅₀ values generated in this PhD (see section 4.3.3), the suggested use of an assessment or bridging factor of 10, to allow toxicological data from *A. mellifera* to be used in other species, as recommended by the European Chemical Agency in 2020, would also not reduce the risks to *M. rotundata* to a safe level [475-477]. The suggestion that either *B. terrestris* or *O. bicornis* (or both) should be added as extra model species to the toxicity trials would also not give *M. rotundata* the protection it needs [472, 645-647]. A better choice might be that *M. rotundata* is added as the model solitary bee species, as it fulfils the criteria required from a new surrogate species for tier one testing of insecticides and due to its use as a commercial pollinator, it is also readily available.

There are more than 1500 species of wild leafcutter bees found globally [64, 639]. The ecosystem services provided by this large, cosmopolitan genus will extend beyond simply pollination. There are more than 30 insect species that target *M. rotundata* alone [52]. They can be divided broadly into: parasitoids, such as Chalcidoidea wasps; cleptoparasites or brood parasites (*Coelioxys* and *Stelis* genera), and predators, such as beetles, predatory Vespiodea wasps and ants [664-666]. There are also many species of bird and rodent that predate *M. rotundata* adults and nests [54, 56, 665]. It is clear from the data on this single *Megachile* species, that for each of the 1500 bees there is likely to be numerous

dependent species. Without further testing and regulation, there could be a significant cost, in terms of ecosystem diversity, if all species of this genus of bee do indeed lack a *CYP9Q/BU* ortholog, and are therefore more vulnerable than other bees to insecticides.

Given these implications, there is a need to develop standardised comparative genomics analyses and targeted *in vitro* functional tests that can inform the current ecotoxicological trials. Acute contact toxicity testing could then be targeted at species or genera as required to help avoid negative outcomes of insecticide use.

7.6 Applications

There is a need to bring pest-specific insecticidal products to the market to meet the challenge of producing enough food to feed the increasing human population [7, 667]. Over the last few decades, insecticides have been withdrawn, and regulatory requirements have been altered, due to environmental concerns about toxicity and persistence of compounds [338, 339]. Concurrently, there has been an increase in pest-species insecticide resistance to existing products [123, 668, 669]. There is therefore a concerted drive to develop new, more pest-selective and less persistent compounds that can be used in the context of integrated pest management (IPM) practices [667, 670-672]. This process is both lengthy and expensive. In 2013, it was estimated that to take a new pesticide through development and registration would cost \$250 million and take eight to twelve years [667]. Given this, the ability to screen out products that would be toxic to non-target beneficial insects, at an early stage of its development, would save both time and money.

The results generated in this PhD can be used to inform the structure or pipeline of work that underpin the concept of a 'tool-kit' to help screen insecticides and to develop the next-generation of bee-safe products, such as that shown in figure 7.3. Techniques from each of the boxes shown (0) – (4) were used in this PhD, although not necessarily in the order shown. With the dramatic increase in genomic and transcriptomic data becoming available for bee species, the need to collect specimens from the wild will not always be a necessary step in the

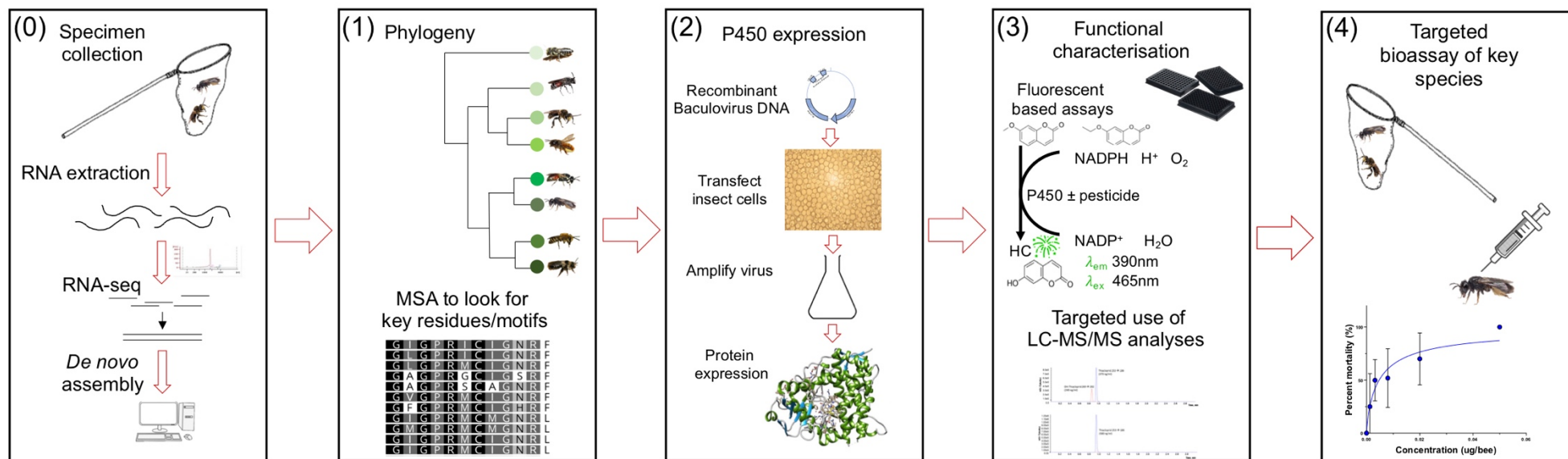


Figure 7.3: Suggested framework or ‘tool-kit’ to predict bee sensitivity to insecticides. Boxes (1) - (3) show the pipe-line of the comparative genomics, heterologous expression of P450 enzymes and metabolic profiling that make up the ‘tool-kit’.

pipe-line as the data may be already be available. Likewise, once the testing stages in boxes (1) to (3) are standardised and reliable enough, acute contact bioassays, box (4), will not always be required.

For the insecticides used in this PhD, P450s belonging to the *CYP9Q/BU/DL* lineage appear to have the potential to act as genetic markers for the expected toxicity level in bees. As such, this type of approach could form the basis of a framework that applies a toxicogenomics approach to assess the likelihood of sensitivity, or tolerance, to insecticides in a bee species [549, 643, 651].

7.7 Future work

- **LC-MS/MS verification of fluorescence-based insecticide mediated inhibition assays**

The results for the Megachilidae expressed CYP9BU-like enzymes (see section 6.3.4.2) should be thought of as indicative only. It would be useful to run further fluorescence-based assays using *tau*-fluvalinate and coumaphos as the inhibitor and then verify the metabolism of all the insecticides (i.e. the neonicotinoids) using LC-MS/MS. Similarly, insecticide incubations and LC-MS/MS analyses should ideally be performed for the expressed *M. rotundata* CYP9 enzymes (see section 5.3.2).

- **A standard panel of fluorescent model substrates needs to be determined**

Alongside the coumarin-based model substrates used in this PhD, resorufins and other fluorophores could be investigated for use in fluorescence-based insecticide mediated inhibition assays. The creation of a standard, available and affordable screening panel would be highly advantageous.

- **Metabolic profiling of functionally expressed candidate Hymenopteran CYP9s**

In a similar fashion to that used in section 6.3.2, key candidate CYP9 from ant and wasp species could be selected for heterologous expression, fluorescence-based assays and LC-MS/MS verification. This would allow the conservation of function of this lineage of potent metabolisers to be explored. Not only would this be a valuable exercise in exploring the evolution of a P450 lineage, it would

also be predictive in regards to insecticide sensitivity in non-bee Hymenoptera. This would allow the exploration of the leaves of the phylogenetic tree used to select candidate CYP9s. To investigate the *CYP9Q/BU/DL* lineage more fully, ancestral sequence reconstruction could be used to investigate the branches and nodes of the phylogeny, this would allow the form of the ancestral enzymes to be predicted. A selection of key nodes could be selected and the ancestral proteins they represented could be functionally expressed, effectively resurrecting them. Using these techniques there is the potential to understand the evolutionary history of *CYP9Q/BU/DL* lineage, and to perhaps trace the origins of their ability to metabolise insecticide.

- **Using site-direct mutagenesis can we determine which are the key amino acid residues in the loss of the ability to metabolise insecticides?**

To explore and understand what mutations and substitutions are involved in the loss or gain of the ability of the CYP9Q/BU enzyme to metabolise insecticides, site-directed mutagenesis, which can introduce targeted amino-acid replacements into a protein, could be used. The functional expression of different CYP9Q/BU mutants would allow their metabolism to be profiled. In this manner, it would be possible to establish *in vitro*, which mutations are the key molecular determinants that impart the ability to metabolise insecticides.

- **What impact does genetic variation within subspecies and populations have on insecticide sensitivity?**

The level of intrapopulation variation in CYP9Q/BU sequence is not clear. The availability of the genomes for several strains of *A. mellifera* will allow any intraspecies sequence mutation to be annotated. This could also be examined by sourcing *O. bicornis* or *M. rotundata* cocoons from geographically distinct suppliers to look for intrapopulation differences. Allelic differences and copy number variation could both potentially have large effects on insecticide sensitivity. Any variations found could be explored using heterologous expression and metabolic profiling.

- **Can we relate CYP-lineage loss to life-history and ecology of the species?**

As a group of organisms, the wide-ranging and diverse life-histories, strategies and ecological traits of bees make it difficult to describe the gain or loss of an enzyme lineage in anything other than qualitative ways. The challenge, as more genomic and transcriptomic information becomes available, will be to attempt to quantify these data. In the future, it might be possible to conceive a statistical model to describe these types of results quantitatively. For example, this might enable the significance of a link between dietary specialisation of a species and repertoire of detoxification enzymes to be established. Table 6.1 was a first attempt at collating the ecology of a species and its phylogeny (i.e. family/subfamily, presence of a specific P450 gene, diet etc).

- **Can we use DeepMind AI or similar to produce more accurate tertiary structure of insect P450s?**

To date, the crystal structure of an insect P450 has not been resolved. Whilst the structural conservation is remarkable across these enzymes, it would be a very useful tool to have a more accurate prediction of the tertiary folding of an insect P450.

Appendices

Appendix chapter two

Table 2.1: Potassium-phosphate buffer (0.1M Na/K-phosphate buffer, pH adjusted to 7.6 with H₃PO₄)

Ingredient	Amount g/l
KH ₂ PO ₄ (0.016M)	2.18
Na ₂ HPO ₄ (0.084M)	30.1
Distilled water	Final volume 1 l (1000 ml)

Table 2.2: Homogenisation buffer (pH adjusted to 7.4 with H₃PO₄)

Ingredient	Amount g/l
KH ₂ PO ₄ (0.016M)	2.18
Na ₂ HPO ₄ (0.084M)	30.1
Ethylendiamintetraacetate (EDTA) (1mM)	0.292
1,4-Dithio-DL-threitol (DTT) (1mM)	0.154
Saccharose (200mM)	68.4
Distilled water	Final volume 1 l (1000 ml)

Table 2.3: Buffer R (pH adjusted to 7.6 with H₃PO₄)

Ingredient	Amount g/l
KH ₂ PO ₄ (0.016M)	2.18
Na ₂ HPO ₄ (0.084M)	30.1
Ethylendiamintetraacetate (EDTA) (0.1mM)	0.029
1,4-Dithio-DL-threitol (DTT) (1mM)	0.154
Glycerol 10%	100 ml
Distilled water	Final volume 1 l (1000 ml)

Table 2.4: Stop solution for use with coumarin model substrate assays

Ingredient	Amount
100% DMSO	4.5 ml
0.05M Tris/HCL buffer, pH10	4.5 ml
Glutathione reductase (0.2U)	88 µl
5mM L-Glutathione oxidized	30.65 mg in 500 µl distilled water
Distilled water	412 µl

Table 2.5: LB agar and LB broth

Ingredient	Amount
LB Broth Powder (Lennox BP1427-500 Molecular Genetics Powder)	10 g
Distilled water	500 ml
Agar (for plates only)	7.5 g

Table 2.6: Primer sequences

Primer name	Sequence
2 step PCR (Invitrogen)	
Universal_12attB_F	GGGGACAAGTTTGTACAAAAAAGCAGGCT
Universal_12attB_R	GGGGACCACTTTGTACAAGAAAGCTGGGT
CYP9DM2_12attB1_F	AAAAGCAGGCTTCATGGAGGAAACGCAGATC
CYP9DM2_12attB2_R	AGAAAGCTGGGTCTCAAGAATTTGCACTAACAGC
CYP9DN1_12attB_F	AAAAGCAGGCTTCATGGATCCGTTTACTTTGACATTG
CYP9DN1_12attB_R	AGAAAGCTGGGTCTTATTTGCGCGAGACCGT
CYP9R1_12attB_F	AAAAGCAGGCTTCATGTGGACAATAATCGCGGGGGTG
CYP9R1_12attB_R	AGAAAGCTGGGTCTCATGCACGTTGTGCTGATGAAC
CYP9P22_12attB_F	AAAAGCAGGCTTCATGGTTTTCGAACTGAGC
CYP9P22_12attB_R	AGAAAGCTGGGTCCTAAGCTCTCTTCTCCAAAC
CYP9P23_12attB_F	AAAAGCAGGCTTCATGGAGATTCTCCTGTCTTGG
CYP9P23_12attB_R	AGAAAGCTGGGTCTTAAACATCCCTTTTCTCCAAG
CYP9R59_12attB_F	AAAAGCAGGCTTCATGGACAGTACAGCAGTTTG
CYP9R59_12attB_R	AGAAAGCTGGGTCTTACATCCCCTTCTAGATTG
CYP9P2_12attB_F	AAAAGCAGGCTTCATGGAGTCCGCGTCGTTTTTC
CYP9P2_12attB_R	AGAAAGCTGGGTCTCAAACAGAACTCCTACTCTCCAG
Primers used in sequencing	
CYP9DM2 F2	GATATCAAGATTATATTGACAAG
CYP9DM2 R2	GCATGTACTGCAACATG
CYP9DN1 F2	CTGTCCAATTTGCTTGGC
CYP9DN1 R2	TCGTCATCTCCTCGATCG
CYP9R1 F2	CTAAACTGTGTAACGCGCTG
CYP9R1 R2	GTCATATCCTCGATGGTGAGC
CYP9P22 F2	GCGGCGTGTGGCGTATGTTG
CYP9P22 R2	CCCTAGCCTGCATCAGAAGGTGG
CYP9P23 F2	CGGGCACCTGTACAACTG
CYP9P23 R2	CTGATGCAGGCTAAGAACAAG
CYP9R59 F2	GCCACGAATCCCACGACG
CYP9R59 R2	GCTGAATCATGTGCGGGTCGC
CYP9P2 F2	CGTCTGATCAAGTTCCTGC
CYP9P2 R2	GCTTCTCTTTATCCCTGGC

Primers used in PCR cloning	
M13_F	GTAAAACGACGGCCAG
M13_R	CAGGAACAGCTATGAC
pJET1.2_F	CGACTCACTATAGGGAGAGCGGC
pJET1.2_R	AAGAACATCGATTTTCCATGGCAG
Polyhedrin_F	AAATGATAACCATCTCGC
V5_R	ACCGAGGAGAGGGTTAGGGAT

Appendix chapter three

Appendix table 3.1: Maximum Likelihood fits of 56 different amino acid substitution models. Bayesian Information Criterion (BIC); Akaike Information Criterion, corrected (AICc) and Maximum Likelihood (InL) are shown. Where applicable, estimates of gamma (G) shape parameter and the estimated fraction of invariant (I) sites are presented.

Model	BIC	AICc	InL	Invariant	Gamma
LG+G	39593.7796	36547.74781	-17870.83983	n/a	1.921500972
LG+G+I	39603.5565	36549.84058	-17870.83983	0	1.921500972
JTT+G+F	39738.16392	36546.17251	-17850.15136	n/a	1.961109612
JTT+G+I+F	39747.94072	36548.27249	-17850.15136	0	1.961109612
JTT+G	39762.39205	36716.36026	-17955.14606	n/a	1.93463646
LG+G+F	39765.78133	36573.77261	-17863.95011	n/a	1.931976408
JTT+G+I	39772.16896	36718.45303	-17955.14606	0	1.93463646
LG+G+I+F	39775.53823	36575.87	-17863.95011	0	1.931976408
WAG+G	39921.44702	36875.41523	-18034.67354	n/a	2.095342052
WAG+G+I	39931.22392	36877.508	-18034.67354	0	2.095342052
WAG+G+F	39955.11164	36763.12292	-17958.62527	n/a	2.038786363
WAG+G+I+F	39964.88855	36765.22032	-17958.62527	0	2.038786363
rREV+G+F	40084.93793	36892.94921	-18023.53841	n/a	2.016903112
rREV+G+I+F	40094.71483	36895.0466	-18023.53841	0	2.016903112
rREV+G	40146.26238	37100.23059	-18147.08122	n/a	2.041780867
rREV+G+I	40156.03928	37102.32336	-18147.08122	0	2.041780867
Dayhoff+G+F	40236.99348	37045.00476	-18099.56619	n/a	1.784515617
Dayhoff+G+I+F	40246.77038	37047.10216	-18099.56619	0	1.784515617
miREV24+G+F	40285.4382	37093.44948	-18123.78855	n/a	1.668375034
miREV24+G+I+F	40295.2151	37095.54688	-18123.78855	0	1.668375034
Dayhoff+G	40366.98829	37320.95649	-18257.44417	n/a	1.859289237
Dayhoff+G+I	40376.76519	37323.04927	-18257.44417	0	1.859289237
cpREV+G	40506.19819	37550.1864	-18372.04913	n/a	1.432003701
cpREV+G+I	40605.9751	37552.25917	-18372.04913	0	1.432003701
LG	40623.09014	37584.74272	-18390.38355	n/a	n/a
LG+I	40632.86267	37586.83088	-18390.38137	0	n/a
JTT+F	40699.48467	37515.1757	-18335.70023	n/a	n/a
JTT+I+F	40709.26157	37517.27285	-18335.70023	0	n/a
JTT	40726.40592	37688.05851	-18442.04144	n/a	n/a
JTT+I	40736.18282	37690.15103	-18442.04144	0	n/a
WAG	40781.31525	37742.96783	-18469.49611	n/a	n/a
WAG+I	40791.09215	37745.06036	-18469.49611	0	n/a
LG+I+F	40818.42531	37626.4366	-18390.28211	0	n/a
LG+F	40818.57342	37634.26445	-18395.24461	n/a	n/a
WAG+F	40834.69486	37650.38589	-18403.30533	n/a	n/a
WAG+I+F	40845.4277	37653.43898	-18403.30533	0	n/a
cpREV+G+F	40891.32963	37699.34091	-18426.73426	n/a	0.823905324
cpREV+G+I+F	40901.10654	37701.43831	-18426.73426	0	0.823905324
rREV+F	41180.67195	37996.36298	-18576.29387	n/a	n/a
rREV	41182.7319	38144.38439	-18670.20438	n/a	n/a
rREV+I+F	41186.14416	37994.15544	-18574.14153	0	n/a
rREV+I	41192.50871	38146.47692	-18670.20438	0	n/a
miREV24+G	41232.86287	38186.85106	-18690.38146	n/a	1.818786967
miREV24+G+I	41242.83977	38188.92385	-18690.38146	0	1.818786967
Dayhoff+F	41274.84733	38090.53836	-18623.38156	n/a	n/a
Dayhoff+I+F	41284.62423	38092.63551	-18623.38156	0	n/a
miREV24+F	41357.1917	38172.88274	-18664.55375	n/a	n/a
Dayhoff	41401.30314	38362.95573	-18779.49005	n/a	n/a
Dayhoff+I	41411.08005	38365.04826	-18779.49005	0	n/a
miREV24+I+F	41571.25458	38379.26586	-18766.69674	0	n/a
cpREV+I	41737.51396	38691.48216	-18942.70701	0	n/a
cpREV	41892.24007	38853.89265	-19024.95852	n/a	n/a
cpREV+I+F	42196.46717	39004.47845	-19079.30303	0	n/a
cpREV+F	42531.07681	39346.76784	-19251.4963	n/a	n/a
miREV24	42545.47299	39507.12558	-19351.57498	n/a	n/a
miREV24+I	42556.60708	39510.57529	-19352.25357	0	n/a

Appendix table 3.2: Maximum Likelihood fits of 56 different amino acid substitution models. Bayesian Information Criterion (BIC); Akaike Information Criterion, corrected (AICc) and Maximum Likelihood (lnL) are shown. Where applicable, estimates of gamma (G) shape parameter and the estimated fraction of invariant (I) sites are presented.

Model	BIC	AICc	lnL	Invariant	Gamma
LG+G	39593.7796	36547.74781	-17870.83983	n/a	1.921500972
LG+G+I	39603.5565	36549.84058	-17870.83983	0	1.921500972
JTT+G+F	39738.16382	36546.1751	-17850.15136	n/a	1.961109612
JTT+G+I+F	39747.94072	36548.27249	-17850.15136	0	1.961109612
JTT+G	39762.39205	36716.36026	-17955.14606	n/a	1.93463646
LG+G+F	39765.76133	36573.77261	-17863.95011	n/a	1.931976408
JTT+G+I	39772.16896	36718.45303	-17955.14606	0	1.93463646
LG+G+I+F	39775.53823	36575.87	-17863.95011	0	1.931976408
WAG+G	39921.44702	36875.41523	-18034.67354	n/a	2.095342052
WAG+G+I	39931.22392	36877.508	-18034.67354	0	2.095342052
WAG+G+F	39955.11164	36763.12292	-17958.62527	n/a	2.038786363
WAG+G+I+F	39964.88855	36765.22032	-17958.62527	0	2.038786363
rREV+G+I	40084.93793	36892.94921	-18023.53841	n/a	2.016903112
rREV+G+I+F	40094.71483	36895.0466	-18023.53841	0	2.016903112
rREV+G	40146.26238	37100.23059	-18147.08122	n/a	2.041780867
rREV+G+I	40156.03928	37102.32336	-18147.08122	0	2.041780867
Dayhoff+G+F	40236.99348	37045.00476	-18099.56619	n/a	1.784515617
Dayhoff+G+I+F	40246.77038	37047.10216	-18099.56619	0	1.784515617
mtREV24+G+F	40285.4382	37093.44948	-18123.78855	n/a	1.668375034
mtREV24+G+I+F	40295.2151	37095.54688	-18123.78855	0	1.668375034
Dayhoff+G	40366.98829	37320.95649	-18257.44417	n/a	1.859289237
Dayhoff+G+I	40376.76519	37323.04927	-18257.44417	0	1.859289237
cpREV+G	40596.19819	37550.1664	-18372.04913	n/a	1.432003701
cpREV+G+I	40605.9751	37552.25917	-18372.04913	0	1.432003701
LG	40623.09014	37584.74272	-18390.38355	n/a	n/a
LG+I	40632.86267	37586.83088	-18390.38137	0	n/a
JTT+F	40699.48467	37515.1757	-18335.70023	n/a	n/a
JTT+I+F	40709.26157	37517.27285	-18335.70023	0	n/a
JTT	40726.40592	37688.05851	-18442.04144	n/a	n/a
JTT+I	40736.18282	37690.15103	-18442.04144	0	n/a
WAG	40781.31525	37742.96783	-18469.49611	n/a	n/a
WAG+I	40791.09215	37745.06036	-18469.49611	0	n/a
LG+I+F	40818.42531	37626.4366	-18390.28211	0	n/a
LG+F	40818.57342	37634.26445	-18395.24461	n/a	n/a
WAG+F	40834.69486	37650.38589	-18403.30533	n/a	n/a
WAG+I+F	40845.4277	37653.43898	-18403.7833	0	n/a
cpREV+G+F	40891.32963	37699.34091	-18426.73426	n/a	0.823905324
cpREV+G+I+F	40901.10654	37701.43831	-18426.73426	0	0.823905324
rREV+F	41180.67195	37996.36298	-18576.29387	n/a	n/a
rREV	41182.7318	38144.38439	-18670.20438	n/a	n/a
rREV+I+F	41186.14416	37994.15544	-18574.14153	0	n/a
rREV+I	41192.50871	38146.47692	-18670.20438	0	n/a
mtREV24+G	41232.86287	38186.83108	-18690.38146	n/a	1.818786967
mtREV24+G+I	41242.63977	38188.92385	-18690.38146	0	1.818786967
Dayhoff+F	41274.84733	38090.53836	-18623.38156	n/a	n/a
Dayhoff+I+F	41284.62423	38092.63551	-18623.38156	0	n/a
mtREV24+F	41357.1917	38172.88274	-18664.55375	n/a	n/a
Dayhoff	41401.30314	38362.95573	-18779.49005	n/a	n/a
Dayhoff+I	41411.08005	38365.04826	-18779.49005	0	n/a
mtREV24+I+F	41571.25458	38379.26586	-18766.69674	0	n/a
cpREV+I	41737.51396	38691.48216	-18942.70701	0	n/a
cpREV	41892.24007	38853.89265	-19024.95852	n/a	n/a
cpREV+I+F	42196.46717	39004.47845	-19079.30303	0	n/a
cpREV+F	42531.07681	39346.76784	-19251.49663	n/a	n/a
mtREV24	42545.47299	39507.12558	-19351.57498	n/a	n/a
mtREV24+I	42556.60708	39510.57529	-19352.25357	0	n/a

Appendix table 3.3: Comparison of CYPomes of four species of managed bee pollinators (*A. mellifera*; *B. terrestris*; *O. bicornis* and *M. rotundata*). All protein sequences accessed from NCBI protein database.

Comparison of the CYPome's of 4 managed bee species									
	<i>Megachile rotundata</i>	<i>Osmia bicornis</i>	<i>Apis mellifera</i>	<i>Bombus terrestris</i>		<i>Megachile rotundata</i>	<i>Osmia bicornis</i>	<i>Apis mellifera</i>	<i>Bombus terrestris</i>
CYP3 CYP6s	CYP6AQ52	CYP6AQ55	CYP6AQ1	CYP6AQ26	CYP3 CYP336As	CYP336A33	CYP336A35	CYP336A1	CYP336A22
	CYP6AQ53			CYP6AQ27		CYP336A34	CYP336A36		CYP336A23
	CYP6AQ54			CYP6AQ28		CYP336M1	CYP336L1		CYP336A24
				CYP6AQ29					CYP336A25
				CYP6AQ31					
				CYP6AQ33					
	CYP6AS108	CYP6AS121	CYP6AS1	CYP6AS5					
	CYP6AS109	CYP6AS122	CYP6AS2	CYP6AS7					
	CYP6AS110	CYP6AS123	CYP6AS3	CYP6AS10					
	CYP6AS111	CYP6AS124	CYP6AS4	CYP6AS12	CYP2	CYP15A1	CYP15A1	CYP15A1	CYP15A1
	CYP6AS112	CYP6AS125	CYP6AS5	CYP6AS13		CYP18A1	CYP18A1	CYP18A1	CYP18A1
	CYP6AS113	CYP6AS126	CYP6AS5P	CYP6AS19		CYP303A1	CYP303A1	CYP303A1	CYP303A1
	CYP6AS114	CYP6AS127	CYP6AS7	CYP6AS72		CYP305D1	CYP305D1	CYP305D1	CYP305D1
	CYP6AS115	CYP6AS128	CYP6AS8	CYP6AS73			CYP306A1	CYP306A1	CYP306A1
	CYP6AS116	CYP6AS129	CYP6AS10	CYP6AS74		CYP307B1	CYP307B1	CYP307B1	CYP307B1
	CYP6AS117	CYP6AS130	CYP6AS11	CYP6AS75		CYP343A1	CYP343A1	CYP343A1	CYP343A1
	CYP6AS118	CYP6AS131	CYP6AS12	CYP6AS76		CYP369A1	CYP369A1	CYP369A1	CYP369A1
	CYP6AS119	CYP6AS132	CYP6AS13	CYP6AS77					
	CYP6AS120	CYP6AS133	CYP6AS14						
		CYP6AS134	CYP6AS15						
	CYP6AS135	CYP6AS16							
	CYP6AS136	CYP6AS17							
	CYP6AS151	CYP6AS18		CYP4	CYP4G11	CYP4G11	CYP4G11	CYP4G11	
		CYP6AS19			CYP4G202	CYP4G202			
CYP6BC1	CYP6BC1	CYP6BC1	CYP6BC1		CYP4AA1	CYP4AA1		CYP4AA1	
CYP6BD1	CYP6BD1	CYP6BD1	CYP6BD1		CYP4AB3	CYP4AB3	CYP4AB3	CYP4AB3	
CYP6BE1	CYP6BE1	CYP6BE1	CYP6BE1	CYP4AV1	CYP4AV1	CYP4AV1	CYP4AV1		
CYP6BE1P									
CYP9DN1	CYP9DN1								
CYP9P2	CYP9P2	CYP9P1	CYP9P1						
CYP9P22	CYP9P22	CYP9P2	CYP9P2						
CYP9P23	CYP9P24								
CYP9DM1	CYP9BU1	CYP9Q1	CYP9Q4						
CYP9DM2	CYP9BU2	CYP9Q2	CYP9Q5	Mitochondrial	CYP301A1	CYP301A1	CYP301A1	CYP301A1	
		CYP9Q3	CYP9Q6		CYP301B1	CYP301B1	CYP301B1	CYP301B1	
CYP9R1	CYP9R1	CYP9R1	CYP9R1		CYP302A1	CYP302A1	CYP302A1	CYP302A1	
CYP9R58	CYP9R38				CYP314A1	CYP314A1	CYP314A1	CYP314A1	
CYP9R59	CYP9R39				CYP315A1	CYP315A1	CYP315A1	CYP315A1	
		CYP9S1	CYP9S1		CYP334A1	CYP334A1	CYP334A1	CYP334A1	

Appendix table 3.4: Genomic region of *A. mellifera* DH4 linkage group LG14, Amel_HAV3.1, WGS containing the CYP9 cluster, compared to scaffolds of three managed bee pollinators: *B. terrestris*, *M. rotundata* and *O. bicornis*. A BLASTn search through the genomes of the other bee species, using this region as a query sequence, was performed to establish which scaffold(s) had possible syntenic regions. Individual scaffolds are coloured differently.

Genes from <i>Apis mellifera</i> LG14 [4526896-5965014] in order	Genomic position <i>B. terrestris</i> BLASTn hits	Position on <i>B. terrestris</i> scaffold (looked up manually on the entire sequence)	Notes	Genomic position <i>M. rotundata</i> BLASTn hits	Position on <i>M. rotundata</i> scaffold (looked up manually on the entire sequence)	Notes	Genomic position <i>O. bicornis</i> BLASTn hits	Position on <i>O. bicornis</i> scaffold (looked up manually on the entire sequence)
mediator of RNA polymerase II transcription factor subunit 16				MROT_1.0_wd_0030	29190-29191	CYP phosphatase (serine/threonine kinase 2, serine/threonine protein kinase-like)		
thioredoxin-like protein 1				MROT_1.0_wd_0030	287217-288422 comp.			
uncharacterised LOC276153								
carboxypeptidase B				MROT_1.0_wd_0030	75497-76148 comp.			
homeobox protein engrailed-2-B								
adenylosuccinate synthetase Phenoloxidase subunit A3				MROT_1.0_wd_0464	75495-83824 comp.	MROT dehydrogenase (NAD-dependent) alpha subunit (serine/threonine kinase receptor-associated protein 1), MCOGA complex subunit MCOGA, then phenoloxidase 2 (below)		
Facilitated trehalose transporter Tret1-2 homolog				MROT_1.0_wd_0464	110385-131083 comp.	uncharacterised region XVIII, basic multi-phosphatase family member 1, protein kinase, probable fructosyltransferase (FRTF) class II (more genes follow, joining to transporter molecules C/D below)		
Facilitated trehalose transporter Tret1				MROT_1.0_wd_0030	749163-776602 comp.	below, then joins to transporter molecules B/C below		
Solute carrier family 12 member 9	LG B01	5216019..5217673 comp.	Fairlie-like proteinase 1, glyoxylate isomerase-like chain 3, chain 2-like, polyphosphate receptor subunit beta-like 1, neuronal polyphosphate receptor subunit alpha?	MROT_1.0_wd_1303	33712-33943	Allosteric Na/Cd transporter regulator, Na/Cd transporter protein, Serine/threonine kinase family 35 member PB, ATP-driven synthase	Ochromis_v3 scaffold00020	792152..796882 comp. 1
Enoyl-fatty-carrier-protein reductase, mitochondrial	LG B01	5223003..5223988		MROT_1.0_wd_1303	31038-33125 comp.	probable trans-2-enoyl-CoA reductase, mitochondrial	Ochromis_v3 scaffold00020	796937..798848
Protein OSCP1	LG B01	5224009..5225549 comp.		MROT_1.0_wd_1303	26913-31555		Ochromis_v3 scaffold00020	798859..801459 comp.
Protein IIB homolog	LG B01	5226630..5229253 comp.		MROT_1.0_wd_1303	27601-28816		Ochromis_v3 scaffold00020	LOC119871281
Nuclear valosin-containing protein-like	LG B01	5238990..5238920		MROT_1.0_wd_1303	17168-21498 comp.		Ochromis_v3 scaffold00020	807724..812295
Serine/threonine-protein kinase Aurora-2	LG B01	5239237..5241504 comp.		MROT_1.0_wd_1303	13832-17096		Ochromis_v3 scaffold00020	816967..818398 comp.
Ferritin subunit	LG B01	5241871..5243873 comp. 1		MROT_1.0_wd_1303	10565-13883 1		Ochromis_v3 scaffold00020	424040-427499
Ferritin heavy polypeptide-like 17								
1-phosphatidylinositol 4,5-bisphosphate phosphodiesterase gamma-1	LG B01	5242081..5256548 comp. 1		MROT_1.0_wd_1303	2217-6876 1	Start of scaffold	Ochromis_v3 scaffold00020	413379-421873
Unchar LOC551741 (blast shows rat guanine nucleotide exchange factor J)	LG B01		Annotated as putative protein leg52 but BLASTn hit is rat guanine nucleotide exchange factor J					
Membralin	LG B01	5259426..5262355 Membralin		MROT_1.0_wd_0244	24414-324133 comp. Membralin (P58335-012882 comp. CYP9P2)	Also these genes did not appear to fly but guanine nucleotide exchange factor J, using guanine nucleotide exchange factor J, also did not fly but guanine nucleotide exchange factor J, also did not fly but guanine nucleotide exchange factor J	Ochromis_v3 scaffold00060	17165..22080 Membralin (177645-179202 CYP9P2)
CYP9P2	LG B01	(5262068-5267867 CYP9P1)		MROT_1.0_wd_0244	(514389-5168102 comp. CYP9P2) 1		Ochromis_v3 scaffold00060	(180428-182658 CYP9P2) 4
CYP9P1	LG B01	5271276..5273576 CYP9P2		MROT_1.0_wd_0244	519350-516492 comp. CYP9P2		Ochromis_v3 scaffold00060	182512-183548 CYP9P2
Unchar LOC107965493	LG B01	5279462..5281239 CYP9P1		MROT_1.0_wd_0244	564756-525694 CYP9P1		Ochromis_v3 scaffold00060	274100-278153 CYP9P1
Mir137								
CYP9Q3		5386650..5386502 comp. CYP9Q6		MROT_1.0_wd_0244	524163-524649 comp. CYP9Q6		Ochromis_v3 scaffold00060	276988-282452 CYP9Q6
CYP9S1		Unplaced scaffold 34733052 comp. CYP9S1	Orphan ---	MROT_1.0_wd_0244	348116-311679 CYP9Q2		Ochromis_v3 scaffold00060	282610-284648 CYP9S1
CYP9R1	LG B01	5386287..5390882 comp. CYP9Q4		MROT_1.0_wd_0244	548517-548837 comp. CYP9R1		Ochromis_v3 scaffold00060	286941-288404 comp. CYP9BLX
CYP9Q2	LG B01	5400739..5402017 comp. CYP9Q6		MROT_1.0_wd_0244	541817-543373 comp. CYP9Q2		Ochromis_v3 scaffold00060	290360-291889 comp. CYP9BLX
Negative elongation factor E								
Myosin IIIb	LG B01	5621617..5630356		MROT_1.0_wd_0100	3388609-3388423 comp.	Orphan ---	Ochromis_v3 scaffold00060	383158-403743
Alpha-cactinin	LG B01	5538185..5620209	After alpha cactinin Bombyx scaffold moves to 3002 proteinase regulatory subunit 6k-D, (BAC) cloning program 6k, beta-1-specific, glyoxylate isomerase-like chain 3-like, glyoxylate kinase, zinc finger protein 583, glutamate receptor ionotropic, kainate 2	MROT_1.0_wd_0244	11820-121492 comp. 1	First uncharacterised gene on the scaffold ---	Ochromis_v3 scaffold00060	404469-463103
Unchar LOC107965484								
Unchar LOC102656562								
NADPH oxidase 5	LG B14	9611468..9625797 comp. orphaned ---	neighbour genes, diacylglycerol kinase, epsilon, B-cell receptor-associated protein 31	MROT_1.0_wd_0154	33914-80594 comp.	Orphan ---	Ochromis_v3 scaffold00077	
Band 4.1-like protein 5	LG B14	9262902..9267388	From 403 ribosomal protein SA 1	MROT_1.0_wd_0100	422348-45489 comp.	From Thylakoid embryonic factor	Ochromis_v3 scaffold00090	102445-105087 comp. 1
Kinase D-interacting substrate of 220kDa	LG B14	3269269..3270522 comp.		MROT_1.0_wd_0100	398269-421718 & 269599-402979		Ochromis_v3 scaffold00090	61348-103291
tRNA (guanine-N(7))-methyltransferase	LG B14	3315144..3316636		MROT_1.0_wd_0100	385336-392012 comp.		Ochromis_v3 scaffold00090	76946-80188 comp. 1
Cilia- and flagella-associated protein 65		3314467..3322292 comp.				379789-381202 coded on domain-containing protein 128-like	Ochromis_v3 scaffold00090	73615-79502 1
Unchar LOC2727650								
Acylamino-acid-releasing enzyme	LG B14	3325108..3328627 comp.		MROT_1.0_wd_0100	373603-377437		Ochromis_v3 scaffold00090	87302-92714 1
EF-hand domain-containing protein D2		3329494..3329293 comp.			261917-529456		Ochromis_v3 scaffold00090	97968-97971 1
Unchar LOC100578049								
E3 ubiquitin-protein ligase parkin	LG B14	3365500..3361200		MROT_1.0_wd_0100	360617-358837		Ochromis_v3 scaffold00090	52352-55866 comp. 1
Brachyury protein	LG B14	3361401..3368427 comp.		MROT_1.0_wd_0100	360389-358483			
Unchar LOC102654696								
Unchar LOC100578654								
Unchar LOC113219234								
dnaJ homolog subfamily B member 13-like		3365449..3368390		MROT_1.0_wd_0100	329169-330376 comp.	Also have dnaJ homolog subfamily B member 13-like, SC7661-028001	Ochromis_v3 scaffold00090	32426-33676 comp. 1
Unchar LOC102654842								
Folliculin-interacting protein 2	LG B14	3368346..3368442	protein 2	MROT_1.0_wd_0100	326306-327493 comp.		Ochromis_v3 scaffold00090	28536-31294 comp. 1

