Identification and Functional Analysis of Mountain Pine Beetle Genes Involved in Pheromone Biosynthesis and Resin Detoxification

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Biochemistry

by

Jeffrey Nadeau

Dr. Claus Tittiger/Dissertation Advisor

December, 2017
We recommend that the dissertation prepared under our supervision by

JEFFREY A. NADEAU

Entitled

Identification And Functional Analysis Of Mountain Pine Beetle Genes Involved In Pheromone Biosynthesis And Resin Detoxification

be accepted in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Claus Tittiger, Ph.D., Advisor

Gary Blomquist, Ph.D., Committee Member

Patricia Ellison, Ph.D., Committee Member

Kathleen Schegg, Ph.D., Committee Member

Robert G. Qualls, Ph.D., Graduate School Representative

David W. Zeh, Ph. D., Dean, Graduate School

December, 2017
Abstract

The mountain pine beetle (MPB, *Dendroctonus ponderosae* Hopkins) and Jeffrey pine beetle (JPB, *Dendroctonus jeffreyi*) are highly destructive pests of pine forests in western North America. During flight to a new host tree and initiation of feeding, MPB and JPB release aggregation pheromones. The biosynthetic pathways of these pheromones are sex-specific and localized in the midgut and fat body, but the enzymes involved have not all been identified or characterized. Additionally, very few resin detoxification enzymes used by bark beetles to survive host tree defenses have been identified or functionally analyzed. This dissertation used combination of a comparative transcriptomic techniques to analyze mRNA levels between fed and unfed male and female MPB midguts and fat bodies, male and female tissue distribution profiles and monoterpane exposure responses in males and females to identify candidate genes involved in pheromone biosynthesis or resin detoxification. Furthermore, functional analyses were used to confirm pheromone biosynthetic or detoxification activity. Quantitative real-time PCR (qRT-PCR) measurements of CYP6DH3 mRNA levels in male and female tissues and in response to monoterpane exposure were inconsistent with a role in either resin detoxification or pheromone biosynthesis, although the CYP6DH3 substrate range was very similar to closely related CYP6DH2, a previously characterized cytochrome P450 monooxygenase (P450) likely involved in resin detoxification. A comparative RNA-Seq analysis between fed and unfed male and female MPB midguts and fat bodies identified all four known pheromone biosynthetic genes, confirmed the tentative identification of four others from a previous study, and suggested nine novel candidates. One P450, CYP6DE3, identified as a possible *exo*-brevicomin-biosynthetic enzyme in this study, was functionally characterized and likely is involved in resin detoxification rather than pheromone biosynthesis. A putative frontalin-biosynthetic enzyme identified in this
study, CYP6DK1, was highly expressed in fed male midguts and catalyzes the production of frontalin from 6-methylhept-6-en-2-one (6MHO) without the need of a cyclase. Furthermore, an alternative product is produced from 6MHO when CYP6DK1 is recombinantly fused to house fly cytochrome P450 reductase (CPR). CYP4G55 and CYP4G56, one of which has been hypothesized to be involved in exo-brevicomin biosynthesis, both catalyzed the conversion of cis-7-decenal to cis-3-nonene. In addition to these analyses, attempts to develop a medium-throughput method for measuring P450 substrate specificity and enzyme kinetics was conducted with mixed results. The project was set aside to concentrate on the other experiments presented in this dissertation. Finally, comparative RNA-Seq analysis between fed and unfed male and female JPB midguts and fat bodies to identify pheromone biosynthetic genes was conducted with the data to be analyzed in the future. Further functional analyses of candidate genes found in this study should lead to the full characterization of MPB pheromone biosynthetic pathways and the identification of molecular targets for possible pest management strategies.
Acknowledgements

Sharon Young contributed to recombinant cytochrome P450 expression and with her expertise in molecular biology in general (Chapters 2-6). Sharon Young also helped editing the manuscript from the MPB comparative transcriptomics study for publication (Chapter 3).

Joseph Bernardo assisted in CYP6DH3 sequencing, expression, and functional assays (Chapter 2). Camei (Ella) Kuang conducted the qRT-PCR for tissue distribution profiling (Chapter 2&4).

Marina MacLean, Misha Fotoohi and Kristina Jung helped in beetle feeding and dissections in the MPB comparative transcriptomic study (Chapter 3). Marina MacLean also expressed and assayed CYP4G55 and CYP4G56 (Chapter 5) and assisted in editing the manuscript in Chapter 3 for publication. Kristina Jung also conducted all qRT-PCR measurements of monoterpane-exposed beetles (Chapters 2-4).

Dr. Karen Schlauch and Dr. Richard Tillett of the Nevada INBRE Bioinformatics Core and Dr. Juli Petereit all contributed to the bioinformatics analysis of the MPB comparative transcriptomics study, preparing figures and editing the manuscript for publication (Chapter 3).

Billie Leong assisted in CYP6DE3 sequencing, expression, and functional assays (Chapter 3).

Dr. Shumin Lyu and Cory Columbini contributed to CYP6DK1 and the CYP6DK1-CPR fusion protein sequencing, expression, and functional assays (Chapter 3).
Yingjuan Shi helped express CYP9T3 and HF-CPR for use in the OxoPlate® assays (Chapter 6).

Dr. Dylan Kosma generously allowed the use of his GC-MS for functional analyses of CYP6DH3, CYP6DE3, CYP6DK1, CYP4G55 and CYP4G56 (Chapters 2-5).

Dr. Gary Blomquist helped devise experiments and assisted in editing the MPB comparative transcriptomic study manuscript for publication (Chapter 3).

Dr. Claus Tittiger provided invaluable guidance and assistance with all experiments, the writing of this dissertation, and in the preparation of Chapter 3 manuscript for publication.
Table of Contents

Abstract .................................................................................................................................i

Acknowledgments ...............................................................................................................iii

Table of Contents ...............................................................................................................v

List of Abbreviations .........................................................................................................vi

Chapter 1. Introduction: Identification and functional analysis of mountain pine beetle genes involved in pheromone biosynthesis and resin detoxification ..............................................1

Chapter 2. Expression profiling and functional analysis of CYP6DH3 ..............................................26

Chapter 3. Comparative transcriptomics of mountain pine beetle pheromone-biosynthetic tissues and functional analysis of CYP6DE3 ..............................................................................56

Chapter 4. Expression profiling and functional analysis of CYP6DK1, a frontalin biosynthetic gene in the mountain pine beetle ...........................................................................102

Chapter 5. Mountain pine beetle CYP4G55 and CYP4G56 exhibit functions consistent with \textit{exo}-brevicomin biosynthesis ........................................................................................................132

Chapter 6. Towards a moderately high-throughput method for measuring cytochrome P450 kinetics and substrate specificities using OxoPlates® ...........................................................................146

Chapter 7. Discussion and future directions ...........................................................................179

Appendix 1. Comparative transcriptomics of the Jeffrey pine beetle midgut and fat body tissues ...............................................................................................................................210
List of Abbreviations

6MHO, 6-methyl-hept-6-en-2-one

a.a., amino acid

bp, base pair

cDNA, complementary deoxyribonucleic acid

CO, carbon monoxide

DF, degrees of freedom

EDTA, ethylenediaminetetraacetic acid

FAD, Flavin adenine dinucleotide

FF, female fed

FMN, Flavin mononucleotide

FU, female unfed

GC-MS, gas chromatography-mass spectrometry

GGF, Georgia Genomics Facility

GGPPS, geranygeranyl diphosphate synthase

GO, gene ontology

GTF, gene transfer format

HF-CPR, house fly cytochrome P450 reductase
HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase

JPB, Jeffrey pine beetle

JH, juvenile hormone

MF, male fed

MS, mean square

MU, male unfed

MPB, mountain pine beetle

mRNA, messenger ribonucleic acid

NADPH, nicotinamide adenine dinucleotide phosphate

nt, nucleotide

ORF, open reading frame

P450, cytochrome P450 monooxygenase

PBO, piperonyl butoxide

PCA, principal component analysis

PCR, polymerase chain reaction

PMSF, phenylmethysulfonyl fluoride;

qRT-PCR, quantitative reverse transcriptase PCR
RNA-Seq, RNA sequencing

SAM, sequence alignment/map

SS, sum of squares

VN, vicinity network
Chapter 1

Introduction: Identification and Functional Analysis of Mountain Pine Beetle Genes
Involved in Pheromone Biosynthesis and Resin Detoxification

I. Pine Bark Beetle Biology

II. Pheromones

III. Cytochrome P450 Monooxygenases

IV. Experimental Rationale

V. References

VI. Figure Legends

I. Pine Bark Beetle Biology

Bark beetles have devastated forests of North America in recent years, killing over forty million acres of trees in the past 15 years in the western United States (United States Department of Agriculture, 2017) and approximately 18.3 million hectares, greater than 53% of the merchantable pine, in western Canada leading to an estimated loss of almost 150 billion Canadian dollars over the next 45 years (Corbett et al., 2016). Although most bark beetle-killed trees still retain some value as lumber, salvaging killed trees may not be economically feasible in all impacted areas. In the western United States, an economic assessment of the viability of salvaging bark beetle-killed trees found that while some states could net a profit, the states with the most salvageable volume, Colorado, Utah, and Wyoming, could not expect to recoup the cost of salvaging activities (Prestemon et al., 2013). Additional economic impacts incurred from bark
beetle outbreaks may also be significant, including losses due to decreased tourism and recreational activities in impacted areas (Rosenberger et al., 2013).

*Dendroctonus* and *Ips* spp. are the main culprits in recent bark beetle outbreaks. Their native ranges span from Alaska to portions of Mexico (Williams and Liebhold, 2002). These pests not only cause grave long-term economic damage, but also cause important changes to the environment, ranging from small scale ecological impacts such as nutrient cycling to possible implications for global climate change (Kurz et al., 2008; Morehouse et al., 2008). Kurz et al. (2008) estimated that the bark beetle outbreak in British Columbia, Canada (BC) would result in the release of about 270 megatonnes of carbon into the atmosphere due to tree decomposition over the 20 year period starting in the year 2000. Coupled with decreased carbon uptake by the forests in BC, these emissions exacerbate the already copious amounts of man-made greenhouse gas emissions that have been implicated in global climate change.

Two bark beetles, the mountain pine beetle (*Dendroctonus ponderosae*; MPB) and the Jeffrey pine beetle (*Dendroctonus jeffreyi*; JPB), spend the majority of their life cycles sequestered beneath the bark of their host trees, only leaving the tree for a few summer days once a year to migrate to a new host, (Gibson et al., 2009; Smith et al., 2009). Both species have a four-stage life cycle including egg, larva, pupa and adult, all occurring within galleries built into the phloem of the tree. MPB mainly infest lodgepole (*Pinus contorta*) and ponderosa pine (*P. ponderosa*) trees but may successfully infest any pine tree, with the exception of Jeffrey pines (*P. jeffreyi*), within their range (Gibson et al., 2009). JPB, a closely related species to MPB, only successfully infests Jeffrey pine trees (Smith et al., 2009). Tree mortality from MPB and JPB beetle attacks is a result both of the beetles feeding on the phloem of the tree, and associated
fungal species that are introduced by the bark beetles, the growth of which results in decreased water flow and a blue staining of the sapwood (Gibson et al., 2009; Smith et al., 2009).

Targeted coniferous trees have been especially impacted in recent years as a result of persistent drought conditions (Williams and Liebhold, 2002; Negrón et al., 2009). Normally, healthy trees with abundant water possess inherent natural defenses, including the production of copious amounts of monoterpeneladen resin, that expels invading beetles and inhibits beetle survival after infestation (Boone et al., 2011). However, drought conditions significantly impede resin production, and bark beetles have thrived as a result. Increasing average global temperatures in the past few decades have also played a role in large bark beetle outbreaks. MPB that normally only have one generation may have two in warm conditions, resulting in more trees being attacked during abnormally warm years (Gibson et al., 2009). Additionally, warmer temperatures have allowed MPB to expand their historic ranges into normally frigid northern climes and east over the high altitude Rocky Mountains, a natural geographic barrier where normally the conditions are too cold in winter for the beetles to survive (Carroll et al., 2003). Some beetles may be adapting to colder conditions and longer dispersal flights through alterations in cellular and metabolic functions (Janes et al., 2014). Taken together, climate change and these adaptations have allowed MPB to threaten the northern boreal forests of Canada with the potential for devastating outbreaks into stands of a novel host, the jack pine (P. banksiana) (Cullingham et al., 2011). MPB are already beginning to attack jack pines and studies have shown their survival is not affected by the new chemistry in these novel hosts (Erbilgin et al., 2014).

Because bark beetles spend most of their life cycles beneath the bark of infested trees, they are largely unaffected by insecticide application. Large scale management attempts have
been almost solely restricted to standard silvicultural practices such as harvesting trees that have been invaded by bark beetles, but have produced very limited success given the massive scale of the problem (Fettig et al., 2007). However, employing pheromones to bait trees treated with pesticides has been marginally effective in protecting stands of trees (Wermelinger, 2004). Therefore, the right combination of silvicultural practices, pesticide application, and pheromone baiting may increase the efficiency of bark beetle management given an ample supply of pheromones.

II. Pheromones

A combination of aggregation pheromone components allows for a coordinated mass attack by MPB to overcome a host tree’s defenses (Blomquist et al., 2010). Female MPB use the aggregation pheromone (–)-trans-verbenol [(1S,2R,5S)-4,6,6-trimethylbicyclo[3.1.1]hept-3-en-2-ol], to attract other males and females to a newly colonized host tree (Pitman et al., 1968). Biosynthesis of (–)-trans-verbenol likely occurs through P450 hydroxylation of the host tree produced monoterpene (–)-α-pinene during colonization (Pureswaran et al., 2000) (Figure 1.1A), although possible sequestration of (–)-trans-verbenol during larval and pupal stages for release during colonization has been hypothesized (Keeling et al., 2016).

In synergy with the female produced (–)-trans-verbenol, newly emerged male MPB produce exo-brevicomin [(1R,5S,7R)-7-Ethyl-5-methyl-6,8-dioxabicyclo[3.2.1] octane] in small concentrations as an aggregation pheromone component, with production decreasing significantly upon reaching the new host tree and mating (Pureswaran et al., 2000). Biosynthesis of exo-brevicomin in the MPB is thought to proceed through β-oxidation of long chain fatty acid precursors to produce a ten carbon aldehyde, subsequent P450 mediated decarbonylation of the
aldehyde to a nine carbon alkene, followed by a P450 mediated hydroxylation to 6(Z)-non-6-en-2-ol (Blomquist et al., 2010) (Figure 1.1B). 6(Z)-non-6-en-2-ol is reduced to 6(Z)-non-6-en-2-one by the enzyme ZnoDH followed by epoxidation to 6,7-epoxy-nonan-2-one by P450 CYP6CR1 (Song, Delaplain, et al., 2014). Finally, 6,7-epoxy-nonan-2-one is likely cyclized by a cyclase to produce exo-brevicomin.

A third component, frontalin [(1R,5R)-1,5-dimethyl-6,8-dioxabicyclo[3.2.1] octane], is produced by MPB males upon feeding or hormone juvenile hormone III (JH) treatment. Frontalin acts as an anti-aggregation signal to halt the attack and prevent overcrowding of the tree (Barkawi et al., 2003). Biosynthesis of frontalin occurs through the mevalonate pathway and includes the 20-carbon geranylgeranyl diphosphate (GGPP) as an intermediate (Keeling, Chiu, et al., 2013) (Figure 1.1C). GGPP is thought to be cleaved to produce a 9-carbon intermediate, 6-methylhept-6-en-2-one (6-MHO), possibly through 6-methylhept-5-en-2-ol and/or 6-methylhept-5-en-2-one precursors, by a P450 or dioxygenase (Barkawi et al., 2003). Epoxidation of 6-methylhept-6-en-2-one is thought to occur by a P450 to form 6,7-epoxy-6-methylheptan-2-one (Blomquist et al., 2010) which, similar to the hypothesized exo-brevicomin biosynthetic pathway, is likely cyclized to frontalin by a yet to be identified cyclase.

Similar to MPB, JPB evidently use a combination of pheromones to coordinate mass attacks on new host trees, although the JPB pheromone system is not as well understood. Females produce 1-heptanol, presumably derived from a P450 mediated hydroxylation of the abundant supply of heptane produced in Jeffrey pines (Paine et al., 1999). The 1-heptanol, in synergy with host produced heptane, attracts male JPB to newly attacked host trees. Like MPB, male JPB produce exo-brevicomin to attract females (Paine et al., 1999), presumably through the same or a very similar pathway to MPB exo-brevicomin biosynthesis. Finally, male JPB also
produce a racemic mixture of frontalin which acts like an anti-aggregation pheromone in sufficient quantities (Paine et al., 1999). Frontalin in JPB is also produced via the mevalonate pathway and biosynthesis occurs in the anterior midgut (Hall et al., 2002). Very little research has occurred on JPB pheromone biosynthetic pathways, so the enzymes involved are mostly unknown. However, because of the close evolutionary relationship between JPB and MPB, many of the enzymes identified in MPB may also be present in JPB and could provide clues into the JPB pheromone biosynthetic pathways.

III. Cytochrome P450 Monooxygenases

P450s are important heme containing proteins found in almost all organisms. They are implicated in diverse functions ranging from the production of metabolic intermediates to detoxification through chemical modifications such as decarbonylation, oxidation, epoxidation and hydroxylation (Denisov et al., 2005). Although known for their oxidation activity, P450s can also catalyze a variety of reactions including reductions, desaturations, and ring formations (Guengerich, 2001). Prokaryotic P450s are soluble while eukaryotic P450s are membrane bound and found either on the outer face of the endoplasmic reticulum or the inner membrane of mitochondria. Over 20,000 distinct P450s have been identified and the current nomenclature guidelines suggest that members of new CYP families share at least 40% amino acid identity, while members of subfamilies must share at least 55% amino acid identity (Nelson et al., 2004; Nelson, 2009). While P450s have great sequence diversity, with identities between amino acid sequences dipping below 20% in some instances, structural homology is remarkably conserved, especially in the core of the protein (Werck-Reichhart and Feyereisen, 2000). The overall structure of P450 consists of a number of α-helices (A-L) and folds with a heme group located in
the conserved core of the protein (helices D, E, I and L) and bound to the protein through a
cysteine thiol bond. The most highly variable regions of P450s are found in the substrate
recognition and binding sites (Denisov et al., 2005).

Most P450s require a redox partner and electron donor and are split into four distinct
classes depending on which redox partner and/or electron donor are used. Class I P450s require a
FMN containing reductase and iron sulfur redoxin to pass electrons from NADPH. Class II uses
a CPR containing both FMN and FAD domains to pass electrons from NADPH. Class III
requires no redox partner or NADPH, and finally Class IV can take electrons directly from
NAPDH (Werck-Reichhart and Feyereisen, 2000). Electrons from NADPH are required in most
cases to reduce the iron in the heme group so that a dioxygen molecule may bind after a substrate
has entered the substrate binding pocket. In the most common P450 reaction, a number
intermediate steps leads to oxidation of a substrate and a molecule of water being released at the
conclusion of the reaction (Groves and Han, 1995). P450s may be very specific for their
substrates, but many can bind multiple different substrates, often yielding a variety of products
from a single substrate (Denisov et al., 2005).

Over 1750 insect P450s have been identified. Most belong to the, CYP6, CYP9, CYP4,
CYP325, CYP12 and CYP314 families (Feyereisen, 2006; Schuler, 2011; Schuler and
Berenbaum, 2013). A large portion of the insect P450s studied are involved in insecticide
resistance or metabolism of natural products such as plant toxins (Schuler, 2011). In many cases,
the evolution of these P450s seems to involve gene duplication events followed by divergent
evolution through mechanisms such as intron loss or gain and positive selection pressures
(Feyereisen, 1999). Transposable elements also play a role in modifying promoter regions to
induce overexpression in P450s involved in insecticide resistance (Chen and Li, 2007; Schmidt et al., 2010).

Bark beetles use a variety of P450s for monoterpene detoxification and pheromone biosynthesis (Blomquist et al., 2010). One bark beetle P450 from *Ips pini*, CYP9T3, catalyzes the hydroxylation of myrcene to produce ipsdienol, an aggregation pheromone (Song et al., 2013). CYP9T3 not only hydroxylates myrcene, a tree resin component that is also produced *de novo* in *Ips pini* (Seybold et al., 1995), but also hydroxylates other compounds found in tree resin including (+) and (–)-α-pinene, 3-carene and R-(+)-limonene (Song et al., 2013). This activity suggests an ancestral role in resin detoxification that evolved into pheromone production (Byers and Birgersson, 2012; Song et al., 2013). Female MPB are hypothesized to use a P450 mediated hydroxylation of a host tree precursor, (–)-α-pinene, to produce the aggregation pheromone component (–)-trans-verbenol (Blomquist et al., 2010). Like CYP9T3, the P450 involved in trans-verbenol production may also have evolved from a detoxification enzyme that has a broad substrate range including other tree resin components. Another important MPB P450 is CYP6CR1, an enzyme that epoxidizes 6(Z)-nonen-2-one to 6,7-epoxynonan-2-one in the production of the aggregation pheromone component *exo*-brevicomin (Song, Delaplain, et al., 2014). Other yet to be identified MPB P450s are hypothesized to be involved in other steps of *exo*-brevicomin and frontalin biosynthesis (Blomquist et al., 2010).

MPB may also use P450s to produce cuticular hydrocarbons to prevent desiccation. Other insects use a family of P450s, called the CYP4G family, which catalyze the oxidative decarbonylation of long chain fatty acids to produce hydrocarbons for transport to the cuticle (Qiu et al., 2012; Balabanidou et al., 2016). Likewise, it is hypothesized that the only two MPB P450s in the CYP4G family, CYP4G55 and CYP4G56, act in a similar fashion. Additionally, the
decarbonylation step that converts cis-7-decenal to cis-3-nonene early in the exo-brevicomin biosynthetic pathway is hypothesized to require a CYP4G (Blomquist et al., 2010) (Figure 1.3), in part because the alternative, an α-oxidation of a β-keto-precursor, is highly unusual. The insect-specific CYP4G are the only class of enzyme known to catalyze decarbonylation reactions to form hydrocarbons from aldehydes (Qiu et al., 2012). Thus, CYP4G55 or CYP4G56 are likely candidates for involvement in the exo-brevicomin biosynthetic pathway.

IV. Experimental Rationale

Bark beetle outbreaks are notoriously difficult to control because the beetles spend most of their lives sequestered beneath the bark of the host tree where they are protected from insecticides and predators (Gibson et al., 2009; Smith et al., 2009). Harvesting infested trees is another option for control, but it can be difficult and costly with mass outbreaks in difficult terrain (Prestemon et al., 2013). Baiting trees with aggregation pheromones to prevent the spread of the pests, as well as protecting high value trees with anti-aggregation pheromones, has met with limited success, but the synthesis of these pheromones is very costly and not suitable for large-scale forest-wide applications (Wermelinger, 2004). Therefore, identifying new targets for control strategies and/or cheap methods for synthesizing pheromones is a potentially valuable field of study.

MPB require aggregation and anti-aggregation pheromones to coordinate mass attacks on new host trees necessary to overcome tree defenses (Blomquist et al., 2010). The pheromone biosynthetic enzymes involved in making these pheromones are likely species specific, making them good targets for possible control strategies. Other potential molecular targets in MPB
include detoxification enzymes, such as P450s, without which the MPB is likely to succumb to
the toxic effects of the host tree’s resin.

Various high-throughput studies, including a draft MPB genome (Keeling, Yuen, et al.,
2013), and various transcriptomic (Keeling et al., 2012; Robert et al., 2013) and proteomic (Pitt
et al., 2014) analyses have been directed at this highly destructive coniferous forest pest in order
to better understand its physiology and to assist development of new management strategies.
Aggregation and anti-aggregation pheromones in the MPB are synthesized in the midguts and fat
bodies in a sex- and feeding status-specific manner (Blomquist et al., 2010; Song et al., 2014).
Therefore, mRNAs encoding enzymes involved in pheromone-biosynthetic pathways may also
be differentially expressed based on sex, feeding status, and tissue. Recently, Keeling et al.
(2016) reported transcriptome, proteome, and metabolome-level responses of unfed MPB treated
with juvenile hormone (JH) III. JH III stimulates aggregation pheromone component production
in bark beetles (Borden, Nair and Slater, 1969; Chen, Borden and Pierce, 1988; Tillman et al.,
1998; Barkawi et al., 2003; Keeling et al., 2006), including frontalin and trans-verbenol in MPB,
and several “pheromone-biosynthetic gene” candidates were identified by this study. However,
the dynamics of the transcriptional response of JH III-treated insects can differ from those of fed
insects (Tittiger, Keeling and Blomquist, 2005; Keeling et al., 2006; Bearfield et al., 2009) and
the study noted that JH III treatment did not affect exo-brevicomin production. Thus, JH
treatment may not induce differential expression of all genes involved in pheromone biosynthesis
for the MPB.

Most of the enzymes catalyzing pheromone biosynthesis remain unidentified and even
fewer are characterized. Two MPB genes, CYP6CR1 and ZnoDH, were identified through
differential gene expression analysis using the first ever microarray for this species and were
shown to have increased transcript levels in unfed males, consistent with \textit{exo}-brevicomin biosynthesis (Aw \textit{et al.}, 2010). Functional analyses through cloning and enzyme assays confirmed that these two enzymes are involved in the biosynthetic pathway leading to \textit{exo}-brevicomin (Song, Delaplain, \textit{et al.}, 2014). Another study confirmed that MPB frontalin biosynthesis proceeds through the mevalonate pathway and employs a geranylgeranyl diphosphate synthase (GGPPS) to produce GGPP (Keeling \textit{et al.}, 2013). Although these enzymes have been characterized, there are still many yet to be identified in the \textit{trans}-verbenol, frontalin and \textit{exo}-brevicomin biosynthetic pathways (Figure 1.1).

The overall goal of this dissertation is to identify the remaining candidate genes involved in MPB pheromone biosynthesis and confirm their function in the pheromone biosynthetic pathways through functional analyses. I therefore analyzed the differential expression of mRNAs in the MPB midguts and fat bodies based on sex and feeding status through RNA-seq using a rigorous bioinformatics approach to identify candidate genes for involvement in pheromone biosynthesis (Chapter 3). I then followed up these preliminary identifications with functional assays of recombinant enzymes to confirm their activities (Chapters 4, 5).

Select candidate genes identified in the comparative transcriptomics study were functionally characterized using cloning and enzyme assays to confirm their involvement in MPB pheromone biosynthesis. One enzyme, CYP6DH3, was functionally characterized because of its relatively close homology to two other enzymes of interest in the same family, CYP6DH1 and CYP6DH2. CYP6DH1 mRNA has an expression profile consistent with \textit{trans}-verbenol production (Robert \textit{et al.}, 2013), while CYP6DH2 has an expression profile indicating possible involvement in detoxification (Song, 2012). Further functional characterization showed that both CYP6DH1 and CYP6DH2 are involved in monoterpane hydroxylation, with CYP6DH1 having a
narrow substrate range selective for (−)-α-pinene, resulting in (−)-trans-verbenol production, while and CYP6DH2 has a broader monoterpenic substrate range, suggesting a role in detoxification (Song, 2012). We hypothesized that CYP6DH3 is likely also involved in monoterpenic hydroxylation because of its sequence similarity to CYP6DH1 and CYP6DH2, both of which hydroxylate monoterpenes. The results of this study are reported in Chapter 2.

A novel high-throughput method to measure P450 substrate specificities and enzyme kinetics was also explored (Chapter 6). Understanding substrate specificities and enzyme kinetics can lead to insights into how structure relates to function in enzymes, and the evolution of kinetic mechanisms (Ulusu, 2015). High throughput methods for measuring P450 kinetics have been developed using radioactive, chromophoric, fluorescent or luminescent substrates or products for particular P450s (Moody et al., 1999; Stresser et al., 2000; Crespi, Miller and Stresser, 2002; Yamamoto, Suzuki and Kohno, 2002; Cali et al., 2006), but these substrates are suitable for very few enzymes. The conventional method for measuring P450 kinetics is time-consuming and labor intensive, involving incubating microsomes containing the relevant P450 with different concentrations of substrate, stopping the reaction at various intervals with a chemical inhibitor, and measuring the change in substrate or product concentrations using a method like liquid chromatography–tandem mass spectrometry (Obach and Reed-Hagen, 2002; Baranczewski, Edlund and Postlind, 2006). Measuring product formation in P450 reactions may also be difficult if the products are unknown and proper standards are unavailable (Obach and Reed-Hagen, 2002). Given these limitations, development of a reliable medium or high-throughput method for measuring P450 substrate specificities and kinetics would be extremely useful.

Most P450 reactions involve the consumption of molecular oxygen to create an oxidized substrate and water. Past studies have used oxygen sensing fluorophores in a 96-well plate
format to measure P450 activity with limited success (Olry et al., 2007; Chang et al., 2011). We tested a sensitive, commercially available oxygen sensing system in a 96-well format called the OxoPlate® in an attempt to develop a moderately high-throughput method for measuring microsomal recombinant P450 substrate specificities and enzyme kinetics.

