DISCOVERY AND CHARACTERIZATION OF OLFACTORY-RELATED GENES IN THE WHEAT STEM SAWFLY, *CEPHUS CINCTUS*, A MAJOR PEST OF WHEAT IN THE NORTHERN PLAINS

by

Joanna Christine Gress

A dissertation submitted in partial fulfillment
of the requirements for the degree

of

Doctor of Philosophy

in

Plant Science

MONTANA STATE UNIVERSITY
Bozeman, Montana

November 2014

DEDICATION

I dedicate this work to Bruno, my faithful companion who has faithfully walked by my side through this journey from the beginning, and to my Mom for her unending support and love.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank everyone who has helped me through this scientific endeavor; it has been a long but fruitful journey. First, thank you to my advisor, Dr. Kevin W. Wanner, who did come through for me in the end! Thank you for your guidance, advice, and support. To my committee members, Drs. David Weaver, Aurélien Mazurie, Andreas Fischer, and Kevin O'Neill, thank you for all your advice.

I would like to thank the Weaver laboratory, especially Megan Hofland. I would also like to thank the Wanner laboratory. I also want to thank my student assistants, Adrienne Brown and Tember Dykgreve, for all their hard work and dedication!

I also want to acknowledge the Montana Wheat and Barley Committee and the Montana Agricultural Experiment Station for funding support. USDA grant 2008-35302-18815 (Hugh Robertson) and the MSU Ag Stimulus Committee (Kevin Wanner and David Weaver) provided funding to sequence and assemble the wheat stem sawfly genome. Thank you to Dr. Charles Carey of the Bioinformatics Core and the Molecular BioSciences Program and Stephanie Cunningham for their support. I also want to thank the Department of Plant Science and Plant Pathology for all their help in the different aspects of my work. A special thank-you to Irene Decker and Tamara Parnell!

A grateful thank-you goes to Kate McInnerney; your mentorship, time, patience, knowledge, and friendship made Montana State University a much more special place. Also, a special thank-you to Dr. Amy Thomas for her support and friendship.

Thank you to my family, I could not have accomplished this without your love and support. I also want to thank Dr. Duran for all of his support through this journey.

VITA

Joanna Christine Gress
Born 1979 Los Angeles, California
Parentage: Rodney M. Lackey and Mary M. Gress

Education

May, 2003. Mills College, Oakland, California B.A. Biology
July 2007. The University of Arizona, Tucson, Arizona. M.S. Plant Science Master's Thesis: A Multidisciplinary Approach for the Control of the Citrus Nematode, *Tylenchulus semipenetrans* in Arizona. Implications in Diagnostics and Management Strategies.
August 2008-present. Montana State University, Bozeman, Montana. Department of Plant Sciences and Plant Pathology. PhD Candidate. Dissertation: Discovery and Characterization of Olfactory-related genes in the Wheat stem sawfly *Cephus cinctus*: A Major Pest of Wheat in the Northern Plains.

Awards and Fellowships

May 1998. College of Santa Fe, New Mexico, Outstanding General Chemistry Student Award.
Summer 1998. Howard Hughes Undergraduate Research Fellow, University of Illinois, Urbana-Champaign.
Summer 1999. NSF Undergraduate Research Fellow, Michigan State University, East Lansing, Michigan.
Summer 2002. Barret undergraduate Research Fellow, Mills College, Oakland, California.
May 2003. Caryl Haskins Prize for Distinction in Life Sciences, Mills College, Oakland, California.
2003-2005. IGERT Fellowship in Evolutionary, Functional, and Computational Genomics. University of Arizona, Tucson, Arizona.
2006. ARCS Foundation Award for Outstanding Graduate Research. University of Arizona, Tucson, Arizona.
2007. ARCS Foundation Award for Outstanding Graduate Research. University of Arizona, Tucson, Arizona.
2008-present. Molecular BioScience Graduate Fellowship. Montana State University Bozeman, Montana.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF FIGURES – CONTINUED

ABSTRACT

The wheat stem sawfly (WSS), *Cephus cinctus* (Hymenoptera: Cephidae), is one of the most important insect pests of wheat in the northern Great Plains region of the United States and Canada, with economic losses exceeding $100 million per year. Traditional pest management strategies including pesticides are generally unsuccessful due to an extended adult flight time and the inaccessible larval stage that feeds within the wheat stem. Research towards integrated pest management strategies based on olfaction has proved promising. However, little is known about the molecular basis of olfaction in this important insect pest. We have identified and annotated 131 members of the major olfactory-related gene families from antennal transcriptome and whole genome sequences, including: 6 odorant binding proteins (OBP), 8 chemosensory proteins (CSP), 53 odorant receptors (OR), 14 ionotropic receptors (IR), 12 carboxylesterases (CCE), 8 gluthatione S-transferases (GST), and 29 cytochrome P450s (P450). Expression levels in the antennae, sawfly bodies, and whole larvae were analyzed using RNA-seq. Gene expression results were used to identify candidate genes for further functional characterization based on higher enriched expression in antennae and/or sex-biased expression in the antennae. These candidate WSS olfactory genes may mediate important pest behaviors and serve as molecular targets for future insect management strategies.

CHAPTER ONE

GENERAL INTRODUCTION

The wheat stem sawfly, *Cephus cinctus* Norton (Hymenoptera: Cephidae) (WSS), is considered one of the most important insect pests of wheat (*Triticum aestivum* L.) grown in Montana and the northern Great Plains region (Morrill et al. 1994; Morrill 1997), causing an estimated $100 million per year in economic damage (Wahl et al. 2007). The WSS was first reported from native *Agropyron* and *Elymus* large-stemmed prairie grasses (Ainslie 1929) and is believed to have expanded its host range to include wheat more than a century ago, when wheat was introduced to the region as a crop (Ainslie 1920; Munro 1945). WSS is distributed throughout most of North America (Ivie, 2001) and, although the majority of damage has been restricted to the Northern Great Plains region, it has spread more recently to Colorado and Nebraska (Irell & Peairs, 2011; Bradshaw, 2013).

The insect sense of olfaction mediates many important pest behaviors. Insects locate host plants, mates, and oviposition sites from a distance based on volatile odors produced at the source. Odors in an insect's environment are detected by the peripheral olfactory system, specialized olfactory neurons housed in sensilla on olfactory organs. The dendrites of olfactory neurons extend into the hollow hair-like sensilla that are filled with aqueous lymph. Olfactory-related proteins in lymph transport and degrade odors while transmembrane receptors on the dendrite detect the odors and initiate nerve impulses. The major olfactory-related proteins involved in the detection of odorants by

insect antennae are odorant binding proteins (OBP), chemosensory proteins (CSP), odorant receptors (OR), ionotropic receptors (IR), and odorant-degrading enzymes (ODE). OBPs contribute to the sensitivity of the olfactory system by transporting pheromones and odorants across the sensilla lymph. They help to solubilize ligands and transport hydrophobic molecules (Leal, 2005). CSPs are thought to be involved in transporting odorants, similar to the OBPs, shuttling aliphatic compounds, esters, and long chain pheromonal compounds through the aqueous sensilla, but much less supporting evidence exists. ORs are a large diverse gene family primarily responsible for the molecular recognition of odors in the insect's environment (reviewed in Rutzler & Zwiebel, 2005; Vosshall & Stocker, 2007; and Leal, 2012). IRs are a family of glutamate ionotropic receptors first identified in *Drosophila* (Benton et al., 2009) that are expressed in a combinatorial way in the sensory neurons that do not express ORs and have been shown to respond to odorants containing amines, carboxylic acids, and aldehydes (Rytz et al., 2012). ODEs function to help rapidly deactivate the chemical signal, once the message has been received, by degrading the odorant (Ishida & Leal, 2005).

The goal of this thesis was to use sequencing technology to increase the molecular-genetic knowledge of this important pest species. Specifically, olfactory-related genes were identified from genome and transcriptome sequences, annotated and characterized, and their expression patterns assayed. The ultimate goal of this research is to identify new molecular targets that that can be used to develop new insect pest management strategies that exploit olfactory behaviors.

References

Ainslie, C. N. 1920. The western grass-stem sawfly. USDA. Bull. No. 841.

Ainslie, C. N. 1929. The western grass-stem sawfly - a pest of small grains. USDA. Bull. No. 157.

Benton, R., Vannice, K.S., Gomez-Diaz, C., Vosshall, L.B. 2009. Variant ionotropic glutamate receptors as chemosensory receptors in Drosophila. Cell 136(1): 149–162.

Bradshaw, J. 2013. Adult Wheat Stem Sawflies Have Begun to Emerge. University of Nebraska *Cropwatch*. 30 May 2013.

Ishida, Y. and Leal, W.S. 2005. Rapid inactivation of a moth pheromone. Proc Natl Acad Sci USA 102: 14075-9.

Leal, W.S. 2005. Pheromone reception. Top. Curr. Commun. **335**: 1044-1050.

Rutzler, M. and Zwiebel, L. J. 2005. Molecular biology of insect olfaction: recent progress and conceptual models. J Comp Physiol A Neuroethol Sens Neural Behav Physiol 191: 777–790.

Leal, W.S. 2012. Odorant Reception in Insects: Roles of Receptors, Binding Proteins, and Degrading Enzymes. Annu Rev Entomol 58: 373–391.

Morrill, W. L., and Kushnak, G. D. 1999. Planting date influence on the wheat stem sawfly (Hymenoptera: Cephidae) in spring wheat. J Agric Entomol 16: 123–128.

Morrill, W. L., Kushnak, G. D. Bruckner, P. L, and Gabor, J. W. 1994. Wheat stem sawfly damage, rates of parasitism, and overwintering survival in resistant wheat lines. J Econ Entomol 87: 1373–1376.

Munro, J. A. 1945. The Wheat Stem Sawfly and Harvest Loss. NDAES Bimonthly Bulletin. 7: 12–16.

Ivie, M. A. 2001. On the Geographic Origin of the Wheat Stem Sawfly (Hymenoptera: Cephidae): A New Hypothesis of Introduction from Northeastern Asia. Am Entomol 47: 84–97.

Irell, B. and Peairs, F. 2011. Wheat Stem Sawfly: A New Pest of Colorado Wheat. http://www.ext.colostate.edu/pubs/insect/05612.html

Rytz, R., Croset, V. and Benton, R. 2013. Ionotropic Receptors (IRs): Chemosensory ionotropic glutamate receptors in *Drosophila* and beyond. Insect Biochem Mol Biol 43: 888–897.

Vosshall, L.B. and Stocker, R. F. 2007. Molecular architecture of smell and taste in Drosophila. Annu Rev Neurosci 30: 505–533.

Wahl, D. V., Shanower, T. G., and Hoelmer, K. A. 2007. A new species of *Collyria* Schiødte (Hymenoptera: Ichneumonidae: Collyriinae), a parasitoid of *Cephus fumipennis* (Hymenoptera: Cephidae) in China, and potential biological control agent for *Cephus cinctus* in North America. J Kans Entomol Soc 80: 43–50.

CHAPTER TWO

A BRIEF INTRODUCTION OF THE WHEAT STEM SAWFLY AND THE OLFACTORY SYSTEM

The wheat stem sawfly (WSS), *Cephus cinctus* Norton (Hymenoptera: Cephidae), has been recognized as one of the most important insect pests of wheat, *Triticum aestivum* L. (Cyperales: Poaceae), in the northern Great Plains of the United States and Canada for almost a century (Weiss and Morrill, 1992). Its status as one of the most damaging insect pests of wheat grown in the Northern Great Plains region results from its persistent and widespread occurrence each year at economic levels (Morrill et al., 1994; Morrill, 1997). In recent years, infestation levels in both winter wheat and spring wheat have approached 100% in some fields (Morrill et al., 1994). This has caused losses of more than $100 million per year in the U.S. and Canada (Wahl et al. 2007).

The WSS currently infests wheat in North America, in a range from Pennsylvania to Georgia to the Pacific Ocean and from the Peace River District of Alberta to Texas (Ivie 2001). The WSS first occurred in the native *Agropyron* and *Elymus* large-stemmed prairie grasses (Ainslie 1929). Then, with the introduction of wheat and small grains to this area about a century ago, the WSS was able to expand its host range and the first infestations on wheat were reported (Ainslie, 1929; Munro, 1945). Most economic damage occurs in the Northern Great Plains region, with recent reports of economic damage to winter wheat in Colorado and Nebraska (Irell and Peairs 2011, Bradshaw 2013).

Taxonomic Classification of the wheat stem sawfly:

Phylum: Arthopoda

Class: Hexapoda (Insecta)

Order: Hymenoptera

Suborder: Cephoidea

Family: Cephidae

Genus: *Cephus*

Species: *Cephus cinctus* Norton

The larvae overwinter in the below-ground portion of the stem with adults emerging in the late spring during crop growth. Males emerge earlier and, after mating, females begin searching for suitable host plants for oviposition (Ainslie, 1920; Weiss and Morill, 1992). Females usually lay one egg within a wheat stem per visit in the uppermost developing internode (Holmes and Peterson, 1960). The larvae feed internally on the parenchyma and vascular tissue, damaging the stem and reducing seed set and kernel weight (Holmes 1977). As the wheat crop matures, the larvae move down to the base of the stem where it cuts a notch at ground level causing the plant to fall over before harvest. The stem is then plugged with frass and the prepupa forms a cocoon within the stubble in preparation for overwintering (Criddle 1923).

Conventional management strategies do not provide adequate control of this insect pest. Planting-tolerant varieties that have solid rather than hollow stems represents the most common tactic, but its success has been limited because of inconsistent performance and reduced yield in the absence of pest pressure (Weaver et al., 2004).

Insecticides generally are ineffective because the immature stages feed within the stem and are not exposed to foliar sprays, and systemic insecticidal seed treatments do not persist long enough. An extended adult flight period, often lasting for more than four weeks, makes foliar sprays targeting the adult stage unfeasible (Beres et al., 2011). In response, integrated pest management (IPM) strategies based on insect behavior and chemical ecology have been developed. Recent studies have suggested that the release of volatile compounds, including the green leaf volatiles (Z)-3-hexenyl acetate, (Z)-3-hexenol, and 6-methyl-5-hepten-2-one from host plants, provide cues for female oviposition (Piesik et al., 2008). The pheromone 9-acetyloxynonanal has also been identified from both sexes of the WSS and demonstrated to be attractive in field bioassays (Cossé et al., 2002). Attractive odors can be used in IPM programs for monitoring population densities and adult emergence periods as well as for trap crop strategies.

Biology of Wheat Stem Sawfly

Adult

Wheat stem sawflies are small, typically 8-13mm in length, with a narrow and elongated abdomen. They have a slightly compressed, black body with three characteristic yellow bands. Their wings are smoke-colored and the legs are yellow. Females are noticeably larger than males and have a long, saw-like ovipositor. Both male and female adults are relatively weak fliers and perform a particular hovering flight windward of a grass plant without alighting (Ainsile, 1920). The males are on the wing

much more than the females, but neither sex will remain in the air while the wind is strong or when it is cool. Adults typically do not feed and usually fly to the nearest host plant (Sing 2002). When resting on a plant stem, adults typically orient with the head downward and legs aligned with their body (Ainslie, 1920; Criddle, 1922). The WSS is univoltine and has only one generation per year. They display haplodiploid sex determination (arrhenotoky) where males are haploid with 9 chromosomes and females are diploid with 18 chromosomes. Unmated females produce male progeny and mated females can produce both male and female progeny (Holmes, 1978).

Adult WSS typically live for five to eight days after emerging in late May to early June, and emergence usually occurs over a three- to four-week period (Munro, 1945). Adult males emerge first (protandry) to ensure that most of the early emerging females will be mated and will be able to produce female larvae. At the end of the emergence period, more eggs are unfertilized producing male larvae. The emerging sawflies are sexually mature and ready for copulation and oviposition. Copulation is brief, lasting less than one minute, during which the male holds onto the back of the female with his abdomen curved down and under the female ovipositor. Females typically lay their eggs several hours after copulation.

Oviposition

Female WSS oviposit into wheat stems during early summer, but environmental conditions can be an important factor influencing the seasonal timing of emergence and oviposition. Temperature, humidity, light, wind, and host plant can all play a role in determining when and where females deposit their eggs. Females usually lay 30-50 eggs

in a lifetime (Criddle, 1922). A female reportedly lays only one egg per stem; however other females can oviposit in the same stem (Nansen et al., 2005a, 2005b; Perez et al., 2006).

Females prefer to lay their egg in the more tender part of the internodes of wheat stems (Holmes and Peterson, 1960). Early in the sawfly emergence period, females select the elongating stems for oviposition and, later in the emergence period, they select the less mature stems to oviposit into. Typically, females oviposit in the second-to-last developing internode and, as the stem grows, eggs are laid in progressively higher internodes. Stem diameter is also an important selection factor for female oviposition. Large stems are preferred over more slender stems (Ainsle, 1920; Morill et al., 1992).

Eggs

Sawfly eggs range from 1.00 to 1.25 mm in length and 0.33 to 0.42 mm in width. The egg is oviposited freely inside the stem either in a stem cavity or in a hollow space produced by the ovipositor of the female. The newly laid egg is crescent-shaped and milky white in color. Usually the eggs hatch in the lumen of the wheat stem within seven days after oviposition, and the larva feeds on the parenchyma tissue (Ainslie, 1920).

Larva

The larvae are initially transparent in color until they begin to feed. After feeding, they become yellow-green in color. Larvae have a pale brown head with dark brown, pointed mandibles and no legs. Larvae are 8 to 14 mm in length and 1 to 2 mm in diameter.

After emerging from the egg, the larva immediately feeds on parenchyma and vascular tissue inside the stem, forming a gallery above and below the space where the egg was deposited. Larvae generally move up the stem feeding on plant tissue, and will cannibalize other larvae and eggs they encounter, typically leaving one individual in each stem by plant maturity (Seamans et al., 1944).

As the wheat plant senesces around late July or August, the larvae move to the lower part of the stem, girdling it near the soil surface. The transmission of visible and infrared light through the stem wall at desiccation is a trigger for the downward movement (Holmes, 1978, 1982). When the larvae reach ground level, they prepare for diapause. First, they cut a V-shaped groove or notch around the inside of the stem, slightly above or below ground level (Ruynon et al., 2002). This cut does not completely sever the stem, but weakens it and allows the wind to break it or blow it over when the plant desiccates. Immediately after the notch is made, the larva forms a plug from frass and other plant material. The stem usually breaks at the notch forming a cut stem or 'stub' that serves as an overwintering chamber. The larva spins a cellophane-like hibernaculum here, where it remains until it emerges as an adult (Ainslie, 1920; Farstad et al., 1949). Obligatory diapause is at least 90 days, and the breaking of diapause will depend on warming weather conditions. Larvae go through five instars. In the spring, the larva becomes active and goes through its pre-pupal and pupal stages before emerging as an adult from the stub (Munro, 1945).

Pupa

The wheat stem sawfly pupa is exarate, initially milky-white in color, and slender, with an average length of 12 mm and diameter of 1.5 mm. After the first day of development, the legs and body begin to turn darker until reaching a black color. The pupal stage lasts 7 to 14 days and, during the initial few days, the pupa is motionless, after which it begins to move within the overwintering chamber (Ainslie, 1920). After completing the pupal stage, the pharate adult emerges and pushes its way through the frass plug, eventually emerging out of the cut stem.

Host Plants

The WSS feeds on and develops within grass hosts, particularly large-stem grasses such as *Agropyron* spp., *Lolium* spp. (rye grasses), *Phleum pratense* L. (timothy grass) and *Bromus inermis* L. (smooth brome) (Ainsle, 1920; Criddle, 1922). The first severe infestation of wheat was recorded in 1922 in western Canada (Criddle, 1923). Criddle (1923) described the substantial damage and economic loss caused by the WSS to wheat as a result of the elimination of natural population checks such as limited host plants and lack of natural enemies after it adapted to the shift in host preference from native grasses to wheat. The development of wheat rust resistance varieties that promoted wheat acreage and drought conditions in the late 1930s led to sustained increases in wheat stem sawfly populations in the northern Great Plains (Morrill, 1983). Currently, extensive damage to wheat caused by *C. cinctus* persists throughout the Northern Great Plains. Additionally, *C. cinctus* has expanded its host range, feeding on nearly all

cultivated grains and many wild grasses with the exception of oat (*Avena sativa* L.) that appears to be resistant to the sawfly.

Pest Management Practices

The biology and lifecycle of WSS make it difficult to control using conventional practices. Currently, there is no single method of control to effectively reduce WSS populations in wheat (Morrill et al., 2001). Partial control has been obtained through a combination of solid stem varieties, planting delays, swathing, and biological control.

Chemical Control

Chemical Control is not effective for the wheat stem sawfly. *Cephus cinctus* has a short flight period and a long adult emergence period and the eggs, larvae, and pupae are enclosed within the stem of the wheat plant, making modern pesticides ineffective (Runyon, 2001). Several studies have looked at the efficacy of insecticides as seed treatments or as foliar applications, including heptachor, lorsban (chlorpyrifos), Furadan (carbofuran), and Warrior (lambdacyhalothrin) (Wallace, 1962; Skoog and Wallace, 1963; Blodgett et al., 1996), but none were found to be effective. Currently, there are no insecticides labeled for WSS management in Montana.