Region ~500Kbp upstream and downstream

T-complex protein 1 subunit delta	LG B14	339592_339599 comp 1		MRDT_1.0 ref_0120	31768-32071		
Protein MON2 homolog	LG B14	339673_340763 L		MRDT_1.0 ref_0120	30215-31760 comp		Obioma_v3 scaffold00050 13790-21919 comp 1
CTD small phosphatase-like protein 2	LG B14	341048_341295 L		MRDT_1.0 ref_0120	30378-30871 comp		Obioma_v3 scaffold00050 9180-13761 comp 1
Melanin-2	LG B14	341432_341953 comp 1		MRDT_1.0 ref_0120	29629-30853		Obioma_v3 scaffold00050 3030-10188 1
60S ribosomal protein L23		341920_342078 L	Gene to T2451 (ribophagy-associated) 1 beta subcomplex subunit 10	MRDT_1.0 ref_0120	29227-29786 comp		Obioma_v3 scaffold00050 238-3719 comp 1
Unchar LOC113219229							
Thyrotroph embryonic factor - end of 500,000bp	LG B14	281181_309988 comp 1	From Protein extract[1]	MRDT_1.0 ref_0120	42867-54358 comp	Is to Bacc1.1 (beta protein 0) from Glutamine oxidase channel but with many genes below[1]	Obioma_v3 scaffold00052 63800-102737 1
Histone H1, gonadal							
Unchar LOC107965504							
Unchar LOC107965502							
J domain-containing protein							
Unchar LOC100576767							
Probable 28S ribosomal protein S16		312750_312815 L		MRDT_1.0 ref_0120	29789-30704 comp	From oligonucleotide-activated T7CT1 promoter protein, regulator protein, sorting-associated protein 41 homolog	Obioma_v3 scaffold00052 3999-38705 comp 1
TRNA1-UAU		313652_313415 comp 1			287628-287697		
Unchar LOC100577506							
Protein POLR1D							Obioma_v3 scaffold00052 35051-28895 comp 1
Pentatricopeptide repeat-containing protein 2, mitochondrial	LG B14	313769_313840 L					Obioma_v3 scaffold00052 30599-32563 comp 1
Glutamate-gated chloride channel	LG B14	314040_318270 L		MRDT_1.0 ref_0120	28351-287025 comp	Gene to hhhhh beta chain, growth differentiation factor 11, four several other genes to Thyrotroph embryonic factor	Obioma_v3 scaffold00052 1442-30035 comp 1
Mir3786							
DNA-directed RNA polymerases I, II and III subunit RPABC1	LG B14	3186130_318712 L		MRDT_1.0 ref_0814	4703-4537 comp 1	-- no further genes from Apla LG14 extract	Obioma_v3 scaffold00046 33684-34835
Protein enhancer of sevenless 2B	LG B14	318740_319208 comp 1		MRDT_1.0 ref_0814	4334-4587 1		Obioma_v3 scaffold00046 339736-343421 comp 1
Unchar LOC102656671							
Unchar LOC410487							
Integrator complex subunit 13							Obioma_v3 scaffold00046 37346-37687 comp 1
40S ribosomal protein SA		329710_328190 L	Gene to hhhhh on B1, Bacc1.1 (beta protein 1)	MRDT_1.0 ref_0814	734-2801 comp 1		Obioma_v3 scaffold00046 37688-37891 ---
Transmembrane and TPR repeat-containing protein CG4050	LG B14	3308191_3511862 L	Gene to URP then eventually on to transcriptional regulator ATRX homolog 1	MRDT_1.0 ref_0484	8127-86848 comp 1	From - Transcriptional regulator ATRX homolog 1	Obioma_v3 scaffold00046 37688-37891 ---
Neurotactin	LG B14	349787_350528 comp 1		MRDT_1.0 ref_0484	8722-9378 L		Obioma_v3 scaffold00046 179031-18585 1
Unchar LOC409563							
Mannose-1-phosphate guanytransferase alpha-A	LG B14	348080_348194 L		MRDT_1.0 ref_0484	9954-101725 comp 1		Obioma_v3 scaffold00046 19015-18291 comp 1
Congested-like trachea protein	LG B14	3484120_3489613 L		MRDT_1.0 ref_0484	101901-108787 comp 1		Obioma_v3 scaffold00046 193094-18808 comp 1
Leucine-rich repeat-containing protein 15							
5'-AMP-activated protein kinase catalytic subunit alpha-2	LG B14	348054_348309 comp 1					
adapter molecule Crk	LG B14	3477284_3480438 comp 1		MRDT_1.0 ref_0484	115003-113356 L ---	Gene to protein C1 (beta12) homolog, ubiquitin-40S ribosomal protein S7A, fibronectin type-III domain protein 182C1, homolog	Obioma_v3 scaffold00046 210708-20488 L
Ubiquitin-like with PHD and ring finger domains 1		3472863_347697 comp 1	E3 ubiquitin-protein ligase UBR1				Obioma_v3 scaffold00046 205630-20915 1
cutA divalent cation tolerance homolog protein Fer2-like		3471284_3472529 L		MRDT_1.0 ref_0530	8623-87943	protein cutA homolog	Obioma_v3 scaffold00046 20983-21049 comp 1
Inositol hexakisphosphate and diphosphoinositol-pentakisphosphate kinase 2	LG B14	3449732_3482428 L		MRDT_1.0 ref_0530	53953-58881 & 45046-53891 comp		Obioma_v3 scaffold00046 211818-212622 comp 1
Dolichylidiphosphate 1	LG B14	3444489_3448886 L		MRDT_1.0 ref_0530	30194-32784		Obioma_v3 scaffold00046 215440-220163 comp 1
Meiosis regulator and mRNA stability factor 1							Obioma_v3 scaffold00046 23988-23994 comp 1
NADH dehydrogenase [ubiquinone] 1 beta subcomplex subunit 10		342042_342072 L	From 60S ribosomal protein L23.1	MRDT_1.0 ref_0530	17818-30088 comp	meiosis arrest female protein 1	Obioma_v3 scaffold00046 233393-24443 1
60S ribosomal protein L3							Obioma_v3 scaffold00046 24474-24538 1
small nucleolar RNA U43							Obioma_v3 scaffold00046 24534-24870 1
Unchar LOC113219235							Obioma_v3 scaffold00046 24856-24925 1
Insulin-like peptide 2							
Unchar LOC724646							
Transcriptional regulator ATRX homolog	LG B14	3520471_3530409 L	From Transmembrane and TPR repeat-containing protein CG4050	MRDT_1.0 ref_0484	86792-74160 comp 1	Gene to transmembrane and TPR repeat-containing protein CG4050-like 1	Obioma_v3 scaffold00046 198858-187379 comp 1
iduronate 2-sulfatase	LG B14	3530851_3532037 comp 1		MRDT_1.0 ref_0484	64201-68028 L		Obioma_v3 scaffold00046 157282-158818 1
Asparagine-rich zinc finger protein AZF1	LG B14	3531136_3541004 L		MRDT_1.0 ref_0484	60214-59483 comp 1		Obioma_v3 scaffold00046 153463-152703 1
Glutamate receptor ionotropic, delta 2							Obioma_v3 scaffold00046 160202-163380 comp 1
Unchar LOC100578443							
YEATS domain-containing protein 2	LG B14	3543334_3548232 comp 1		MRDT_1.0 ref_0484	49502-54477 L		
OTU domain-containing protein 5-B	LG B14	3547916_3553889 L		MRDT_1.0 ref_0484	43887-44885 comp 1		Obioma_v3 scaffold00046 34826-42181 comp 1
Serine/threonine-protein kinase Nek5		3556100_3568337 L			41873-42887 comp 1		
MAM and LDL-receptor class A domain-containing protein 1	LG B14	3558310_3559848 L		MRDT_1.0 ref_0484	37782-41872 comp 1		Obioma_v3 scaffold00046 26227-34863 comp 1
rRNA methyltransferase 3, mitochondrial	LG B14	3558566_2961731 comp 1		MRDT_1.0 ref_0484	30884-37716 L		Obioma_v3 scaffold00046 37796-39461 1
Unchar LOC100577121							
Probable ATP-dependent RNA helicase DDX17	LG B14	3578992_3880581 comp 1		MRDT_1.0 ref_0484	17266-24070 v		Obioma_v3 scaffold00046 32468-18221 1
Unchar LOC724121							

o investigation due to syntenic break that occurs in both Megachilidae spp - membrales neighbours alostatrin A receptor

Region added to	Endoplasmic reticulum mannosyl-oligosaccharide 1,2-alpha-mannosidase	LG B14	350318..350692 comp ;	Gene in uncharacterised region added to scaffold 1 --> see more Apps LG14 ...	MR07_1.0.scf_0464	4335-4330 ;	First gene on the scaffold -->	Obicoma_v3 scaffold00101	977-5362 ;
	Starazin related protein STG-1								
	Unchar LOC107965519								
	Unchar LOC107965520								
	Esterase A2								
	Esterase E4-like							Obicoma_v3 scaffold00138	
	Collagen alpha-1(IV) chain								
	Allostatin A receptor	LG B14	24111..51145 comp orphaned -->	1985 annotated gene on scaffold --> links to beta-casein protein case, voltage-dependent calcium channel gamma2, tubulin, multiple insulin polypeptides alpha2, alpha3	MR07_1.0.scf_0244	41734341101 ;		Obicoma_v3 scaffold00090	91779-117717 comp ;
	Trehalase		1804100..1804384 ;	From TPNAV1 LGSU	MR07_1.0.scf_0244	122251109898 ;		Obicoma_v3 scaffold00090	19888-49421 comp ;
	Unchar LOC109577953								
	Unchar LOC107965499		1640216..1643424 comp ;		MR07_1.0.scf_0244	719911747600-comp ;		Obicoma_v3 scaffold00090	13287-15688 ;
	B9 domain-containing protein1-like		1645607..1653333 comp ;		MR07_1.0.scf_0244	162654749484-comp ;		Obicoma_v3 scaffold00090	6252-13183 ;
	Probable aconitate hydratase, mitochondrial	LG B14	1680107..1681462 ;		MR07_1.0.scf_0244	828294-832376 ;		Obicoma_v3 scaffold00129	41784-43165 comp ; -->
	Mucin-19	LG B14	1702744..1727093 comp ;	MR07_1.0.scf_0244	864026-864493-comp ;			Obicoma_v3 scaffold00129	381422-408873 ;
	C1q-like venom protein		1705225..1707790 ;	MR07_1.0.scf_0244	864026-864493-comp ;			Obicoma_v3 scaffold00129	40204-40443 comp ;
	Adenylosuccinate lyase	LG B14	1705225..1707790 ;	MR07_1.0.scf_0244	864026-864493-comp ;			Obicoma_v3 scaffold00129	40204-40443 comp ;
	Protein snakeskin								
	Unchar LOC100577682								
	Unchar LOC102656664								
	Unchar LOC107964014								
	Unchar LOC100576934								
	Unconventional myosin-XV				MR07_1.0.scf_0244	86098101-86098100-comp ;	see uncharacterised region from MR07 scaffold1	Obicoma_v3 scaffold00129	30946-31101 ;

TOTAL = 98 (plus 34 uncharacterised)

18 non-CYPs in 500,000bp stretch upstream

9/18 in B01 = 50%

20 non-CYPs in 500,000bp stretch downstream

2/3 in B01 = 66.7%
15/17 in B14 = 88.24%
Together 17/20 = 85%

Total over 2 scaffolds

26/38 = 68.42%

With CYPs add 7 for Am and 6 for Bt

32/45 = 71.11%

18 non-CYPs in 500,000bp stretch upstream

8/18 in scf_1303
1/18 in scf_0244
TOTAL = 9/18 = 50%

20 non-CYPs in 500,000bp stretch downstream

2/20 in scf_0244
15/20 in scf_0120
TOTAL = 17/20 = 85%

Total over 3 scaffolds

26/38 = 68.42%

With CYPs add 7 for Am and 6 for Mr

8/45 in scf_1303
11/45 in scf_0244
15/45 in scf_0120
Total 34/45 = 75.55%

18 non-CYPs in 500,000bp stretch upstream

5/18 in scf_00020
1/18 in scf_00060
TOTAL = 6/18 = 33.33%

20 non-CYPs in 500,000bp stretch downstream

2/20 in scf_00060
13/20 in scf_00090
TOTAL = 15/20 = 75%

Total over 3 scaffolds

21/38 = 55.26

With CYPs add 7 for Am and 6 for Ob

5/45 in scf_00020
11/45 in scf_00060
13/45 in scf_0120
Total 29/45 = 64.44%

Appendix table 3.5: Genomic region of *A. mellifera* DH4 linkage group LG13, Amel_HAV3.1, WGS containing the CYP6AS cluster, compared to scaffolds of three managed bee pollinators: *B. terrestris*, *M. rotundata*, *O. bicornis* and *D. novaeangliae*. A BLASTn search through the genomes of the other bee species, using this region as a query sequence, was performed to establish which scaffold(s) had possible syntenic regions. Individual scaffolds are coloured differently.

Genes from <i>Apis mellifera</i> LG13 [1-1082529] in order	<i>Bombus terrestris</i> scaffold	Genes from <i>B. terrestris</i> scaffolds [the order is different to <i>Apis</i>]	<i>Megachile rotundata</i> scaffold	Genes from <i>M. rotundata</i> scaffolds	<i>Osmia bicornis</i> scaffold	Genes from <i>O. bicornis</i> scaffolds	<i>Dufourea novaeangliae</i> scaffold	Genes from <i>D. novaeangliae</i> scaffolds
globin 1	LG B13	globin 1						
MD-2-related lipid-recognition protein	LG B13	MD-2-related lipid-recognition protein	scf_0448 NW_003797462	MD-2-related lipid-recognition protein				
TRNA _S CGA	LG B13	TRNA _S CGA	scf_0449 NW_003797462	TRNA _S CGA				
echinoderm microtubule-associated protein like 2	LG B13	echinoderm microtubule-associated protein like 2	scf_0448 NW_003797462	echinoderm microtubule-associated protein like 2				
protein TIS11								
methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	LG B13	methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial	scf_0147 NW_003797193	methylmalonate-semialdehyde dehydrogenase [acylating], mitochondrial				
cell wall integrity and stress response component 2	LG B13	cell wall integrity and stress response component 2						
pyruvate carboxylase, mitochondrial	LG B13	pyruvate carboxylase, mitochondrial	scf_0147 NW_003797193	pyruvate carboxylase, mitochondrial				
bumetanide-sensitive sodium-(potassium)-chloride transporter	LG B13	bumetanide-sensitive sodium-(potassium)-chloride transporter	scf_0147 NW_003797193	bumetanide-sensitive sodium-(potassium)-chloride transporter				
ATP-dependent (S)-NAD(P)H-hydrate dehydratase	LG B13	ATP-dependent (S)-NAD(P)H-hydrate dehydratase	scf_0147 NW_003797193	ATP-dependent (S)-NAD(P)H-hydrate dehydratase				
P450 6a14	LG B13	P450 6A1	scf_0147 NW_003797193	P450 6a14				
8 more P450 6 family members	LG B13	6 more P450 6 family members	scf_0128 NW_003797174	8 more P450 6 family members				
P450 6a17	LG B13	P450 6A2	scf_0147 NW_003797193	P450 6B1-like				
PI-PLC X domain-containing protein 3	LG B13	PI-PLC X domain-containing protein 3	scf_0128 NW_003797174	PI-PLC X domain-containing protein 3				
coiled-coil domain-containing protein 47	LG B13	coiled-coil domain-containing protein 47	scf_0128 NW_003797174	coiled-coil domain-containing protein 47				
kinesin-like protein KIF3B	LG B13	kinesin-like protein KIF3B	scf_0128 NW_003797174	kinesin-like protein KIF3B				
potassium voltage-gated channel protein Shaw	LG B13	potassium voltage-gated channel protein Shaw	scf_0128 NW_003797174	potassium voltage-gated channel protein Shaw				
potassium voltage-gated channel protein Shaw	LG B13	potassium voltage-gated channel protein Shaw-like						
4 more P450 6 family members								
protein bowel	LG B13	protein bowel	scf_0128 NW_003797174	protein bowel				
protein bowel-like								
spermidine synthase	LG B13	spermidine synthase	scf_0147 NW_003797193	spermidine synthase				
lat1-1	LG B13	lat1-1	scf_0147 NW_003797193	lat1-1				
protein ERGIC-53	LG B13	protein ERGIC-53	scf_0147 NW_003797193	protein ERGIC-53				
vesicle transport through interaction with t-SNAREs homolog	LG B13	vesicle transport through interaction with t-SNAREs homolog 1A	scf_0147 NW_003797193	vesicle transport through interaction with t-SNAREs 1A				
long-chain-fatty-acid-CoA ligase 6	LG B13	long chain-fatty-acid-CoA ligase 1	scf_0147 NW_003797193	long chain-fatty-acid-CoA ligase 1				
growth/differentiation factor 2	LG B13	microsomal triglyceride transfer protein large subunit	scf_0147 NW_003797193	microsomal triglyceride transfer protein large subunit				
microsomal triglyceride transfer protein large subunit			scf_0128 NW_003797174	disk large-associated protein 5-like				
slitk large-associated protein 5	LG B13	paired mesoderm homeobox protein 2B						
paired mesoderm homeobox protein 2B								
TOTAL 41 genes (plus 10 uncharacterised)		TOTAL 30 genes (plus 11 uncharacterised)		TOTAL 20 genes scf_0147 TOTAL 8 genes scf_0128				
26 non CYP genes		22 non CYP genes		10 non CYP genes scf_0147 6 non CYP genes scf_0128		1 non CYP gene Scaffold000161 1 non CYP gene Scaffold000374		
		Number of potential homologs with Am = 22 22/26 = 84.62%		Number of potential homologs (across the 2 scfs) with Am = 16 61.53%		Number of potential homologs (across the 12 scfs) with Am = 2 7.69%		

Solo (orphan)scaffolds:
Scaffold00032
Scaffold01090
Scaffold00935
Scaffold00730
Scaffold00726
Scaffold00707
Scaffold00936
Scaffold00695
Scaffold00476
Scaffold00462
Scaffold00654

Non-orphan scaffolds:
Scaffold00014
Scaffold00039
Scaffold00091
Scaffold00198
Scaffold00161
Scaffold00274
Scaffold00321
Scaffold00358
Scaffold00374

11 scaffolds have orphaned P450 6A sequences. 29 non-orphaned scaffolds have shared flanking genes.

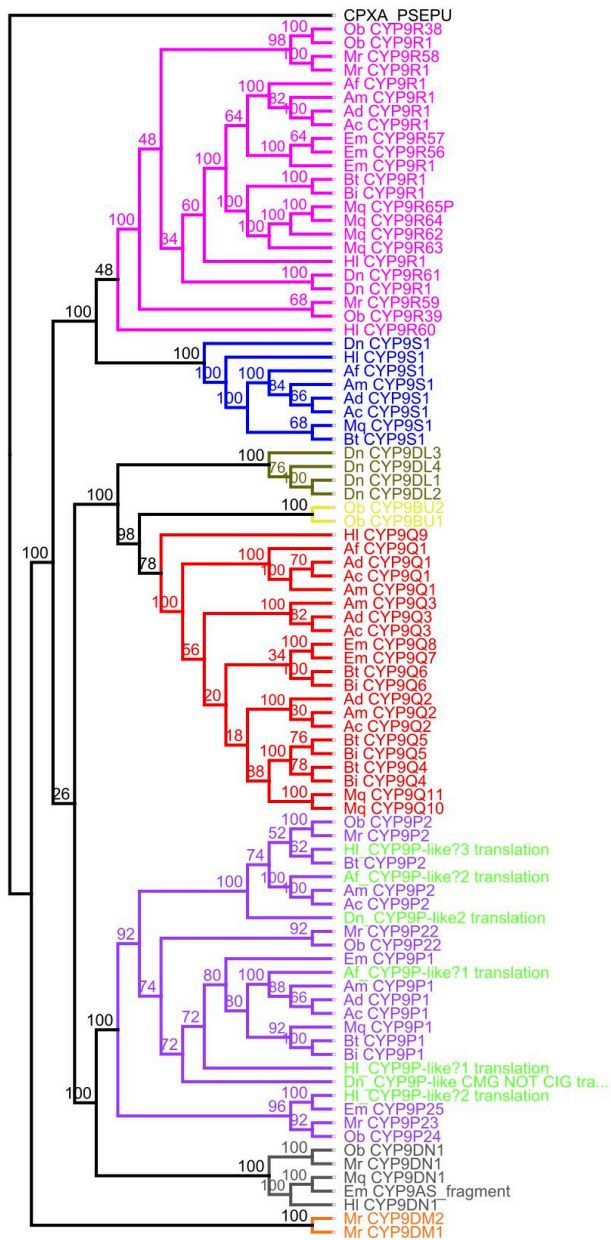
Appendix table 3.6: Distance matrix of the conserved motifs and SRSs of CYP9 proteins from the CYP9 cluster of four managed bee pollinators (*A. mellifera*, *B. terrestris*, *O. bicornis* and *M. rotundata*). The values for *A. mellifera* CYP9P1, CYP9R1 and CYP9Q3 (to compare % values for each CYP9 lineage) are outlined in red.

	Mr CYP9DM1	Mr CYP9DM2	4D62_1[Chai]	Ob CYP98U1	Ob CYP98U2	Am CYP9Q1	Am CYP9Q2	Bt CYP9Q4	Am CYP9Q5	Am CYP9Q3	Bt CYP9Q6	Mr CYP9DN1	Ob CYP9DN1	Am CYP9S1	Bt CYP9S1	Mr CYP9P2	Ob CYP9P2	Am CYP9P2	Bt CYP9P2	Am CYP9P1	Bt CYP9P1	Ob CYP9P2	Mr CYP9R2	Ob CYP9R2	Mr CYP9R59	Ob CYP9R39	Am CYP9R1	Bt CYP9R1	Mr CYP9R1	Mr CYP9R58	Ob CYP9R1	Ob CYP9R38			
Mr CYP9DM1	87.963	30.909	29.091	35.455	37.273	30.909	33.636	32.727	30.909	29.091	32.727	31.818	33.636	32.432	35.455	39.091	40	40.909	39.091	37.273	43	40	29.091	36.364	37.273	37.273	40	40	43.636	44.545	44.545	43.636			
Mr CYP9DM2	87.963	30	30	36.364	34.545	29.091	32.727	30.909	30.909	29.091	32.727	30	29.091	32.432	36.364	40	40	39.091	40	37.273	36.364	40	39.091	31.818	35.455	37.273	40	39.091	42.727	43.636	44.545	42.727	41.818		
4D62_1[Chai]	30.909	30	30	36.364	36.364	31.818	36.364	30.909	30.909	31.818	31.818	40	44.545	37.838	40	36.364	37.273	37.273	37.273	39.091	39.091	42	40	37.273	38.182	36.364	38.182	34.545	34.545	36.364	37.273	34.545	33.636		
Ob CYP98U1	29.091	30	30	36.364	56.364	45.455	47.273	41.818	40	47.273	51.818	42.727	40	38.739	37.273	44.545	46.364	45.455	43.636	43.636	46	44.545	41.818	42.727	40.909	42.727	42.727	40.909	43.636	42.727	41.818	42.727	41.818	42.727	
Ob CYP98U2	35.455	36.364	36.364	56.364	40	43.636	45.455	44.545	45.455	45.455	44.545	49.091	43.636	36.937	31.818	42.727	41.818	42.727	43.636	41.818	48	45.455	40	39.091	43.636	43.636	43.636	45.455	43.636	43.636	43.636	41.818	42.727	41.818	42.727
Am CYP9Q1	37.273	34.545	31.818	45.455	40	60	52.727	53.636	51.818	52.727	43.636	40.909	42.342	39.091	41.818	40.909	43.636	42.727	48.182	49.091	46	44.545	38.182	43.636	40.909	36.364	44.545	41.818	40	40.909	42.727	41.818	42.727	41.818	
Am CYP9Q2	30.909	29.091	36.364	47.273	43.636	60	60	63.636	64.545	58.182	59.091	43.636	42.727	39.64	36.364	49.091	50	50	48.182	45.455	47	46.364	41.818	45.455	41.818	42.727	45.455	48.182	48.182	49.091	47.273	46.364	46.364		
Bt CYP9Q4	33.636	32.727	30.909	41.818	45.455	52.727	63.636	90	57.273	56.364	47.273	44.545	44.144	36.364	43.636	43.636	44.545	43.636	43.636	40.909	43	44.545	39.091	41.818	40.909	43.636	47.273	50.909	50.909	50.909	53.636	52.727	52.727		
Bt CYP9Q5	32.727	30.909	30.909	40	44.545	53.636	64.545	90	56.364	56.364	46.364	42.727	44.144	35.455	45.455	45.455	46.364	45.455	46.364	41.818	45	47.273	37.273	42.727	40.909	44.545	47.273	50	50.909	50	52.727	52.727	52.727		
Am CYP9Q3	29.091	30.909	31.818	47.273	45.455	51.818	58.182	57.273	56.364	60	42.727	39.091	39.64	36.364	50	49.091	50	50	49.091	45.455	46	43.636	38.182	42.727	41.818	44.545	49.091	50.909	44.545	46.364	47.273	46.364	46.364		
Bt CYP9Q6	32.727	32.727	31.818	51.818	44.545	52.727	59.091	56.364	56.364	60	43.636	40.909	36.036	32.727	44.545	44.545	45.455	44.545	45.455	46.364	46	41.818	40.909	40.909	42.727	43.636	46.364	47.273	43.636	43.636	44.545	43.636	44.545	43.636	
Mr CYP9DN1	31.818	30	40	42.727	49.091	43.636	43.636	47.273	46.364	42.727	43.636	65.455	45.946	41.818	51.818	51.818	54.545	54.545	54.545	50.909	57	51.818	44.545	50	47.273	43.636	48.182	50.909	47.273	49.091	49.091	49.091	49.091		
Ob CYP9DN1	33.636	29.091	44.545	40	43.636	40.909	42.727	44.545	42.727	39.091	40.909	65.455	46.847	43.636	48.182	48.182	50.909	50.909	48.182	55	46.364	42.727	48.182	45.455	43.636	46.364	50.909	44.545	44.545	47.273	46.364	46.364			
Am CYP9S1	32.432	32.432	37.838	38.739	36.937	42.342	39.64	44.144	39.64	44.144	39.64	36.036	45.946	46.847	70.909	44.144	43.243	46.847	45.946	45.045	44.144	47.525	44.144	38.739	48.182	49.091	56.364	52.727	55.455	52.727	57.273	57.273	57.273		
Bt CYP9S1	35.455	36.364	40	37.273	31.818	39.091	36.364	36.364	35.455	36.364	32.727	41.818	43.636	70.909	48.182	48.182	50	50.909	48.182	47.273	47	49.091	37.615	41.818	41.284	46.789	53.211	47.706	48.624	49.541	51.376	50.459	50.459		
Mr CYP9P2	39.091	40	36.364	44.545	42.727	41.818	49.091	43.636	45.455	50	44.545	51.818	48.182	44.144	48.182	97.273	87.273	90	71.818	65.455	72	71.818	57.273	61.818	49.091	55.455	58.182	58.182	58.182	59.091	59.091	57.273	57.273		
Ob CYP9P2	39.091	40	37.273	44.545	41.818	40.909	50	43.636	45.455	49.091	44.545	51.818	48.182	43.243	48.182	97.273	89.091	91.818	71.818	66.364	74	72.727	58.182	62.727	49.091	56.364	57.273	58.182	59.091	60	58.182	56.364	56.364		
Am CYP9P2	40	39.091	37.273	46.364	42.727	43.636	50	44.545	46.364	50	45.455	54.545	50.909	46.847	50	87.273	89.091	97.273	74.545	70.909	78	74.545	56.364	61.818	48.182	54.545	55.455	57.273	55.455	56.364	59.091	56.364	56.364		
Bt CYP9P2	40.909	40	37.273	45.455	43.636	42.727	50	43.636	45.455	50	44.545	54.545	50.909	45.946	50.909	90	91.818	97.273	74.545	70.909	78	74.545	56.364	62.727	49.091	54.545	56.364	57.273	56.364	57.273	60	57.273	57.273		
Am CYP9P1	39.091	37.273	39.091	43.636	43.636	48.182	48.182	43.636	46.364	49.091	46.364	54.545	48.182	45.045	48.182	71.818	71.818	74.545	74.545	82.727	79	77.273	57.273	60	50	51.818	54.545	51.818	50.909	51.818	55.455	54.545	54.545		
Bt CYP9P1	37.273	36.364	39.091	43.636	41.818	49.091	45.455	40.909	41.818	45.455	46.364	50.909	50.909	44.144	47.273	65.455	66.364	70.909	70.909	82.727	77	72.727	56.364	56.364	50	49.091	50.909	50.909	49.091	50	50.909	49.091	49.091		
Ob CYP9P2	43	40	42	46	48	46	47	43	45	46	46	57	55	47.525	47	72	74	78	78	79	77	82	82	59	62	53	53	53	54	55	54	54			
Mr CYP9R2	40	39.091	40	44.545	44.545	46.364	44.545	47.273	43.636	41.818	51.818	46.364	44.144	40.909	41.818	72.727	74.545	74.545	77.273	72.727	82	72.727	55.455	60	46.364	50	56.364	51.818	54.545	55.455	55.455	55.455			
Mr CYP9R23	29.091	31.818	37.273	41.818	40	38.182	41.818	39.091	37.273	38.182	40.909	44.545	42.727	38.182	37.615	57.273	58.182	56.364	56.364	57.273	56.364	59	55.455	43.119	47.706	48.624	44.037	45.872	44.954	45.872	44.037	44.037	44.037		
Ob CYP9R24	36.364	35.455	38.182	42.727	39.091	43.636	45.455	41.818	42.727	42.727	40.909	50	48.182	38.739	41.818	61.818	61.818	62.727	61.818	62.727	60	56.364	62	60	63.636	45.455	48.182	46.364	46.364	49.091	50	49.091	47.273		
Ob CYP9R59	37.273	37.273	36.364	40.909	39.091	40.909	41.818	40.909	41.818	42.727	47.273	45.455	48.182	41.284	49.091	49.091	48.182	49.091	50	50	53	46.364	43.119	45.455	60.55	58.716	60.55	58.716	58.716	60.55	61.468	60.55	61.468		
Ob CYP9R39	37.273	37.273	38.182	42.727	43.636	36.364	42.727	43.636	44.545	44.545	43.636	43.636	43.636	43.636	49.091	46.789	55.455	56.364	54.545	51.818	49.091	53	50	47.706	48.182	60.55	58.716	62.385	61.468	64.22	65.138	64.22	65.138		
Am CYP9R1	40	39.091	34.545	42.727	43.636	44.545	45.455	47.273	47.273	49.091	46.364	48.182	46.364	53.211	58.182	57.273	55.455	56.364	54.545	50.909	53	56.364	48.624	46.364	58.716	65.138	75.229	69.725	74.312	74.312	74.312	74.312			
Bt CYP9R1	40	42.727	34.545	40.909	45.455	41.818	48.182	50.909	50	50.909	47.273	50.909	50.909	52.727	47.706	58.182	58.182	57.273	57.273	51.818	53	51.818	44.037	46.364	60.55	63.303	75.229	69.725	74.312	79.817	76.147	76.147			
Mr CYP9R1	43.636	43.636	36.364	43.636	43.636	40	48.182	50.909	50.909	44.545	43.636	47.273	44.545	55.455	48.624	58.182	59.091	55.455	56.364	50.909	49.091	54	54.545	45.872	49.09										

Appendix table 3.7: Distribution of CYP9 lineages across 12 species of bee.

Figures in brackets indicate additional sequences curated manually in this PhD.

Species	CYP9R lineage	CYP9S lineage	CYP9Q lineage	CYP9DN lineage	CYP9P lineage	CYP9DM lineage	Total
<i>A. cerana</i>	1	1	3	0	2	0	7
<i>A. dorsata</i>	1	1	3	0	1	0	6
<i>A. florea</i>	1	1	1	0	0 (2)	0	3 (5)
<i>A. mellifera</i>	1	1	3	0	2	0	7
<i>B. impatiens</i>	1	0	3	0	1	0	5
<i>B. terrestris</i>	1	1	3	0	2	0	7
<i>D. novaeangliae</i>	2	1	4 DLs	0	0 (2)	0	7(9)
<i>E. mexicana</i>	3	0	2	1	2	0	8
<i>H. laboriosa</i>	2	1	1	1	0 (3)	0	5(8)
<i>M. quadrifasciata</i>	4	1	2	1	1	0	9
<i>M. rotundata</i>	3	0	0	1	3	2	9
<i>O. bicornis</i>	3	0	2 BUs	1	3	0	9



Appendix figure 3.1: Phylogeny of CYP9 amino acid sequences from 12 bee species (*A. mellifera*; *A. cerana*; *A. dorsata*; *A. florea*; *B. terrestris*; *B. impatiens*; *D. novaeangliae*; *E. mexicana*; *H. laboriosa*; *M. quadrifasciata*; *M. rotundata* and *O. bicornis*). Phylogeny estimated using PhyML Maximum likelihood algorithm and substitution model LG+G, with branch support of 50 bootstraps, shown as %, rooted on camphor hydroxylase (P450cam; *P. putida*). Candidate CYP9P sequences from *A. florea* (2), *D. novaeangliae* (2) and *H. laboriosa* (3), are shown in green.

Appendix chapter four

Appendix table 4.1: Response of *M. rotundata* to topically applied neonicotinoid insecticides at 24h, 48h, 72h and 96h. [95% confidence intervals calculated using a heterogeneity factor (the chi-square value divided by the df) as the p value for the Pearson goodness-of fit chi-square test was significant (p<0.150)].

*results calculated without converting doses to logarithms.

Compound tested	LD ₅₀ 24h (µg ai/bee)	Lower CI 95%	Upper CI 95%	Hill Slope	SE ±
Thiacloprid	0.024	0.018	0.039	1.963	0.271
Imidacloprid	0.003	0.002	0.004	423.319*	82.168
Acetamiprid	0.183	0.110	0.410	1.471	0.214
Compound tested	LD ₅₀ 48h (µg ai/bee)	Lower CI 95%	Upper CI 95%	Hill slope	SE ±
Thiacloprid	0.013	0.012	0.017	3.997	0.585
Imidacloprid	0.001	0.001	0.002	915.807*	206.532
Acetamiprid	0.179	0.115	0.351	1.523	0.219
Compound tested	LD ₅₀ 72h (µg ai/bee)	Lower CI 95%	Upper CI 95%	Hill slope	SE ±
Thiacloprid	0.010	0.008	0.012	3.816	0.715
Imidacloprid	0.001	0.001	0.002	827.567*	215.286
Acetamiprid	0.224	0.134	0.315	6.965*	1.272
Compound tested	LD ₅₀ 96h (µg ai/bee)	Lower CI 95%	Upper CI 95%	Hill slope	SE ±
Thiacloprid	0.003	-0.003	0.006	86.489*	14.522
Imidacloprid	0.001	0.001	0.002	827.567*	215.286
Acetamiprid	0.181	0.112	0.258	9.031*	2.027

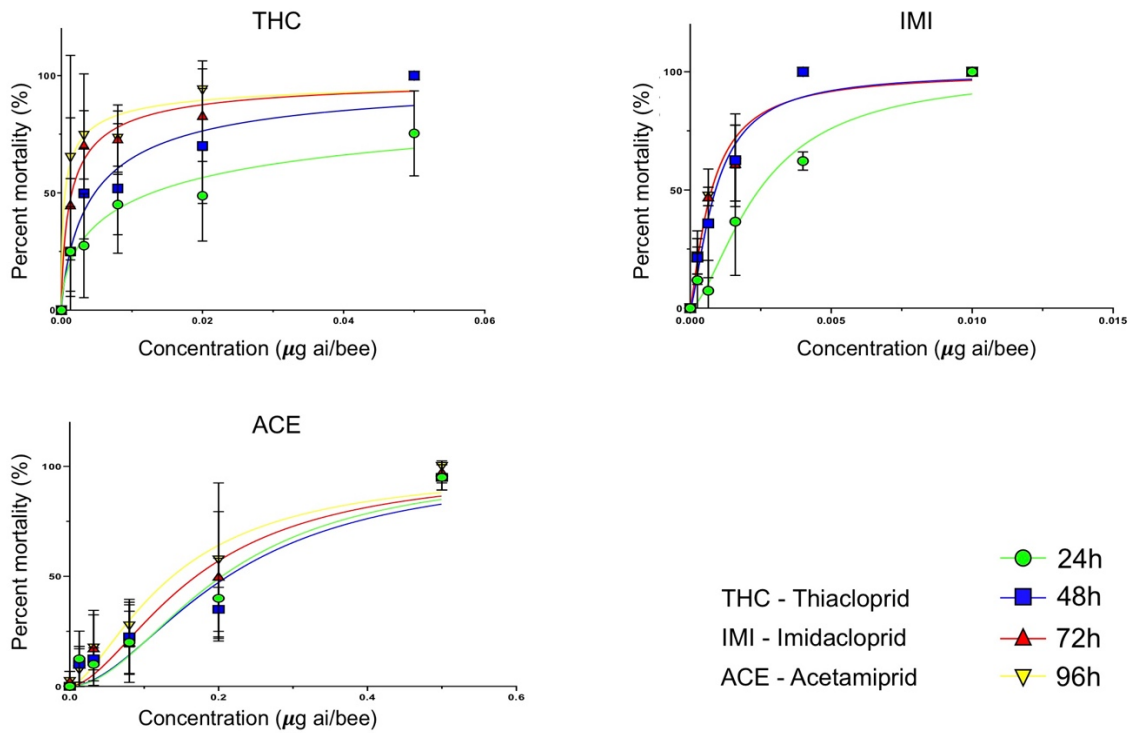
Appendix able 4.2: Response of *M. rotundata* to topically applied pyrethroid insecticides at 24h, 48h, 72h and 96h. [95% confidence intervals calculated using a heterogeneity factor (the chi-square value divided by the df) as the p value for the Pearson goodness-of fit chi-square test was significant (p<0.150)].