Finally, little is known about JPB pheromone biosynthesis. Like MPB, JPB pheromones appear to be produced in a sex- and feeding status-specific manner. Therefore, we designed a comparative transcriptomics study to ascertain what genes may be involved in JPB pheromone biosynthesis (Appendix 1). Because the JPB genome has not sequenced, we used the closely related MPB genome as a reference for transcript assembly.

V. References


VI. Figure Legends

**Figure 1.1. MPB pheromone biosynthetic pathways.** (A) (−)-trans-Verbenol is produced by a single P450-mediated hydroxylation of (−)-α-pinene. (B) Frontalin biosynthesis occurs through the mevalonate pathway to geranylgeranyl diphosphate followed by multiple steps likely catalyzed by P450s, a dioxygenase, and a cyclase. (C) *exo*-Brevicomin production from long chain fatty acid precursors in the fat body of unfed males involves steps catalyzed by P450s, a short chain dehydrogenase, and a cyclase. Previously characterized enzymes are colored blue. From Nadeau et al. (2017)(Chapter 3).
Figure 1.1. MPB pheromone biosynthetic pathways.
Chapter 2
Expression profiling and functional analysis of CYP6DH3

I. Abstract

The mountain pine beetle (MPB, *Dendroctonus ponderosae* Hopkins) is a highly destructive pest of pine forests in western North America. When colonizing a new host tree, MPB coordinate a mass attack to overcome the tree’s defenses using aggregation pheromones. Once feeding commences, detoxification enzymes are required to survive toxic resins secreted by the tree. Expression patterns and substrate profiles can be helpful in determining whether an enzyme is used in pheromone biosynthesis or resin detoxification. Two closely related MPB cytochrome P450s (P450s), CYP6DH1 and CYP6DH2, were previously characterized to have expression patterns and substrate profiles consistent with pheromone biosynthesis and resin detoxification, respectively, through hydroxylation of host monoterpenes. A third member of the same P450 family, CYP6DH3, was hypothesized to have similar monoligneous hydroxylating activity because of the close amino acid sequence similarity between it and CYP6DH1/2 (approximately
68% identity). Quantitative real-time PCR (qRT-PCR) measurements of CYP6DH3 mRNA levels in male and female tissues and in response to monoterpane exposure were inconsistent with a role in either resin detoxification or pheromone biosynthesis, although its substrate range was very similar to CYP6DH2 with the substrates (+)-α-pinene, 3-carene and R-(+)-limonene yielding one or more products. Not all products could be positively identified through GC-MS but they are likely to be hydroxylation products based on the activity of closely related CYP6DH1/2. The expression patterns of CYP6DH1, CYP6DH2 and CYP6DH3 suggest very different regulatory mechanisms for each of these closely related P450s, highlighting the complex nature of these enzymes and their evolutionary adaptability to a number of roles within the host organism.

II. Introduction

The mountain pine beetle is a highly destructive pest of western pine forests of North America and has killed millions of acres of trees over the last few decades (United States Department of Agriculture, 2017). The primary host tree for MPB is the lodgepole pine (Pinus contorta), although MPB may attack any pine trees in their range (Gibson et al., 2009). In coordination with symbiotic fungal partners, MPB kill their host trees by feeding on the phloem and blocking water transport in the xylem (Gibson et al., 2009). MPB use coordinated mass attacks to overcome host tree defenses (Blomquist et al., 2010). Under normal conditions the host trees are able to produce copious amounts or resin, composed of toxic mono-, di- and sesquiterpenes, allowing for the tree to “pitch out” attacking MPB (Boone et al., 2011). However, during drought conditions, the trees are unable to produce sufficient resin, allowing for
successful MPB attacks and often leading to mass outbreaks of MPB (Williams and Liebhold, 2002; Negrón et al., 2009).

Successful MPB attacks require coordinated information from three pheromone components that, in synergy with host monoterpenes, create aggregation and anti-aggregation signals (Blomquist et al., 2010). The first component produced during an attack is (−)-trans-verbenol [(1S,2R,5S)-4,6,6-trimethylbicyclo[3.1.1]hept-3-en-2-ol], which is synthesized by feeding (pioneer) females. (−)-trans-Verbenol is an aggregation pheromone that attracts other males and females to a newly colonized host tree (Pitman et al., 1968). In synergy with the female produced (−)-trans-verbenol, newly emerged male MPB produce exo-brevicomin [(1R,5S,7R)-7-Ethyl-5-methyl-6,8-dioxabicyclo[3.2.1]octane] in small concentrations as an aggregation pheromone component, with production decreasing significantly upon reaching the new host tree and mating (Pureswaran et al., 2000). The third component, frontalin [(1R,5R)-1,5-dimethyl-6,8-dioxabicyclo[3.2.1]octane], is produced by males upon feeding and acts as an anti-aggregation signal to halt the attack and prevent overcrowding of the tree (Barkawi et al., 2003).

While aggregation pheromones are critical for mass attacks and successful reproduction, detoxification enzymes are also essential for MPB to survive the toxic environment of the host tree (Keeling and Bohlmann, 2006; Robert et al., 2013). Insect P450s are involved in oxidation of toxic terpenoids to allow for easier excretion (Li et al., 2004). It has been hypothesized that pheromone biosynthetic P450s evolved from detoxification P450s because hydroxylated monoterpenes are both detoxification products and pheromone components. Enzymes involved in the two processes may be distinguished by expression patterns and substrate profiles (Seybold, Bohlmann and Raffa, 2000; Blomquist et al., 2010). For instance, detoxification P450s are predicted to have a broad substrate range and to be induced by exposure to toxic terpenoids.
through feeding or inhalation, while pheromone biosynthetic P450s are likely to have a narrower substrate range with strong preference for the pheromone precursor and be induced by feeding or hormones (Blomquist et al., 2010; Song et al., 2013).

Biosynthesis of (-)-trans-verbenol likely occurs through P450-catalyzed hydroxylation of the host tree produced monoterpene (-)-α-pinene (Figure 3.1) during colonization (Pureswaran et al., 2000), although possible sequestration of (-)-trans-verbenol during larval and pupal stages for release during colonization has been hypothesized (Keeling et al., 2016). Despite (-)-trans-verbenol being the aggregation pheromone, female MPB also have the ability to produce (+)-trans-verbenol from (+)-α-pinene, and evidence suggests two different enzymes may be involved, one that hydroxylates both (-)-α-pinene and (+)-α-pinene to produce (-)-trans-verbenol and (+)-trans-verbenol, respectively, while another is specific for (-)-α-pinene and produces only (-)-trans-verbenol (Pierce et al., 1987). This suggests that the enzyme system that accepts both enantiomers is part of a detoxification capability that is distinct from the pheromone-biosynthetic enzyme(s) accepting only the (-)-enantiomer.

As part of an effort to identify the enzymes catalyzing α-pinene hydroxylation, (Robert et al., 2013) noted that CYP6DH1 expression is highest in feeding female MPB, which correlates with a role in (-)-trans-verbenol production. Subsequent functional analyses of recombinant CYP6DH1 confirmed this enzyme accepted (-)-α-pinene and (+)-α-pinene as substrates, producing trans-verbenol from each, but did not accept other monoterpene substrates (Song, 2012). These findings suggest a role in pheromone biosynthesis for CYP6DH1 because of its narrow substrate range and high expression in feeding females. CYP6DH2 is a close paralog with only a 16 amino acid difference with CYP6DH1 (Aw et al., 2010; Gorzalski, 2010; Song, 2012). Microarray and qRT-PCR analyses showed that CYP6DH2 is highly expressed in both
male and female MPB and expression is induced by monoterpene exposure (Aw et al., 2010; Gorzalski, 2010). Functional characterization of recombinant CYP6DH2 showed products produced from the monoterpene substrates (−)-α-pinene, (+)-α-pinene, 3-carene and R-(+)-limonene, with (−)-trans-verbenol and myrtenol being the products obtained from (−)-α-pinene and (+)-α-pinene respectively (Song, 2012). These findings suggest CYP6DH2 is likely a detoxification P450 because of its broad substrate range, expression in both males and females, and induction upon monoterpene exposure. Furthermore, the data support the interpretation that CYP6DH1 may have evolved from CYP6DH2 through gene duplication and divergent evolution, and explain the biochemical basis for the two distinct metabolic pathways inferred by Pierce et al (1987).

The interpretation above implies that CYP6DH1 and CYP6DH2 are distinct loci, but the situation is complicated by the possibility that CYP6DH1 and CYP6DH2 are allelic. CYP6DH2 has not been reported in northern (Canadian) populations, while extensive efforts to identify and characterize CYP6DH1 failed to detect this gene in in southern (US) populations (Tittiger, unpublished observation). Thus, the two genes appear to be geographically isolated. This leads to the suggestion that other enzyme(s) provide the missing activities for each population. One obvious candidate for this role is CYP6DH3, which shares 68% amino acid identity with CYP6DH1 and CYP6DH2, and is found in both northern and southern populations.

This study examines CYP6DH3. We hypothesized that CYP6DH3 is also involved in monoterpene hydroxylation because of its sequence similarity to CYP6DH1 and CYP6DH2, both of which hydroxylate monoterpens. Tissue distribution profiles of CYP6DH3 mRNA were conducted for both sexes and feeding status using qRT-PCR. CYP6DH3 expression profiles for monoterpene-exposed male and female MPB were also analyzed using qRT-PCR. Finally,
functional characterization of recombinant CYP6DH3 was accomplished using a baculoviral expression system in Sf9 cells and GC-MS to analyze reactions with various substrates. The data obtained in this study will be useful for understanding P450 structure and function and P450 evolution in mountain pine beetles.

III. Methods

III.1 Insects
Sections of mountain pine beetle-attacked lodgepole pine were collected from the Whittell Forest, located in the Carson Range on the east slope of the Sierra Nevada (approx. N 39°16’29” W 119°52’43’”). The beetles overwintered in the bolts and emerging adults were collected and sexed as reported previously (Aw et al., 2010; Song et al., 2014). For feeding experiments, fresh lodgepole pine bolts were obtained from the Whittell Forest and stored at 4°C prior to use. Females were fed by drilling small holes just beneath the bark, inserting the beetles head first, stapling a wire mesh over the occupied hole, incubating the bolt vertically for 24 h in the dark at room temperature and collecting the live, fed beetles by stripping the bark. Fresh frass indicated that the beetles had fed. Males were fed using the same method except females were first placed head first into the holes for 24 h, followed by insertion of the males head first for 24 h and subsequent collection and sexing of both. Unfed beetles were incubated for 24 h in 2 oz. plastic cups with perforated lids in a dark drawer kept humid with small flasks filled with water and a wet paper towel.
III.2 Tissue Expression Profiling

The tissue distribution of CYP6DH3 mRNA was assayed by quantitative (real time) PCR (qRT-PCR). All beetles were dissected following treatment to collect head, carcass, fat body, midgut, posterior midgut and hindgut tissues. Four replicates of each pooled tissue from 10 beetles were collected for each of the four treatments (fed and unfed males and females) for a total of 96 samples. Tissues were immediately frozen in liquid nitrogen and stored at -80°C for subsequent RNA extraction. RNA was extracted from the tissues using an RNeasy Plant Mini Kit (Qiagen) as per the manufacturer’s instructions. Aliquots consisting of approximately 500 ng of total RNA from each sample were used to make cDNA using iScript™ Reverse Transcription Supermix (Bio-Rad, Hercules, CA) as per the manufacturer’s instructions. Relative levels of CYP6DH3 mRNA were analyzed for each tissue from four different treatment groups (Fed Male (FM), Fed Female (FF), Unfed Male (UM), and Unfed Female (UF)) using qRT-PCR. PCR was conducted in a 20 µL reaction consisting of iTaq™ Universal SYBR® Green Supermix (Bio-Rad) and 2 µL of template for 40 cycles of 95°C for 5 s and 60°C for 30 s on a Bio-Rad CFX96 Real-Time PCR Machine (Bio-Rad). qRT-PCR Primers for CYP6DH3 (Table 2.1) were designed using IDT Primer Quest and melt curves were inspected to ensure primer specificity and proper PCR temperature cycling parameters. For each cDNA sample the PCR reactions were conducted in triplicate and relative target gene expression was normalized to that of Tubulin, measured using DpoTubulinF1 and DpoTubulinR1 (Table 2.1), to produce the ΔCT value. The effects of gender, treatment (fed or unfed) and tissue were examined by three-way ANOVA using GraphPad Prism version 7.03 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. Statistical significance was determined based on α=0.05.
III.3 qRT-PCR of Monoterpene Exposed Beetles

Beetles were separated by sex and placed in two oz. plastic cups with perforated lids and incubated in a humidified dark drawer for 24 h as described previously (Song et al. 2014) to ensure they were unfed at the beginning of the monoterpene exposures. Small clumps of glass wool were placed in four pyrolyzed 500 mL glass jars, two of which contained two mL vials capped with 500 µL of a selected monoterpene and a cotton mesh lid, or two jars with no vials containing monoterpenes as a control. Eight males or females were transferred into each jar so that each sex had monoterpene-exposed and control treatments. The jars were incubated in the dark for 24 h, and the beetles (two or three from each of three replicate incubations) were then placed in microcentrifuge tubes and flash frozen in liquid nitrogen. Six different monoterpenes were tested: 3-carene, R-(+)-limonene, myrcene, (+)-α-pinene, (–)-α-pinene, terpinolene. We also tested a monoterpene cocktail containing all six listed monoterpenes. RNA was extracted from the whole beetles using the RNeasy Plant Mini Kit from Qiagen as described above with an on-column DNase treatment. qRT-PCR was conducted using CYP6DH3 primers and normalized to Tubulin as described above. The effects of gender and treatment (monoterpene exposure) were examined by two-way ANOVA using GraphPad Prism version 7.03 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. Statistical significance was determined based on α=0.05.

III.4 Recombinant CYP6DH3 Production

The CYP6DH3 (GenBank i.d. AF145039.1) open reading frame (ORF) was amplified by PCR using CYP6DH3F6 and CYP6DH3R3 primers (Table 1) and CloneAmp HiFi PCR Premix (Takara Bio USA, Inc., Mountain View, CA) from first strand cDNA prepared from mature
male MPB, cloned into pENTR4 modified to remove the NcoI site (Sandstrom et al., 2006) by Gibson assembly using the In-Fusion® HD Cloning Kit (Takara Bio USA, Inc.), and transformed into *E. coli* Stellar™ Competent Cells (Takara Bio USA, Inc.). Recombinant plasmid was confirmed by sequencing prior to recombination into BaculoDirect™ C-Term Linear DNA (Invitrogen) by LR Clonase™ II (Invitrogen). The recombinant BaculoDirect clone was transferred into Sf9 cells by transfection using Cellfectin II (Invitrogen) and amplified by successive infections of P1 and P2 viral stocks to a high-titer P3 viral stock. Protein expression was initiated by infecting 50 mL of 10⁶ cells/mL Sf9 cells in Sf-900 II SFM culture media supplemented with 10% (vol/vol) fetal bovine serum (Atlas Biologicals, Fort Collins, CO), 0.3 mM δ-aminolevulinic acid, and 0.1 mM ferric citrate with 50 µL of the P3 viral stock and incubating at 27°C for 72 h. Recombinant CYP6DH3 and HF-CPR (Song, Delaplain, et al., 2014) were harvested 72 h post infection in a cell lysis buffer (100 mM potassium phosphate, pH 7.6, 20% (vol/vol) glycerol, 1.1 mM EDTA, 200 µM PMSF, and protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO) and the microsomes were isolated by differential centrifugation essentially as per Nadeau et al. (2017). The microsomal fraction was tested for functional CYP6DH3 using a CO-difference spectrum analysis (Choi et al., 2003) as described previously (Song et al., 2013).

### III.5 Enzyme Assays

3-Carene, *R-(+)-limonene, S-(–)-limonene, (+)-α-pinene, (–)-α-pinene, γ-terpinene, β-phellandrene, terpinolene, *p*-cymene, abeitic acid and myrcene were obtained from Sigma-Aldrich (St. Louis, MO). Reaction mixtures consisted of 200 µL of the recombinant P450 microsomal preparation, 40 µL of HF-CPR microsomal preparation, 100 mM sodium phosphate
buffer pH 7.6, 1.5 mM NADPH and 21 mM substrate (final concentrations) in a total volume of 602 μL. Control reactions containing only HF-CPR were identical to the experimental reactions except that reaction buffer was substituted for the CYP6DH3 microsomal preparation. Reactions were incubated in a capped 5 mL glass vial and rotated lengthwise at 30 °C in a Fisher Biotech Hybridization Incubator (Thermo Fisher Scientific, Waltham, MA) for three hours. The reactions were terminated and extracted twice using 750 μL of pentane:ether (1:1) and dried down to approximately 100 μL under a gentle stream of nitrogen. The extracts were analyzed by GC-MS on a HP-5ms capillary column (Agilent; Santa Clara, CA) using an Agilent 7890B gas chromatograph coupled to a 5977A mass spectrum detector. The instrument running parameters were: initial temperature of 40 °C with a one min hold, 5 °C/min ramp to 240 °C and 15 °C/min ramp to 300 °C with a 5 min hold. The MS detector was a single quadrapole with an electron ionization source and a molecular weight scanning range of 40 to 700 atomic mass units (amu) and an ionization potential of 70 eV. Mass spectra were compared to the main EI MS library (NIST08) to confirm the identity of the products. Each functional assay was repeated twice to confirm products.

IV. Results

IV.1 Tissue Expression Profiles

Relative CYP6DH3 mRNA levels were measured in fed and unfed male and female MPB tissues using qRT-PCR. Three-way ANOVA analysis of CYP6DH3 mRNA levels normalized to Tubulin revealed significant variation in different tissues (F_{5,72}=2.67, p=0.0287) and between treatments (fed or unfed) (F_{1,72}=38.28, p<0.0001) (Table 2.2). Gender alone did not yield
significant variation in expression levels but tissue x treatment (F\textsubscript{5,72}=9.79, p<0.0001) and tissue x gender (F\textsubscript{5,72}=5.92, p=0.0001) interactions were significant (Table 2.2). CYP6DH3 expression levels were lowest in fed male and female midguts, posterior midguts and hindguts and in unfed female fat bodies (Figure 2.2).

IV.2 Monoterpene Exposure

The effects of exposure to various monoterpenes on relative CYP6DH3 mRNA levels in unfed male and female MPB were measured by qRT-PCR. Analysis of CYP6DH3 mRNA levels normalized to Tubulin revealed a significant variation in CYP6DH3 mRNA levels after monoterpene exposure (Two-way ANOVA F\textsubscript{7,48}=15.86, p<0.0001) and the interaction between monoterpene exposure and gender (F\textsubscript{7,48}=67.59, p<0.0001) (Table 2.3). Relative CYP6DH3 mRNA levels in un-exposed (control) males and females were not significantly different. However, CYP6DH3 mRNA levels were highest in males and lowest in females following exposure to (+)-α-pinene (Figure 2.3). Differences in males and females were also observed following exposure to a mixture of monoterpenes (“mono-cocktail”), but exposure to all tested monoterpenes except (+)-α-pinene showed little or no difference in CYP6DH3 mRNA levels (Figure 2.3).

IV.3 Functional Analysis of CYP6DH3

A BLASTP analysis of CYP6DH3 against CYP6DH1 (GenBank i.d. AFI45037.1) and CYP6DH2 (GenBank i.d. AFI45038.1) showed a 69% and 68% amino acid identity respectively (BLASTP 2.7.1; https://blast.ncbi.nlm.nih.gov, Altschul et al. 1997). Recombinant CYP6DH3 showed the 450 nm CO difference peak characteristic of active P450s (Figure 2.4). Recombinant
CYP6DH3 yielded no unique products when compared to control incubations with CPR alone when incubated with S-(−)-limonene, (−)-α-pinene, γ-terpinene, β-phellandrene, terpinolene, p-cymene, abeitic acid and myrcene as potential substrates (data not shown). Incubations of CYP6DH3 with (+)-α-pinene yielded one unique product at 15.23 minutes compared to control reactions, but the product could not be identified based on the mass spectrum data (Figures 2.4A-C). Incubations with 3-carene yielded two unique products at 15.15 and 16.15 minutes when compared to controls; again, the identities of these products could not be determined using MS data (Figures 2.5A-D). Incubations with R-(+)-limonene yielded three unique products at 16.02, 16.24 and 19.88 minutes when compared to CPR only reactions (Figures 2.6A & B). The mass spectra for the GC peaks at 16.02 and 16.24 minutes were consistent with those for p-mentha-1(7),8(10)-dien-9-ol and (S)-perillyl alcohol, respectively (Figures 2.6C & D). The third unique peak at 19.88 minutes could not be identified using MS data (Figure 2.6E).

V. Discussion

The CYP6DH family of P450s in the MPB has three relatively closely related members, two of which, CYP6DH1 and CYP6DH2, have been characterized based on expression and catalytic activity. They are likely involved in pheromone biosynthesis and resin detoxification, respectively (Aw et al., 2010; Gorzalski, 2010; Song, 2012). As both CYP6DH1 and CYP6DH2 catalyze the hydroxylation of monoterpenes, we hypothesized that the third member encoded by this family, CYP6DH3, also hydroxylates monoterpenes. We found that CYP6DH3 does in fact hydroxylate some, but not all, the monoterpenes we investigated (Figures 2.4 – 6).
mRNA expression profiles can be indicative of involvement in pheromone production or resin detoxification in bark beetles (Blomquist et al., 2010; Robert et al., 2013; Song et al., 2013; Keeling et al., 2016; Nadeau et al., 2017). Differences in mRNA levels based on tissue, gender and/or feeding status can indicate involvement in pheromone biosynthesis, while elevated mRNA levels in response to monoterpene exposures suggest likely involvement in resin detoxification. We found that CYP6DH3 mRNA levels are unaffected by differences in gender but are overall significantly higher in unfed compared to fed beetles (Figure 2.1), with decreased levels in the midgut, posterior midgut and hindgut of feeding beetles compared to the same tissues of unfed beetles. This pattern is not consistent with pheromonal (−)-trans-verbenol production. Similarly, while the neutral nature of CYP6DH3 expression based on gender is indicative of resin detoxification, increased expression in unfed beetles and lower expression in the digestive tract of the feeding beetles does not necessarily coincide with that conclusion.

CYP6DH3 mRNA profiles in response to monoterpene exposure of unfed beetles also suggest a complex regulatory mechanism for CYP6DH3 that may not agree with the usual dogma of expression profiles for pheromone biosynthetic or detoxification enzymes. Monoterpene exposure caused a significant increase in CYP6DH3 mRNA levels in both males and females with no significant difference overall between genders. However, there was a highly significant interaction between monoterpene exposure and gender as is evident from exposure to (+)-α-pinene, with males having increased CYP6DH3 expression compared to control beetles and females having highly decreased expression levels (Figure 2.2). Alternatively, when exposed to a monoterpene cocktail that combined all the tested monoterpenes, mRNA levels in females were higher than those in controls, while males had lower CYP6DH3 mRNA levels in general. The complex regulatory regimes affecting CYP6DH3 mRNA levels are consistent with gender-based
differences in expression of P450s in response to monoterpene exposures for both *Dendroctonus armandi* (Dai *et al.*, 2015) and MPB (Nadeau *et al.*, 2017). Taking both tissue and monoterpene exposure profiles into account, CYP6DH3 expression does not correlate well with the expected expression pattern of either pheromone biosynthesis (gender and tissue specific expression increases) or resin detoxification (increased expression for monoterpene exposures compared to the control).

Although expression patterns do not correlate well with those expected of a resin detoxification enzyme, functional analysis of CYP6DH3 shows a relatively broad substrate range similar to CYP6DH2. CYP6DH2 is thought to be a resin detoxification P450 and catalyzes the hydroxylation of (–)-α-pinene, (+)-α-pinene, 3-carene and R-(+)-limonene (Song, 2012), while CYP6DH3 similarly accepts the same monoterpenes, except (–)-α-pinene, as substrates. Many of the CYP6DH3 products could not be identified solely based on the MS data. Two identified products from R-(+)limonene were hydroxylation products, so it is likely that the other substrates are hydroxylated as well, though an epoxidation across the double bond is also possible. Pierce *et al.* (1987) noted that monoterpene hydroxylations localize to allylic carbons adjacent to carbon-carbon double bonds. The two major hydroxylation products of α-pinene based on this rule are *trans*-verbenol and myrtenol, and these were observed by Song (2012) as products of incubations with CYP6DH1 and CYP6DH2. Based on the MS data, the product of (+)-α-pinene following incubation with CYP6DH3 in this study is neither verbenol nor myrtenol. Hydroxylation of carbon 1 to form the tertiary alcohol is also formally possible, however the spectrum for the observed product is not consistent with this. Given that ring rearrangements are relatively common (Rücker *et al.*, 1990), it is difficult to confidently identify the product of *trans*-verbenol beyond saying that is likely is neither verbenol or myrtenol.
It is not necessarily surprising that even though CYP6DH3 is closely related (68% amino acid identity) to CYP6DH1/2, their substrate ranges and/or products are not identical. Small differences in the amino acid composition of the solvent channel, substrate binding site, or even other locations that affect tertiary structure can cause vast differences in P450 substrate specificity and product formation (Li et al., 2004; Denisov et al., 2005; Bernhardt, 2006). The expression pattern of CYP6DH1, CYP6DH2 and CYP6DH3 also suggest very different regulatory mechanisms for each of these closely related P450s, highlighting the complex nature of these enzymes and their evolutionary adaptability to a number of roles within the host organism. In this respect, it is worth remembering that distinguishing between “pheromone biosynthetic” and “resin detoxifying” as role descriptors may not be accurate in this case. Unlike other insects that have well-defined cells or glands dedicated to pheromone production (Tillman et al., 1998), both processes occur in the MPB midgut along with digestion. (−)-trans-Verbenol production requires only a single metabolic step, compared to the fully-developed pathways for frontalin, exo-brevicomin (Figure 1.1), and ipsdienol (Sandstrom et al., 2006). The amount of (−)-trans-verbenol produced (less than 40 ng/female) (Pureswaran et al., 2000) is ten-fold less than that for ipsdienol in male Ips pini. Thus, with the metabolic load being lower in midgut cells compared to Ips bark beetles or pheromone-producing cells in other insects, evolutionary pressure to “focus” enzyme activity on pheromone production over detoxification is likely lower. This may explain the complex regulatory processes evident for the CYP6DH genes in MPB. Further work is needed to positively identify all the products of CYP6DH3 and kinetic data would also be helpful in comparing the activities of the whole CYP6DH family.
VI. References


mechanisms.’, *PloS one*. Public Library of Science, 8(11), p. e77777. doi: 10.1371/journal.pone.0077777.


VII. Figure Legends

**Figure 2.1. Tissue distribution profile.** Tissue distribution profiles of CYP6DH3 mRNA in fed or unfed males and females using qRT-PCR. Data are normalized to Tubulin and presented as the mean +/- standard deviation, n=4.
**Figure 2.2. Monoterpene exposures.** CYP6DH3 mRNA levels in male and female MPB following monoterpene exposures. Values are normalized to Tubulin and presented as the mean +/- the standard deviation, n=4.

**Figure 2.3. CO difference spectrum.** CO difference spectrum of CYP6DH3 showing the characteristic absorbance peak at 450 nm for active P450s.

**Figure 2.4. Functional assay with (+)-α-pinene.** GC chromatograms of incubations with (+)-α-pinene and microsomes containing (A) CYP6DH3 or (B) CPR only. The large peak at 9.93 min is a 10 ng/µL n-octanol standard. (C) MS scan of the 15.23 minute peak in (A) representing an unknown product.

**Figure 2.5. Functional assay with 3-carene.** GC chromatograms of incubations with 3-carene and microsomes containing (A) CYP6DH3 or (B) CPR only. The large peak at 9.93 min is a 10 ng/µL n-octanol standard. MS scans for the peaks at (C) 15.15 minutes and (D) 16.15 minutes could not be matched in the (describe library).