Cultural Control

Plant Resistance: Planting tolerant varieties that have solid rather than hollow stems represents the most common tactic for control of WSS damage, but its success has been limited because of inconsistent performance and reduced yield in the absence of

pest pressure (Weaver et al., 2004). Some wheat varieties develop different amounts of pith in their lumen, which makes them less susceptible to damage by *C. cinctus* (Farstad, 1940). Solid-stem cultivars in the Canadian Red Western Spring Wheat class are ‘AC Eatonia,’ ‘AC Abbey,’ and ‘AC Lillian’ (DePauw et al., 1994, 2000, 2005). Solid-stem spring wheat cultivars for Montana include ‘Fortuna,’ ‘Lew,’ and ‘Choteau’ (McBride, 1989). Semi-solid varieties are also available and include ‘Ambion,’ ‘Glenman,’ ‘Conan,’ and ‘Scholar’ cultivars, but these are not as favored owing to their significant decrease in efficacy when facing high sawfly pressure. Solid stem pith provides a mechanical resistance to larval feeding, where the small larvae die from desiccation or starvation due to the physical impediment (Wallace and McNeal, 1966). However, these varieties have some drawbacks including reduced yield and lower grain quality and reduced disease resistance (Weiss and Morrill, 1992). Further, it has been determined that the amount of pith produced in the stems of solid wheat varieties is dependent on weather conditions. Wet weather and cloudy days result in less pith and more hollow stems and reduced resistance whereas dry and sunny weather contributes to more pith and more solid stems with increased resistance to sawfly infestation (Roemhild, 1954).

Planting Delays: Delayed planting of spring wheat can reduce sawfly infestation by breaking the synchrony between host plant stem elongation and WSS emergence and oviposition. However, the practice of delayed planting has consequences of making less efficient use of critical soil moisture, which can result in a lower yield (Morrill and Kushnak, 1996).

Swathing: Actions taken prior to harvest in infested fields can help to reduce the severity of yield loss from cut stems. Swathing heavily infested fields can help ensure stems are collected in a windrow before stem cutting by the larvae (Criddle, 1915). Goosey (1999) reported that swathing helped reduce stem cutting by 23% to 33% in fields at 41% to 48% grain moisture with no reduction in crop value. Swathing does not affect sawfly survival as most larvae have already traveled to the base of the plant to prepare for overwintering.

Tillage: This practice has been evaluated extensively as the soil around the infested wheat stub protects the hibernating larva against extreme temperature and water loss. Tillage involves exposing the sawfly-cut stems to the surface of the soil, to conditions that result in death of the larvae either by freezing or desiccation (Callenbauch and Hansmeier, 1944). Weiss et al. (1987) reported 90% sawfly mortality in North Dakota using fall tillage. McBride et al. (1989) showed 25% mortality and Holmes (1978) showed 33-70% mortality using spring tillage. Callenbach and Hansmeier (1944) reported between 35-44% mortality for fall tillage. The results of these trials demonstrated that, although there was mortality to overwintering larvae from tillage, mortality levels were not high enough to effectively control the sawfly populations. Also, no-tillage practices have become widely adopted by growers to reduce soil erosion and water loss (Quisenberry et al., 2000; Cochran et al., 2006).

Biological Control

Biological control of the wheat stem sawfly has been under investigation for many years. There are nine species of Hymenoptera that parasitize *C. cinctus* (Morrill et al., 1998; Meers, 2005). The parasitoids have been slower in their shift in host preference to wheat in comparison to the sawfly. Only two of the nine parasitoids have been recorded in *C. cinctus* populations in wheat (Morrill et al., 1998): *Bracon cephi* (Gahan) and *Bracon lissogaster* (Muesubeck). Criddle (1923) suggested that *B. cephi*, a sympatric idiobiont ectoparastoid (Runyon et al., 2001) has potential for wheat stem sawfly control due to its high levels of larval parasitism that can reach up to 85% in grasses. Emergence of the first generation of *B. cephi* and of sawfly adults is almost synchronous, but the second generation of parasitoids emerges in August when early harvest can negatively impact their populations.

Bracon cephi has become the most important parasitoid of *C. cinctus* in Canada (Nelson and Farstad, 1953) and North Dakota (Meers, 2005). The female locates a sawfly host larva by traversing a stem and, after sensing the presence of a host larva, she inserts her ovipositor. The ovipositor is used to inject venom that immobilizes the larva as the egg is deposited. Upon hatching, the parasitoid larva searches for and attaches itself to the paralyzed *C. cinctus* larva and feeds on it. The host is consumed within 10 days, after which the fully developed parasitoid larva spins a cylindrical cocoon to enter diapause. The adults of the second generation emerge in August by chewing a circular hole through the stem (Nelson and Farstad, 1953). *Bracon lissogaster* is the second major parasitoid of *C. cinctus* in wheat (Meers, 2005). Its lifecycle is similar to that of *B. cephi*, but it can

complete a second generation in late fall more readily, an ability attributed to the immediate oviposition of females directly after emergence (Somsen and Luginbill, 1956).

Plant Semiochemicals

Plants and insects can interact through a complex network of chemical signals termed semiochemicals. Plant-produced signals can play a crucial role in the survival of the plant in response to predators and pathogens and can help to warn other plants about impending attack through plant to plant communication (Baldwin and Schultz, 1983). Plants can release volatiles organic compounds (VOC) from vegetative organs in response to damage by herbivores. Several compounds have been studied extensively including the common plant-produced signals ethylene and methyl jasmonate. Methyl jasmonate released by sagebrush (*Artemisia tridentata*) was the first compound shown to render intact plants resistant to herbivores by increasing proteinase inhibitor production (Farmer and Ryan, 1990). Additional studies have shown that other VOCs emitted by damaged plants influence the neighboring plants, regardless of whether or not they were conspecifics (Tscharntke et al., 2001; Preston et al., 2004). There are two main types of plant response to herbivory: a direct and an indirect response. The direct response involves the production of compounds that exert an immediate effect on the herbivore or pathogen and include toxic secondary compounds like nicotine, caffeine or furanocoumarins (Rietz and Trumble, 1997), defensin proteins, and oviposition deterrents (Hilker and Meiners, 2002). The indirect response involves the production of compounds

that do not act directly on the insect but instead attract parasitoids or predators (Turlings et al., 1990; van Loon et al., 2000; Fritzsche-Hoballah and Turlings, 2001).

There has been considerable research conducted to identify and study VOCs produced by wheat (Hamilton-Kemp and Andersen, 1984, 1986; Buttery et al., 1985; Gianoli and Niemeyer, 1998; Quiroz et al., 1997; Fuentes-Contreras and Niemeyer, 1998; Jimenez-Martinez et al., 2004; Peck, 2004; Piesik et al., 2006; Lenardis et al., 2007), but less research has looked at induced compounds. In tobacco (*Nicotiana tabacum L.*), cotton (*Gossypium hirsutum L.*), and maize (*Zea mays L.*), plant herbivory by the tobacco budworm, *Heliothis virescens* Fabricius, was found to induce the production of compounds that attract the specialist parasitic wasp *Cardiochiles nigriceps* Viereck (De Moraes et al., 1998). Lima bean (*Phaseolus lunatus L.*) responds to infestation by the two-spotted spider mite (*Tetranychus urticae* Koch) by producing a blend of VOCs including several terpenoids and phenolics that attract the predatory mite (*Phytoseilus persimilis* Evans), which feeds on the spider mites (Dicke et al., 1990).

Plants produce chemical volatiles for defense, but these semiochemicals can also be used by insects as pheromones, kairomones, or allomones. Some insects use plant volatiles and pheromones, and only the perception of both types of volatile signals leads to successful attraction of conspecific mates (Landolt and Phillips, 1997; Reinecke et al., 2002; Ruther et al., 2002; Reddy and Guerrero, 2004). Plant volatiles can also help modulate the response of insect pheromones at the sensory levels (Party et al., 2009; Rouyar et al., 2011). In the scarab beetle (*Melolontha melolontha* L.), the green leafy alcohol from plants attract males. These leaf alcohols function as sexual kairomones and

attract males to leaves where females are feeding and help to synergistically increase the attractiveness of the sex pheromone (Ruther et al., 2001, 2002; Reinecke et al., 2002).

Semiochemicals and the Wheat Stem Sawfly

The chemical ecology of the wheat stem sawfly and its host plant has been investigated in recent years. Cossé et al. (2002) identified 13 WSS volatiles that stimulated both female and male antennae using antennal electrograms. There were some notable quantitative differences between the sexes, with relatively higher amounts of 9-acetyloxynonanol, phenylacetic acid, and tetradecanal produced by males and more hexadecanal produced by females. In field bioassays, both males and females were found to be attracted to the compound 9-acetyloxynonanol. Attractive odors can be used in IPM programs for monitoring population densities and adult emergence periods.

The cuticular hydrocarbons of the wheat stem sawfly were determined by Bartelt et al. (2002) and were found to contain mostly n-alkenes and n-alkanes. Alkenes were found to be the most abundant, and sexually dimorphic. (Z)-9-tricosene was the most abundant and accounted for half of the total hydrocarbons in males and was almost absent from females. In females, the dominant alkenes were (Z)-9-pentacosene and (Z)-9-heptacosene. The pheromone communication system of the WSS may be more complex compared to some other insect species. The major components of the male and female cuticular extracts only produced electrophysiological activity after they were oxidized (Cossé et al., 2002).

The wheat stem sawfly responds to a variety of wheat volatiles. Studies by Piesik et al. (2008) showed behavioral responses of WSS to wheat odors in a Y-tube olfaction

system. Females were found to be highly attracted to (Z)-3-hexenyl acetate, (E)-β-ocimene and (Z)-3-hexenol and repelled by 6-methyl-5-hepten-2-one. Males did not respond to any of these compounds. These responses were also found to be concentration dependent.

This response to wheat volatiles can be exploited for sawfly management through the development of trap crops. Weaver et al. (2009) found that the wheat variety Reeder releases greater amounts of the attractant compound (Z)-3-hexenyl acetate as compared to the variety Conan, and Reeder was significantly preferred by ovipositing females in both laboratory and field trials. Trap cropping could be optimized through careful selection of cultivars that differ in the production of attractant volatiles. Less attractive varieties can be planted as the main crop and more attractive varieties as the trap around the border. This system could also be enhanced through the use of synthetic wheat stem sawfly pheromones or additional wheat kairomones applied to the trap area (Bartelt et al., 2002; Cossé et al. 2002).

Insect Olfaction

Olfaction is the ability of an organism to detect and discriminate volatile compounds from the vast range of odors present in the environment. In animals, this sensory ability is used to find mates, locate food, and detect predators. Chemical cues are used extensively by animals as a means of communication both within species and between species (Hartlieb and Anderson, 1999). Intraspecific chemical signals, such as pheromones, are released and detected by members of the same species and can act as

either releasers or primers. Releasers are pheromones that take effect immediately; eliciting behaviors involved in kin recognition, mating, alarms against predators, oviposition, territory marking, or nest building. Primers are pheromones that cause changes in development of the organism, such as sexual maturation (Howse, 1998). Interspecific chemical signals that are released by one species and detected by another can be a kairomone, where the signal benefits the receiver or an allomone, where the sender or both the receiver and sender are benefited. Examples of kairomones include pheromones, toxins, and metabolites used in host/prey location and floral scents used to locate host plants and food sources. Allomones include defense secretions, repellents, and floral scents (Jones, 1998).

Olfaction allows insects to respond rapidly to changes in environmental cues and to detect external biological compounds via a chemical sensor. The olfactory system senses volatile odorants derived from prey, predators, host plants, and conspecific individuals. Any plant organ (leaves, stems, flowers, roots, and fruit) is a potentially rich source of volatile and non-volatile semiochemicals (VOCs) that can mediate plant-insect interactions (Pichersky et al., 2006). As a group, insects can detect and respond to most, if not all, volatiles produced by plant, secondary, metabolic pathways. These volatiles are used by insects for a wide range of purposes. Pollinators use them to locate flowers for pollen and nectar uptake (Pichersky and Gershenzon, 2002; Raguso, 2008; Schiestl, 2010). Herbivores use VOCs to locate host plants (Visser, 1986; Bruce et al., 2005; Bruce and Pickett, 2011) and to avoid non-host plants (Anderson et al., 2009), as well as to assess plant quality. Plant volatiles induced by insect feeding and damage or from

oviposition may inform herbivorous insects about the presence of competitors or inform predators about prey location (Kessler and Baldwin, 2001). There is also growing evidence that background odors, including plant volatiles, can help inform insects about the quality of a habitat (Schroder and Hilker, 2008). Herbivores may use these volatiles to aggregate on host plants as well as for finding mates. Volatile blends and the context of their detection can be an important factor required to elicit a behavioral response. For example, some insects require both plant volatiles and pheromones for successful attraction to a conspecific insect (Reinecke et al., 2002; Ruther et al., 2002; Reddy and Guerrero, 2004).

The general route of detection is the same, regardless of the origin of the odorant. The entire olfactory system relies on different types of receptors expressed in peripheral olfactory receptor neurons (ORNs).

The Olfactory System

The head of adult insects has two main sets of olfactory organs, the antennae and the maxillary palps. Antennae are the major odor detecting appendages in insects, including the wheat stem sawfly. Hair-like projections on the surface of the antennae called sensilla house specialized bipolar chemosensory neurons; the dendrite extends into the hollow sensilla and the axon extends to the central nervous system (CNS). Each sensillum is filled with a potassium- and protein-rich fluid called sensillum lymph. The small pits or pores on the cuticle surface of the sensilla allow chemicals from the external environment to access to the neuron dendrite via the lymph. Chemosensory neurons are further classified as gustatory neurons involved in contact chemosensation (taste and

gustation) and olfactory neurons that detect odorants (volatile molecules) present in the environment (Schneider, 1964). The axons of gustatory neurons project to the subesophageal ganglion that processes taste information, whereas the axons of olfactory neurons project to the antennal lobe that first processes olfactory cues.

An olfactory sensillum has a porous outer wall and the inner layer is filled with an aqueous fluid, the sensillum lymph (Breer, 1997). The dendrite of the bipolar olfactory receptor neuron projects into the sensillum lymph while the axon projects to the antennal lobe (Keil, 1999). The cell body of the ORN is surrounded by three auxiliary cells: the tormogen, trichogen, and thecogen (Steinbrecht et al., 1992). These cells participate in sensillum structural development and in maintaining the physiological composition of the sensillum lymph. Proteins and enzymes within the sensillum lymph, such as soluble odorant binding proteins (OBPs) and odorant degrading enzymes (ODEs), are important physiological components of odor detection (Vogt, 2005). The actual odorant receptors (ORs) are transmembrane proteins embedded in the dendrite membrane.

Chemosensory sensilla can exhibit a wide range of morphologies and are characterized by their shape, number, and location of pores on their surfaces and by the bipolar sensory cells in their lumens (Altner, 1977). Contact chemosensilla involved in taste have a single terminal pore and are innervated by four gustatory neurons, whereas olfactory sensilla tend to be multiporous and can be innervated by a few or up to hundreds of olfactory neurons. In Hymenoptera, six types of sensilla involved in chemo- and mechanoreception are the most common: *sensilla trichodea, sensills chaetica, sensilla basiconica, sensilla coeloconia, sensilla placodea*, and *sensilla ampullaceal*

(Feng et al., 1992; Zhang and Xiao, 1992). In the web spinning sawfly, *Acantholyda posticalis* Matsumura, research has demonstrated that some *sensilla chaetica* are mechanoreceptors involved in the perception of texture, movement, and wind (Yuan et al., 2013; Isidoro et al., 1996). *Sensilla trichodea* serve a similar function in some parasitic wasps (Gao et al., 2007; Van Baaren et al., 1999) and, due to the structure and location within the antennae of *A. posticalis*, Yuan et al. (2013) suggested that *sensilla trichodea* serve as contact chemoreceptors with a tactile function at the tip, used to touch host plants or potential mates. In other insects trichoid sensilla are multiporous and have an olfactory function. *Sensilla basiconica* have numerous pores, and the dendrites of the olfactory neurons are branched, indicating a role in olfaction (Altner and Prillinger, 1980; Zacharuk, 1985; Yuan et al., 2013). *Sensilla coelonica* have been characterized as having deep longitudinal grooves and a single terminal pore with more than five ORNs, suggesting they detect long distance olfactory stimuli (Altner et al., 1983; Ochieng et al., 2000; Roux et al., 2005; Yuan et al., 2013).

The cascade of molecular events that begins with the detection of an odor in the peripheral olfactory system and leads to its conversion into a neuronal signal to the antennal lobe of the insect brain is still not completely understood. The current model of olfaction begins with an odorant molecule contacting the cuticle of the sensilla on the antenna, after which it migrates along the cuticle surface to a pore that allows entry into the sensillum (Steinbrecht, 1997). There is no physical barrier that separates the sensillum lymph from the external environment at the pore entrance. Soluble odors simply dissolve in the sensillum lymph. Hydrophobic odors that are not water-soluble require transport

across the hydrophilic sensillum lymph. These odors are bound and transported by a family of odorant binding proteins (OBP) across the aqueous lymph to the odorant receptor (OR) located in the membrane of the sensory dendrite (Vogt and Riddiford, 1981; Vogt, 2003; Tegoni et al., 2004). Upon binding of the ligand or the ligand-OBP complex, the OR will change conformation and trigger the signaling mechanism leading to a nerve impulse. Odorant degrading enzymes (ODE) metabolize the stimulating molecule, thereby "resetting" the sensory neuron allowing it to respond to further stimulation (Vogt and Riddiford, 1981; Vogt, 2003).

Odorant Binding Proteins

Odorant binding proteins (OBP) are small, soluble, and globular proteins consisting of about 150 amino acids that are abundant in the sensillum lymph. OBPs are translated in the support cells surrounding the ORN and secreted into the sensillum lymph (Vogt and Riddford, 1981). The first OBP discovered was a male-specific antennal protein unique to the pheromone-sensitive sensilla that detect female-produced sex pheromones in the wild silk moth, *Antheraea polyphemus* (Cramer) (Vogt and Riddiford, 1981). Due to its specificity for sex pheromones, it was termed a pheromone-binding protein (PBP). By the early 1990s, several OBPs had been discovered and it became clear that they are a multigene family unique to insects. Additional members were characterized as general odorant binding proteins that did not exhibit sex bias (Vogt et al., 1991; Kreiger et al., 1993, 1996). Diverse biochemical roles have been proposed for OBPs, including solubilization of odors in the sensillum lymph, the transport of odors through the lymph, interaction of the complex with specific ORs, protection of the

odorant ligand from degradation by formation of an OBP-ligand complex, and deactivation of the ligand.

Two models have been proposed to explain how OBPs transport ligands across the aqueous lymph. The first model was discovered and inferred from biochemical, biophysical, structural, and kinetic studies of the PBPs Lepidoptera including the silk moth, *Bombyx mori* L. The 3-D structure of *B. mori* PBP1 changes in response to pH and the presence of ligand. At low pH the C-terminus domain is inserted into the binding pocket of the PBP, whereas at higher pH, it forms an unstructured tail leaving the binding pocket open. This phenomenon led to the hypothesis that a transition from a neutral pH where the PBP is in the open form, to a lower pH where it is in a closed form, is physiologically relevant (Wojtasek and Leal, 1999). Kinetic studies demonstrated that release of bombykol from *B. mori* PBP1 is 10,000-fold faster at low pH (Leal et al., 2005). Collectively, these experimental results support the hypothesis that PBPs are able to bind pheromones under the neutral pH conditions of the lymph and transport and protect them from odor-degrading enzymes. When the PBP-pheromone complex comes in close proximity to the dendrite membrane, the pheromone is released because the pH near the membrane lipid is lower compared to the lymph, and the low pH triggers the conformational change where the C-terminus domain inserts into the binding pocket (Leal, 2013).

The second mode of action has been suggested for the OBP LUSH, from *Drosophila melanogaster* L. (Xu et al., 2005). Mutant flies that lack LUSH are defective and do not avoid high concentrations of ethanol. LUSH is expressed in pheromone-

sensitive sensilla that detect the male-specific lipid 11-cis vaccenyl acetate (cVA) that mediates aggregation behavior in wild type flies. This behavior is lost in mutant flies that lack LUSH. The activity of this pheromone was restored by transgenic rescue, where expression of the LUSH protein was restored, suggesting that the loss of function in the mutants was specifically caused by the absence of LUSH in pheromone sensitive sensilla (Zhou et al., 2004; Xu et al., 2005). By mutating specific amino acid residues of the LUSH protein, Laughlin et al. (2008) demonstrated that the OBP itself was activating the OR and subsequent nerve impulse. CVA binding to LUSH causes a conformational change in the 3-D structure of the OBP transforming it into an active form that can activate the OR directly. This conformational change can be induced in the absence of cVA by mutating specific amino acids, and this mutant LUSH can activate the pheromone sensitive ORN in the absence of the pheromone (Zhou et al., 2004; Xu et al., 2005).

Chemosensory Proteins

Chemosensory proteins (CSPs) represent a second class of soluble transport proteins that may function in olfaction. Similar to OBPs, they are small, globular proteins that have been implicated in the binding of odorants and pheromones to transport them across the sensillum lymph (Forêt et al., 2007). The first CSP was identified from *D. melanogaster* using subtractive hybridization experiments to enrich for antennal specific cDNA (McKenna et al., 1994). Some studies have localized CSPs to the lymph surrounding the dendrites of olfactory and taste sensilla, supporting their role as odorant carriers (Angeli et al., 1999; Monteforti et al., 2002; Jin et al., 2005). Recent studies have

confirmed that CSPs are able to bind a range of aliphatic compounds, esters, and long-chain fatty acids that are typical pheromone compounds (Lartigue et al., 2002; Briand et al., 2002). The 3-D structure of a CSP from *Mamestra brassicae* (L.) forms a globular configuration with six alpha helices surrounding an internal hydrophobic binding pocket (Mosbash et al., 2003). Forêt et al. (2007) suggest that CSPs are more highly conserved and ancient within arthropod lineages compared to the OBPs.

Odorant Receptors

Odorant receptors (ORs) are seven-transmembrane (7TM) domain proteins localized to the dendritic membrane of the ORNs (Benton et al., 2006). ORs are activated by odorants after they are transported to the dendritic membrane. Their activation initiates a signaling cascade causing the ORN to depolarize and a nerve impulse to be relayed via the axon to the antennal lobe in the brain.