*results calculated without converting doses to logarithms.

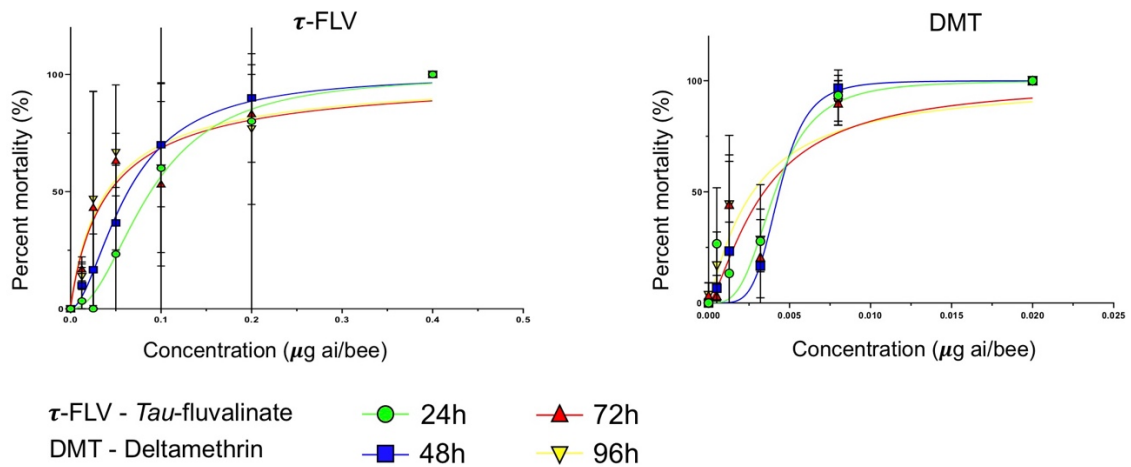
Compound tested	LD ₅₀ 24h (µg ai/bee)	Lower CI 95%	Upper CI 95%	Hill slope	SE ±
<i>tau</i> -fluvalinate	0.099	0.060	0.270	2.517	0.407
Deltamethrin	0.004	0.003	0.007	284.076*	48.360
Compound tested	LD ₅₀ 48h (µg ai/bee)	Lower CI 95%	Upper CI 95%	Hill slope	SE ±
<i>tau</i> -fluvalinate	0.061	0.044	0.092	2.208	0.339
Deltamethrin	0.004	0.003	0.008	359.620*	2.9930
Compound tested	LD ₅₀ 72h (µg ai/bee)	Lower CI 95%	Upper CI 95%	Hill slope	SE ±
<i>tau</i> -fluvalinate	0.040	0.015	0.082	1.4163	0.273
Deltamethrin	0.004	0.003	0.007	273.662*	0.4024
Compound tested	LD ₅₀ 96h (µg ai/bee)	Lower CI 95%	Upper CI 95%	Hill slope	SE ±
<i>tau</i> -fluvalinate	0.037	0.014	0.070	1.632	0.281
Deltamethrin	0.005	0.003	0.008	251.340*	0.3305

Appendix table 4.3: Response of *M. rotundata* to topically applied organophosphate insecticides at 24h, 48h, 72h and 96h. [95% confidence intervals calculated using a heterogeneity factor (the chi-square value divided by the df) as the p value for the Pearson goodness-of fit chi-square test was significant (p<0.150)]

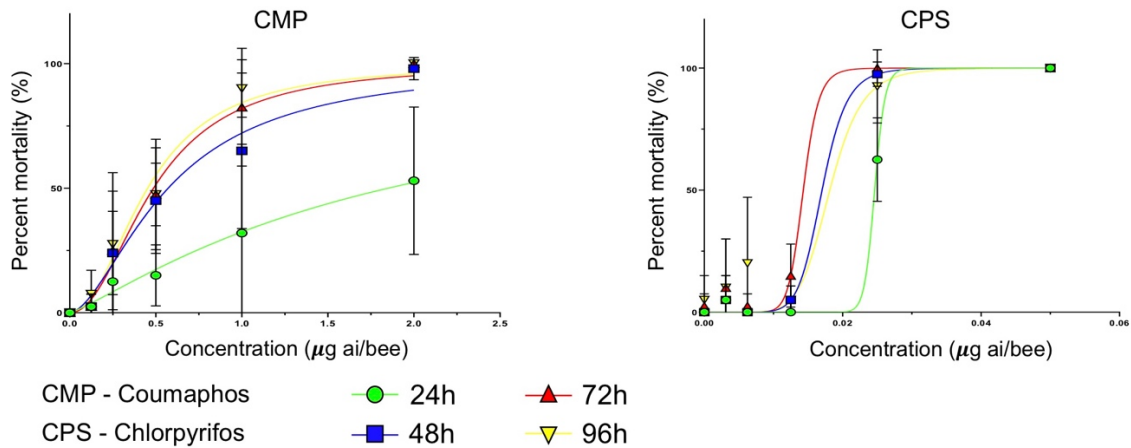
Compound tested	LD ₅₀ 24h (µg ai/bee)	Lower CI 95%	Upper CI 95%	Hill slope	SE ±
Coumaphos	1.827	1.108	6.085	1.573	0.264
Chlorpyrifos	0.022	0.018	0.029	6.777	0.853
Compound tested	LD ₅₀ 48h (µg ai/bee)	Lower CI 95%	Upper CI 95%	Hill slope	SE ±
Coumaphos	0.557	0.433	0.720	2.700	0.298
Chlorpyrifos	0.017	0.015	0.020	10.277	1.274
Compound tested	LD ₅₀ 72h (µg ai/bee)	Lower CI 95%	Upper CI 95%	Hill slope	SE ±
Coumaphos	0.471	0.368	0.601	3.183	0.367
Chlorpyrifos	0.015	0.013	0.020	10.070	1.582
Compound tested	LD ₅₀ 96h (µg ai/bee)	Lower CI 95%	Upper CI 95%	Hill slope	SE ±
Coumaphos	0.419	0.324	0.542	3.063	0.353
Chlorpyrifos	0.016	0.013	0.028	5.760	0.839



Appendix figure 4.1: Acute contact toxicity dose-response curves for one nitro-group, imidacloprid (IMI) and two cyano-group neonicotinoid insecticides, thiachloprid (THC) and cacetamiprid (ACE) against *M. rotundata* at each observation time (24h, 48h, 72h and 96h).



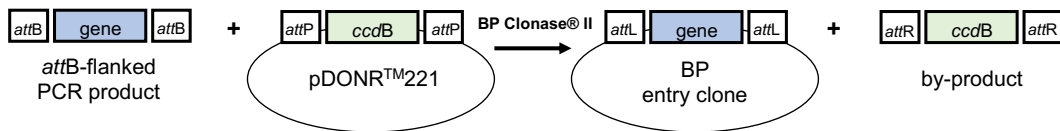
Appendix figure 4.2: Acute contact toxicity dose-response curves for two pyrethroid insecticides (τ -fluvalinate (τ -FLV); deltamethrin (DMT)) against *M. rotundata* at each observation time (24h, 48h, 72h and 96h).



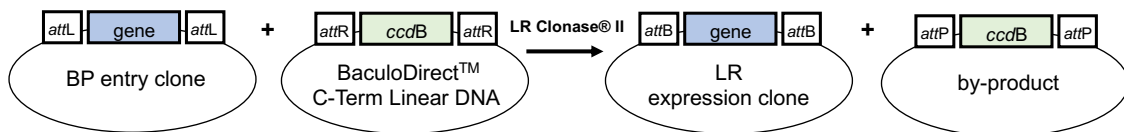
Appendix figure 4.3: Acute contact toxicity dose-response curves for two organophosphate insecticides, coumaphos (CMP) and chlorpyrifos (CPS) against *M. rotundata* at each observation time (24h, 48h, 72h and 96h).

Appendix chapter five

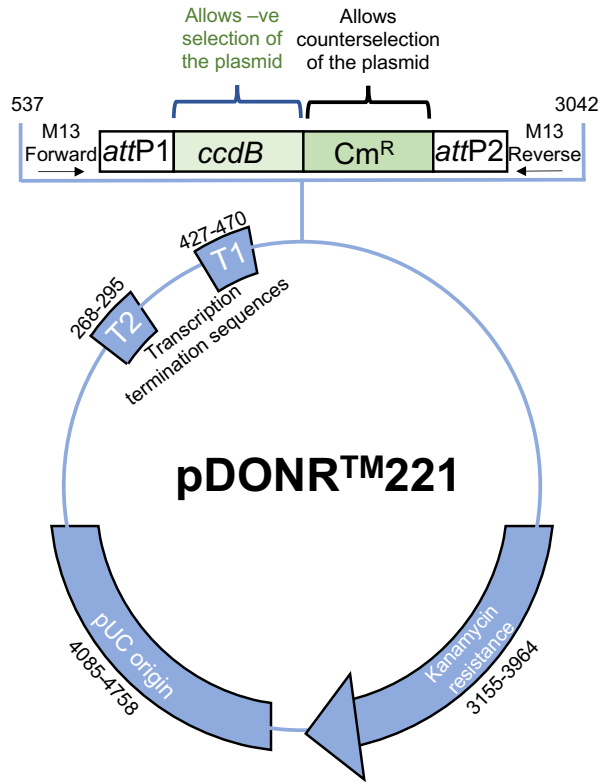
A) BP Recombination reaction



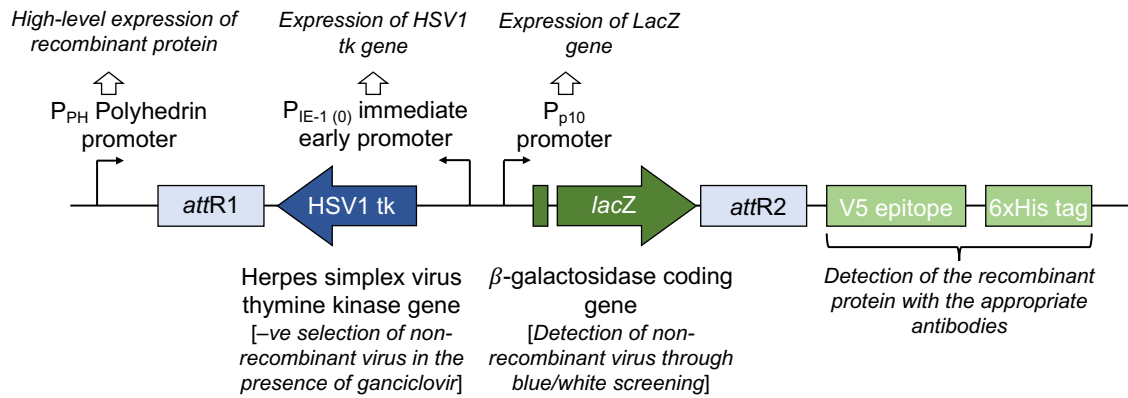
B) LR Recombination reaction



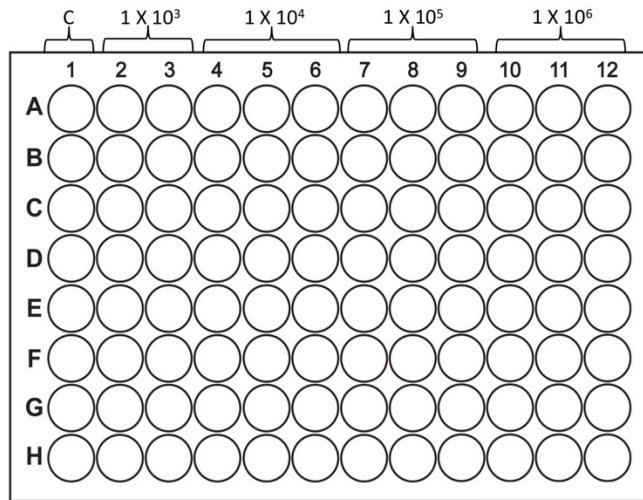
Appendix figure 5.1: The two recombination reactions. A) BP recombination reaction (lysogenic pathway: $attB \times attP \rightarrow attL \times attR$); B) LR recombination reaction (lytic pathway: $attL \times attR \rightarrow attB \times attP$)



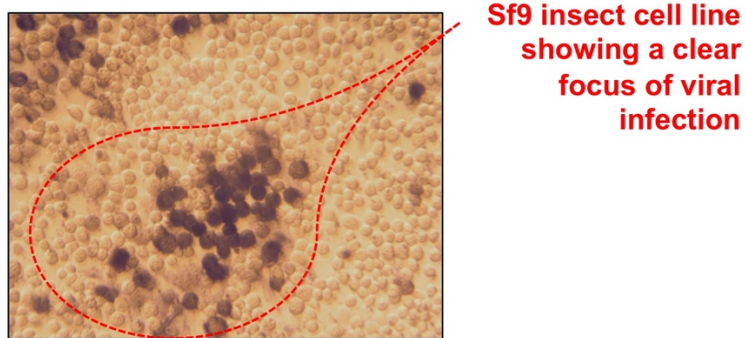
Appendix figure 5.2: A map of the attP-containing-donor vector pDONR221



Appendix figure 5.3: Map of the Gateway® cassette elements of the BaculoDirect™ C-Term linear DNA



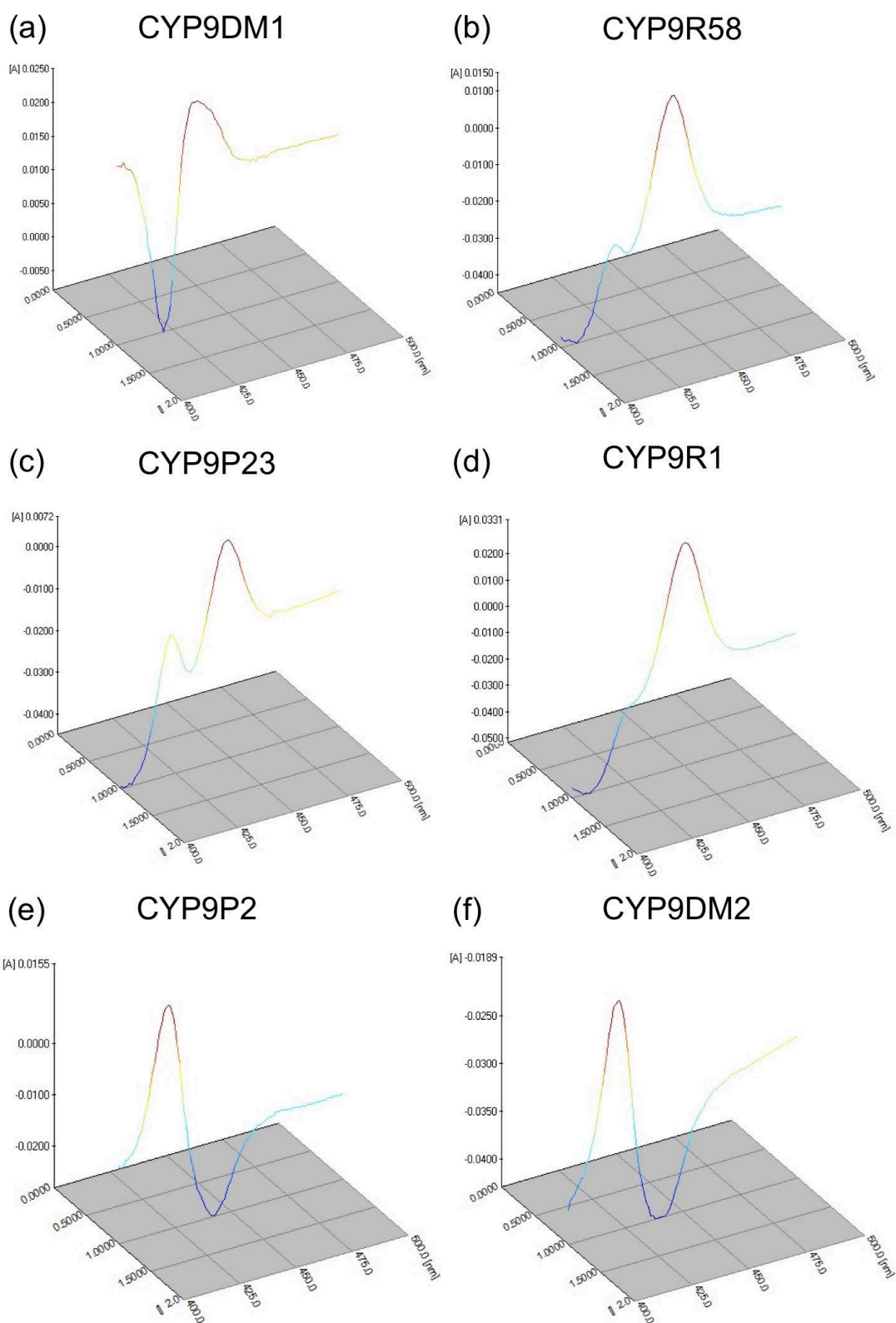
Appendix figure 5.4: 96-well plate set up for quantification of viral titre



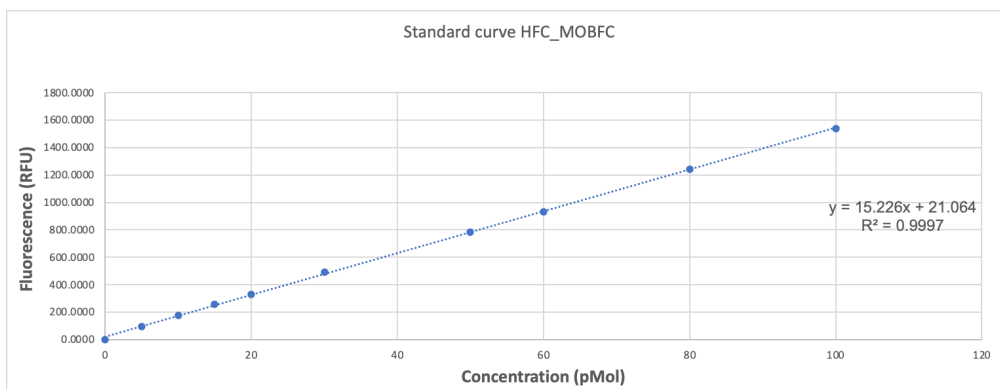
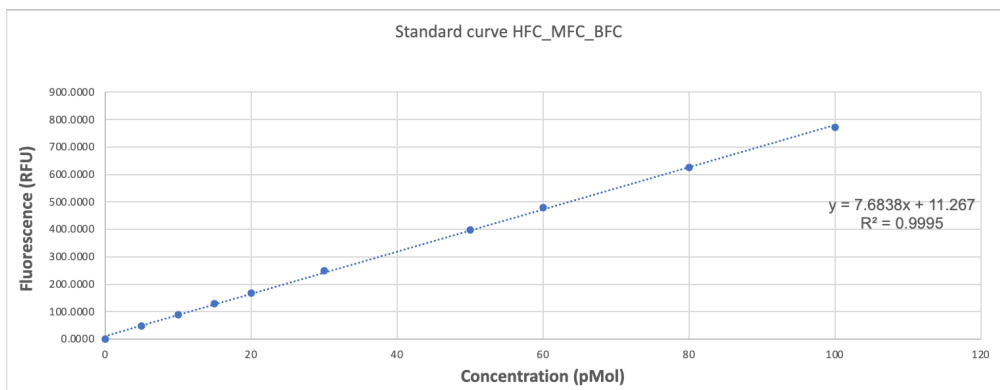
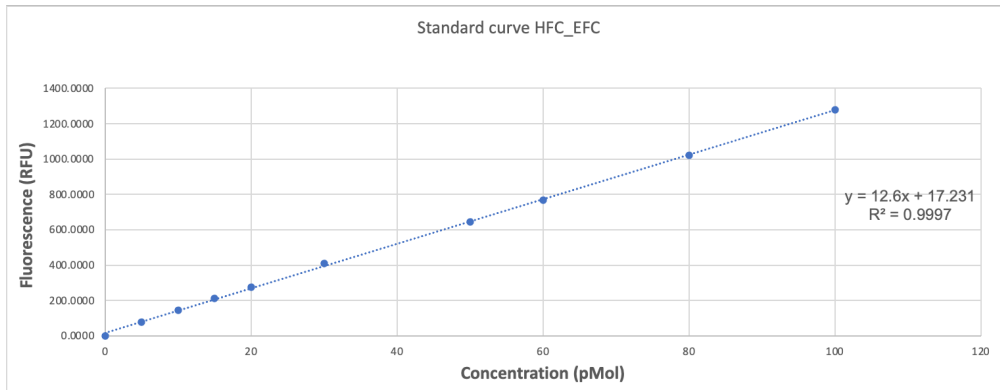
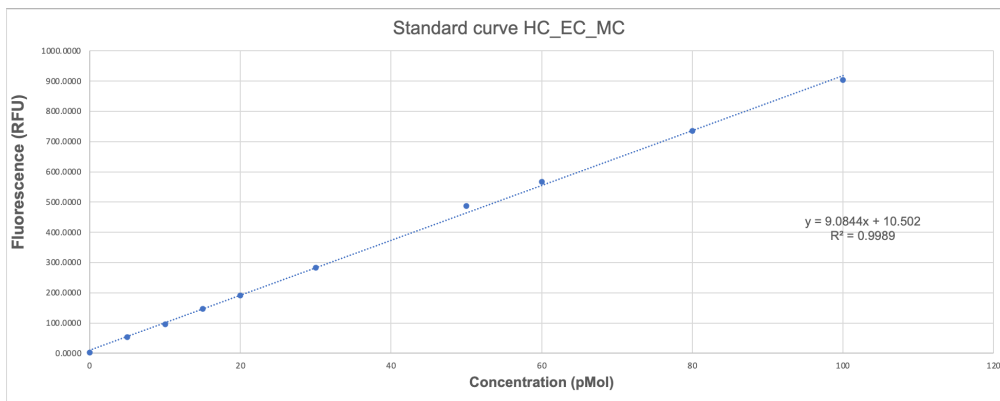
Appendix figure 5.5: A viral focus in Sf9 insect cells produced in an ELISA to quantify viral titre.



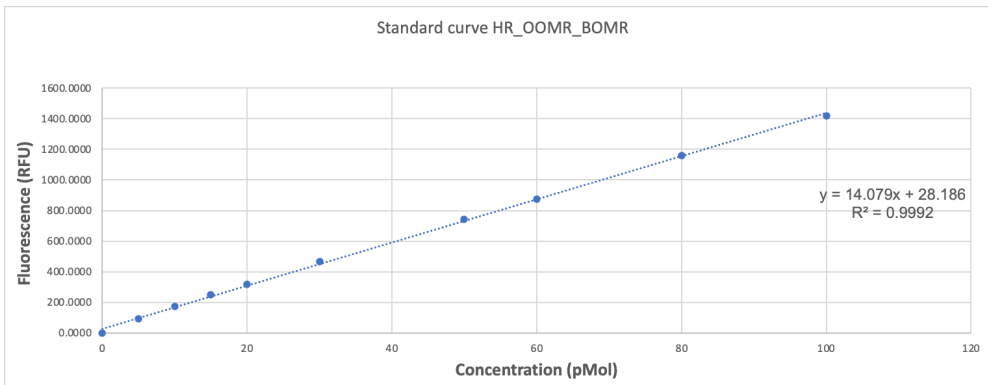
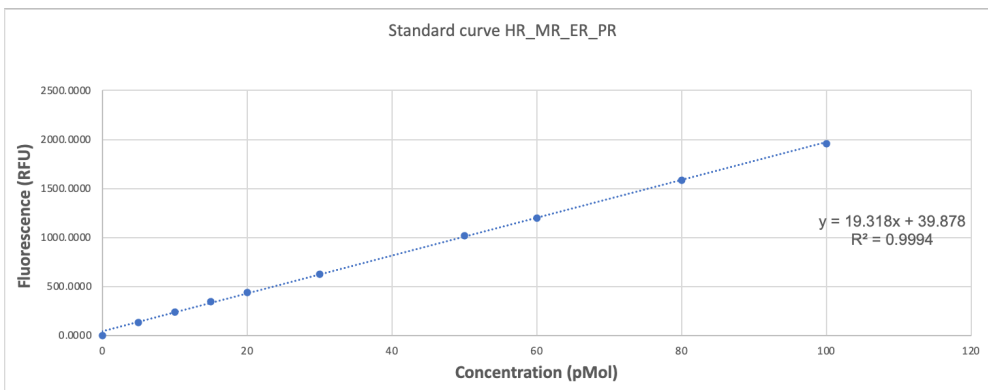
Appendix figure 5.6: 1% Agarose gel of 2-step 12attB PCR using Phusion® high-fidelity DNA polymerase (NEB) of seven CYP9 family P450s: CYP9DN1, CYP9P22, CYP9P23, CYP9R59, CYP9DM2, CYP9R1 and CYP9P2. Measured against GeneRuler™ 1 kb Plus DNA ladder (Thermo Fisher).



Appendix figure 5.7: CO-difference spectrum of recombinant CYP9 proteins: (a) CYP9DM1, (b) CYP9R58, (c) CYP9P23, (d) CYP9R1, (e) CYP9P2, (f) CYP9DM2. The P450s were reduced with sodium dithionite ($\text{Na}_2\text{S}_2\text{O}_4$).



Appendix figure 5.8: Model substrate standard curves – coumarins



Appendix figure 5.9: Model substrate standard curves – resorufins

Appendix table 5.1: Reagents used in the quantification of viral titre by ELISA

REAGENT	COMPONENTS
Formyl buffered acetone (Place at -20°C for at least 3 hours before use)	5.4 ml PBS (Sigma) 8.1 ml 100% Acetone (Sigma) 4.5 ml 37% Formaldehyde solution (Sigma)
PBS + 0.05% Tween 20	500 ml PBS 250 µl Tween 20 (Sigma)
Diluted Normal Goat Serum	480 µl Normal Goat Serum (Sigma) 13 ml PBS + 0.05% Tween 20
Mouse gp64 Antibody (1:666)	6 µl gp64 Antibody (AbCam) 3994 µl Diluted Normal Goat Serum
Goat Anti-mouse Antibody/HRP Conjugate (1:2000)	2.5 µl Anti-mouse Conjugate (AbCam) 4990 µl Diluted Normal Goat Serum

Appendix table 5.2: Kinetics data for MC metabolism (resulting in HC) by CYP9BU1 and CYP9DM2 co-incubated with imidacloprid at two different concentrations

MC + IMI	Conc IMI (µM)	K _M (µM)	95% CI	V _{max} (pmol/min ⁻¹ mg protein ⁻¹)	95% CI	Adjusted R ²
CYP9BU1	100	144.9	87.64-271.3	10.40	8.011-15.11	0.9255
	10	148.4	93.94-259.0	10.56	8.308-14.7	0.9404
	0	200.4	123.0-378.4	13.82	10.47-20.97	0.9493
CYP9DM2	100	200.7	114.7-429.8	15.80	11.49-26.27	0.9223
	10	307.0	137.1-1645	22.39	13.56-86.54	0.8874
	0	245.9	142.0-534.5	19.00	13.63-32.86	0.9361

Appendix table 5.3: LC-MS/MS results showing metabolite (OH) production after 1 h incubation with *M. rotundata* native microsomes (results shown are for the 160ug microsomal protein used in the incubation)

Sample description (max. 30 characters)	parent area ng/ml	IM-2-1 ng/ml	OH-Imidacloprid ng/ml	OH-Thiacloprid ng/ml
10 µM Acetamiprid M Mr mcs rep 1	350.0	22.84		
10 µM Acetamiprid M Mr mcs rep 2	368.4	24.04		
10 µM Acetamiprid M Mr mcs rep 3	345.7	20.45		
10 µM Acetamiprid M Mr mcs NADPH-	368.0	<LOQ		
10 µM Imidacloprid M Mr mcs rep 1	274.5		9.95	
10 µM Imidacloprid M Mr mcs rep 2	286.8		9.00	
10 µM Imidacloprid M Mr mcs rep 3	266.7		8.37	
10 µM Imidacloprid M Mr mcs NADPH-	308.2		<LOQ	
10 µM Thiacloprid M Mr mcs rep 1	336.0			6.42
10 µM Thiacloprid M Mr mcs rep 2	343.1			6.85
10 µM Thiacloprid M Mr mcs rep 3	360.6			6.97
10 µM Thiacloprid M Mr mcs NADPH-	380.5			<LOQ
Sample description (max. 30 characters)	parent area ng/ml	SES11154-2-3 ng/ml	VRF3162-3-1 ng/ml	
10 µM Flupyradifurone rep 1	491	<LOQ	0.16	
10 µM Flupyradifurone rep 2	484	<LOQ	0.18	
10 µM Flupyradifurone rep 3	544	<LOQ	0.19	
10 µM Flupyradifurone NADPH- rep 1	507	<LOQ	<LOQ	
10 µM Flupyradifurone NADPH- rep 2	457	<LOQ	<LOQ	
10 µM Flupyradifurone NADPH- rep 3	482	<LOQ	<LOQ	
Sample description (max. 30 characters)	parent area ng/ml	OH-Deltamethrin ng/ml		
10 µM Deltamethrin rep 1	318	0.33		
10 µM Deltamethrin rep 2	288	0.31		
10 µM Deltamethrin rep 3	280	0.31		
10 µM Deltamethrin NADPH-	295	<LOQ		
Sample description (max. 30 characters)	parent area ng/ml	Cotinine ng/ml		
10uM Nicotine NADPH+ rep. 1	133	6.25		
10uM nicotine NADPH+ rep. 2	143	6.26		
10uM Nicotine NADPH+ rep. 3	133	6.72		
10uM Nicotine NADPH- rep. 1	226	1.31		
10uM Nicotine NADPH- rep. 2	231	1.29		
10uM Nicotine NADPH- rep. 3	223	1.31		

Appendix table 5.4: Concentration of diamide insecticides used in range finding tests (only one repetition of test per compound performed)

Compound	Concentration µg/ µl (µg/bee)
Flubendiamide	50, 12.5, 6.250, 3.125, 1.5625
Chlorantraniliprole	10, 5, 1, 0.5

Appendix chapter six

Appendix table 6.1: Hit table of Blastn results – query sequence *A. mellifera* CYP9Q3 through the NCBI TSA database gated by bees (taxid:34735)