**Figure 2.6. Functional assay with R-(+)-limonene.** GC chromatograms of reactions with R-(+)-limonene and microsomes containing (A) CYP6DH3 or (B) CPR only. The large peak at 9.93 min is a 10 ng/µL n-octanol standard. MS scans for the peaks at (C) 16.02 min and (D) 16.24 min are most similar to those for p-mentha-1(7),8(10)-dien-9-ol and (S)-perillyl alcohol, respectively. The MS scan shown in (E) for the peak at 19.88 min is an unknown product.
Table 2.1. Cloning and qRT-PCR primers.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cloning Primers</strong></td>
<td></td>
</tr>
<tr>
<td>CYP6DH3F6</td>
<td>TCACTTGCAGTTGAGGCATTC</td>
</tr>
<tr>
<td>CYP6DH3R3</td>
<td>TTTCACTTCACTTCAACTCGC</td>
</tr>
<tr>
<td><strong>qRT-PCR Primers</strong></td>
<td></td>
</tr>
<tr>
<td>YQE_01840F</td>
<td>GATTGCTGCCTGCTTCTTTG</td>
</tr>
<tr>
<td>YQE_01840R</td>
<td>CGTGGTTCGTATTACATTTCC</td>
</tr>
<tr>
<td>DpoTubulinF1</td>
<td>GACAATCTGGAGCAGGAAACAT</td>
</tr>
<tr>
<td>DpoTubulinR1</td>
<td>TGCCCTTTTTCGCAAACATC</td>
</tr>
</tbody>
</table>
Table 2.2. Three-way ANOVA of $\Delta$CT(Tubulin-CYP6DH3) values for CYP6DH3 tissue distribution analysis. SS=sum of squares, DF=degrees of freedom, MS=mean square.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F (DFn, DFd)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tissue</td>
<td>18.53</td>
<td>5</td>
<td>3.705</td>
<td>$F(5, 72) = 2.668$</td>
<td>P=0.0287</td>
</tr>
<tr>
<td>Treatment</td>
<td>53.16</td>
<td>1</td>
<td>53.16</td>
<td>$F(1, 72) = 38.28$</td>
<td>P&lt;0.0001</td>
</tr>
<tr>
<td>Gender</td>
<td>3.375</td>
<td>1</td>
<td>3.375</td>
<td>$F(1, 72) = 2.43$</td>
<td>P=0.1234</td>
</tr>
<tr>
<td>Tissue x Treatment</td>
<td>67.95</td>
<td>5</td>
<td>13.59</td>
<td>$F(5, 72) = 9.787$</td>
<td>P&lt;0.0001</td>
</tr>
<tr>
<td>Tissue x Gender</td>
<td>41.08</td>
<td>5</td>
<td>8.217</td>
<td>$F(5, 72) = 5.917$</td>
<td>P=0.0001</td>
</tr>
<tr>
<td>Treatment x Gender</td>
<td>0.8363</td>
<td>1</td>
<td>0.8363</td>
<td>$F(1, 72) = 0.6022$</td>
<td>P=0.4403</td>
</tr>
<tr>
<td>Tissue x Treatment x Gender</td>
<td>11.68</td>
<td>5</td>
<td>2.336</td>
<td>$F(5, 72) = 1.682$</td>
<td>P=0.1499</td>
</tr>
<tr>
<td>Residual</td>
<td>99.98</td>
<td>72</td>
<td>1.389</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.3. Two-way ANOVA of ΔCT(Tubulin-CYP6DH3) values for CYP6DH3 induction by monoterpenes exposures.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F (DFn, DFd)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interaction</td>
<td>120.2</td>
<td>7</td>
<td>17.18</td>
<td>F (7, 48) = 67.59</td>
<td>P&lt;0.0001</td>
</tr>
<tr>
<td>Monoterpene Exposure</td>
<td>28.21</td>
<td>7</td>
<td>4.03</td>
<td>F (7, 48) = 15.86</td>
<td>P&lt;0.0001</td>
</tr>
<tr>
<td>Gender</td>
<td>0.4647</td>
<td>1</td>
<td>0.4647</td>
<td>F (1, 48) = 1.828</td>
<td>P=0.1826</td>
</tr>
<tr>
<td>Residual</td>
<td>12.2</td>
<td>48</td>
<td>0.2541</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 2.1. Tissue distribution profile.
Figure 2.2. Monoterpene exposures.
Figure 2.3. CO difference spectrum.
Figure 2.4. Functional assay with (+)-α-pinene.
Figure 2.5. Functional assay with 3-carene.
Figure 2.6. Functional assay with R-(+)-limonene.
Chapter 3

This chapter is presented as published in:


Title:

Comparative transcriptomics of mountain pine beetle pheromone-biosynthetic tissues and functional analysis of CYP6DE3.

Authors:

Nadeau, J. A.¹, jnadeau20@sbcglobal.net

J. Petereit², jpetereit@unr.edu

R. L. Tillett³, rtillett@unr.edu

K. Jung¹, khjung37@yahoo.com

M. Fotoohi¹, mishafotoohi@yahoo.com

M. MacLean¹, marinam@usamedia.tv

S. Young¹, sharony@unr.edu

K. Schlauch¹, schlauch@unr.edu

G. J. Blomquist¹, garyb@cabnr.unr.edu

and C. Tittiger¹*, crt@unr.edu

¹ Department of Biochemistry and Molecular Biology, University of Nevada, Reno, NV, 89557, USA.
ABSTRACT

Background: The mountain pine beetle (MPB, *Dendroctonus ponderosae* Hopkins) is a highly destructive pest of pine forests in western North America. During flight to a new host tree and initiation of feeding, mountain pine beetles release aggregation pheromones. The biosynthetic pathways of these pheromones are sex-specific and localized in the midgut and fat body, but the enzymes involved have not all been identified or characterized.

Results: We used a comparative RNA-Seq analysis between fed and unfed male and female MPB midguts and fat bodies to identify candidate genes involved in pheromone biosynthesis. The 13,407 potentially unique transcripts showed clear separation based on feeding state and gender. Gene co-expression network construction and examination using *petal* identified gene groups that were tightly connected. This, as well as other co-expression and gene ontology analyses, identified all four known pheromone biosynthetic genes, confirmed the tentative identification of four others from a previous study, and suggested nine novel candidates. One cytochrome P450 monooxygenase, CYP6DE3, identified as a possible *exo*-brevicomin-biosynthetic enzyme in this study, was functionally characterized and likely is involved in resin detoxification rather than pheromone biosynthesis.

Conclusions: Our analysis supported previously characterized pheromone-biosynthetic genes involved in *exo*-brevicomin and frontalin biosynthesis and identified a number of candidate
cytochrome P450 monooxygenases and a putative cyclase for further studies. Functional analyses of CYP6DE3 suggest its role in resin detoxification and underscore the limitation of using high-throughput data to tentatively identify candidate genes. Further functional analyses of candidate genes found in this study should lead to the full characterization of MPB pheromone biosynthetic pathways and the identification of molecular targets for possible pest management strategies.

KEYWORDS
transcriptomics, RNA-Seq, bark beetle, Dendroctonus, pheromone, P450

BACKGROUND

The mountain pine beetle (MPB) (Dendroctonus ponderosae Hopkins) uses three main pheromone components to coordinate the “mass attack” necessary to overcome a host tree’s defenses. Each component has a distinctive role and is produced from a different metabolic pathway (Figure 1). Females produce the aggregation pheromone (−)-trans-verbenol [(1S,2R,5S)-4,6,6-trimethylbicyclo[3.1.1]hept-3-en-2-ol], to attract other males and females to a newly colonized host tree [1]. Biosynthesis of (−)-trans-verbenol likely requires cytochrome P450-mediated hydroxylation of the host tree produced monoterpene, (−)-α-pinene. Newly emerged males produce the aggregation pheromone exo-brevicomin [(1R,5S,7R)-7-ethyl-5-methyl-6,8-dioxabicyclo[3.2.1]octane] from long chain fatty acid precursors in the fat body [2,3]. exo-Brevicomin production decreases substantially upon reaching the new host tree and mating [4]. The third major component, frontalin [(1R,5R)-1,5-dimethyl-6,8-dioxabicyclo[3.2.1]octane], is produced by feeding males and acts as an anti-aggregation signal to halt the attack and prevent overcrowding of the tree [4]. Frontalin biosynthesis uses the mevalonate pathway [5].
Various high-throughput studies, including a draft MPB genome [6], and various transcriptomic [7,8] and proteomic [9] analyses have been directed at this highly destructive coniferous forest pest in order to better understand its physiology and to assist development of new management strategies. Aggregation and anti-aggregation pheromones in the MPB are synthesized in the midguts and fat bodies in a sex- and feeding status-specific manner [10]. Therefore, mRNAs encoding enzymes involved in pheromone-biosynthetic pathways may also be differentially expressed based on sex, feeding status, and tissue. Recently, Keeling et al. (2016) [11] reported transcriptome, proteome, and metabolome-level responses of unfed MPB treated with juvenile hormone (JH) III. JH III stimulates aggregation pheromone component production in bark beetles [5,12–15], including frontalin and trans-verbenol in MPB, and several “pheromone-biosynthetic gene” candidates were identified by this study. However, the dynamics of the transcriptional response of JH III-treated insects can differ from those of fed insects [15–17] and the study noted that JH III treatment did not affect exo-brevicomin production. Thus, JH treatment may not induce differential expression of all genes involved in pheromone biosynthesis for the MPB.

Here, we report the results of an RNA-Seq study of MPB midguts and fat bodies based on sex and feeding status. We employed network, differential gene, and gene ontology enrichment analyses to isolate candidates involved in pheromone biosynthesis. Additionally, functional analysis of a candidate gene, \textit{CYP6DE3}, identified by our RNA-Seq analysis for exo-brevicomin biosynthesis, indicates a role in resin detoxification rather than pheromone biosynthesis, reinforcing the need for functional assays to confirm possible roles of enzymes identified via bioinformatics approaches.
MATERIALS AND METHODS

Tissue Collection

Sections of mountain pine beetle-attacked lodgepole pine, *Pinus contorta*, were collected from Truckee, CA, USA, near the Northwoods Clubhouse in the Tahoe Donner housing subdivision (approx. N 39°20'37" W 120°12'54") on September 30, 2013. The beetles overwintered in the bolts and emerging adult beetles were collected and sexed as reported previously [3,18]. For feeding experiments, fresh lodgepole pine bolts were obtained from the Whittell Forest, located in the Carson Range on the east slope of the Sierra Nevada (approx. N 39°16'29" W 119°52'43") in June, 2014 and stored at 4°C prior to use. Females were fed by drilling small holes just beneath the bark, inserting the beetles head first, stapling a wire mesh over the occupied hole, incubating the bolt vertically for 24 h in the dark at room temperature and collecting the live beetles by stripping the bark. Fresh frass indicated that the beetles had fed. Males were fed using the same method except females were first placed head first into the holes for 24 h, followed by insertion of the males head first for 24 h and subsequent collection and sexing of the live beetles. Unfed beetles were incubated in 2 oz plastic cups with perforated lids in a dark drawer kept humid with small flasks filled with water and a paper towel for 24 h. All beetles were dissected following treatment to collect midgut and fat body tissue. Four replicates of pooled tissue from 10 beetles were collected for each of the four treatments (fed and unfed males and females) for a total of 16 samples. Midguts and fat bodies were immediately frozen in liquid nitrogen and stored at -80°C for subsequent RNA extraction.
RNA Extraction and Sequencing

Total RNA was isolated from the midguts and fat bodies pooled from 10 beetles per sample using an RNeasy Plant Mini Kit from Qiagen (Valencia, CA) and treated with RNase-Free DNase Set from Qiagen as described by the supplier's manual. RNA was quantified using Quant-iT™ RiboGreen® reagent and a Labsystems Fluoroskan Ascent fluorescence plate reader. RNA integrity for each sample was determined using an Agilent 2100 Bioanalyzer and a Eukaryote Total RNA Nano Series II chip. Only samples with an RNA Integrity Number of eight or higher were used for sequencing [19]. RNA was precipitated and provided to the Georgia Genomics Facility (GGF) for library preparation and sequencing. GGF confirmed the quality of the total RNA using an Agilent 2100 Bioanalyzer, prepared barcoded cDNA libraries using a Kapa Stranded mRNA-Seq Kit (Wilmington, MA), and sequenced them on the Illumina NextSeq 500 using paired-end sequencing with a NextSeq 2x75 High Output Flow Cell.

Sequence Quality Control

Sequence quality for each sample was characterized using FastQC (v. 0.11.2; http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Sequence pairs were trimmed and filtered for nucleotide-base quality and Illumina adapter sequences using Cutadapt v. 1.8.1 [20], with options set as follows: trimming of low-quality (phred quality ≤ 10) and “N” base calls from both ends of each read and removal of sequences with trimmed length < 35 nt.

Sequence Alignment and Expression Quantification

Before assembly, trimmed sequence pairs were compared to one another using the MaSuRCa `superreads.pl’ script, and if found to intersect (minimum k-mer 41), were combined
into single “super-reads” [21]. Read pairs with no intersection were retained as separate paired-end reads. Single reads and read pairs were aligned to the Ensembl Metazoa (release-25; [22]) _D. ponderosae_ reference genome [6] using the HISAT spliced read alignment tool (v. 0.1.6-beta; [23]). The coordinates of each known _Dendroctonus_ gene and its exons were extracted from the Ensembl Metazoa Gene Transfer Format (GTF) file and supplied to HISAT at time of alignment (via HISAT’s `--known-splicesite-infile’ option). Resultant alignments were compressed from the sequence alignment/map (SAM) format to the binary BAM format [24]. Upon alignment, the raw counts of reads and read pairs aligned to each gene were totaled using the featureCounts tool of the subread package (v. 1.4.6-p1; [25]). Reads were counted once per pair and summarized for gene loci, with only read pairs aligned to a unique transcribed location included in the count totals.

Transcripts underwent two filtering steps. First, those with no counts for thirteen or more out of the sixteen (75%) samples were excluded. Then, all transcripts with less than 10 fragments (counts) observed in all sixteen samples were excluded. Data were normalized using the standard median ratio method for RNA-Seq data [26]. Principal component analysis (PCA) was performed on the normalized and filtered zero-centered counts per million data using singular value decomposition to validate clear separation between gender and feeding status of the biological replicates of MPBs.

**Annotation**

_Dendroctonus ponderosae_ gene descriptors and annotated Interpro protein domains [27] were obtained from Ensembl Metazoa, via the BioMart interface [28]. Interpro2GO file (v. 2016/03/19 11:04:26) was used to map Interpro IDs to gene ontology (GO) terms.
Validation of RNA-Seq Data by Quantitative PCR

To validate RNA-Seq data, the transcript levels of 15 genes (Table 2) were examined by quantitative reverse transcriptase PCR (qRT-PCR). Genes were chosen by their notable differential expression between feeding states in male beetles. Aliquots consisting of approximately 500 ng of total RNA from a subset of samples (fed male replicate 4, unfed male replicate 4, fed female replicate 4, unfed female replicate 4, fed male replicate 1, unfed male replicate 3, unfed female replicate 2) were used to make cDNA using iScript™ Reverse Transcription Supermix (Bio-Rad, Hercules, CA). PCR was conducted in a 20 µL reaction consisting of iTaq™ Universal SYBR® Green Supermix (Bio-Rad) and 2 µL of template for 40 cycles of 95°C for 5 s and 60°C for 30 s on a Bio-Rad CFX96 Real-Time PCR Machine (Bio-Rad). Primers designed to amplify specific transcripts of these genes were designed using IDT Primer Quest and melt curves were produced to ensure primer specificity and proper PCR temperature cycling parameters (Supplemental File 1). For each cDNA sample the PCR reactions were conducted in triplicate and relative target gene expression was normalized to that of YQE_05788, which encodes ribosomal protein S3P. Ribosomal protein S3 is a relatively more stable normalizing gene for qRT-PCR in another beetle, Tribolium castaneum, compared to the more routinely used actin or tubulin genes [29]. Fold change was calculated for each normalized gene in relation to the expression of the unfed male treatment using the $2^{-\Delta\Delta CT}$ method [30]. For each gene, Pearson and Spearman Correlation Coefficients were computed between the seven samples measured by qRT-PCR and RNA-Seq.
Co-Expression Network

Gene co-expression networks are node-edge graphs. Nodes represent genes that are connected by edges if there is an association between genes as defined by a co-expression measure [31]. Structural components of co-expression networks are used to identify densely connected subgraph, called gene modules. Genes within a module share similar expression patterns, thus they are hypothesized to have similar gene function, to share pathway membership, or to be co-regulated. A co-expression network of the filtered and normalized counts per million of 11,342 mountain pine beetle genes was generated via *petal*, a co-expression network construction and analysis tool [32]. The entire dataset of 11,342 genes over 16 measures was loaded, along with a list of previously confirmed and hypothesized pheromone biosynthetic genes. No other input was specified.

Differential Gene Analysis

Differential gene expression between the feeding conditions and the genders were examined using DESeq2 [26]. Four comparisons, male fed (MF) vs. female fed (FF), MF vs. male unfed (MU), MU vs. female unfed (FU) and FF vs. FU, were considered using simple contrasts. A multiple testing correction was performed for each of the four comparisons to adjust for the false discovery rate [33]. The two other possible comparisons (MU vs. FF and MF vs. FU) were not considered because they are less likely to inform regarding putative pheromone-biosynthetic genes. Genes with absolute value of the log$_2$ fold change greater than one and an adjusted p-value less than 0.01 were retained for further analysis. Venn diagrams were prepared within R to visualize the intersection of the statistically-significant differentially-expressed genes between the considered comparisons (Supplemental File 2).
Gene Ontology Enrichment Analysis

Gene Ontology (GO) enrichment analyses of statistically-significant differentially-expressed gene groups were conducted to identify over-represented molecular functions and metabolic processes. BiNGO (v. 3.0.3) [34] within Cytoscape (v. 3.3.0) and GO file (v. 1.2 2016/03/01) were utilized. GO terms with adjusted significance values less than 0.05 upon a Benjamini-Hochberg adjustment [33] were considered for further investigation.

Recombinant CYP6DE3 Expression

RNA was extracted from two whole beetles using a RNeasy Plant Mini Kit (Qiagen) as per the manufacturer’s instructions. First strand cDNA was synthesized using Superscript III Reverse Transcriptase (Invitrogen, Carlsbad, CA) as per the manual. The CYP6DE3 open reading frame (ORF) was amplified by PCR using CYP6DE3F1 and CYP6DE3R1 primers (Supplemental File 1) and CloneAmp HiFi PCR Premix (Takara Bio USA, Inc., Mountain View, CA), cloned into pENTR4 modified to remove the NcoI site [35], and transformed into Stellar™ Competent Cells (Takara Bio USA, Inc.). Recombinant plasmid was confirmed by sequencing prior to recombination into BaculoDirect™ C-Term Linear DNA (Invitrogen) by LR Clonase™ II (Invitrogen). The recombinant BaculoDirect clone was transferred into Sf9 cells by transfection using Cellfectin II (Invitrogen) and amplified by successive infections of P1 and P2 viral stocks to a high-titer P3 viral stock. Protein expression was initiated by infecting 50 mL of 10^6 cells/mL Sf9 cells in Sf-900 II SFM culture media supplemented with 10% (vol/vol) fetal bovine serum (Atlas Biologicals, Fort Collins, CO), 0.3 mM δ-aminolevulinic acid, and 0.1 mM ferric citrate with 50 µL of the P3 viral stock and incubating at 27°C for 72 h. Recombinant
CYP6DE3 and housefly cytochrome P450 reductase (HF-CPR) [2] were harvested 72 h post infection in a cell lysis buffer (100 mM sodium phosphate, pH 7.6, 20% (vol/vol) glycerol, 1.1 mM EDTA, 200 µM PMSF, and protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO) and the microsomes were isolated by differential centrifugation. The microsomal fraction was tested for functional CYP6DE3 using a CO-difference spectrum analysis [36].

**Enzyme Assays**

3-Carene, R-(+)-limonene, (+)-α-pinene and (Z)-dec-7-enal were obtained from Sigma-Aldrich. cis-3-Nonene was obtained from MP Biomedicals (Santa Ana, CA). Reaction mixtures consisted of 200 µL of the CYP6DE3 microsomal fraction, 40 µL of HF-CPR microsomal fraction, 100 mM sodium phosphate buffer pH 7.6, 1.5 mM NADPH (Sigma-Aldrich) and 21 mM of substrate in a total volume of 602 µL. Control reactions containing only HF-CPR were identical to the experimental reactions except that the reaction buffer was substituted for the CYP6DE3 microsomal fraction. Reactions were incubated in a capped 5 mL vial and rotated lengthwise at 30 °C in a FisherBiotech Hybridization Incubator (Thermo Fisher Scientific, Waltham, MA) for three hours. The reactions were terminated and extracted using pentane:ether (1:1). The extracts were analyzed by GC-MS on a HP-5ms capillary column (Agilent) using an Agilent (Santa Clara, CA) 7890B gas chromatograph coupled to a 5977A mass spectrum detector. The instrument running parameters were: initial temperature of 40 °C with a one min hold, 5 °C/min to 240 °C and 15 °C/min to 300 °C with a 5 min hold. The MS detector was a single quadrupole with an electron ionization source and a molecular weight scanning range of 40 to 700 atomic mass units and an ionization potential of 70 eV.
qRT-PCR of Monoterpene Exposed Beetles

Beetles were separated by sex and placed in two oz plastic cups with perforated lids and incubated in a humidified dark drawer for 24 h, as described previously [3] to ensure they were unfed at the beginning of the monoterpene exposures. Small clumps of glass wool were placed in four pyrolyzed glass jars, two of which contained two mL vials capped with 500 µL of a selected monoterpene and a cotton mesh lid, or two jars with no vials as a control. Eight live males or females were transferred into each of the four jars so that each sex had monoterpene-exposed and control treatments. The jars were incubated in the dark for 24 h followed by placing 2-3 beetles in each of three replicate microcentrifuge tubes for each treatment and flash freezing them in liquid nitrogen. Six different monoterpenes were tested: 3-carene, \( R-(+)-\)limonene, myrcene, \( (+)-\alpha\)-pinene, \( (-)-\alpha\)-pinene, terpinolene. We also tested a monoterpene cocktail containing all six listed monoterpenes. RNA was extracted from the whole beetles using the RNeasy Plant Mini Kit from Qiagen as described above with an on-column DNase treatment. qRT-PCR was conducted using CYP6DE3 primers and normalized to YQE_05788 (rpS3P) as described above. Statistically significant differences between the means of relative expression between males and females for each gene were measured using an unpaired two-sample t-test at p<0.05.

RESULTS

RNA-Seq Quality Control and Validation

A total of 424,776,657 paired-end reads consisting of at least 76 bp were recovered from the 16 libraries, with reads per library ranging from 18,659,429 to 33,943,439 with a mean of 26,548,541. After processing, the number of reads aligned to the reference genome was
317,944,928 (Table 1), representing 13,407 potentially different transcripts. Of these transcripts a total of 11,342 transcripts passed the two filtering steps as explained in the Materials and Methods. Verification of RNA-Seq by qRT-PCR analysis confirmed similar expression measures between the two platforms. For the 15 genes in the seven samples measured by qRT-PCR and RNA-Seq, Pearson and Spearman Correlation Coefficients averaged to 0.924 and 0.878, respectively (Table 2). The PCA showed a clear separation between feeding states and genders, with almost 70% of the variance explained by feeding states (Figure 3.2).
<table>
<thead>
<tr>
<th>Sample ID</th>
<th>Sample Description</th>
<th>Biological Replicate</th>
<th>Index 1</th>
<th>Index 2</th>
<th>Total Read Processed</th>
<th>Read Parts Processed</th>
<th>Alligned Contigs</th>
<th>Alleled Contigs</th>
<th>Library Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>69</td>
<td>Mountain pine beetle libraries</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 1. Illumina NextSeq500 read processing and mapping results from RNA-seq.
**Table 2.** Correlation coefficients of RNA-Seq and qRT-PCR measured expression levels across seven samples and 15 genes.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>YQE_04793</td>
<td>0.966</td>
<td>0.964</td>
</tr>
<tr>
<td>YQE_11756</td>
<td>0.989</td>
<td>0.821</td>
</tr>
<tr>
<td>YQE_06376</td>
<td>0.992</td>
<td>0.927</td>
</tr>
<tr>
<td>YQE_05062</td>
<td>0.990</td>
<td>0.991</td>
</tr>
<tr>
<td>YQE_11643</td>
<td>0.931</td>
<td>0.929</td>
</tr>
<tr>
<td>YQE_09540</td>
<td>0.747</td>
<td>0.714</td>
</tr>
<tr>
<td>YQE_06803</td>
<td>0.923</td>
<td>1.000</td>
</tr>
<tr>
<td>YQE_06028</td>
<td>0.984</td>
<td>1.000</td>
</tr>
<tr>
<td>YQE_01079</td>
<td>0.802</td>
<td>0.607</td>
</tr>
<tr>
<td>YQE_06277</td>
<td>0.875</td>
<td>0.714</td>
</tr>
<tr>
<td>YQE_01901</td>
<td>0.925</td>
<td>0.893</td>
</tr>
<tr>
<td>YQE_02812</td>
<td>0.975</td>
<td>0.889</td>
</tr>
<tr>
<td>YQE_01868</td>
<td>0.994</td>
<td>0.991</td>
</tr>
<tr>
<td>YQE_01611</td>
<td>0.996</td>
<td>1.000</td>
</tr>
<tr>
<td>YQE_04799</td>
<td>0.945</td>
<td>0.893</td>
</tr>
<tr>
<td><strong>AVERAGE</strong></td>
<td><strong>0.935</strong></td>
<td><strong>0.889</strong></td>
</tr>
</tbody>
</table>

**Co-Expression Network**

*petal* generated a co-expression network model based on Spearman Correlation Coefficient and similarity threshold of 0.808. This model includes 10,661 transcripts (94% of genes from the entire dataset) and follows the well-established biological network structure characteristics: small-world and scale-free [32]. From this model, closely connected gene groups based upon the genes of interest were extracted, resulting in thirteen vicinity networks (VN). Here, a vicinity network (VN) is defined by the genes of interests and their common neighbors. For more detail refer to [32].

Based on the results from *petal*, further in-depth analyses were conducted leading to three distinct gene modules, one representing candidates involved in *exo*-brevicomin biosynthesis (purple), and two representing candidates involved in frontalin biosynthesis...
Gene membership of each module is listed in Supplemental File 3.

**Gene Ontology Enrichment Analysis of Differential Expressed Genes**

Based on the differential gene analysis 3,799 differentially expressed transcripts had a statistically significant ($\log_2$-fold change $>1$) differential gene expression at a significance level of 0.01 after multiple testing adjustment (Table 3, Supplemental File 4). The greatest number of genes with a differential expression matching these criteria occurred in the MU > FU, MU > MF, and MF > MU comparisons with 894, 698, and 638 genes, respectively. The smallest numbers of genes with a differential expression matching our criteria were in the comparisons of FU > MU and FF > MF with 74 and 108 genes respectively.

**Table 3.** Summary of genes differentially expressed with greater than two fold change and adjusted p-value $\leq 0.01$.

<table>
<thead>
<tr>
<th>Comparison</th>
<th># of Genes</th>
<th>Comparison</th>
<th># of Genes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male Fed &gt; Male Unfed</td>
<td>638</td>
<td>Male Unfed &gt; Male Fed</td>
<td>698</td>
</tr>
<tr>
<td>Male Fed &gt; Female Fed</td>
<td>446</td>
<td>Female Fed &gt; Male Fed</td>
<td>108</td>
</tr>
<tr>
<td>Male Unfed &gt; Female Unfed</td>
<td>894</td>
<td>Female Unfed &gt; Male Unfed</td>
<td>74</td>
</tr>
<tr>
<td>Female Fed &gt; Female Unfed</td>
<td>581</td>
<td>Female Unfed &gt; Female Fed</td>
<td>360</td>
</tr>
</tbody>
</table>

Two areas highlighted in the Venn diagram (Figure 3.4) were of particular interest as they represented pools of potential candidate genes for pheromone biosynthesis: 1) the 1225 genes with statistically significant differential expression between MU vs. FU beetles are *exo*-brevicomin-biosynthetic candidates; 2) the 217 genes with statistically significant differential expression in MF vs. MU and MF vs. FF, excluding differentially expressed genes in MU vs. FU, are potential frontaldehyde-biosynthetic candidates.
To confirm potential candidate genes for pheromone biosynthesis, GO enrichment analysis was performed. Of the final 11,342 transcripts, 7,603 had at least one GO term. For the two pools of statistically significant differentially-expressed genes identified in the Venn diagram (possible *exo*-brevicomin- and frontalin-biosynthetic genes), a GO enrichment analysis using BiNGO within Cytoscape identified over-represented molecular functions and metabolic processes with an adjusted significance value of less than 0.05 (Figure 3.5). Enriched GO terms in the possible *exo*-brevicomin-biosynthetic genes included P450-associated terms such as iron ion binding, heme binding and increased monooxygenase activity. Within the possible frontalin-biosynthetic genes, a wider variety of enriched metabolic processes are represented: uroporphyrinogen-III synthase and tetapyrrole biosynthetic activity, monooxygenase activity, carbohydrate metabolic processes, isoprenoid biosynthetic processes, and two processes involved in P450 biosynthesis. Both gene groups showed enrichment in P450 related processes, however GO terms related to the mevalonate pathway were more highly enriched in the putative frontalin-synthesizing candidates.