ORs were first discovered from *D. melanogaster* using a bioinformatics-based approach designed to detect G-protein-coupled receptor (GPCR)-like olfactory genes from the recently sequenced genome (Clyne et al., 1999). Both GPCRs and insect ORs have 7TM domains; however, the insect ORs have a reversed topology compared to GPCRs with an intracellular N-terminus. Evidence for this reversed topology came from epitope tagging of the C- and N-termini and the predicted loop regions of *Drosophila* OR22a to visualize intracellular and extracellular domains (Smart et al., 2008). These studies suggested that insect ORs are unique and different from GPCRs and use insect-specific signaling pathways. Since the initial discovery of insect ORs, knowledge of the molecular basis of insect olfaction has expanded greatly. The genomes of several insect

species have been sequenced facilitating the identification of large OR gene families from moths, mosquitoes, ants, honeybees, wasps, and beetles (Fox et al., 2001; Nakagawa et al., 2005; Wurm et al., 2010; Robertson & Wanner, 2006; Robertson et al., 2010; Engsontia et al., 2008).

Odorant Receptor Co-receptor

The odorant receptor co-receptor (ORco) was first described from *D. melanogaster* as OR83b, and characterized as the only highly conserved OR that retains orthology across insect orders (Vosshall and Hansson, 2011). Although OR sequence diversity within and between species is high, ORco is highly conserved in insects such as *D. melanogaster*, *B. mori*, *Apis mellifera* L., and *Nasonia vitripennis* L. (Vosshall et al., 1999; Hill et al., 2002; Robertson and Wanner, 2006; Robertson et al., 2010). Its broad sequence conservation is explained by its critical function as a co-receptor for all other ORs. Fruit fly mutants lacking ORco are unable to detect most odors; their olfactory ability is broadly impaired by the lack of this single OR gene. Orthologs of ORco from other insect species were hypothesized to have a similar functional role (Larrson et al., 2004; Neuhaus et al., 2005; Nakagawa et al., 2005).

Recent studies have determined that ORco functions as a ligand-gated ion channel when it forms a heterodimer with regular ORs that confer ligand sensitivity. Sato et al. (2008) expressed a number of different insect OR receptors in HeLa cells, together with the co-receptor ORco, and demonstrated their activation by odors using changes in intracellular calcium concentration and whole-cell currents as indicators of activation. OR expressing cells were activated by odorant stimulation 10-fold faster compared to

cells expressing GPCRs, providing the first evidence of a novel insect ion channel independent of known G-protein-coupled second-messenger pathways. This response was independent of secondary messengers like cGMP and cAMP that are common with GPCR signaling (Smart et al., 2008). Collectively, these pharmacological studies clearly demonstrated that the ORco-OR complex functions as a cation-non-selective ligand-gated ion channel (Sato et al., 2008).

Odorant Receptors in Hymenoptera

The first ORs were identified from *D. melanogaster* in 1999 by combined experimental and bioinformatic approaches (Vosshall et al., 1999; Clyne et al., 1999). OR sequences are highly divergent with little sequence homology both within and between species. For example, *D. melanogaster* ORs share only 17–26% amino-acid sequence identity (Vosshall and Stocker, 2007) and *A. mellifera* ORs can have as little as 20% amino-acid identity (Robertson and Wanner, 2006). Initially, the OR gene family in *Drosophila* comprised 57 receptors (Vosshal et al., 2000), but subsequent genome annotation has brought the final number to 60 OR genes (Robertson et al., 2003). The recent sequencing of additional insect genomes has allowed for OR identification in several Hymenoptera species, including the honeybee, jewel wasp, and two ant species. The *A.mellifera* genome sequence revealed a total of 163 functional OR genes and 7 pseudogenes (Robertson and Wanner, 2006), the *N. vitripennis* genome encodes 225 functional OR genes and 76 pseudogenes (Robertson et al., 2010), and the *Solenopsis invicta* Buren genome contains at least 400 ORs, of which 297 are intact (Wurm et al., 2011). The expansion of ORs from *Drosophila* to *Apis* was originally interpreted as

reflecting increased gene duplication to mediate the detection of social pheromones and floral scents of bees (Robertson and Wanner, 2006). The life history and complexity of an insect's chemical environment has been used to explain the number of ORs encoded in their genome. However, the jewel wasp, *Nasonia,* was later found to have more ORs than the social honeybee, the result of several lineage specific gene expansions (Robertson et al., 2010). This suggests that the number of ORs encoded in an insect's genome may not necessarily be directly correlated with its chemical ecology as first thought.

Ionotropic Receptors

The ORs do not reflect the entire repertoire of chemosensory genes that are expressed specifically in the olfactory neurons. Benton et al. (2009) recently discovered a divergent family of integral membrane proteins called ionotropic receptors, or IRs, from the *D. melanogaster* genome, that are related to the glutamate ionotropic family of receptors. These IRs were found expressed in a subset of ORNs innervating coeloconic sensilla that did not express ORs, but were able to detect and respond to diverse odorants, including acids, ammonia, and water (Yao et al., 2005).

Analysis of the *Drosophila melanogaster* genome identified 66 IR genes. Comprehensive expression analysis by qRT-PCR, fluorescence RNA *in situ* hybridization, and transgenic reporters showed that 16 of these IRs are expressed in the antenna (Benton et al., 2009). Spatial mapping of these 16 IR genes in *D. melanogaster* showed that their distribution corresponds to physiologically characterized coeloconic sensilla, providing indirect evidence that these receptors define the odor responses for these neurons (Yao et al., 2005; Benton et al., 2009). IRs contain a predicted extracellular

N-terminus, a bipartite ligand-binding domain (LBD) whose two lobes (S1, S2) are separated by an ion-channel domain and a short cytoplasmic C-terminal region (Rytz et al., 2013). Demonstration that IRs are necessary and sufficient for odor-evoked responses came from loss- and gain-of function studies (Ai et al., 2010; Grosjean et al., 2011). Functional assays of IRs in *Xenopus* oocytes have suggested that IR84a and IR8a form a heteromeric complex composed of two subunits including the odor specific receptor (IR84a) and the co-receptor (IR8a). The co-receptor may form a structural scaffold onto which IR84a or other odor-specific IRs can insert (Rytz et al., 2013). Croset et al. (2010) have identified additional IRs from a wide range of insect species, including 10 IRs in *A. mellifera* and *N. vitripennis*, using bioinformatic analysis.

Odorant Degrading Enzymes

Odorant degrading enzymes (ODEs) are an additional class of soluble proteins present in the sensillum lymph (Vogt and Riddiford, 1981). Once the odorant signal has been relayed to the neuron, the odorant needs to be cleared from the lymph so that the sensitivity of the olfactory system is maintained and new signals can be detected, since continuous firing of ORNs can lead to loss of sensitivity. Many different odors may enter the sensillum lymph and many will not be cognate ligands for the ORN. These background odors may also be toxic to the ORN and their removal from the sensillum lymph is important. Three protein families that function in biotransformation reactions, such as detoxification, have evolved secondary olfactory functions. These include the carboxylesterases (CCEs), glutathione S-transferases (GSTs), and cytochrome P450s (P450s). CCEs include antennal-specific esterases that degrade pheromones (Ishida and

Leal, 2005; Durand et al., 2011). The best-studied ODE is the esterase, ApolPDE, from the moth, *A. polyphenus.* ApolPDE is expressed in the pheromone-detecting sensilla and is secreted into the sensillum lymph (Ishida and Leal, 2005). Kinetic studies using both native and recombinant enzymes show that ApolPDE rapidly degrades the main component of the sex pheromone, E6Z11-16OAc (Ishiha and Leal, 2005; Vogt, 2005). GSTs are enzymes that often function as deotixifcation enzymes, complexing venobitoics to glutathione and thereby rendering them harmless to the insect (Fournier et al., 1992; Synder, 1995). GSTs can have features suggesting a dual role of attacking both xenobiotics and odor moleculates. The best characterized is MsexGSTolf, an antennal-specific GST from *M. sexta* capable of transforming aldehyde odorants (Rogers et al., 1999). MsexGSTolf has been found in male and female antennae, but in male antennae, it is restricted to the cells underlying the pheromone-sensitive trichoid sensilla, suggesting that this GST plays a role in pheromone inactivation (Rogers et al., 1999). P450 enzymes are generally known for their role in xenobiotic metabolism in insects; functional evidence for P450s as ODEs is still scant, but one study from Wojtasek and Leal (1999b) showed a P450 enzyme specifically produced in the male antennae of the pale brown chafer, *P. diversa*, was able to metabolize an alkaloid sex pheromone.

There is little known about the molecular biology and genetics of the WSS. Two studies published to date focused on using molecular tools to study WSS taxonomy and population biology. Hartel et al. (2003) isolated and characterized several microsatellites from both the WSS and its related species. Lou et al. (1998) conducted a random, amplified polymorphic DNA (RAPD) analysis on different geographic populations of the

WSS. Beyond these two studies, little molecular biology work has been conducted. The economic importance of this pest combined with the increasing availability of genomic and molecular biology tools for basic research has provided the rationale for this study, increasing our knowledge of the molecular basis of wheat stem sawfly olfaction. This study identifies and characterizes the major gene families involved in WSS peripheral olfaction and identifies candidate genes that might mediate important pest behaviors and, thus, serve as molecular targets for new management strategies for this important pest.

References

Ai, M., Min, S., Grosjean, Y. Leblanc, C., Bell, R., Benton, R. and Suh, G.S. 2010. Acid sensing by the *Drosophila* olfactory system. *Nature* 468: 691–695.

Ainslie, C.N. 1920. The western grass-stem sawfly. *USDA. Bull. No.* 841.

Ainslie, C.N. 1929. The western grass-stem sawfly - a pest of small grains. *USDA. Bull. No*. 157.

Altner, H. 1977. Insect sensillum specificity and structure: an approach to new typology. *Olfaction and Taste* 6: 295–303.

Altner, H. and. Prillinger, L. 1980. Ultrastructure of invertebrate chemo-, thermo-, and hygroreceptors and its functional significance. *Intl Rev Cytol* 67: 69–139.

Altner, I., Hatt, H. and Altner, H. 1983. Structural properties of bimodal chemo- and mechanosensitive setae on the pereiopod chelae of the crayfish, *Austropotamobius torrentium*. *Cell Tissue Res* 242: 357–374.

Anderson, A.R., Wanner, K.W., Trowell, S.C., Warr, C.G., Jaquin-Joly, E., Zagatti, P., Robertson, H. and Newcomb, R. D. 2009. Molecular basis of female-specific odorant responses in *Bombyx mori*. *Insect Biochem. Mol. Biol*. 39: 189–197.

Angeli, S., Ceron, F., Scaloni, A., Monti, M., Monteforti, G., Minnocci, A., Petacchi, R. and Pelosi, P. 1999. Purification, structural characterization, cloning and immunocytochemical localization of chemoreception proteins from *Schistocerca gregaria*. *Eur. J. Biochem* 262: 745–754.

Baldwin, I.T. and Schultz, J.C. 1983. Rapid changes in tree leaf chemistry induced by damage: evidence for communication between plants. *Science* 221: 277–279

Bartlet, R.J., Cossé, A.A., Petroski, R. J. and Weaver, D. K. 2002. Cuticular hydrocarbons and novel alkenediol diacetates from wheat stem sawfly (*Cephus cinctus):* natural oxidation to pheromone components. *J Chem Ecol* 28: 385-405.

Benton, R., Sachse, S., Michnick, S.W. and. Vosshall, L.B. 2006. Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo. *PLoS Biol* 4: e20.

Benton, R., Vannice, K.S., Gomez-Diaz, C., Vosshall, L.B. 2009. Variant ionotropic glutamate receptors as chemosensory receptors in Drosophila. *Cell* 136(1): 149–162.

Beres, B.L., Dosdall, L.M., Weaver, D.K. Cárcamo, H.A. and Spaner, D.M. 2011. Biology and integrated management of wheat stem sawfly and the need for continuing research. *Can Entomol* 143: 105–125.

Blodgett, S. L., Goosey, H.B.,Waters, D.K., Tharp, C. I. and Johnson, G. D. 1996. Wheat stem sawfly control on winter wheat. *Arthropod Manag Tests* 22: 331-332.

Bradshaw, J. 2013. Adult Wheat Stem Sawflies Have Begun to Emerge. University of Nebraska *Cropwatch*. 30 May 2013.

Breer, H. 1997. Molecular mechanisms of pheromone reception in insect antennae. In Insect Pheromone Research New Directions (ed. R. T. Cardé and A. K. Minks), pp. 115–130. New York: Chapman & Hall.

Briand, L., Swasdipan, N., Nespoulous, C., Bezirard, V., Blon, F., Huet, J.C., Ebert, P. and Penollet, J. C. 2002. Characterization of a chemosensory protein (ASP3c) from honeybee (*Apis mellifera* L.) as a brood pheromone carrier. *Eur J Biochem* 269: 4586–4596.

Bruce, T. J. A. and Pickett, J. A. 2011. Perception of plant volatile blends by herbivorous insects - finding the right mix. *Phytochemistry* 72: 1605–1611.

Bruce, T. J. A., Wadhams, L. J. and Woodcock, C. M. 2005. Insect host location: a volatile situation. *Trends Plant Sci* 10: 269–274.

Buttery, R. G., Xu, C. and Ling. L. C. 1985. Volatile components of wheat leaves (and stems): possible insect attractants. *J Agri Food Chem* 33: 115–117 .

Callenbach, J. A. and Hansmeier, M. P. 1944. Wheat stem sawfly control in severely infested areas. *Montana Extension Service in Agriculture and Home Economics*, Circular No. 156.

Clyne, P. J., Warr, C. G., Freeman, M. R., Lessing, D., Kim, J., and Carlson, J. R. 1999. A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. *Neuron* 22: 327–338.

Cochran, V., Danielson, J., Kolberg, R., and Miller, P. 2006. Dryland cropping in the Canadian prairies and the US Great Plains. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America. Dryland Agriculture, 2nd ed., Agronomy Monograph 23.

Cossé, A. A., Bartelt, R. J., Weaver, D. K., and Zilkowski, B. W. 2002. Pheromone Components of the Wheat Stem Sawfly: Identification, Electrophysiology, and Field Bioassay. *J Chem Ecol* 28(2): 407-423.

Criddle, N. 1915. The Hessian fly and the western wheat stem sawfly in Manitoba, Saskatchewan and Alberta. *Canadian Department of Agriculture, Entomology Branch Bulletin* 11: 1-23.

Criddle, N. 1922. The western wheat-stem sawfly and its Control. *Dominion of Canada Department of Agriculture Pamphlet* 6- New Series 3-8.

Criddle, N. 1923. The life habits of *Cephus cinctus* Norton in Manitoba. *Can Entomol* 55: 1-4.

Croset, V., Ritz, R., Cummins, S. F., Budd, A., Brawand, D., Kaessmann, H., Gibson, T. J., and Benton, R. 2010. Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet* 6: e1001064.

De Moraes, C. M., Lewis, W. J., Paré, P. W., and Tumlinson, J. H. 1998. Herbivore infested plants selectively attract parasitoids. *Nature* 393: 570-574.

DePauw, R. M., Clarke, J. M., Knox, R. E., Fernandez, M. R., McCraig, T. N., and McLeod, J. G. 2000. AC Abbey hard red spring wheat. *Can J Plant Sci* 80: 123-127.

DePauw, R. M., McLeod, J. G., Clarke, J. M., McCaig, T. N., Fernandez, M. R., and Knox, R. E. 1994. AC Eatonia hard red spring wheat. *Can J Plant Sci* 74: 821–823.

DePauw, R. M., Townley-Smith, T. F., Humphreys, G., Knox, R. E., Clarke, F. R., and Clarke, J. M. 2005. Lillian hard red spring wheat. *Can J Plant Sci* 85: 397–401.

Dicke, M., Sabelis, M. W., Takabayashi, J., Bruin, J., and Posthumus, M. A. 1990. Plant strategies of manipulating predator-prey interactions through allelochemicals: prospects for application in pest control. *J Chem Ecol* 16: 3091-3118

Durand N., Carot-Sans, G., Bozzolan, F., Rosell, G., Siaussat, D., Debernard, S., et al. 2011. Degradation of Pheromone and Plant Volatile Components by a Same Odorant-Degrading Enzyme in the Cotton Leafworm, *Spodoptera littoralis*. *PLoS ONE* 6(12): e29147.

Farmer, E. E. and Ryan, C.A. 1990. Interplant communication: airborne methyl jasmonate induces synthesis of proteinase inhibitors in plant leaves. *Proc. Natl. Acad. Sci. USA.* 87: 7713-7716.

Farstad, C. 1940. The development of western wheat stem sawfly (*Cephus cinctus* Nort.) in various host plants as an index of resistance [PhD dissertation]. Iowa State College, Ames, IA. P 83.

Farstad, C., Platt, A. W., and McGinnis, A. J. 1949. Influence of wheat varieties on the sex ratio of the wheat stem sawfly, *Cephus cinctus* Nort. (Hymenoptera: Cephidae). *80th Annual Report of the Entomological Society of Ontario.*

Feng, H. L., Zhang, D. X., Chang, T. R., Guo, W. Y., Zhang, R. C., and Xin, Y. 1992. Scanning electron microscopy studies of antennal sensilla and ommateum eye of ant. *Entomol Knowledge* 29: 292–294.

Forêt, S., Wanner, K. W., and Maleszka, R. 2007. Chemosensory proteins in the honey bee: insights from the annotated genome, comparative analyses and expressional profiling. *Insect Biochem Mol Biol* 37: 19–28.

Fox, A. N., Pitts, R. J., and Zwiebel, L. J. 2002. A cluster of candidate odorant receptors from the malaria vector mosquito, Anopheles gambiae. *Chemical senses*, 27(5), 453–459.

Fritzsche-Hoballah, M. E. and Turlings, T. C. J. 2001. Experimental evidence that plants under caterpillar attack may benefit from attracting parasitoids. *Evol Ecol Res* 3: 553–565.

Fuentes-Contreras, E. and Niemeyer, H. M. 1998. Dimboa glucoside, a wheat chemical defense, affects host acceptance and suitability of *Sitobian avanae* to the cereal aphid parasitoid *Aphidius rhopalosiphi*. *J Chem Ecol* 24: 371-381.

Gao, Y., Luo, L. Z., and Hammond, A. 2007. Antennal morphology, structure and sensilla distribution in *Microplitis pallidipes* (Hymenoptera:Braconidae). *Micron* 38: 684–693

Gianoli, E. and Niemeyer, H. M. 1998. Allocation of herbivory-induced hydroxamic acids in the wild wheat *Triticum uniaristatum*. *Chemoecology* 8: 19-23.

Goosey, H. B. 1999. In field distributions of the wheat stem sawfly, (Hymenoptera: Cephidae), and evaluation of selected tactics for an integrated management program. [Master's Thesis]. Montana State University, Bozeman, Montana.

Grosjean, Y., Rytz, R., Farine, J. P., Abuin, L., Cortot, J., Jefferis, G. S., and Benton, R. 2011. An olfactory receptor for food-derived odours promotes male courtship in Drosophila. *Nature* 478: 236–240.

Hamilton-Kemp, T. R. and Anderson, R. A. 1984. Volatile compounds from *Triticum aestivum*. *Phytochem* 23: 1176–1177.

Hamilton-Kemp, T. R. and Anderson, R. A. 1986. Volatiles from winter wheat: identification of additional compounds and effects of tissue source. *Phytochem* 25**:** 241–243.

Hartel, K. D., Frederick, B. A., and Shanower, T. G. 2003. Isolation and characterization of microsatellite loci in wheat stem sawfly *Cephus cinctus* and cross-species amplification in related species. *Mol Ecol Notes* 3(1): 85-87.

Hartlieb, E. and Anderson, P. 1999. Olfactory-released behaviours. In Hansson, B.S. (ed.), Insect Olfaction. Springer, Berlin, pp.315– 349.

Hilker, M. and Meiners, T. 2002. Induction of plant responses towards oviposition and feeding of herbivorous arthropods: a comparison. *Entomol Exp Appl* 104: 181-192.

Hill, E. S., Iwano, M., Gatellier, L., and Kanzaki, R. 2002. Morphology and physiology of the serotonin-immunoreactive putative antennal lobe feedback neuron in the male silk moth Bombyx mori. *Chem Senses* 27: 475-483.

Holmes, N. D. 1977. The effect of the wheat stem sawfly, *Cephus cinctus* (Hymenoptera: Cephidae), on the yield and quality of wheat. *Can Entomol* 109: 1591-1598.

Holmes, N. D. 1978. The wheat stem sawfly. *Proc 26th Annu Mtg Entomol Soc Alberta*, pp. 2-13.

Holmes, N. D. 1982. Population dynamics of the wheat stem sawfly, *Cephus cinctus* (Hymenoptera Cephidae), in wheat. *Can Entomol* 114: 775-788.

Holmes, N. D. and Peterson, L. K. 1960. The influence of the host on oviposition by the wheat stem sawfly, *Cephus cinctus* Nort. (Hymenoptera: Cephidae) II. Resistance to the larva. *Can Entomol* 94: 348-365.

Howse, P. E. 1998. Pheromones and Behaviour. Insect Pheromones and their Use in Pest Management. P. E. Howse, I. D. R. Stevens and O. T. Jones. London, Chapman & Hall.

Irell, B. and Peairs, F. 2011. Wheat Stem Sawfly: A New Pest of Colorado Wheat. *Colorado State University Extension*. Circular No. 5.612.

Ishida, Y. and Leal, W. S. 2005. Rapid inactivation of a moth pheromone. *Proc. Natl. Acad. Sci. USA* 102: 14075–14079.

Isidoro, N., Bin, F., Colazza, S., and Vinson, S. B. 1996. Morphology of antennal gustatory sensilla and glands in some parasitoids Hymenoptera with hypothesis on their role in sex and host recognition. *J Hym Research* 5: 206–239.

Ivie, M. A. (2001) On the Geographic Origin of the Wheat Stem Sawfly (Hymenoptera: Cephidae): A New Hypothesis of Introduction from Northeastern Asia. *Am Entomol* 47: 84-97.