Query acc ver	Subject acc ver	Subject name	% id	alignment length	mismatches	gap opens	Q start	Q end	S start	S end	e-value	bit score	
XM_006562300.3	JR040720.1	<i>A. cerana</i>	69.699	1462	424	12	162	1617	1731	1621	9.43E-169	603	
XM_006562300.3	GIH01010925.1	<i>H. illama</i>	69.392	1529	441	19	85	1604	1311	1701	7.25E-164	587	
XM_006562300.3	GBTLO1078920.1	<i>T. carbonaria</i>	69.17	1531	438	19	85	1604	192	1699	3.76E-161	579	
XM_006562300.3	GBTLO1078919.1	<i>T. carbonaria</i>	68.867	1545	448	18	85	1615	169	1694	6.79E-158	568	
XM_006562300.3	GBTLO1077204.1	<i>E. dilemma</i>	69.03	1495	430	20	125	1604	107	1593	1.82E-152	549	
XM_006562300.3	GBMPO101343.1	<i>X. violacea</i>	67.716	1493	437	25	85	1559	1722	307	4.95E-122	448	
XM_006562300.3	GBUM01016761.1	<i>X. violacea</i>	67.229	1483	431	22	147	1606	460	3510	1.01E-117	434	
XM_006562300.3	GBMPO101342.1	<i>E. dilemma</i>	67.425	1495	438	25	85	1559	1722	307	3.52E-117	433	
XM_006562300.3	HP64608.1	<i>E. cordata</i>	66.86	1551	467	23	85	1616	81	1603	2.25E-113	419	
XM_006562300.3	GBMPO101344.1	<i>E. dilemma</i>	66.406	1533	476	22	85	1602	1804	296	1.95E-101	379	
XM_006562300.3	HP60169.1	<i>F. varia</i>	70.855	795	212	10	439	1220	17	788	6.36E-99	358	
XM_006562300.3	GBMPO100059.1	<i>B. rufipes</i>	66.475	1390	410	20	145	1522	1787	422	3.30E-92	349	
XM_006562300.3	HP64744.1	<i>A. florea</i>	68.058	958	284	8	162	1113	229	1170	2.54E-87	333	
XM_006562300.3	GBM0381.1	<i>B. terrestris</i>	64.375	1527	524	8	85	1604	18	1531	1.15E-72	284	
XM_006562300.3	GHHB0100084.1	<i>B. terrestris</i>	64.375	1527	524	8	85	1604	239	878	1.15E-72	284	
XM_006562300.3	HP66008.1	<i>Enigta</i>	67.4	954	288	13	328	1272	14	951	1.41E-71	281	
XM_006562300.3	GBUM01016763.1	<i>X. violacea</i>	66.73	1049	312	16	147	1181	1028	5	1.41E-71	280	
XM_006562300.3	GIH01051385.1	<i>H. illama</i>	72.453	530	134	7	1081	1604	1471	948	1.71E-70	278	
XM_006562300.3	LS42865.1	<i>B. terrestris</i>	64.244	1262	526	9	85	1604	2555	1042	2.54E-68	270	
XM_006562300.3	GHPFO10101076.1	<i>B. pascuorum</i>	64.822	1319	449	9	273	1490	273	1427	1.98E-66	266	
XM_006562300.3	GBMPO1021203.1	<i>D. cincta</i>	65.476	1280	387	22	284	1523	277	1508	5.60E-64	256	
XM_006562300.3	GBMPO1021201.1	<i>D. cincta</i>	65.476	1280	387	22	284	1523	277	1508	5.60E-64	256	
XM_006562300.3	GHHDD1008977.1	<i>B. cryptarum</i>	65.527	1132	359	15	439	1542	1530	419	5.60E-64	255	
XM_006562300.3	L71244.1	<i>B. terrestris</i>	65.813	1057	309	10	85	1604	5365	3845	5.60E-64	255	
XM_006562300.3	L71244.1	<i>B. terrestris</i>	65.083	478	164	1	85	559	85	562	1.83E-19	108	
XM_006562300.3	GBMPO1010335.1	<i>Ceratina chalybea</i>	65.717	1028	348	11	289	1368	2255	155	1.95E-63	254	
XM_006562300.3	GBTMO1027256.1	<i>Anthophora plumipes</i>	65.247	1338	441	24	289	1608	601	1910	1.95E-63	254	
XM_006562300.3	JD44861.1	<i>B. terrestris</i>	64.046	1488	518	8	121	1604	96	1570	6.82E-63	251	
XM_006562300.3	L779752.1	<i>B. terrestris</i>	63.916	1527	528	9	85	1604	100	1610	2.38E-62	251	
XM_006562300.3	GBTMO1027287.1	<i>Anthophora plumipes</i>	65.175	1243	411	11	691	1604	691	1910	6.82E-61	247	
XM_006562300.3	GBUM01016762.1	<i>X. violacea</i>	70.175	570	162	5	1041	1606	796	231	2.90E-61	247	
XM_006562300.3	GBUM01016762.1	<i>X. violacea</i>	69.136	597	175	0	147	713	1656	1090	1.83E-57	234	
XM_006562300.3	GBMPO1010931.1	<i>B. terrestris</i>	64.419	1427	439	48	439	1599	48	1529	4.9E-59	240	
XM_006562300.3	GHHDD000144.1	<i>B. cryptarum</i>	65.271	1146	372	12	466	1604	1232	103	4.30E-59	240	
XM_006562300.3	J122838.1	<i>B. impatiens</i>	65.161	1151	371	14	466	1604	2322	1230	6.39E-57	233	
XM_006562300.3	GBMPO103059.1	<i>B. rufipes</i>	65.025	1144	378	11	466	1604	2064	838	2.23E-56	231	
XM_006562300.3	GBMPO103058.1	<i>B. rufipes</i>	64.948	1144	382	10	466	1604	2067	938	7.78E-56	228	
XM_006562300.3	GBMPO103057.1	<i>B. rufipes</i>	65.183	1146	373	11	466	1604	2067	938	7.78E-56	228	
XM_006562300.3	GBMPO101897.1	<i>M. haemorrhoidalis</i>	64.407	1023	349	11	162	1604	206	142	1572	3.31E-54	223
XM_006562300.3	J122490.1	<i>B. impatiens</i>	65.863	897	297	5	289	1184	444	1330	1.41E-52	217	
XM_006562300.3	HP68185.1	<i>F. varia</i>	71.234	425	122	0	55	479	0	416	52E-52	217	
XM_006562300.3	J121292.1	<i>B. impatiens</i>	67.589	629	201	1	85	710	1364	736	1.71E-51	214	
XM_006562300.3	J121292.1	<i>B. impatiens</i>	64.205	362	117	6	955	1301	501	154	0.15	49.1	
XM_006562300.3	J121624.1	<i>C. flavifrons</i>	63.719	439	145	0	466	1096	948	24	2.98E-51	211	
XM_006562300.3	GBMPO1017182.1	<i>H. illama</i>	64.508	1220	395	7	270	1473	1993	796	2.54E-49	206	
XM_006562300.3	GBMPO1017183.1	<i>Halictus quadricinctus</i>	64.508	1220	395	17	270	1473	2515	1318	2.54E-49	206	
XM_006562300.3	GBTMO1027232.1	<i>Anthophora plumipes</i>	64.43	1243	417	29	289	1604	891	1904	4.94E-49	206	
XM_006562300.3	GBTMO1027243.1	<i>Anthophora plumipes</i>	64.456	1242	414	27	288	1604	600	1903	2.54E-49	206	
XM_006562300.3	GHPFO1026338.1	<i>B. pycnosoma</i>	65.43	1024	326	16	586	1601	1381	378	8.88E-49	205	
XM_006562300.3	HP64634.1	<i>E. nigrita</i>	74.363	604	81	0	401	691	0	536	1.05E-48	202	
XM_006562300.3	GHPFO1026330.1	<i>B. terrestris</i>	67.091	629	204	1	85	710	1025	397	3.78E-47	200	
XM_006562300.3	L67496.1	<i>B. terrestris</i>	76.786	58	13	0	955	1536	0	64	4.4	20.7	
XM_006562300.3	GBMPO1019334.1	<i>M. luvipes</i>	65.24	981	313	15	629	1601	1146	186	1.96E-44	194	
XM_006562300.3	HP69154.1	<i>A. florea</i>	63.901	1338	441	19	168	1493	145	1452	6.83E-44	188	
XM_006562300.3	L675174.1	<i>B. terrestris</i>	68.111	505	149	5	145	645	214	2614	6.83E-44	188	
XM_006562300.3	HP69129.1	<i>A. florea</i>	69.93	429	124	2	166	591	188	614	6.83E-44	188	
XM_006562300.3	GBMPO1033076.1	<i>T. nigropus</i>	64.327	1025	300	19	601	1604	1080	860	3.88E-43	187	
XM_006562300.3	HP623416.1	<i>E. nigrita</i>	73.702	289	76	0	1328	1616	21	309	3.54E-41	179	
XM_006562300.3	L715382.1	<i>A. florea</i>	70.888	357	102	2	289	644	370	15	6.99E-38	168	
XM_006562300.3	GBMPO1015837.1	<i>N. diverripes</i>	68.191	289	146	0	289	496	496	884	4.23E-37	167	
XM_006562300.3	GBMPO1015837.1	<i>N. diverripes</i>	68.103	116	37	0	1060	1175	1165	1280	6.4	43.7	
XM_006562300.3	GHPFO1013522.1	<i>B. terrestris</i>	64.028	1162	372	19	454	1601	1219	90	2.23E-37	168	
XM_006562300.3	GHPFO1013521.1	<i>B. pascuorum</i>	65.99	100	31	0	1060	1175	1165	1280	6.4	43.7	
XM_006562300.3	GBLVO1001364.1	<i>P. tumigera</i>	64.015	1233	396	23	161	1374	231	1435	7.79E-37	166	
XM_006562300.3	GBMPO1008832.1	<i>Ceratina chalybea</i>	63.338	1026	334	11	289	1301	678	681	1.79E-37	167	
XM_006562300.3	GBLRO1017457.1	<i>T. tabulae</i>	64.075	1223	396	23	165	1370	219	1415	9.49E-36	162	
XM_006562300.3	GBMPO1021142.1	<i>Halictus quadricinctus</i>	68.894	440	123	6	283	715	649	217	9.49E-36	162	
XM_006562300.3	GBTMO102721.1	<i>E. nigrita</i>	77.723	202	65	0	1060	1175	1165	1280	6.4	43.7	
XM_006562300.3	GBMPO102643.1	<i>T. malvae</i>	66.667	558	180	4	161	715	154	708	4.03E-34	156	
XM_006562300.3	GBMPO102643.1	<i>T. malvae</i>	68.376	234	70	4	1143	1374	1127	1358	1.08E-09	75.2	
XM_006562300.3	HP64570.1	<i>A. florea</i>	67.078	441	130	0	1082	1616	1	536	4.03E-34	156	
XM_006562300.3	J038510.1	<i>A. cerana</i>	68.202	456	141	4	271	724	509	962	1.41E-33	155	
XM_006562300.3	J038510.1	<i>A. cerana</i>	68.545	213	63	2	1338	1548	1567	1777	3.30E-10	77.9	
XM_006562300.3	J038511.1	<i>A. cerana</i>	68.202	456	141	4	271	724	509	962	1.41E-33	155	
XM_006562300.3	J038511.1	<i>A. cerana</i>	68.545	213	63	2	1338	1548	1567	1777	3.30E-10	77.9	
XM_006562300.3	GHPWO1022581.1	<i>A. cineraria</i>	63.428	1266	436	15	270	1526	1815	968	4.91E-33	153	
XM_006562300.3	GBLUO1045421.1	<i>Colletes curvicaulus</i>	67.652	287	147	0	270	496	496	1740	4.91E-33	153	
XM_006562300.3	GBLUO1045421.1	<i>Colletes curvicaulus</i>	64.532	406	132	6	962	1361	1495	1096	1.61E-07	68	
XM_006562300.3	GBLVO1024539.1	<i>A. vaga</i>	65.556	910									

Appendix table 6.1 (continued):

XM_00562300.3	GBQK01014348.1	<i>H truncorum</i>	75.776	161	39	0	500	660	2291	2131	1.24E-21	115
XM_00562300.3	GBQK01014348.1	<i>H truncorum</i>	65.975	482	156	8	1046	1523	1754	1277	7.80E-18	102
XM_00562300.3	GBQK01014348.1	<i>H truncorum</i>	66.443	149	50	0	176	326	203	2465	1.8	44.6
XM_00562300.3	GBUN01015814.1	<i>X violacea</i>	66.366	138	143	1	262	719	1937	1921	1.24E-21	115
XM_00562300.3	GBUN01015814.1	<i>X violacea</i>	68.584	226	63	4	958	1179	1693	1472	3.10E-10	77.9
XM_00562300.3	GHUO10204046.1	<i>A vega</i>	65.162	142	1	0	971	1522	2135	766	4.31E-21	113
XM_00562300.3	GHUO10204046.1	<i>A haemorrhoea</i>	66.357	431	142	3	289	718	1953	1525	4.31E-21	113
XM_00562300.3	GHUO10204046.1	<i>A haemorrhoea</i>	72.932	133	33	3	1423	1553	905	774	1.61E-07	68
XM_00562300.3	GHCJ01020383.1	<i>A haemorrhoea</i>	66.357	142	1	0	269	718	2781	2353	4.31E-21	113
XM_00562300.3	GHCJ01020383.1	<i>O cornuta</i>	72.932	133	33	3	1423	1553	1659	1528	1.61E-07	68
XM_00562300.3	HPR20966.1	<i>A. florea</i>	71.429	224	58	2	178	368	522	422	4.31E-21	113
XM_00562300.3	GHHV01051389.1	<i>H. flama</i>	67.684	393	111	9	624	1011	2	385	1.95E-20	111
XM_00562300.3	J118120.1	<i>D. crinita</i>	65.726	531	167	9	1081	1601	63	588	1.95E-20	111
XM_00562300.3	GBWV01023618.1	<i>S. punctulifalissima</i>	66.395	140	4	0	464	889			5.25E-20	109
XM_00562300.3	GBWV01023618.1	<i>S. punctulifalissima</i>	81.633	49	9	0	962	1010	1129	1177	0.15	49.1
XM_00562300.3	GHFO01016779.1	<i>B. lempstris</i>	67.446	364	117	6	1075	1454	395	16	5.25E-20	110
XM_00562300.3	HPR61422.1	<i>M. quadrifasciata</i>	67.213	166	116	3	126	488	132	485	5.25E-20	110
XM_00562300.3	GBDI01021291.1	<i>C. sacum</i>	65.898	428	144	2	288	714	3258	2832	1.63E-19	107
XM_00562300.3	GBUN01017886.1	<i>X. violacea</i>	66.478	249	76	0	158	406	171	419	1.83E-19	107
XM_00562300.3	GBVK01019373.1	<i>M. haemorrhoidalis</i>	63.607	915	305	15	270	1176	295	1189	1.83E-19	107
XM_00562300.3	GBVK01019373.1	<i>M. haemorrhoidalis</i>	68.182	154	43	4	1330	1480	1343	1493	1.8	45.5
XM_00562300.3	GHFPO1033432.1	<i>O. cornuta</i>	72.932	144	54	0	520	693	221	28	1.61E-19	107
XM_00562300.3	GMEB01021546.1	<i>P. dentipes</i>	66.516	442	141	7	281	719	584	1021	6.40E-19	106
XM_00562300.3	GBJ001016178.1	<i>A. maniculatus</i>	66.513	433	139	6	281	710	418	587	6.40E-19	106
XM_00562300.3	GBJ001016178.1	<i>Nomiodex sp.</i>	66.983	421	129	8	1069	1484	1333	1918	6.40E-19	106
XM_00562300.3	GHFV01000822.1	<i>B. pasorum</i>	67.29	321	105	0	289	609	341	21	6.40E-19	106
XM_00562300.3	GBPU01012575.1	<i>C. chalybea</i>	66.516	448	127	11	283	719	3222	2787	2.23E-18	104
XM_00562300.3	GBPU01012575.1	<i>C. chalybea</i>	73.043	115	29	2	1429	1542	2044	1981	2.39E-05	61.7
XM_00562300.3	GBPU01012575.1	<i>C. chalybea</i>	67.879	33	4	0	978	1010	2545	2213	6.4	42.8
XM_00562300.3	GBPU01012575.1	<i>C. chalybea</i>	73.043	115	29	2	1429	1542	2196	2082	2.39E-05	61.7
XM_00562300.3	GBPU01012575.1	<i>C. chalybea</i>	67.879	33	4	0	978	1010	2546	2014	6.4	42.8
XM_00562300.3	GBWV01015835.1	<i>N. diveripes</i>	66.329	395	124	4	271	662	391	779	2.23E-18	104
XM_00562300.3	GBWV01015835.1	<i>N. diveripes</i>	68.103	116	37	0	1060	1175	1148	1263	6.4	43.7
XM_00562300.3	GELL01062484.1	<i>M. genalis</i>	64.24	130	137	11	283	719	3323	2888	2.23E-18	104
XM_00562300.3	J000501.1	<i>C. flavipes</i>	68.539	267	84	0	337	603	268	2	2.23E-18	104
XM_00562300.3	GELL01072538.1	<i>M. genalis</i>	66.228	458	143	6	267	715	342	803	7.80E-18	100
XM_00562300.3	GBE01022451.1	<i>S. abditus</i>	69.75	272	79	3	456	724	1331	1063	1.73E-17	100
XM_00562300.3	GHFR01001661.1	<i>A. filius</i>	65.601	431	145	3	289	718	2612	2184	2.72E-17	100
XM_00562300.3	GHUO10103013.1	<i>A. cineraria</i>	65.107	150	210	0	270	713	738	295	2.72E-17	100
XM_00562300.3	GHGA01050651.1	<i>A. haemorrhoea</i>	65.661	447	150	2	270	713	738	295	2.72E-17	100
XM_00562300.3	GBMX01017377.1	<i>A. syrticus</i>	66.842	380	120	5	991	1367	7227	8951	2.72E-17	101
XM_00562300.3	GBMX01017377.1	<i>A. syrticus</i>	64.662	139	89	6	289	686	7098	7511	3.10E-13	77.9
XM_00562300.3	GBMX01017377.1	<i>A. syrticus</i>	66.842	380	120	5	991	1367	7300	6924	2.72E-17	101
XM_00562300.3	GBMX01017377.1	<i>A. syrticus</i>	64.662	139	89	6	289	686	7098	7584	3.10E-13	77.9
XM_00562300.3	GHG001002769.1	<i>B. cryptum</i>	63.804	652	229	4	955	1604	749	103	2.72E-17	101
XM_00562300.3	GHGD01020769.1	<i>B. cryptum</i>	79.242	53	11	0	658	710	1037	985	0.53	47.3
XM_00562300.3	GELL01072535.1	<i>M. genalis</i>	66.528	448	143	10	621	1077	862	609	8.82E-17	98.7
XM_00562300.3	GHT101003818.1	<i>O. cornuta</i>	71.134	194	56	0	500	693	299	106	9.50E-17	98.7
XM_00562300.3	GHFZ01030804.1	<i>B. pasorum</i>	65.498	273	84	0	1331	1602	500	78	9.50E-17	98.7
XM_00562300.3	GHV01028886.1	<i>Nomiodex sp.</i>	66.398	372	120	3	355	722	1986	1096	9.50E-17	98.7
XM_00562300.3	GHV01028886.1	<i>Nomiodex sp.</i>	67.317	205	53	5	979	1176	1348	1151	8.33E-05	59
XM_00562300.3	GHFO010200241.1	<i>B. pasorum</i>	64.616	160	108	6	934	1176	1348	1277	9.50E-17	98.6
XM_00562300.3	GHV01016592.1	<i>P. dentipes</i>	62.726	1328	454	23	294	1605	1968	666	3.31E-16	97.8
XM_00562300.3	GBDI01020547.1	<i>C. sacum</i>	65.434	454	138	5	283	710	250	677	1.16E-15	96
XM_00562300.3	GHV01020547.1	<i>C. sacum</i>	66.076	83	274	8	1298	1476	1183	1431	3.10E-15	96
XM_00562300.3	GHFX01027681.1	<i>O. cornuta</i>	73.125	160	43	0	500	659	184	25	1.16E-15	96
XM_00562300.3	GHV01014681.1	<i>P. dentipes</i>	70.423	203	84	0	480	715	24	249	4.04E-15	94.2
XM_00562300.3	GBPP01027867.1	<i>D. crinita</i>	66.829	353	117	2	295	646	2167	1816	4.04E-15	94.2
XM_00562300.3	GHFPO10102788.1	<i>O. cornuta</i>	65.286	436	138	7	1042	1473	1006	578	4.04E-15	94.2
XM_00562300.3	GHT10101929.1	<i>O. cornuta</i>	65.626	138	138	1	1042	1473	1006	578	4.04E-15	94.2
XM_00562300.3	GHX01011335.1	<i>O. cornuta</i>	65.626	436	138	7	1042	1473	705	277	4.04E-15	94.2
XM_00562300.3	GHX01011335.1	<i>O. cornuta</i>	69.074	152	45	2	764	914	871	821	0.001	58.3
XM_00562300.3	GHFZ01010843.1	<i>B. pasorum</i>	74.306	37	4	0	466	609	642	899	4.04E-15	94.2
XM_00562300.3	HPR4016.1	<i>A. florea</i>	67.384	279	91	0	271	549	300	578	4.04E-15	94.2
XM_00562300.3	GHV01015329.1	<i>M. fulvipes</i>	65.582	79	161	0	460	519	598	856	1.41E-15	94.2
XM_00562300.3	GHV01015329.1	<i>M. fulvipes</i>	61.296	108	31	0	1429	1536	1655	1682	0.001	56.3
XM_00562300.3	GBDP01017181.1	<i>M. fulvipes</i>	67.403	335	99	7	1144	1473	1125	796	1.41E-14	91.4
XM_00562300.3	GHFZ01009662.1	<i>B. pasorum</i>	63.333	130	120	0	1230	1600	1186	1541	6.16E-14	91.4
XM_00562300.3	GBWV01017988.1	<i>N. diveripes</i>	65.035	429	142	3	281	705	962	1386	1.41E-14	91.4
XM_00562300.3	GBWV01017988.1	<i>N. diveripes</i>	64.522	419	158	0	980	1480	1552	2152	1.61E-07	68.9
XM_00562300.3	GBWV01017990.1	<i>N. diveripes</i>	65.035	429	142	3	281	705	962	1386	1.41E-14	91.4
XM_00562300.3	GBPS01016330.1	<i>H. viregatus</i>	64.365	449	160	0	270	718	626	716	4.92E-14	89.7
XM_00562300.3	GFOP1065051.1	<i>H. viregatus</i>	74.658	31	44	0	61	227	471	613	4.92E-14	89.7
XM_00562300.3	GHC01001868.1	<i>O. bicornis</i>	65.596	456	139	7	1042	1473	1540	1112	4.92E-14	89.7
XM_00562300.3	GHG0101867.1	<i>D. icornis</i>	65.076	45	45	2	764	914	1606	1656	0.001	58.3
XM_00562300.3	GMEB01024450.1	<i>P. dentipes</i>	65.542	119	61	4	269	701	414	119	4.92E-14	89.6
XM_00562300.3	GBP201030797.1	<i>C. concolia</i>	65.873	378	121	4	337	710	413	786	4.92E-14	89.6
XM_00562300.3	GBP201030797.1	<i>C. concolia</i>	78.23	23	23	0	1058	1176	1322	1240	7.12E-14	89.6
XM_00562300.3	GHUO1032733.1	<i>C. cunicularis</i>	65.674	427	137	6	288	709	664	1085	4.92E-14	89.6
XM_00562300.3	GHUO1030090.1	<i>A. cineraria</i>	64.802	429	145	2	288	713	2005	1580	4.92E-14	89.6
XM_00562300.3	GHUO1030090.1	<i>A. cineraria</i>	70.866	35	35	0	1423	1548	952	970	0.001	58.3
XM_00562300.3	HPR6872.1	<i>E. nigrita</i>	66.23	382	124	5	340	719	1117	739	4.92E-14	89.6
XM_00562300.3	GHV01017454.1	<i>T. tabulata</i>	67.318	403	89	5	428	618	608	897	1.73E-13	87.8
XM_00562300.3	GHFR01012421.1	<i>A. filius</i>	71.144	201	54	4	978	1176	309	111	1.73E-13	87.8
XM_00562300.3	GHFV01000821.1	<i>B. pasorum</i>	62.759	87	13	1	624	710	639	555	1.73E-13	87.8
XM_00562300.3	GHFV01000821.1	<i>B. pasorum</i>	65.93	117	69	0	951	1001	377	31	1.73E-13	87.8
XM_00562300.3	GHV0101869.1	<i>O. bicornis</i>	72.848	151	41	0	500	650	154	4	1.73E-13	88.7
XM_00562300.3	GHV0101869.1	<i>O. bicornis</i>	64.302	437	156	0	269	707	3196	2730	5.99E-13	86
XM_00562300.3	GHV0101869.1	<i>O. bicornis</i>	73.832	227	67	3	1074	1077	2774	2270	8.33E-13	

Appendix table 6.1 (continued):

XM_00952200.3	HP870571.1	<i>M quadrifasciata</i>	78.121	91	18	1	629	719	748	837	1,085.09	76.1
XM_00952200.3	HP870571.1	<i>M quadrifasciata</i>	65.116	258	88	2	133	389	252	508	0.004	53.6
XM_00952200.3	GBWFO1007387.1	<i>S curvicornis</i>	66.957	230	76	0	283	512	371	142	3.78E-09	73.4
XM_00952200.3	GBWFO1007387.1	<i>S curvicornis</i>	64.296	361	133	0	283	645	270	238	1.32E-06	74.3
XM_00952200.3	GBWFO1007387.1	<i>S curvicornis</i>	65.704	277	89	5	1079	1352	1914	1641	0.001	55.4
XM_00952200.3	GBWFO1019844.1	<i>N diversipes</i>	66.154	260	88	0	457	716	523	782	3.78E-09	73.4
XM_00952200.3	GBWFO1019844.1	<i>N diversipes</i>	67.074	81	21	0	318	398	387	467	0.012	52.7
XM_00952200.3	GHFRO1001565.1	<i>A fulva</i>	68.085	235	65	5	1294	1523	407	178	3.78E-09	73.4
XM_00952200.3	GHIH0102207.1	<i>H itama</i>	68.852	183	53	1	1255	1437	28	206	3.78E-09	73.4
XM_00952200.3	GELLO1017761.1	<i>L xanthopus</i>	64.912	389	132	5	319	713	5414	5220	3.78E-07	74.3
XM_00952200.3	GELLO1017761.1	<i>L xanthopus</i>	72.581	124	30	3	1058	1179	4887	4566	2.39E-05	61.7
XM_00952200.3	GCGPO1023899.1	<i>B pycnosoma</i>	64.336	429	150	3	292	719	5120	4694	3.78E-09	74.3
XM_00952200.3	GELLO1071922.1	<i>M genalis</i>	61.214	70	17	0	452	548	852	865	1.32E-06	71.6
XM_00952200.3	GELLO1071922.1	<i>M genalis</i>	75.714	70	17	0	1403	1472	1696	1765	0.043	50.9
XM_00952200.3	GELLO1071922.1	<i>M genalis</i>	66.397	247	83	0	1227	1473	1	247	1.32E-08	72.5
XM_00952200.3	GELLO1071922.1	<i>O cornuta</i>	64.085	426	151	2	289	713	826	402	1.32E-08	72.5
XM_00952200.3	GELLO1071922.1	<i>O cornuta</i>	67.839	159	64	0	339	537	4	202	1.32E-06	71.6
XM_00952200.3	GELLO1071922.1	<i>H variegatus</i>	65.588	273	94	0	291	563	2071	1799	4.61E-08	69.8
XM_00952200.3	GELLO1071922.1	<i>H variegatus</i>	78.182	55	12	0	1126	1180	1254	1220	0.53	46.4
XM_00952200.3	GELLO1071922.1	<i>B lemestris</i>	65.337	326	109	3	978	1301	87	410	4.61E-08	69.8
XM_00952200.3	GELLO1071922.1	<i>A haemorrhoea</i>	73.228	127	32	2	1423	1548	1058	931	4.61E-08	69.8
XM_00952200.3	GELLO1071922.1	<i>A haemorrhoea</i>	74.034	38	2	0	289	395	1832	1706	1.61E-07	68
XM_00952200.3	GAGHO1023552.1	<i>C chalybea</i>	63.781	693	205	20	335	1010	462	1125	4.61E-08	70.7
XM_00952200.3	HP88232.1	<i>C chalybea</i>	69.504	141	43	0	1407	1547	1522	1662	2.39E-05	61.7
XM_00952200.3	HP88232.1	<i>M willughbiella</i>	71.654	127	36	0	289	395	1832	1706	1.61E-07	68
XM_00952200.3	HP88232.1	<i>N diversipes</i>	65.759	257	88	0	457	713	190	446	1.61E-07	68
XM_00952200.3	HP88232.1	<i>C rorisome</i>	64.267	389	133	3	166	551	1674	1289	1.61E-07	68
XM_00952200.3	HP88232.1	<i>D cincib</i>	76.289	17	22	1	622	718	772	867	6.82E-06	57.2
XM_00952200.3	HP88232.1	<i>B cryptocram</i>	69.945	183	51	4	466	646	1993	1813	1.61E-07	68.9
XM_00952200.3	HP88232.1	<i>B cryptocram</i>	73.043	115	31	0	292	406	2291	2177	1.61E-07	68.9
XM_00952200.3	HP88232.1	<i>A florea</i>	67.428	93	27	1	958	1316	1865	1593	0.024	52.6
XM_00952200.3	HP88232.1	<i>C ravisrons</i>	70.748	147	37	2	955	1098	35	178	1.61E-07	68.9
XM_00952200.3	HP88232.1	<i>C ravisrons</i>	71.034	145	38	2	1338	1480	162	20	1.61E-07	68
XM_00952200.3	HP88232.1	<i>A florea</i>	71.034	145	38	2	1338	1480	162	20	1.61E-07	68
XM_00952200.3	HP88232.1	<i>T orbatus</i>	75.728	103	19	3	1079	1178	725	626	5.61E-07	66.2
XM_00952200.3	HP88232.1	<i>E dilemma</i>	64.176	455	156	6	268	710	882	1132	5.61E-07	66.2
XM_00952200.3	HP88232.1	<i>D cincib</i>	72.872	60	10	0	620	1010	4276	4168	5.61E-07	66.2
XM_00952200.3	HP88232.1	<i>T carbonaria</i>	82.353	51	9	0	1126	1176	3762	3712	0.012	52.7
XM_00952200.3	HP88232.1	<i>S punctulatislima</i>	76.744	86	20	0	625	710	697	782	5.61E-07	66.2
XM_00952200.3	HP88232.1	<i>M genalis</i>	71.901	281	94	0	292	410	3051	2921	0.012	52.7
XM_00952200.3	HP88232.1	<i>A haemorrhoea</i>	75.824	91	22	0	620	710	2	92	5.61E-07	66.2
XM_00952200.3	HP88232.1	<i>A haemorrhoea</i>	65.714	280	94	2	270	548	3210	2932	5.61E-07	66.2
XM_00952200.3	HP88232.1	<i>A haemorrhoea</i>	70.079	127	36	0	1423	1548	2069	1944	6.82E-06	57.2
XM_00952200.3	HP88232.1	<i>M genalis</i>	75.532	94	23	0	1058	1151	1156	1249	5.61E-07	67.1
XM_00952200.3	HP88232.1	<i>M genalis</i>	67.453	63	33	3	337	545	487	655	1.96E-06	64.4
XM_00952200.3	HP88232.1	<i>E variegatus</i>	67.05	261	73	0	466	202	608	861	1.96E-06	64.4
XM_00952200.3	HP88232.1	<i>S punctulatislima</i>	67.123	219	70	2	493	710	664	881	1.96E-06	64.4
XM_00952200.3	HP88232.1	<i>S punctulatislima</i>	67.123	219	70	2	493	710	664	881	1.96E-06	64.4
XM_00952200.3	HP88232.1	<i>S punctulatislima</i>	67.123	219	70	2	493	710	664	881	1.96E-06	64.4
XM_00952200.3	HP88232.1	<i>O cornuta</i>	74.49	98	25	0	500	597	533	630	1.96E-06	65.3
XM_00952200.3	HP88232.1	<i>C rorisome</i>	73.146	28	9	0	289	396	143	450	1.96E-06	65.3
XM_00952200.3	HP88232.1	<i>C rorisome</i>	64.8	375	116	7	335	701	315	681	1.96E-06	65.3
XM_00952200.3	HP88232.1	<i>C rorisome</i>	69.8	375	116	7	335	701	315	681	1.96E-06	65.3
XM_00952200.3	HP88232.1	<i>M genalis</i>	64.416	381	131	4	330	512	330	111	4	65.3
XM_00952200.3	HP88232.1	<i>M genalis</i>	71.429	112	26	3	1074	1182	1121	1229	0.043	50.9
XM_00952200.3	HP88232.1	<i>B pascorum</i>	63.869	429	152	3	292	719	2062	1636	1.96E-06	65.3
XM_00952200.3	HP88232.1	<i>A haemorrhoea</i>	71.108	63	19	0	468	488	448	448	1.96E-06	65.3
XM_00952200.3	HP88232.1	<i>B cryptocram</i>	67.172	198	65	0	289	486	199	2	1.96E-06	65.3
XM_00952200.3	HP88232.1	<i>O cornuta</i>	64.22	436	145	7	1042	1473	403	831	6.84E-06	62.6
XM_00952200.3	HP88232.1	<i>M genalis</i>	66.507	209	70	0	508	716	3	211	6.84E-06	62.6
XM_00952200.3	HP88232.1	<i>M genalis</i>	70	120	32	2	1058	1175	541	658	0.043	50
XM_00952200.3	HP88232.1	<i>M genalis</i>	66.507	209	70	0	508	716	3	211	6.84E-06	62.6
XM_00952200.3	HP88232.1	<i>M genalis</i>	70	120	32	2	1058	1175	541	658	0.043	50
XM_00952200.3	HP88232.1	<i>M genalis</i>	70	120	32	2	1058	1175	541	658	0.043	50
XM_00952200.3	HP88232.1	<i>B lemestris</i>	65.992	247	80	2	1123	1397	292	44	6.84E-06	62.6
XM_00952200.3	HP88232.1	<i>O bicornis</i>	64.22	145	47	3	1332	1432	1670	1670	0.03	46.4
XM_00952200.3	HP88232.1	<i>O bicornis</i>	73.333	75	20	0	619	693	1744	1670	0.03	46.4
XM_00952200.3	HP88232.1	<i>C chalybea</i>	73.216	116	29	2	600	714	1117	1003	6.84E-06	63.5
XM_00952200.3	HP88232.1	<i>H itama</i>	69.6	200	69	0	469	609	313	294	6.84E-06	63.5
XM_00952200.3	HP88232.1	<i>E dilemma</i>	64.074	438	162	0	289	726	3131	2694	2.39E-05	61.7
XM_00952200.3	HP88232.1	<i>E dilemma</i>	63.074	438	162	0	289	726	3131	2694	2.39E-05	61.7
XM_00952200.3	HP88232.1	<i>E dilemma</i>	73.469	88	26	0	625	722	1063	966	2.39E-05	61.7
XM_00952200.3	HP88232.1	<i>E dilemma</i>	78.846	52	11	0	1429	1480	288	217	1.8	45.5
XM_00952200.3	HP88232.1	<i>H brunconum</i>	71.214	103	24	0	622	714	976	1068	2.39E-05	60.8
XM_00952200.3	HP88232.1	<i>T nigritops</i>	74.194	93	24	0	622	714	976	1068	2.39E-05	60.8
XM_00952200.3	HP88232.1	<i>S punctulatislima</i>	72.222	108	30	0	292	399	362	469	2.39E-05	60.8
XM_00952200.3	HP88232.1	<i>B pascorum</i>	63.636	153	3	0	292	399	362	469	2.39E-05	60.8
XM_00952200.3	HP88232.1	<i>B pascorum</i>	63.636	153	3	0	292	399	362	469	2.39E-05	60.8
XM_00952200.3	HP88232.1	<i>B pascorum</i>	63.636	153	3	0	292	399	362	469	2.39E-05	60.8
XM_00952200.3	HP88232.1	<i>A cineraria</i>	65.079	315	104	4	335	646	2127	1816	2.39E-05	60.8
XM_00952200.3	HP88232.1	<i>A cineraria</i>	63.074	429	153	3	292	719	2062	1636	2.39E-05	60.8
XM_00952200.3	HP88232.1	<i>P denipes</i>	66.963	221	75	0	1153	1373	224	444	2.39E-05	61.7
XM_00952200.3	HP88232.1	<i>D hirsipes</i>	68	250	83	2	289	537	1623	137		

Appendix table 6.2: Maximum Likelihood fits of 56 different amino acid substitution models for the genomic and transcriptomic data for bee species. Bayesian Information Criterion (BIC); Akaike Information Criterion, corrected (AICc) and Maximum Likelihood (lnL) are shown. Where applicable, estimates of gamma (G) shape parameter and the estimated fraction of invariant (I) sites are presented.

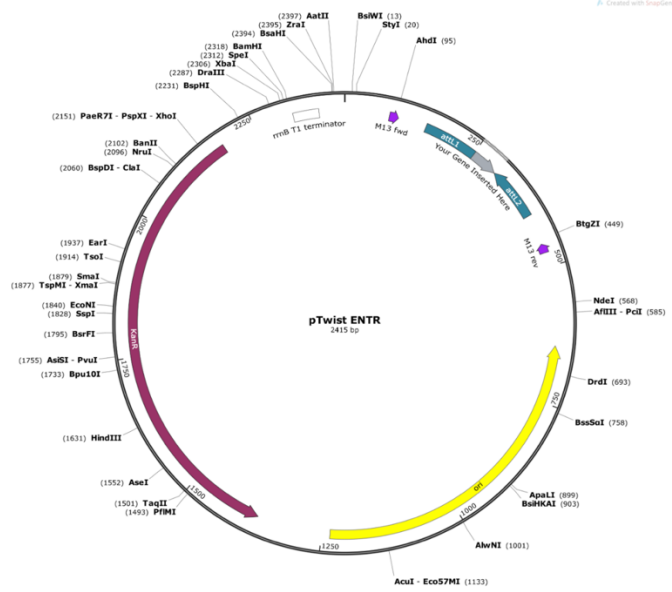
Model	BIC	AICc	lnL	Invariant	Gamma
LG+G	9246.66063	5186.13081	-351.68008	n/a	3.94316748
rtREV+G	9247.27652	5186.7467	-351.98803	n/a	4.10491761
LG+G+I	9254.3573	5194.28946	-351.68008	0	3.94316748
rtREV+G+I	9254.97318	5194.90535	-351.98803	0	4.10491761
WAG+G	9255.44304	5194.91322	-356.07129	n/a	4.18826406
JTT+G	9256.04719	5195.51738	-356.37337	n/a	3.53046869
WAG+G+I	9263.13971	5203.07188	-356.07129	0	4.18826406
JTT+G+I	9263.74386	5203.67603	-356.37337	0	3.53046869
Dayhoff+G	9270.83731	5210.30749	-363.76842	n/a	3.1928926
mtREV24+G	9276.82031	5216.29049	-366.75992	n/a	3.6021515
Dayhoff+G+I	9278.53397	5218.46614	-363.76842	0	3.1928926
rtREV	9282.07819	5221.10135	-373.2372	n/a	n/a
LG	9282.71341	5221.73657	-373.55481	n/a	n/a
mtREV24+G+	9284.51697	5224.44914	-366.75992	0	3.6021515
rtREV+I	9289.81612	5229.2863	-373.25783	0	n/a
WAG	9290.40555	5229.42871	-377.40088	n/a	n/a
LG+I	9290.45965	5229.92983	-373.57959	0	n/a
WAG+I	9297.9861	5237.45628	-377.34282	0	n/a
cpREV+G	9299.95745	5239.42763	-378.32849	n/a	1.80048871
JTT	9305.54578	5244.56894	-384.97099	n/a	n/a
cpREV+G+I	9307.65412	5247.58628	-378.32849	0	1.80048871
JTT+I	9313.2984	5252.76858	-384.99897	0	n/a
Dayhoff	9325.75048	5264.77363	-395.07334	n/a	n/a
mtREV24	9326.52426	5265.54742	-395.46023	n/a	n/a
Dayhoff+I	9333.41503	5272.88521	-395.05729	0	n/a
mtREV24+I	9334.28137	5273.75155	-395.49046	0	n/a
cpREV	9340.92436	5279.94751	-402.66028	n/a	n/a
cpREV+I	9348.62103	5288.09121	-402.66028	0	n/a
LG+G+F	9503.63221	5454.48537	-407.04754	n/a	1.1723021
JTT+G+F	9505.92753	5456.78069	-408.1952	n/a	1.08026198
WAG+G+F	9509.68851	5460.54166	-410.07569	n/a	1.13895936
LG+G+I+F	9511.32888	5462.9362	-407.04754	0	1.1723021
JTT+G+I+F	9513.6242	5465.23152	-408.1952	0	1.08026198
rtREV+G+F	9514.42915	5465.2823	-412.44601	n/a	4.44859519
WAG+G+I+F	9517.38517	5468.99249	-410.07569	0	1.13895936
LG+F	9520.52177	5470.63653	-419.34065	n/a	n/a
mtREV24+G+	9520.65044	5471.50359	-415.55665	n/a	1.15434743
rtREV+G+I+F	9522.12582	5473.73314	-412.44601	0	4.44859519
Dayhoff+G+F	9526.29718	5477.15034	-418.38002	n/a	1.04538785
rtREV+F	9526.80582	5476.92057	-422.48267	n/a	n/a
LG+I+F	9528.21581	5479.06896	-419.33934	0	n/a
mtREV24+G+	9528.3471	5479.95442	-415.55665	0	1.15434743
JTT+F	9532.27079	5482.38555	-425.21516	n/a	n/a
WAG+F	9532.30258	5482.41734	-425.23106	n/a	n/a
Dayhoff+G+I+	9533.99385	5485.60117	-418.38002	0	1.04538785
rtREV+I+F	9534.48788	5485.34103	-422.47537	0	n/a
WAG+I+F	9539.9603	5490.81345	-425.21158	0	n/a
JTT+I+F	9539.96746	5490.82061	-425.21516	0	n/a
mtREV24+F	9543.82236	5493.93712	-430.99095	n/a	n/a
mtREV24+I+F	9551.51805	5502.3712	-430.99046	0	n/a
Dayhoff+F	9559.52647	5509.64123	-438.843	n/a	n/a
cpREV+G+F	9559.82493	5510.67809	-435.1439	n/a	2.63786768
Dayhoff+I+F	9567.21505	5518.06821	-438.83896	0	n/a
cpREV+G+I+	9567.5216	5519.12892	-435.1439	0	2.63786768
cpREV+F	9578.56698	5528.68174	-448.36326	n/a	n/a