**Pheromone Biosynthesis Candidate Genes**

**Putative *trans*-verbenol-biosynthetic genes**

*trans*-Verbenol biosynthesis in female MPB likely takes place via P450-mediated hydroxylation of host (−)-α-pinene [37]. Three out of four P450 genes, *CYP6DJ1* (YQE_04799), *CYP6DJ2* (YQE_04800) and *CYP349B2* (YQE_02158) showed relatively high mRNA levels in fed females (Figure 3.6A). One other P450, *CYP4BD4*
(YQE_07200), showed mRNA levels in both fed and unfed females higher than those in males (Figure 3.6A).

**Putative exo-brevicomin-biosynthetic genes**

*exo*-Brevicomin production from long chain fatty acid precursors in the fat body of unfed males [3] involves steps catalyzed by a desaturase, P450s, a short chain dehydrogenase, and a cyclase (Figure 3.1C). The mRNA levels of two genes previously identified as active in the biosynthetic pathway, *CYP6CR1* and *ZnoDH* [2], were elevated in unfed males (Figure 6B). Other P450 genes with high expression levels in unfed males included *CYP6DE3* (YQE_02812), *CYP6BW4* (YQE_01441), *CYP6DF1* (YQE_11788), *CYP4EX1* (YQE_01611), *CYP4CV2* (YQE_05823), a short chain dehydrogenase (YQE_04359) and one putative terpenoid cyclase (YQE_04789) (Figure 3.6B). The expression profile of YQE_03851, which encodes a P450 with 98% a.a. identity to CYP4G56, similarly had increased expression levels in unfed males.

The purple gene module in Figure 3.3A includes seven genes hypothesized or confirmed to be involved in *exo*-brevicomin biosynthesis (*CYP6CR1*, *ZnoDH*, *CYP6DE3*, *CYP4EX1*, a cyclase, an oxidoreductase and a hydrolase) and 22 common neighbors. This module is densely connected, with density equal to 0.987 and all genes having greater expression values in males than in females (Figure 3.3C, Supplemental File 3). All gene members are statistically-significantly differentially-expressed (log₂-fold change>1) at a significance level of 0.01 (see Supplemental File 2).
Putative frontalin-biosynthetic genes

Steps for frontalin biosynthesis downstream of geranylgeranyl diphosphate are likely catalyzed by a dioxygenase or P450, and at least one additional P450 and a cyclase (Figure 3.1B). Genes with higher relative expression in fed males were considered as candidates for frontalin biosynthesis. These included four P450s *CYP6DK1* (YQE_01078), *CYP6DE4* (YQE_01868), *CYP345F1* (YQE_06277) and *CYP6BW3* (YQE_02884) (Figure 3.6C). No putative dioxygenase-encoding genes with elevated transcript levels in fed males compared to other treatments were identified. One putative short chain dehydrogenase, *YQE_11963*, and one putative terpenoid cyclase, *YQE_04789*, also had increased mRNA levels in fed males relative to the other groups (Figures 3.6B and 3.6C).

The light blue gene module in Figure 3B is based on seven genes previously hypothesized to play a role in frontalin biosynthesis: *CYP6DE4, CYP6BW3*, a kinase, two oxidoreductases, HMG-CoA reductase (*HMGR*), and geranylgeranyl diphosphate synthase (*GGPPS*). These seven genes and their 31 common neighbors created the densely connected light blue module (density = 0.983). The majority (26 of 38 genes) were statistically significantly over-expressed in fed males compared to unfed males and fed females (Figure 3.3C).

Another interesting grouping is presented in the orange module. This module also includes seven putative candidate genes for frontalin biosynthesis: *CYP345F1, CYP6DK1, CYP6DH2*, a *CYP6DK1*-like P450, a prenyltransferase, an acyl-CoA oxidase, and a dehydrogenase (Figure 3.3B). These seven candidate genes and their 77 common immediate neighbors are almost perfectly intra-connected (density = 0.994). Similar to
the light blue module, the majority of genes (45) have a statistically significant higher expression (log<sub>2</sub>-fold change>1) in fed male than unfed male and fed female (Figure 3.3C).

**Functional Analysis of CYP6DE3**

CYP6DE3 mRNA levels were elevated in all samples of unfed male and female beetles exposed to a variety of monoterpenes for 24 hours compared to the controls. In general, monoterpene exposure elevated CYP6DE3 mRNA levels in females more strongly than in males, with the exception that (+)-α-pinene exposure elevated mRNA levels more strongly in males (Figure 3.7).

Recombinant CYP6DE3 was assayed for activity with hypothesized intermediates in the *exo*-brevicomin biosynthetic pathway, (Z)-dec-7-enal and (Z)-3-nonene (Figure 3.1), as well as a variety of monoterpenes. Assays with (Z)-dec-7-enal or (Z)-3-nonene showed no unique products when analyzed by GC-MS compared to the negative controls, and neither were (Z)-3-nonene nor 6(Z)-nonen-2-ol, the hypothesized products, respectively, detected. The monoterpene substrates, however, showed unique or substantially increased GC-MS retention peaks compared to controls. Products resulting from (+)-α-pinene were tentatively identified as 3-oxatricyclo [4.1.1.0(2, 4)] octane (10.88 min) and *trans*-verbenol (12.15 min). Two other products (at 15.42 and 15.65 min) remain unidentified (Figure 3.8A). Similarly, incubations with 3-carene as a substrate yielded two unknown products at 14.78 and 15.92 min (Figure 3.8B), and incubations
with (+)-limonene as a substrate also produced two unknown products (16.00 and 16.12 min; Figure 3.8C).

**DISCUSSION**

Comparative transcriptome analysis to tentatively identify MPB genes encoding enzymes active in pheromone-biosynthetic pathways is based on the hypothesis that the genes for each pathway are coordinately regulated. It has proven useful for prior studies [15,18]. We extended this approach here by comparing expression profiles in pheromone-biosynthetic tissues of fed and unfed female and male mountain pine beetles. Extensive RNA-Seq profiling yielded nearly 425 million paired-end reads, with 74.9% aligning to the reference genome. The clear separation of gender and feeding status shown by PCA underscores the remarkable shifts in genome usage exhibited by these beetles [7,8]. The close correlation between RNA-Seq and qRT-PCR data (Table 2) supports that the expression values reported here reliably indicate in vivo mRNA levels.

We used a combination of bioinformatics analyses to narrow the pool of candidate pheromone-biosynthetic genes, beginning with four straight-forward comparisons of relative expression levels between physiological conditions. For example, *exo*-brevicomin-biosynthetic genes would be expected to have elevated expression levels in unfed males compared to both unfed females and fed males, whereas frontalin-biosynthetic genes would be higher in fed males compared to both unfed males and fed females. This produced preliminary pools of only ~200 – 1,200 candidates, depending on the analysis (Table 3; Figure 3.3). These pools were enriched for P450 production and activity, consistent with increased metabolic activity upon feeding [7,8,18], which
complicates identifying enzymes involved primarily in pheromone biosynthesis. Nevertheless, mevalonate pathway enzymes are more predominantly represented in fed males compared to females, consistent with frontalin production (Figure 3.5). In parallel, co-expression network analysis by petal using a Spearman Correlation Coefficient and similarity threshold of 0.808 also isolated candidate genes. As these expression data were not normally distributed, the Spearman Correlation Coefficient supplied a robust non-parametric alternative to the standard Pearson Correlation Coefficient. Twenty-two of the final genes selected by the petal analysis appear relevant to exo-brevicomin biosynthesis, while another 71 may be involved in frontalin production.

(–)-trans-Verbenol is produced by a single P450-mediated hydroxylation of (–)-α-pinene, a reaction that may be catalyzed by multiple enzymes as part of a detoxification process [37]. Thus, a “pheromone-biosynthetic” P450 that specifically produces trans-verbenol in females may be an artificial designation. Our current study notes three P450s with relatively high expression levels in fed females compared to unfed females and males (Figure 3.6A). A fourth P450, CYP4BD4, showed the highest mRNA levels in females compared to males, though in a pattern that is not consistent with feeding-induced trans-verbenol production.

exo-Brevicomin production from long chain fatty acid precursors in the fat body of unfed males [3] involves steps catalyzed by P450s, a short chain dehydrogenase, and a cyclase (Figure 3.1C). High probability candidate genes for exo-brevicomin biosynthesis are likely in the same petal group containing CYP6CR1 and ZnoDH (Figure 3.3A). Interestingly, this gene group includes a putative cyclase (YQE_04789) that may catalyze the terminal reaction. The two P450s (CYP6DE3 and CYP4EXI) may be active upstream
of ZnoDH to produce and/or hydroxylate 3-nonene. In this respect, the CYP4G56-like P450 (YQE_03851) was not part of the gene group but is of interest given its similar expression profile (Figure 3.6B) and identity as a CYP4G. While predicting P450 function from sequences is very difficult, CYP4G family P450s appear to be insect-specific and function as oxidative decarbonylases – yielding hydrocarbons from long chain fatty aldehydes [38,39]. Thus, YQE_03851 may contribute to 3-nonene production.

Frontalin-biosynthetic steps through the mevalonate pathway to geranylgeranyl diphosphate are well established in fed and JH treated MPB males [40]. Our analysis also identified mRNAs for mevalonate pathway enzymes, including HMGR and GGPPS, to be elevated in fed males compared to other treatment groups. Later steps are likely catalyzed by P450s, a dioxygenase, and a cyclase that should group together with HMGR and GGPPS in the petal analysis. Two P450 genes, CYP6DE4 (YQE_01868) and CYP6BW3 (YQE_02884), did group with HMGR and GGPS (light blue VN in Figure 3.3B) while four other P450 genes, CYP345F1 (YQE_06277), CYP6DK1 (YQE_01078), CYP6DH2 (YQE_01329) and a CYP6DK1-like P450 (YQE_01079), grouped into one different VN (orange VN in Figure 3.3B). The two VNs are connected directly by two links, and both gene groups portray increased expression in fed males (Figure 3.3C), a pattern consistent with frontalin biosynthesis. Interestingly, a putative dioxygenase was not identified, which may suggest alternative activities on a GGPP precursor, perhaps catalyzed by a cytochrome P450. It is also noteworthy that the cyclase identified in the “exo-brevicomin cluster” (YQE_04789) also shows elevated mRNA levels in fed males (Figure 3.6B). Given the structural similarities of the epoxide precursors for both exo-
brevicomin and frontalin, it is possible that a single cyclase could serve the terminal steps in both pathways.

While comparative transcriptomics is invaluable to preliminarily identify putative pheromone-biosynthetic genes, a more accurate assessment requires additional information. [10]. For MPB, our transcriptomic analyses return more candidate genes than there are reactions to catalyze. We hypothesized that those with elevated expression upon exposure to monoterpenes are more likely to contribute to resin detoxification than pheromone component production (except for the case of trans-Verbenol, as noted above). We therefore measured relative mRNA levels for CYP6DE3, which we had tagged as a potential exo-brevicomin biosynthetic enzyme, in beetles that had been exposed to atmospheres saturated with various monoterpenes. The clear elevation observed for all cases (Figure 3.7) suggests that CYP6DE3 is induced by monoterpane exposure, particularly in females, implying a resin-detoxifying role. The absence of this induction in fed insects further implicates that CYP6DE3 regulation is complex. The monoterpane-dependent difference in response in males and females is curious, but has been exhibited in another study reporting similar sex-specific transcriptional responses of various D. armandi P450 genes in response to monoterpenes [41]. A detoxification role for CYP6DE3 is supported by functional assays of the recombinant enzyme which showed that it oxidized a variety of monoterpenes, but did not appear to accept exo-brevicomin precursors as substrates (Figure 3.8 and data not shown). Interestingly, the products at 15.65 minutes for (+)-α-pinene and 14.78 minutes for 3-carene have a m/z peak at 168 suggesting these substrates were oxidized twice (Figures 3.8A and 3.8B).
De novo pheromone component biosynthesis in pine bark beetles is affected by sex, feeding status, environment, and JH III [10], with JH III treatment sometimes being sufficient to elevate mRNA levels of pheromone-biosynthetic genes even in insects that otherwise require feeding to trigger pheromone production [17]. Indeed, JH III stimulates both frontalin [5,11] and trans-verbenol biosynthesis, but not exo-brevicomin biosynthesis [11] in MPB. Our study complements those of Robert et al. (2013) [8], who compared fed and JH III-treated whole insects and concentrated on a survey of detoxification mechanisms, and Keeling et al. (2016) [11], who compared starved and JH III-treated midguts and fat bodies. Our study differs in that we focused on midgut and fat body tissues of fed and unfed insects rather than JH III-treated insects because of the evident complexity in regulating production of these three main pheromone components. Several putative pheromone-biosynthetic genes identified in our study agree with those reported by Keeling et al. (2016) [11] (Table 4), and the increased confidence resulting from this concurrence makes the common genes high priorities for functional assays. It is also noteworthy that CYP6DE4 does not accept pheromone precursors despite being induced by JH III [11]. The discrepancies in the list of candidate enzymes are likely due to a combination of factors, including differences in experimental design and data analysis. Given that the populations used by Keeling et al. (2016) [11] and us appear to be geographically and genetically isolated [42], it is also possible that their responses to different conditions also differ ([2], unpublished data).
Table 4. Candidate genes for MPB pheromone biosynthesis identified by RNA-seq using feeding status or JH treatment.

<table>
<thead>
<tr>
<th>Pheromone biosynthetic pathway</th>
<th>Feeding Status and Sex</th>
<th>JH Treatment and Sex (Keeling et al. 2016)</th>
<th>Proposed in Both Studies</th>
</tr>
</thead>
<tbody>
<tr>
<td>trans-Verbenol</td>
<td>CYP349B2</td>
<td>CYP6DJ1</td>
<td>CYP6DJ1</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CYP6DJ2</td>
</tr>
<tr>
<td>exo-Brevicomin</td>
<td>CYP6DE3</td>
<td>CYP6DE4</td>
<td>CYP4EX1</td>
</tr>
<tr>
<td></td>
<td>CYP6BW4</td>
<td>CYP18A1</td>
<td>CYP6DF1</td>
</tr>
<tr>
<td></td>
<td>CYP4CV2</td>
<td>CYP4BQ1</td>
<td>CYP6CR1</td>
</tr>
<tr>
<td></td>
<td>P450</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>YQE_03851</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cyclase</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>YQE_04789</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Frontalin</td>
<td>CYP6DE4</td>
<td>CYP6BW1</td>
<td>HMGR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>YQE_02503</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>GGPPS</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>YQE_09494</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>CYP6DK1</td>
</tr>
<tr>
<td></td>
<td>CYP6BW3</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>CYP345F1</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Cyclase</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>YQE_04789</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

CONCLUSION

This study identified a number of candidate genes for involvement in MPB aggregation and anti-aggregation pheromone biosynthesis through differential gene expression analysis based on sex and feeding status. However, as Keeling et al. (2016) [11] noted, caution should be employed when using comparative transcriptomic data to identify putative pheromone-biosynthetic genes. As evidenced by the functional analysis of an identified candidate gene, CYP6DE3, for exo-brevicomin biosynthesis in MPB, expression profiles are not always indicative of a specific role for a gene of interest. Further functional analysis of the genes identified in this study should lead to the
discovery of most, if not all, of the unknown enzymes involved in MPB aggregation and anti-aggregation pheromone biosynthesis.

**Declarations:**

**List of Abbreviations:**
a.a., amino acid; bp, base pair; cDNA, complementary deoxyribonucleic acid; EDTA, ethylenediaminetetraacetic acid; FF, female fed; FU, female unfed; GC-MS, gas chromatography-mass spectrometry; GGF, Georgia Genomics Facility; GGPPS, geranygeranyl diphosphate synthase; GO, gene ontology; GTF, gene transfer format; HF-CPR, house fly cytochrome P450 reductase; HMGR, 3-hydroxy-3-methylglutaryl-CoA reductase; JH, juvenile hormone; MF, male fed; MU, male unfed; MPB, mountain pine beetle; mRNA, messenger ribonucleic acid; NADPH, nicotinamide adenine dinucleotide phosphate; nt, nucleotide; ORF, open reading frame; P450, cytochrome P450 monooxygenase; PCA, principal component analysis; PCR, polymerase chain reaction; PMSF, phenylmethylsulfonyl fluoride; qRT-PCR, quantitative reverse transcriptase PCR; RNA-Seq, RNA sequencing; SAM, sequence alignment/map; VN, vicinity network

**Ethics approval and consent to participate:**
Mountain Pine Beetles were collected in the Tahoe Donner housing subdivision in Truckee, CA with permission from the Tahoe Donner Forestry Department. No other permissions were necessary.
Consent to Publish:

not applicable

Competing Interests

The authors declare no competing interests.

Funding:

The project described was supported by awards from USDA-NIFA (2014-67013-21748), The State of Nevada (NV00381) and National Institute of General Medical Sciences (P20GM103440) from the National Institutes of Health.

Authors’ Contributions:

CT, JN, KS and GJB conceived and designed the project. JN, KJ, MF, MM, and SY participated in data acquisition. JN, JP, RJT, and KS analyzed the data. JN, KS, GJB, JP, and CT prepared figures and wrote and edited the manuscript. All authors reviewed and approved the final manuscript.

Availability of Data and Materials

The dataset supporting the conclusions of this article is available in the NCBI Sequence Read Archive, https://www.ncbi.nlm.nih.gov/sra[ncbi.nlm.nih.gov], at SRA Project Accession SRP072778, with additional project and sample descriptions accessible at
NCBI BioProject PRJNA317010, and BioSamples SAMN04595104-SAMN04595119, respectively.

Acknowledgements
We thank the Georgia Genomics Facility for the Illumina sequencing services and D. Kosma for access to the GC-MS instrument.

References
5. Barkawi LS, Francke W, Blomquist GJ, Seybold SJ. Frontalin: De novo biosynthesis of an aggregation pheromone component by Dendroctonus spp. bark beetles (Coleoptera:


FIGURES

Figure 1. Pheromone-biosynthetic pathways.

MPB pheromone biosynthetic pathways. (A) (−)-trans-Verbenol is produced by a single P450-mediated hydroxylation of (−)-α-pinene. (B) Frontalin biosynthesis occurs through the mevalonate pathway to geranylgeranyl diphosphate followed by multiple steps likely catalyzed by P450s, a dioxygenase, and a cyclase. (C) exo-Brevicomin production from long chain fatty acid precursors in the fat body of unfed males involves steps catalyzed by P450s, a short chain dehydrogenase, and a cyclase. Chemical structures were generated using PubChem Sketcher V2.4, https://pubchem.ncbi.nlm.nih.gov/edit2/index.html.
Figure 2: Principal Component Analysis.

Principal component analysis of the filtered and normalized RNA-Seq expression levels of 11,342 transcripts shows a clear separation between feeding states. Almost 70% of the total variation is attributed to difference in feeding state.

Figure 3: petal co-expression network analysis. Co-expression network model of a subnetwork of 23 genes hypothesized or confirmed to be involved in pheromone biosynthesis. Within the subnetwork, very densely connected groups, vicinity networks (VNs), are identified for (A) exo-brevicomin, highlighted purple, and (B) frontalin, highlighted light blue and orange (showing only the genes of interest.) The gene expression profiles of the intersections of VNs associated to the three colored gene groups are shown in (C) with a total of 31 genes in purple, 38 genes in blue and 84 genes in orange (see Supplemental File 3).

Figure 4: Venn diagram of over-expressed genes.

A Venn diagram showing the intersections of statistically-significant differentially over-expressed genes (p<0.01). The first treatment in each grouping has greater expression than the second treatment (i.e. FF>FU in the FFvsFU grouping). Two areas of interest included 1,422 genes: 1,225 genes as possible exo-brevicomin-biosynthetic genes (blue outline) and 217 genes as possible frontalin-biosynthetic genes (red outline).

Figure 5: BiNGO analysis for candidate exo-brevicomin and frontalin biosynthetic genes. A BiNGO analysis showing gene ontology enrichment for (A) differentially over-
expressed genes for MU>FU (possible \textit{exo}-brevicomin-biosynthetic genes), p<0.01, and (B) for both MF>MU and MF>FF (possible frontalin-biosynthetic genes), p<0.01. Both comparisons show enrichment in P450 biosynthesis and oxidoreductase activity.

**Figure 6: Candidate \textit{trans}-verbenol, \textit{exo}-brevicomin and frontalin biosynthetic genes.** Expression profiles in regard to feeding status and sex for (A) four candidate genes hypothesized to be involved in \textit{trans}-verbenol, (B) nine candidate genes hypothesized to be involved in \textit{exo}-brevicomin and (C) seven candidate genes hypothesized to be involved in frontalin biosynthesis. Values represent the mean ± the standard deviation, n=4.

**Figure 7: CYP6DE3 expression in response to monoterpenes exposures.** qRT-PCR analysis to measure relative CYP6DE3 mRNA levels in groups of adults (3 individuals per sample) exposed to atmosphere saturated with the indicated monoterpenes for 24 h. Normalized to RP S3, n=3. Values represent the range of CYP6DE3 expression levels relative to the no exposure control. One asterisk indicates a statistically significant difference between the means of female ΔΔCt values compared to the means of male ΔΔCt values at p<0.05.

**Figure 8: CYP6DE3 enzyme assays.** GC chromatograms of pentane:ether (1:1) extracts from recombinant CYP6DE3 incubated for one hour in the presence of 6 mM (A) (+)-\textit{α}-pinene, (B) 3-carene or (C) (+)-limonene, 3 mM NADPH and recombinant house fly CPR or with CPR alone (negative control). Below are mass spectra for selected retention
peaks showing (A) production of 3-Oxatricyclo[4.1.1.0(2,4)]octane at 10.88 minutes, trans-verbenol at 12.15 minutes and two unknown products at 15.42 and 15.65 minutes for (+)-α-pinene, (B) two unknown products at 14.78 and 15.92 minutes for 3-carene and (C) to unknown products 16.00 and 16.12 for (+)-limonene. The 15.65 minute product for (+)-α-pinene and the 14.78 minute product for 3-carene have m/z peaks at 168 indicating a possible double oxidation of (+)-α-pinene and 3-carene.
SUPPLEMENTAL FILES

Supplemental File 1: Microsoft Excel file, additionalfile1.xlsx
Primer Table: A list of primers used for qRT-PCR and CYP6DE3 cloning.

Supplemental File 2: Microsoft PowerPoint file, additionalfile2.pptx
petal Venn Diagrams: Venn diagrams showing the intersections of differentially over-expressed genes with $a<0.01$ for the (A) purple, (B) light blue and (C) orange gene groups identified in the petal analysis. “gr” indicates greater than in each of the comparisons.

Supplemental File 3: Microsoft Excel file, additionalfile3.xlsx
petal Module Gene Annotations: A list of genes and their annotations found in each petal module.

Supplemental File 4: Microsoft Excel file, additionalfile4.xlsx
Differentially Expressed Genes for Each Comparison: A list of all the genes that have greater than two fold differential expression for each considered comparison of sex and feeding status.
Figure 3.1.
Figure 3.2.
Figure 3.3.
Figure 3.4.
Figure 3.5.
Figure 3.6.
Figure 3.7.

CYP6DE3 Expression Levels
Figure 3.8
Chapter 4

Expression Profiling and Functional Analysis of CYP6DK1, a Frontalin Biosynthetic Gene in the Mountain Pine Beetle

I. Abstract

II. Introduction

III. Methods

IV. Results

V. Discussion

VI. Conclusions

VII. References

VIII. Figure legends

I. Abstract

(+)-Frontalin [(1R,5R)-1,5-dimethyl-6,8-dioxabicyclo [3.2.1]octane] is a chemical produced in various bark beetles (Dendroctonus spp.) and functions as an aggregation or anti-aggregation pheromone. The Mountain Pine Beetle (MPB; Dendroctonus ponderosae Hopkins), a highly destructive pest in North American forests, uses frontalin as one of three major pheromone components when colonizing a host tree. In MPB, frontalin is synthesized in the male midgut via the mevalonate pathway and released in the frass. The final steps are hypothesized to consist of the epoxidation of the C6-C7 double bond in 6-methylhept-6-en-2-one (6MHO) by a cytochrome P450
monooxygenase (P450) to form 6,7-epoxy-6-methylheptan-2-one followed by cyclization, possibly enzyme mediated, to (+)-frontalin. A candidate P450, CYP6DK1, previously identified for possible involvement in frontalin biosynthesis was examined using expression profiling and functional analyses. Tissue distribution profiling and mRNA levels following exposure to monoterpenes were measured using qRT-PCR and showed that CYP6DK1 mRNA levels were highest in fed male midguts and increased in both males and females in response to monoterpene exposure, a pattern suggestive of both pheromone production and resin detoxification activities. Functional analysis showed CYP6DK1 catalyzes the production of frontalin from 6-methylhept-6-en-2-one (6MHO) without the need of a cyclase. Furthermore, an alternative product is produced from 6MHO when CYP6DK1 is recombinantly fused to housefly cytochrome P450 reductase (CPR).

II. Introduction

(+)‐Frontalin [(1R,5R)-1,5-dimethyl-6,8-dioxabicyclo[3.2.1]octane] is a bicyclic acetal produced by various bark beetles (Dendroctonus spp.) (Barkawi et al., 2003) and Asian elephants (Elephas maximus) (Rasmussen and Greenwood, 2003). Asian elephants use frontalin as a pheromone during mating to attract ovulating females and as an anti-attractant to young male adults (Rasmussen and Greenwood, 2003). Bark beetles use frontalin as an aggregation or anti-aggregation pheromone to help coordinate mass attacks to overcome the defenses of new host trees (Barkawi et al., 2003). MPB, a highly destructive pest in the forests on North America, uses frontalin as one of three major
pheromone components when colonizing a host tree (Pureswaran et al., 2000). (-)-trans-Verbenol [(1S,2R,5S)-4,6,6-trimethylbicyclo[3.1.1]hept-3-en-2-ol] and exo-brevicomin [(1R,5S,7R)-7-ethyl-5-methyl-6,8-dioxabicyclo[3.2.1]octane], the other two major pheromone components, are produced by feeding female and emerged male MPBs, respectively, and act in synergy as aggregation pheromones (Pitman et al., 1968; Pureswaran et al., 2000; Song, Gorzalski, et al., 2014). Frontalin is produced by feeding adult MPBs and, in high enough concentrations, acts as an anti-aggregation signal to prevent overcrowding in a new host tree (Borden et al., 1987).

In the MPB, frontalin is synthesized in the male midgut via the mevalonate pathway and released in the frass (Barkawi et al., 2003; Blomquist et al., 2010) (Figure 3.1). Frontalin biosynthesis in the MPB involves the C5 precursors isopentenyl diphosphate (IPP) and dimethylallyl diphosphate (DMAPP), which are condensed stepwise by isoprenyl diphosphate synthases (IDSs), including one farnesyl diphosphate synthase (FPPS) that has geranyl diphosphate synthase (GPPS) activity, to form the C20 frontalin precursor geranylgeranyl diphosphate (GGPP) (Keeling et al., 2013). GGPP is then thought to be cleaved by a dioxygenase or P450 to form 6-methylhept-5-en-2-one, which is subsequently isomerized to the eight carbon frontalin precursor 6MHO (Blomquist et al., 2010). The final steps are hypothesized to include the epoxidation of the C6,7 double bond in 6MHO by a P450 to form 6,7-epoxy-6-methylheptan-2-one, followed by cyclization, possibly enzyme mediated, to (+)-frontalin (Blomquist et al., 2010) (Figure 3.1).

Based on a previous comparative transcriptomic study using sex and feeding status as treatments, we identified candidate P450s involved in frontalin biosynthesis
(Chapter 3) (Nadeau et al., 2017). Expression profiling is a useful tool to identify pheromone-biosynthetic genes (e.g. Keeling et al., 2006; Aw et al., 2010), but functional assays are often necessary to confidently assess the role of a candidate P450, especially when genetic manipulations are impractical. To further investigate the potential role for CYP6DK1 in frontalin biosynthesis, we characterized its expression pattern and enzyme activity. As a putative frontalin-biosynthetic enzyme, we expected CYP6DK1 to be highly expressed in the midguts of fed males, where frontalin is synthesized (Barkawi et al., 2003) and to potentially epoxidize 6MHO (Figure 3.1). Here we show that CYP6DK1 is highly expressed in fed male posterior midguts and encodes an enzyme that catalyzes the production of frontalin from 6MHO without the need of a cyclase. Furthermore, we show that an alternative product is produced from 6MHO when CYP6DK1 is recombinantly fused to CPR.

III. Methods

III.1 Insects

Sections of mountain pine beetle-attacked lodgepole pine, Pinus contorta, were collected from the Whittell Forest, located in the Carson Range on the east slope of the Sierra Nevada (approx. N 39°16’29” W 119°52’43”). The beetles overwintered in the bolts and emerging adult beetles were collected and sexed as reported previously (Aw et al., 2010; Song et al., 2014). For feeding experiments, fresh lodgepole pine bolts were obtained from the Whittell Forest and stored at 4°C prior to use. Females were fed by drilling small holes just beneath the bark, inserting the beetles head first, stapling a wire mesh
over the occupied hole, incubating the bolt vertically for 24 h in the dark at room temperature and collecting the live beetles by stripping the bark. Fresh frass indicated that the beetles had fed. Males were fed using the same method except females were first placed head first into the holes for 24 h, followed by insertion of the males head first for 24 h and subsequent collection and sexing of the live beetles. Unfed beetles were incubated in 2 oz. plastic cups with perforated lids in a dark drawer kept humid with small flasks filled with water and a paper towel for 24 h.