Jiménez-Martínez, E., Bosque-Pérez, S. N. A., Berger, P. H., Zemetra, R. S., Ding, H., and Eigenbrode, S. D. 2004. Volatile cues influence the response of *Rhopalosiphum padi* (Homoptera: Apididae) to barley yellow dwarf virus-infected transgenic and untransformed wheat. *Environ Entomol* 33: 1207-1216.

Jin, X., Brandazza, A., Navarrini, A., Ban, L., Zhang, S., Steinbrecht, R. A., Zhang, L., and Pelosi, P. 2005. Expression and immunolocalisation of odorant-binding and chemosensory proteins in locusts. *Cell Mol Life Sci* 62:1156–1166.

Jones, O. T. 1998. Practical Applications of Pheromones and Other Semiochemicals. Insect Pheromones and their Use in Pest Management. P. E. Howse, I. D. R. Stevens and O. T. Jones. London, Chapman & Hall.

Keil T. A. 1999. Morphology and development of the peripheral olfactory organs. In: Insect Olfaction, edited by Hansson BS, editor. Berlin: Springer. p. 6–47.

Kessler, A. and Baldwin, T. 2001. Defensive function of herbivore-induced plant volatile emissions in nature. *Science* 291: 2141-2144.

Krieger, J., Ganssle, H., Raming, K., and Breer, H. 1993. Odorant binding proteins of Heliothis virescens. *Insect Biochem Mol Biol* 23: 449-456.

Krieger, J., von Nickisch-Rosengk, E., Mameli, M., Pelosi, P., and Breer, H. 1996. Binding proteins from the antennae of *Bombyx mori*. *Insect Biochem Mol Biol* 26: 297-307.

Landolt, P. J. and Phillips, T. W. 1997. Host plant influences on sex pheromone behavior of phytophagous insects. *Ann Rev Entomol* 42**:** 371-391.

Larsson, M. C., Domingos, A. I., Jones, W. D., Chiappe, M. E., Amrein, H., and Vosshall, L. B. 2004. Or83b encodes a broadly expressed odorant receptor essential for Drosophila olfaction. *Neuron* 43: 703–714.

Lartigue, A., Campanacci, V., Roussel, A., Larsson, A. M., Jones, T. A., Tegoni, M., and Cambillau, C. 2002. X-ray structure and ligand binding study of a moth chemosensory protein. *J Biol Chem* 277: 32,094–32,098.

Laughlin, J., Ha, T., Jones, D., and Smith, D. 2008. Activation of pheromone sensitive neurons is mediated by conformational activation of pheromone binding protein. *Cell* 133: 1255–1265.

Leal, W. S. (2013) Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol* 58: 373-391.

Leal, W. S., Chen, A. M., Ishida, Y., Chiang, V. P., Erickson, M. L., Morgan, T. I., Tsuruda, J. M. 2005. Kinetics and molecular properties of pheromone binding and release. *Proc Natl Acad Sci USA* 102: 5386–5391.

Lenardis, A. E., vand Baren, C., Di Leo Lira, P., and Ghersa, C. M. 2007. Plant-soil interactions in wheat and coriander crops driving arthropod assemblies through volatile compounds. *Europ J Agronomy* 26: 410-417.

Lou, K. F., Weiss, M. J., Bruckner, P. L., Morrill, W. L., Talbert, L. E., and Martin, J. M. 1998. RAPD Variation Within and Among Geographic Populations of Wheat Stem Sawfly (*Cephus cinctus* Norton). *J Heredity* 89(4): 329-335.

McBride, D. K., Koop, D. D., and Nyegaard, C. W. 1989. Wheat stem insect pests and management practices. North Dakota State University Extension Service, Fargo, ND. E-680.

McKenna, M. P., Hekmat-Scafe, D. S., Gaines, P., Carlson, J. R., and Putative, D. 1994. Melanogaster pheromone-binding proteins expressed in a subregion of the olfactory system. *J Biol Chem* 269: 16340-16347.

Meers, S. B. 2005. Impact of harvest operations on parasitism of the wheat stem sawfly, *Cephus cinctus* Norton (Hymenoptera: Cephidae). [Master's Thesis]. Entomology, Montana State University, Bozeman, Montana. 129 p.

Monteforti, G., Angeli, S., Petacchi, R., and Minnocci, A. 2002. Ultrastructural characterization of antennal sensilla and immunocytochemical localization of a chemosensory protein in *Carausius morosus* BRÜNNER (Phasmida: Phasmatidae). *Arthropod Struct Dev* 30: 195–205.

Morrill, W. L. 1983. Early history of cereal grain insect pests in Montana. *Bull Entomol Soc Amer* 29: 24–28.

Morrill, W. L. 1997. The wheat stem sawfly, *Cephus cinctus* Nort. (Hymenoptera: Cephidae) and associated parasitoids in the northern Great Plains of North America. *Trends Entomol*. 1: 171-174.

Morrill, W. L., Gabor, J. W., and Kushnak, G. D. 1992. Wheat stem sawfly (Hymenoptera: Cephidae) at low temperatures. *Environ Entomol*. 22: 1358-1361.

Morrill, W. L. and Kushnak, G. D. 1996. Wheat stem sawfly (Hymenoptera: Cephidae) adaptation to winter wheat. *Environ Entomol*. 25: 1128-1132.

Morrill, W. L., and Kushnak, G. D. 1999. Planting date influence on the wheat stem sawfly (Hymenoptera: Cephidae) in spring wheat. *J Agric Entomol*. 16: 123-128.

Morrill, W. L., Kushnak, G. D., Bruckner, P. L., and Gabor, J. W. 1994. Wheat stem sawfly damage, rates of parasitism, and overwintering survival in resistant wheat lines. *J Econ Entomol*. 87: 1373-1376.

Morrill, W. L., Kushnak, G. D., and Gabor, J. W. 1998.. Parasitism of the wheat stem sawfly (Hymenoptera: Cephidae) in Montana. *Biol Control* 12:159-163.

Morrill, W. L., Weaver, D. K., and Johnson, G. D. 2001. Trap strip and field modification for management of the wheat stem sawfly (Hymenoptera: Cephidae). *J Entomol Sci*. 36: 34-45.

Mosbah, A., Campanacci, V., Lartigue, A., Tegoni, M., Cambillau, C., Darbon, H. 2003. Solution structure of a chemosensory protein from the moth Mamestra brassicae. *J Biochem* 369: 39–44.

Munro, J. A. 1945. The Wheat Stem Sawfly and Harvest Loss. *North Dakota Agricultural Experimental Station*. Bimonthly Bulletin. 7: 12-16.

Munro, J. A. 1947. Wheat sawfly is on the wing. *Co-Op Grain Quarterly, National Confederation of Grain Cooperatives*. 5: 39-42.

Nakagawa, T., Sakurai, T., Nishioka, T., and Touhara, K. 2005. Insect sex-pheromone signals mediated by specific combinations of olfactory receptors. *Science* 307: 1638–1642.

Nansen, C., Macedo, T. B., Weaver, D. K., Peterson, R. K. D. 2005a. Spatiotemporal distributions of wheat stem sawfly eggs and larvae in dryland wheat fields. *Can Entomol*. 137: 428-440.

Nansen, C., Weaver, D. K., Sing, S. E., Runyon, J. B., Morrill, W. L., Grieshop, M. J., Shannon, C. L., and Johnson, M. L. 2005b. Within-field spatial distribution of *Cephus cinctus* (Hymenoptera: Cephidae) larvae in Montana wheat fields. *Can Entomol*. 137: 202-214.

Nelson, W. A. and Farstad, C. W. 1953. Biology of *Bracon cephi* (Gahan) (Hymenoptera: Braconidae), an important native parasite of the wheat stem sawfly, *Cephus cinctus* Nort. (Hymenoptera: Cephidae), in Western Canada. *Can Entomol*. 85: 103-107.

Neuhaus, E. M., Gisselmann, G., Zhang, W., Dooley, R., Stortkuhl, K., and Hatt, H. 2005. Odorant receptor heterodimerization in the olfactory system of Drosophila melanogaster. *Nat Neurosci*. 8: 15–17.

Ochieng, S. A., Park, K. C., Zhu, J. W., and Baker, T. C. 2000. Functional morphology of antennal chemoreceptors of the parasitoid Microplitis croceipes (Hymenoptera: Braconidae). *Arthro Struct Dev* 29: 231–240.

Party, V., C. Hanot, I. Said, D. Rochat, and M. Renou. 2009. Plant terpenes affect intensity and temporal parameters of pheromone detection in a moth. *Chem Senses* 34: 763-774

Peck, G. E. 2004. Potential semiochemicals of wheat (*Triticum aestivum* L.) induced by oviposition and feeding of the wheat stem sawfly, *Cephus cinctus* Norton (Hymenoptera: Cephidae) [Master's Thesis]. Montana State University, Bozeman, Montana. 65 p.

Perez, J., Weaver, D. K., and Morrill, W. L. 2006. Infestation of wheat and downy brome grass by wheat stem sawfly and subsequent larval performance. *Environ Entomol*. 35: 1279-1285.

Pichersky, E. and Gershenzon, J. 2002. The formation and function of plant volatiles: perfumes for pollinator attraction and defense. *Curr Opin Plant Biol*. 5: 237-243.

Pichersky, E., Noel, J. P., and Dudareva, N. 2006. Biosynthesis of plant volatiles: nature's diversity and ingenuity. *Science* 311: 808-811

Piesik, D., Weaver, D. K., Peck, G. E., and Morrill, W. L. 2006. Mechanically-injured wheat plants release greater amounts of linalool and linalool oxide. *J Plant Protection Res* 46: 29–39.

Piesik, D., Weaver, D. K., Runyon, J. B., Buteler, M., Peck, G. E., and Morrill, W. L. 2008. Behavioural responses of wheat stem sawflies to wheat volatiles. *Agric For Entomol*. 10: 245-253.

Preston, C. A., Laue, G., and Baldwin, I. T. 2004. Plant-plant signaling: application of trans- or cis-methyl jasmonate equivalent to sagebrush releases does not elicit direct defenses in native tobacco. *J Chem Ecol* 30: 2193–2214.

Quiroz, A., Pettersson, J., Pickett, J. A., Wadhams, L. J., Niemeyer, H. M. 1997. Semiochemicals mediating spacing behavior of bird cherry-oat aphid, *Rhopalosiphum padi*, feeding on cereals. Journal of Chemical Ecology 23: 2599–2607.

Quisenberry, S. S., Schotzko, D. J., Lamb, P. F., and Young, F. L. 2000. Insect distribution in spring pea-winter wheat-spring barley crop rotation system. *J Entomol Sci*. 35: 327-333.

Raguso, R. A. 2008. Start making scents: the challenge of integrating chemistry into pollination ecology. *Entomol Exp Appl*. 128(1): 196–207.

Reddy, G. V. P. and Guerrero, A. 2004. Interactions of insect pheromones and plant semiochemicals. *Trends Plant Sci*. 9: 253-261.

Reinecke, A., Ruther, J., Tolasch, T., Francke, W., and Hilker, M. 2002. Alcoholism in cockchafers: orientation of male *Melolontha melolontah* towards green leaf alcohols. *Naturwissenschaften* 89: 265-269.

Reitz, S. R., and Trumble, J. T. 1997. Effects of linear furanocoumarins on the hervibore *Spodoptera exiqua* and the parasitoid *Archytas marmoratus:* host quality and the parasitoid success. *Entomol Exp Appl*. 84: 9-16.

Robertson, H. M., Gadau, J., and Wanner, K. W. 2010. The insect chemoreceptor superfamily of the parastiod jewel wasp *Nasonia vitripennis*. *Insect Mol Biol*. 19: 121–136.

Robertson, H. M. and Wanner, K. W. 2006. The chemoreceptor superfamily in the honey bee, *Apis mellifera*: expansion of the odorant, but not gustatory, receptor family. *Genome Res* 16: 1395–1403.

Robertson, H. M., Warr, C. G., and Carlson, J. R. 2003. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 100: 14537–14542.

Roemhild, G. R. 1954. Morphological resistance of some of the Gramineae to the wheat stem sawfly *(Cephus cinctus* Norton) [Master's Thesis]. Montana State College, Bozeman, MT. 42 p.

Rogers, M. E., Jani, M. K., and Vogt, R. G. 1999. An olfactory-specific glutathione-S-transferase in the sphinx moth Manduca sexta. *J exp biol 202*(12): 1625–1637.

Roux, O., Van Baaren, J., Gers, C., Arvanitakis, L., and Legal, L. 2005. Antennal structure and oviposition behavior of the *Plutella xylostella* specialist parasitoid: *Cotesia plutella*e. *Micros Res Techn* 68: 36–44.

Rouyar, A., Party, V., Presern, J., Blejec, A., Renou, M. 2011. A general odorant background affects the coding of pheromone stimulus intermittency in specialist olfactory receptor neurons. *PLoS ONE* 6:e26443 10.1371/journal.pone.0026443.

Runyon, J. B., Hurley, R.L., Morrill, W.L., and Weaver, D. K. 2001. Distinguishing adults of *Bracon cephi* and *Bracon lissogaster*(Hymenoptera: Braconidae), parasitoids of the wheat stem sawfly (Hymenoptera: Cephidae). *Can Entomol*. 133: 215–217

Runyon, J. B., Morrill, W. L., Weaver, D. K., and Miller, P. R. 2002. Parasitism of the wheat stem sawfly (Hymenoptera: Cephidae) by *Bracon cephi* and *B. lissogaster* (Hymenoptera: Braconidae) in wheat fields bordering tilled and untilled fallow in Montana. *J Econ Entomol*. 95: 1130-1134.

Ruther, J., Reinecke, A., and Hilker, M. 2002. Plant volatiles in the sexual communication of *Melolontha hippocastani*: response towards time-dependent bouquets and novel function of (Z)-3-hexen-1-ol as a sexual kairomone. *Ecol Entomol*. 27.

Ruther, J., Reinecke, A., Tolasch, T., and Hilker, M. 2001. Make love not war: a common arthropod defense compound as sex pheromone in the forest cockchafer *Melolontha hippocastani*. *Oecologia* 128: 44-47.

Rytz, R., Croset, V., and Benton, R. 2013. Ionotropic receptors (IRs): chemosensory ionotropic glutamate receptors in Drosophila and beyond. *Insect Biochem Mol Biol*. 43:888-897.

Sakurai, T., Nakagawa, T., Mitsuno, H., Mori, H., Endo, Y., Tanoue, S., et al. 2004. Identification and functional characterization of a sex pheromone receptor in the silkmoth *Bombyx mori*. *Proc Natl Acad Sci USA* 101: 16653–16658.

Sato, K., Pellegrino, M., Nakagawa, T., Nakagawa, T., Vosshall, L. B., and Touhara, K. 2008. Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* 452: 1002–1006.

Schiestl, F. 2010. The evolution of floral scent and insect chemical communication. *Ecol Lett*. 13: 643–656.

Schneider, D. 1964. Insect Antennae. *Ann Rev Entomol* 9: 103-122.

Schröder, R. and Hilker, M. 2008. The relevance of background odor in resource location by insects: a behavioral approach. *BioScience* 58: 308-316.

Seamans, H. L., Manson, G. F., and Farstad, C. W. 1944. The effect of the wheat stem sawfly (*Cephus cinctus* Nort.) on the heads and grains of infested stems. *Ontario Entomol Soc Annu Rpt*. 75: 10-15.

Sing, S. E. 2002. Spatial and biotic Interactions of the wheat stem sawfly with wild oat and Montana dryland spring wheat. Ph.D. dissertation, Montana State University, Bozeman, Montana.

Skoog, F. E. and Wallace, L. E. 1963. Application of systemic insecticides as seed treatment to protect wheat plants against grasshoppers and wheat stem sawfly. *J Econ Entomol*. 57: 199-205.

Smart, R., Kiely, A., Beale, M., Vargas, E., Carraher, C., Kralicek, A. V., et al. 2008. *Drosophila* odorant receptors are novel seven transmembrane domain proteins that can signal independently of heterotrimeric G proteins. *Insect Biochem Mol Biol*. 38: 770–780.

Somsen, H. W. and Luginbill, P. 1956. Bracon lissogaster Mues. a parasite of the wheat stem sawfly. *USDA Technical Bulletin* No. 1153. pp. 1-7.

Steinbrecht, R. A. 1997. Pore structures in insect olfactory sensilla: a review of data and concepts. *Int J Insect Morphol Embryol*. 26: 229–245.

Steinbrecht, R. A., Ozaki, M., and Ziegelberger, G. 1992. Immunocytochemical localization of pheromone-binding protein in moth antennae. *Cell Tissue Res*. 270: 287–302.

Tegoni, M., Campanacci, V., and Cambillau, C. 2004. Structural aspects of sexual attraction and chemical communication in insects. *Trends Biochem Sci*. 29: 257–264.

Tscharntke, T., Thiessen, S., Dolch, R., Boland, W. 2001. Herbivory, induced resistance, and interplant signal transfer in Alnus glutinosa. *Biochem Syst Ecol*. 29: 1025–1047.

Turlings, T. C., Tumlinson, J. H., and Lewis, W.J. 1990. Exploitation of hervibore-induced plant odors by host-seeking parasitic wasps. *Science* 250: 1251-1253.

Van Baaren, J., Boivin, G., Le Lannic, J., and Nénon, J. P. 1999. Comparison of antennal sensilla of *Anaphes victus* and *A. listronoti* (Hymenoptera, Mymaridae), egg parasitoids of Curculionidae. *Zoomorphology* 119: 1–8.

Van Loon, J. J. A., de Boer, J. G., and Dicke, M. 2000. Parasitoid-plant mutualism: parasitoid attack of herbivore increases plant reproduction. *Entomologia Experimentalis et Applicata* 97: 219–227.

Visser, J. H. 1986. Host odor perception in phytophagous insects. *Ann Rev Entomol*. 31: 121-144.

Vogt, R. G. 2003. Biochemical diversity of odor detection: OBPs, ODEs and SNMPs. In Insect Pheromone Biochemistry and Molecular Biology G. J. Blomquist, R. G. Vogt (eds) pp391-446. Elsevier Academic Press.

Vogt, R. G. 2005. Molecular basis of pheromone detection in insects. In: L. Gilbert, K. Iatrou, S. Gill, editors. Comprehensive Insect Physiology, Biochemistry, Pharmacology and Molecular Biology. London: Elsevier. pp. 753–804.

Vogt R. G., Prestwich, G. D., and Lerner, M. R. 1991. Odorant-binding-protein subfamilies associate with distinct classes of olfactory receptor neurons in insects. *J Neurobiol*. 22: 74-84.

Vogt, R. G. and Riddiford, L. M. 1981. Pheromone binding and inactivation by moth antennae. *Nature* 293: 161–163.

Vogt, R. G., Riddiford, L. M., and Prestwich, G. D. 1985. Kinetic properties of pheromone degrading enzyme: the sensillar esterase of *Antheraea polyphemus*. *Proc Natl Acad Sci USA* 82: 8827–8831.

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A., and Axel, R. 1999. A spatial map of the olfactory receptor expression in the *Drosophila* antenna. *Cell* 96: 725-736.

Vosshall, L. B. and Hansson, B. S. 2011. A unified nomenclature system for the insect olfactory coreceptor. *Chem Senses* 36: 497–498.

Vosshall, L. B. and Stocker, R. F. 2007. Molecular architecture of smell and taste in Drosophila. *Annu Rev Neurosci.* 30: 505–533.

Vosshall, L. B., Wong, A. M., and Axel, R. 2000. An olfactory sensory map in the fly brain. *Cell* 102(2):147-59

Wahl, D. V., Shanower, T. G., and Hoelmer, K. A. 2007. A new species of *Collyria* Schiødte (Hymenoptera: Ichneumonidae: Collyriinae), a parasitoid of *Cephus fumipennis* (Hymenoptera: Cephidae) in China, and potential biological control agent for *Cephus cinctus* in North America. *J Kans Entomol*. Soc. 80: 43–50.

Wallace, L. 1962. Field-plot tests of chemicals for wheat stem sawfly control. *J Econ Entomol*. 55: 909-912.

Wallace, L. E. and McNeal, F. H. 1966. Stem sawflies of economic importance in grain crops in the United States. Agricultural Research Service United States Department of Agriculture (in Cooperation with Montana Agricultural Experimental Station), Technical Bulletin 1350.

Weaver D. K., Buteler, M., Hofland, M. L., Runyon, J. B., Nansen, C. et al. 2009. Cultivar preferences of ovipositing wheat stem sawflies as influenced by the amount of volatile attractant. *J Econ Entomol* 102: 1009–1017.

Weaver, D. K., Sing, S. E., Runyon, J. B., and Morrill, W. L. 2004. Potential Impact of Cultural Practices on Wheat Stem Sawfly (Hymenoptera: Cephidae) and Associated Parasitoids. Montana State University Department of Entomology. *J Agric Urban Entomol*. 21(4): 271-287.

Weiss, M. J. and Morrill, W. L. 1987. Influence of planting date and spring tillage on the wheat stem sawfly. *Mont Agr Res*. 4: 2-5.

Weiss, M.J. and Morrill, W. L. 1992**.** Wheat stem sawfly (Hymenoptera: Cephidae) revisited. *Am Entomol*. 241-245.

Wicher, D., Schafer, R., Bauernfeind, R., Stensmyr, M. C., Heller, R., Heinemann, S. H., and Hansson, B. S. 2008. *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature* 452: 1007–1011.

Wojtasek, H. and Leal, W. S. 1999. Conformational change in the pheromone-binding protein from Bombyx mori induced by pH and by interaction with membranes. *J Biol Chem* 274: 30950–30956.

Wojtasek, H. and Leal, W. S. 1999. Degradation of an alkaloid pheromone from the pale-brown chafer, *Phyllophertha diversa* (Coleoptera: Scarabaeidae), by an insect olfactory cytochrome P450. *FEBS Lett*. 458: 333–336.