Appendix table 6.3: Data quality summary of RNA Illumina sequencing

Species	Raw reads	Raw data	Effective (%)	Error (%)	Q20 (%)	Q30 (%)	GC (%)
<i>M. centuncularis</i>	304495898	45674384700	98.27	0.02	98.18	94.64	42.88
<i>M. leachella</i>	301152054	45172808100	96.9	0.03	97.79	94.06	44.99
<i>M. lapponica</i>	218518922	32777838300	98.02	0.02	98.2	94.54	41.73
<i>M. willughbiella</i>	253327736	37999160400	98.77	0.02	98.21	94.64	42.66

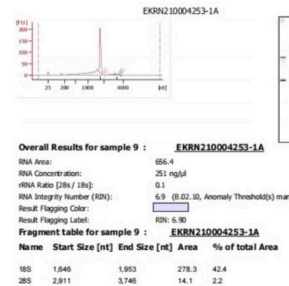
Appendix table 6.4: Maximum Likelihood fits of 56 different amino acid substitution models for the Hymenoptera protein sequence data. Bayesian Information Criterion (BIC); Akaike Information Criterion, corrected (AICc) and Maximum Likelihood (lnL) are shown. Where applicable, estimates of gamma (G) shape parameter and the estimated fraction of invariant (I) sites are presented

Model	BIC	AICc	lnL	Invariant	Gamma
JTT+G+I+F	545739.932	528569.676	-262685.76	0.01079914	1.49687398
JTT+G+F	545889.96	528730.48	-262767.18	n/a	1.35890084
JTT+G+I	546821.043	529855.54	-263347.86	0.01079914	1.35800034
LG+G+I	546854.279	529888.777	-263364.48	0.01079914	1.48420508
JTT+G	546971.516	530016.79	-263429.5	n/a	1.29610041
LG+G	547015.834	530061.108	-263451.66	n/a	1.40632825
LG+G+I+F	547019.047	529848.791	-263325.32	0.01079914	1.48459447
LG+G+F	547158.696	529999.215	-263401.54	n/a	1.40467843
WAG+G+I+F	550064.235	532893.978	-264847.92	0.01079914	1.43043464
WAG+G+F	550209.847	533050.367	-264927.12	n/a	1.35841998
WAG+G+I	550820.352	533854.85	-265347.52	0.01079914	1.42009577
WAG+G	550975.658	534020.932	-265431.57	n/a	1.32832988
mtREV24+G+	553880.898	536710.642	-266756.25	0.01079914	1.30547187
mtREV24+G	553929.862	536770.382	-266787.13	n/a	1.27345176
rtREV+G+I+F	554660.814	537490.557	-267146.21	0.01079914	1.43459511
rtREV+G+F	554801.208	537641.728	-267222.8	n/a	1.35134827
Dayhoff+G+I	554807.967	537637.711	-267219.78	0.01079914	1.32131604
Dayhoff+G+F	554935.608	537776.128	-267290	n/a	1.26171845
cpREV+G	556508.708	539553.982	-268198.09	n/a	0.77991725
cpREV+G+I	557312.888	540347.385	-268593.79	0.01079914	0.70489677
Dayhoff+G+I	557676.707	540711.204	-268775.7	0.01079914	1.26369571
cpREV+G+F	557707.47	540547.989	-268675.93	n/a	0.78501313
rtREV+G+I	557737.463	540771.961	-268806.08	0.01079914	1.46006751
Dayhoff+G	557805.224	540850.498	-268846.35	n/a	1.22803513
rtREV+G	557900.98	540946.254	-268894.23	n/a	1.36466803
cpREV+G+I	558773.696	541603.439	-269202.65	0.01079914	0.69964258
JTT+I+F	565965.149	548805.669	-272804.77	0.01079914	n/a
mtREV24+G+	566092.113	549126.61	-272983.4	0.01079914	1.20665726
mtREV24+G	566117.718	549162.992	-273002.6	n/a	1.18310823
JTT+F	566300.035	549151.332	-272978.61	n/a	n/a
JTT+I	567230.29	550275.564	-273558.89	0.01079914	n/a
JTT	567632.598	550688.648	-273766.44	n/a	n/a
LG+I	567757.671	550802.945	-273822.58	0.01079914	n/a
LG	568119.133	551175.184	-274009.7	n/a	n/a
LG+I+F	568525.944	551366.464	-274085.17	0.01079914	n/a
LG+F	568814.233	551665.53	-274235.71	n/a	n/a
WAG+I+F	569603.628	552444.148	-274624.01	0.01079914	n/a
WAG+F	569866.443	552717.739	-274761.81	n/a	n/a
WAG+I	570363.017	553408.291	-275125.25	0.01079914	n/a
WAG	570738.687	553794.737	-275319.48	n/a	n/a
Dayhoff+I+F	575496.684	558337.203	-277570.54	0.01079914	n/a
Dayhoff+F	575659.949	558511.245	-277658.57	n/a	n/a
mtREV24+F	575934.291	558785.587	-277795.74	n/a	n/a
mtREV24+I+F	576070.089	558910.608	-277857.24	0.01079914	n/a
rtREV+I+F	576931.646	559772.166	-278288.02	0.01079914	n/a
rtREV+F	577276.763	560128.06	-278466.97	n/a	n/a
rtREV+I	578972.998	562018.272	-279430.24	0.01079914	n/a
Dayhoff+I	579171.891	562217.165	-279529.69	0.01079914	n/a
rtREV	579444.611	562500.662	-279672.44	n/a	n/a
Dayhoff	579492.802	562548.853	-279696.54	n/a	n/a
cpREV	580914.732	563970.783	-280407.5	n/a	n/a
cpREV+F	582160.507	565011.804	-280908.85	n/a	n/a
cpREV+I	582370.054	565415.328	-281128.77	0.01079914	n/a
cpREV+I+F	584143.086	566983.606	-281893.74	0.01079914	n/a
mtREV24	591106.605	574162.656	-285503.44	n/a	n/a

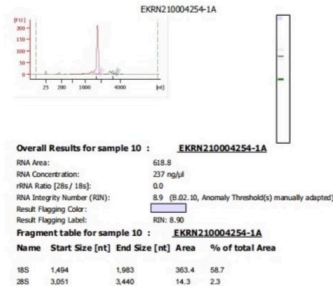


Appendix figure 6.1: A map of the attL-containing-donor vector pTwistENTR

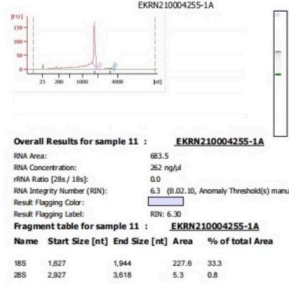
Sample 1: *M. centuncularis*



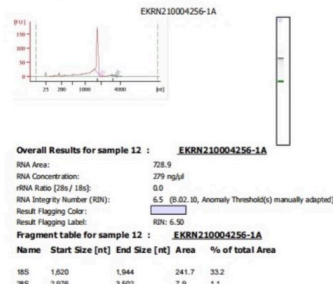
Sample 2: *M. leachella*



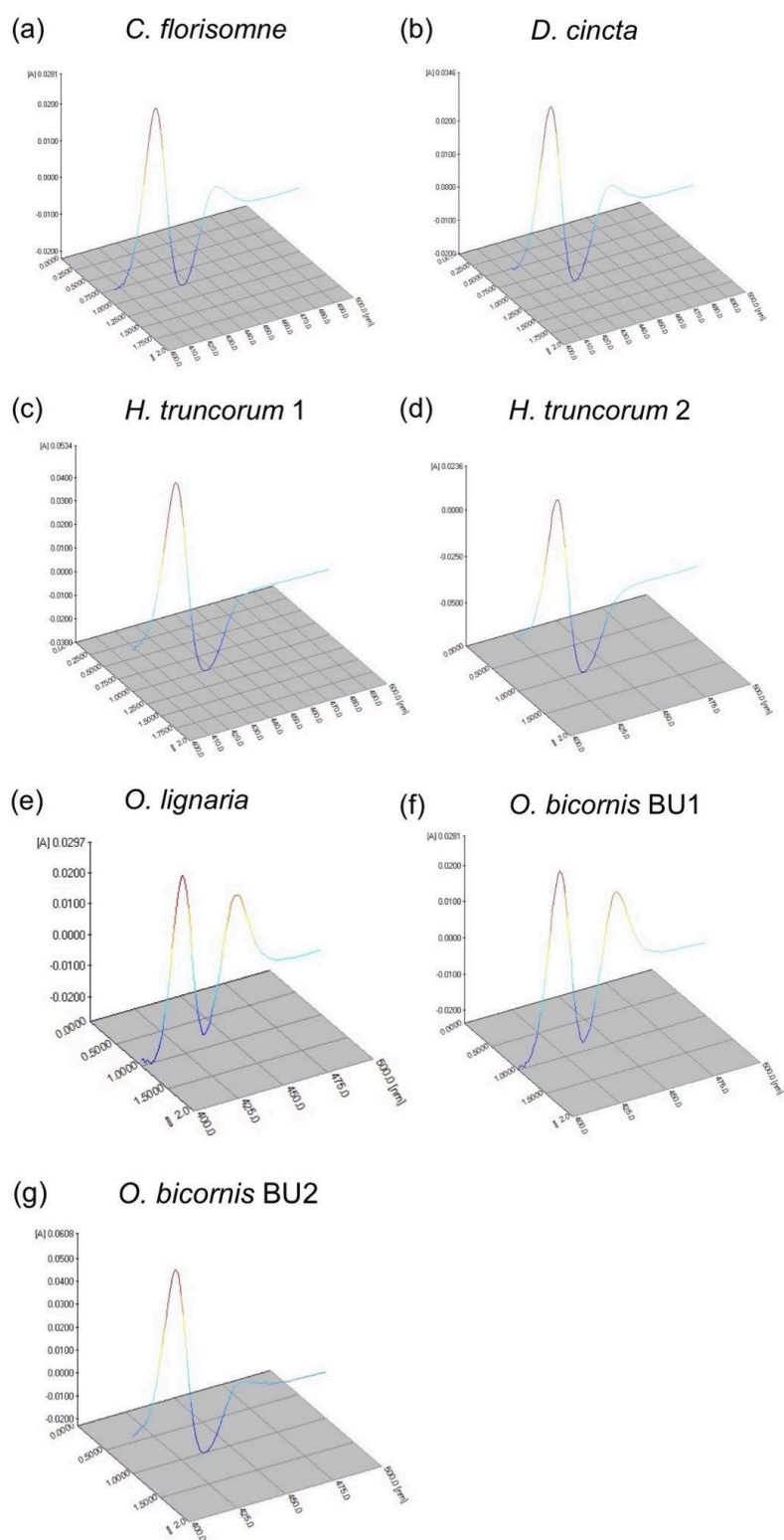
Sample 3: *M. lapponica*



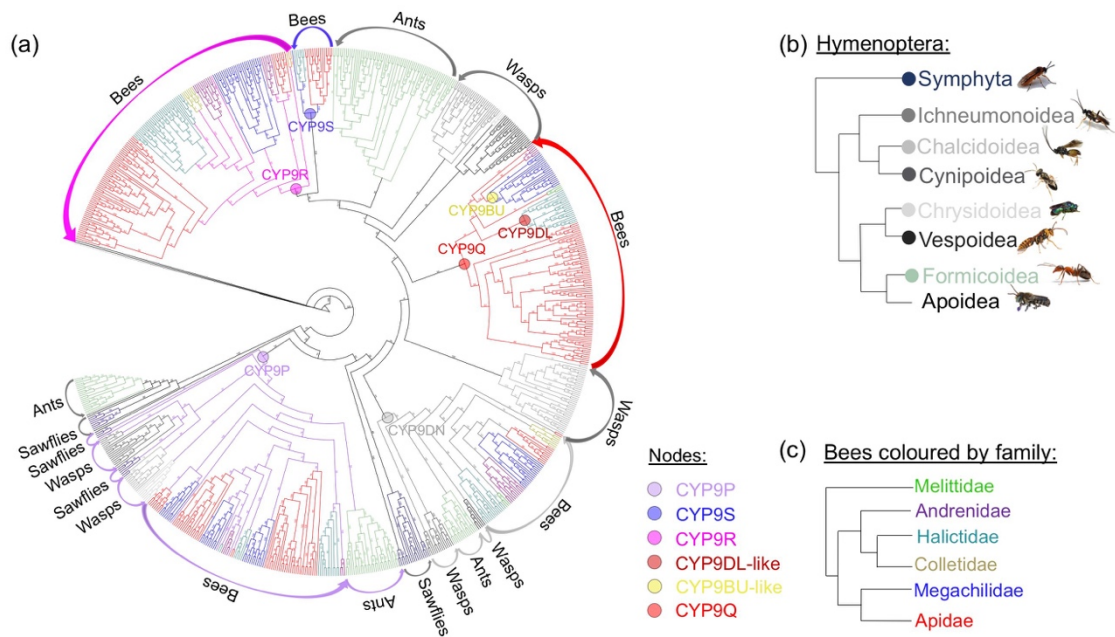
Sample 4: *M. willughbiella*



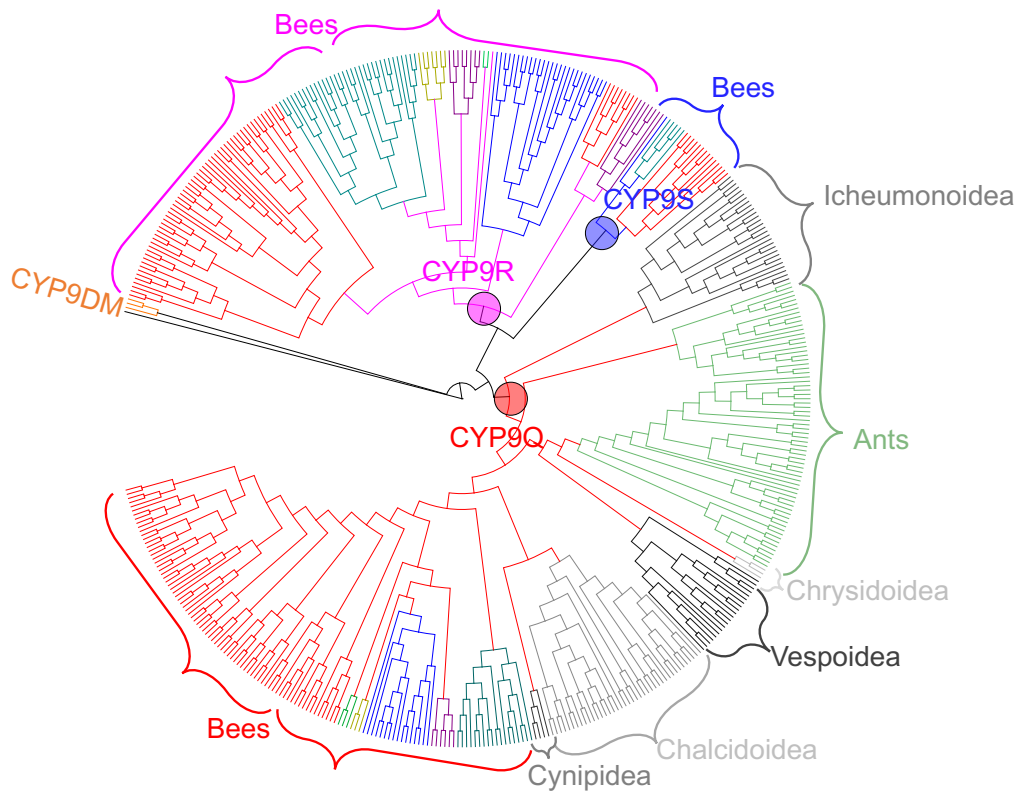
Appendix figure 6.2: RNA quality control results from Novogene UK for three UK native and one Canadian *Megachile* species using Agilent 2100



Appendix figure 6.3: CO-difference spectrum of recombinant Megachilidae CYP9 proteins: (a) *C. florissomne* CYP9-like, (b) *D. cincta* CYP9-like, (c) *H. truncorum* CYP9-like1, (d) *H. truncorum* CYP9-like1, (e) *O. lignaria* CYP9-like, (f) *O. bicornis* CYP9BU1, (g) *O. bicornis* CYP9BU2. The P450s were reduced with sodium dithionite ($\text{Na}_2\text{S}_2\text{O}_4$).

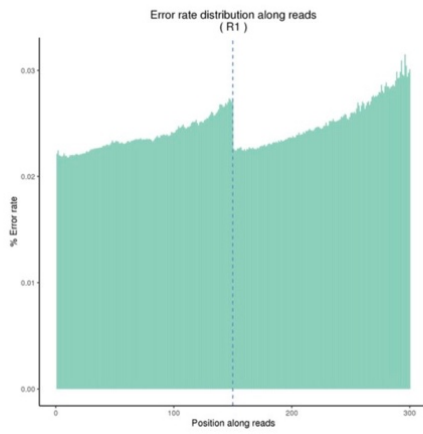


Appendix figure 6.4: (a) PhyML maximum likelihood phylogeny (substitution model JTT+G, with branch support of 50 bootstraps) of the CYP subfamily across the Order: Hymenoptera. Tree rooted on *M. domestica* CYP4ae1. Sequences are coloured by superfamily (see table 6.8), apart from bees which are coloured by family. All nucleotide sequences accessed from NCBI databases. Nodes that mark the divergence of superfamilies are denoted by coloured circles. (b) Schematic of the phylogenetic relationship of the Hymenoptera. (c) Schematic of the phylogenetic relationship between bee families.

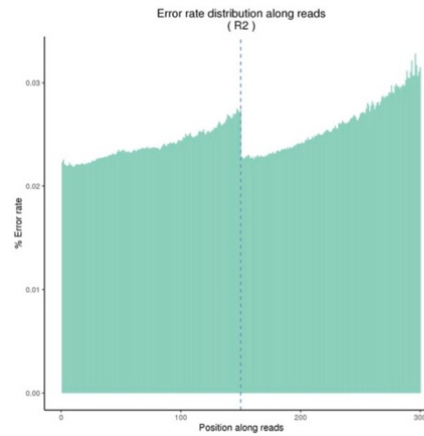


Appendix figure 6.5: PhyML maximum likelihood phylogeny (substitution model JTT+G, with branch support of 50 bootstraps) of the CYP9R, CYP9S and CYP9Q/BU/DL lineages across the Order: Hymenoptera. Tree rooted on *M. domestica* CYP4ae1. Sequences are coloured by superfamily (see table 6.8), apart from bees which are coloured by family. All nucleotide sequences accessed from NCBI databases. Nodes that mark the divergence of CYP9 lineages are denoted by coloured circles.

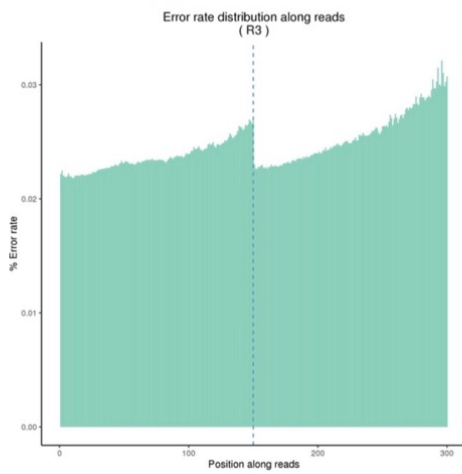
M. centuncularis



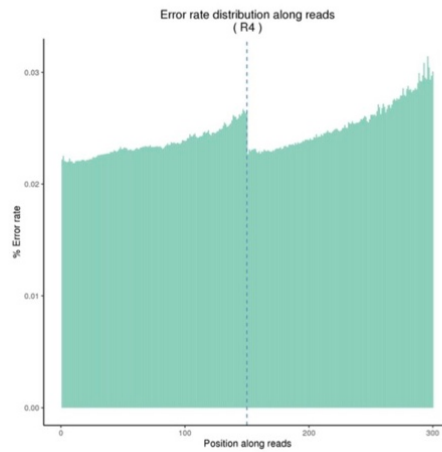
M. leachella



M. lapponica



M. willughbiella



Appendix figure 6.7: Error rate in the RNA-seq data from Novogene. The error rate distribution along reads is shown ($\leq 0.03\%$). The base position is shown on the horizontal axis and the single base error rate is on the vertical axis.

Bibliography

1. FAO, *The State of Food Insecurity in the World 2001*. 2001, FAO, Rome: Rome.
2. Gaud, W.S. *The Green revolution: Accomplishments and Apprehensions*. in *The Society for International Development*. 1968. Shorehan Hotel, Washington, DC.
3. Popp, J., K. Peto, and J. Nagy, *Pesticide productivity and food security. A review*. *Agronomy for Sustainable Development*, 2013. **33**(1): p. 243-255.
4. Foley, J.A., et al., *Global consequences of land use*. *Science*, 2005. **309**(5734): p. 570-574.
5. Godfray, H.C.J., et al., *Food Security: The Challenge of Feeding 9 Billion People*. *Science*, 2010. **327**(5967): p. 812-818.
6. Ganivet, E., *Growth in human population and consumption both need to be addressed to reach an ecologically sustainable future*. *Environment Development and Sustainability*, 2020. **22**(6): p. 4979-4998.
7. Godfray, H.C.J., *The challenge of feeding 9-10 billion people equitably and sustainably*. *Journal of Agricultural Science*, 2014. **152**: p. S2-S8.
8. FAO; IFAD; WFP, *The State of Food Insecurity in the World 2015. Meeting the 2015 international hunger targets: taking stock of uneven progress*. 2015, Rome: FAO Rome.
9. FAO, et al., *The state of food security and nutrition in the world 2020. Transforming food systems for affordable healthy diets*. 2020, Food and Agriculture Organization of the United Nations.: Rome.
10. Ollerton, J., *Pollinator Diversity: Distribution, Ecological Function, and Conservation*. *Annual Review of Ecology, Evolution, and Systematics*, Vol 48, 2017. **48**: p. 353-376.
11. Woodcock, B.A., et al., *Meta-analysis reveals that pollinator functional diversity and abundance enhance crop pollination and yield*. *Nature Communications*, 2019. **10**.
12. IPBES; S.G. Potts, V.L.I.-F., H. T. Ngo, J. C. Biesmeijer, T. D. Breeze, L. V., L.A.G. Dicks, R. Hill, J. Settele, A. J. Vanbergen, M. A. Aizen, S. A. Cunningham, C., and B.M.F. Eardley, N. Gallai, P. G. Kevan, A. Kovács-Hostyánszki, P. K. Kwapong, J. Li, X. Li, D.J. Martins, G. Nates-Parra, J. S. Pettis, R. Rader, and B. F. Viana (eds.), *Summary for policymakers of the assessment report of the Intergovernmental Science-Policy Platform on Biodiversity and Ecosystem Services on pollinators, pollination and food production*. 2016, IPBES: Bonn, Germany.
13. Rader, R., et al., *Non-Bee Insects as Visitors and Pollinators of Crops: Biology, Ecology, and Management*. *Annual Review of Entomology*, Vol 65, 2020. **65**: p. 391-407.
14. Chaplin-Kramer, R., et al., *Global malnutrition overlaps with pollinator-dependent micronutrient production*. *Proceedings of the Royal Society B-Biological Sciences*, 2014. **281**(1794).
15. Eilers, E.J., et al., *Contribution of Pollinator-Mediated Crops to Nutrients in the Human Food Supply*. *Plos One*, 2011. **6**(6).
16. Biesmeijer, J.C., et al., *Parallel declines in pollinators and insect-pollinated plants in Britain and the Netherlands*. *Science*, 2006. **313**(5785): p. 351-354.
17. NRC, *Committee on the Status of Pollinators in North America. 2007. Status of Pollinators in North America*. 2007, Washington, D.C.: National Research Council of the National Academies. National Academies Press.
18. Leather, S.R., *"Ecological Armageddon" - more evidence for the drastic decline in insect numbers*. *Annals of Applied Biology*, 2018. **172**(1): p. 1-3.
19. Seibold, S., et al., *Arthropod decline in grasslands and forests is associated with landscape-level drivers*. *Nature*, 2019. **574**(7780): p. 671-674.
20. Wagner, D.L., *Insect Declines in the Anthropocene*. *Annual Review of Entomology*, Vol 65, 2020. **65**: p. 457-480.

21. FAO, *Rapid assessment of pollinators' status. A contribution to the International Initiative for the Conservation and Sustainable Use of Pollinators.*, ed. . 2008, Rome, Italy.: United Nations Food and Agriculture Organization (FAO).
22. Nieto, A., et al., *European Red List of bees.*, ed. IUCN. 2014, Luxembourg. EU: IUCN. Publication office of the European Union.
23. Cameron, S.A., et al., *Patterns of widespread decline in North American bumble bees.* Proceedings of the National Academy of Sciences of the United States of America, 2011. **108**(2): p. 662-667.
24. Williams, P.H. and J.L. Osborne, *Bumblebee vulnerability and conservation world-wide.* Apidologie, 2009. **40**(3): p. 367-387.
25. Goulson, D., et al., *Causes of rarity in bumblebees.* Biological Conservation, 2005. **122**(1): p. 1-8.
26. Goulson, D., et al., *Bee declines driven by combined stress from parasites, pesticides, and lack of flowers.* Science, 2015. **347**(6229).
27. Carvell, C., et al., *Declines in forage availability for bumblebees at a national scale.* Biological Conservation, 2006. **132**(4): p. 481-489.
28. Wood, T.J., et al., *Managed honey bees as a radar for wild bee decline?* Apidologie, 2020. **51**(6): p. 1100-1116.
29. Hardstone, M.C. and J.G. Scott, *Is Apis mellifera more sensitive to insecticides than other insects?* Pest Management Science, 2010. **66**(11): p. 1171-1180.
30. Johnson, R.M., et al., *Pesticides and honey bee toxicity - USA.* Apidologie, 2010. **41**(3): p. 312-331.
31. Sanchez-Bayo, F., et al., *Are bee diseases linked to pesticides? - A brief review.* Environment International, 2016. **89-90**: p. 7-11.
32. Briggs, H., *New pesticides 'may have risks for bees'.* in *BBC Science and Environment.* 2018, BBC.
33. Gabbatiss, J., *New 'safe' pesticides to replace banned chemicals still hurt bees, scientists say.*, in *Independent.* 2019.
34. Aratani, L., *Pesticide widely used in US particularly harmful to bees, study finds.*, in *The Guardian.* 2019.
35. Allen-Wardell, G., et al., *The potential consequences of pollinator declines on the conservation of biodiversity and stability of food crop yields.* Conservation Biology, 1998. **12**(1): p. 8-17.
36. Klein, A.M., et al., *Importance of pollinators in changing landscapes for world crops.* Proceedings of the Royal Society B-Biological Sciences, 2007. **274**(1608): p. 303-313.
37. Osborne, J.L., *ECOLOGY Bumblebees and pesticides.* Nature, 2012. **491**(7422): p. 43-45.
38. Lehmann, D.M. and A.A. Camp, *A systematic scoping review of the methodological approaches and effects of pesticide exposure on solitary bees.* Plos One, 2021. **16**(5).
39. Vasiliev, D. and S. Greenwood, *Pollinator biodiversity and crop pollination in temperate ecosystems, implications for national pollinator conservation strategies: Mini review.* Science of the Total Environment, 2020. **744**.
40. Whitmee, S., et al., *Safeguarding human health in the Anthropocene epoch: report of The Rockefeller Foundation-Lancet Commission on planetary health.* Lancet, 2015. **386**(10007): p. 1973-2028.
41. Myers, S.S., et al., *Climate Change and Global Food Systems: Potential Impacts on Food Security and Undernutrition*, in *Annual Review of Public Health, Vol 38*, J.E. Fielding, Editor. 2017. p. 259-277.
42. IPCC, *IPCC Summary for Policymakers. In Global warming of 1.5°C. An IPCC Special Report on the impacts of global warming of 1.5°C above pre-industrial levels and related global greenhouse gas emission pathways, in the context of strengthening the global response to the threat of climate change, sustainable development, and efforts to eradicate poverty.*, ed. V. Masson-Delmotte, et al. 2018, Geneva, Switzerland: World Meteorological Organization,. 32pp.

43. Deutsch, C.A., et al., *Increase in crop losses to insect pests in a warming climate*. *Science*, 2018. **361**(6405): p. 916-919.
44. Patel, R., *Food sovereignty*. *Journal of Peasant Studies*, 2009. **36**(3): p. 663-673.
45. Messer, E. and M.J. Cohen, *Conflict, food insecurity and globalization*. *Food, Culture & Society*, 2007. **10**(2): p. 297-315.
46. Alston, J.M. and P.G. Pardey, *Agriculture in the Global Economy*. *Journal of Economic Perspectives*, 2014. **28**(1): p. 121-146.
47. Gomes, H.D., et al., *A socio-environmental perspective on pesticide use and food production*. *Ecotoxicology and Environmental Safety*, 2020. **197**.
48. Richards, K.W., *Alfalfa Leafcutter Bee Management in Western Canada*. . 1984, Canada: Agriculture Canada Publications
49. Zoltowska, K., R. Fraczek, and Z. Lipinski, *Hydrolases of developing worker brood and newly emerged worker of *Apis mellifera carnica**. . *Journal of Apicultural Science*, 2011. **55**(1): p. 27-36.
50. Mitchell, T.B., *Bees of the eastern United States II. [Megachilidae, Anthophoridae, Apidae s.s.]*. Technical bulletin (North Carolina Agricultural Experiment Station), 1962. **152**: p. 1-557.
51. Brewer, M., *Alfalfa leafcutting bee, *Megachile rotundata**. University of Wyoming CES, 1995. **B1013.5**: p. 1-2.
52. Bohart, G.E., *Management of wild bees for pollination of crops*. *Annual Review of Entomology*, 1972. **17**: p. 287-&.
53. Pitts-Singer, T.L., *Management of Alfalfa Bees. Chapter 7, In Bee Pollination in Agricultural Ecosystems*. . 2008, New York.: Oxford University Press.
54. Bohart, G.E., *Commercial production and management of wild bees – a new entomological industry*. *Bulletin of the Entomological Society of America*, 1970. **6**(1): p. 8.
55. Kemp, W.P. and J. Bosch, *Development and emergence of the alfalfa pollinator *Megachile rotundata* (Hymenoptera : megachilidae)*. *Annals of the Entomological Society of America*, 2000. **93**(4): p. 904-911.
56. Pitts-Singer, T.L. and J.H. Cane, *The Alfalfa Leafcutting Bee, *Megachile rotundata*: The World's Most Intensively Managed Solitary Bee*. *Annual Review of Entomology*, Vol 56, 2011. **56**: p. 221-237.
57. Richards, K.W., *Alfalfa leafcutter bee management in Canada*. *Bee World*, 1987. **68**(4): p. 168-178.
58. Tepedino, V.J. and D.R. Frohlich, *Fratricide in *Megachile rotundata*, a non-social Megachilid bee - impartial treatment of sibs and non-sibs*. *Behavioral Ecology and Sociobiology*, 1984. **15**(1): p. 19-23.
59. Eickwort, G.C. and H.S. Ginsberg, *Foraging and mating behavior in Apoidea*. *Annual Review of Entomology*, 1980. **25**: p. 421-446.
60. Gerber, H.S. and E.C. Klostermeyer, *Sex control by bees. A voluntary act of egg fertilization during oviposition*. *Science*, 1970. **167**(3914): p. 82-+.
61. Richards, K.W., *Ovarian development in the alfalfa leafcutter bee, *Megachile rotundata**. *Journal of Apicultural Research*, 1994. **33**(4): p. 199-203.
62. Trostle, G. and P.F. Torchio, *Comparative nesting behavior and immature development of *Megachile rotundata* (Fabricius) and *Megachile apicalis spinola* (Hymenoptera, Megachilidae)*. *Journal of the Kansas Entomological Society*, 1994. **67**(1): p. 53-72.
63. Stephen, W.P. and C.E. Osgood, *Influence of tunnel size and nesting medium on sex ratios in leafcutter bee *Megachile rotundata**. *Journal of Economic Entomology*, 1965. **58**(5): p. 965-&.
64. Litman, J.R., et al., *Why do leafcutter bees cut leaves? New insights into the early evolution of bees*. *Proceedings of the Royal Society B-Biological Sciences*, 2011. **278**(1724): p. 3593-3600.
65. Mitra, A., *Function of the Dufour's gland in solitary and social Hymenoptera*. *Journal of Hymenoptera Research*, 2013. **35**: p. 33-58.

66. Klostermeyer, E.C. and H.S. Gerber, *Nesting behavior of Megachile rotundata (Hymenoptera - Megachilidae) monitored with an event recorder*. Annals of the Entomological Society of America, 1969. **62**(6): p. 1321-+.
67. Bohart, G.E., *How to Manage the Alfalfa Leaf-cutting Bee (Megachile rotundata Fabr.) for Alfalfa Pollination*. Utah Agricultural Experiment Station Circular, 1962. **144**.
68. Klostermeyer, E.C., S.J.J. Mech, and W.B. Rasmussen, *Sex and Weight of Megachile rotundata (Hymenoptera: Megachilidae) Progeny Associated with Provision Weights*. Journal of the Kansas Entomological Society, 1973. **46**(4): p. 536-548.
69. Peterson, J. and B. Roitberg, *Impact of resource levels on sex ratio and resource allocation in the solitary bee, Megachile rotundata*. Environmental Entomology, 2006. **35**(5): p. 1404-1410.
70. Johansen, C.A. and J.D. Eves, *Effects of Chilling, Humidity and Seasonal Conditions on Emergence of the Alfalfa Leafcutting Bee*. Environmental entomology., 1973. **2**: p. 23-26.
71. Murrell, D., *Alfalfa Seed and Leafcutting Bee Production in Saskatchewan*. 1997, SASPA Extension Publication.
72. Michener, C.D., *The Bees of the World*. 2007, Baltimore, Maryland, USA.: The John Hopkins University Press.
73. Wcislo, W.T., *Nest localization and recognition in a solitary bee, Lasioglossum (Dialictus) figueresi Wcislo (Hymenoptera, Halictidae), in relation to sociality*. Ethology, 1992. **92**(2): p. 108-123.
74. Guedot, C., et al., *Olfactory cues and nest recognition in the solitary bee Osmia lignaria*. Physiological Entomology, 2006. **31**(2): p. 110-119.
75. Guedot, C., et al., *Nest Marking Behavior and Chemical Composition of Olfactory Cues Involved in Nest Recognition in Megachile rotundata*. Environmental Entomology, 2013. **42**(4): p. 779-789.
76. Klostermeyer, L.E. and C.F.S. Hoo, *Chromosome number in leafcutter bee Megachile rotundata*. Annals of the Entomological Society of America, 1968. **61**(3): p. 782-+.
77. McCorquodale, D.B. and R.E. Owen, *Laying sequence, diploid males and nest usurpation in the leafcutter bee, Megachile rotundata (Hymenoptera, Megachilidae)*. Journal of Insect Behavior, 1994. **7**(5): p. 731-738.
78. Cowan, D.P. and J.K. Stahlhut, *Functionally reproductive diploid and haploid males in an inbreeding hymenopteran with complementary sex determination*. Proceedings of the National Academy of Sciences of the United States of America, 2004. **101**(28): p. 10374-10379.
79. Tepedino, V.J. and F.D. Parker, *Alternation of sex-ratio in a partially bivoltine bee, Megachile rotundata (Hymenoptera, Megachilidae)*. Annals of the Entomological Society of America, 1988. **81**(3): p. 467-476.
80. Pitts-Singer, T. and R. James, *Emergence success and sex ratio of commercial alfalfa leafcutting bees from the United States and Canada*. Journal of Economic Entomology, 2005. **98**(6): p. 1785-1790.
81. Rossi, B.H., P. Nonacs, and T.L. Pitts-Singer, *Sexual harassment by males reduces female fecundity in the alfalfa leafcutting bee, Megachile rotundata*. Animal Behaviour, 2010. **79**(1): p. 165-171.
82. O'Neill, K.M., et al., *Composition of pollen loads of Megachile rotundata in relation to flower diversity (Hymenoptera : Megachilidae)*. Journal of the Kansas Entomological Society, 2004. **77**(4): p. 619-625.
83. Ogilvie, J.E. and J.R.K. Forrest, *Interactions between bee foraging and floral resource phenology shape bee populations and communities*. Current Opinion in Insect Science, 2017. **21**: p. 75-82.
84. Cane, J.H., *Pollinating bees (Hymenoptera : Apiformes) of US alfalfa compared for rates of pod and seed set*. Journal of Economic Entomology, 2002. **95**(1): p. 22-27.