III.2 Tissue Expression Profiling

Templates for tissue distribution and monoterpene exposure mRNA profiling were prepared previously (Chapter 2). qRT-PCR reactions and analyses were performed as described previously (Chapter 2), with the exception that CYP6DK1 (target gene) primers were used.

III.3 qRT-PCR of Monoterpene Exposed Beetles

Beetles were separated by sex and placed in two oz. plastic cups with perforated lids and incubated in a humidified dark drawer for 24 h as described previously (Song et al. 2014) to ensure they were unfed at the beginning of the monoterpene exposures. Small clumps of glass wool were placed in four pyrolyzed 500 mL glass jars, two of which contained two mL vials with 500 µL of a selected monoterpene, and capped with a cotton mesh lid. Two jars with no vials were used as a control. Eight live males or females were transferred into each jar so that each sex had monoterpene-exposed and control treatments. The jars were incubated in the dark for 24 h, and the beetles (two or three
from each of three replicate incubations) were then placed in microcentrifuge tubes and flash frozen in liquid nitrogen. Six different monoterpenes were tested: 3-carene, R-(+)-limonene, myrcene, (+)-α-pinene, (−)-α-pinene, terpinolene. We also tested a monoterpane cocktail containing all six listed monoterpenes. RNA was extracted from the whole beetles using the RNeasy Plant Mini Kit from Qiagen as described above with an on-column DNase treatment. qRT-PCR was conducted using CYP6DK1 primers and normalized to Tubulin as described above. The effects of gender and treatment (monoterpane exposure) were examined by two-way ANOVA using GraphPad Prism version 7.03 for Windows, GraphPad Software, La Jolla California USA, www.graphpad.com. Statistical significance was determined based on α=0.05.

**III.4 Recombinant CYP6DK1 Expression**

The CYP6DK1 open reading frame (ORF) (GenBank accession no. JQ855679.1) was amplified by PCR using YQE_01079F1 and YQE_01079R1 primers (Table 4.1) and CloneAmp HiFi PCR Premix (Takara Bio USA, Inc., Mountain View, CA) from first strand cDNA prepared from mature male MPB, cloned into pENTR4 modified to remove the _NcoI_ site (Sandstrom _et al._, 2006) by Gibson assembly using the In-Fusion® HD Cloning Kit (Takara Bio USA, Inc.) to create CYP6DK1_pENTR4, and transformed into _E. coli_ Stellar™ Competent Cells (Takara Bio USA, Inc.). For the CYP6DK1_CPR fusion protein, the full length CYP6DK1 ORF was modified to remove the stop codon through amplification with CYP6DK1F1 and CYP6DK1R1 primers (Table 4.1) and fused via a Ser-Ser dipeptide to the catalytic domain of housefly cytochrome P450 reductase (amino acids 51-671; GenBank accession no. AAA29295.1). A
CYP4G2_CPR_pENTR4 (NeoI-) (Sandstrom et al., 2006; Qiu et al., 2012) plasmid was divergently amplified to remove CYP4G2 using Q5® High-Fidelity DNA Polymerase (New England BioLabs Inc., Ipswich, MA), CPRfus_F2 forward and pENTR4_R5 reverse primers (Table 4.1) to create CPR_pENTR4. The full-length CYP6DK1 ORF, without the stop codon, was cloned into CPR_pENTR4 by Gibson assembly using the In-Fusion® HD Cloning Kit (Takara Bio USA, Inc.) to create CYP6DK1_CPR_pENTR4, and transformed into \textit{E. coli} Stellar™ Competent Cells (Takara Bio USA, Inc.). Recombinant CYP6DK1_pENTR4 and CYP6DK1_CPR_pENTR4 were confirmed by sequencing prior to recombination into BaculoDirectTM C-Term Linear DNA (Invitrogen) by LR ClonaseTM II (Invitrogen). The recombinant BaculoDirect clones were transferred into Sf9 cells by transfection using Cellfectin II (Invitrogen) and amplified by successive infections of P1 and P2 viral stocks to a high-titer P3 viral stock. High-titer P3 viral stocks for CYP9T3 and CPR were from an earlier study (Song et al., 2013). Protein expression was initiated by infecting 50 mL of $10^6$ cells/mL Sf9 cells in Sf-900 II SFM culture media supplemented with 10% (vol/vol) fetal bovine serum (Atlas Biologicals, Fort Collins, CO), 0.3 mM δ-aminolevulinic acid, and 0.1 mM ferric citrate with 50 µL of the P3 viral stock and incubating at 27°C for 72 h. Recombinant CYP6DK1_CPR, CYP6DK1, CYP9T3 and CPR were harvested 72 h post infection in a cell lysis buffer (100 mM potassium phosphate, pH 7.6, 20% (vol/vol) glycerol, 1.1 mM EDTA, 200 µM PMSF, and protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO) and the microsomes were isolated by differential centrifugation essentially as per (Nadeau et al., 2017). The microsomal fraction was tested for functional CYP6DK1_CPR
and CYP6DK1 using a CO-difference spectrum analysis (Choi et al., 2003) as described previously (Song et al., 2013).

**III.5 Enzyme Assays**

6-Methylhept-5-en-2-one, geranylgeranyl diphosphate, and NADPH were from Sigma-Aldrich (St. Louis, MO). 6-Methylhept-6-en-2-one (6MHO) was obtained from AKos Consulting & Solutions Deutschland GmbH (Steinen, Germany). Reaction mixtures consisted of 200 µL of the recombinant P450 microsomal preparation, 40 µL of HF-CPR microsomal preparation (only in CYP6DK1 reactions), 100 mM potassium phosphate buffer pH 7.6, 1.5 mM NADPH and 21 mM substrate (final concentrations) in a total volume of 602 µL. Control reactions containing only HF-CPR were identical to the experimental reactions except that reaction buffer was substituted for the CYP6DK1_CPR or CYP6DK1 microsomal preparation. Reactions were incubated in separate capped 5 mL glass vials and rotated lengthwise at 30 °C in a FisherBiotech Hybridization Incubator (Thermo Fisher Scientific, Waltham, MA) for three hours. The reactions were terminated and extracted using 1 mL pentane:ether (1:1) with a 10 µg/mL n-octanol internal standard. The extracts were analyzed by GC-MS on an HP-5ms capillary column (Agilent) using two different Agilent (Santa Clara, CA) 7890B gas chromatographs coupled to a 5977A mass spectrum detector. The instrument running parameters were: initial temperature of 40 °C with a one min hold, 5 °C/min to 240 °C and 15 °C/min to 300 °C with a 5 min hold. The MS detector was a single quadrapole with an electron ionization source and a molecular weight scanning range of 40 to 700 atomic mass units (amu) and an ionization potential of 70 eV. Mass spectra were
compared to the main EI MS library (NIST08) to confirm the identity of the products. Each functional assay was repeated twice to confirm products.

IV. Results

IV.1 Tissue Expression Profiles

Relative CYP6DK1 expression levels were measured in fed and unfed male and female MPB tissues using qRT-PCR. Three-way ANOVA analysis of CYP6DK1 tissue expression levels normalized to Tubulin revealed significant variation of CYP6DK1 expression between tissues (F_{5,72}=83.21, p<0.0001) (Table 4.2). Gender alone did not yield significant variation in expression levels, but tissue and treatment (fed or unfed) (F_{5,72}=9.73, p<0.0001), tissue and gender (F_{5,72}=5.49, p=0.0002), and treatment and gender (F_{1,72}=5.72, p=0.02) interactions were significant (Table 2). CYP6DK1 expression levels were highest in unfed male midguts, fed male and female midguts and fed female posterior midguts (Figure 4.1).

IV.2 Monoterpene Exposure

The effects of exposure to various monoterpines on relative CYP6DK1 expression levels in unfed male and female MPB were measured by qRT-PCR. Two-way ANOVA analysis of CYP6DK1 expression levels normalized to Tubulin revealed significant variation in response to monoterpene exposure (F_{7,32}=136.30, p<0.0001), gender (F_{1,32}=2391, p<0.0001) and the interaction between monoterpene exposure and gender (F_{7,32}=32.13,
p<0.0001) (Table 4.3). CYP6DK1 expression levels were overall higher in males compared to females and increased in response to exposure to all monoterpenes when compared to no exposure controls, except that no significant change in mRNA levels was observed in females following exposure to (+)-α-pinene (Figure 4.2).

IV.3 Functional Analysis of CYP6DK1_CPR Fusion and CYP6DK1

The CO difference spectra of both recombinant CYP6DK1_CPR and recombinant CYP6DK1 showed the characteristic 450 nm peak of active P450s (Figure 4.3). Functional assays using putative frontalin precursors (geranylgeranyl diphosphate, 6-methylhept-5-en-2-one and 6MHO) as substrates were carried out with both enzymes. Incubations of recombinant CYP6DK1_CPR fusion enzyme with geranylgeranyl diphosphate and 6-methylhept-5-en-2-one yielded no unique products when compared to CPR only reactions (data not shown), however incubations of the fusion enzyme with 6MHO yielded two unique products at 6.28 and 9.24 minutes when compared to CPR only reactions (Figures 4.4A & B). Another peak at 7.62 minutes (Figure 4.4A) was observed in both the CYP6DK1_CPR and CPR only (control) incubations, but not in the 6MHO standard (not shown). The mass spectra of the 6.28 and 9.24 minute peaks were consistent with frontalin and 6-methylheptan-2-one, respectively (Figures 4.4C & E). The mass spectrum for the 7.62 minute peak was most similar to that of 6-methyl-5-hepten-2-ol (Figure 4.4D). Incubations of recombinant CYP6DK1 (not fused to CPR) in the presence of CPR and 6MHO yielded only one unique product (at 5.92 minutes) when compared to incubations of recombinant CYP9T3 and CPR only (Figures 4.5A-C). The
mass spectrum of the 5.92 minute product for CYP6DK1 and MHO was identical to that of frontalin (Figure 4.5D). Incubations of either CYP6DK1 or CYP9T3 with 6MHO both yielded a product at 7.23 minutes (Figures 4.5A & B). This peak was not present in the 6MHO standard and produced a mass spectrum consistent with 6-methyl-5-hepten-2-ol (Figure 4.5E).

V. Discussion

Two previous comparative transcriptomic studies of the MPB using gender and feeding-status (Nadeau et al., 2017) or gender and juvenile hormone III (JH) (Keeling et al., 2016) as differentiating factors identified CYP6DK1 as a candidate for frontalin biosynthesis. Nadeau et al. (2017) found that CYP6DK1 mRNA levels are elevated in feeding males, a pattern suggestive of involvement in frontalin biosynthesis, while Keeling et al. (2016) showed that treatment of males with JH III, a hormone known to stimulate pheromone biosynthesis (Blomquist et al., 2010), induces elevated CYP6DK1 mRNA levels. In this study, we further explored CYP6DK1 expression patterns and function based on the hypothesis that it is involved in frontalin biosynthesis.

Frontalin production via the mevalonate pathway takes place in the midgut of fed male MPB (Barkawi et al., 2003; Blomquist et al., 2010). We hypothesized that if CYP6DK1 is involved in frontalin biosynthesis, its mRNA levels would be highest in fed male midguts compared to other tissues. Here, we show that CYP6DK1 mRNA levels are indeed highest in the midgut of fed males (Figure 4.1). However, there was increased expression of CYP6DK1 in unfed male and fed female midguts as well when compared
to other tissues, indicating a possible additional function to frontalin production. Female MPB do not naturally produce frontalin, but they can produce frontalin upon exposure to 6MHO (Perez et al., 1996). Thus, the sex-specificity of frontalin production lies upstream of CYP6DK1, possibly at the step of 6MHO biosynthesis from GGPPS. Indeed, significant production of 6MHO only occurs in feeding male MPB (Keeling et al., 2016) so the fact that unfed males and fed females do not naturally produce frontalin is not surprising even though both express relatively similar levels of CYP6DK1.

An additional function for CYP6DK1 is further supported when looking at the expression patterns in response to monoterpene exposures. Monoterpenes are highly toxic compounds trees use in defense to herbivory, and bark beetles cope with monoterpenes in part by expressing a suite of detoxifying enzymes including P450s (Keeling and Bohlmann, 2006; Robert et al., 2013). P450s that are involved in resin detoxification should have increased expression when exposed to monoterpenes (Blomquist et al., 2010), so we tested the hypothesis that CYP6DK1 might be involved in detoxification by exposing unfed male and female MPB to monoterpenes and monitoring expression levels of CYP6DK1. We found that both unfed male and female MPB had significantly increased expression following exposure to all (male) or most (female) of the monoterpenes tested (Figure 4.2), suggesting that CYP6DK1 may function in resin detoxification. Such a role is consistent with elevated mRNA levels for the gene in alimentary canals of feeding females (Figure 4.1). Alternatively, the change in CYP6DK1 mRNA levels upon monoterpene exposure may be due to monoterpene-related alterations in endogenous metabolism. It is also possible that CYP6DK1 functions in both frontalin production and resin detoxification. This possible dual role has precedence in Ips pini,
where CYP9T3 produces ipsdienol from myrcene as a pheromone-biosynthetic reaction, but can also accept other monoterpenes as substrates (Song et al., 2013). These data support the hypothesis that pheromone biosynthetic P450s evolved from detoxification P450s (Seybold, Bohlmann and Raffa, 2000; Blomquist et al., 2010). Functional assays of CYP6DK1 activity on different monoterpenic substrates are beyond the scope of the current study, but would clarify our understanding of CYP6DK1 function.

Given that the expression patterns of \textit{CYP6DK1} from both this and previous studies support the possibility of involvement in frontalin biosynthesis, we conducted a functional analysis of recombinant CYP6DK1 using geranylgeranyl diphosphate, 6-methylhept-5-en-2-one and 6MHO as possible substrates because they are all compounds hypothesized to be a part of the frontalin biosynthetic pathway (Figure 3.1). In our first attempt at functional analysis, we used CYP6DK1 fused to the catalytically active domain of CPR to conduct reactions. Previous attempts to express recombinant P450s, especially in the CYP4G family, have required a fusion of the P450 to CPR for functionality (Qiu et al. 2012; Balabanidou et al. 2016; S. Young and M. MacLean personal communication), and P450-CPR fusions can assist functional assays of different P450s (e.g. Dodhia et al., 2006; Leonard and Koffas, 2007; Lundemo et al., 2016). Although CYP6DK1 is not in the CYP4G family, we assumed a CYP6DK1_CPR fusion would yield a functional P450 that we could use in assays.Expressing the CYP6DK1_CPR fusion did in fact yield a functional P450 (Figure 4.3A) that generated products in reactions with 6MHO but not geranylgeranyl diphosphate or 6-methylhept-5-en-2-one (Figure 4.4). The differing retention times for the same products (e.g. 6.28 and 5.92 minutes for frontalin, Figures 4.4 and 4.5) is likely due to the GC analyses being
done on different dates and on different machines. From the previously hypothesized pathway, we expected formation of 6,7-epoxy-6-methylheptan-2-one and the need of a cyclase to convert the epoxide to frontalin. Interestingly, the products observed from the CYP6DK1_CPR and 6MHO reactions were frontalin and 6-methylheptan-2-one, confirming a role for CYP6DK1 in frontalin production and indicating that a cyclase is not necessary. If the epoxide product of the reaction is relatively unstable, it is possible that cyclization may occur while it is still bound to CYP6DK1, assuming that the enzyme holds the epoxide in a conformation that facilitates the cyclization reaction. Indeed, other P450s have cyclase activity (Guengerich, 2001) so CYP6DK1 is not unique in that aspect.

The other product, 6-methylheptan-2-one, was produced in much greater quantities than frontalin and appears to be a result of the reduction of the 6,7 double bond in 6MHO. This suggests CYP6DK1_CPR catalyzes a reduction in contrast to the normal P450 oxidase activity. We hypothesize a possible explanation for this unusual reaction is that CYP6DK1, being fused to CPR, may occasionally uncouple the normal catalytic process such that electrons are passed directly onto the substrate instead of being passed to oxygen to form water. In fact, it has been hypothesized that even in the presence of oxygen, transfer of an electron to the substrate may be competitive with oxidation in some P450 systems (Guengerich, 2001). To test this hypothesis, we expressed recombinant CYP6DK1 that was not fused to CPR and conducted assays of this “unfused” enzyme with 6MHO in the presence of full-length, recombinant CPR. Fortunately, the un-fused enzyme was functional (Figure 4.3B), and it similarly produced frontalin while production of 6-methylheptan-2-one was not observed (Figure 4.5). This
suggests that production of 6-methylheptan-2-one is indeed a consequence of fusing CPR to CYP6DK1 and likely not a natural occurrence under physiological conditions. In addition, 6-methyl-5-hepten-2-ol was produced in all assays, including the negative controls. This chemical was not present in the 6MHO standard, suggesting its production by endogenous microsomal enzyme activity of the Sf9 cells.

VI. Conclusions

We found that CYP6DK1 has expression patterns not only suggestive of frontalin biosynthesis but also possible resin-detoxification activities. Functional analysis of CYP6DK1 showed it could catalyze the formation of frontalin from 6MHO, indicating an intrinsic cyclase activity. Finally, fusing CYP6DK1 with CPR yielded an unexpected byproduct resulting from a reduction of 6MHO, indicating P450 and CPR fusion proteins may cause uncoupling of the normal P450 reaction and suggesting caution should be employed when analyzing products from P450/CPR fusion proteins. Further functional analysis of CYP6DK1 using monoterpenes as substrates should be conducted to confirm any involvement in resin detoxification and RNAi knockdown of CYP6DK1 could be used to further confirm its involvement in frontalin biosynthesis.

VII. References


VIII. Figure Legends

Figure 4.1. Tissue distribution profile. MPB tissue distribution profile of fed or unfed males and females for CYP6DK1 mRNA by qRT-PCR. Data are normalized to Tubulin and presented as the mean with the error bars representing the standard deviation, n=4.
**Figure 4.2. Monoterpene exposures.** CYP6DK1 mRNA levels in male and female MPB in response to monoterpene exposures measured by qRT-PCR. Data are normalized to Tubulin and presented as the mean with the error bars representing the standard deviation, n=4.

**Figure 4.3. CO difference spectra.** CO difference spectra of (A) CYP6DK1_CPR fusion protein and (B) CYP6DK1 showing the characteristic absorbance peak at 450 nm for active P450s.

**Figure 4.4. Fusion protein functional assay.** GC chromatograms of extracts from microsomes bearing (A) CYP6DK1_CPR fusion protein or (B) CPR only, each incubated with 6MHO. The large peak at 9.94 minutes is a 10 ng/µL n-octanol standard and the smaller peak at 6.28 minutes is frontalin. (C) structure of frontalin and an MS scan of the 6.28 minute peak in (A) (red) compared to frontalin from the NIST08 library (blue). (D) and (E) structures of the products found at 7.62 and 9.24 minutes respectively and the corresponding mass spectra of the peaks in (A) compared to the NIST08 library MS for the identified products.

**Figure 4.5. Unfused protein functional assay.** GC chromatograms of extracts from microsomes containing (A) CYP6DK1, (B) CYP9T3 or (C) CPR only, each incubated with 6MHO. The large peak at 9.55 minutes is a 10 ng/µL n-octanol standard and the smaller peak at 5.92 minutes is frontalin. (D) structure of frontalin and a MS scan of the 5.92 minute peak in (A) (red) compared to frontalin from the NIST08 library(blue). (E)
structure of the product found at 7.23 minutes and the corresponding mass spectra of the 7.23 minute peaks in (A) compared to the NIST08 library MS for the identified product.
### Table 4.1. Cloning and qRT-PCR primers.

<table>
<thead>
<tr>
<th>Primer Name</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Cloning Primers</strong></td>
<td></td>
</tr>
<tr>
<td>YQE_01079F1</td>
<td>GCGGTCAGTAACCTGAGCCAGA</td>
</tr>
<tr>
<td>YQE_01079R1</td>
<td>GACCCATTATGTCTGGCTGGG</td>
</tr>
<tr>
<td>CYP6DK1(-Stop)F1</td>
<td>ATGTTAGCATTATATGC</td>
</tr>
<tr>
<td>CYP6DK1(-Stop)R1</td>
<td>AATATAATTATTTTCCAATGTCTAC</td>
</tr>
<tr>
<td>CPRfus_F2</td>
<td>TCAA CTACATCTATAACACCC</td>
</tr>
<tr>
<td>pENTR4_R5</td>
<td>CCGGATCCAGTGACTGAAT</td>
</tr>
<tr>
<td><strong>qRT-PCR Primers</strong></td>
<td></td>
</tr>
<tr>
<td>YQE_01079F</td>
<td>GGAACCTTTTCGCTGCTGATTCC</td>
</tr>
<tr>
<td>YQE_01079R</td>
<td>TTCTCCAAATGGGAAAAACG</td>
</tr>
<tr>
<td>DpoTubulinF1</td>
<td>GACAATCTGGAGCAGGAAACAAAT</td>
</tr>
<tr>
<td>DpoTubulinR1</td>
<td>TGCCCTTTTCGCAACACATC</td>
</tr>
</tbody>
</table>
Table 4.2. Three-way ANOVA of ΔCT(Tubulin-CYP6DK1) values for CYP6DK1 tissue distribution analysis.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F (DFn, DFd)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tissue</td>
<td>1178</td>
<td>5</td>
<td>235.7</td>
<td>F (5, 72) = 83.21</td>
<td>P&lt;0.0001</td>
</tr>
<tr>
<td>Treatment</td>
<td>0.1121</td>
<td>1</td>
<td>0.1121</td>
<td>F (1, 72) = 0.03957</td>
<td>P=0.8429</td>
</tr>
<tr>
<td>Gender</td>
<td>3.466</td>
<td>1</td>
<td>3.466</td>
<td>F (1, 72) = 1.224</td>
<td>P=0.2723</td>
</tr>
<tr>
<td>Tissue x Treatment</td>
<td>137.8</td>
<td>5</td>
<td>27.55</td>
<td>F (5, 72) = 9.729</td>
<td>P&lt;0.0001</td>
</tr>
<tr>
<td>Tissue x Gender</td>
<td>77.68</td>
<td>5</td>
<td>15.54</td>
<td>F (5, 72) = 5.486</td>
<td>P=0.0002</td>
</tr>
<tr>
<td>Treatment x Gender</td>
<td>16.2</td>
<td>1</td>
<td>16.2</td>
<td>F (1, 72) = 5.721</td>
<td>P=0.0194</td>
</tr>
<tr>
<td>Tissue x Treatment x Gender</td>
<td>13.37</td>
<td>5</td>
<td>2.675</td>
<td>F (5, 72) = 0.9444</td>
<td>P=0.4578</td>
</tr>
<tr>
<td>Residual</td>
<td>203.9</td>
<td>72</td>
<td>2.832</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 4.3. Two-way ANOVA of ΔCT(Tubulin-CYP6DK1) values for CYP6DK1 induction by monoterpene exposures.

<table>
<thead>
<tr>
<th>Source of Variation</th>
<th>SS</th>
<th>DF</th>
<th>MS</th>
<th>F (DFn, DFd)</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Interaction</td>
<td>16.54</td>
<td>7</td>
<td>2.362</td>
<td>F (7, 32) = 32.13</td>
<td>P &lt; 0.0001</td>
</tr>
<tr>
<td>Monoterpene Exposure</td>
<td>70.17</td>
<td>7</td>
<td>10.02</td>
<td>F (7, 32) = 136.3</td>
<td>P &lt; 0.0001</td>
</tr>
<tr>
<td>Gender</td>
<td>175.8</td>
<td>1</td>
<td>175.8</td>
<td>F (1, 32) = 2391</td>
<td>P &lt; 0.0001</td>
</tr>
<tr>
<td>Residual</td>
<td>2.353</td>
<td>32</td>
<td>0.07353</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Figure 4.1. Tissue distribution profile.
Figure 4.2. Monoterpenes exposures.
Figure 4.3. CO difference spectra.
Figure 4.4. Fused protein functional assay.
Figure 4.5. Unfused protein functional assay.
Chapter 5

Mountain Pine Beetle CYP4G55 and CYP4G56 Exhibit Functions Consistent with 

*exo*-Brevicomin Biosynthesis

I. Prologue

II. Abstract

III. Introduction

IV. Methods

V. Results

VI. Discussion

VII. References

VIII. Figure legends

I. Prologue

This portion of the dissertation contributes to a larger study by the Blomquist lab on the function of the CYP4G family of cytochrome P450 monooxygenases in the mountain pine beetle.

II. Abstract

*exo*-Brevicomin is a bicyclic acetal produced by mountain pine beetle (MPB; *Dendroctonus ponderosae* Hopkins), a highly destructive pest in the forests on North
America, and is used as one of three major pheromone components when colonizing a host tree. The early steps in the biosynthesis of exo-brevicomin in MPB are thought to proceed through β-oxidation and reduction of long chain fatty acid precursors to produce a ten carbon aldehyde (cis-7-decenal) with subsequent cytochrome P450 (P450) mediated decarbonylation of the aldehyde to a nine carbon alkene (cis-3-nonene). One of the two CYP4G P450s found in MPB, CYP4G55 and CYP4G56, are hypothesized to catalyze the decarbonylation reaction of cis-7-decenal to cis-3-nonene. We found that both enzymes catalyzed the conversion of cis-7-decenal to cis-3-nonene.

III. Introduction

*exo*-Brevicomin [(1R,5S,7R)-7-Ethyl-5-methyl-6,8-dioxabicyclo [3.2.1]octane] is a bicyclic acetal produced by MPB (Blomquist *et al.*, 2010). MPB is a highly destructive pest in the forests on North America and uses *exo*-brevicomin as one of three major pheromone components when colonizing a host tree (Pureswaran *et al.*, 2000). (–)-*trans*-Verbenol [(1S,2R,5S)-4,6,6-trimethylbicyclo[3.1.1]hept-3-en-2-ol] and frontalin [(1R,5R)-1,5-dimethyl- 6,8-dioxabicyclo[3.2.1]octane], the other two major pheromone components, are produced by feeding female and feeding male MPB, respectively, and act as aggregation and anti-aggregation pheromones, respectively (Pitman *et al.*, 1968; Borden *et al.*, 1987). In synergy with the female produced (–)-*trans*-verbenol, newly emerged male MPB produce the *exo*-brevicomin in small concentrations as an aggregation pheromone component, with production decreasing significantly upon reaching the new host tree and mating (Pureswaran *et al.*, 2000).
Given the importance of pheromone signaling to MPB reproduction, interfering with exo-brevicomin production may contribute to managing MPB populations. However, much of the biosynthetic pathway remains uncharacterized (Figure 3.1). Biosynthesis of exo-brevicomin in the MPB is thought to proceed through β-oxidation and reduction of long chain fatty acid precursors to produce cis-7-decenal, subsequent P450 mediated decarbonylation of the aldehyde to cis-3-nonene, followed by a cytochrome P450 mediated hydroxylation to 6(Z)-non-6-en-2-ol (Blomquist et al., 2010). 6(Z)-Non-6-en-2-ol is oxidized to 6(Z)-non-6-en-2-one by the enzyme 6(Z)-nonen-2-ol dehydrogenase (ZnoDH), and the ketone is then epoxidized to 6,7-epoxy-nonan-2-one by cytochrome P450 (Song et al. 2014). Finally, 6,7-epoxy-nonan-2-one is likely cyclized by a cyclase to produce exo-brevicomin although non-enzyme mediated cyclization is also possible (Figure 3.1).

The only known enzymes specific for exo-brevicomin synthesis are ZnoDH and CYP6CR1 (Song et al., 2014). The decarbonylation step that converts cis-7-decenal to cis-3-nonene early in the pathway was hypothesized by Blomquist et al. (2010) in part because the hypothesized alternative, an α-oxidation of a β-keto-precursor, is highly unusual (Skiba and Jackson, 1994). Our hypothesis predicts that a CYP4G enzyme catalyzes the reaction, as these are the only class of enzyme known to catalyze decarbonylation reactions to form hydrocarbons in insects (Qiu et al., 2012). By extension, we hypothesize that one of the two CYP4G P450s found in MPB, CYP4G55 and CYP4G56, catalyzes the decarbonylation reaction of cis-7-decenal to cis-3-nonene. Previous work has shown these two CYP4Gs catalyze the decarbonylation of long chain (i.e. C18) alcohols and aldehydes (M. MacLean personal communication), however
assays with shorter chain aldehydes like cis-7-decenal have not been conducted. Here we assay recombinantly expressed CYP4G55 and CYP4G56 fused to housefly cytochrome P450 reductase (CPR) for cis-7-decenal decarbonylation activity, using GC-MS to monitor production of cis-3-nonene.