Wurm, Y., Wang, J., Riba-Grognuz, O., Corona, M., Nygaard, S., Hunt, B. G., et al. 2011. The genome of the fire ant Solenopsis invicta. *Proc Natl Acad Sci USA* 108: 5679–5684.

Xu, P., Atkinson, R., Jones, D. N., and Smith, D. P. 2005. *Drosophila* OBP LUSH is required for activity of pheromone-sensitive neurons. *Neuron* 45: 193-200.

Yao, C. A., Ignell, R., Carlson, J. R. 2005. Chemosensory coding by neurons in the coeloconic sensilla of the *Drosophila* antenna*. J Neurosci*. 25(37): 8359-8367.

Yuan, X., Zhang, S., Zhang, Z., Kong, X., Wang, H., Shen, G., and Zhang, H. 2013. Antennal morphology and sensilla ultrastructure of the web-spinning sawfly *Acantholyda posticalis* Matsumura (Hymenoptera: Pamphiliidae). *Micron* 50: 20-28.

Zacharuk, R. Y. 1985. Antennae and sensilla. G.A. Kerkut, L.I. Gilbert (Eds.), Comprehensive Insect Physiology, Biochemistry and Pharmacology, Pergamon Press, London, pp. 1–70.

Zhang, G. J. and Xiao, G. X. 1992. Scanning electron microscopy studies of antennal sensilla of *Spalangia endius* and *Nasonia vitripennis*. *Entomol Knowledge* 35: 154–159.

Zhou, J. J., Zhang, G. A., Huang, W., Birkett, M. A., Field, L. M., Pickett, J. A., and Pelosi, P. 2004. Revisiting the odorant-binding protein LUSH of *Drosophila melanogaster*: evidence for odour recognition and discrimination. *FEBS Lett*. 558: 23–26.

CHAPTER THREE

ODORANT RECEPTORS OF A PRIMITIVE HYMENOPTERAN PEST, THE WHEAT STEM SAWFLY

Contribution of Authors and Co-Authors

Manuscript in Chapter 3

Author: Joanna C. Gress

Contributions: Conceived and implemented the study design and data collection. Analyzed the data and wrote the manuscript.

Co-Author: Hugh M. Robertson

Contributions: Laboratory collaborated in sequencing *C. cinctus* 454 transcriptome. Provided feedback on data analysis.

Co-Author: Mensur Dlakic

Contributions: Assembled 454 transcriptome; provided feedback on data analysis.

Co-Author: David K. Weaver

Contributions: Provided specimens for sequencing and comments on the manuscript.

Co-Author: Kevin W. Wanner

Contributions: Provided feedback for study design and data analysis and comments on the manuscript.

Manuscript Information Page

Joanna C. Gress, Hugh M. Robertson, David K. Weaver, Mensur Dlakić, and Kevin W. Wanner
Insect Molecular Biology
Status of Manuscript: (Put an x in one of the options below)
___ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-review journal
____ Accepted by a peer-reviewed journal
__X__ Published in a peer-reviewed journal

Published by the Royal Entomological Society
Insect Molecular Biology
Volume 22, Issue 6, pages 659-667 December 2013

Abstract

The wheat stem sawfly, *Cephus cinctus*, is an herbivorous hymenopteran that feeds exclusively on members of the Gramineae. Synanthropically, it has become one of the most important insect pests of wheat grown in the Northern Great Plains region of the United States and Canada. Insecticides are generally ineffective due to an extended adult flight period and the inaccessible larval stage that feeds within the wheat stems, making it virtually intractable to most pest management strategies. While research towards integrated pest management strategies based on insect olfaction has proved promising, nothing is known about the molecular basis of olfaction in this important pest species. In this study we identified 28 unique odorant receptor (Or) transcripts from an antennal transcriptome. A phylogenetic analysis with the predicted Ors from the honeybee and jewel wasp genomes revealed at least four clades conserved amongst all three Hymenoptera species. Antennal expression levels were analyzed using quantitative real-time PCR and one male-biased and five female-biased Ors were identified. This study provides the basis for future functional analyses to identify behaviorally active odors that can be used to help develop olfactory mediated pest management tools.

Introduction

The wheat stem sawfly (WSS), *Cephus cinctus* Norton (Hymenoptera: Cephidea), has been recognized as one of the most important insect pests of wheat, *Triticum aestivum* L. (Cyperales: Poaceae), in the northern Great Plains of the United States and Canada for almost a century (Weiss & Morrill, 1992). Annual economic losses caused by this insect are estimated to be more than $350 million US dollars in the affected region (Beres et al., 2011). Larvae overwinter in the below ground portion of dead stems and the adults emerge in late spring to attack developing wheat crops. Males emerge earlier, and after mating, females begin searching for suitable host plants for oviposition (Ainslie, 1929; Weiss et al., 1992). Females usually lay one egg within a wheat stem per visit, reportedly in the uppermost developing internode (Holmes & Peterson, 1960), but there can be multiple eggs deposited in a stem by different females (Buteler et al., 2009). Larvae feed internally on the parenchyma and vascular tissue, damaging the stem and reducing kernel weight up to 35% (Holmes, 1977; Delaney et al., 2010). As the wheat plant matures, the larva moves down to the base of the stem where it cuts a notch at ground level, causing the plant to fall over before harvest and creating further yield loss (Beres et al., 2007). The cut stem is plugged with frass and the prepupa forms a hibernaculum within the stubble in preparation for overwintering (Criddle, 1923).

Conventional management strategies have not provided adequate control of this important pest. Planting tolerant varieties that have solid rather than hollow stems represents the most common tactic, but its success has been limited because of inconsistent performance and reduced yield in the absence of pest pressure (Weaver et

al., 2004; Beres et al., 2011). Insecticides generally are ineffective (Knodel et al., 2009) because the immature stage feeds within the stem and is not exposed to foliar sprays, and systemic insecticidal seed treatments do not persist long enough. An extended adult flight period often lasting for more than four weeks makes foliar sprays targeting the adult stage unfeasible (Beres et al., 2011). In response, integrated pest management (IPM) strategies based on insect behavior and chemical ecology has been developed. Recent studies have demonstrated that host plant green leaf volatiles such as (Z)-3-hexenyl acetate, (Z)-3-hexenol and 6-methyl-5-hepten-2-one provide cues for female oviposition (Piesik et al., 2008) and that the amounts of these emitted by plants may play a role in oviposition (Weaver et al., 2009). A single pheromone compound, 9-acetyloxynonanal, has been identified from both sexes of the WSS and demonstrated to be attractive to both sexes in field bioassays (Cossé et al., 2002). This compound is among a suite of electrophysiologically active compounds arising from oxidation of cuticular components in male and female WSS (Bartelt et al., 2002; Cossé et al., 2002). Attractive odors can be used in IPM programs for monitoring population densities and adult emergence periods as well as for trap crop strategies (Piesik et al., 2008).

Olfaction mediates many important insect pest behaviors, including locating host plants, mates and oviposition sites. Odors in the environment are detected by odorant receptors (Or) expressed in the dendrite membrane of olfactory neurons that extend into the lymph filled interior of sensilla located primarily on the antennae. Ors are a large diverse gene family primarily responsible for the molecular recognition of odors in the insect's environment (reviewed in Rutzler & Zwiebel, 2005; Vosshall & Stocker, 2007;

and Leal, 2012). Insect Ors, first identified from the *Drosophila melanogaster* genome, represent a novel form of chemoreceptor that function as ligand-gated ion channels (Clyne et al., 1999; Sato et al., 2008; Wicher et al., 2008; Smart et al., 2008). The Or co-receptor (Orco) is a highly conserved insect Or that acts as a chaperone and dimerization partner for other Ors that impart ligand specificity (Benton et al., 2006; Vosshall & Hansson, 2011). Together Orco+Or_x form the ligand-gated ion channel. Supporting its critical function as a partner, olfaction in mutant fruit flies lacking Orco is broadly impaired (Larsson et al., 2004).

Ors have been annotated from the sequenced genomes of several insect species, and range in number from 62 *Or* genes in *D. melanogaster* to 265 in *Tribolium castenum* (Robertson et al., 2003; Engsontia et al., 2008; Richards et al., 2008). In several cases, specific Ors that detect behaviorally important odors have been identified, including the receptor for queen substance 9-oxo-2-decenoic acid in *Apis mellifera* (Wanner et al., 2007b), the bombykol sex pheromone receptor in *Bombyx mori* (Sakurai et al., 2004) and the receptor in *A. gambiae* that responds to components of human odor (Carey et al., 2010). *A. mellifera* and *Nasonia vitripennis*, also in the order Hymenoptera, have 174 and 225 *Or* genes respectively (Robertson & Wanner, 2006; Robertson et al., 2010). Remarkably, annotations from recently sequenced ant genomes have revealed as many as 407 Or genes (Zhou et al., 2012). In this study we sequenced the antennal transcriptome of the WSS and identified 28 unique *Or* transcripts and analyzed their phylogenetic relationship within the Hymenoptera. Ors have been annotated from six hymenopteran species that all belong to the monophyletic “true” wasp suborder Apocrita (Grimaldi &

Engel, 2005; Zhou et al., 2012). Sawflies, members of the basal, paraphyletic suborder Symphyta (Grimaldi & Engel, 2005), are primitively phytophagous and considerably diverged from the more modern Apocrita wasp and bee lineages, suggesting they have divergent Ors. We identified one male-biased and five female-biased *Ors* by assaying their antennal expression levels using quantitative real-time PCR (qPCR). This study provides the basis to prioritize WSS *Ors* for future functional analysis towards identifying behaviorally active odors that can be used to help develop olfactory mediated pest management strategies.

Experimental procedures

Insects and RNA extraction

Adult male and female wheat stem sawflies were collected from infested wheat fields near Conrad or Amsterdam MT during July 2009. Antennae were dissected from male and female adults within 24h of collection along with headless bodies. All tissues were frozen on dry ice and stored at -80°C. For transcriptome sequencing, total RNA was extracted from 2,000 antennal pairs collected from both male and female insects. For gene expression studies antennae were collected in three batches each consisting of 130-175 male and 130-175 female antennal pairs. Total RNA was extracted from frozen tissues using a Dounce homogenizer and an RNeasy Mini Kit (Qiagen, Valenica, CA) and quantified and assayed for purity by absorbance at 260nm, 280nm and 230nm using a Nanodrop 1000 Spectrophotometer (Thermo Scientific, Waltham, MA).

Antennal transcriptome sequencing and Or identification

cDNA was prepared by the University of Illinois Urbana-Champaign W.M. Keck Center for Comparative and Functional Genomics from 200µg of Total RNA isolated from pooled male and female antennae. The cDNA was sequenced using a Roche 454 GS-FLX instrument and the sequence reads assembled into contigs using miraEST Assembler (Chevreux et al., 2004). A FASTA file of the non-redundant contigs was formatted as a BLASTable database and searched using a PC version of standalone BLAST (Altschul et al., 1997). A subset of *A. mellifera* and *D. melanogaster* Or sequences were used as queries in tBLASTn searches to identify contigs with homology to known insect odorant receptors. The contigs were translated to identify the longest region with homology to insect Ors. Protein sequences of at least 250 amino acids were used for phylogenetic and gene expression analysis. These protein sequences were aligned using MEGA 5 (Tamura et al., 2011) and ClustalW (Larkin et al., 2007) and imported into Jalview (Waterhouse et al., 2009). TMpred was used to predict the number of transmembrane domains in the full length protein Or sequences (Hofmann & Stoffel, 1993).

Phylogenetic analysis of WSS Ors

Protein sequences were combined into a multiple alignment file with a representative subset of *N. vitripennis* and *A. mellifera* Ors for phylogenetic analysis in MAFFT using a BLOSUM62 matrix (Katoh et al., 2005) and manually adjusted to minimize gaps in Jalview (Clamp et al., 2004; Waterhouse et al., 2009). The alignment

was imported into BEAUTI and BEAST (Drummond & Rambaut, 2007) for Bayesian phylogenetic analysis. A consensus tree was created using default settings with a gamma 8 base with relaxed clock (Drummond et al., 2006). The consensus tree was inferred from 30 million iterations and 20% of trees were burned in. Trees were rooted using Orco (Or83b family) orthologs as the outgroup. Posterior probabilities are provided at branch points, in some cases followed by bootstrap values from a second analysis of a Maximum likelihood tree based on the JTT matrix-based model in MEGA 5 (Tamura et al., 2011).

Gene expression

The expression of 25 *Ors* and *Orco*, normalized to two different ribosomal protein encoding genes (*RPS3* and *RL31*), was assayed by qPCR. The primer sets for three of 28 Or transcripts failed. Genomic DNA was digested from Total RNA used for gene expression with the TURBO DNA-free kit (Applied Biosystems, Foster City, CA). cDNA was synthesized from 1µg of Total RNA using Superscript III Reverse Transcriptase (Invitrogen) and 50µM anchored Oligo(dT)20 primer and incubated for 50 min at 50°C followed by inactivation at 85°C for 5 min. qPCR primers for the 25 *Or* genes, *Orco* and two control genes were designed using Primer3 open source software with the following criteria: primers 15-30 bp in length, annealing temperatures 58-60°C, an amplicon size between 75-120 nt, and a G/C content around 50% (Supplementary Table 1). Each primer set was validated by calculating standard curves with a 10X serial dilution of WSS cDNA as template (three technical replicates for each template dose). The threshold cycle (Ct) was plotted against the log of the template dilution and primers with slopes ranging from 3.0 to 3.7 were used, where a slope of -3.33 corresponds to 100% efficiency.

qPCR experiments were performed using 100µL strip cap tubes (Qiagen, Valenica, CA), a 72 well disk (Qiagen), the RotorGeneQ (Qiagen) and SsoFast EvaGreen Supermix (Bio-Rad, Hercules, CA). Each 20µl reaction was done in triplicate. Cycling conditions were as follows: 95°C for 1 minute, 45 cycles of 95°C for 7 seconds, 57°C for 45 seconds, followed by melting temperature analysis: 64°C-94°C hold 5 seconds for each degree. Baseline cycle and threshold values were calculated automatically using default settings. No-template and no-reverse transcriptase controls were included in each experiment.

qPCR data was analyzed using the relative ΔΔCt method (Livak & Schmittgen, 2001). The threshold cycle (Ct) was calculated for each sample using default parameters. ΔCt values were obtained by subtracting the Ct value of the *Or* gene from the Ct value of the endogenous control (*RPS3* or *RL31*) for each tissue tested (Ct endogenous control – Ct target Or). ΔΔCt values were calculated by subtracting female ΔCt values from male ΔCt values for each *Or* (ΔCt male target Or - ΔCt female target Or). Finally, the relative amount of transcript in male and female antennae was calculated using $2^{-\Delta\Delta Ct}$ for each endogenous control gene.

Results

28 complete or near full length Or transcripts from WSS antennae

To identify candidate WSS Ors an antennal transcriptome was created by high-throughput pyrophosphate sequencing of antennal cDNA. Approximately 500,000 reads averaging 360 bp were obtained and assembled into 55,115 non-redundant contigs

(unpublished results). A total of 205 cDNA sequences with amino acid homology to known *A. mellifera* and *D. melanogaster* Ors were identified, but most were incomplete fragments that did not overlap. Complete or near full length Ors, as judged by amino acid alignment with select published *A. mellifera* and *N. vitripennis* Ors (Fig. 1). The obligate co-receptor and ortholog of DmelOrco was named CcinOrco according to Vosshall and Hansson (2011). A total of 28 complete, or near complete Ors (250 residues or larger), were identified from the antennal transcriptome (Fig. 1, GenBank accession no. KC778499 - KC778527). Suffixes were used to denote partial sequences: N (N-terminus missing), C (C-terminus missing) and E (internal exon missing). The 28 candidate Ors have BLAST similarity to published Or sequences. Full length CcinOrs also possess typical features of the family, including 7 or 8 predicted transmembrane domain motifs and conserved residues at the C-terminus (Fig. 1; Robertson et al., 2010).

The remaining 177 transcripts that did not encode proteins greater than 250 amino acids in length were aligned to a set of preliminary scaffolds from an ongoing whole genome sequencing project. The transcripts mapped to 98 unique scaffold positions, suggesting that the WSS genome could encode as many as 127 Or genes.

The 28 WSS Ors were analyzed with a subset of 76 *N. vitripennis* and *A. mellifera* Ors, selected to represent major lineages published in Robertson et al. (2010), for patterns within the Hymenoptera. The phylogenetic tree (Fig. 2) is characterized by at least four groups of homologous receptors each represented by all three species. The WSS Ors were named using an alphanumeric system based upon their location within clades 1-4 (1a-d, 2a-j, 3a-l and 4a-b, Fig. 2). This naming convention was selected to make it easier to

adapt to complete repertoire of Ors that will be annotated from a whole genome sequencing project currently under way. The first clade includes CcinOrs 1a-d that group with several *N. vitripennis* and *A. mellifera* Ors with 64% posterior probability. The second clade includes CcinOrs 2a-j that group with six *A. mellifera* Ors and three *N. vitripennis* Ors with 70% posterior probability. All three hymenopteran species are also represented in clades 3 and 4 that have 80% and 99% posterior probability. In general the overall amino acid identity of the 28 WSS Ors was diverse, ranging from 12 to 92% and averaging 20%. One cluster of Ors showing greater identity and conservation, CcinOrs 3e-g, 3i and 3j, may represent an example of an expansion through gene duplication after divergence of the WSS lineage (Fig. 2). Within this cluster, CcinOr3h and NvitOr288 share 41% identity, an example of two more highly conserved Ors considering the phylogenetic divergence of these two species. CcinOr3c groups with 100% confidence with four Nvit Ors that represent an expanded lineage of more than 20 Nvit Ors located on a single genomic scaffold, a lineage that has apparently been lost from the honeybee genome (Robertson et al., 2010). Consistent with their essential role as co-receptors, the Orco branch is the only lineage to demonstrate a more simple orthologous relationship (Fig. 2).

One male-biased and five female-biased Ors

The expression level of 25 WSS *Ors*, normalized to the control genes *CcinRPS3* and *CcinRL31* and averaged over 3 biological replicates, was determined in male and female antennae by qPCR (Table 1). Gene expression values were arbitrarily categorized based on untransformed Ct values, into very high (<17), moderately high to moderate

(17-21), and low (>21) groups. The most abundant transcripts were expressed 1000-fold greater relative to the least abundant, a difference of 10 or more Ct cycles. Consistent with its functional role, *CcinOrco* was highly expressed equally in male and female antennae, at levels comparable to the control genes *RPS3* and *RL31*. *CcinOr2a* was expressed at the highest levels, similar to that of Orco. *CcinOr2i* was among the five most highly expressed transcripts and also exhibited sex bias, expressed 15 times higher in male compared to the female antennae (Table 1.) Five Or transcripts were expressed at levels 5-15 times higher in female compared to male antennae, *CcinOrs 1a*, $1d_{[CN]}$, $2e_{[C]}$, $3d_{[C]}$ and *4b*. Sexual differences in expression levels were all statistically significant (Table 1). Expression of CcinOrs in body tissues such as legs and abdomens was not detected at significant levels (data not shown). Negative controls using Total RNA and water as template did not produce expression signals.

Discussion

In this study we identify 28 Ors from an antennal transcriptome of the WSS, the most important insect pest of wheat grown in the northern Great Plains region of the United States and Canada. Both sexes of the WSS produce and respond behaviorally to the pheromone 9-acetyloxynonanal (Cosse et al., 2002). CcinOr2i, highly expressed and 15 times greater in male compared to female antennae, is a candidate receptor for this pheromone. However, since 9-acetyloxynonanal is attractive to both sexes, it may be detected by one of the Ors abundant in both sexes, such as Or2a. If Or2i does not detect 9-acetyloxynonanal it would suggest that additional odors that mediate male-specific behaviors remain to be identified.

Several green leaf volatiles emitted by wheat plants have been identified as oviposition cues for female sawflies (Piesik et al., 2008). Ors expressed highly in the antennae of both sexes may detect host plant odors that are generally used to locate their habitat. However, CcinOrs 1a, $1d_{[CN]}$, $2e_{[C]}$, $3d_{[C]}$ and 4b that are expressed 5-15 times higher in female antennae may detect host plant odors specifically used for host recognition for oviposition.

The phylogenetics of the insect Or family is characterized by extensive lineage-specific gene duplication and gene loss. With the advent of whole genome sequencing and the analysis of several insect genomes, large regions of clustered duplicated Or genes have been identified (Robertson and Wanner, 2006; Robertson et al., 2010; Richards et al., 2008). While similar examples of sawfly specific gene duplication were evident in the phylogenetic analysis, the 28 WSS Ors identified from the antennal transcriptome were surprisingly diverse and exhibited homology with many of the more modern hymenopteran Apocrita lineages represented by *A. mellifera* and *N. vitripennis*. This pattern suggests that the ancestor of sawflies and modern Hymenoptera possessed a diversified repertoire of *Or* lineages that were not lost in the more modern species of Hymenoptera.

Hymenoptera genomes have revealed large numbers of Or genes, 174 and 225 from the honeybee and jewel wasp genomes, for example (Robertson et al., 2010). Mapping the partial Or EST fragments to a preliminary assembly of an ongoing WSS genome project suggest this species may have as many as 127 Or genes. This may reflect the basal position of sawflies in the Hymenoptera phylogeny and/or less complex life

histories, however, a complete annotation of the genome will be required to determine the final complement of Or genes.