85. Jabbour, R. and S. Noy, *The Promise of a Multi-Disciplinary, Mixed-Methods Approach to Inform Insect Pest Management: Evidence From Wyoming Alfalfa*. Frontiers in Sustainable Food Systems, 2020. **4**.
86. O'Neal, S., *Pest management strategic plan - with a special focus on pollinator protection - for alfalfa seed production in the Western United States. Summary of a workshop held on January 30-31 2017*. 2017, Western Integrated Pest Management Centre. : Las Vegas, NV.
87. Walsh, D., E. Johansen, and S. O'Neal, *Washington state managed pollinator protection plan for alfalfa seed production.*, W.S.U.a.W.S.D.o. Agriculture., Editor. 2017, Washington State University and Washington State Department of Agriculture.
88. Hooven, L., Sagili R. and Johansen E., *How to reduce bee poisoning from pesticides*. 2006, Revised 2013.
89. Zukoff, S., et al., *Alfalfa insect management 2019*, in *K-State Research and Extension*, K.S. University, Editor. 2019, Kansas State University: Kansas USA.
90. Scott-Dupree, C.D., L. Conroy, and C.R. Harris, *Impact of Currently Used or Potentially Useful Insecticides for Canola Agroecosystems on *Bombus impatiens* (Hymenoptera: Apidae), *Megachile rotundata* (Hymenoptera: Megachilidae), and *Osmia lignaria* (Hymenoptera: Megachilidae)*. Journal of Economic Entomology, 2009. **102**(1): p. 177-182.
91. Gradish, A.E., C.D. Scott-Dupree, and G.C. Cutler, *Susceptibility of *Megachile rotundata* to insecticides used in wild blueberry production in Atlantic Canada*. Journal of Pest Science, 2012. **85**(1): p. 133-140.
92. EPA, *Guidance for inspecting alleged cases of pesticide-related bee incidents.*, U.E.P. Agency., Editor. 2013, EPA: Washington DC.
93. Alberta.ca, *How to reduce bee poisonings from pesticides.*, D.o.A.a. Forestry., Editor. 2018, Government of Alberta, Canada.
94. Canada.ca, *Update on Canadian Bee Incident Reports 2012-2016.*, H. Canada., Editor. 2017, Government of Canada.: Canada.
95. USGAO, *United States Government Accountability Office Report to Congressional Requesters: Bee Health. USDA and EPA should take additional actions to address threats to bee populations.*, U. GAO, Editor. 2016, US GAO: Washington DC.
96. Morandin, L., et al., *Practices to reduce bee poisoning from agricultural pesticides in Canada.*, D.o.A.a.A.-F. Canada., Editor. 2018, Pollinator Partnership Canada.: Canada.
97. Johansen, E., *How to reduce bee poisoning from pesticides.*, in *Pacific Northwest Insect Management Handbook*. 2014, PNW: Idaho, US.
98. Westcott, L. and D. Nelson, *Canola pollination: an update*. Bee World, 2001. **82**(3): p. 115-129.
99. Tepedino, V.J., *A comparison of the alfalfa leafcutting bee (*Megachile rotundata*) and the honeybee (*Apis mellifera*) as pollinators for hybrid carrot seed in field cages*. Seventh International Symposium on Pollination - Pollination: from Theory to Practise, 1997(437): p. 457-461.
100. Soroka, J.J., et al., *Alfalfa leaf cutting bee (Hymenoptera : Megachilidae) pollination of oilseed rape (*Brassica napus* L.) under isolation tents for hybrid seed production*. Canadian Journal of Plant Science, 2001. **81**(1): p. 199-204.
101. Richards, K.W., *The alfalfa leafcutter bee, *Megachile rotundata*: a potential pollinator for some annual forage clovers*. Journal of Apicultural Research, 1995. **34**(3): p. 115-121.
102. Richards, K.W., *Effectiveness of the alfalfa leafcutter bee *Megachile rotundata* Fab. to pollinate perennial clovers*. Journal of Apicultural Research, 2016. **55**(3): p. 259-267.
103. Stubbs, C.S. and F.A. Drummond, *Management of the alfalfa leafcutting bee, *Megachile rotundata* (Hymenoptera : Megachilidae), for pollination of wild lowbush blueberry*. Journal of the Kansas Entomological Society, 1997. **70**(2): p. 81-93.

104. Cane, J.H., *An evaluation of pollination mechanisms for purple prairie-clover, Dalea purpurea (Fabaceae : Amorphaea)*. American Midland Naturalist, 2006. **156**(1): p. 193-197.
105. Cane, J.H., *Pollinating bees crucial to farming wildflower seed for US habitat restoration. In Bee pollination in agricultural ecosystems.*, ed. R.R. James and T.L. Pitts-Singer. 2008, New York. USA: Oxford University Press, Inc.
106. Pitts-Singer, T.L., *Past and present management of alfalfa bees. In Bee pollination in agricultural ecosystems.*, ed. R.R. James and T.L. Pitts-Singer. 2008, New York. USA: Oxford University Press, Inc.
107. CIFA, *Alfalfa leafcutting bee producer guide to the national bee farm-level biosecurity standard*. CFIA P0866E-13 ed, ed. C.F.I. Agency. 2013, Canada: CFIA.
108. Calderone, N.W., *Insect Pollinated Crops, Insect Pollinators and US Agriculture: Trend Analysis of Aggregate Data for the Period 1992-2009*. Plos One, 2012. **7**(5).
109. Oerke, E.C., *Crop losses to pests*. Journal of Agricultural Science, 2006. **144**: p. 31-43.
110. Flood, J., *The importance of plant health to food security*. Food Security, 2010. **2**(3): p. 215-231.
111. Beddington, J., *Food security: contributions from science to a new and greener revolution*. Philosophical Transactions of the Royal Society B-Biological Sciences, 2010. **365**(1537): p. 61-71.
112. Hagstrum, D.W. and T.W. Phillips, *Evolution of Stored-Product Entomology: Protecting the World Food Supply*. Annual Review of Entomology, Vol 62, 2017. **62**: p. 379-397.
113. Birch, A.N.E., G.S. Begg, and G.R. Squire, *How agro-ecological research helps to address food security issues under new IPM and pesticide reduction policies for global crop production systems*. Journal of Experimental Botany, 2011. **62**(10): p. 3251-3261.
114. FAOSTAT. 2020 [cited 2020 18/08/20]; Available from: <http://www.fao.org/faostat/en/#data/RP/visualize>.
115. Carvalho, F.P., *Agriculture, pesticides, food security and food safety*. Environmental Science & Policy, 2006. **9**(7-8): p. 685-692.
116. Casida, J.E. and G.B. Quistad, *Golden age of insecticide research: Past, present, or future?* Annual Review of Entomology, 1998. **43**: p. 1-16.
117. Isman, M.B., *Botanical insecticides, deterrents, and repellents in modern agriculture and an increasingly regulated world*. Annual Review of Entomology, 2006. **51**: p. 45-66.
118. Tomizawa, M. and J.E. Casida, *Neonicotinoid insecticide toxicology: Mechanisms of selective action*. Annual Review of Pharmacology and Toxicology, 2005. **45**: p. 247-+.
119. Tomizawa, M. and J.E. Casida, *Molecular Recognition of Neonicotinoid Insecticides: The Determinants of Life or Death*. Accounts of Chemical Research, 2009. **42**(2): p. 260-269.
120. Panini, M., et al., *An overview of the main pathways of metabolic resistance in insects*. Isj-Invertebrate Survival Journal, 2016. **13**: p. 326-335.
121. Jeanguenat, A., *The story of a new insecticidal chemistry class: the diamides*. Pest Management Science, 2013. **69**(1): p. 7-14.
122. Casida, J.E., *Neonicotinoids and Other Insect Nicotinic Receptor Competitive Modulators: Progress and Prospects*. Annual Review of Entomology, Vol 63, 2018. **63**: p. 125-144.
123. Sparks, T.C. and R. Nauen, *IRAC: Mode of action classification and insecticide resistance management*. Pesticide Biochemistry and Physiology, 2015. **121**: p. 122-128.
124. IRAC, *IRAC mode of action classification scheme. Version 9.4*. 2020. p. 1-30.

125. Stephenson, G.R., et al., *Glossary of terms relating to pesticides - (IUPAC recommendations 2006)*. Pure and Applied Chemistry, 2006. **78**(11): p. 2075-2154.
126. Bloomquist, J.R., *Chloride channels as tools for developing selective insecticides*. Archives of Insect Biochemistry and Physiology, 2003. **54**(4): p. 145-156.
127. Ware, G.W. and D.M. Whitacre, *An introduction to insecticides (4th Edition)*. In *The pesticide book*. 6th Edition ed. 2004, Willoughby, Ohio. USA: MeisterPro Information Resources, Meister Media Worldwide.
128. EPA, *New pesticide fact sheet - Fipronil.737-F-96-005*. 1996, US Environmental Protection Agency, Office of Pesticide Programs, US, Government Printing Office, Washington DC: Washington DC USA. p. 1-10.
129. Hainzl, D., L.M. Cole, and J.E. Casida, *Mechanisms for selective toxicity of fipronil insecticide and its sulfone metabolite and desulfinyl photoproduct*. Chemical Research in Toxicology, 1998. **11**(12): p. 1529-1535.
130. Satpathy, S., B.S. Gotyal, and V.R. Babu, *Role of novel insecticides in crop protection and their selectivity to natural enemies: A review*. Journal of Environmental Biology, 2020. **41**(2): p. 149-160.
131. Mayer, D.F. and J.D. Lunden, *Field and laboratory tests of the effects of fipronil on adult female bees of Apis mellifera, Megachile rotundata and Nomia melanderi*. Journal of Apicultural Research, 1999. **38**(3-4): p. 191-197.
132. Li, X.Q., et al., *Toxicities of fipronil enantiomers to the honeybee Apis Mellifera L. and enantiomeric compositions of fipronil in honey plant flowers*. Environmental Toxicology and Chemistry, 2010. **29**(1): p. 127-132.
133. Pisa, L.W., et al., *Effects of neonicotinoids and fipronil on non-target invertebrates*. Environmental Science and Pollution Research, 2015. **22**(1): p. 68-102.
134. Aldridge, W.N. and A.N. Davison, *The inhibition of erythrocyte cholinesterase by tri-esters of phosphoric acid .1. Diethyl P-nitrophenyl phosphate (E600) and analogues*. Biochemical Journal, 1952. **51**(1): p. 62-70.
135. Wilson, I.B., M.A. Hatch, and S. Ginsburg, *CARBAMYLATION OF ACETYLCHOLINESTERASE*. Journal of Biological Chemistry, 1960. **235**(8): p. 2312-2315.
136. Fukuto, T.R., *Mechanism of action of organophosphorus and carbamate insecticides*. Environmental Health Perspectives, 1990. **87**: p. 245-254.
137. Reid, R.J., et al., *Assessing the acute toxicity of insecticides to the buff-tailed bumblebee (Bombus terrestris audax)*. Pesticide Biochemistry and Physiology, 2020. **166**.
138. Johnson, R.M., H.S. Pollock, and M.R. Berenbaum, *Synergistic Interactions Between In-Hive Miticides in Apis mellifera*. Journal of Economic Entomology, 2009. **102**(2): p. 474-479.
139. PubChem. *2D structure image of Chlorpyrifos*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/2730>].
140. PubChem. *2D structure image of Coumaphos*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/2871>].
141. Mansour, S.A. and M.K. Al-Jalili, *Determination of residues of some insecticides in clover flowers - A bioassay method using honeybee adults*. Journal of Apicultural Research, 1985. **24**(3): p. 195-198.
142. Marletto, F., A. Patetta, and A. Manino, *Laboratory assessment of pesticide toxicity to bumblebees*. Bulletin of Insectology, 2003. **56**: p. 155-158.
143. Casida, J.E., *Pyrethrum flowers and pyrethroid insecticides*. Environmental Health Perspectives, 1980. **34**(FEB): p. 189-202.
144. Gajendiran, A. and J. Abraham, *An overview of pyrethroid insecticides*. Frontiers of Biology, 2018. **13**: p. 79-90.
145. Carvalho, S.M., et al., *Enzymatic biomarkers as tools to assess environmental quality: A case study of exposure of the honeybee Apis mellifera to insecticides*. Environmental Toxicology and Chemistry, 2013. **32**(9): p. 2117-2124.

146. Sanchez-Bayo, F. and K. Goka, *Pesticide Residues and Bees - A Risk Assessment*. Plos One, 2014. **9**(4).
147. PubChem. *2D structure image of Deltamethrin*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/40585>].
148. PubChem. *2D structure image of tau-Fluvalinate*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/91768>].
149. PubChem. *2D structure image of alpha-Cypermethrin*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/2912>].
150. Kagabu, S., *Discovery of chloronicotinyl insecticides. Chapter 4: In Nicotinoid insecticides and the nicotinic acetylcholine receptor*. 1999, Tokyo, Japan: Springer
151. Salgado, V.L., *Studies on the mode of action of spinosad: Insect symptoms and physiological correlates*. Pesticide Biochemistry and Physiology, 1998. **60**(2): p. 91-102.
152. Casida, J.E. and K.A. Durkin, *Neuroactive Insecticides: Targets, Selectivity, Resistance, and Secondary Effects*. Annual Review of Entomology, Vol 58, 2013. **58**: p. 99-117.
153. Bass, C. and L.M. Field, *Neonicotinoids*. Current Biology, 2018. **28**(14): p. R772-R773.
154. Feuer, H. and J.P. Lawrence, *Alkyl nitrate nitration of active methylene compounds .6. A new synthesis of alpha-nitroalkyl heterocyclics*. Journal of the American Chemical Society, 1969. **91**(7): p. 1856-&.
155. Iwasa, T., et al., *Mechanism for the differential toxicity of neonicotinoid insecticides in the honey bee, Apis mellifera*. Crop Protection, 2004. **23**(5): p. 371-378.
156. Laurino, D., et al., *Toxicity of neonicotinoid insecticides to honey bees: laboratory tests*. Bulletin of Insectology, 2011. **64**(1): p. 107-113.
157. Tomizawa, M. and J.E. Casida, *Selective toxicity of neonicotinoids attributable to specificity of insect and mammalian nicotinic receptors*. Annual Review of Entomology, 2003. **48**: p. 339-364.
158. Goulson, D., *REVIEW: An overview of the environmental risks posed by neonicotinoid insecticides*. Journal of Applied Ecology, 2013. **50**(4): p. 977-987.
159. PubChem. *2D structure image of Nithiazine*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/3033155>].
160. PubChem. *2D structure image of Thiacloprid*. Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/115224>.
161. PubChem. *2D structure image of Acetamiprid*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/213021>].
162. PubChem. *2D structure image of Imidacloprid*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/86287518>].
163. PubChem. *2D structure image of Clothianidin*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/86287519>].
164. van der Steen, J.J.M., *Review of the methods to determine the hazard and toxicity of pesticides to bumblebees*. Apidologie, 2001. **32**(5): p. 399-406.
165. van der Sluijs, J.P., et al., *Neonicotinoids, bee disorders and the sustainability of pollinator services*. Current Opinion in Environmental Sustainability, 2013. **5**(3-4): p. 293-305.
166. Medrzycki, P., et al., *Standard methods for toxicology research in Apis mellifera*. Journal of Apicultural Research, 2013. **52**(4).
167. Brunet, J.L., A. Badiou, and L.P. Belzunces, *In vivo metabolic fate of C-14 - acetamiprid in six biological compartments of the honeybee, Apis mellifera L.* Pest Management Science, 2005. **61**(8): p. 742-748.
168. Manjon, C., et al., *Unravelling the Molecular Determinants of Bee Sensitivity to Neonicotinoid Insecticides*. Current Biology, 2018. **28**(7): p. 1137-+.
169. Nauen, R., et al., *Whitefly-active metabolites of imidacloprid: biological efficacy and translocation in cotton plants*. Pesticide Science, 1999. **55**(3): p. 265-271.

170. Spurgeon, D., et al., *Chronic oral lethal and sub-lethal toxicities of different binary mixtures of pesticides and contaminants in bees (Apis mellifera, Osmia bicornis and Bombus terrestris)*. 2016, EFSA. p. 1-66.
171. Nauen, R., et al., *Flupyradifurone: a brief profile of a new butenolide insecticide*. Pest Management Science, 2015. **71**(6): p. 850-862.
172. PubChem. *2D structure image of Flupyradifurone*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/16752772>].
173. Chen, X.D., et al., *Risk assessment of various insecticides used for management of Asian citrus psyllid, Diaphorina citri in Florida citrus, against honey bee, Apis mellifera*. Ecotoxicology, 2017. **26**(3): p. 351-359.
174. Lanner, J.T., et al., *Ryanodine Receptors: Structure, Expression, Molecular Details, and Function in Calcium Release*. Cold Spring Harbor Perspectives in Biology, 2010. **2**(11).
175. Lahm, G.P., D. Cordova, and J.D. Barry, *New and selective ryanodine receptor activators for insect control*. Bioorganic & Medicinal Chemistry, 2009. **17**(12): p. 4127-4133.
176. Cordova, D., et al., *Anthranilic diamides: A new class of insecticides with a novel mode of action, ryanodine receptor activation*. Pesticide Biochemistry and Physiology, 2006. **84**(3): p. 196-214.
177. Teixeira, L.A. and J.T. Andaloro, *Diamide insecticides: Global efforts to address insect resistance stewardship challenges*. Pesticide Biochemistry and Physiology, 2013. **106**(3): p. 76-78.
178. Wade, A., et al., *Combined Toxicity of Insecticides and Fungicides Applied to California Almond Orchards to Honey Bee Larvae and Adults*. Insects, 2019. **10**(1).
179. EPA, *Pesticide Ecotoxicity Database (formerly: Environmental Effects Database (EEDB))*, U.E.P. Agency, Editor. 2000: USA.
180. Tohnishi, M., et al., *Flubendiamide, a novel insecticide highly active against lepidopterous insect pests*. Journal of Pesticide Science, 2005. **30**(4): p. 354-360.
181. PubChem. *2D structure image of Flubendiamide*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/124203139>].
182. PubChem. *2D structure image of Chlorantraniliprole*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/11271640>].
183. Wink, M., *Plant-breeding - importance of plant secondary metabolites for protection against pathogens and herbivores*. Theoretical and Applied Genetics, 1988. **75**(2): p. 225-233.
184. Detzel, A. and M. Wink, *Attraction, deterrence or intoxication of bees (Apis mellifera) by plant allelochemicals*. Chemoecology, 1993. **4**(1): p. 8-18.
185. Mao, W.F., et al., *Allelic variation in the Depressaria pastinacella CYP6AB3 protein enhances metabolism of plant allelochemicals by altering a proximal surface residue and potential interactions with cytochrome P450 reductase*. Journal of Biological Chemistry, 2007. **282**(14): p. 10544-10552.
186. Wink, M., *Allelochemical properties or the raison d'etre of Alkaloids*. In *The Alkaloids, Volume 43*, ed. J. Cordell. Vol. 43. 1993, New York: Academic Press.
187. Chomel, M., et al., *Plant secondary metabolites: a key driver of litter decomposition and soil nutrient cycling*. Journal of Ecology, 2016. **104**(6): p. 1527-1541.
188. Stevenson, P.C., S.W. Nicolson, and G.A. Wright, *Plant secondary metabolites in nectar: impacts on pollinators and ecological functions*. Functional Ecology, 2017. **31**(1): p. 65-75.
189. PubChem. *2D structure image of Nicotine*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/89594>].
190. PubChem. *2D structure image of Atropine*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/174174>].
191. PubChem. *2D structure image of Cytisine*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/10235>].

192. PubChem. *2D structure image of Anabasine*. [cited 2020; Available from: <https://pubchem.ncbi.nlm.nih.gov/compound/2181>].
193. Fatur, K. and S. Kreft, *Common anticholinergic solanaceous plants of temperate Europe - A review of intoxications from the literature (1966-2018)*. *Toxicon*, 2020. **177**: p. 52-88.
194. Zatonski, W. and M. Zatonski, *Cytisine versus Nicotine for Smoking Cessation*. *New England Journal of Medicine*, 2015. **372**(11): p. 1072-1072.
195. Lee, S.T., et al., *Relative toxicities and neuromuscular nicotinic receptor agonistic potencies of anabasine enantiomers and anabaseine*. *Neurotoxicology and Teratology*, 2006. **28**(2): p. 220-228.
196. Tabashnik, B.E., et al., *Defining Terms for Proactive Management of Resistance to Bt Crops and Pesticides*. *Journal of Economic Entomology*, 2014. **107**(2): p. 496-507.
197. Melander, A.L., *Can insects become resistant to sprays?* *Journal of Economic Entomology*, 1914. **7**: p. 167-173.
198. Labbe, P., et al., *Chapter 14: Evolution of resistance to insecticide in disease vectors*. In: *Genetics and Evolution of Infectious Disease*. First Edition ed. 2011: Elsevier Inc.
199. Buhler, W. *Introduction to pesticide resistance*. [cited 2020 18/08/20]; Available from: <https://pesticidestewardship.org/resistance/>.
200. Whalon, M., D. Mota-Sanchez, and R.M. Hollingworth, *Analysis of global pesticide resistance in arthropods*. In *Global pesticide resistance in arthropods*. 2008, CABI Head Office, Wallingford, Oxfordshire. UK OX10 8DE: CAB International.
201. Carrasco, D., et al., *Behavioural adaptations of mosquito vectors to insecticide control*. *Current Opinion in Insect Science*, 2019. **34**: p. 48-54.
202. Gong, Y.H. and Q.Y. Diao, *Current knowledge of detoxification mechanisms of xenobiotic in honey bees*. *Ecotoxicology*, 2017. **26**(1): p. 1-12.
203. Glendinning, J.I., *How do herbivorous insects cope with noxious secondary plant compounds in their diet?* *Entomologia Experimentalis Et Applicata*, 2002. **104**(1): p. 15-25.
204. Li, X.C., M.A. Schuler, and M.R. Berenbaum, *Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics*. *Annual Review of Entomology*, 2007. **52**: p. 231-253.
205. Meyer, U.A., *Overview of enzymes of drug metabolism*. *Journal of Pharmacokinetics and Biopharmaceutics*, 1996. **24**(5): p. 449-459.
206. Berenbaum, M.R. and R.M. Johnson, *Xenobiotic detoxification pathways in honey bees*. *Current Opinion in Insect Science*, 2015. **10**: p. 51-58.
207. Misra, J.R., et al., *Transcriptional regulation of xenobiotic detoxification in Drosophila*. *Genes & Development*, 2011. **25**(17): p. 1796-1806.
208. Khan, S., et al., *Mechanism of Insecticide Resistance in Insects/Pests*. *Polish Journal of Environmental Studies*, 2020. **29**(3): p. 2023-2030.
209. Gatehouse, J.A., *Plant resistance towards insect herbivores: a dynamic interaction*. *New Phytologist*, 2002. **156**(2): p. 145-169.
210. Despres, L., J.P. David, and C. Gallet, *The evolutionary ecology of insect resistance to plant chemicals*. *Trends in Ecology & Evolution*, 2007. **22**(6): p. 298-307.
211. Claudianos, C., et al., *A deficit of detoxification enzymes: pesticide sensitivity and environmental response in the honeybee*. *Insect Molecular Biology*, 2006. **15**(5): p. 615-636.
212. Ranson, H., et al., *Evolution of supergene families associated with insecticide resistance*. *Science*, 2002. **298**(5591): p. 179-181.
213. Lampert, E.C., L.A. Dyer, and M.D. Bowers, *Dietary specialization and the effects of plant species on potential multitrophic interactions of three species of nymphaline caterpillars*. *Entomologia Experimentalis Et Applicata*, 2014. **153**(3): p. 207-216.

214. Petschenka, G. and A.A. Agrawal, *How herbivores coopt plant defenses: natural selection, specialization, and sequestration*. *Current Opinion in Insect Science*, 2016. **14**: p. 17-24.
215. Erb, M. and C.A.M. Robert, *Sequestration of plant secondary metabolites by insect herbivores: molecular mechanisms and ecological consequences*. *Current Opinion in Insect Science*, 2016. **14**: p. 8-11.
216. Pittendrigh, B.R., et al., *Chapter 3: Resistance in the Post-Genomics Age*. In: *Insecticide Resistance Management*. 1st Edition ed, ed. D.W. Onstad. 2008: Elsevier Ltd, Academic Press. 30.
217. Berenbaum, M.R., *Postgenomic chemical ecology: From genetic code to ecological interactions*. *Journal of Chemical Ecology*, 2002. **28**(5): p. 873-896.
218. Feyereisen, R., *Insect CYP genes and P450 enzymes*. In *Insect molecular biology and biochemistry*. 2012: Academic Press (Elsevier).
219. Werck-Reichhart, D. and R. Feyereisen, *Cytochromes P450: a success story*. *Genome Biology*, 2000. **1**(6): p. 9.
220. Omura, T. and R. Sato, *The Carbon Monoxide-binding Pigment of Liver Microsomes: I. Evidence for its hemoprotein nature*. *Journal of Biological Chemistry*, 1964. **239**(7): p. 2370-2378.
221. Lee, S.H., et al., *Decreased detoxification genes and genome size make the human body louse an efficient model to study xenobiotic metabolism*. *Insect Molecular Biology*, 2010. **19**(5): p. 599-615.
222. Tijet, N., C. Helvig, and R. Feyereisen, *The cytochrome P450 gene superfamily in Drosophila melanogaster: Annotation, intron-exon organization and phylogeny*. *Gene*, 2001. **262**(1-2): p. 189-198.
223. Arensburger, P., et al., *Sequencing of Culex quinquefasciatus Establishes a Platform for Mosquito Comparative Genomics*. *Science*, 2010. **330**(6000): p. 86-88.
224. Nelson, D.R., et al., *Comparison of cytochrome P450 (CYP) genes from the mouse and human genomes, including nomenclature recommendations for genes, pseudogenes and alternative-splice variants*. *Pharmacogenetics*, 2004. **14**(1): p. 1-18.
225. Zanger, U.M. and M. Schwab, *Cytochrome P450 enzymes in drug metabolism: Regulation of gene expression, enzyme activities, and impact of genetic variation*. *Pharmacology & Therapeutics*, 2013. **138**(1): p. 103-141.
226. Beadle, K., et al., *Genomic insights into neonicotinoid sensitivity in the solitary bee Osmia bicornis*. *Plos Genetics*, 2019. **15**(2).
227. Nebert, D.W., et al., *The P450 gene superfamily - recommended nomenclature*. *DNA-a Journal of Molecular & Cellular Biology*, 1987. **6**(1): p. 1-11.
228. Nelson, D.R., et al., *The P450 superfamily - update on new sequences, gene-mapping, accession numbers, early trivial names of enzymes, and nomenclature*. *DNA and Cell Biology*, 1993. **12**(1): p. 1-51.
229. Nelson, D.R., et al., *P450 superfamily: Update on new sequences, gene mapping, accession numbers and nomenclature*. *Pharmacogenetics*, 1996. **6**(1): p. 1-42.
230. Holt, R.A., et al., *The genome sequence of the malaria mosquito Anopheles gambiae*. *Science*, 2002. **298**(5591): p. 129-+.
231. Ingelman-Sundberg, M., *Genetic polymorphisms of cytochrome P450 2D6 (CYP2D6): clinical consequences, evolutionary aspects and functional diversity*. *Pharmacogenomics Journal*, 2005. **5**(1): p. 6-13.
232. Emerson, J.J., et al., *Natural selection shapes genome-wide patterns of copy-number polymorphism in Drosophila melanogaster*. *Science*, 2008. **320**(5883): p. 1629-1631.
233. Ingelman-Sundberg, M., *Polymorphism of cytochrome P450 and xenobiotic toxicity*. *Toxicology*, 2002. **181**: p. 447-452.
234. Nelson, D.R., *Metazoan cytochrome P450 evolution*. *Comparative Biochemistry and Physiology C-Pharmacology Toxicology & Endocrinology*, 1998. **121**(1-3): p. 15-22.

235. Feyereisen, R., *Evolution of insect P450*. Biochemical Society Transactions, 2006. **34**: p. 1252-1255.
236. Nelson, D.R., *Cytochrome P450 and the individuality of species*. Archives of Biochemistry and Biophysics, 1999. **369**(1): p. 1-10.
237. Cohen, M.B. and R. Feyereisen, *A cluster of cytochrome-P450 genes of the CYP6 family in the house-fly*. DNA and Cell Biology, 1995. **14**(1): p. 73-82.
238. Berge, J.B., R. Feyereisen, and M. Amichot, *Cytochrome P450 monooxygenases and insecticide resistance in insects*. Philosophical Transactions of the Royal Society of London Series B-Biological Sciences, 1998. **353**(1376): p. 1701-1705.
239. Amichot, M., et al., *Point mutations associated with insecticide resistance in the Drosophila cytochrome P450 Cyp6a2 enable DDT metabolism*. European Journal of Biochemistry, 2004. **271**(7): p. 1250-1257.
240. Pittendrigh, B., et al., *Cytochrome P450 genes from Helicoverpa armigera: Expression in a pyrethroid-susceptible and -resistant strain*. Insect Biochemistry and Molecular Biology, 1997. **27**(6): p. 507-512.
241. David, J.P., et al., *The Anopheles gambiae detoxification chip: A highly specific microarray to study metabolic-based insecticide resistance in malaria vectors*. Proceedings of the National Academy of Sciences of the United States of America, 2005. **102**(11): p. 4080-4084.
242. Scharf, M.E., et al., *Expression and induction of three family 4 cytochrome P450 (CYP4)* genes identified from insecticide-resistant and susceptible western corn rootworms, Diabrotica virgifera virgifera*. Insect Molecular Biology, 2001. **10**(2): p. 139-146.
243. Graham, S.E. and J.A. Peterson, *How similar are P450s and what can their differences teach us?* Archives of Biochemistry and Biophysics, 1999. **369**(1): p. 24-29.
244. Gotoh, O., *Substrate recognition sites in cytochrome-P450 family-2 (CYP2) proteins inferred from comparative analyses of amino-acid and coding nucleotide-sequences*. Journal of Biological Chemistry, 1992. **267**(1): p. 83-90.
245. Sezutsu, H., G. Le Goff, and R. Feyereisen, *Origins of P450 diversity*. Philosophical Transactions of the Royal Society B-Biological Sciences, 2013. **368**(1612).
246. Zimmer, C.T., et al., *Neofunctionalization of Duplicated P450 Genes Drives the Evolution of Insecticide Resistance in the Brown Planthopper*. Current Biology, 2018. **28**(2): p. 268-+.
247. Troczka, B.J., et al., *Identification and functional characterisation of a novel N-cyanoamidine neonicotinoid metabolising cytochrome P450, CYP9Q6, from the buff-tailed bumblebee Bombus terrestris*. Insect Biochemistry and Molecular Biology, 2019. **111**.
248. Poulos, T.L., B.C. Finzel, and A.J. Howard, *High-resolution crystal-structure of cytochrome-P450cam*. Journal of Molecular Biology, 1987. **195**(3): p. 687-700.
249. Peterson, J.A. and S.E. Graham, *A close family resemblance: the importance of structure in understanding cytochromes P450*. Structure, 1998. **6**(9): p. 1079-1085.
250. Yano, J.K., et al., *The structure of human microsomal cytochrome P450 3A4 determined by X-ray crystallography to 2.05-angstrom resolution*. Journal of Biological Chemistry, 2004. **279**(37): p. 38091-38094.
251. Baudry, J., S. Rupasinghe, and M.A. Schuler, *Class-dependent sequence alignment strategy improves the structural and functional modeling of P450s*. Protein Engineering Design & Selection, 2006. **19**(8): p. 345-353.
252. Rupasinghe, S., et al., *The cytochrome P450 gene family CYP157 does not contain EXXR in the K-helix reducing the absolute conserved P450 residues to a single cysteine*. Febs Letters, 2006. **580**(27): p. 6338-6342.
253. Schuler, M.A. and M.R. Berenbaum, *Structure and Function of Cytochrome P450S in Insect Adaptation to Natural and Synthetic Toxins: Insights Gained*

- from *Molecular Modeling*. *Journal of Chemical Ecology*, 2013. **39**(9): p. 1232-1245.
254. Kaur, P., et al., *Structure-Based Inhibitor Design for Evaluation of a CYP3A4 Pharmacophore Model*. *Journal of Medicinal Chemistry*, 2016. **59**(9): p. 4210-4220.
 255. Hollingworth, R.M. and K. Dong, *Chapter 3: The Biochemical and molecular genetic basis of resistance to pesticides in arthropods*. In: *Global pesticide resistance in arthropods*. 2008, Wallingford, Oxfordshire. UK: CAB International.
 256. Zhu, Y.C. and R. Luttrell, *Altered gene regulation and potential association with metabolic resistance development to imidacloprid in the tarnished plant bug, *Lygus lineolaris**. *Pest Management Science*, 2015. **71**(1): p. 40-57.
 257. Yang, X., et al., *Glutathione S-transferases are involved in thiamethoxam resistance in the field whitefly *Bemisia tabaci* Q (Hemiptera: Aleyrodidae)*. *Pesticide Biochemistry and Physiology*, 2016. **134**: p. 73-78.
 258. Enayati, A.A., H. Ranson, and J. Hemingway, *Insect glutathione transferases and insecticide resistance*. *Insect Molecular Biology*, 2005. **14**(1): p. 3-8.
 259. Higgins, C.F., *ABC transporters - from microorganisms to man*. *Annual Review of Cell Biology*, 1992. **8**: p. 67-113.
 260. Dermauw, W., et al., *A link between host plant adaptation and pesticide resistance in the polyphagous spider mite *Tetranychus urticae**. *Proceedings of the National Academy of Sciences of the United States of America*, 2013. **110**(2): p. E113-E122.
 261. Buss, D.S. and A. Callaghan, *Interaction of pesticides with p-glycoprotein and other ABC proteins: A survey of the possible importance to insecticide, herbicide and fungicide resistance*. *Pesticide Biochemistry and Physiology*, 2008. **90**(3): p. 141-153.
 262. Lanning, C.L., et al., *Tobacco budworm P-glycoprotein: Biochemical characterization and its involvement in pesticide resistance*. *Biochimica Et Biophysica Acta-General Subjects*, 1996. **1291**(2): p. 155-162.
 263. Epis, S., et al., *ABC transporters are involved in defense against permethrin insecticide in the malaria vector *Anopheles stephensi**. *Parasites & Vectors*, 2014. **7**.
 264. Strycharz, J.P., et al., *Resistance in the highly DDT-resistant 91-R strain of *Drosophila melanogaster* involves decreased penetration, increased metabolism, and direct excretion*. *Pesticide Biochemistry and Physiology*, 2013. **107**(2): p. 207-217.
 265. Hawthorne, D.J. and G.P. Dively, *Killing Them with Kindness? In-Hive Medications May Inhibit Xenobiotic Efflux Transporters and Endanger Honey Bees*. *Plos One*, 2011. **6**(11).
 266. IRAC. *Established insecticide target site mutations (Ver 2.4)*. 2020 [cited 2020; Available from: <https://www.iraconline.org/documents/established-insecticide-target-site-mutations/?ext=xls>.
 267. Balabanidou, V., L. Grigoraki, and J. Vontas, *Insect cuticle: a critical determinant of insecticide resistance*. *Current Opinion in Insect Science*, 2018. **27**: p. 68-74.
 268. Bull, D.L. and R.S. Patterson, *Characterization of pyrethroid resistance in a strain of the German-cockroach (*Dictyoptera, Blattellidae*)*. *Journal of Economic Entomology*, 1993. **86**(1): p. 20-25.
 269. Puinean, A.M., et al., *Amplification of a Cytochrome P450 Gene Is Associated with Resistance to Neonicotinoid Insecticides in the Aphid *Myzus persicae**. *Plos Genetics*, 2010. **6**(6).
 270. Danka, R.G., et al., *Comparative toxicities of 4 topically applied insecticides to Africanized and European honey-bees (*Hymenoptera, Apidae*)*. *Journal of Economic Entomology*, 1986. **79**(1): p. 18-21.
 271. IRAC. *Resistance: Mechanisms*. 2020; Available from: <https://iraconline.org/about/resistance/mechanisms/>.

272. Kessler, S.C., et al., *Bees prefer foods containing neonicotinoid pesticides*. *Nature*, 2015. **521**(7550): p. 74-U145.
273. Liao, L.H., W.Y. Wu, and M.R. Berenbaum, *Behavioral responses of honey bees (*Apis mellifera*) to natural and synthetic xenobiotics in food*. *Scientific Reports*, 2017. **7**.
274. Muth, F. and A.S. Leonard, *A neonicotinoid pesticide impairs foraging, but not learning, in free-flying bumblebees*. *Scientific Reports*, 2019. **9**.
275. Jaffe, B.D., A.N. Lois, and C. Guedot, *Effect of Fungicide on Pollen Foraging by Honeybees (*Hymenoptera: Apidae*) in Cranberry Differs by Fungicide Type*. *Journal of Economic Entomology*, 2019. **112**(1): p. 499-503.
276. Gatton, M.L., et al., *The importance of mosquito behavioural adaptations to malaria control in Africa*. *Evolution*, 2013. **67**(4): p. 1218-1230.
277. Zalucki, M.P. and M.J. Furlong, *Behavior as a mechanism of insecticide resistance: evaluation of the evidence*. *Current Opinion in Insect Science*, 2017. **21**: p. 19-25.
278. Olaya-Arenas, P., M.E. Scharf, and I. Kaplan, *Do pollinators prefer pesticide-free plants? An experimental test with monarchs and milkweeds*. *Journal of Applied Ecology*, 2020. **57**(10): p. 2019-2030.
279. Klein, S., et al., *Why Bees Are So Vulnerable to Environmental Stressors*. *Trends in Ecology & Evolution*, 2017. **32**(4): p. 268-278.
280. Schmehl, D.R., et al., *Genomic analysis of the interaction between pesticide exposure and nutrition in honey bees (*Apis mellifera*)*. *Journal of Insect Physiology*, 2014. **71**: p. 177-190.
281. du Rand, E.E., et al., *Detoxification mechanisms of honey bees (*Apis mellifera*) resulting in tolerance of dietary nicotine*. *Scientific Reports*, 2015. **5**.
282. Johnson, R.M., et al., *Mediation of pyrethroid insecticide toxicity to honey bees (*Hymenoptera : Apidae*) by cytochrome P450 monooxygenases*. *Journal of Economic Entomology*, 2006. **99**(4): p. 1046-1050.
283. Johnson, R.M., et al., *Acaricide, Fungicide and Drug Interactions in Honey Bees (*Apis mellifera*)*. *Plos One*, 2013. **8**(1).
284. Mao, W.F., M.A. Schuler, and M.R. Berenbaum, *CYP9Q-mediated detoxification of acaricides in the honey bee (*Apis mellifera*)*. *Proceedings of the National Academy of Sciences of the United States of America*, 2011. **108**(31): p. 12657-12662.
285. Ranson, H., et al., *Identification of a novel class of insect glutathione S-transferases involved in resistance to DDT in the malaria vector *Anopheles gambiae**. *Biochemical Journal*, 2001. **359**: p. 295-304.
286. Johnson, R.M., et al., *Ecologically Appropriate Xenobiotics Induce Cytochrome P450s in *Apis mellifera**. *Plos One*, 2012. **7**(2).
287. Derecka, K., et al., *Transient Exposure to Low Levels of Insecticide Affects Metabolic Networks of Honeybee Larvae*. *Plos One*, 2013. **8**(7).
288. Niu, G.D., R.M. Johnson, and M.R. Berenbaum, *Toxicity of mycotoxins to honeybees and its amelioration by propolis*. *Apidologie*, 2011. **42**(1): p. 79-87.
289. Losey, J.E. and M. Vaughan, *The economic value of ecological services provided by insects*. *Bioscience*, 2006. **56**(4): p. 311-323.
290. Pilling, E.D., et al., *Mechanism of synergism between the pyrethroid insecticide lambda-Cyhalothrin and the imidazole fungicide prochloraz, in the honeybee (*Apis mellifera* L)*. *Pesticide Biochemistry and Physiology*, 1995. **51**(1): p. 1-11.
291. Boncristiani, H., et al., *Direct effect of acaricides on pathogen loads and gene expression levels in honey bees *Apis mellifera**. *Journal of Insect Physiology*, 2012. **58**(5): p. 613-620.
292. Mao, W.F., et al., *Quercetin-metabolizing CYP6AS enzymes of the pollinator *Apis mellifera* (*Hymenoptera: Apidae*)*. *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology*, 2009. **154**(4): p. 427-434.
293. Mao, W., M.A. Schuler, and M.R. Berenbaum, *Honey constituents up-regulate detoxification and immunity genes in the western honey bee *Apis mellifera**.

- Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(22): p. 8842-8846.
294. Gill, R.J., O. Ramos-Rodriguez, and N.E. Raine, *Combined pesticide exposure severely affects individual- and colony-level traits in bees*. *Nature*, 2012. **491**(7422): p. 105-U119.
 295. Potts, S.G., et al., *Declines of managed honey bees and beekeepers in Europe*. *Journal of Apicultural Research*, 2010. **49**(1): p. 15-22.
 296. Ollerton, J., et al., *Extinctions of aculeate pollinators in Britain and the role of large-scale agricultural changes*. *Science*, 2014. **346**(6215): p. 1360-1362.
 297. Gerlach, J., et al., *Prioritizing non-marine invertebrate taxa for Red Listing*. *Journal of Insect Conservation*, 2014. **18**(4): p. 573-586.
 298. Hallmann, C.A., et al., *More than 75 percent decline over 27 years in total flying insect biomass in protected areas*. *Plos One*, 2017. **12**(10).
 299. Lister, B.C. and A. Garcia, *Climate-driven declines in arthropod abundance restructure a rainforest food web*. *Proceedings of the National Academy of Sciences of the United States of America*, 2018. **115**(44): p. E10397-E10406.
 300. Sanchez-Bayo, F. and K.A.G. Wyckhuys, *Worldwide decline of the entomofauna: A review of its drivers*. *Biological Conservation*, 2019. **232**: p. 8-27.
 301. Gallai, N., et al., *Economic valuation of the vulnerability of world agriculture confronted with pollinator decline*. *Ecological Economics*, 2009. **68**(3): p. 810-821.
 302. Kremen, C., *The value of pollinator species diversity*. *Science*, 2018. **359**(6377): p. 741-742.
 303. Marshman, J., A. Blay-Palmer, and K. Landman, *Anthropocene Crisis: Climate Change, Pollinators, and Food Security*. *Environments*, 2019. **6**(2).
 304. Potts, S.G., et al., *Global pollinator declines: trends, impacts and drivers*. *Trends in Ecology & Evolution*, 2010. **25**(6): p. 345-353.
 305. Vanbergen, A.J., et al., *Threats to an ecosystem service: pressures on pollinators*. *Frontiers in Ecology and the Environment*, 2013. **11**(5): p. 251-259.
 306. Pellmyr, O., *Evolution of insect pollination and angiosperm diversification*. *Trends in Ecology & Evolution*, 1992. **7**(2): p. 46-49.
 307. Danforth, B.N., et al., *The history of early bee diversification based on five genes plus morphology*. *Proceedings of the National Academy of Sciences of the United States of America*, 2006. **103**(41): p. 15118-15123.
 308. Cardinal, S. and B.N. Danforth, *Bees diversified in the age of eudicots*. *Proceedings of the Royal Society B-Biological Sciences*, 2013. **280**(1755).
 309. Kaiser-Bunbury, C.N., et al., *The robustness of pollination networks to the loss of species and interactions: a quantitative approach incorporating pollinator behaviour*. *Ecology Letters*, 2010. **13**(4): p. 442-452.
 310. Papanikolaou, A.D., et al., *Wild bee and floral diversity co-vary in response to the direct and indirect impacts of land use*. *Ecosphere*, 2017. **8**(11).
 311. Koh, I., et al., *Modeling the status, trends, and impacts of wild bee abundance in the United States*. *Proceedings of the National Academy of Sciences of the United States of America*, 2016. **113**(1): p. 140-145.
 312. Phillips, B.B., et al., *Road verges support pollinators in agricultural landscapes, but are diminished by heavy traffic and summer cutting*. *Journal of Applied Ecology*, 2019. **56**(10): p. 2316-2327.
 313. Ramer, H. and K.C. Nelson, *Applying 'action situation' concepts to public land managers' perceptions of flowering bee lawns in urban parks*. *Urban Forestry & Urban Greening*, 2020. **53**.
 314. Phillips, B.B., et al., *Enhancing road verges to aid pollinator conservation: A review*. *Biological Conservation*, 2020. **250**.
 315. Ritchie, H. and M. Roser, *Land use*. 2013: OurWorldInData.org.
 316. Foley, J.A., et al., *Solutions for a cultivated planet*. *Nature*, 2011. **478**(7369): p. 337-342.

317. Wood, S., K. Sebastian, and S.J. Scherr, *Pilot analysis of global ecosystems (PAGE): Agroecosystems*. 2000, Washington, DC USA: International Food Policy Research Institute and World Resources Institute.
318. Matson, P.A., et al., *Agricultural intensification and ecosystem properties*. *Science*, 1997. **277**(5325): p. 504-509.
319. Williams, N.M., et al., *Ecological and life-history traits predict bee species responses to environmental disturbances*. *Biological Conservation*, 2010. **143**(10): p. 2280-2291.
320. Kline, O. and N.K. Joshi, *Mitigating the Effects of Habitat Loss on Solitary Bees in Agricultural Ecosystems*. Agriculture-Basel, 2020. **10**(4).
321. Gathmann, A. and T. Tschamtko, *Foraging ranges of solitary bees*. *Journal of Animal Ecology*, 2002. **71**(5): p. 757-764.
322. Peterson, J.H. and B.D. Roitberg, *Variable flight distance to resources results in changing sex allocation decisions, *Megachile rotundata**. *Behavioral Ecology and Sociobiology*, 2016. **70**(2): p. 247-253.
323. Brittain, C.A., et al., *Impacts of a pesticide on pollinator species richness at different spatial scales*. *Basic and Applied Ecology*, 2010. **11**(2): p. 106-115.
324. Johansen, C.A., *Pesticides and pollinators*. *Annual Review of Entomology*, 1977. **22**: p. 177-192.
325. Fletcher, M. and L. Barnett, *Bee pesticide poisoning incidents in the United Kingdom*. *Bulletin of Insectology*, 2003. **56**: p. 141-145.
326. Barnett, E.A., A.J. Charlton, and M.R. Fletcher, *Incidents of bee poisoning with pesticides in the United Kingdom, 1994-2003*. *Pest Management Science*, 2007. **63**(11): p. 1051-1057.
327. SASA. *Pesticide poisoning of animals*. 2006- 2019 26/09/2020]; Available from: <https://www.sasa.gov.uk/animal-poisoning-reports>.
328. Blane, E., *Wildlife Incident Investigation Scheme (WIIS) Pesticide poisoning of bees.*, A. Hayward, Editor. 2020, Natural England.
329. Brasse, D., *Poisoning incidents involving honeybees in Germany (1999-2002) and new problems for beekeeping*. *Bulletin of Insectology*, 2003. **56**: p. 199.
330. WIIS, *Pesticides poisoning of animals*. 2006-2007, Wildlife Incident Investigation Scheme - HSE: UK.
331. Mullin, C.A., et al., *High Levels of Miticides and Agrochemicals in North American Apiaries: Implications for Honey Bee Health*. *Plos One*, 2010. **5**(3).
332. Wu, J.Y., C.M. Anelli, and W.S. Sheppard, *Sub-Lethal Effects of Pesticide Residues in Brood Comb on Worker Honey Bee (*Apis mellifera*) Development and Longevity*. *Plos One*, 2011. **6**(2).
333. Wood, T.J. and D. Goulson, *The environmental risks of neonicotinoid pesticides: a review of the evidence post 2013*. *Environmental Science and Pollution Research*, 2017. **24**(21): p. 17285-17325.
334. Dively, G.P. and A. Kamel, *Insecticide Residues in Pollen and Nectar of a Cucurbit Crop and Their Potential Exposure to Pollinators*. *Journal of Agricultural and Food Chemistry*, 2012. **60**(18): p. 4449-4456.
335. Botias, C., et al., *Neonicotinoid Residues in Wildflowers, a Potential Route of Chronic Exposure for Bees*. *Environmental Science & Technology*, 2015. **49**(21): p. 12731-12740.
336. Krupke, C.H., et al., *Multiple Routes of Pesticide Exposure for Honey Bees Living Near Agricultural Fields*. *Plos One*, 2012. **7**(1).
337. Bishop, C.A., et al., *Determination of neonicotinoids and butenolide residues in avian and insect pollinators and their ambient environment in Western Canada (2017,2018)*. *Science of the Total Environment*, 2020. **737**.
338. European Food Safety, A., *Conclusion on the peer review of the pesticide risk assessment of the active substance clothianidin*. *EFSA Journal*, 2013. **11**(3066).
339. European Food Safety, A., *Conclusion on the peer review of the pesticide risk assessment of the active substance imidacloprid*. *EFSA Journal*, 2013. **11**(3068).

340. European Food Safety, A., *Conclusion on the peer review of the pesticide risk assessment of the active substance thiamethoxam*. EFSA Journal, 2013. **11**(3067).
341. Kathage, J., et al., *The impact of restrictions on neonicotinoid and fipronil insecticides on pest management in maize, oilseed rape and sunflower in eight European Union regions*. Pest Management Science, 2018. **74**(1): p. 88-99.
342. Fent, K., M. Schmid, and V. Christen, *Global transcriptome analysis reveals relevant effects at environmental concentrations of cypermethrin in honey bees (Apis mellifera)*. Environmental Pollution, 2020. **259**.
343. Jactel, H., et al., *Alternatives to neonicotinoids*. Environment International, 2019. **129**: p. 423-429.
344. EPA, *Pesticide Ecotoxicity Database [Formerly: Environmental Effects Database (EEDB)] Fate and Effects Division*. 2000: Washington DC. USA.
345. Thompson, H.M., *Assessing the exposure and toxicity of pesticides to bumblebees (Bombus sp.)*. Apidologie, 2001. **32**(4): p. 305-321.
346. Carrasco-Letelier, L., Y. Mendoza-Spina, and M.B. Branchiccela, *Acute contact toxicity test of insecticides (Cipermetrina 25, Lorsban 48E, Thionex 35) on honeybees in the southwestern zone of Uruguay*. Chemosphere, 2012. **88**(4): p. 439-444.
347. Botias, C., et al., *Multiple stressors interact to impair the performance of bumblebee (Bombus terrestris) colonies*. Journal of Animal Ecology, 2020.
348. NAS, *National Academy of Sciences. Insect-Pest Management and Control*. 1969, Washington D.C. USA: The National Academies Press.
349. DAERA, *Integrated pest management guide*. 2017.
350. IPCC, *IPCC, 2014: Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Core Writing Group, R.K.Pacauri and L.A. Meyer (eds)]*. 2014, IPCC: Geneva, Switzerland. p. 151pp.
351. Heller, N.E. and E.S. Zavaleta, *Biodiversity management in the face of climate change: A review of 22 years of recommendations*. Biological Conservation, 2009. **142**(1): p. 14-32.
352. Jentsch, A., J. Kreyling, and C. Beierkuhnlein, *A new generation of climate-change experiments: events, not trends*. Frontiers in Ecology and the Environment, 2007. **5**(7): p. 365-374.
353. Scheffer, M. and S.R. Carpenter, *Catastrophic regime shifts in ecosystems: linking theory to observation*. Trends in Ecology & Evolution, 2003. **18**(12): p. 648-656.
354. Burhan, A., et al., *Environmental policies to protect pollinators: attributes and actions needed to avert climate borne crisis of oil seed agriculture in Pakistan*. Aims Agriculture and Food, 2017. **2**(3): p. 233-250.
355. Le Conte, Y. and M. Navajas, *Climate change: impact on honeybee populations and diseases*. Revue scientifique et technique (International Office of Epizootics), 2008. **27**: p. 499-510.
356. Giannini, T.C., et al., *Climate change in the Eastern Amazon: crop-pollinator and occurrence-restricted bees are potentially more affected*. Regional Environmental Change, 2020. **20**(1).
357. Menendez, R., et al., *Species richness changes lag behind climate change*. Proceedings of the Royal Society B-Biological Sciences, 2006. **273**(1593): p. 1465-1470.
358. Bale, J.S., et al., *Herbivory in global climate change research: direct effects of rising temperature on insect herbivores*. Global Change Biology, 2002. **8**(1): p. 1-16.
359. Frazier, M.R., R.B. Huey, and D. Berrigan, *Thermodynamics constrains the evolution of insect population growth rates: "Warmer is better"*. American Naturalist, 2006. **168**(4): p. 512-520.
360. Lehmann, P., et al., *Complex responses of global insect pests to climate warming*. Frontiers in Ecology and the Environment, 2020. **18**(3): p. 141-149.

361. Boggs, C.L., *The fingerprints of global climate change on insect populations*. Current Opinion in Insect Science, 2016. **17**: p. 69-73.
362. Heinrich, B. and H. Esch, *Thermoregulation in bees*. American Scientist, 1994. **82**(2): p. 164-170.
363. Sun, B.J., et al., *Nocturnal dispersal flight of crickets: Behavioural and physiological responses to cool environmental temperatures*. Functional Ecology, 2020. **34**(9): p. 1907-1920.
364. Gonzalez-Tokman, D., et al., *Insect responses to heat: physiological mechanisms, evolution and ecological implications in a warming world*. Biological Reviews, 2020. **95**(3): p. 802-821.
365. Heinrich, B., *Thermoregulation in bumblebees .2. Energetics of warm-up and free flight*. Journal of Comparative Physiology, 1975. **96**(2): p. 155-166.
366. Heinrich, B., *How the honey bee regulates its body temperature*. Bee World, 1996. **77**(3): p. 130-137.
367. Gallego, B., J.R. Verdu, and J.M. Lobo, *Comparative thermoregulation between different species of dung beetles (Coleoptera: Geotrupinae)*. Journal of Thermal Biology, 2018. **74**: p. 84-91.
368. Deutsch, C.A., et al., *Impacts of climate warming on terrestrial ectotherms across latitude*. Proceedings of the National Academy of Sciences of the United States of America, 2008. **105**(18): p. 6668-6672.
369. Chen, I.C., et al., *Rapid Range Shifts of Species Associated with High Levels of Climate Warming*. Science, 2011. **333**(6045): p. 1024-1026.
370. Sunday, J.M., A.E. Bates, and N.K. Dulvy, *Thermal tolerance and the global redistribution of animals*. Nature Climate Change, 2012. **2**(9): p. 686-690.
371. Kerr, J.T., et al., *Climate change impacts on bumblebees converge across continents*. Science, 2015. **349**(6244): p. 177-180.
372. Parmesan, C., et al., *Poleward shifts in geographical ranges of butterfly species associated with regional warming*. Nature, 1999. **399**(6736): p. 579-583.
373. Parmesan, C. and G. Yohe, *A globally coherent fingerprint of climate change impacts across natural systems*. Nature, 2003. **421**(6918): p. 37-42.
374. Ho, W.C. and J.Z. Zhang, *Evolutionary adaptations to new environments generally reverse plastic phenotypic changes*. Nature Communications, 2018. **9**.
375. Parmesan, C., *Ecological and evolutionary responses to recent climate change*. Annual Review of Ecology Evolution and Systematics, 2006. **37**: p. 637-669.
376. Kingsolver, J.G. and L.B. Buckley, *How do phenology, plasticity, and evolution determine the fitness consequences of climate change for montane butterflies?* Evolutionary Applications, 2018. **11**(8): p. 1231-1244.
377. Renner, S.S. and C.M. Zohner, *Climate Change and Phenological Mismatch in Trophic Interactions Among Plants, Insects, and Vertebrates*. Annual Review of Ecology, Evolution, and Systematics, Vol 49, 2018. **49**: p. 165-182.
378. Kearney, M.R., et al., *Early emergence in a butterfly causally linked to anthropogenic warming*. Biology Letters, 2010. **6**(5): p. 674-677.
379. Gallinat, A.S., R.B. Primack, and D.L. Wagner, *Autumn, the neglected season in climate change research*. Trends in Ecology & Evolution, 2015. **30**(3): p. 169-176.
380. Frund, J., S.L. Zieger, and T. Tschardt, *Response diversity of wild bees to overwintering temperatures*. Oecologia, 2013. **173**(4): p. 1639-1648.
381. Stemkovski, M., et al., *Bee phenology is predicted by climatic variation and functional traits*. Ecology Letters, 2020. **23**(11): p. 1589-1598.
382. Memmott, J., et al., *Global warming and the disruption of plant-pollinator interactions*. Ecology Letters, 2007. **10**(8): p. 710-717.
383. de Jong, P.W. and P.M. Brakefield, *Climate and change in clines for melanism in the two-spot ladybird, Adalia bipunctata (Coleoptera : Coccinellidae)*. Proceedings of the Royal Society B-Biological Sciences, 1998. **265**(1390): p. 39-43.

384. Bradshaw, W.E. and C.M. Holzapfel, *Genetic shift in photoperiodic response correlated with global warming*. Proceedings of the National Academy of Sciences of the United States of America, 2001. **98**(25): p. 14509-14511.
385. Rosenkranz, P., P. Aumeier, and B. Ziegelmann, *Biology and control of Varroa destructor*. Journal of Invertebrate Pathology, 2010. **103**: p. S96-S119.
386. Furst, M.A., et al., *Disease associations between honeybees and bumblebees as a threat to wild pollinators*. Nature, 2014. **506**(7488): p. 364-+.
387. Tehel, A., M.J.F. Brown, and R.J. Paxton, *Impact of managed honey bee viruses on wild bees*. Current Opinion in Virology, 2016. **19**: p. 16-22.
388. Fisher, M.C., et al., *Emerging fungal threats to animal, plant and ecosystem health*. Nature, 2012. **484**(7393): p. 186-194.
389. Cunningham, A.A., P. Daszak, and J.L.N. Wood, *One Health, emerging infectious diseases and wildlife: two decades of progress? Philosophical Transactions of the Royal Society B-Biological Sciences*, 2017. **372**(1725).
390. Grozinger, C.M. and M.L. Flenniken, *Bee Viruses: Ecology, Pathogenicity, and Impacts*. Annual Review of Entomology, Vol 64, 2019. **64**: p. 205-226.
391. Ravoet, J., et al., *Widespread occurrence of honey bee pathogens in solitary bees*. Journal of Invertebrate Pathology, 2014. **122**: p. 55-58.
392. Vandenberg, J.D. and W.P. Stephen, *Etiology and symptomatology of chalkbrood in the alfalfa leafcutting bee, Megachile rotundata*. Journal of Invertebrate Pathology, 1982. **39**(2): p. 133-137.
393. Johansen, C.A., et al., *Pesticides and bees*. Environmental Entomology, 1983. **12**(5): p. 1513-1518.
394. Devillers, J., et al., *Comparative toxicity and hazards of pesticides to APIS and non-APIS bees. A chemometrical study*. Sar and Qsar in Environmental Research, 2003. **14**(5-6): p. 389-403.
395. Thompson, H., *Extrapolation of Acute Toxicity Across Bee Species*. Integrated Environmental Assessment and Management, 2016. **12**(4): p. 622-626.
396. Fischer, D.M. and T. Moriarty, *Pesticide risk assessment for pollinators: Summary of a SETAC Pellston Workshop*. 2011, Pensacola FL USA: Society of Environmental Toxicology and Chemistry (SETAC).
397. Lucchetti, M.A., et al., *Nursing protects honeybee larvae from secondary metabolites of pollen*. Proceedings of the Royal Society B-Biological Sciences, 2018. **285**(1875).
398. Babendreier, D., et al., *Pollen consumption in honey bee larvae: a step forward in the risk assessment of transgenic plants*. Apidologie, 2004. **35**(3): p. 293-300.
399. Hedtke, S.M., S. Patiny, and B.N. Danforth, *The bee tree of life: a supermatrix approach to apoid phylogeny and biogeography*. BMC Evolutionary Biology, 2013. **13**.
400. Branstetter, M.G., et al., *Phylogenomic Insights into the Evolution of Stinging Wasps and the Origins of Ants and Bees*. Current Biology, 2017. **27**(7): p. 1019-1025.
401. Shahan, T., *Photo image: Western Honey Bee (Apis mellifera)(Photo)*. 2020, Oregon Department of Agriculture.: <https://www.inaturalist.org/taxa/47219-Apis-mellifera>.
402. Isselee, E.P., *Photo image: Buff-tailed bumblebee, Bombus terrestris*. 2020: https://www.123rf.com/photo_25987523_buff-tailed-bumblebee-bombus-terrestris-isolated-on-white.html.
403. Hoyer, R., *Photo image: Melipona quadrifasciata*. 2018: <https://www.flickr.com/photos/birdernaturalist/30543062937>.
404. Harleston, C., *Photo image: Mexican Orchid Bee (Eufriesea mexicana)*. 2020, <https://www.inaturalist.org/taxa/307525-Eufriesea-mexicana>.
405. Yurika, A., *Photo image: Bumble bee - Habropoda laboriosa - male*. 2013, <https://bugguide.net/node/view/750808>.
406. Woodmore, L., *Photo image: Alfalfa Leafcutter Bee (Megachile rotundata)*. 2006, <https://davesgarden.com/guides/bf/showimage/742/#b>.

407. Early, J., *Photo image: Female Osmia rufa in flight carrying pollen, Reigate, Surrey.* 2020, <https://www.bwars.com/bee/megachilidae/osmia-bicornis>.
408. Early, J., *Photo image: Female Heriades truncorum flying towards nest in cut reed carrying pollen, Reigate, Surrey.* 2020, <https://www.bwars.com/bee/megachilidae/heriades-truncorum>.
409. Early, J., *Photo image: Andrena cineraria female, Reigate, Surrey.* 2020, <https://www.bwars.com/bee/andrenidae/andrena-cineraria>.
410. Early, J., *Photo image: Melitta haemorrhoidalis* http://www.natureconservationimaging.com/Pages/nature_conservation_imagin_g_download1_bees.php. 2020, Nature Conservation Imaging.
411. Westrich, P., *Photo image: Dioxys cincta.* 2020, https://www.wildbienen.info/steckbriefe/fotos/dioxys_cincta_m_09_3885.jpg.
412. Dvorak, J., *Photo image: Chelostoma florissomne.* 2020, <https://www.biolib.cz/en/image/id154199/>.
413. Dvorak, J., *Photo image: Hylaeus variegatus.* 2020, <https://www.biolib.cz/en/image/id158318/>.
414. Vereecken, N.J., *Photo image: Colletes cunicularis.* 2020, <https://alchetron.com/cdn/colletes-c826cb59-fd13-43fe-a712-8bcc92c32db-resize-750.jpeg>.
415. Packer, L., *Photo image: Ctenocolletes smaragdinus Male.* 2014, https://www.discoverlife.org/mp/20p?see=I_YORKU544&res=640.
416. Life, D., *Photo image: Stenotritus pubescens, F, side, Australia.* 2020, https://www.discoverlife.org/mp/20p?see=I_SD14801&res=640: Discover life.
417. Dumesh, S. and C.S. Sheffield, *Photo image: Dufourea novaeangliae (Robertson, 1897).* 2020, Canadian Journal of Arthropod Identification.: Key to the *Dufourea* of Canada.
418. Gray, K., *Photo image: Alkali bee Nomia melanderi Cockerell Female adult.* 2020, <https://pnwhandbooks.org/insect/bee-protection/alkali-bee>.
419. Fritzsche, M. and M. Fritzsche, *Photo image: Gelbbein-Furchenbiene (Lasioglossum xanthopus).* 2015, Beekeeping Association Dresden V.
420. Fritzsche, M. and M. Fritzsche, *Photo image: Panurgus calcaratus, female, Teufelsmauer (Harz), July 2016.* 2016, Beekeeping Association Dresden V.
421. Commons, C., *Photo image: Camptopoeum sacrum.* 2016, Creative commons. Bold Systems.
422. De Greef, S., *Photo image: Macropis fulvipes* <https://www.facebook.com/WildBnB/photos/en-while-were-into-macropis-bees-there-are-two-species-belonging-to-that-genus-i/681519915656038>. 2019, Wildbnb: Facebook.
423. De Greef, S., *Photo image: Dasygoda hirtipes (Melittidae) female* <https://twitter.com/hashtag/BeesOfBrussels?src=hash>. 2019, Wildbnb: Twitter.
424. Ascher, J.S. and J. Pickering. *Discover Life bee species guide and world checklist (Hymenoptera: Apoidea: Anthophila).* 2017; Available from: https://www.discoverlife.org/mp/20q?guide=Apoidea_species&flags=HAS:.
425. Treangen, T.J. and S.L. Salzberg, *Repetitive DNA and next-generation sequencing: computational challenges and solutions.* Nature Reviews Genetics, 2012. **13**(1): p. 36-46.
426. Uricaru, R., et al., *YOC, A new strategy for pairwise alignment of collinear genomes.* BMC Bioinformatics, 2015. **16**.
427. Liu, D., M. Hunt, and I.J. Tsai, *Inferring synteny between genome assemblies: a systematic evaluation.* BMC Bioinformatics, 2018. **19**.
428. Ohno, S., U. Wolf, and N.B. Atkin, *Evolution from fish to mammals by gene duplication.* Hereditas-Genetiskt Arkiv, 1968. **59**(1): p. 169-+.
429. Zhang, J.Z., *Evolution by gene duplication: an update.* Trends in Ecology & Evolution, 2003. **18**(6): p. 292-298.
430. Zdobnov, E.M. and P. Bork, *Quantification of insect genome divergence.* Trends in Genetics, 2007. **23**(1): p. 16-20.

431. Chen, X.Y. and M. Tompa, *Comparative assessment of methods for aligning multiple genome sequences*. Nature Biotechnology, 2010. **28**(6): p. 567-U53.
432. Ehrlich, J., D. Sankoff, and J.H. Nadeau, *Synteny conservation and chromosome rearrangements during mammalian evolution*. Genetics, 1997. **147**(1): p. 289-296.
433. Renwick, J.H., *Mapping of human chromosomes*. Annual Review of Genetics, 1971. **5**: p. 81-+.
434. Thomasova, D., et al., *Comparative genomic analysis in the region of a major Plasmodium-refractoriness locus of Anopheles gambiae*. Proceedings of the National Academy of Sciences of the United States of America, 2002. **99**(12): p. 8179-8184.
435. Nadeau, J.H. and B.A. Taylor, *Lengths of chromosomal segments conserved since divergence of man and mouse*. Proceedings of the National Academy of Sciences of the United States of America-Biological Sciences, 1984. **81**(3): p. 814-818.
436. Chipman, A.D., et al., *The First Myriapod Genome Sequence Reveals Conservative Arthropod Gene Content and Genome Organisation in the Centipede Strigamia maritima*. Plos Biology, 2014. **12**(11).
437. Zdobnov, E.M., et al., *Comparative genome and proteome analysis of anopheles gambiae and Drosophila melanogaster*. Science, 2002. **298**(5591): p. 149-159.
438. Hall, B., *Phylogenetic trees made easy: A how-to manual*. 4th edition ed. 2011, Sunderland, MA 01375 USA: Sinauer Associates, Inc.
439. Philippe, H., et al., *Resolving Difficult Phylogenetic Questions: Why More Sequences Are Not Enough*. Plos Biology, 2011. **9**(3).
440. Yang, Z.H. and B. Rannala, *Molecular phylogenetics: principles and practice*. Nature Reviews Genetics, 2012. **13**(5): p. 303-314.
441. Goodman, M., et al., *Fitting the gene lineage into its species lineage, a parsimony strategy illustrated by cladograms constructed from globin sequences*. Systematic Zoology, 1979. **28**(2): p. 132-163.
442. Abascal, F., R. Zardoya, and D. Posada, *ProtTest: selection of best-fit models of protein evolution*. Bioinformatics, 2005. **21**(9): p. 2104-2105.
443. Yang, Z., *Computational Molecular Evolution*. 2006, Oxford, OX2 6DP: Oxford University Press.
444. Guindon, S. and O. Gascuel, *A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood*. Systematic Biology, 2003. **52**(5): p. 696-704.
445. Huelsenbeck, J.P. and F. Ronquist, *MRBAYES: Bayesian inference of phylogenetic trees*. Bioinformatics, 2001. **17**(8): p. 754-755.
446. Nauen, R., U. Ebbinghaus-Kintscher, and R. Schmuck, *Toxicity and nicotinic acetylcholine receptor interaction of imidacloprid and its metabolites in Apis mellifera (Hymenoptera : Apidae)*. Pest Management Science, 2001. **57**(7): p. 577-586.
447. Hayward, A., et al., *The leafcutter bee, Megachile rotundata, is more sensitive to N-cyanoamidine neonicotinoid and butenolide insecticides than other managed bees*. Nature Ecology & Evolution, 2019. **3**(11): p. 1521-1524.
448. Singh, K.S., *Methods used in NCBI interrogation: G1, G2 and G3*. 2021.
449. Emms, D.M. and S. Kelly, *OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy*. Genome Biology, 2015. **16**.
450. Altschul, S.F., et al., *Basic Local Alignment Search Tool*. Journal of Molecular Biology, 1990. **215**(3): p. 403-410.
451. Edgar, R.C., *MUSCLE: multiple sequence alignment with high accuracy and high throughput*. Nucleic Acids Research, 2004. **32**(5): p. 1792-1797.
452. Jukes, T.H. and C.R. Cantor, *Evolution of protein molecules*. Mammalian protein metabolism., ed. H.M. Munro. Vol. 3. 1969, New York: Academic Press.

453. Guindon, S., et al., *New Algorithms and Methods to Estimate Maximum-Likelihood Phylogenies: Assessing the Performance of PhyML 3.0*. Systematic Biology, 2010. **59**(3): p. 307-321.
454. Le, S.Q. and O. Gascuel, *An improved general amino acid replacement matrix*. Molecular Biology and Evolution, 2008. **25**(7): p. 1307-1320.
455. Kumar, S., et al., *MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms*. Molecular Biology and Evolution, 2018. **35**(6): p. 1547-1549.
456. Tamura, K., et al., *Estimating divergence times in large molecular phylogenies*. Proceedings of the National Academy of Sciences of the United States of America, 2012. **109**(47): p. 19333-19338.
457. Pridgeon, J.W., L. Zhang, and N.N. Liu, *Overexpression of CYP4G19 associated with a pyrethroid-resistant strain of the German cockroach, *Blattella germanica* (L.)*. Gene, 2003. **314**: p. 157-163.
458. Consortium, H.G.S., et al., *Insights into social insects from the genome of the honeybee *Apis mellifera**. Nature, 2006. **443**(7114): p. 931-949.
459. Johnson, R.M., et al., *Genomic footprint of evolution of eusociality in bees: floral food use and CYPome "blooms"*. Insectes Sociaux, 2018. **65**(3): p. 445-454.
460. Schmickl, T. and K. Crailsheim, *Inner nest homeostasis in a changing environment with special emphasis on honey bee brood nursing and pollen supply*. Apidologie, 2004. **35**(3): p. 249-263.
461. Feyereisen, R., *Arthropod CYPomes illustrate the tempo and mode in P450 evolution*. Biochimica Et Biophysica Acta-Proteins and Proteomics, 2011. **1814**(1): p. 19-28.
462. Petersen, M., et al., *Diversity and evolution of the transposable element repertoire in arthropods with particular reference to insects*. BMC Evolutionary Biology, 2019. **19**.
463. Beadle, K., *Understanding the molecular and biochemical basis of insecticide selectivity against solitary bee pollinators*. 2018, University of Exeter: University of Exeter, UK.
464. Calla, B., et al., *Cytochrome P450 diversification and hostplant utilization patterns in specialist and generalist moths: Birth, death and adaptation*. Molecular Ecology, 2017. **26**(21): p. 6021-6035.
465. Gonzalez, F.J. and D.W. Nebert, *Evolution of the P450-gene superfamily - animal plant warfare, molecular drive and human genetic-differences in drug oxidation*. Trends in Genetics, 1990. **6**(6): p. 182-186.
466. d'Alencon, E., et al., *Extensive synteny conservation of holocentric chromosomes in Lepidoptera despite high rates of local genome rearrangements*. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(17): p. 7680-7685.
467. Beye, M., et al., *Exceptionally high levels of recombination across the honey bee genome*. Genome Research, 2006. **16**(11): p. 1339-1344.
468. Stolle, E., et al., *A second generation genetic map of the bumblebee *Bombus terrestris* (Linnaeus, 1758) reveals slow genome and chromosome evolution in the Apidae*. BMC Genomics, 2011. **12**.
469. Jones, J.C., et al., *Extreme Differences in Recombination Rate between the Genomes of a Solitary and a Social Bee*. Molecular Biology and Evolution, 2019. **36**(10): p. 2277-2291.
470. Damalas, C.A. and I.G. Eleftherohorinos, *Pesticide Exposure, Safety Issues, and Risk Assessment Indicators*. International Journal of Environmental Research and Public Health, 2011. **8**(5): p. 1402-1419.
471. EPPO, *Efficacy evaluation of plant protection products. Side-effects on honeybees*. EPPO Bulletin, 2010. **40**(3): p. 313-319.
472. EFSA, *Guidance on the risk assessment of plant protection products on bees (*Apis mellifera*, *Bombus* spp. and solitary bees)*. EFSA Journal, 2013. **11**(7): p. 3295.

473. EPA, *Guidance on Exposure and Effects Testing for Assessing Risks to Bees*, O.o.P. Programs, Editor. 2016, U.S. Environmental Protection Agency: Washington, D.C. 20460 U.S.A.
474. ECPA, *Proposal for a protective and workable regulatory European bee risk assessment scheme based on the EFSA bee guidance and other new data and available approaches.*, E.C.P. Agency, Editor. 2017, European Crop Protection Agency: EU.
475. ECHA, *Preliminary considerations for ECHA's guidance on the "Methodology to assess the risk to bees and other non-target arthropod pollinators from the use of biocides". Summary of consultation feedback.*, ed. E.C. Agency. Vol. ED-04-20-730-EN-N. 2020, Helsinki, Finland: European Chemicals Agency.
476. Pamminger, T., *Extrapolating Acute Contact Bee Sensitivity to Insecticides Based on Body Weight Using a Phylogenetically Informed Interspecies Scaling Framework*. *Environmental Toxicology and Chemistry*, 2021. **40**(7): p. 2044-2052.
477. Arena, M. and F. Sgolastra, *A meta-analysis comparing the sensitivity of bees to pesticides*. *Ecotoxicology*, 2014. **23**(3): p. 324-334.
478. Uhl, P., et al., *Is Osmia bicornis an adequate regulatory surrogate? Comparing its acute contact sensitivity to Apis mellifera*. *Plos One*, 2019.
479. Davidson, I.W.F., J.C. Parker, and R.P. Beliles, *Biological basis for extrapolation across mammalian-species*. *Regulatory Toxicology and Pharmacology*, 1986. **6**(3): p. 211-237.
480. West, G.B., J.H. Brown, and B.J. Enquist, *The fourth dimension of life: Fractal geometry and allometric scaling of organisms*. *Science*, 1999. **284**(5420): p. 1677-1679.
481. Wagner, D.L., et al., *Insect decline in the Anthropocene: Death by a thousand cuts*. *Proceedings of the National Academy of Sciences of the United States of America*, 2021. **118**.
482. OECD/OCDE, *OECD Test No. 214: Honeybees, acute contact toxicity test. OECD guideline for the testing of chemicals.*, OECD, Editor. 1998, OECD Publishing: Paris.
483. OECD/OCDE, *OECD Test No. 213: Honeybees, acute oral toxicity test. OECD guideline for the testing of chemicals.*, OECD, Editor. 1998, OECD Publishing: Paris.
484. OECD/OCDE, *OECD Test No. 246: Bumblebee, acute contact toxicity test. OECD guideline for the testing of chemicals, Section 2.*, OECD, Editor. 2017, OECD Publishing: Paris.
485. OECD/OCDE, *OECD Test No. 247: Bumblebee, acute oral toxicity test. OECD guideline for the testing of chemicals, Section 2.*, OECD, Editor. 2017, OECD Publishing: Paris.
486. Phan, N.T., et al., *A new ingestion bioassay protocol for assessing pesticide toxicity to the adult Japanese orchard bee (Osmia cornifrons)*. *Scientific Reports*, 2020. **10**(1).
487. Roessink, I., J.J.M. van der Steen, and N. Hanewald, *Solitary bee, acute contact toxicity test. Version: March 2016*, in *ICPPR workgroup non-Apis bees*. 2016.
488. Johansen, C., E.R. Jaycox, and R. Hutt, *The effect of pesticides on the alfalfa leaf cutting bee Megachile rotundata*. *Station Circular*. Washington Agricultural Experiment Station, 1963. **No 418**: p. pp12.
489. Torchio, P.F., *Relative Toxicity of Insecticides to the Honey Bee, Alkali Bee, and Alfalfa Leafcutting Bee (Hymenoptera: Apidae, Halictidae, Megachilidae)*. *Journal of the Kansas Entomological Society*, 1973. **46**(4): p. 446--453.
490. Helson, B.V., K.N. Barber, and P.D. Kingsbury, *Laboratory toxicology of 6 forestry insecticides to 4 species of bee (Hymenoptera, Apoidea)*. *Archives of Environmental Contamination and Toxicology*, 1994. **27**(1): p. 107-114.