IV. Methods

IV.1 Functional expression of CYP4Gs

The two CYP4Gs used in this study were fusion protein constructs consisting of a full-length P450 fused via a Ser-Ser dipeptide to the catalytic domain of housefly cytochrome P450 reductase (HF-CPR) (amino acids 51-671; GenBank accession no. AAA29295.1). The coding sequence was constructed by InFusion (Clontech, Mountain View, CA) cloning as outlined below. A CYP4G2_CPR_pENTR4 (NcoI-) (Sandstrom et al., 2006; Qiu et al., 2012) plasmid was divergently amplified to remove CYP4G2 using Q5® High-Fidelity DNA Polymerase (New England BioLabs Inc., Ipswich, MA), CPRfus_F2 forward (TCAACTACATCTATACAACCC) and pENTR4_R5 reverse (CCGGATCCAGTCGACTGAAT) primers. Each 50 µL reaction contained 100 ng template, 25 pmol forward and reverse primers, 10 nmol dNTPs, 10 µL 5X Q5 Reaction buffer, and 0.5 µL Q5® High-Fidelity DNA Polymerase. The cycling profile was 98°C for 60 seconds, 30 cycles of 98°C for 10 seconds, 50°C for 30 seconds, and 72°C for 135 seconds, with a final extension at 72°C for 2 minutes. The ORFs of CYP4G55 and CYP4G56 (Accession numbers: JQ855658.1 and JQ855659.1), codon optimized for Sf9 cells and with flanking primers sites pENTR4_R5ic (ATTCAGTCGACTGGATCCGG)
and CPRfusF2ic (GGGTTGTATAGATGTAGTTGA), were synthesized by Biomatik (Wilmington, DE). The pBSK(+)_CYP4G55_Sf9 bias and pBSK(+)_CYP4G56_Sf9 bias plasmids were reconstituted and transformed into NEB-5α competent cells (New England Biolabs) using standard transformation. Colonies containing pBSK(+)_CYP4G55_Sf9 bias and pBSK(+)_CYP4G56_Sf9 bias were positively identified by colony PCR and one of each was selected to amplify pBSK(+)_CYP4G55_Sf9 bias and pBSK(+)_CYP4G56_Sf9 bias using CloneAmp HiFi PCR premix (Clontech) to generate high fidelity coding regions. The 25 µL reactions contained 300 pmol pENTR4_R5ic forward and CPRFus_F2ic reverse primers, 90 ng pBSK(+)_CYP4G55_Sf9 bias or 90 ng pBSK(+)_CYP4G56_Sf9 bias, and 12.5 µL CloneAmp HiFi PCR premix. The cycling profile was 98°C for 10 sec, 30 cycles of 98°C for 10 seconds, 55°C for 10 seconds, and 72°C for 9 seconds. The pENTR4_CPR/CYP4G55_Sf9 bias, and pENTR4_CPR/CYP4G56_Sf9 bias fragments were verified by gel electrophoresis, copurified as pENTR4_CPR/CYP4G55_Sf9 bias or pENTR4_CPR/CYP4G56_Sf9 bias fragment pairs with a NucleoSpin Extract II column (Macherey-Nagel, Bethlehem, PA), ligated by In-Fusion HD® Ecody™ Mix (Clontech) in 10 µL reactions, transformed into Stellar competent Escherichia coli (Clontech), and selected on LB kanamycin plates. Recombinant clones were identified by PCR and sequenced to confirm the integrity of the construct.

IV.2 Recombinant Protein Production

Recombinant CYP4G55-CPR and CYP4G56-CPR were produced in Sf9 cells using the Baculo-Direct expression system (Invitrogen, Carlsbad, CA). Briefly, the insert was
transferred from pENTR4(\textit{NcoI}-) into the linearized BaculoDirect™ C-term vector by LR recombination. A high-titer P3 viral stock was prepared by successive amplifications of P1 and P2 stocks. Serum-adapted cells were infected for 72 h in 50-mL cultures (10^6 cells/mL) in Sf900 II serum-free media (Gibco, Grand Island, NY) supplemented with 10% (vol/vol) FBS (Atlas Biologicals, Fort Collins, CO), 0.3 mM δ-aminolevulinic acid, and 0.2 mM ferric citrate. Recombinant CYP4G2-CPR and HF-CPR from previously produced P3 viral stocks (Qiu \textit{et al.}, 2012) were used in Sf9 cells under similar conditions. Cells expressing CYP4G55-CPR, CYP4G56-CPR, CYP4G2-CPR or HF-CPR were then harvested with Cell Lysis Buffer (100 mM sodium phosphate, pH 7.6, 20% (vol/vol) glycerol, and 1.1 mM EDTA) supplemented with 100 μM DTT, 200 μM PMSF, and Protease Inhibitor Cocktail (Sigma-Aldrich, St. Louis, MO). Cells were lysed by sonication at an output of 4 Watt for 1 second on, 1 second off for 40 times, and cell debris and larger organelles were pelleted at 10,000 x g for 20 minutes at 4°C. Microsome pellets were obtained from centrifugation of lysate at 150,000 x g for 90 minutes at 4°C and solubilized in Cell Lysis Buffer. Functional CYP4G55-CPR and CYP4G56-CPR were quantitated by CO-difference spectrum analysis (Song \textit{et al.} 2013). Microsome preparations containing CYP4G55-CPR and CYP4G56-CPR were flash frozen and stored at -80°C for up to 1 month prior to assay.

\textbf{IV.3 Enzyme Assays}

cis-7-Decenal and NADPH were obtained from Sigma-Aldrich. Assays with cis-7-decenal were performed in duplicate 800 μL total volume reactions in pyrolyzed 5 ml glass vials. Each reaction contained 25 pmol of microsomal cytochrome P450, CYP4G
reaction buffer (0.02 M potassium phosphate, 0.05 mM sucrose, pH 7.4), 1.2 mM MgCl₂, 1.2 mM CaCl₂, 0.4 mM DTT, 25 µg substrate and 1.0 mM NADPH. The reactions were incubated at 30 °C overnight in a hybridization chamber with rotation and quenched with 200 µL 2 M HCl. Approximately 25 mg of NaCl was added to each reaction and products extracted with 500 µL hexane:ether (1:1) spiked with 10 ng/µL n-octanol. Extracts from each duplicate reaction were pooled prior to analysis by GC-MS. A trace chromatograph containing a 30-m, 25-µm film thickness DB-5 capillary column (Agilent) was programmed with the following parameters: initial temperature of 40 °C with no hold, ramp at 5 °C/min to 300 °C with a 5 min hold. The detector was an Agilent 5977A MSD with a molecular weight scanning range of 50 to 550 atomic mass units. Mass spectra were compared to the main EI MS library (NIST08) to confirm the identity of the product.

V. Results

The functionality of CYP4G55-CPR and CYP4G56-CPR fusion proteins in the microsomal preparations were measured by a CO difference spectrum. Figure 5.1 shows that both fusion proteins had the characteristic absorbance peak at 450 nm for active P450s. Reactions consisting of each P450-CPR fusion with NADPH and cis-7-decenal were performed and monitored for product by GC-MS. Both CYP4G55-CPR and CYP4G56-CPR reactions showed a peak at 5.02 minutes on the corresponding GC chromatograms, with the CYP4G55-CPR reaction exhibiting a larger peak, while reactions with CPR alone showed no peak at 5.02 minutes (Figures 5.2A-C). A
comparison of the mass spectrum from the CYP4G55-CPR reaction for the GC chromatogram peak at 5.02 minutes with the NIST08 MS library indicated a close match with cis-3-nonene (Figure 5.2D). Finally, a GC chromatogram of a cis-3-nonene standard showed a peak at 5.02 minutes (Figure 5.2E).

VI. Discussion

P450s in the CYP4G family are insect specific and have been associated with cuticular hydrocarbon biosynthesis and insecticide resistance (Feyereisen, 1999; Karatolos et al., 2012; Qiu et al., 2012; Balabanidou et al., 2016), but to date they have not been implicated in small molecule reactions. CYP4Gs have not been shown to directly metabolize insecticides and likely increase insecticide resistance by blocking uptake through increasing the amount of cuticular hydrocarbons (Balabanidou et al., 2016). In cuticular hydrocarbon biosynthesis, CYP4Gs are oxidative decarbonylases that catalyze the final step in the process by converting an aldehyde to a hydrocarbon (Qiu et al., 2012). Here we show that MPB CYP4Gs can catalyze the decarbonylation of cis-7-decenal to cis-3-nonene, a proposed step in the production of the MPB pheromone exobrevicomin.

Active recombinant CYP4Gs are difficult to produce. Indeed, a chimera consisting of the full-length CYP4G fused to the catalytic domain of CPR appears necessary as all active recombinant CYP4Gs have been produced in this form (Qiu et al., 2012; Balabanidou et al., 2016; S. Young and M. MacLean, personal communication). Here we used the same approach and found both MPB CYP4Gs could be recombinantly
expressed as active P450s when fused to the CPR catalytic domain. Moreover, we found that both CYP4G55 and CYP4G56 catalyzed the decarbonylation of cis-7-decenal to cis-3-nonene. The larger cis-3-nonene peak found in the GC chromatogram for CYP4G55 when compared to CYP4G56 is likely due to differing amounts of the P450 in each microsomal preparation. Careful preparation of the reactions to ensure equal amounts of each P450 and conducting enzyme kinetic assays in the future will allow for clarification of whether the reaction rates are different and may provide further insights into the structure and function of these two closely related P450s.

It is not clear from these experiments which MPB CYP4G is involved in exo-brevicomin biosynthesis in vivo, although both appear to accept longer chain aldehydes and alcohols as substrates and are likely involved in cuticular hydrocarbon biosynthesis (M. MacLean personal communication). Nadeau et al. (2017)(Chapter 3) showed a differential expression pattern based on sex and feeding status consistent with exo-brevicomin biosynthesis (increased expression in unfed males) for a CYP4G56-like protein, but whether this CYP4G56-like protein is in fact CYP4G56 or a yet to be identified CYP4G in MPB is still to be determined. Preliminary qRT-PCR investigations into CYP4G56 also showed a differential expression pattern with increased expression in unfed males whereas CYP4G55 did not (data not shown). Altogether, the work presented in this chapter provides the first evidence of a P450 in the CYP4G family being involved in biosynthesis of a short-chain pheromone component.
VII. References


VIII. Figure Legends

Figure 5.1. CO difference spectrum. CO difference spectrum of CYP4G55-CPR and CYP4G56-CPR fusion proteins showing the characteristic absorbance peak at 450 nm for active P450s.

Figure 5.2. Functional assays with cis-7-decenal. GC chromatograms of reactions with cis-7-decenal and microsomes containing (A) CYP4G55-CPR fusion, (B) CYP4G56-CPR fusion, and (C) CPR only. The large peak at 9.56 min is a 10 ng/µL n-octanol standard and the smaller peak at 5.02 minutes is cis-3-nonene. (D) shows the structure of cis-3-nonene and an MS scan of the 5.02 minute peak in (A) colored in red compared to cis-3-nonene from the NIST08 library in blue. A GC chromatogram of a cis-3-nonene standard is shown in (E).
Figure 5.1. CO difference spectrum.
Figure 5.2. Functional assays with cis-7-decenal.
I. Abstract

Cytochrome p450 monooxygenases (P450s) are important heme containing proteins found in almost all organisms and with functions ranging from the biosynthesis of metabolic intermediates to the detoxification of xenobiotic compounds. Understanding substrate specificities and enzyme kinetics of P450s can lead to insights into how structure relates to function in enzymes, the evolution of kinetic mechanisms, and factors affecting drug metabolism. The conventional method for measuring P450 kinetics is time-consuming and labor intensive, involving incubating microsomes containing the relevant P450 with different concentrations of substrate, stopping the reaction at various intervals with a chemical inhibitor, and measuring the change in substrate or product
concentrations using a method like liquid chromatography–tandem mass spectrometry. The typical P450 reaction, in addition to product formation, consumes NADPH and oxygen to form water. Here we test the OxoPlate® system, a 96-well plate with oxygen sensitive fluorophores, for measuring P450 kinetics and substrate specificities through rates of oxygen depletion. We found that under optimized conditions, the OxoPlate® is useful for measuring substrate specificities but is so far unreliable in measuring enzyme kinetics. Although we had limited success, the research presented here may lay the groundwork for further refinement of using the OxoPlate® system to monitor P450 reactions.

II. Introduction

Cytochromes P450 (P450s) are important heme containing proteins found in almost all organisms. They are implicated in a diverse range of functions, from production of metabolic intermediates to detoxification, through chemical modifications such as decarbonylation, epoxidation and hydroxylation (Denisov et al., 2005). In animals, P450s are membrane bound proteins that require an electron donor, often NADPH, and a redox partner, often the co-enzyme cytochrome P450 reductase (CPR), to allow for reduction of the heme group iron, a step necessary for binding an oxygen molecule and the subsequent donation of oxygen to the substrate (Werck-Reichhart and Feyereisen, 2000). Although substrate oxidations are the most common reactions, P450s can also catalyze a variety of other reactions including reductions, desaturations, and ring formations (Guengerich, 2001).
Over 20,000 distinct P450s have been identified. Current nomenclature guidelines suggest that members of new CYP families share at least 40% amino acid identity, while members of subfamilies must share at least 55% amino acid identity (Nelson et al., 2004; Nelson, 2009). P450s have great sequence diversity, with amino acid sequence identities dipping below 20% in some instances. Structural homology, however, is remarkably conserved, especially in the core of the protein (Werck-Reichhart and Feyereisen, 2000). The overall structure of P450s consists of a number of α-helices (A-L) and folds with a heme group located in the conserved core of the protein (helices D, E, I and L) and bound to the protein through a cysteine thiol bond (Figure 6.1). The most highly variable regions are in the substrate recognition and binding sites (Denisov et al. 2005). The substrates and products of most P450s are unknown and cannot be elucidated from sequence or structural homology alone (Denisov et al., 2005). P450s may be very specific for their substrates, but many can bind multiple different substrates, often yielding a variety of products from a single substrate (Denisov et al. 2005).

Most P450s require a redox partner and electron donor to function, and are organized into four distinct classes depending on which redox partner and/or electron donor are used. Class I P450s require an FMN containing reductase and iron sulfur reoxin to pass electrons from NADPH. Class II uses a cytochrome P450 reductase (CPR) containing both FMN and FAD domains to pass electrons from NADPH to the P450. Class III requires no redox partner or NADPH at all. Finally, Class IV enzymes take electrons directly from NAPDH (Werck-Reichhart and Feyereisen, 2000). Electrons from NADPH are required in most cases to reduce the iron in the heme group so that a dioxygen molecule may bind after a substrate has entered the substrate binding pocket. In
the most common P450 reaction, a number intermediate steps leads to oxygen being passed on to the substrate with an oxidized substrate and a molecule of water being released at the conclusion of the reaction (Groves and Han, 1995).

Understanding substrate specificities and enzyme kinetics can lead to insights into how structure relates to function in enzymes, and the evolution of kinetic mechanisms (Ulusu, 2015). P450 enzyme kinetics are also useful for studying drug metabolism to predict proper drug dosing and clearance times in the treatment of various maladies in both humans and animals (Kramer and Tracy, 2008). Most P450 reactions proceed with standard Michaelis-Menten enzyme kinetics and kinetic constants are easily calculated, however a number of atypical reactions can occur and models have been developed to calculate kinetic parameters for those reactions (Shou et al., 2001). High throughput methods for measuring P450 kinetics have been developed using radioactive, chromophoric, fluorescent or luminescent substrates or products for particular P450s (Moody et al., 1999; Stresser et al., 2000; Crespi, Miller and Stresser, 2002; Yamamoto, Suzuki and Kohno, 2002; Cali et al., 2006), but these substrates are suitable for very few enzymes. The conventional method for measuring P450 kinetics is time-consuming and labor intensive, involving incubating microsomes containing the relevant P450 with different concentrations of substrate, stopping the reaction at various intervals with a chemical inhibitor, and measuring the change in substrate or product concentrations using a method like liquid chromatography–tandem mass spectrometry (Obach and Reed-Hagen, 2002; Baranczewski, Edlund and Postlind, 2006). Measuring product formation in P450 reactions may also be difficult if the products are unknown and proper standards are unavailable (Obach and Reed-Hagen, 2002). Given these limitations, development of a
reliable medium or high-throughput method for measuring P450 substrate specificities and kinetics would be extremely useful.

Given the difficulties often associated with measuring substrate or product concentrations, there are strategies that monitor other components in order to follow reaction progress. Considering the overall reaction equation of the most common P450 reaction, NADPH and oxygen are also substrates:

\[
R-H + O_2 + \text{NADPH} + \text{H}^+ \rightarrow R-OH + H_2O + \text{NADP}^+ \quad \text{(Eq. 1)}
\]

NADPH and oxygen depletion are possible ways to measure P450 kinetics. Both can be measured in real-time with existing technologies. In fact, NADPH consumption coupled with product formation was used to measure enzyme kinetics for P450-mediated steroid hydroxylation (Kille et al., 2011). Medium-throughput methods for measuring P450 enzyme kinetics through oxygen depletion have also been developed using fluorescent probes (Olry et al., 2007; Zitova et al., 2010; Chang et al., 2012). Here, we test a 96-well microplate system that uses an oxygen sensitive fluorophore very similar to the system used in Olry et al. (2007) to measure oxygen depletion during P450 reactions. The microplate system used in Olry et al. (2007) is no longer commercially available, but an alternative, OxoPlate®, based on the same technology, is available. We used a bark beetle P450 from *Ips pini*, CYP9T3, which is relatively easy to produce in recombinant form and which hydroxylates a broad range of substrates (Song et al., 2013), to test whether the OxoPlate® is a viable solution for measuring P450 substrate specificity and enzyme kinetics in a medium-throughput format.
III. Methods

III.1 Recombinant CYP9T3 and HF-CPR Expression

CYP9T3 and house fly cytochrome P450 reductase (HF-CPR) high-titer P3 viral stocks produced using the BaculoDirect™ Baculovirus Expression System (Invitrogen, Carlsbad, CA) were produced in an earlier study by Song et al. (2013). Protein expression was initiated by infecting 50 mL of $10^6$ cells/mL Sf9 cells in Sf-900 II SFM culture media supplemented with 10% (vol/vol) fetal bovine serum (Atlas Biologicals, Fort Collins, CO), 0.3 mM δ-aminolevulinic acid, and 0.1 mM ferric citrate with 50 µL of the P3 viral stock and incubating at 27°C for 72 h. Recombinant CYP9T3 and HF-CPR were harvested 72 h post infection in cell lysis buffer (100 mM potassium phosphate, pH 7.6, 20% (vol/vol) glycerol, 1.1 mM EDTA, 200 µM PMSF, and protease inhibitor cocktail (Sigma-Aldrich, St. Louis, MO) and the microsomes were isolated by differential centrifugation essentially as per Nadeau et al. (2017). The microsomal fraction was tested for functional CYP9T3 using a CO-difference spectrum analysis (Choi et al., 2003) as described previously (Song et al., 2013). CYP9T3 concentrations ranged from 1.5 to 2 µM in the microsomal preparations.

III.2 Reaction Mixture and Incubation

Monoterpenes substrates (3-Carene, S-(-)-limonene, (+)-α-pinene, α-phellandrene, γ-terpinene and myrcene), bovine liver catalase, and the P450 inhibitor, piperonyl butoxide (PBO) were from Sigma-Aldrich (St. Louis, MO). Reaction mixtures consisted of 200 µL
of the CYP9T3 microsomal fraction, 50 µL of HF-CPR microsomal fraction, 100 mM sodium phosphate buffer pH 7.6 or 100 mM potassium phosphate buffer pH 7.6 and various concentrations of substrate in a total volume of 550 µL. Control reactions containing only HF-CPR were identical to the experimental reactions except that reaction buffer was substituted for the CYP9T3 microsomal fraction. Mixtures were incubated in capped 5 mL glass vials and rotated lengthwise at 30 °C in a FisherBiotech Hybridization Incubator (Thermo Fisher Scientific, Waltham, MA) for one hour to ensure microsomal uptake of the hydrophobic monoterpene substrates. For some experiments, 50 µL of 1,000 units/mL catalase in reaction buffer and/or 1 µL of PBO were included in the reaction mixtures before incubation.

III.3 Oxygen Depletion Measurements

OxoPlate® OP96U (PreSens Precision Sensing GmbH, Regensburg, Germany) were loaded with samples in a darkened room at room temperature. 150 µL of each reaction mixture was loaded into three wells of the microplate for a technical triplicate. The plate was then incubated in a SpectraMax M5 Multi-Mode Microplate Reader (Molecular Devices, Sunnyvale, CA) at 30 °C for 20 minutes to allow for fluorophore sensor and temperature equilibration. To initiate the reaction, 50 µL of preheated (30 °C) 4 mM or various concentrations of NADPH (Sigma-Aldrich) in reaction buffer were placed in each well and the microplate lid replaced. Fluorescence was measured from the bottom side of the plate in dual kinetic mode with two different filter pairs, the Indicator dye excitation and emission filters being 540 nm and 650 nm, respectively, and for the Reference dye 540 nm and 570 nm, respectively. Measurements were taken every minute
for up to 45 minutes in time-resolved mode with a 0 µs integration start and 500 µs integration time.

The OxoPlates® were calibrated using two different oxygen calibration standards as per the manual. One calibration solution contained oxygen saturated water at room temperature (k_{100}) and the other water containing no oxygen (k_0). For oxygen free water, 100 mg of sodium sulfite were dissolved in 10 mL of water in a capped glass vial with little head space. Air-saturated water was obtained by filling a 10 mL glass vial halfway, capping the vial and shaking vigorously for 2 minutes. Calibration standard solutions were prepared immediately before the initiation of measurements for each plate and the calibration standards were measured concurrently with other samples in the plate. The wells containing water with no oxygen were not sealed with adhesive foil as per the manual because preliminary measurements did not provide as strong a signal as when the wells were left unsealed.

The referenced signal (I_R) was calculated by dividing the measurements of the indicator dye (I_{indicator}) by the reference dye measurements (I_{reference}):

\[
I_R = \frac{I_{indicator}}{I_{reference}}
\]  
(Eq. 2)

The I_R is used to normalize the amount of indicator dye in each well by the amount of reference dye and allows for optimal well to well reproducibility.

III.4 Kinetic Measurements
For enzyme kinetic measurements, ten reaction mixtures with single monoterpenes concentrations ranging from 250 µM to 6 mM were assayed as described above. 3-Carene, myrcene and (+)-α-pinene were chosen as substrates for the kinetic assays based on strong oxygen consumption measurements from substrate specificity assays. Reactions were initiated by addition of 50 µL of 4 mM NADPH as described above and measurements were taken every minute for up to 45 minutes or until oxygen consumption ceased. Maximum rates of oxygen depletion were calculated for each reaction mixture using the “least squares” method by the command “linest” in Microsoft EXCEL. An EXCEL template, anemone.xlt, was used to calculate Michaelis-Menten kinetic parameters using non-linear regression (Hernández and Ruiz, 1998).

IV. Results

IV.1 Optimizing Reaction Conditions

P450 and substrate reaction conditions in the OxoPlate® were analyzed by varying the amounts of the P450 microsomal fraction in the reaction mixture (Figure 6.2A), adjusting the concentration of NADPH added to the reactions (Figure 2B), adding catalase to the reactions (Figure 6.3), and varying the amount of the HF-CPR microsomal fraction (Figure 6.4). There was a clear increase in oxygen depletion over time when increasing amounts of the CYP9T3 microsomal fraction are added, with the 150 µL and 200 µL amounts giving a relatively strong signal while 250 µL gave the best signal (Figure 6.2A). Figure 6.2B shows an increase in the rate of oxygen depletion over time as
well, with increasing concentrations of NADPH in the reaction mixture and a negative rate of oxygen depletion at the lowest concentration added (500 µM).

Adding catalase to the reactions containing both CYP9T3 and HF-CPR with (+)-α-pinene increased the rate of oxygen depletion when compared to the same reaction without catalase and a negative control containing only catalase, HF-CPR and the substrate (no CYP9T3) (Figure 6.3). When catalase was not added to the reactions, CYP9T3/HF-CPR with (+)-α-pinene showed very similar reaction rates to HF-CPR alone with (+)-α-pinene. Adding catalase to HF-CPR with (+)-α-pinene did not seem to affect the rate of oxygen depletion when compared to the same reaction mixture without catalase (Fig. 6.3). HF-CPR incubated with (+)-α-pinene showed very similar rates of oxygen depletion when compared to HF-CPR with 3-carene, however HF-CPR with myrcene showed very slow rates of oxygen depletion and was comparable to HF-CPR with no substrate.

Figure 6.4 shows reaction mixtures with varying amounts of the HF-CPR microsomal fraction with or without CYP9T3 and (+)-α-pinene as the substrate. Reactions with CYP9T3 and 400 µL of HF-CPR showed relatively high rates of oxygen depletion but were similar to reaction mixtures containing only 400 µL of HF-CPR (1.3:1 ratio). Decreasing the amount of HF-CPR in the reaction mixtures decreased the rate of oxygen depletion by a large amount, while the CYP9T3/HF-CPR reactions had higher rates than the HF-CPR only (1.8:1 ratio). CYP9T3 and 50 µL HF-CPR reactions showed very similar or slightly greater oxygen depletion rates when compared to CYP9T3 and 100 µL HF-CPR reactions. 50 µL HF-CPR reactions without CYP9T3 exhibited very low rates of oxygen depletion (Figure 6.4) and the ratio of the rates of oxygen depletion in the
CYP9T3 and 50 µL HF-CPR reactions compared to 50 µL HF-CPR control reactions was 7.3 to 1.

Finally, the P450 inhibitor piperonyl butoxide (PBO) was added to reaction mixtures with HF-CPR and (+)-α-pinene to assess whether endogenous P450s in the HF-CPR microsomal fraction were reacting with (+)-α-pinene. The addition of PBO did cause a decrease in the rate of oxygen depletion when compared to the same reaction without PBO, however the rates of oxygen depletion were still much higher when compared to HF-CPR with myrcene or the no substrate control (Figure 6.5).

IV.2 Substrate Profiling and Enzyme Kinetics

Using reaction mixtures optimized for rates of oxygen depletion (200 µL of CYP9T3, 4 mM NADPH and 50 µL of HF-CPR) and taking into account the availability of the components, a substrate profile for CYP9T3 was created using the OxoPlate® (Figure 6.6). Six different monoterpenes were tested ((+)-α-pinene, 3-Carene, myrcene, S-(-)-limonene, α-phellandrene and γ-terpinene) by adding 1 µL of each to separate reaction mixtures. Figure 6.6 shows that (+)-α-pinene, 3-Carene and myrcene all showed similar rates of oxygen depletion while S-(-)-limonene, α-phellandrene and γ-terpinene showed essentially no oxygen depletion.

A preliminary analysis of enzyme kinetics using the Michaelis-Menten model was attempted for CYP9T3 and the substrates myrcene, 3-carene and (+)-α-pinene (Figure 6.7). Although lower substrate concentrations in the reactions of CYP9T3 with myrcene or 3-carene showed abnormally high reaction velocities, the model was still able to approximate $K_m$ and $V_{max}$ values of 0.89 ± 0.49 mM and 0.01 mM/minute respectively.
for myrcene and $16.92 \pm 0.94$ mM and $0.11 \pm 0.02$ mM/minute respectively for 3-carene. The model fit much better for the CYP9T3 and (+)-α-pinene reactions, yielding $K_m$ and $V_{max}$ values of $1.32 \pm 0.16$ mM and 0.02 mM/minute respectively. Unfortunately, attempts to repeat the experiment for CYP9T3 and (+)-α-pinene (Figure 6.8) and 3-carene (data not shown) were unsuccessful. Figure 6.8 shows that the rates of oxygen depletion in subsequent CYP9T3 reactions with (+)-α-pinene were erratic and did not seem to respond in a predictable manner to different substrate concentrations.