The potential of "reverse chemical ecology", using the proteins that detect odors in peripheral olfactory system to identify behaviorally active compounds, has been demonstrated by Leal et al. (2008). The discovery of the first insect *Or* sequences from the partially sequenced *D. melanogaster* genome (Clyne et al., 1999; Vosshall et al., 1999) and the demonstration of their preeminent role in detecting and discriminating odors (Hallem et al., 2004) has facilitated progress in non-model pest insect species. Gene expression levels have been used to identify and prioritize candidate Ors that detect behaviorally important odors. Typically, Ors that are expressed predominantly in the antennae of one sex (sex-biased expression) respond to odors that mediate behaviors specific to that sex. Lepidoptera sex pheromones, produced by females to attract males for mating, represent one of the best examples. Several moth sex pheromone receptors have now been functionally characterized, and most are expressed at higher levels in the male antennae (Krieger et al., 2004; Nakagawa et al., 2005; Miura et al., 2010; and Wanner et al., 2010, for example). This approach also proved effective in identifying AmelOr11 as the receptor for the queen honeybee pheromone (Wanner et al., 2007b) and for identifying female-biased silkworm Ors that detect host plant volatiles or male produced pheromones (Wanner et al., 2007a; Anderson et al., 2009). However, in some cases, insect pheromones are attractive to, and can be detected by, both sexes. Rather than using sex-biased expression, Mitchell et al. (2012) took advantage of the fact that sex pheromone receptors also tend to be highly expressed relative to other *Ors*. A total of 57

Or transcripts were identified from the antennae of the cerambycid beetle *Megacyllene caryae*, and by functionally testing only the five most common transcripts the authors identified two pheromone receptors.

The next step in this project is to clone and functionally screen candidate Ors against known insect and host plant odors. Highly expressed or sex-biased Ors that cannot be deorphanized using these known odors will become candidates for reverse chemical ecology approaches to identify candidate ligands. Field testing of candidate odors for behavioral activity (Leal et al., 2008) will represent the final goal of this research.

Acknowledgements

We thank Drs. Aurélien Mazurie and Charles Carey for assistance with the statistical analysis of the qPCR experiments and Kate McInnerney for assistance with the qPCR experimental set up. We also thank Dr. Marcella McClure for assistance with the Bayesian analysis. HMR acknowledges USDA grant 2008-35302-18815 and KWW and DKW acknowledge the Montana Wheat and Barley Committee for funding support.

Conflicts of Interest

None

Table 3.1. Sex-biased expression of Or genes in male and female wheat stem sawfly antennae determined by quantitative real-time PCR. Fold Difference was calculated using the $2^{-\Delta\Delta Ct}$ equation, Ct values are averages of three biological replicates. Or expression was normalized using two different endogeneous control genes, RPS3 or RL31. $\Delta\Delta Ct = \Delta Ct_{male} - \Delta Ct_{female}$ where $\Delta Ct_{male} = Ct_{male\ Control} - Ct_{male\ Or}$ and $\Delta Ct_{female} = Ct_{female\ Control} - Ct_{female\ Or}$. All negative fold difference values were transformed to their inverse value. Positive values represent higher expression in male antennae compared to females; Ors with more than five-fold sex-bias are bolded. Results are ranked by their untransformed expression levels; Ct values < 17 were considered very high, 17-20 moderately high to moderate, and >21 low.

	Untransformed Ct Values			
Or No.	Male Antennae (Mean ± SEM)	Female Antennae (Mean ± SEM)	Fold Difference ($2^{-\Delta\Delta Ct}$; RPS3)	Fold Difference ($2^{-\Delta\Delta Ct}$; RL31)
CcinOr2a	14.5 ±0.4	16.6 ±0.4	2.6	2.9
CcinOrco	15.9 ±0.3	16.1 ±0.4	-1.1	-1.0
CcinOr2h[$_{CN}$]	17.3 ±0.3	17.6 ±0.4	1.1	-1.2
CcinOr3l[$_{E}$]	17.4 ±0.5	19.1 ±0.6	2.1	-5
CcinOr2i	18.1 ±1.0	22.7 ±0.8	**14.7**	**16.2**
CcinOr3g	18.3 ±0.6	20.0 ±0.6	1.9	2.1
CcinOr3h	18.3 ±0.4	18.6 ±0.6	-1.3	-1.2
CcinOr2j[$_{N}$]	18.8 ±0.4	18.6 ±0.3	-2	-1.2
CcinOr3f	19.0 ±0.6	18.3 ±0.4	-2.7	-2.4
CcinOr3c	19.0 ±0.3	18.5 ±0.1	-2.4	-2.1
CcinOr3b[$_{C}$]	19.1 ±0.6	17.4 ±0.3	-5.5	2.3
CcinOr2g[$_{N}$]	19.2 ±0.2	19.2 ±0.5	-1.2	-1.5
CcinOr3a[$_{C}$]	19.3 ±0.4	19.0 ±0..6	-2	-1.8
CcinOr2d[$_{N}$]	19.9 ±0.4	19.7 ±0.7	-1.8	-1.7
CcinOr1c	20.5 ±0.5	19.3 ±0.6	-3.7	-3.3
CcinOr2c[$_{N}$]	20.5 ±0.7	20.8 ±1.1	-1.3	-1.2
CcinOr4b	21.2 ±0.5	19.5 ±0.4	**-5**	**-4.5**
CcinOr1a	21.2 ±0.2	18.9 ±0.3	**-7.8**	**-7.1**
CcinOr2e[$_{C}$]	22.7 ±0.3	19.4 ±0.6	**-15.8**	**-14.3**
CcinOr2b	21.1 ±0.1	21.3 ±0.6	-1.4	-1.3
CcinOr4a	21.6 ±0.8	20.5 ±1.0	-3.5	-3.2
CcinOr3e	22.1 ±0.5	20.9 ±0.2	-3.7	-3.4
CcinOr1b	22.1 ±0.1	20.6 ±0.6	-2.5	-4.2
CcinOr1d[$_{CN}$]	22.6 ±0.6	20.8 ±0.4	**-5.7**	**-5.1**
CcinOr3d[$_{C}$]	22.7 ±0.3	20.6 ±0.1	**-6.6**	**-6**
CcinOr2f[$_{N}$]	24.3 ±1.1	26.6 ±1.2	2.9	3.2

Sex-biased expression is supported by ANOVA of normalized Ct values, P=0.036, 0.003, 0.014, 0.007, 0.028, and 0.043, CcinOrs 2i, 4b, 1a, 2e, 1d and 3d, respectively.

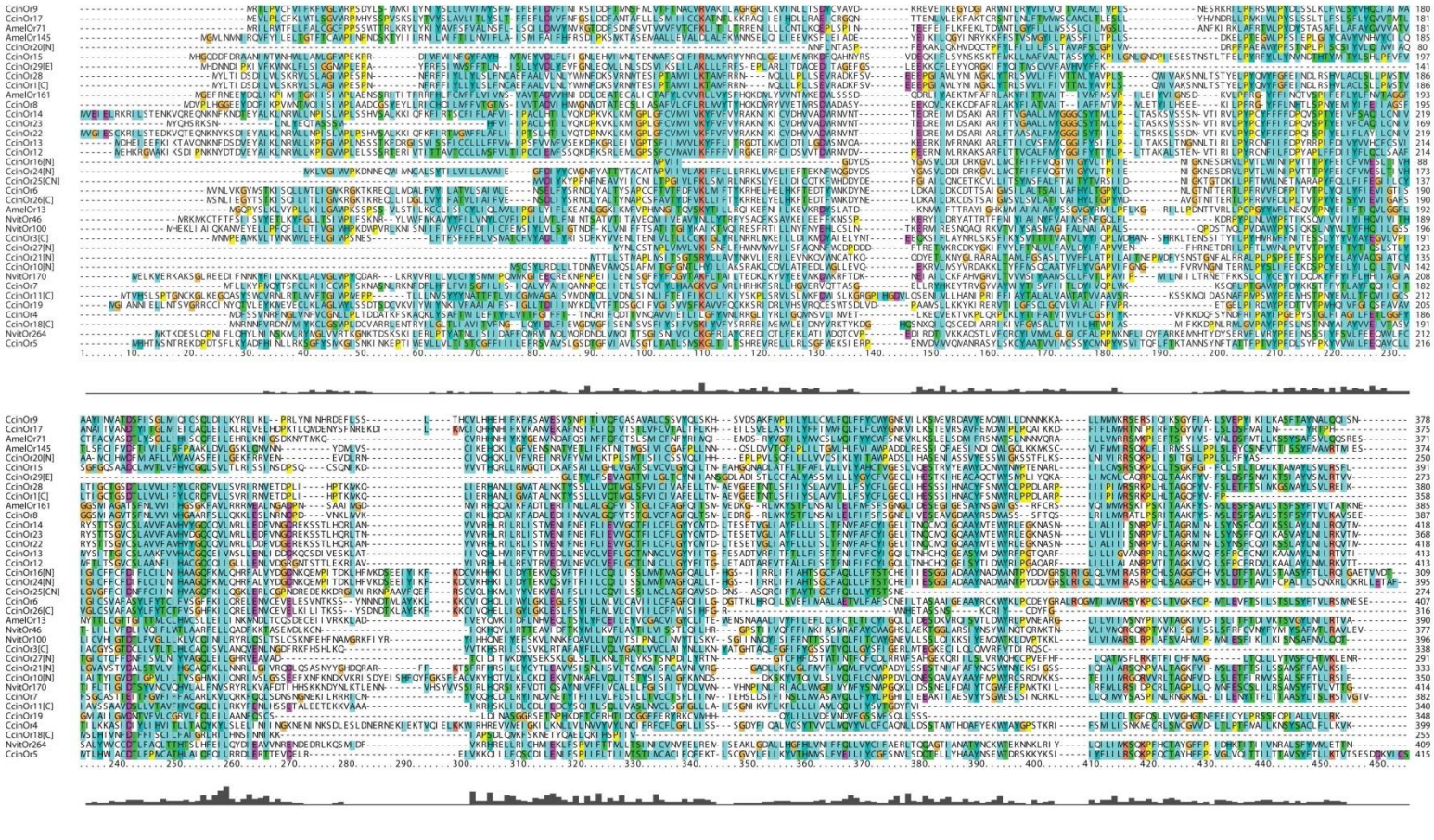


Figure 3.1. Amino acid sequence alignment of WSS Ors annotated from the antennal transcriptome. A total of 25 WSS Ors were aligned with the published sequences of four full-length jewel wasp and honeybee Ors using ClustalX set to default parameters, and the alignment was used to judge completeness of WSS Ors. Suffixes were used to denote partial sequences: N (N-terminus missing), C (C-terminus missing) and E (internal exon missing).

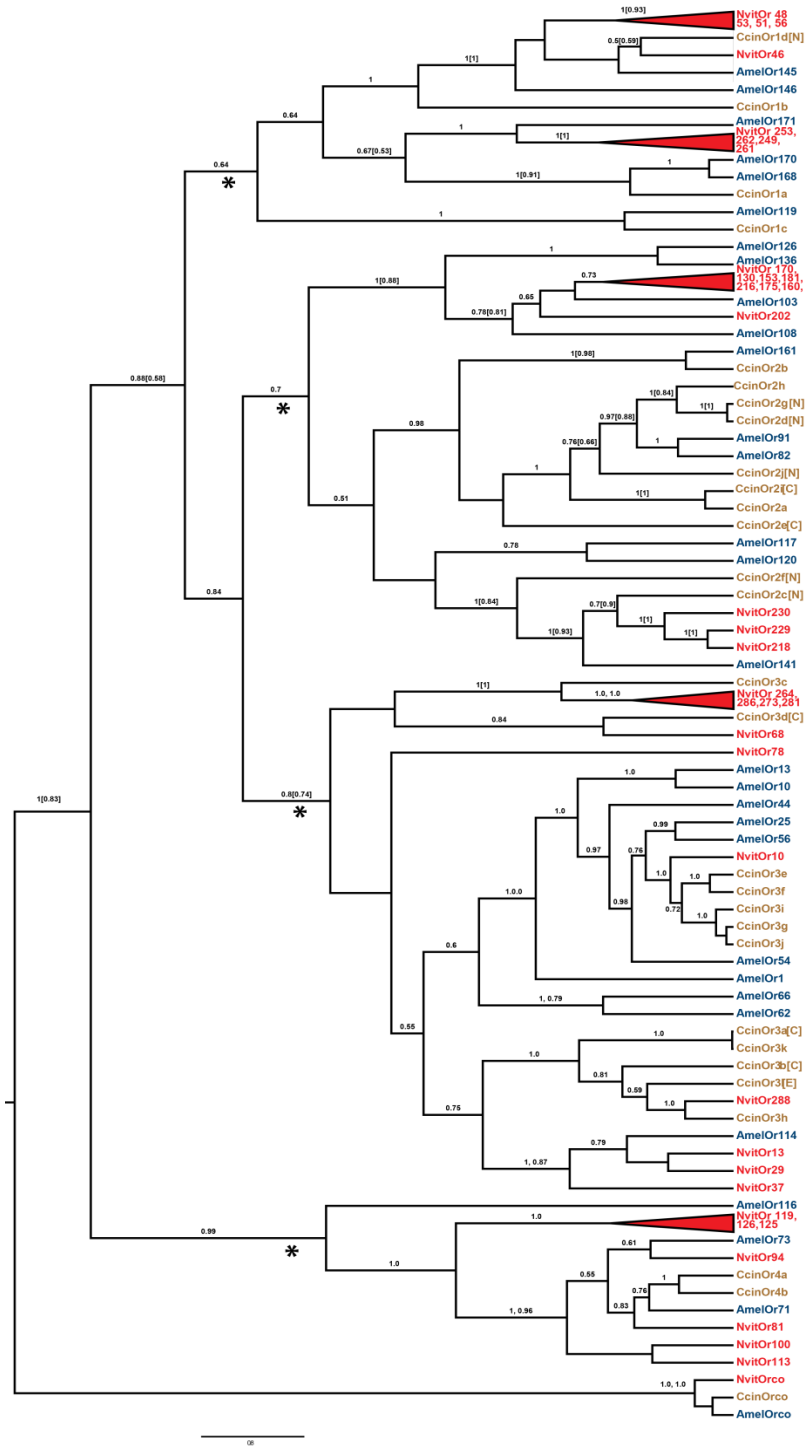


Figure 3.2. Phylogenetic relationships of 25 WSS Ors with representative jewel wasp and honeybee Ors inferred from Bayesian analysis with posterior probabilities values provided at significant branch points. The tree was rooted using WSS, jewel wasp and honeybee Orco sequences as the outgroup based on their basal position in the Or family (Robertson et al., 2003). Bootstrap values obtained from a Maximum likelihood analysis are provided in brackets after Bayesian values to provide additional branch support. Jewel wasp, red font; honeybee, blue font; and wheat stem sawfly, brown font. An asterisk marks the base of clades 1-4.

Supplementary Table 3.1. Wheat Stem Sawfly qPCR primer sequences for all 25 Ors, Orco and both control genes. The slope and R^2 value for each Or standard curve are listed.

CcinOr	qPrimer Forward	qPrimer Reverse	Slope	R^2
Orco	ACCGATCATCTGAGGGACAC	CCCACCATTGAAATTCATCC	3.17	.996
1	GGGATCGGATACATTGTTGG	TTTAAAGCTGTGGCCATTCC	3.24	.960
3	TTGGTGACACTGACGCTACA	GCAAACGCGGTACGTAACTT	3.52	.981
4	CGTCTACGTCTACCTGAATTGC	TGAGCGCAAAAGCATAACAC	2.95	.961
5	TCAATCTTGCGACATTCTGG	CCACGAAAAGCGACATCATA	3.36	.985
6	TTGTACTTTATTGAGGTGATTGGAA	TTCTAGTTCTACCTCGCACACG	3.37	.962
7	TGTTCAACAACACTGCGACA	TGGAGTCCAATGAATGTTCG	3.36	.996
8	GAGCAGCGAGGTTCAGCTT	CAGCGCCACTACGTTTATGA	3.58	.987
9	ATCAAGCTGCCACGTTTGTA	GAGCATAATGCAACGGCACT	3.26	.986
10	TCTTTCGAAGCCTGTGTCAA	CTTCGATTTGTCGGAATCGT	3.39	.994
11	CCTCTTGGGTATCCGTTCGAG	ATCCCAATTGACCGCAAAC	3.18	.928
12	GTTGTCATCGTGCGACACTT	GCAGTGTCAGCTGTCTCCAA	3.61	.963
13	TGGACAGTGCGAAATCGTTA	TAGTGCAGCCCAAAAATTCC	3.90	.997
14	GCAGCTTGGCTGTCGTATTT	GCCTAAGAATCCGCAAGTGA	3.77	.990
15	TAGCGACCCGTCACAATGT	TATCCGAGCACACAAACCAA	3.60	.972
16	GCAGGATTTCAAGCACTTC	TCAGCGTTGTAAGCAGCATC	3.29	.988
17	TGCATTCAGCATCACAATCA	CACACTGGCCAATTCAACAC	3.91	.992
18	TGTTGGAGTTCCTGCATATT	TCCTATGGCAAAAAGGTGC	3.27	.989
19	CCGTCTTCGTTCTTTGTCCT	GCCTCCGCAGTCTATTGTGT	3.25	.985

Supplementary Table 3.1

CcinOr	qPrimer Forward	qPrimer Reverse	Slope	R^2
20	GTTCAGCTGATTCACCACGA	TGATTCGTAAACCGACGAAG	3.45	.991
21	GATCTCATTCTGCGCTGTGA	AAAAGCAAACGCAATGTTGT	3.54	.982
24	GATGACCATGGCAGGATTC	ATCGGCAATACCACCACTTT	3.29	.975
25	CTCCAAGGCAAATTGGAAAG	GCTTCGACCTTTTCCACGTA	3.21	.936
26	TACTGGGATGCCGTAGGAAC	GAATGCTACGCTGCACAGAC	3.69	.956
27	CCATGAAGGATTACGTGTCG	CGCAGCTCACAAAACTGGAA	3.45	.963
29	GTGTGGTGTTCGCTGTTCAC	TACAGCGCAAAGCAGCATAA	3.59	.974
RPS3	GTGGTCAACGTGAAAATCA	CCAGTTTCATCACAAGGAAGC	3.41	.990
RL31	GCACCACGTGCAATTAAAGA	CATCATCGTTCCTTCTTCTGC	3.34	.995

References

Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman D.J. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucl Acids Res* **25**: 3389–3402.

Anderson, A. R., Wanner, K. W., Trowell, S. C., Warr, C. G., Jaquin-Joly, E., Zagatti, P., Robertson, H. and Newcomb, R. D. 2009. Molecular basis of female-specific odorant responses in Bombyx mori. *Insect Biochem Mol Bio* **39**: 189–197.

Anislie, C.N. 1929. The Western Grass-Stem Sawfly: A Pest of Small Grains. *U.S. Dept of Agric Techn Bul No* **57**: 1–24.

Bartelt, R. J., Cossé, A. A., Petroski, R. J. and Weaver, D. K. 2002. Cuticular hydrocarbons and novel alkenediol diacetates from wheat stem sawfly (*Cephus cinctus*): Natural oxidation to pheromone components. *J Chem Ecol* **28:** 385–406.

Benton R., Sachse S., Michnick S.W. and Vosshall L.B. 2006. Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo. *PLoS Biol* **4**: e20.

Beres, B. L., Cárcamo, H. A. and Byers, J. R. 2007. Effect of wheat stem sawfly damage on yield and quality of selected Canadian spring wheat. *J Econ Entomol* **100:** 79–87.

Beres, B.L., Dosdall, L.M., Weaver, D.K., Spaner, D.M. and Cárcamo, H.A. 2011. The biology and integrated management of wheat stem sawfly, *Cephus cinctus* (Hymenoptera: Cephidae), and the need for continuing research. *Can Entomol* **143**: 105–125.

Buteler, M., Weaver, D. K. and Peterson, R. K. D. 2009. Exploring the oviposition behavior of the wheat stem sawfly when encountering plants infested with cryptic conspecifics. *Environ Entomol* **38**: 1707–1715.

Carey, A., Wang, G, Su, C., Zwiebel, L.J. and Carlson, J.R. 2010. Odourant reception in the malaria mosquito *Anopheles gambiae*. *Nature* **464**: 66–71.

Chevereux, B., Pfisterer, T., Drescher, B., Driesel, A.J., Müller, W.E.G., Wetter, T. and Suhai, S. 2004. Using the miraEST Assembler for Reliable and Automated mRNA Transcript Assembly and SNP Detection in Sequenced ESTs. *Genome Res* **14**: 1147–1159.

Clamp, M., Cuff, J., Searle, S. M. and Barton, G. J. 2004. The Jalview Java Alignment Editor. *Bioinformatics* **20**: 426–467.

Clyne, P.J., Warr, C.G., Freeman, M.R., Lessing, D., Kim, J. and Carlson, J.R. 1999. A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. *Neuron* **22**: 327–338.

Cossé, A.A., Bartlet, R.J., Weaver, D.K. and Zilkowski, B.W. 2002. Pheromone components of the wheat stem sawfly: Identification, electrophysiology, and field bioassay. *J Chem Ecol* **28**: 407–423.

Criddle N. 1923. Popular and Pratical Entomolgy: The Life Habits of *Cephus Cinctus* Nort. in Manitoba. *Can Entomol* **55**: 1–4.

Delaney, K. J., Weaver, D. K. and Peterson, R. K. D. 2010. Wheat photosynthesis and yield reductions: wheat stem sawfly (Hymenoptera: Cephidae) herbivory interacts with abiotic conditions and host plant resistance. *J Econ Entomol* **103**: 516–524.

Drummond , A.J., Ho, S.Y.W., Phillips, M.J. and Rambaut, A. 2006. Relaxed Phylogenetics and Dating with Confidence. *PLoS Biology* **4**: e88.

Drummond, A.J. and Rambaut, A. 2007. BEAST: Bayesian evolutionary analysis by sampling trees. *BMC Evol Biol* **7**: 214.

Engsontia, P., Sanderson, A.P., Cobb, M., Walden, K.K.O., Robertson, H. and Brown, S. 2008. The red flour beetle's large nose: an expanded odorant receptor gene family in *Tribolium casteneum*. *Insect Biochem Mol Biol* **38**: 387–397.

Grimaldi, D. and Engle, M.S. 2005. Evolution of the insects, Cambridge University Press. pp 755.