491. Mayer, D.F., G. Kovacs, and J.D. Lunden, *Field and laboratory tests on the effects of cyhalothrin on adults of Apis mellifera, Megachile rotundata and Nomia melanderi*. Journal of Apicultural Research, 1998. **37**(1): p. 33-37.
492. Laurino, D., et al., *Acute oral toxicity of neonicotinoids on different honey bee strains*. Redia-Giornale Di Zoologia, 2010. **93**: p. 99-102.
493. EPA, *Guidance for assessing pesticide risks to bees.*, O.o.P. Programs, Editor. 2014, U.S. Environmental Protection Agency: Washington, D.C. 20460 U.S.A.
494. Finney, D.J., *Probit Analysis*. 3rd edition ed. 1971: Cambridge, London and New York.
495. Abbott, W.S., *A method of computing the effectiveness of an insecticide*. Journal of Economic Entomology, 1925. **18**: p. 265-267.
496. Johnson, R.M., Z.Y. Huang, and M.R. Berenbaum, *Role of detoxification in Varroa destructor (Acari: Parasitidae) tolerance of the miticide tau-fluvalinate*. International Journal of Acarology, 2010. **36**(1): p. 1-6.
497. Calatayud-Vernich, P., et al., *Influence of pesticide use in fruit orchards during blooming on honeybee mortality in 4 experimental apiaries*. Science of the Total Environment, 2016. **541**: p. 33-41.
498. Haas, J., et al., *A toxicogenomics approach reveals characteristics supporting the honey bee (Apis mellifera L.) safety profile of the butenolide insecticide flupyradifurone*. Ecotoxicology and Environmental Safety, 2021. **217**.
499. EU, *SANCO/10329/2002 Rev 2. Guidance document on terrestrial ecotoxicology under Council Directive 91/414/EEC*, E.C.H.a.C. Protection, Editor. 2002, European Union.
500. Reich, P.B., et al., *Universal scaling of respiratory metabolism, size and nitrogen in plants*. Nature, 2006. **439**(7075): p. 457-461.
501. Chown, S.L., et al., *Scaling of insect metabolic rate is inconsistent with the nutrient supply network model*. Functional Ecology, 2007. **21**(2): p. 282-290.
502. Riveros, A.J. and B.J. Enquist, *Metabolic scaling in insects supports the predictions of the WBE model*. Journal of Insect Physiology, 2011. **57**(6): p. 688-693.
503. Polilov, A.A. and A.A. Makarova, *The scaling and allometry of organ size associated with miniaturization in insects: A case study for Coleoptera and Hymenoptera*. Scientific Reports, 2017. **7**.
504. OECD/OCDE, *OECD Test No. 237: Honey bee (Apis mellifera) larval toxicity test, single exposure. OECD guidelines for the testing of chemicals.*, OECD, Editor. 2013, OECD Publishing: Paris.
505. David, A., et al., *Widespread contamination of wildflower and bee-collected pollen with complex mixtures of neonicotinoids and fungicides commonly applied to crops*. Environment International, 2016. **88**: p. 169-178.
506. Yang, E.C., et al., *Abnormal Foraging Behavior Induced by Sublethal Dosage of Imidacloprid in the Honey Bee (Hymenoptera: Apidae)*. Journal of Economic Entomology, 2008. **101**(6): p. 1743-1748.
507. Stanley, D.A. and N.E. Raine, *Bumblebee colony development following chronic exposure to field-realistic levels of the neonicotinoid pesticide thiamethoxam under laboratory conditions*. Scientific Reports, 2017. **7**.
508. Tsvetkov, N., et al., *Chronic exposure to neonicotinoids reduces honey bee health near corn crops*. Science, 2017. **356**(6345): p. 1395-+.
509. Decourtye, A., et al., *Comparative sublethal toxicity of nine pesticides on olfactory learning performances of the honeybee Apis mellifera*. Archives of Environmental Contamination and Toxicology, 2005. **48**(2): p. 242-250.
510. Arce, A.N., et al., *Impact of controlled neonicotinoid exposure on bumblebees in a realistic field setting*. Journal of Applied Ecology, 2017. **54**(4): p. 1199-1208.
511. Lu, C.S., Y.T. Hu, and Q. Cheng, *A Review of Sub-lethal Neonicotinoid Insecticides Exposure and Effects on Pollinators*. Current Pollution Reports, 2020. **6**(2): p. 137-151.

512. Alaux, C., et al., *Interactions between Nosema microspores and a neonicotinoid weaken honeybees (Apis mellifera)*. Environmental Microbiology, 2010. **12**(3): p. 774-782.
513. Fauser-Misslin, A., et al., *Influence of combined pesticide and parasite exposure on bumblebee colony traits in the laboratory*. Journal of Applied Ecology, 2014. **51**(2): p. 450-459.
514. Gradish, A.E., et al., *Lethal and sublethal effects of some insecticides recommended for wild blueberry on the pollinator Bombus impatiens*. Canadian Entomologist, 2012. **144**(3): p. 478-486.
515. Feyereisen, R., *Insect P450 enzymes*. Annual Review of Entomology, 1999. **44**: p. 507-533.
516. Hannemann, F., et al., *Cytochrome P450 systems - biological variations of electron transport chains*. Biochimica Et Biophysica Acta-General Subjects, 2007. **1770**(3): p. 330-344.
517. Riddick, D.S., et al., *NADPH-Cytochrome P450 Oxidoreductase: Roles in Physiology, Pharmacology, and Toxicology*. Drug Metabolism and Disposition, 2013. **41**(1): p. 12-23.
518. Halon, E., et al., *Only a minority of broad-range detoxification genes respond to a variety of phytotoxins in generalist Bemisia tabaci species*. Scientific Reports, 2015. **5**.
519. Ehrlich, P.R. and P.H. Raven, *Butterflies and plants - a study in coevolution*. Evolution, 1964. **18**(4): p. 586-608.
520. Chung, H., et al., *Characterization of Drosophila melanogaster cytochrome P450 genes*. Proceedings of the National Academy of Sciences of the United States of America, 2009. **106**(14): p. 5731-5736.
521. Willoughby, L., et al., *A comparison of Drosophila melanogaster detoxification gene induction responses for six insecticides, caffeine and phenobarbital*. Insect Biochemistry and Molecular Biology, 2006. **36**(12): p. 934-942.
522. Liu, N.N., et al., *Cytochrome P450s-Their expression, regulation, and role in insecticide resistance*. Pesticide Biochemistry and Physiology, 2015. **120**: p. 77-81.
523. Cohen, M.B., M.A. Schuler, and M.R. Berenbaum, *A host-inducible cytochrome-P-450 from a host-specific caterpillar - molecular-cloning and evolution*. Proceedings of the National Academy of Sciences of the United States of America, 1992. **89**(22): p. 10920-10924.
524. Hung, C.F., et al., *CYP6B3 - A 2nd furanocoumarin-inducible cytochrome-P450 expressed in Papilio polyxenes*. Insect Molecular Biology, 1995. **4**(3): p. 149-160.
525. Giraudo, M., et al., *Regulation of cytochrome P450 expression in Drosophila: Genomic insights*. Pesticide Biochemistry and Physiology, 2010. **97**(2): p. 115-122.
526. Taneja, G., et al., *Transcriptomic profiling identifies novel mechanisms of transcriptional regulation of the cytochrome P450 (Cyp)3a11 gene*. Scientific Reports, 2019. **9**.
527. Obach, R.S. and A.E. Reed-Hagen, *Measurement of Michaelis constants for cytochrome P450-mediated biotransformation reactions using a substrate depletion approach*. Drug Metabolism and Disposition, 2002. **30**(7): p. 831-837.
528. Atkins, W.M., *Non-Michaelis-Menten kinetics in cytochrome P450-catalyzed reactions*. Annual Review of Pharmacology and Toxicology, 2005. **45**: p. 291-310.
529. Crowe, J. and T. Bradshaw, *Chemistry for the Biosciences. The essential concepts*. Third Edition. ed. 2014, Oxford. U.K.: Oxford University Press.
530. Wheelock, G.D. and J.G. Scott, *Rapid microsome preparation from limited numbers of house-flies*. Entomologia Experimentalis Et Applicata, 1991. **61**(3): p. 295-299.
531. Brodie, B.B., et al., *Detoxification of drugs and other foreign compounds by liver microsomes*. Science, 1955. **121**(3147): p. 603-604.

532. O'Brien, R.D., *Activation of thionophosphates by liver microsomes*. . Nature, 1959. **183**(4654): p. 121-122.
533. Agosin, M., et al., *A new DDT-metabolizing enzyme in German cockroach*. Journal of Economic Entomology, 1961. **54**(2): p. 340-&.
534. Zaworra, M. and R. Nauen, *New approaches to old problems: Removal of phospholipase A(2) results in highly active microsomal membranes from the honey bee, Apis mellifera*. Pesticide Biochemistry and Physiology, 2019. **161**: p. 68-76.
535. Gilbert, M.D. and C.F. Wilkinson, *Microsomal oxidases in honey bee, Apis mellifera (L.)*. Pesticide Biochemistry and Physiology, 1974. **4**(1): p. 56-66.
536. Gilbert, M.D. and C.F. Wilkinson, *Inhibitor of microsomal oxidation from gut tissues of honey bee (Apis mellifera)*. Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology, 1975. **50**(NB4): p. 613-619.
537. Macalintal, E.A. and C.K. Starr, *Comparative morphology of the stinger in the social wasp genus Ropalidia (Hymenoptera: Vespidae)*. Memoirs of the Entomological Society of Washington, 1996. **17**: p. 108-115.
538. Kumpanenko, A., D. Gladun, and L. Vilhelmsen, *Functional morphology and evolution of the sting sheaths in Aculeata (Hymenoptera)*. Arthropod Systematics & Phylogeny, 2019. **77**(2): p. 325-338.
539. Invitrogen, *Gateway Technology with Clonase II. A universal technology to clone DNA sequences for functional analysis and expression in multiple systems. User Guide*. 2012, Invitrogen. Life Technologies.
540. Zimmer, C.T., et al., *Molecular and functional characterization of CYP6BQ23, a cytochrome P450 conferring resistance to pyrethroids in European populations of pollen beetle, Meligethes aeneus*. Insect Biochemistry and Molecular Biology, 2014. **45**: p. 18-29.
541. Lu, H., et al., *Effects of heme precursors on CYP1A2 and POR expression in the baculovirus/Spodoptera frugiperda system*. Journal of Biomedical Research, 2010. **24**: p. 242-249.
542. Hasemann, C.A., et al., *Structure and function of cytochromes-P450 - a comparative-analysis of 3 crystal-structures*. Structure, 1995. **3**(1): p. 41-62.
543. Pitt, J.J., *Principles and Applications of Liquid Chromatography- Mass Spectrometry in Clinical Biochemistry*. Clinical Biochemist Reviews, 2009. **30**: p. 19-34.
544. Suchail, S., et al., *In vivo distribution and metabolisation of C-14-imidacloprid in different compartments of Apis mellifera L*. Pest Management Science, 2004. **60**(11): p. 1056-+.
545. Burke, M.D. and R.T. Mayer, *Ethoxyresorufin - direct fluorimetric assay of a microsomal O-dealkylation which is preferentially inducible by 3-methylcholanthrene*. Drug Metabolism and Disposition, 1974. **2**(6): p. 583-588.
546. Boulenc, X., et al., *Regulation of cytochrome-P450IA1 gene-expression in a human intestinal-cell line, CACO-2*. Journal of Pharmacology and Experimental Therapeutics, 1992. **263**(3): p. 1471-1478.
547. Haasch, M.L., et al., *Use of 7-alkoxyphenoxazones, 7-alkoxycoumarins and 7-alkoxyquinolines as fluorescent substrates for rainbow-trout hepatic microsomes after treatment with various inducers*. Biochemical Pharmacology, 1994. **47**(5): p. 893-903.
548. McLaughlin, L.A., et al., *Characterization of inhibitors and substrates of Anopheles gambiae CYP6Z2*. Insect Molecular Biology, 2008. **17**(2): p. 125-135.
549. Haas, J. and R. Nauen, *Pesticide risk assessment at the molecular level using honey bee cytochrome P450 enzymes: A complementary approach*. Environment International, 2021. **147**.
550. Haas, J., et al., *A toxicogenomics approach reveals characteristics supporting the honey bee (Apis mellifera L.) safety profile of the butenolide insecticide flupyradifurone*. Ecotoxicology and Environmental Safety, 2021. **217**.

551. Bradford, M.M., *Rapid and sensitive method for quantitation of microgram quantities of protein utilizing principle of protein-dye binding*. Analytical Biochemistry, 1976. **72**(1-2): p. 248-254.
552. Pitts-Singer, T.L., et al., *Structural Examination of the Dufour's Gland of the Solitary Bees *Osmia lignaria* and *Megachile rotundata* (Hymenoptera: Megachilidae)*. Annals of the Entomological Society of America, 2012. **105**(1): p. 103-110.
553. King, L.A. and R.D. Possee, *The Baculovirus Expression System: A laboratory Guide*. 1992, London: Chapman and Hall.
554. Invitrogen, *BaculoDirect Baculovirus Expression System. For cloning and high-level expression of recombinant proteins using Gateway-adapted Baculovirus DNA. User Guide*. 2012, Invitrogen. Life Technologies.
555. Contech, *BacPAK Baculovirus Rapid Titer Kit. User Manual*. 2011, Clontech Laboratories, Inc.: California. U.S.A.
556. Guengerich, F.P., et al., *Measurement of cytochrome P450 and NADPH-cytochrome P450 reductase*. Nature Protocols, 2009. **4**(9): p. 1245-1251.
557. Zhu, F., et al., *A brain-specific cytochrome P450 responsible for the majority of deltamethrin resistance in the QTC279 strain of *Tribolium castaneum**. Proceedings of the National Academy of Sciences of the United States of America, 2010. **107**(19): p. 8557-8562.
558. Welch, B.L., *The generalization of students problem when several different population variances are involved*. Biometrika, 1947. **34**(1-2): p. 28-35.
559. Srinivasan, B., et al., *Allosteric regulation and substrate activation in cytosolic nucleotidase II from *Legionella pneumophila**. Febs Journal, 2014. **281**(6): p. 1613-1628.
560. Parkinson, A., *An overview of current cytochrome P450 technology for assessing the safety and efficacy of new materials*. Toxicologic Pathology, 1996. **24**(1): p. 45-57.
561. Fowler, S. and H.J. Zhang, *In vitro evaluation of reversible and irreversible cytochrome P450 inhibition: Current status on methodologies and their utility for predicting drug-drug interactions*. Aaps Journal, 2008. **10**(2): p. 410-424.
562. du Rand, E.E., et al., *The metabolic fate of nectar nicotine in worker honey bees*. Journal of Insect Physiology, 2017. **98**: p. 14-22.
563. Tricker, A.R., *Nicotine metabolism, human drug metabolism polymorphisms, and smoking behaviour*. Toxicology, 2003. **183**(1-3): p. 151-173.
564. Delijewski, M., et al., *Genetically determined metabolism of nicotine and its clinical significance*. Acta Biochimica Polonica, 2019. **66**(4): p. 375-381.
565. Schep, L.J., R.J. Slaughter, and D.M.G. Beasley, *Nicotinic plant poisoning*. Clinical Toxicology, 2009. **47**(8): p. 771-781.
566. Bass, C., et al., *Gene amplification and microsatellite polymorphism underlie a recent insect host shift*. Proceedings of the National Academy of Sciences of the United States of America, 2013. **110**(48): p. 19460-19465.
567. Klot, A., et al., *Adaptation to nicotine in the facultative tobacco-feeding hemipteran *Bemisia tabaci**. Pest Management Science, 2014. **70**(10): p. 1595-1603.
568. Singaravelan, N., et al., *The effects of nectar-nicotine on colony fitness of caged honeybees*. Journal of Chemical Ecology, 2006. **32**(1): p. 49-58.
569. Zammit, M., et al., *The effects of anabasine and the alkaloid extract of *Nicotiana glauca* on Lepidopterous larvae*. International Journal of Biology, 2014. **6**(No 3).
570. Xu, X., M.M. Iba, and C.P. Weisel, *Simultaneous and sensitive measurement of anabasine, nicotine, and nicotine metabolites in human urine by liquid chromatography-tandem mass spectrometry*. Clinical Chemistry, 2004. **50**(12): p. 2323-2330.
571. Denton, T.T., X.D. Zhang, and J.R. Cashman, *Nicotine-related alkaloids and metabolites as inhibitors of human cytochrome P-450 2A6*. Biochemical Pharmacology, 2004. **67**(4): p. 751-756.

572. Horne, M., *Pollen preference and its relationship to nesting success of Megachile rotundata (Hymenoptera, Megachilidae)*. Annals of the Entomological Society of America, 1995. **88**(6): p. 862-867.
573. Sinu, P.A. and J.L. Bronstein, *Foraging preferences of leafcutter bees in three contrasting geographical zones*. Diversity and Distributions, 2018. **24**(5): p. 621-628.
574. Kunjwal, N., et al., *Notes on the nesting ecology of the Megachile bees from North India*. Journal of Apicultural Research.
575. Walker, N., et al., *Cytisine versus Nicotine for Smoking Cessation*. New England Journal of Medicine, 2014. **371**(25): p. 2353-2362.
576. Jeong, S.H., et al., *Pharmacokinetics of cytisine, an (42) nicotinic receptor partial agonist, in healthy smokers following a single dose*. Drug Testing and Analysis, 2015. **7**(6): p. 475-482.
577. Wustenberg, D.G. and B. Grunewald, *Pharmacology of the neuronal nicotinic acetylcholine receptor of cultured Kenyon cells of the honeybee, Apis mellifera*. Journal of Comparative Physiology a-Neuroethology Sensory Neural and Behavioral Physiology, 2004. **190**(10): p. 807-821.
578. Tutka, P. and W. Zatonski, *Cytisine for the treatment of nicotine addiction: from a molecule to therapeutic efficacy*. Pharmacological Reports, 2006. **58**(6): p. 777-798.
579. Moyer, B.G., et al., *COROLLIN, CORONILLIN AND CORONARIAN - 3 NEW 3-NITROPROPANOYL-D-GLUCOPYRANOSIDES FROM CORONILLA-VARIA*. Phytochemistry, 1977. **16**(3): p. 375-377.
580. Sun, Z.M., et al., *Biosynthesis and regulation of cyanogenic glycoside production in forage plants*. Applied Microbiology and Biotechnology, 2018. **102**(1): p. 9-16.
581. Egnell, A.C., B. Houston, and S. Boyer, *In vivo CYP3A4 heteroactivation is a possible mechanism for the drug interaction between felbamate and carbamazepine*. Journal of Pharmacology and Experimental Therapeutics, 2003. **305**(3): p. 1251-1262.
582. Shou, M., et al., *Enzyme kinetics of cytochrome P450-mediated reactions*. Current Drug Metabolism, 2001. **2**(1): p. 17-36.
583. Heraty, J., et al., *Evolution of the hymenopteran megaradiation*. Molecular Phylogenetics and Evolution, 2011. **60**(1): p. 73-88.
584. Klopstein, S., et al., *The Hymenopteran Tree of Life: Evidence from Protein-Coding Genes and Objectively Aligned Ribosomal Data*. Plos One, 2013. **8**(8).
585. Peters, R.S., et al., *Evolutionary History of the Hymenoptera*. Current Biology, 2017. **27**(7): p. 1013-1018.
586. Demain, J.G., *Hymenoptera allergy and anaphylaxis: are warmer temperatures changing the impact?* Current Opinion in Allergy and Clinical Immunology, 2020. **20**(5): p. 438-444.
587. Branstetter, M.G., et al., *Genomes of the Hymenoptera*. Current Opinion in Insect Science, 2018. **25**: p. 65-75.
588. Vilhelmsen, L., I. Miko, and L. Krogmann, *Beyond the wasp-waist: structural diversity and phylogenetic significance of the mesosoma in apocritan wasps (Insecta: Hymenoptera)*. Zoological Journal of the Linnean Society, 2010. **159**(1): p. 22-194.
589. Misof, B., et al., *Phylogenomics resolves the timing and pattern of insect evolution*. Science, 2014. **346**(6210): p. 763-767.
590. Eising-Birding, M., *Photo image: Athalia rosae (Sawfly sp.)*. 2019: <https://www.markeisingbirding.com/index/athalia-rosae-sawfly-sp>.
591. CBG, *Photo image: Image of a stem sawfly (Cephus cincta)*. Centre for Biodiversity Genomics Photography Group (CBG). 2009: <https://eol.org/pages/604065/media>.
592. Pathogen, P.P.a., *Photo image: Adult Cotesia vestalis*. Pacific Pest and Pathogen - Full size fact sheets. Biocontrol - Cotesia (287). 2019:

- https://apps.lucidcentral.org/ppp/text/web_full/entities/biocontrol_cotesia_287.htm.
593. Rosenfield, J., *Photo image: Wasp - Nasonia*. 2014: <https://bugguide.net/node/view/936473/bgimage>.
 594. Commons, C., *Photo image: Specimen of Microplitis demolitor*. 2011: https://v3.boldsystems.org/index.php/Taxbrowser_Taxonpage?taxid=106518.
 595. McCann, S., *Photo image: Chrysis (Cuckoo wasp)*. Creative commons non-commercial share-alike version 2.0. 2006: <http://tolweb.org/Chrysoidea/11186>.
 596. Koide, Y., *Photo image: Polistes rothneyi DSC 0322*. 2011: https://en.wikipedia.org/wiki/Polistes_rothneyi.
 597. Lief, J., *Photo image: Ant - individual*. 2012: <https://jonlieffmd.com/wp-content/uploads/2012/02/Ant-individual-jpg>.
 598. Johnson, B.R., et al., *Phylogenomics Resolves Evolutionary Relationships among Ants, Bees, and Wasps*. Current Biology, 2013. **23**(20): p. 2058-2062.
 599. Hines, H.M., et al., *Multigene phylogeny reveals eusociality evolved twice in vespid wasps*. Proceedings of the National Academy of Sciences of the United States of America, 2007. **104**(9): p. 3295-3299.
 600. Kapheim, K.M., et al., *Genomic signatures of evolutionary transitions from solitary to group living*. Science, 2015. **348**(6239): p. 1139-1143.
 601. Grimaldi, D., *The co-radiations of pollinating insects and angiosperms in the Cretaceous*. Annals of the Missouri Botanical Garden, 1999. **86**(2): p. 373-406.
 602. Murray, E.A., S. Bossert, and B.N. Danforth, *Pollinivory and the diversification dynamics of bees*. Biology Letters, 2018. **14**(11).
 603. Wcislo, W.T. and J.H. Cane, *Floral resource utilization by solitary bees (Hymenoptera: Apoidea) and exploitation of their stored foods by natural enemies*. Annual Review of Entomology, 1996. **41**: p. 257-286.
 604. Noll, F.B., et al., *Food collection and maturation in the necrophagous stingless bee, Trigona hypogea (Hymenoptera: Meliponinae)*. Journal of the Kansas Entomological Society, 1996. **69**(4): p. 287-293.
 605. Roubik, D.W., *Obligate necrophagy in a social bee*. Science, 1982. **217**(4564): p. 1059-1060.
 606. Mateus, S. and F.B. Noll, *Predatory behavior in a necrophagous bee Trigona hypogea (Hymenoptera: Apidae, Meliponini)*. Naturwissenschaften, 2004. **91**(2): p. 94-96.
 607. Heimpel, G.E. and J.G. de Boer, *Sex determination in the Hymenoptera*. Annual Review of Entomology, 2008. **53**: p. 209-230.
 608. Normark, B.B., *The evolution of alternative genetic systems in insects*. Annual Review of Entomology, 2003. **48**: p. 397-423.
 609. Blackmon, H., N.B. Hardy, and L. Ross, *The evolutionary dynamics of haplodiploidy: Genome architecture and haploid viability*. Evolution, 2015. **69**(11): p. 2971-2978.
 610. Sharkey, M.J., *Phylogeny and classification of Hymenoptera*. Zootaxa, 2007(1668): p. 521-548.
 611. Wang, Z.Z., et al., *Parasitoid wasps as effective biological control agents*. Journal of Integrative Agriculture, 2019. **18**(4): p. 705-715.
 612. Nelson, D.R., *Progress in tracing the evolutionary paths of cytochrome P450*. Biochimica Et Biophysica Acta-Proteins and Proteomics, 2011. **1814**(1): p. 14-18.
 613. Putnam, N.H., et al., *The amphioxus genome and the evolution of the chordate karyotype*. Nature, 2008. **453**(7198): p. 1064-U3.
 614. Luo, M., et al., *The evolution of insect metallothioneins*. Proceedings of the Royal Society B-Biological Sciences, 2020. **287**(1937).
 615. Nardelli, A., et al., *The Evolutionary History and Functional Divergence of Trehalase (treh) Genes in Insects*. Frontiers in Physiology, 2019. **10**.
 616. Shah, N., et al., *Evolution of a Large, Conserved, and Syntenic Gene Family in Insects*. G3-Genes Genomes Genetics, 2012. **2**(2): p. 313-319.

617. Wang, Z., M. Gerstein, and M. Snyder, *RNA-Seq: a revolutionary tool for transcriptomics*. Nature Reviews Genetics, 2009. **10**(1): p. 57-63.
618. Stoler, N. and A. Nekrutenko, *Sequencing error profiles of Illumina sequencing instruments*. NAR Genomics and Bioinformatics., 2021. **3**.
619. Sheffield, C.S., et al., *Leafcutter and mason bees of the genus Megachile Latreille (Hymenoptera: Megachilidae) in Canada and Alaska.*, in *Canadian Journal of Arthropod Identification*. 2011.
620. Baker, J.R., E.D. Kuhn, and S.B. Bambara, *Nests and immature stages of leafcutter bees (Hymenoptera, Megachilidae)*. Journal of the Kansas Entomological Society, 1985. **58**(2): p. 290-313.
621. Falk, S., *Field guide to the bees of Great Britain and Ireland*. 2016, London: Bloomsbury wildlife, Bloomsbury Publishing Plc. London. WC1B 3DP UK.
622. Owens, N., *Photo image: Megachile centuncularis female, Weybourne, Norfolk.* <https://www.bwars.com/bee/megachilidae/megachile-centuncularis>. 2011, Bees, Wasps and Ants Recording Society (BWARS).
623. Falk, S., *Photo image: Megachile dorsalis female. Hengistbury, Dorset.* <https://www.bwars.com/bee/megachilidae/megachile-dorsalis>. 2020, Bees, Wasps and Ants Recording Society (BWARS).
624. Falk, S., *Photo image: Megachile willughbiella female. Avon Dassett, Warwickshire.* <https://www.bwars.com/bee/megachilidae/megachile-willughbiella>. 2020, Bees, Wasps and Ants Recording Society (BWARS).
625. Grabherr, M.G., et al., *Full-length transcriptome assembly from RNA-Seq data without a reference genome*. Nature Biotechnology, 2011. **29**(7): p. 644-U130.
626. Bolger, A.M., M. Lohse, and B. Usadel, *Trimmomatic: a flexible trimmer for Illumina sequence data*. Bioinformatics, 2014. **30**(15): p. 2114-2120.
627. Jones, D.T., W.R. Taylor, and J.M. Thornton, *The rapid generation of mutation data matrices from protein sequences*. Computer Applications in the Biosciences, 1992. **8**(3): p. 275-282.
628. Wild, A., *Photo image: Cyphomyrmex costatus* <https://www.alexanderwild.com/Ants/Taxonomic-List-of-Ant-Genera/Cyphomyrmex/i-G4LMrBX/A>. 2007.
629. Wild, A., *Photo image: Camponotus floridanus* <https://www.alexanderwild.com/Ants/Taxonomic-List-of-Ant-Genera/Camponotus/i-9BwfNxX/A>. 2009.
630. Surcica, A., *Photo image: European paper wasp (worker) - Polistes dominula - female* <https://bugguide.net/node/view/553102>. 2011.
631. Koide, Y., *Photo image: Queen of Japanese giant hornet, Vespa mandarinia* https://cs.m.wikipedia.org/wiki/Soubor:20200512-P1100071_Vespa_mandarinia_japonica.jpg. 2020.
632. Kamping, A., et al., *Inheritance of gynandromorphism in the parasitic wasp Nasonia vitripennis*. Genetics, 2007. **175**(3): p. 1321-1333.
633. Simfruit, *Photo image: Trichogramma pretiosum* <https://www.simfruit.cl/buscan-controlar-la-lobesia-botrana-en-zonas-urbanas-usando-enemigo-natural-de-la-plaga/>. 2016.
634. DeSantis, T., *Photo image: Neodiprion lecontei* <https://bugguide.net/node/view/736109>. 2013.
635. Singh, K.S., et al., *Extension of Partial Gene Transcripts by Iterative Mapping of RNA-Seq Raw Reads*. IEEE-Acm Transactions on Computational Biology and Bioinformatics, 2019. **16**(3): p. 1036-1041.
636. Gonzalez, V.H., *Phylogeny and classification of the bee tribe Megachilini (Hymenoptera: Apoidea, Megachilidae), with emphasis on the genus Megachile.*, in *Department of Ecology and Evolutionary Biology*. 2008, University of Kansas.
637. Galetin, A., S.E. Clarke, and J.B. Houston, *Multisite kinetic analysis of interactions between prototypical CYP3A4 subgroup substrates: Midazolam, testosterone, and nifedipine*. Drug Metabolism and Disposition, 2003. **31**(9): p. 1108-1116.

638. Raw, A., *An annotated catalogue of the leafcutter and mason bees (genus Megachile) of the Neotropics*. Zootaxa, 2007(1601): p. 1-127.
639. Young, B.E., et al., *Conservation and management of North American leafcutter bees.*, NatureServe, Editor. 2016, NatureServe: Arlington, VA.
640. Gonzalez, V.H., G.T. Gustafson, and M.S. Engel, *Morphological phylogeny of Megachilini and the evolution of leaf-cutter behavior in bees (Hymenoptera: Megachilidae)*. Journal of Melittology, 2019. **85**: p. 1-123.
641. Danforth, B.N., R.L. Minckley, and J.L. Neff, *The solitary bees. Biology, Evolution, Conservation*. 2019, Princeton, New Jersey, USA: Princeton University Press.
642. Kumar, V. and K.M. Kumarang, *Diversity and conservation of leaf cutter bees (Hymenoptera: Megachilidae)*. Advances in Plants and Agriculture Research, 2018. **8**(1): p. 53-54.
643. Lopez-Osorio, F. and Y. Wurm, *Healthy Pollinators: Evaluating Pesticides with Molecular Medicine Approaches*. Trends in Ecology & Evolution, 2020. **35**(5): p. 380-383.
644. Gill, R.J., et al., *Protecting an Ecosystem Service: Approaches to Understanding and Mitigating Threats to Wild Insect Pollinators*. Ecosystem Services: from Biodiversity to Society, Pt 2, 2016. **54**: p. 135-+.
645. Uhl, P., et al., *Is Osmia bicornis an adequate regulatory surrogate? Comparing its acute contact sensitivity to Apis mellifera*. Plos One, 2019. **14**(8).
646. Sgolastra, F., et al., *Pesticide Exposure Assessment Paradigm for Solitary Bees*. Environmental Entomology, 2019. **48**(1): p. 22-35.
647. Sgolastra, F., et al., *Bees and pesticide regulation: Lessons from the neonicotinoid experience*. Biological Conservation, 2020. **241**.
648. Stoner, K.A., *Current Pesticide Risk Assessment Protocols Do Not Adequately Address Differences between Honey Bees (Apis mellifera) and Bumble Bees (Bombus spp.)*. Frontiers in Environmental Science, 2016. **4**.
649. Kopit, A.M. and T.L. Pitts-Singer, *Routes of Pesticide Exposure in Solitary, Cavity-Nesting Bees*. Environmental Entomology, 2018. **47**(3): p. 499-510.
650. PHTF, *Pollinator Health Task Force: Pollinator Partnership Action Plan*. 2016, Washington D.C. U.S.A.: The White House.
651. Liu, Z.C., et al., *Toxicogenomics: A 2020 Vision*. Trends in Pharmacological Sciences, 2019. **40**(2): p. 92-103.
652. Sharanowski, B.J., et al., *Phylogenomics of Ichneumonoidea (Hymenoptera) and implications for evolution of mode of parasitism and viral endogenization*. Molecular Phylogenetics and Evolution, 2021. **156**.
653. Blaimer, B.B., et al., *Comprehensive phylogenomic analyses re-write the evolution of parasitism within cynipoid wasps*. BMC Evolutionary Biology, 2020. **20**(1).
654. Doenecke, D. and W. Albig, *Intronless genes*. In *Encyclopedia of life sciences*. 2005, www.els.net: John Wiley and Sons, Ltd.
655. Diaz, M.O., *The human Type-I Interferon gene-cluster*. Seminars in Virology, 1995. **6**(3): p. 143-149.
656. Grzybowska, E.A., *Human intronless genes: Functional groups, associated diseases, evolution, and mRNA processing in absence of splicing*. Biochemical and Biophysical Research Communications, 2012. **424**(1): p. 1-6.
657. Shou, M., et al., *Activation of CYP3A4 - Evidence for the simultaneous binding of 2 substrates in a cytochrome-P450 active-site*. Biochemistry, 1994. **33**(21): p. 6450-6455.
658. Galetin, A., et al., *CYP3A4 substrate selection and substitution in the prediction of potential drug-drug interactions*. Journal of Pharmacology and Experimental Therapeutics, 2005. **314**(1): p. 180-190.
659. Kiani, Y.S., et al., *Molecular Dynamics Simulation Framework to Probe the Binding Hypothesis of CYP3A4 Inhibitors*. International Journal of Molecular Sciences, 2019. **20**(18).

660. Gonzalez, V.H., et al., *Phylogeny of the bee family Megachilidae (Hymenoptera: Apoidea) based on adult morphology*. Systematic Entomology, 2012. **37**(2): p. 261-286.
661. War, A.R., et al., *Plant defence against herbivory and insect adaptations*. Aob Plants, 2018. **10**(4).
662. Bonaventure, G., *Perception of insect feeding by plants*. Plant Biology, 2012. **14**(6): p. 872-880.
663. Appel, H.M. and R.B. Cocroft, *Plants respond to leaf vibrations caused by insect herbivore chewing*. Oecologia, 2014. **175**(4): p. 1257-1266.
664. Hill, B.D., K.W. Richards, and G.B. Schaalje, *Use of Dichlorvos resin strips to reduce parasitism of alfalfa leafcutter bee (Hymenoptera, Megachilidae) cocoons during incubation*. Journal of Economic Entomology, 1984. **77**(5): p. 1307-1312.
665. James, R.R. and T.L. Pitts-Singer, *Health Status of Alfalfa Leafcutting Bee Larvae (Hymenoptera: Megachilidae) in United States Alfalfa Seed Fields*. Environmental Entomology, 2013. **42**(6): p. 1166-1173.
666. Eves, J., D.F. Mayer, and C.A. Johansen, *Parasites, Predators, & Nest Destroyers of the Alfalfa Leafcutting Bee, Megachile rotundata*. Western Regional Extension Publication, 1980. **32**: p. 1-16.
667. Sparks, T.C., *Insecticide discovery: An evaluation and analysis*. Pesticide Biochemistry and Physiology, 2013. **107**(1): p. 8-17.
668. Grant, C., et al., *The evolution of multiple-insecticide resistance in UK populations of tomato leafminer, Tuta absoluta*. Pest Management Science, 2019. **75**(8): p. 2079-2085.
669. Sparks, T.C., et al., *Insecticide resistance management and industry: the origins and evolution of the Insecticide Resistance Action Committee (IRAC) and the mode of action classification scheme*. Pest Management Science, 2021. **77**(6): p. 2609-2619.
670. Cuthbertson, A.G.S., *Special Issue: Integrated Pest Management in Arable and Open Field Horticultural Crops*. Insects, 2020. **11**(2).
671. Ternest, J.J., et al., *Comparing Prophylactic Versus Threshold-Based Insecticide Programs for Striped Cucumber Beetle (Coleoptera: Chrysomelidae) Management in Watermelon*. Journal of Economic Entomology, 2020. **113**(2): p. 872-881.
672. Peterson, R.K.D., L.G. Higley, and L.P. Pedigo, *Whatever happened to IPM?* American Entomologist, 2018. **64**: p. 146-150.