V. Discussion

Here, we tested the suitability the oxygen sensing OxoPlate® 96-well microplate oxygen-sensing system to monitor P450 substrate specificities and enzyme kinetics. Although we had limited success, this research may lay the groundwork for further refinement of using the OxoPlate® system to monitor P450 reactions. Previous work has shown that 96-well microplate oxygen-sensing systems are viable for measuring P450 substrate profiles, kinetics and response to inhibitors, (Yamamoto, Suzuki and Kohno, 2002; Olry et al., 2007). The OxoPlate® used in this study is superior to previous 96-well microplate systems used to study P450 reactions because it is highly sensitive, easy to calibrate, has a fast response time and high well to well reproducibility (Arain et al. 2006). However, P450 reaction systems are complex, in most instances requiring the P450, a redox partner, an electron donor, access to oxygen and a substrate in a reaction mixture with buffer composition and temperature affecting reaction rates (Werck-Reichhart and Feyereisen, 2000; Denisov et al., 2005). Microsomal preparations of the
P450 and redox partner, CPR, also tend to exacerbate the complexity of the system because endogenous proteins from the expression vehicle (insect cells in this case) are still present and may react with components of the P450 reaction system. P450s also may interact with themselves or other P450s and these interactions can further affect reaction mechanics including accessibility of the substrate channel or ability to interact with CPR (Scott et al., 2016). In our system, the hydrophobic nature of the monoterpane substrates made accurate concentration measurement difficult, and also called into question the availability of the substrates to the enzyme, as a portion of the substrates is likely sequestered in the microsomal membranes. Finally, not all P450 carry out the same chemistry, sometimes oxidizing the substrate in the form of hydroxyl or epoxide formations, while other times reducing and cyclizing the substrate or even a combination of multiple reactions (Guengerich, 2001). Therefore, accurate measures of substrate specificities, enzyme kinetics and inhibitor effects of P450s require optimization of numerous factors beyond the simple or complex mechanics of the measurement system.

Initial tests showed that increasing either P450 or NADPH concentrations in reactions correlated with increasing oxygen depletion rates when substrate concentrations were high (~16 mM) (Figure 6.2). This suggests, as expected, that the reaction rate was limited by the amount of P450 and NADPH present. Thus, the highest tested concentrations of each of these components were chosen for further experiments to ensure adequate resolution when monitoring the reactions.

At higher concentrations, HF-CPR controls exhibited oxygen depletion rates very similar to those found when the P450 was present in a 1.3 to 1 ratio with CPR compared to HF-CPR controls (Figure 6.4). This suggests an uncoupling of the oxygen
consumption reaction from the P450 reaction we were attempting to measure. To elucidate whether endogenous P450s in the HF-CPR microsomal fraction were reacting with our substrate, the insect P450 inhibitor PBO was added to reaction mixtures containing just the HF-CPR microsomes. Indeed, lower rates of oxygen depletion were observed when PBO was present compared to when it was not, although not to the degree of when no substrate or the substrate myrcene were present (Figure 6.5). These data suggest that endogenous (Sf9 cell) P450s in the HF-CPR microsomal fraction react with at least some of our substrates. Higher concentrations of PBO in the reaction mixture were not tested but it is likely that greater inhibition would occur with increasing concentration. This, retrospectively, is not surprising given that these microsomes are derived from insect cells that should have detoxification enzymes including P450s that can metabolize toxic compounds like the monoterpene substrates used in this study. Interestingly, myrcene did not appear to be a substrate for the microsomal endogenous P450s indicating a possible preference of cyclical monoterpenes as substrates.

The effects of HF-CPR concentrations on the reaction rates were explored. To minimize oxygen depletion due to endogenous P450 activity, the effects of lowering the concentration of HF-CPR microsomal fraction were tested while keeping the P450 microsomal fraction constant. The rate of oxygen depletion in CP9T3 reactions with (+)-α-pinene and only 50 µL HF-CPR was similar to or greater than the rate observed when double that amount of HF-CPR microsomal fraction (100 µL) was added (Figure 6.4), however incubations of the same amount (50 µL) of the HF-CPR microsomal fraction in a 550 µL reaction without the P450 microsomal fraction showed very little observable oxygen depletion (Figure 6.5). This suggested that the oxygen depletion from
endogenous sources could be minimized while still attaining high enough activity levels for accurate measurements of the P450. Therefore, 50 µL of the HF-CPR microsomal fraction was used for all subsequent assays.

While NADPH and oxygen depletion measurements compared to GC/MS seem to be quick, less labor intensive and able to be measured in real-time, it is unclear that the measurements really correlate with substrate depletion because uncoupled P450 reactions that produce superoxide radicals and hydrogen peroxide would consume O₂ but not oxidized product (Anseed and Thakker, 2004; Traylor et al., 2011). To overcome this limitation, Traylor et al. (2011) used superoxide dismutase and catalase to form water from superoxide and hydrogen peroxide, allowing for direct measurement of oxygen consumption from substrate oxidation alone. In our experiments, catalase increased the rate of oxygen depletion when both the HF-CPR and P450 microsomal fractions were present but did not affect oxygen depletion rates with the HF-CPR microsomal fraction alone (Figure 6.3). This may represent a shift from peroxide formation to other uncoupling reactions that produce superoxide radicals. Further experimentation with additions of catalase and superoxide dismutase are needed to ascertain the effects on measuring P450 activity using the OxoPlate® system as the experiments conducted in this study are insufficient to draw any conclusions.

Screening the substrate specificity of P450s is time consuming using GC/MS, whereas using the OxoPlate® system can significantly decrease that time through the ability to screen dozens of compounds simultaneously in less than an hour. Six potential monoterpene substrates for CYP9T3 were tested using the OxoPlate® and the methods described above. A substrate profile for CYP9T3 using most of the same monoterpenes
was conducted previously by Song et al. (2013) using GC/MS and the results were confirmed in this study (Figure 6.6). Both this study and Song et al. (2013) showed that myrcene, 3-carene and (+)-α-pinene were likely substrates while α-phellandrene and γ-terpinene were not. S-(−)-Limonene also showed no oxygen depletion using the OxoPlate® method and is likely not a substrate of CYP9T3. Given this data, the OxoPlate® seems to be a viable method to measure P450 substrate specificities but measuring enzyme kinetics may be unreliable due to the uncoupling reactions mentioned above.

The final assay conducted in this study was to use the OxoPlate® method to measure CYP9T3 enzyme kinetics for the substrates myrcene, 3-carene and (+)-α-pinene. Initial attempts yielded moderately successful results with data amendable to fitting the Michaelis-Menten model (Figure 6.7). At the lower concentrations of myrcene and 3-carene, however, the fit was not very reliable and could have been the result of the hydrophobic nature of these substrates. Monoterpenes in general have very low solubility and it is likely that the lower dilutions of myrcene and 3-carene were not mixed sufficiently to ensure the correct delivery of the desired concentration to the reaction mixture. Nonetheless, the kinetic data suggest CYP9T3 has a higher affinity for myrcene (Km = 0.89 ± 0.49 mM) and α-pinene (Km = 1.32 ± 0.16 mM) than for 3-carene (Km = 16.92 ± 0.94 mM). Confirmation of these kinetic parameters by the GC/MS method would be useful. Unfortunately, subsequent attempts to use the OxoPlate® method for measuring CYP9T3 kinetics were unsuccessful. Figure 6.8 shows one of the more recent attempts with CYP9T3 and (+)-α-pinene has no discernable pattern of rates of oxygen depletion in response to the amount of substrate added. The reasons why the later
experiments did not work are unknown but perhaps the substrates were no longer viable because of oxidation over time in the storage vials or some settings for the plate reader were changed after maintenance was performed. Work on this project was set aside to concentrate on the other experiments presented in this dissertation. Although we had limited success, the research presented in this chapter may lay the groundwork for further refinement of using the OxoPlate® system to monitor P450 reactions.

VI. References


chemicals, 30(7), pp. 831–7.


Stresser, D. M., Turner, S. D., McNamara, J., Stocker, P., Miller, V. P., Crespi, C. L. and Patten, C. J. (2000) ‘A High-Throughput Screen to Identify Inhibitors of


VII. Figure Legends

**Figure 6.1. P450 structure.** The tertiary structure of cytochrome P450s is highly conserved and shown in a ribbon representation (distal face). Substrate recognition sequence (SRS) regions are black and labeled. R-Helixes mentioned in the text are labeled with capital letters. Reprinted with permission from Denisov, I.G. et al., 2005. Structure and chemistry of cytochrome P450. Chemical reviews, 105(6), pp.2253–77. Copyright 2005 American Chemical Society.

**Figure 6.2. Optimizing P450 and NADPH concentrations.** OxoPlate® measurements of the maximum rate of oxygen depletion in 550 µL reaction mixtures containing 250 µL HF-CPR, 100 mM sodium phosphate buffer pH 7.6, 2 µL of (+)-α-pinene and varying amounts of (A) the CYP9T3 microsomal fraction with 50 µL of 4 mM NADPH or (B) NADPH with 200 µL of CYP9T3 microsomal fraction. NADPH was added to each well after 150 µL of each reaction mixture was allocated to one of three wells on the plate. Values were calculated based on linear regression of oxygen levels in three averaged samples over time and the error bars represent the standard error of the regression. The arbitrary units are I_R/minute.

**Figure 6.3. Catalase additions.** OxoPlate® measurements of the maximum rate of oxygen depletion in 550 µL reaction mixtures containing 250 µL HF-CPR, 100 mM sodium phosphate buffer pH 7.6, 2 µL of (+)-α-pinene, 3-carene, myrcene or no substrate and with or without 200 µL of the CYP9T3 microsomal fraction. 50 µL of 4mM
NADPH was added to each well after 150 µL of each reaction mixture was allocated to one of three wells on the plate. Approximately 50 units of catalase was added to select reactions before addition of NADPH. Values were calculated based on linear regression of oxygen levels in three averaged samples over time and the error bars represent the standard error of the regression. The arbitrary units are Relative Fluorescent Units/minute.

Figure 6.4. Optimizing CPR concentrations. OxoPlate® measurements of the maximum rate of oxygen depletion in 550 µL reaction mixtures containing 50-400 µL HF-CPR, 100 mM sodium phosphate buffer pH 7.6, 2 µL of (+)-α-pinene, and with or without 200 µL of the CYP9T3 microsomal fraction. 50 µL of 4mM NADPH was added to each well after 150 µL of each reaction mixture was allocated to one of three wells on the plate. Values were calculated based on linear regression of oxygen levels in three averaged samples over time and the error bars represent the standard error of the regression. The arbitrary units are I_R/minute.

Figure 6.5. PBO addition. OxoPlate® measurements of the maximum rate of oxygen depletion in 550 µL reaction mixtures containing 250 µL HF-CPR, 100 mM sodium phosphate buffer pH 7.6, 2 µL of (+)-α-pinene, 3-carene, myrcene or no substrate and with or without 200 µL of the CYP9T3 microsomal fraction. 50 µL of 4mM NADPH was added to each well after 150 µL of each reaction mixture was allocated to one of three wells on the plate. 1 µL of PBO was added to a CPR only reaction before addition of NADPH. Values were calculated based on linear regression of oxygen levels in three
averaged samples over time and the error bars represent the standard error of the regression. The arbitrary units are Relative Fluorescent Units/minute.

**Figure 6.6. Substrate profiling.** OxoPlate® measurements of the maximum rate of oxygen depletion in 550 µL reaction mixtures containing 50 µL HF-CPR, 200 µL of the CYP9T3 microsomal fraction, 100 mM sodium phosphate buffer pH 7.6, and 1 µL of (+)-α-pinene, 3-carene, myrcene, S-(-)-limonene, α-phellandrene or γ-terpinene. 50 µL of 4mM NADPH was added to each well after 150 µL of each reaction mixture was allocated to one of three wells on the plate. Values were calculated based on linear regression of oxygen levels in three averaged samples over time and the error bars represent the standard error of the regression. The arbitrary units are I_R/minute.

**Figure 6.7. Enzyme kinetics measurements.** Michaelis-Menten plots and calculated K_m and V_max values (mM) for CYP9T3 reactions with (A) myrcene, (B) 3-carene and (C) (+)-α-pinene. The reaction mixtures were a total volume of 550 µL reaction mixtures containing 50 µL HF-CPR, 200 µL of the CYP9T3 microsomal fraction, 100 mM sodium phosphate buffer pH 7.6 and 8 different concentrations of (+)-α-pinene or 10 different concentrations of 3-carene or myrcene. 50 µL of 4mM NADPH was added to each well after 150 µL of each reaction mixture was allocated to one of three wells on the plate. Observed velocities were calculated based on linear regression of oxygen levels in three averaged samples over time. The arbitrary units are I_R/minute.
Figure 6.8. Repeating enzyme kinetics measurements. OxoPlate® measurements of the maximum rate of oxygen depletion in 550 µL reaction mixtures containing 50 µL HF-CPR, 200 µL of the CYP9T3 microsomal fraction, 100 mM potassium phosphate buffer pH 7.6, and 10 different concentrations of (+)-α-pinene. 50 µL of 4mM NADPH was added to each well after 150 µL of each reaction mixture was allocated to one of three wells on the plate. Values were calculated based on linear regression of oxygen levels in three averaged samples over time and the error bars represent the standard error of the regression. The arbitrary units are $I_R$/minute.
Figure 6.1. P450 structure.
Figure 6.2. Optimizing P450 and NADPH concentrations.
Figure 6.3. Catalase additions.
Figure 6.4. Optimizing CPR concentrations.
Figure 6.5. PBO addition.
Figure 6.6. Substrate profiling.
Figure 6.7. Enzyme kinetics measurements.
Figure 6.8. Repeating enzyme kinetics measurements.
Chapter 7
Discussion and future directions

I. Discussion

II. Future directions

III. References

I. Discussion

I.1 Background

Bark beetles are highly destructive pests and have devastated forests of North America in recent years, killing over forty million acres of trees in the past 15 years in the western United States (United States Department of Agriculture, 2017) and approximately 18.3 million hectares, greater than 53% of the merchantable pine, in western Canada (Corbett et al., 2016). Two of these destructive bark beetles, *Dendroctonus ponderosae* Hopkins (mountain pine beetle, MPB), and *Dendroctonus jeffreyi* (Jeffrey pine beetle, JPB) spend the majority of their life cycles sequestered beneath the bark of their host trees, making them extremely difficult to target for typical pest control strategies (Gibson et al., 2009; Smith et al., 2009). MPB are generalists, mainly infest lodgepole and ponderosa pine trees but may successfully infest any pine tree within their range (Gibson et al., 2009). JPB, which is genetically closely related to MPB, is a specialist that only successfully infests Jeffrey pine trees (Smith et al., 2009).

Successful MPB attacks require coordinated information from three pheromone components that, in synergy with host monoterpenes, create aggregation and anti-
aggregation signals (Blomquist et al., 2010). The first component produced during an attack is \((-\text{-}\text{trans-})\text{-verbenol} \[(1S,2R,5S)-4,6,6\text{-trimethylbicyclo[3.1.1]hept-3-en-2-ol}\], which is synthesized by feeding (pioneer) females. \((-\text{-}\text{trans-})\text{-Verbenol is an aggregation pheromone that attracts other males and females to a newly colonized host tree (Pitman et al., 1968). In synergy with the female produced \((-\text{-}\text{trans-})\text{-verbenol, newly emerged male MPB produce exo-brevicomin} \[(1R,5S,7R)-7\text{-Ethyl-5-methyl-6,8-dioxabicyclo[3.2.1]octane}\] in small concentrations as an aggregation pheromone component, with production decreasing significantly upon reaching the new host tree and mating (Pureswaran et al., 2000). The third component, frontaldehyde \[(1R,5R)-1,5\text{-dimethyl- 6,8-dioxabicyclo[3.2.1]octane}\], is produced by MPB males upon feeding and acts as an anti-aggregation signal to halt the attack and prevent overcrowding of the tree (Barkawi et al., 2003).

While aggregation pheromones are critical for mass attacks and successful reproduction, detoxification enzymes are also essential for MPB to survive the toxic environment of the host tree (Keeling and Bohlmann, 2006; Robert et al., 2013). Insect P450s are involved in oxidation of toxic terpenoids to allow for easier excretion (Li et al., 2004). It has been hypothesized that pheromone biosynthetic P450s evolved from detoxification P450s because some hydroxylated monoterpenes are both detoxification products and pheromone components. Enzymes involved in the two processes may be distinguished by expression patterns and substrate profiles (Seybold, Bohlmann and Raffa, 2000; Blomquist et al., 2010). For instance, detoxification P450s are predicted to have a broad substrate range and be induced by exposure to toxic terpenoids through feeding or inhalation, while pheromone biosynthetic P450s are likely to have a narrower
substrate range with strong preference for the pheromone precursor, and be induced by feeding or hormones (Blomquist et al., 2010; Song et al., 2013).

Because bark beetles spend most of their life cycles beneath the bark of infested trees they are largely unaffected by insecticide application. Large scale management attempts have been almost solely restricted to standard silvicultural practices such as harvesting trees that have been invaded by bark beetles, but have produced very limited success given the massive scale of the problem (Fettig et al., 2007). However, employing pheromones to bait trees treated with pesticides has been marginally effective in protecting stands of trees (Wermelinger, 2004). Therefore, the right combination of silvicultural practices, pesticide application, and pheromone baiting may increase the efficiency of bark beetle management given an ample supply of pheromones. Additionally, identifying bark beetle-specific enzymes involved in resin detoxification and pheromone biosynthesis may allow for species specific molecular targets for pest management strategies.

Very few of the MPB resin detoxification P450s and enzymes involved in pheromone biosynthesis have been identified and functionally characterized to date. The major components and proposed biosynthetic pathways for MPB pheromones have been known for years (Blomquist et al., 2010) (Figure 3.1), but the specific enzymes involved in the individual steps are still unknown in many cases. Two MPB genes, *CYP6CR1* and *ZnoDH*, were identified through differential gene expression analysis using the first ever microarray for this species and were shown to have increased transcript levels in unfed males, consistent with *exo*-brevicomin biosynthesis (Aw et al., 2010). Functional analysis through cloning and enzyme assays confirmed these two enzymes are involved in the
biosynthetic pathway leading to \textit{exo}-brevicomin (Song \textit{et al.}, 2014). Another study confirmed that MPB frontalin biosynthesis proceeds through the mevalonate pathway and employs a geranylgeranyl diphosphate synthase to produce the 20-carbon intermediate geranylgeranyl diphosphate (GGPP) (Keeling \textit{et al.}, 2013). Finally, a recent study conducted by Keeling \textit{et al.} (2016) examined the transcriptome, proteome and metabolome in a sex-specific manner of MPB midgut and fat body tissue in response to JH treatments to identify candidate genes for pheromone biosynthesis, but no functional characterization of identified enzymes was conducted (Keeling \textit{et al.}, 2016). Functional assays to confirm the activities of candidate enzymes are necessary as genetic screens are impractical, although Keeling \textit{et al.} (2013) was able to develop an RNAi-based screen to substantiate the finding that GGPP is an intermediate in frontalin biosynthesis. Thus, while many candidate enzymes have been provisionally identified, most are unconfirmed.

This dissertation used a comparative transcriptomics approach backed by functional analyses to identify genes likely involved in resin detoxification and pheromone biosynthesis in the MPB. Novel candidate genes –in addition to those proposed by Keeling \textit{et al.} (2016)– were identified for MPB pheromone biosynthesis, five of which encode P450s that were functionally characterized for the first time. Two of the P450s analyzed had expression patterns and activity consistent with involvement in \textit{exo}-brevicomin and frontalin biosynthesis. This work represents not only a significant leap forward in ascertaining the identity of MPB pheromone biosynthetic genes, but also provides insights into the evolutionary history between detoxification and pheromone biosynthetic P450s and lays the ground work for a better understanding of P450 structure and function in general.
Chapter 2 investigated the expression patterns and function of CYP6DH3, a member of the CYP6DH family of MPB P450s. The CYP6DH family has three relatively closely related members, two of which, CYP6DH1 and CYP6DH2 are probably allelic and have been previously characterized based on expression and catalytic activity. They are likely involved in pheromone biosynthesis and resin detoxification, respectively (Aw et al., 2010; Gorzalski, 2010; Song, 2012). As both CYP6DH1 and CYP6DH2 catalyze the hydroxylation of monoterpenes, we hypothesized that the third member encoded by this family, CYP6DH3, also hydroxylates monoterpenes.

We found that CYP6DH3 mRNA levels are unaffected by differences in gender but are overall significantly higher in unfed compared to fed beetles, with decreased levels in the midgut, posterior midgut and hindgut of feeding beetles compared to the same tissues of unfed beetles (Figure 2.1). Monoterpene exposure caused a significant increase in CYP6DH3 mRNA levels in both males and females with no significant difference overall between genders. However, there was a highly significant interaction between monoterpene exposure and gender as is evident from exposure to (+)-α-pinene, with males having increased CYP6DH3 expression compared to control beetles, and females having highly decreased expression levels (Figure 2.2). Alternatively, when exposed to a monoterpene cocktail with all the tested monoterpenes included, mRNA levels in females were higher than those in controls, while males had lower CYP6DH3 levels in general. The complex regulatory regimes affecting CYP6DH3 mRNA levels are consistent with gender-based differences in expression of P450s in response to
monoterpene exposures for both *Dendroctonus armandi* (Dai *et al.*, 2015) and MPB (Chapter 3, Nadeau *et al.*, 2017). Taking both tissue and monoterpene exposure profiles into account, CYP6DH3 expression does not correlate well with the expected expression pattern of either pheromone biosynthesis or resin detoxification.

Although expression patterns do not correlate well with those expected of a resin detoxification enzyme, functional analysis of CYP6DH3 showed a relatively broad substrate range similar to CYP6DH2. CYP6DH2 is thought to be a resin detoxification P450 and catalyzes the hydroxylation of (–)-α-pinene, (+)-α-pinene, 3-carene and R- (+)-limonene (Song, 2012), while CYP6DH3 similarly accepts the same monoterpenes, except (–)-α-pinene, as substrates. Interestingly, the identity of the products produced by CYP6DH3 did not correlate well with the products produced by CYP6DH2 based on MS data. It is not necessarily surprising that even though CYP6DH3 is closely related (68% amino acid identity) to CYP6DH1/2, their substrate ranges and/or products are not identical. Small differences in the amino acid composition of the solvent channel, substrate binding site, or even other locations that affect tertiary structure can cause vast differences in P450 substrate specificity and product formation (Li *et al.*, 2004; Denisov *et al.*, 2005; Bernhardt, 2006).

The data provided here thus help with understanding the evolutionary history of the *CYP6DH* family and provide insight into the structure and function of P450s in general. The expression patterns of *CYP6DH1, CYP6DH2* and *CYP6DH3* suggest very different regulatory mechanisms for each of these closely related P450s, highlighting the complex nature of these enzymes and their evolutionary adaptability to a number of roles within the host organism. In this respect, it is worth remembering that distinguishing
between “pheromone biosynthetic” and “resin detoxifying” as role descriptors may not be accurate in this case. Unlike other insects that have well-defined cells or glands dedicated to pheromone production (Tillman et al., 1998), both processes occur in the MPB midgut along with digestion. (−)-trans-Verbenol production requires only a single metabolic step, compared to the fully-developed pathways for frontalin, exo-brevicomin (Figure 1.1), and ipsdienol (Sandstrom et al., 2006). The mount of (−)-trans-verbenol produced (less than 40 ng/female) (Pureswaran et al., 2000) is ten-fold less than that for ipsdienol in male Ips pini. Thus, with the metabolic load being lower in midgut cells compared to Ips bark beetles or pheromone-producing cells in other insects, evolutionary pressure to “focus” enzyme activity on pheromone production over detoxification is likely lower. This may explain the complex regulatory processes evident for the CYP6DH genes in MPB.

I.3 Chapter 3. Comparative transcriptomics of mountain pine beetle pheromone-biosynthetic tissues and functional analysis of CYP6DE3

In Chapter 3 a comparative RNA-Seq analysis between fed and unfed male and female MPB midguts and fat bodies was used to identify candidate genes involved in pheromone biosynthesis. We employed network, differential gene, and gene ontology enrichment analyses to isolate candidates involved in pheromone biosynthesis. Additionally, functional analysis of a candidate gene, CYP6DE3, identified by our RNA-Seq analysis for exo-brevicomin biosynthesis, was conducted. Recently, Keeling et al. (2016) reported transcriptome, proteome, and metabolome-level responses of unfed MPB treated with juvenile hormone (JH) III. JH III stimulates aggregation pheromone
component production in bark beetles (Borden, Nair and Slater, 1969; Chen, Borden and Pierce, 1988; Tillman et al., 1998; Barkawi et al., 2003; Keeling et al., 2006), including frontalin and trans-verbenol in MPB, and several “pheromone-biosynthetic gene” candidates were identified by this study. However, the dynamics of the transcriptional response of JH III-treated insects can differ from those of fed insects (Tittiger, Keeling and Blomquist, 2005; Keeling et al., 2006; Bearfield et al., 2009) and the study noted that JH III treatment did not affect exo-brevicomin production. Thus, JH treatment may not induce differential expression of all genes involved in pheromone biosynthesis for the MPB. Our study differed in that we focused on midgut and fat body tissues of fed and unfed insects rather than JH III-treated insects. We identified all four known MPB pheromone biosynthetic genes, confirmed the tentative identification of four others from Keeling et al. (2016), and suggested nine novel candidates (Table 3.4).

One candidate enzyme arising from the RNA-Seq analysis was CYP6DE3. mRNA levels in response to monoterpane exposures and functional analysis indicated CYP6DE3 is likely involved in resin detoxification rather than pheromone biosynthesis. The clear elevation observed following exposure to all tested monoterpenes (Figure 3.7) suggests that CYP6DE3 is induced by monoterpane exposure, particularly in females, implying a resin-detoxifying role. The absence of this induction in fed insects further implicates that CYP6DE3 regulation is complex. The monoterpane-dependent difference in response in males and females is curious, but has been exhibited in another study reporting similar sex-specific transcriptional responses by CYP6DH2 in MPB (Song, 2012) and of various D. armandi P450 genes in response to monoterpenes (Dai et al., 2015). A detoxification role for CYP6DE3 is supported by functional assays of the
recombinant enzyme which showed that it oxidized a variety of monoterpenes, but did not appear to accept exo-brevicomin precursors as substrates (Figure 3.8 and data not shown).

I.4 Chapter 4. Expression Profiling and Functional Analysis of CYP6DK1, a Frontalin Biosynthetic Gene in the Mountain Pine Beetle

Chapter 4 focused on exploring the expression patterns and functional analysis of CYP6DK1, a candidate frontalin biosynthetic gene identified in Chapter 3. Frontalin production via the mevalonate pathway takes place in the midgut of fed male MPB (Barkawi et al., 2003; Blomquist et al., 2010). We hypothesized that if CYP6DK1 is involved in frontalin biosynthesis, its mRNA levels would be highest in fed male midguts compared to other tissues. CYP6DK1 mRNA levels were indeed highest in the midgut of fed males (Figure 4.1). However, there was increased expression of CYP6DK1 in unfed male and fed female midguts as well when compared to other tissues, indicating a possible additional function to frontalin production. Female MPB do not naturally produce frontalin, but they can produce frontalin upon exposure to 6-methylhept-6-en-2-one (6MHO) (Perez et al., 1996). Thus, the sex-specificity of frontalin production lies upstream of CYP6DK1, possibly at the step of 6MHO biosynthesis from GGPPS. Indeed, significant production of 6MHO only occurs in feeding male MPB (Keeling et al., 2016) so the fact that unfed males and fed females do not naturally produce frontalin is not surprising even though both express relatively similar levels of CYP6DK1.

An additional function for CYP6DK1 is suggested by elevated mRNA levels in response to monoterpene exposures. We tested the hypothesis that CYP6DK1 is involved
in detoxification by monitoring expression levels of CYP6DK1 following exposure of unfed male and female MPB to monoterpenes. We found that both unfed male and female MPB had significantly increased CYP6DK1 mRNA following exposure to all (male) or most (female) of the monoterpenes tested (Figure 4.2), suggesting that CYP6DK1 may function in resin detoxification. Such a role is consistent with elevated mRNA levels for the gene in alimentary canals of feeding females (Figure 4.1). Alternatively, the change in CYP6DK1 mRNA levels upon monoterpane exposure may be due to monoterpane-related alterations in endogenous metabolism. It is also possible that CYP6DK1 functions in both frontalin production and resin detoxification. This possible dual role has precedence in *Ips pini*, where CYP9T3 produces ipsdienol from myrcene as a pheromone-biosynthetic reaction, but can also accept other monoterpenes as substrates (Song *et al.*, 2013). These data support the hypothesis that pheromone biosynthetic P450s evolved from detoxification P450s (Seybold *et al.*, 2000; Blomquist *et al.*, 2010).

Given that the expression patterns of *CYP6DK1* from both this and previous studies support the possibility of involvement in frontalin biosynthesis, we conducted a functional analysis of recombinant CYP6DK1 using geranylgeranyl diphosphate, 6-methylhept-5-en-2-one and 6MHO as possible substrates because they are all compounds hypothesized to be a part of the frontalin biosynthetic pathway (Figure 1.1B). In our first attempt at functional analysis, we used CYP6DK1 fused to the catalytically active domain of housefly cytochrome P450 reductase (CPR). Previous attempts to express recombinant P450s, especially in the CYP4G family, required a fusion of the P450 to CPR for functionality (Qiu *et al.* 2012; Balabanidou *et al.* 2016; S. Young and M.
MacLean personal communication), and P450-CPR fusions can assist functional assays of different P450s (e.g. Dodhia et al., 2006; Leonard and Koffas, 2007; Lundemo et al., 2016). Although CYP6DK1 is not in the CYP4G family, we assumed a CYP6DK1_CPR fusion would yield a functional P450 that we could use in assays. Expressing the CYP6DK1_CPR fusion did in fact yield a functional P450 (Figure 4.3A) that generated products when incubated with 6MHO but not geranylgeranyl diphosphate or 6-mehtylhept-5-en-2-one (Figure 4.4, data not shown). From the hypothesized pathway (Figure 1.1B), we expected formation of 6,7-epoxy-6-methylheptan-2-one and the need of a cyclase to convert the epoxide to frontalin. Interestingly, the products observed from the CYP6DK1_CPR and 6MHO reactions were frontalin and 6-methylheptan-2-one, confirming a role for CYP6DK1 in frontalin production and indicating that a cyclase is not necessary. If the epoxide product of the reaction is relatively unstable, it is possible that cyclization may occur while it is still bound to CYP6DK1, assuming that the enzyme holds the epoxide in a conformation that facilitates the cyclization reaction. Indeed, other P450s have cyclase activity (Guengerich, 2001) so CYP6DK1 is not unique in that aspect.