Hallem, E.A., Ho, M.G. and Carlson, J.R. 2004. The molecular basis of odor coding in the Drosophila antenna. *Cell* **117**: 965–979.

Hofmann, K. and Stoffel, W. 1993. TMbase-A database of membrane spanning segments. *Biol Chem Hoppe-Seyler* **374**: 166

Holmes, N.D. 1977. The effect of the wheat stem sawfly, *Cephus cinctus* Norton (Hymenoptera: Cephiadae), on the yield and quality of wheat. *Can Entomol* **109**: 1591–1598.

Holmes, N.D. and Peterson, L.K. 1960. The influence of the host on oviposition by the wheat stem sawfly *Cephus cinctus* Nort. (Hymenoptera:Cephidae). *Can J Plant Sci* **40**: 29–46.

Katoh, K., Kuma, K., Toh, H. and Miyata, T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucl Acids Res* **33**: 511–518.

Knodel, J. J., Beauzay, P. B., Eriksmoen, E. D. and Peterson J. D. 2009. Pest management of wheat stem maggot (Diptera: Chloropidae) and wheat stem sawfly (Hymenoptera: Cephidae) using insecticides in spring wheat. *J Agri Urban Entomol* **26**: 183–197.

Krieger, J., Grosse-Wilde, E., Gohl, T., Dewer, Y. M. E., Raming, K. and Breer, H. 2004. Genes encoding candidate pheromone receptors in a moth (Heliothis virescens) *Proc Natl Acad Sci USA* **101**: 11845-–11850.

Larkin, M.A, Blackshields, G., Brown, N.P., Chenna, R. McGettigan, P.A., McWilliam, H., *et al*. 2007. ClustalW and ClustalX version 2. *Bioinformatics* **23**: 2947–2948.

Larsson, M., Domingos, A. I., Jones, W. D., Chiappa, M. E., Amrein, H., and Vosshall, L. B. 2004. Or83b encodes a broadly expressed odorant receptor essential for *Drosophila* olfaction. *Neuron* **43**: 703–714.

Leal, W.S. 2012. Odorant Reception in Insects: Roles of Receptors, Binding Proteins, and Degrading Enzymes. *Annu Rev Entomol* Sep 27. [Epub ahead of print]

Leal, W.S., Barbosa, R.M., Xu, W., Ishida, Y., Syed, Z., Latte, N., *et al*. 2008. Reverse and conventional chemical ecology approaches for the development of oviposition attractants for *Culex* mosquitoes. *PLoS One* **3**: e3045.

Livak. K.J. and Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**: 402–408.

Mitchell, R.F., Hughes, D.T., Luetje, C.W., Millar, J.G., Soriano-Agatón, F., Hanks, L.M., *et al*. 2012. Sequencing and characterizing odorant receptors of the cerambycid beetle *Megacyllene caryae*. *Insect Biochem Mol Biol* **42**: 499–505.

Miura, N., Nakagawa, T., Touhara, K. and Ishikawa, Y. 2010. Broadly and narrowly tuned odorant receptors are involved in female sex pheromone reception in Ostrinia moths. *Insect Biochem Mol Biol* **40**: 64–73.

Nakagawa, T., Sakurai, T., Nishioka, T. and Touhara, K. 2005. Insect sex-pheromone signals mediated by specific combinations of olfactory receptors. *Science* **307**: 1638–1642.

Piesik, D., Weaver, D.K., Runyon, J.B., Buteler, M., Peck, G.E. and Morrill, W.L. 2008. Behavioural responses of wheat stem sawflies to wheat volatiles. *Agric Forest Entomol* **10**: 245–253.

Richards, S., Gibbs, R.A., Weinstock, G.M., Brown, S.J., Denell, R., Beeman, R.W. *et al.* 2008. The genome of the model beetle and pest *Tribolium castaneum*. *Nature* **452**: 949–955.

Robertson, H.M., Gadau, J. and Wanner, K.W. 2010. The insect chemoreceptor superfamily of the parastiod jewel wasp *Nasonia vitripennis*. *Insect Mol Biol* **19**: 121–136.

Robertson, H.M. and Wanner, K.W. 2006. The chemoreceptor superfamily in the honey bee, *Apis mellifera*: expansion of the odorant, but not gustatory, receptor family. *Genome Res* **16**: 1395–1403.

Robertson H.M., Warr, C.G. and Carlson, J.R. 2003. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **100**: 14537–14542.

Rutzler, M. and Zwiebel, L. J. 2005. Molecular biology of insect olfaction: recent progress and conceptual models. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* **191**: 777–790.

Sakurai, T., Nakagawa, T., Mitsuno, H., Mori, H., Endo, Y., Tanoue, S., *et al*. 2004. Identification and functional characterization of a sex pheromone receptor in the silkmoth *Bombyx mori*. *Proc Natl Acad Sci USA* **101**: 16653–16658.

Sanchez-Gracia, A., Vieira, F. G. and Rozas, J. 2009. Molecular evolution of the major chemosensory gene families in insects. *Heredity* **103**: 208–216.

Sato, K., Pellegrino, M., Nakagawa, T., Nakagawa, T., Vosshall, L.B. and Touhara, K. 2008. Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **452**: 1002–1006.

Smart, R., Kiely, A., Beale, M., Vargas, E., Carraher, C., Kralicek, A.V., *et al*. 2008. *Drosophila* odorant receptors are novel seven transmembrane domain proteins that can signal independently of heterotrimeric G proteins. *Insect Biochem Mol Biol* **38**: 770–780.

Tamura ,K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol* **28**: 2731–2739.

Vosshall L.B., Amrein H., Morozov P.S., Rzhetsky A. and Axel, R. 1999. A spatial map of olfactory receptor expression in the *Drosophila* antenna. *Cell* **96**: 725–736.

Vosshall, L.B. and Hansson, B.S. 2011. A unified nomenclature system for the insect olfactory coreceptor. *Chem Senses* **36**: 497–498.

Vosshall, L.B. and Stocker, R. F. 2007. Molecular architecture of smell and taste in Drosophila. *Annu Rev Neurosci* **30**: 505–533.

Wahl, D.V., Shanover, T.G. and Hoelmer, K.A. 2007. A new species of *Collyria Sciodte* (Hymenoptera: Ichneumonida: Collriinae), a parasitoid of *Cephus fumipennis* (Hymenoptera: Cephidae) in China, and potential biological control agent for *Cephus cinctus* in North America. *J Kans Entomol Soc* **80**: 43–50.

Wanner, K.W., Anderson, A.R., Trowell, S.C., Theilmann, D.A., Robertson, H.M. and Newcomb, R.D. 2007a. Female-biased expression of odor receptor genes in the adult antennae of the silkworm, *Bombyx mori*. *Insect Mol Biol* **16**: 107–19.

Wanner, K.W., Nichols, A. S., Allen, J. E., Bunger, P. L., Garczynski, S.F., Linn, C. E., *et al*. 2010. Sex Pheromone Receptor Specificity in the European Corn Borer Moth, Ostrinia nubilalis. *PloS One* **5**: e8685.

Wanner, K.W., Nichols, A.S., Walden, K.K., Brockmann, A., Luetje C.W. and Robertson, H.M. 2007b. A honey bee odorant receptor for the queen substance 9-*oxo*-2-decenoic acid. *Proc Natl Acad Sci USA* **104**: 14383–14388.

Waterhouse, A.M., Procter, J.B., Martin, D.M.A, Clamp, M. and Barton, G. J. 2009. Jalview Version 2 - a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**: 1189–1191.

Weaver, D. K., Buteler, M., Hofland, M. L., Runyon, J. B., Nansen, C., Talbert, L. E., *et al*. 2009. Cultivar preferences of ovipositing wheat stem sawflies as influenced by the amount of volatile attractant. *J Econ Entomol* **102**: 1009–1017.

Weaver, D. K., Sing, S. E., Runyon, J. B. and Morrill, W.L. 2004. Potential impact of cultural practices on wheat stem sawfly (Hymenoptera: Cephidae) and associated parasitoids. *J Agri Urban Entomol* **21**: 271–287.

Weiss, M.J., Morrill, L.W. and Reitz, L.L. 1992. Wheat stem sawfly (Hymenoptera: Cephidae) revisted. *Am Entomol* **38**: 241–245.

Wicher, D., Schafer, R., Bauernfeind, R., Stensmyr, M.C., Heller, R., Heinemann, S.H. and Hansson, B.S. 2008. *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature* **452**: 1007–1011.

Zhou, X., Slone, J.D., Rokas, A., Berger, S.L., Liebig, J., Ray, A., *et al*. 2012. Phylogenetic and transcriptomic analysis of chemosensory receptors in a pair of divergent ant species reveals sex-specific signatures of odor coding. *PLoS Genet* **8**: e1002930.

CHAPTER FOUR

OLFACTORY-RELATED GENES ANNOTATED FROM THE WHEAT STEM SAWFLY, *CEPHUS CINCTUS*

Contribution of Authors and Co-Authors

Manuscript in Chapter 4

Author: Joanna C. Gress

Contributions: Conceived and implemented the study design and data collection. Collected and analyzed data. Wrote the manuscript.

Co-Author: Hugh M. Robertson

Contributions: Laboratory collaborated in sequencing *C. cinctus* genome and transcriptome. Provided feedback on data analysis.

Co-Author: Kim Walden

Contributions: Collaborated in sequencing and assembly of *C. cinctus* genomic data. Provided feedback on data analysis.

Co-Author: Kevin W. Wanner

Contributions: Provided feedback on study design, data analysis, and comments on the manuscript.

Manuscript Information Page

Joanna C. Gress, Hugh M. Robertson, Kim Walden, and Kevin. W. Wanner
Insect Molecular Biology
Status of Manuscript: (Put an x in one of the options below)
__X_ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-review journal
____ Accepted by a peer-reviewed journal
____ Published in a peer-reviewed journal

Published by the Royal Entomological Society

Abstract:

Next generation DNA sequencing has revolutionized basic and applied genetic studies of biological organisms. The wheat stem sawfly, *Cephus cinctus*, is one of the most economically-damaging insect pests of wheat grown in the Northern Great Plains region of North America. Olfactory-related proteins expressed in the peripheral sensory system mediate many important pest behaviors, such as mate and host-seeking and selection. The genes encoding these proteins represent new targets for novel pest control strategies. A transcriptome from male and female WSS antennae along with whole genome sequences was used to annotate 131 olfactory-related genes from the WSS, including 11 odorant binding proteins (OBPs), 8 chemosensory proteins (CSPs), 53 odorant receptors (ORs), 10 ionotropic receptors (IRs), 12 carboxylesterases (CCEs), 8 glutathione S-transerases (GSTs) and 29 cytochrome P450s (P450s). The annotated genes form the basis for functional studies to identify critical olfactory genes that mediate important pest behaviors.

Introduction

Insects use their olfactory system to detect chemical cues in their environment, to search for habitat, food, mates and hosts for oviposition. In the peripheral sensory system, odorant molecules first interact with binding proteins that transport hydrophobic odors across the aqueous sensillum lymph to the olfactory receptor neurons (ORNs). Among these, odorant binding proteins (OBPs) and chemosensory proteins (CSPs) are proposed to bind general compounds like host volatiles (Rutzler & Zwibel, 2005; Vogt, 2003), although the role of CSPs in chemoreception remains unclear. Pheromone binding proteins (PBPs), a subfamily of OBP, specifically transport insect pheromones used in chemical communication within a species. After crossing the lymph, odorant molecules interact with receptors embedded within the dendrite membrane of the ORN. Two distinctly different families of olfactory receptors have been described in insects, the odorant receptors (ORs) and the ionotropic receptors (IRs) (Nakagawa & Vosshall, 2009). Research using the model organism *Drosophila melanogaster* has demonstrated that these two families detect different classes of odor molecules (Benton et al., 2006; Larrson et al., 2004; Sato et al., 2008; Wicher at al., 2008). While ORs are seven transmembrane domain protein receptors, they are specific to insects, and have an inverted membrane topology compared to G-protein coupled olfactory receptors described from vertebrate animal species (Benton et al., 2006). OR genes are very divergent in their sequences, within the same species or between different species. ORs are novel olfactory receptors in that they function as a heterodimer with a highly

conserved OR co-receptor (ORco) that forms a ligand-gated ion channel (Vosshall & Hanson 2011). IRs form an evolutionarily distinct family of chemosensory receptors that are far more ancient compared to the ORs, being conserved throughout protostome species lineages (Croset et al., 2010). They are related to ionotropic glutamate receptors but harbor a divergent ligand-binding domain (Benton et al., 2009). Like the ORs, they function as ion channels and form heterodimers with conserved co-receptors (Silbering et al. 2011). IRs were first identified from *D. melanogaster* where they were found to play a role in food odor detection as well as detecting odors involved in mating and reproduction (Benton et al., 2009; Grosjean et al., 2011).

While OBPs, CSPs, ORs and IRs transport and detect odor molecules in the environment, biotransformation enzymes expressed specifically in the sensillum lymph are responsible for odor degradation. These odor-degrading enzymes (ODEs) can keep the sensillum lymph free of deleterious odors, thus reducing background noise and improving signal detection. ODEs also terminate olfactory signals by degrading the odor after it has activated receptors on the dendrite membrane, rapidly removing odorant molecules from the vicinity of the ORs to allow the detection of new stimuli (Leal, 2013). ODEs are members of large conserved gene families involved in metabolic and detoxification pathways, that have evolved roles in olfaction, including carboxylesterases (CES), glutathione S-transferases (GSTs) and cytochrome P450s (CYPs). Collectively, these genes involved in the transport, detection and degradation of odors can be termed olfactory-related genes.

Olfactory-related genes have been described from several Hymenoptera species including the honeybee and jewel wasp (Robertson & Wanner, 2010; Robertson et al. 2010; Foret et al. 2007, Vieira et al. 2012), as well as several ant species (Wurm et al. 2010; Smith et al. 2011; Smith et al. 2011b), by annotating tissue-specific transcriptomes or whole genome sequences. Because the olfactory system mediates many important pest behaviors, olfactory-related genes have become a new molecular target for the development of alternative pest control strategies (Wanner et al. 2007). The wheat stem sawfly (WSS), *Cephus cinctus*, is one of the most damaging insect pests of the Northern Great Plains region (Beres et al. 2011). Recently, Gress et al. (2013) identified 28 ORs from antennal transcripts of the WSS, using 454 pyrosequencing technology (Chapter 3). When this project was initiated in 2010, 454 pyrosequencing was the state of the art high throughput technology for transcriptome sequencing. Prior to 454 pyrosequencing, automated Sanger sequencing was the standard method of transcriptome sequencing. 454 pyrosequencing raised the sequencing throughput dramatically, from an estimated 166 kb /hr for automated Sanger sequencing to 20-30 Mb / hr for pyrosequencing (Morozova and Marra, 2008). Pyrosequencing utilizes “sequencing by synthesis”, where DNA polymerase is monitored one base at a time by the activity of a chemiluminescent enzyme as it synthesizes the strand complementary to the template DNA (Ronaghi et al. 1998). Original pyrosequencers produced reads averaging 250 bp in length, but this was soon increased to 500 and 1000 bp read lengths, now producing about 700 Mb of sequence information in a single run. In 2010, we obtained approximately 500,000 reads averaging 360 bp in length (180 Mb sequence) from transcripts isolated from WSS antennae (Gress

et al. 2013). Next generation sequencing (NGS) has evolved quickly, and Illumina technology has now become the state of the art high throughput sequencing technology. Illumina read lengths are shorter (50-300 bp), but as many as 3 billion reads are produced in a single run, producing as much as 1.8 Tb (10^{12}) of sequence data.

Illumina sequencing technology has dramatically reduced the cost of DNA sequencing, and whole genome sequencing (wgs) can now be considered affordable basic molecular biological knowledge of an insect. In collaboration with the University of Illinois Urbana-Champaign the wgs of the WSS was sequenced from the genomic DNA isolated from a single male insect (Robertson et al., unpublished results). Illumina technology provides considerably more sequencing coverage depth of transcriptomes compared to the pyrosequencing technology. Total RNA isolated from male and female WSS antennae was sequenced using Illumina technology. Antennal transcriptome data in combination with genomic data was used to annotate all olfactory related genes from the WSS. A total of 131 candidate olfactory-related genes were annotated, including 26 new ORs, 10 new IRs, 11 new OBPs, 8 new CSPs and 48 ODEs. With access to the WSS wgs the genomic location and gene structure of the candidate olfactory-related genes was also mapped, producing a high quality annotation. This investigation is the first to be conducted on sawfly and provides the molecular bases to better understand *C. cinctus* olfaction.

Materials and Methods

RNA Preparation and Illumina Transcriptome Sequencing

Adult male and female wheat stem sawflies were collected from infested wheat fields near Amsterdam MT during July 2012. Antennae were dissected from male and female adults within 24h of collection along with whole bodies. Larvae were collected from wheat and smooth brome near Conrad and Moccasin, MT in 2011. All tissues were frozen on dry ice and stored at -80°C. For transcriptome sequencing, total RNA was extracted from 500 antennal pairs collected from both male and female insects respectively. Total RNA was extracted from male and female WSS antennae using a Dounce homogenizer and purified using an RNeasy Mini Kit (Qiagen, Valencia, CA). RNA quality was verified using a 2100 Bioanalyzer RNA Nanochip (Agilent, Santa Clara, CA) and samples had RNA Integrity Number (RIN) values higher than 8.5. The RNA was quantified using a NanoDrop ND-1000 Spectrophotometer (NanoDrop, Wilmington, DE).

Illumina paired end sequencing was performed at the University of Illinois Urbana-Champaign W.M. Keck Center for Comparative and Functional Genomics from 20 ug of male and 20 ug of female antennal RNA. First, reads with adaptors or reads containing more than 5% unknown nucleotides (Ns) were removed computationally from the data set. Secondly, low-quality reads containing more than 20% suspect-nucleotides of Phred Quality Score less than 10 were filtered out. Finally, both ends of each individual read were evaluated for three successive suspect-nucleotides and these were trimmed. Each cleaned sequence dataset from male and female antenna was assembled

separately into non redundant contigs using Trinity r2012-06-08 software (De novo assembly using paired reads mode and default parameters) (Grabherr et al. 2011). The Trinity outputs were clustered by TGICL (Pertea et al. 2003). The consensus cluster sequences and singletons composed the unigenes Illumina dataset used in this study. The WSS genome was sequenced in collaboration with the University of Illinois Urbana-Champaign. Genomic DNA (gDNA) from a single haploid male sawfly collected near Amsterdam, MT was used to sequence 500bp and 1.5kb shotgun read libraries. gDNA from additional male sawflies was used to sequence 3kb and 5 kb mate-pair read libraries.

Annotating Olfactory-Related Genes

The unigenes Illumina dataset was formatted into a BLASTable database a PC version of standalone BLAST (Altschul et al. 1997). Olfactory-related proteins identified from Hymenoptera (*Apis mellifera* and *Nasonia vitripennis*) were downloaded from the NCBI Genbank (http://www.ncbi.nlm.nih.gov/genbank) and used as queries for tBLASTn searches of the formatted unigenes Illumina dataset. WSS antennal transcripts with homology to known olfactory-related proteins were extracted from the unigenes Illumina dataset and curated into a new FASTA file list. This file was used as a query to search the NCBI nonredundant protein database (20.03.2012) using BLASTX, with a 0.005 p value threshold. Blast matches were noted and used to evaluate the completeness of the open reading frame (ORF) of the candidate olfactory-related transcripts. Each olfactory-related transcript was mapped to a unique position within the WSS genomic scaffolds. The genomic regions containing candidate olfactory-related genes were

extracted along with an extra 1000 bp upstream and downstream of the predicted region. A FASTA file of these genomic regions was created for use with manual annotation to identify the complete ORF of genes whose transcripts were not complete. In addition, the genomic regions were used to create gene models with exon-intron boundaries. ORFs were translated to peptides using the BioPHP DNA to Protein translation tool (BioPHP.org). OBPs and CSPs were searched for the presence of a signal peptide using SignalP 4.1 (Peterson et al. 2011) and transmembrane domains of novel candidate ORs and IRs were predicted using TMpred (Hofmann & Stoffel, 1993).

Results

Approximately 71,800,000 reads were obtained from male antennae and assembled into 70, 253 non-redundant contigs averaging 1,022 base pairs (bp) in length (range, 100 – 27,873) (Table 4.1). For female antenna 84,000,000 reads were obtained and assembled into 82,862 non-redundant contigs averaging 1,014 bp (range, 100 - 32,742) (Robertson et al. unpublished results). The 320,000,000 reads obtained from genomic DNA isolated from a single haploid male were assembled into 27,447 scaffolds averaging 6,067 bp (scaffold N50 = 598,740 bp; longest scaffold = 4,355,184 bp) (NCBI Genbank #s KB465430.1 - KB467404.1). These sequence resources were combined with an antennal transcriptome previously reported from the WSS (Gress et al. 2013) (Table 4.1) and collectively used to annotate the coding regions and gene models of olfactory-related genes.

Annotating *C. cinctus* Olfactory-Related Genes

Using an iterative process 131 olfactory-related transcripts and genes were identified and annotated from the WSS. First, the antennal transcriptome was used to identify olfactory-related transcripts (Table 4.2). In many cases the transcripts were aberrant or incomplete. Aberrant transcripts included concatenated genes, frame shifts, non-translating DNA, miss-spliced intron-exon boundaries, missing exons and incomplete sequences at the 5' prime and 3' prime transcript ends. Such errors can result from incomplete processing of RNA transcripts prior to RNA isolation, RNA degradation and/or incorrect assembly of the sequence reads into contigs. Transcriptomes sequenced from RNA pooled from many different individuals can contain polymorphisms as a result of multiple alleles originating from a single genomic locus. These polymorphisms can interfere with accurate contig assembly. Here, the availability of the WGS sequence was an invaluable resource for producing high quality gene annotations. In an iterative process, transcripts were aligned to genomic regions to identify gene models (Table 4.3) and to identify full length and accurate ORFs of the olfactory-related proteins.