The other product, 6-methylheptan-2-one, was produced in much greater quantities than frontalin and appears to be a result of the reduction of the 6,7 double bond in 6MHO. This suggests CYP6DK1_CPR catalyzes a reduction in contrast to the normal P450 oxidase activity. A possible explanation for this unusual reaction is that CYP6DK1, being fused to CPR, may occasionally uncouple the normal catalytic cycle whereby electrons are passed directly onto the substrate instead of being passed to oxygen to form water. To test this hypothesis, we expressed recombinant CYP6DK1 that was not fused to
CPR and conducted assays of this “unfused” (native) enzyme with 6MHO in the presence of full-length, recombinant CPR. Fortunately, the un-fused enzyme was functional (Figure 4.3B), and it similarly produced frontalin following incubation with 6MHO, while production of 6-methylheptan-2-one was not observed (Figure 4.5). This suggests that production of 6-methylheptan-2-one is indeed a consequence of fusing CPR to CYP6DK1 and likely not a natural occurrence under physiological conditions. In addition, 6-methyl-5-hepten-2-ol was produced in all assays, including the negative controls. This chemical was not present in the 6MHO standard, suggesting its production by endogenous microsomal enzyme activity of the Sf9 cells.

I.5 Chapter 5. Mountain Pine Beetle CYP4G55 and CYP4G56 Exhibit Functions Consistent with exo-Brevicomin Biosynthesis

We further investigated the possibility that one of the two MPB P450s in the CYP4G family is involved in exo-brevicomin biosynthesis. This portion of the dissertation contributes to a larger study by the Blomquist lab on the function of this P450 family in MPB. Biosynthesis of exo-brevicomin is thought to proceed through β-oxidation and reduction of long chain fatty acid precursors to produce a ten carbon aldehyde (cis-7-decenal), subsequent cytochrome P450 (P450) mediated decarbonylation of the aldehyde to a nine carbon alkene (cis-3-nonene), followed by a P450 mediated hydroxylation to 6(Z)-non-6-en-2-ol (Blomquist et al., 2010). 6(Z)-Non-6-en-2-ol is oxidized to 6(Z)-non-6-en-2-one by the enzyme 6Z-nonenol dehydrogenase (ZnoDH), and the ketone is then epoxidized to 6,7-epoxy-nonan-2-one by a P450 (Song et al., 2014). Finally, 6,7-epoxy-nonan-2-one is likely cyclized by a cyclase to produce exo-
brevicomin although non-enzyme mediated cyclization is also possible (Figure 1.1C). The decarbonylation step converting cis-7-decenal to cis-3-nonene early in the pathway was hypothesized by Blomquist et al. (2010) in part because the alternative, an α-oxidation of a β-keto-precursor, is highly unusual. This hypothesis predicts that an insect-specific CYP4G enzyme catalyzes the reaction, as CYP4Gs are the only class of enzyme known to catalyze decarbonylation reactions to form hydrocarbons from aldehydes in insects (Qiu et al., 2012). By extension, we hypothesized that one of the two CYP4G P450s found in MPB, CYP4G55 and CYP4G56, catalyze the decarbonylation reaction of cis-7-decenal to cis-3-nonene. Assays of recombinantly expressed CYP4G55-CPR and CYP4G56-CPR fusion proteins for cis-7-decenal decarbonylation activity did in fact show that both produced cis-3-nonene (Figure 5.2) when incubated with 7-decen-1-ol. It is not clear from these experiments which MPB CYP4G is involved in exo-brevicomin biosynthesis in vivo, although both appear to accept longer chain aldehydes and alcohols as substrates and are likely involved in cuticular hydrocarbon biosynthesis (M. MacLean personal communication). Nadeau et al. (2017) (Chapter 3) showed a differential expression pattern based on sex and feeding status consistent with exo-brevicomin biosynthesis (increased expression in unfed males) for a CYP4G56-like protein, but whether this CYP4G56-like protein is in fact CYP4G56 or a yet to be identified CYP4G in MPB is still to be determined. Preliminary qRT-PCR investigations into CYP4G56 also showed a differential expression pattern with increased expression in unfed males whereas CYP4G55 did not (data not shown). Altogether, the work presented in this chapter provides the first evidence of a P450 in the CYP4G family being involved in pheromone biosynthesis.
I.6 Chapter 6. Towards a moderately high-throughput method for measuring cytochrome P450 kinetics and substrate specificities using OxoPlates®

In chapter 6 the suitability the oxygen sensing OxoPlate® 96-well microplate oxygen-sensing system to monitor P450 substrate specificities and enzyme kinetics was tested. Kinetic analyses of P450-catalyzed reactions would be valuable to inform on structure-function relationships, especially in the evolutionary context of resin detoxification and pheromone production in pine bark beetles. Unfortunately, while high throughput methods for measuring P450 kinetics have been developed using radioactive, chromophoric, fluorescent or luminescent substrates or products for particular P450s (Moody et al., 1999; Stresser et al., 2000; Crespi, Miller and Stresser, 2002; Yamamoto, Suzuki and Kohno, 2002; Cali et al., 2006), these substrates are suitable for very few enzymes. The conventional method for measuring P450 kinetics is time-consuming and labor intensive, involving incubating microsomes containing the relevant P450 with different concentrations of substrate, stopping the reaction at various intervals with a chemical inhibitor, and measuring the change in substrate or product concentrations using a method like liquid chromatography–tandem mass spectrometry (Obach and Reed-Hagen, 2002; Baranczewski, Edlund and Postlind, 2006). Measuring product formation in P450 reactions may also be difficult if the products are unknown and proper standards are unavailable (Obach and Reed-Hagen, 2002). Given these limitations, development of a reliable medium or high-throughput method for measuring P450 substrate specificities and kinetics would be extremely useful. Medium-throughput methods for measuring P450 enzyme kinetics through oxygen depletion have been developed using fluorescent
probes (Olry et al., 2007; Zitova et al., 2010; Chang et al., 2012). In this chapter we tested a 96-well microplate system that uses an oxygen sensitive fluorophore very similar to the system used in Olry et al. (2007) to measure oxygen depletion during P450 reactions. The microplate system used in Olry et al. (2007) is no longer commercially available, but an alternative, OxoPlate®, based on the same technology can be purchased from PreSens Precision Sensing GmbH, Regensburg, Germany. We used P450 from *Ips pini*, CYP9T3, which is relatively easy to produce in recombinant form and hydroxylates a broad range of substrates (Song et al., 2013), to test whether OxoPlate® are a viable solution for measuring P450 substrate specificity and enzyme kinetics in a medium-throughput format.

Initial tests showed that increasing either P450 or NADPH concentrations in reactions correlated with increasing oxygen depletion rates when substrate concentrations were kept constant (Figure 6.2). This suggests, as expected, that the reaction rate is limited by the amount of P450 and NADPH present. Thus, the highest tested concentrations of each of these components were chosen for further experiments to ensure adequate resolution when monitoring the reactions. Reactions with CPR alone exhibited oxygen depletion rates very similar to those found when the P450 was present (Figure 6.4), suggesting oxygen consumption reactions in addition to those from the P450 reaction we were attempting to measure. To elucidate whether endogenous P450s in the CPR microsomal fraction were reacting with our substrate, the insect P450 inhibitor PBO was added to reaction mixtures containing just microsomes bearing CPR. Indeed, lower rates of oxygen depletion were observed when PBO was present compared to when it was not, and oxygen depletion rates were lower still if no substrate or the substrate myrcene
were present (Figure 6.5). These data suggest that endogenous (Sf9 cell) P450s in the microsomal fraction react with at least some of our substrates. Higher concentrations of PBO in the reaction mixture were not tested but it is likely that greater inhibition would occur with increasing concentration.

The effects of HF-CPR concentrations on the reaction rates were explored. To minimize oxygen depletion due to endogenous P450 activity, the effects of lowering the concentration of HF-CPR microsomal fraction was tested while keeping the P450 microsomal fraction constant. The rate of oxygen depletion in CP9T3 reactions with (+)-α-pinene and only 50 µL HF-CPR was similar to or greater than rate of when double that amount of HF-CPR microsomal fraction (100 µL) was added (Figure 6.4). At the 50 µL of the HF-CPR microsomal fraction volume in a 550 µL reaction without the P450 microsomal fraction present there was very little observable oxygen depletion (Figure 6.5). This suggested that the oxygen depletion from endogenous sources can be minimized while still attaining high enough activity levels for accurate measurements of the P450 we are testing. Therefore, 50 µL of the HF-CPR microsomal fraction was used for all subsequent assays.

While NADPH and oxygen depletion measurements seem to be quick, less labor intensive and able to be measured in real-time when compared to GC/MS methods, it is unclear that the measurements really correlate with substrate depletion because uncoupled P450 reactions that produce superoxide radicals and hydrogen peroxide would consume O₂ but not the oxidized product. (Ansed and Thakker, 2004; Traylor et al., 2011). To overcome this limitation, Traylor et al. (2011) used superoxide dismutase and catalase to form water from superoxide and hydrogen peroxide, allowing for direct
measurement of oxygen consumption from substrate oxidation alone. In our experiments, catalase increased the rate of oxygen depletion when both the CPR and P450 microsomal fractions were present but did not affect oxygen depletion rates with the CPR microsomal fraction alone (Figure 6.3). This may represent a shift from peroxide formation to other uncoupling reactions that produce superoxide radicals. Further experimentation with additions of catalase and superoxide dismutase are needed to ascertain the effects on measuring P450 activity using the OxoPlate® system as the experiments conducted in this study are insufficient to draw any conclusions.

Screening the substrate specificity of P450s is time consuming using GC/MS, whereas using the OxoPlate® system can significantly decrease that time through the ability to screen dozens of compounds simultaneously in less than an hour. Six potential monoterpene substrates for CYP9T3 were tested using the OxoPlate® and the methods described above. A substrate profile for CYP9T3 using most of the same monoterpenes was conducted previously by Song et al. (2013) using GC/MS and the results were confirmed in this study (Figure 6.6). Both this study and Song et al. (2013) showed that myrcene, 3-carene and (+)-α-pinene were likely substrates while α-phellandrene and γ-terpinene were not. S-(-)-Limonene also showed no oxygen depletion using the OxoPlate® method and is likely not a substrate of CYP9T3. Given this data, the OxoPlate® seems to be a viable method to measure P450 substrate specificities but measuring enzyme kinetics may be unreliable due to the uncoupling reactions mentioned above.

The final assay conducted in this chapter was to use the OxoPlate® method to measure CYP9T3 enzyme kinetics for the substrates myrcene, 3-carene and (+)-α-pinene.
Initial attempts encouragingly produced data amendable to fitting the Michaelis-Menten model (Figure 6.7). The kinetic data from this initial attempt suggests CYP9T3 has a higher affinity for myrcene ($K_m = 0.89 \pm 0.49 \text{ mM}$) and $\alpha$-pinene ($K_m = 1.32 \pm 0.16 \text{ mM}$) than for 3-carene ($K_m = 16.92 \pm 0.94 \text{ mM}$). Unfortunately, subsequent attempts to use the OxoPlate® to measure CYP9T3 kinetics were unsuccessful for unknown reasons. Figure 6.8 shows one of the more recent attempts with CYP9T3 and (+)-$\alpha$-pinene has no discernable pattern of rates of oxygen depletion in response to the amount of substrate added. The reasons why the later experiments did not work are unknown but perhaps the substrates were no longer viable because of oxidation over time in the storage vials or some settings for the plate reader got changed after maintenance was performed. Work on this project was set aside to concentrate on the other experiments presented in this dissertation. Although we had limited success, the research presented in this chapter may lay the groundwork for further refinement of using the OxoPlate® system to monitor P450 reactions.

I.7 Appendix 1. Comparative transcriptomics of the Jeffrey pine beetle midgut and fat body tissues

Appendix 1 describes efforts to identify JPB pheromone biosynthetic enzymes. Similar to MPB, JPB evidently use a combination of pheromones to coordinate mass attacks on new host trees, although the JPB pheromone system is not as well understood. Females produce 1-heptanol, presumably derived by P450 mediated hydroxylation of the abundant heptane produced in Jeffrey pines (Paine et al., 1999). The 1-heptanol, in synergy with host produced heptane, attracts male JPB to newly attacked host trees. Like
MPB, male JPB produce *exo*-brevicomin to attract females (Paine *et al.*, 1999), presumably through the same or a very similar pathway to MPB *exo*-brevicomin biosynthesis. Finally, male JPB also produce a racemic mixture of frontalin which acts like an anti-aggregation pheromone in sufficient quantities (Paine *et al.*, 1999). Frontalin in JPB is also produced via the mevalonate pathway and biosynthesis occurs in the anterior midgut (Hall *et al.*, 2002). Very little research has been published on JPB pheromone biosynthetic pathways so the enzymes involved are mostly unknown. However, because of the close evolutionary relationship between JPB and MPB, many of the enzymes identified in MPB may also be present in JPB and could provide clues into the JPB pheromone biosynthetic pathways. We conducted RNA-seq on fed and unfed male and female JPB midguts and fat bodies to identify candidate genes involved in pheromone biosynthesis. The sequences were aligned to the MPB genome and raw counts have been obtained for putative gene expression following sequence quality control. The data must be normalized before further analysis can be conducted using differential expression and co-expression networks. Although valuable information, this dissertation focused on identifying and functionally characterizing MPB pheromone biosynthetic and detoxification enzymes and leaves analysis of this JPB data as a project for the future.

II. Future Directions

Both Keeling *et al.* (2016) and Chapter 3 of this dissertation (Nadeau *et al.*, 2017) identified a number of potential MPB pheromone biosynthetic genes, only three of which
were examined and functionally characterized in this dissertation (CYP6DE3, CYP6DK1 and CYP4G56). Further functional analyses of the other candidate enzymes (Table 3.4) should be conducted to identify the uncharacterized steps hypothesized to be a part of MPB pheromone biosynthesis. These include the single P450 involved in \textit{trans}-verbenol biosynthesis, the P450 or dioxygenase that converts geranylgeranyldiphosphate to 6-methylhept-5-en-2-one and the isomerase that converts 6-methylhept-5-en-2-one to 6MHO in the proposed frontalin biosynthetic pathway, and the P450 that converts cis-3-nonene to \(6(Z)\)-nonen-2-ol and the cyclase that catalyzes the final step in the proposed \textit{exo}-brevicomin biosynthetic pathway (Figure 3.1). Although CYP6DH1 was identified as the possible P450 involved in \((-\textit{trans})\)-verbenol biosynthesis (Song, 2012; Robert \textit{et al.}, 2013), it has not been found in populations outside of Canada (Tittiger, unpublished observation). Examining the different MPB populations in northern Canada and the western United States for expression of P450s in the CYP6DH family may shed further light on the function and evolutionary history of these enzymes.

In Chapter 4 we hypothesized that CYP6DK1 may have resin detoxification activities in addition to its evident pheromone biosynthetic activity. Functional assays with various monoterpenes would be useful in determining if this additional function is indeed present. Additionally, RNAi knock down in fed males and monitoring frontalin secretion would be useful in confirming CYP6DK1 function as a pheromone biosynthetic P450 in MPB. Finally, kinetic analysis of CYP6DK1 with any identified substrates would be useful in understanding the structure and function of this P450 and determining optimal reaction conditions for possible commercial applications.
The other P450s characterized in this dissertation, (CYP6DH3, CYP6DE3, CYP4G55 and CYP4G56) would also be prime candidates for kinetic analyses coupled with protein modelling to help better understand their function and evolution. Work is also needed to determine which CYP4G is involved in *exo*-brevicomin biosynthesis. Closer examination of expression patterns in fat body tissues and localization within those tissues may help elucidate which CYP4G P450 participates in pheromone biosynthesis.

Further experimentation with the OxoPlate® is needed to ascertain its fitness in measuring kinetic activity of P450s. Previous studies showed that 96-well microplate oxygen-sensing systems are viable for measuring P450 substrate profiles, kinetics and response to inhibitors (Yamamoto, Suzuki and Kohno, 2002; Olry *et al.*, 2007) so further adjustments to the OxoPlate® assay should likewise yield the desired results. Future additions of catalase and superoxide dismutase in all reactions should help alleviate background oxygen consumption caused by the uncoupling of the P450 reaction (Traylor *et al.*, 2011). If the OxoPlate® assay can be accomplished as intended, and the results confirmed by the GC/MS method, it will provide a medium-throughput method for measuring the enzyme kinetics of the P450s identified in this dissertation.

Finally, the data from the comparative transcriptomics study for JPB must be normalized and studied. Analyses similar to those used in the MPB comparative transcriptomics study in Chapter 3 should be used to identify candidate JPB pheromone biosynthetic genes. Furthermore, confirmed and candidate MPB pheromone biosynthetic genes can be used as a starting point when searching for candidate genes in the JPB because of the close evolutionary relationship between these two species. Any JPB
candidate genes identified will then need to undergo functional analysis to confirm involvement in pheromone biosynthesis.

III. References


Song, M., Kim, A. C., Gorzalski, A. J., MacLean, M., Young, S., Ginzel, M. D.,


United States Department of Agriculture, F. S. (2017) *Major forest insect and disease conditions in the united states: 2015*, Forest Health Protection. Available at:


Appendix A

Comparative transcriptomics of Jeffrey pine beetle midgut and fat body tissues

I. Abstract

Bark beetles are highly destructive pests in western North America. One bark beetle, *Dendroctonus jeffreyi* (Jeffrey pine beetle, JPB) infests exclusively Jeffrey pine trees (*Pinus jeffreyi*). While it is established that JPB use a combination of pheromones to coordinate mass attacks on new host trees, the JPB pheromone system is not well understood. Females produce 1-heptanol, presumably derived from a P450 mediated hydroxylation of abundant heptane in Jeffrey pines. The 1-heptanol, in synergy with host produced heptane, attracts male JPB to newly attacked host trees. Male JPB produce *exo*-brevicomin to attract females and also produce a racemic mixture of frontalin which acts like an anti-aggregation pheromone in sufficient quantities. Very little research has occurred on JPB pheromone biosynthetic pathways so the enzymes involved are mostly unknown. We used a comparative RNA-Seq analysis between fed and unfed male and female JPB midguts and fat bodies to identify candidate genes involved in pheromone
biosynthesis. The sequences obtained have been subjected to quality control and aligned to the genome of *Dendroctonus ponderosae* Hopkins (mountain pine beetle, MPB), a closely related species to JPB. The data will be analyzed in the future to identify JPB candidate pheromone biosynthetic genes.

II. Introduction

Bark beetles have devastated forests of western North America in recent years, killing over forty million acres of trees in the past 15 years in the western United States (United States Department of Agriculture, 2017). *Dendroctonus* and *Ips* spp. are the main culprits in recent outbreaks and have a native range from Alaska to portions of Mexico (Williams and Liebhold, 2002). JPB infest exclusively Jeffrey pine trees and spend the majority of their life cycles sequestered beneath the bark of their host trees, only leaving the tree for a few summer days once a year to migrate to a new host (Smith *et al.*, 2009). JPB is a closely related species to MPB, a bark beetle that attacks a much broader range of pine tree species. Tree mortality from MPB and JPB beetle attacks is a result both of the beetles feeding on the phloem of the tree, and associated fungal species that are introduced by the bark beetles, the growth of which results in decreased water flow and a blue staining of the sapwood (Gibson *et al.*, 2009; Smith *et al.*, 2009).

Targeted coniferous trees have been especially impacted in recent years as a result of persistent drought conditions (Williams and Liebhold, 2002; Negrón *et al.*, 2009). Normally, healthy trees with abundant water possess inherent natural defenses, including the production of copious amounts of monoterpane-laden resin, that expel invading
beetles and inhibit beetle survival after infestation (Boone et al., 2011). However, drought conditions significantly impede resin production, and bark beetles have thrived as a result.

Because bark beetles spend most of their life cycles beneath the bark of infested trees they are largely unaffected by insecticide application. Large scale management attempts have been almost solely restricted to standard silvicultural practices such as harvesting trees that have been invaded by bark beetles, but have produced very limited success given the massive scale of the problem (Fettig et al., 2007). However, employing pheromones to bait trees treated with pesticides has been marginally effective in protecting stands of trees (Wermelinger, 2004). Therefore, the right combination of silvicultural practices, pesticide application, and pheromone baiting may increase the efficiency of bark beetle management given an ample supply of pheromones.

Similar to MPB, JPB evidently use a combination of pheromones to coordinate mass attacks on new host trees, although the JPB pheromone system is not well understood. Females produce 1-heptanol, presumably derived from a P450 mediated hydroxylation of abundant heptane in Jeffrey pines (Paine et al., 1999). The 1-heptanol, in synergy with host produced heptane, attracts male JPB to newly attacked host trees. Like MPB, male JPB produce exo-brevicomin to attract females (Paine et al., 1999), presumably through the same or a very similar pathway to MPB exo-brevicomin biosynthesis. Finally, male JPB also produce a racemic mixture of frontalin which acts like an anti-aggregation pheromone in sufficient quantities (Paine et al., 1999). Frontalin in JPB is also produced via the mevalonate pathway and biosynthesis occurs in the anterior midgut (Hall et al., 2002). Very little research has occurred on JPB pheromone
biosynthetic pathways so the enzymes involved are mostly unknown. However, because of the close evolutionary relationship between JPB and MPB, many of the enzymes identified in MPB may also be present in JPB and could provide clues into the JPB pheromone biosynthetic pathways.

Differences in mRNA levels based on gender and/or feeding status can indicate involvement in pheromone biosynthesis in bark beetles (Blomquist et al., 2010; Robert et al., 2013; Song et al., 2013; Keeling et al., 2016; Nadeau et al., 2017). Here we conducted a comparative transcriptomics study measuring mRNA levels in male and female, fed or unfed JPB using RNAseq. Future analysis of this data should provide insights into the enzymes involved in JPB pheromone biosynthesis and the evolutionary relatedness of the pheromone biosynthetic pathways in JPB and MPB. Although valuable information, this dissertation focused on identifying and functionally characterizing MPB pheromone biosynthetic and detoxification enzymes and leaves analysis of this JPB data as a project for the future.

III. Methods

III.1 Tissue Collection

Sections of JPB-attacked Jeffrey pine were collected from Truckee, CA, USA, near the Sagehen Creek Field Station (approx. 39° 25′ 57″ N, 120° 14′ 13″ W) on September 30, 2013. The beetles overwintered in the bolts and emerging adult beetles were collected and sexed as reported previously (Aw et al., 2010; Song et al., 2014). For feeding experiments, fresh Jeffrey pine bolts were obtained from Truckee, CA, USA, near the
Sagehen Creek Field Station and stored at 4°C prior to use. Females were fed by drilling small holes just beneath the bark, inserting the beetles head first, stapling a wire mesh over the occupied hole, incubating the bolt vertically for 24 h in the dark at room temperature and collecting the live beetles by stripping the bark. Fresh frass indicated that the beetles had fed. Males were fed using the same method except females were first placed head first into the holes for 24 h, followed by insertion of the males head first for 24 h and subsequent collection and sexing of the live beetles. Unfed beetles were incubated in 2 oz plastic cups with perforated lids in a dark drawer kept humid with small flasks filled with water and a paper towel for 24 h. All beetles were dissected following treatment to collect midgut and fat body tissue. Three replicates of pooled tissue from 10 beetles were collected for each of the four treatments (fed and unfed males and females) for a total of 12 samples. Midguts and fat bodies were immediately frozen in liquid nitrogen and stored at -80°C for subsequent RNA extraction.

### III.2 RNA Extraction and Sequencing

Total RNA was isolated from the midguts and fat bodies pooled from 10 beetles per sample using an RNeasy Plant Mini Kit from Qiagen (Valencia, CA) and treated with RNase-Free DNase Set from Qiagen as described by the supplier’s manual. RNA was quantified using Quant-iT™ RiboGreen® reagent and a Labsystems Fluoroskan Ascent fluorescence plate reader. RNA integrity for each sample was determined using an Agilent 2100 Bioanalyzer and a Eukaryote Total RNA Nano Series II chip. Only samples with an RNA Integrity Number of eight or higher were used for sequencing (Fleige and Pfaffl, 2006). RNA was precipitated and provided to the Georgia Genomics Facility
GGF confirmed the quality of the total RNA using an Agilent 2100 Bioanalyzer, prepared barcoded cDNA libraries using a Kapa Stranded mRNA-Seq Kit (Wilmington, MA), and sequenced them on the Illumina NextSeq 500 using paired-end sequencing with a NextSeq 2x75 High Output Flow Cell.

III.3 Sequence Quality Control

Sequence quality for each sample was characterized using FastQC (v. 0.11.2; http://www.bioinformatics.babraham.ac.uk/projects/fastqc/). Sequence pairs were trimmed and filtered for nucleotide-base quality and Illumina adapter sequences using Cutadapt v. 1.8.1 (Martin, 2011), with options set as follows: trimming of low-quality (phred quality ≤ 10) and “N” base calls from both ends of each read and removal of sequences with trimmed length < 35 nt.

III.4 Sequence Alignment and Expression Quantification

Before assembly, trimmed sequence pairs were compared to one another using the MaSuRCa `superreads.pl` script, and if found to intersect (minimum k-mer 41), were combined into single “super-reads” (Zimin et al., 2013). Read pairs with no intersection were retained as separate paired-end reads. Single reads and read pairs were aligned to the Ensembl Metazoa (release-25; Kersey et al., 2015) D. ponderosae reference genome (Keeling et al., 2013) using the HISAT spliced read alignment tool (v. 0.1.6-beta; Kim et al., 2015). The coordinates of each known Dendroctonus gene and its exons were extracted from the Ensembl Metazoa Gene Transfer Format (GTF) file and supplied to HISAT at time of alignment (via HISAT’s `--known-splicesite-infile` option). Resultant
alignments were compressed from the sequence alignment/map (SAM) format to the binary BAM format (Li et al., 2009). Upon alignment, the raw counts of reads and read pairs aligned to each gene were totaled using the featureCounts tool of the subread package (v. 1.4.6-p1; Liao et al., 2014). Reads were counted once per pair and summarized for gene loci, with only read pairs aligned to a unique transcriped location included in the count totals.

III.5 Annotation

*Dendroctonus ponderosae* gene descriptors and annotated Interpro protein domains (Mitchell et al., 2014) were obtained from Ensembl Metazoa, via the BioMart interface (Kinsella et al., 2011). Interpro2GO file (v. 2016/03/19 11:04:26) was used to map Interpro IDs to gene ontology (GO) terms.

IV. Results and Future Directions

The number of quality reads for each treatment ranged from 14,587,308 to 32,228,511 with an average of 20,680,036 and standard deviation of 4,958,472 (Table 1). No valid comparisons between treatments were conducted because the data has not been normalized. Once normalized, principle component and co-expression network analyses may be conducted to analyze the effects of the treatments on gene expression. MPB pheromone biosynthesis candidate genes from the Nadeau et al. (2017) study may be used as a starting point because the two species are closely related and likely have many of the same genes with similar functions. Analysis of this data set will yield JPB
candidate pheromone biosynthetic genes and lead to a better understanding of the evolutionary relationship between MPB and JPB.

V. References


### Table A.1. Summary statistics for JPB RNAseq reads subjected to quality control.

<table>
<thead>
<tr>
<th>Treatment</th>
<th>Replicate</th>
<th>Quality Reads</th>
</tr>
</thead>
<tbody>
<tr>
<td>FedFemale</td>
<td>1</td>
<td>23,117,598</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>16,087,426</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>23,032,683</td>
</tr>
<tr>
<td>FedMale</td>
<td>1</td>
<td>18,860,742</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>25,186,531</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>22,424,725</td>
</tr>
<tr>
<td>UnfedFemale</td>
<td>1</td>
<td>32,228,511</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>19,022,065</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>20,958,071</td>
</tr>
<tr>
<td>UnfedMale</td>
<td>1</td>
<td>14,587,308</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>16,411,118</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>16,243,659</td>
</tr>
</tbody>
</table>

Average 20,680,036  
Std. Dev. 4,958,472