OBPs and CSPs

Eleven OBPs and 8 CSPs were annotated from the WSS transcriptome and genome sequences. The OBPs were named CcOBP1-7, 9, 10, 12 & 13. The CSPs were named CcCSP1, 2a, 2b and 3-7 (Table 4.2). All 11 OBPs had significant BLAST similarity with OBP sequences from other species previously deposited in Genbank. Most shared more than 40% amino acid identity with OBPs reported from other Hymenoptera species. Their ORFs ranged from 393-570 bp coding for proteins with 130-173 amino

acids (Table 4.3). All OBPs were judged to be complete sequences and all included a predicted secretory signal peptide. The OBP genes have 4-7 exons (Table 4.3). The 8 CSP ORFs ranged from 393 – 570 bp in length, encoding proteins with 103 – 138 amino acids that shared 45 – 82% identity with other CSPs already present in Genbank. All CSPs were judged to be complete and all included a predicted secretory signal peptide (Table 4.3). All of the CSP genes included 2 exons, with the exception of CcCSP4 & 6 that included 3 exons.

ORs and IRs

Fifty-three OR and 10 IR genes were annotated from the WSS transcriptome and genome sequences. The 52 ORs include the conserved OR co-receptor (ORco) and 25 ORs previously reported by Gress et al. (2013). They were named CcOR1, 3-6, 7-9, 11-21, 32, 34, 36-54, 56, 58-60 & ORco. The ORFs of most ORs are close to 1,200 bp encoding proteins with a typical length of ~400 amino acids with BLAST similarity to other ORs (Table 4.2). The first 23 ORs listed in Table 4.3 share high amino acid identity with the previously reported WSS ORs that have been entered in Genbank (Gress et al. 2013). The remaining ORs in Table 4.2 have moderate sequence identity of 30-66% to ORs reported from other Hymenoptera species. The longest ORF was 1,269 bp, with the exception of the conserved ORco, that has a unique insertion compared to all other ORs. The ORco ORF is 1,443 bp long. ORs do not have predicted signal peptides, but commonly have 6 or 7 predicted transmembrane domains (Table 4.3). Almost all of the ORs have 5-7 exons. Interestingly, CcOR31 & 52 have four exons. The ORco gene has 8 exons. All but 9 of the OR genes were judged to be complete (Table 4.3). Six of 10 IRs

(CcIR2, 3, 7, 8, 11 &15) were judged to be complete genes; CcIR1, 9 and 17 are incomplete at the N-terminus while CcIR4 is missing part of the C-terminal end (Table 4.3). The ORFs of the complete IRs ranged from 1,482 – 2,808 bp in length, encoding proteins with 493 – 935 amino acids with significant similarity to other insect IR genes (Table 4.2). All of the complete IRs have three or four predicted transmembrane domains, no predicted signal peptide and anywhere from four to 12 exons.

Odorant Degrading Enzymes (ODEs)

A total of 49 candidate ODEs were annotated from the WSS antennal transcriptome and genomic sequences, belonging to three different protein families. Twelve carboxylesterase (CCE) genes were annotated and named CcCCE1-4, 6, 7-9 & 10a-10d. Eight glutathione s-transferase (GST) genes were annotated and named CcGST1a-1c, 2 & 4-7. Finally, 29 cytochrome P450 (P450) genes were annotated and named CcP450_1-3, 5-28, 30 & 31 (Table 4.2).

Of the 12 CCE genes annotated, 10 were judged to be complete. The ORFs of the complete genes ranged from 1,575 – 1,788 bp in length, encoding proteins ranging from 524 – 595 aa in length. The proteins all have good BLAST similarity with CCEs from other insect species and did not have predicted signal peptides or transmembrane domains (Table 4.2). The gene models were comprised of 5 to 11 exons.

Seven of 8 GSTs were complete gene models, CcGST2 is incomplete at the N-terminus. The ORFs range from 618 – 729 bp in length, encoding proteins with 205 – 242 amino acids (Table 4.2). Signal peptides and transmembrane domains were not predicted,

and the proteins exhibited high sequence identity with other insect GSTs. Six of 8 GSTs have more than 60% aa identity with GSTs from other Hymenoptera species. The GST genes all contain 4 or 5 exons (Table 4.3).

Twenty-one of 29 P450 gene models were complete. Five are incomplete at the N-terminus while three are incomplete at the C-terminus. The open reading frames range from 1,017 – 1,794 bp in length encoding 338 – 597 amino acids. All but three of the P450s have more than 50% amino acid identity with P450s from other Hymenoptera (Table 4.2). No signal peptide or transmembrane domain sequences were predicted. Varying numbers of exons comprised the different P450 gene models, anywhere from one to 12 (Table 4.3).

Discussion

High throughput next generation sequencing has transformed the scientific discipline of genetics. In many cases the WGS of an organism has simply become another aspect of its basic biology. One current challenge of the genomics revolution is bridging the knowledge gap between DNA sequence and organismal biology or in other words, gaining the analytical ability to predict aspects of an organism's biology simply from its genetic blueprint. The first step in this process is the annotation of genes encoded within the genome. While the field of bioinformatics has significantly improved the ability to predict and model genes from a genome computationally, manual curating is still required to produce high quality annotations. A second challenge of the genomics revolution is the use of genetic information to solve applied problems to benefit society. Much of this effort has focused on human genomics and its application to medicine,

including the concept of personalized care based on an individual's genome sequence. In entomology, insect genomes are forming the genetic basis to understand pest biology, and to develop new insect control strategies. One approach is to target insect genes that interact with its environment, such as gut genes involved in digesting host plants. A recent example is the development of transgenic corn plants that express double stranded RNA that inhibits a gene in the gut of its insect pest, the corn rootworm (Baum et al. 2007). In this present study, 131 olfactory-related genes are annotated from WSS antennal transcriptome and whole genome sequences. These annotations will help provide a genetic understanding of WSS olfaction, a sensory system that mediates many critical pest behaviors.

The peripheral olfactory system of the WSS includes proteins that transport odors across the sensillum lymph to the olfactory neuron membrane (OBPs and CSPs), proteins that span the membrane of the olfactory neuron dendrite and that initiate nerve impulses by detecting odors (ORs and IRs), and odor degrading enzymes that clear the sensillum lymph (CCEs, GSTs and P450s). A total of 131 olfactory-related genes were annotated from the WSS, including 11 OBPs, 8 CSPs, 53 ORs, 10 IRs, 12 CCEs, 8 GSTs and 29 P450s. Of these 107 genes (82%) were judged to encode complete ORFs. This level of completeness was achieved by iterative manual curating that used experimental expression data (gene transcripts) to model genes encoded in genomic scaffolds. Genes with fewer introns that encode proteins with conserved sequences produce better annotations. Genes with more introns that encode divergent proteins are the most difficult to annotate. Gene annotation largely relies on the use of homologous proteins from

related species to detect protein-coding exons. These sequence similarity searches are constrained by the reference species used and their phylogenetic distance from the species of interest. WSS is located basally in the phylogeny of the Hymenoptera and is separated by more than 200 million years of evolutionary time from the more derived Apocrita, that include the jewel wasp *Nasonia vitripennis* and the honeybee *Apis mellifera*. Further, the accuracy of this process is negatively related to the size and number of exons, and the degree of protein sequence conservation. Small exons with low sequence conservation are inherently more difficult to annotate compared to larger exons with high sequence conservation. The protein-coding exons, often separated in the genome by long non-coding introns, must be assembled in the correct order and with the correct exon-intron splice junctions that produce an in-frame ORF. Genes with many exons are more prone to mistakes in annotation. OR genes, for example, encode divergent protein sequences, and often have 5-7 exons. The first and last exons of OR genes are typically very short, encoding only 10-20 amino acids (Wanner and Robertson, 2010). The last exon can be annotated easily because it encodes a highly conserved amino acid domain, whereas the first exon can be very difficult to detect since the N-terminus of ORs is highly divergent.

Transcriptomes are powerful tools for gene annotation, since they provide direct experimental evidence that can identify and delineate exons, introns and their splice junctions. However, transcriptomes are limited by experimental constraints, including low expression levels for some genes, transcript degradation during RNA isolation, incomplete transcript processing, alternative splicing and incorrect assembly of transcript

reads. A compounding factor is that some genes have become pseudogenes, and pseudogenes can be transcribed. Rarely are computationally-predicted genes scrutinized to this degree. A total of 107 of 131 genes judged to be complete by manual annotation can be considered a high quality annotation. An annotated list of olfactory-related proteins provides the basis for a phylogenetic analysis presented in Chapter 5. The gene models provide the basis for RNAseq gene expression studies presented in Chapter 6.

The 131 olfactory-related WSS genes represent molecular targets for pest behaviors mediated by olfaction. To date this represents the most complete set of chemoreception genes identified from a basal Hymenoptera species. Further experimental approaches are needed to determine which olfactory-related genes are most critical in the sawfly chemoreception system; such genes could be manipulated for management strategies based on chemical ecology.

Table 4.1. Summary of sequencing strategies and data used to annotate olfactory-related genes from the wheat stem sawfly.

Sequencing Strategy	**454 Transcriptome[1]**	**Illumina Transcriptome**		**Illumina WGS[2]**
Tissue	Male and Female Antennae	Male Antennae	Female Antennae	Single Male
No. Sequence Reads	500,000	71,800,000	84,000,000	320,000,000
No. Contigs	55,115	70,253	82,862	27,447
Ave. Contig size (bp)	805	1,022	1,014	67,670
Range (bp)	40-8,165	100-27,873	100-32,742	3,539-4,355,184

[1]Gress et al. 2013 and Chapter 1.
[2]Sizes refer to genome scaffolds, not contigs.

Table 4.2. Olfactory-related transcripts annotated from *C. cinctus* antennal transcriptomes. Each transcript was searched by BLASTx against the NCBI non redundant protein data base, and the accession number and name of the closest match listed, along the e-value score and percent amino acid identity.

Gene Name	Illumina EST Name	ORF Length (bp)	NCBI BLASTx Hit	e-value	%AA Identity
CcOBP1	comp87_c1_seq1	429	AGI05201.1 odorant-binding protein 2 [Osmia cornuta]	3.00E-34	39%
CcOBP2	comp5_c0_seq1	570	XP_006620448.1 PREDICTED: pheromone-binding protein-related protein 3-like isoform X1 [Apis dorsata]	7.00E-78	79%
CcOBP3	comp2099_c0_seq1	453	CCD17833.1 putative odorant binding protein 64 [Nasonia vitripennis]	5.00E-36	60%
CcOBP4	comp27_c0_seq1	420	ACI30688.1 odorant-binding protein [Solenopsis invicta]	1.00E-36	46%
CcOBP5	comp127618_c0_seq 1	393	AEB54591.1 OBP7 [Helicoverpa armigera]	5.00E-02	19%
CcOBP6	comp7_c1_seq1	495	EFA04594.1 odorant binding protein 6 [Tribolium castaneum]	1.00E-34	46%
CcOBP7	comp36942_c0_seq1	483	XP_008215619.1 PREDICTED: general odorant-binding protein 83a-like isoform X2 [Nasonia vitripennis]	4.00E-03	25%
CcOBP9	comp157475_c0_seq 1	504	AEX33164 OBP8 protein, partial [Locusta migratoria]	5.00E-24	42%
CcOBP10	comp131_c0_seq1	408	ADX94403.1 OBP7 precursor, partial [Solenopsis invicta]	9.00E-41	53%
CcOBP12	comp104227_c0_seq 1	402	ADX94405.1 OBP9 precursor, partial [Solenopsis invicta]	4.00E-46	61%
CcOBP13	comp23_c0_seq1	423	XP_008543434.1 PREDICTED: pheromone-binding protein-related protein 6-like [Microplitis demolitor]	3.00E-26	40%

Table 4.2 Cont.

Gene Name	Illumina EST Name	ORF Length (bp)	NCBI BLASTx Hit	e-value	%AA Identity
CcCSP1	comp112647_c0_seq1	339	NP_001071278.1 chemosensory protein 2 precursor [Apis mellifera]	2.00E-54	82%
CcCSP2a	comp1_c2_seq1	387	AID61324.1 chemosensory protein [Calliphora stygia]	8.00E-45	54%
CcCSP2b	comp108_c0_seq1	387	AFQ07769.1 chemosensory protein [Apis cerana cerana]	6.00E-42	58%
CcCSP3	comp311421_c0_seq1	312	NP_001072129.1 chemosensory protein 5 precursor [Apis mellifera]	3.00E-23	55%
CcCSP4	comp15725_c0_seq1	378	AGD80083.1 chemosensory protein 3 [Apolygus lucorum]	1.00E-45	57%
CcCSP5	comp106112_c0_seq1	378	ABE68832.1 putative chemosensory protein 1 [Sclerodermus guani]	1.00E-41	50%
CcCSP6	comp69735_c0_seq1	417	AAM77025.1 chemosensory protein [Rhyparobia maderae]	4.00E-34	45%
CcCSP7	comp2_c3_seq1	375	NP_001011583.1 chemosensory protein 3 precursor [Apis mellifera]	7.00E-33	47%
CcOr1	comp6429_c0_seq1	1188	AGS43070.1 odorant receptor Or3k [Cephus cinctus]	0.00E+00	99%
CcOr3	comp2888_c0_seq1	1152	AGS43061.1 odorant receptor Or3b, partial [Cephus cinctus]	0.00E+00	99%
CcOr4	comp8800_c2_seq1	1200	AGS43046.1 odorant receptor Or1a [Cephus cinctus]	0.00E+00	100%
CcOr5	comp7346_c0_seq1	1248	AGS43062.1 odorant receptor Or3c [Cephus cinctus]	0.00E+00	100%
CcOr6	comp323_c0_seq1	1224	AGS43050.1 odorant receptor Or2a [Cephus cinctus]	0.00E+00	100%
CcOr7	comp10448_c0_seq1	1149	AGS43047.1 odorant receptor Or1b [Cephus cinctus]	0.00E+00	100%
CcOr8	comp21168_c0_seq1	1164	AGS43051.1 odorant receptor Or2b [Cephus cinctus]	0.00E+00	100%
CcOr9	comp14565_c0_seq1	1137	AGS43072.1 odorant receptor Or4a [Cephus cinctus]	0.00E+00	100%
CcOr11	comp8015_c0_seq1	1251	AGS43063.1 odorant receptor Or3d [Cephus cinctus]	0.00E+00	99%

Table 4.2 Cont.

Gene Name	Illumina EST Name	ORF Length (bp)	NCBI BLASTx Hit	e-value	%AA Identity
CcOr12	comp8616_c0_seq1	1242	AGS43064.1 odorant receptor Or3e [Cephus cinctus]	0.00E+00	99%
CcOr13	comp6789_c0_seq1	1242	AGS43065.1 odorant receptor Or3f [Cephus cinctus]	0.00E+00	99%
CcOr14	comp1886_c0_seq1	1257	AGS43066.1 odorant receptor Or3g [Cephus cinctus]	0.00E+00	100%
CcOr15	comp3501_c0_seq1	1137	AGS43067.1 odorant receptor Or3h, partial [Cephus cinctus]	0.00E+00	100%
CcOr16	comp1585_c0_seq1	1236	AGS43056.1 odorant receptor Or2g [Cephus cinctus]	0.00E+00	97%
CcOr17	comp17795_c0_seq1	1128	AGS43073.1 odorant receptor Or4b [Cephus cinctus]	0.00E+00	100%
CcOr18	comp21494_c0_seq1	1182	AGS43054.1 odorant receptor Or2e [Cephus cinctus]	2.00E-165	99%
CcOr19	comp15394_c0_seq1	1269	AGS43048.1 odorant receptor Or1c [Cephus cinctus]	0.00E+00	96%
CcOr20	comp17176_c0_seq1	1119	AGS43049.1 odorant receptor Or1d, partial [Cephus cinctus]	0.00E+00	99%
CcOr21	comp1886_c0_seq1	1257	AGS43066.1 odorant receptor Or3g [Cephus cinctus]	0.00E+00	100%
CcOr25	comp1585_c0_seq1	1224	AGS43057.1 odorant receptor Or2h [Cephus cinctus]	0.00E+00	100%
CcOr26	comp323_c0_seq1	1224	AGS43058.1 odorant receptor Or2i [Cephus cinctus]	0.00E+00	100%
CcOr26b	comp323_c0_seq1	1224	AGS43050.1 odorant receptor Or2a [Cephus cinctus]	0.00E+00	90%
CcOr27	comp6068_c0_seq1	1224	AGS43059.1 odorant receptor Or2j [Cephus cinctus]	4.00E-130	77%
CcOr29	comp2557_c0_seq1	1158	XP_008546680.1 PREDICTED: odorant receptor 43a-like [Microplitis demolitor]	8.00E-90	40%
CcOr30	comp14002_c0_seq1	1200	AGS43052.1 odorant receptor Or2c [Cephus cinctus]	0.00E+00	99%
CcOr31	comp1886_c0_seq1	1266	AGS43068.1 odorant receptor Or3i [Cephus cinctus]	0.00E+00	87%
CcOr32	comp98492_c0_seq1	1152	EFN85521.1 Putative odorant receptor 13a [Harpegnathos saltator]	2.00E-82	36%
CcOr34	comp12487_c0_seq1	1260	XP_008550942.1 PREDICTED: odorant receptor 13a-like [Microplitis demolitor]	2.00E-84	39%

Table 4.2 Cont.

Gene Name	Illumina EST Name	ORF Length (bp)	NCBI BLASTx Hit	e-value	%AA Identity
CcOr36	comp7610_c0_seq1	1215	XP_008548306.1 PREDICTED: odorant receptor 43a-like [Microplitis demolitor]	6.00E-60	31%
CcOr37	comp7610_c0_seq1	1200	AGS43050.1 odorant receptor Or2a [Cephus cinctus]	1.00E-55	31%
CcOr38	comp193785_c0_seq1	888	XP_008549314.1 PREDICTED: odorant receptor 67a-like [Microplitis demolitor]	1.00E-44	38%
CcOr39	comp137124_c0_seq1	1215	NP_001177473.1 odorant receptor 17 [Nasonia vitripennis]	1.00E-49	30%
CcOr40	comp15282_c0_seq1	1233	AGS43065.1 odorant receptor Or3f [Cephus cinctus]	5.00E-178	59%
CcOr41	comp154586_c0_seq1	1065	NP_001177472.1 odorant receptor 16 [Nasonia vitripennis]	4.00E-42	30%
CcOr42	comp5452_c0_seq1	1200	XP_006618840.1 PREDICTED: putative odorant receptor 22c-like [Apis dorsata]	1.00E-62	35%
CcOr43	comp47948_c0_seq1	1146	NP_001177509.1 odorant receptor 69 [Nasonia vitripennis]	2.00E-78	38%
CcOr44	comp6789_c0_seq1	1248	AGS43065.1 odorant receptor Or3f [Cephus cinctus]	0.00E+00	79%
CcOr45	comp57843_c0_seq1	1191	NP_001177510.1 odorant receptor 78 [Nasonia vitripennis]	1.00E-114	45%
CcOr46	comp25787_c0_seq1	1149	AGS43072.1 odorant receptor Or4a [Cephus cinctus]	9.00E-108	46%
CcOr47	comp137124_c0_seq1	1158	NP_001177473.1 odorant receptor 17 [Nasonia vitripennis]	1.00E-49	30%
CcOr48	comp2557_c0_seq1	1230	AGS43061.1 odorant receptor Or3b, partial [Cephus cinctus]	1.00E-75	42%
CcOr49	comp67593_c0_seq1	1242	XP_003707057.1 PREDICTED: odorant receptor Or1-like [Megachile rotundata]	9.00E-89	35%
CcOr50	comp1539_c0_seq1	1235	XP_003707057.1 PREDICTED: odorant receptor Or1-like [Megachile rotundata]	9.00E-123	44%
CcOr51	comp17795_c0_seq1	1154	AGS43073.1 odorant receptor Or4b [Cephus cinctus]	2.00E-132	54%
CcOr52	comp34999_c0_seq1	1131	AGS43073.1 odorant receptor Or4b [Cephus cinctus]	8.00E-135	56%
CcOr53	comp21745_c0_seq1	897	EEZ99229.1 odorant receptor 37 [Tribolium castaneum]	2.00E-34	32%
CcOr54	comp2557_c0_seq1	941	AGS43061.1 odorant receptor Or3b, partial [Cephus cinctus]	2.00E-79	42%

Table 4.2 Cont.

Gene Name	Illumina EST Name	ORF Length (bp)	NCBI BLASTx Hit	e-value	%AA Identity
CcOr56	comp54365_c0_seq1	1164	AGS43061.1 odorant receptor Or3b, partial [Cephus cinctus]	1.00E-82	42%
CcOr58	comp5452_c0_seq1	1224	XP_006618840.1 PREDICTED: putative odorant receptor 22c-like [Apis dorsata]	1.00E-62	34%
CcOr59	comp6789_c0_seq1	1019	AGS43065.1 odorant receptor Or3f [Cephus cinctus]	1.00E-129	56%
CcOr60	comp13725_c0_seq1	1104	AGS43067.1 odorant receptor Or3h, partial [Cephus cinctus]	1.00E-80	35%
CcOrco	comp562_c0_seq1	1443	AGS43074.1 odorant receptor Orco [Cephus cinctus]	0.00E+00	100%
CcIR1	comp4278_c0_seq1	1853	AHA80144.1 ionotropic receptor 8a [Schistocerca gregaria]	0.00E+00	66%
CcIR2	comp1341_c0_seq1	2808	AII01112.1 ionotropic receptors [Dendrolimus houi]	0.00E+00	59%
CcIR3	comp7625_c0_seq1	2508	NP_650924 ionotropic receptor 93a [Drosophila melanogaster]	0.00E+00	56%
CcIR4	comp153931_c0_seq1	1786	ADR64678.1 putative chemosensory ionotropic receptor IR21a [Spodoptera littoralis]	5.50E-156	43%
CcIR7	comp5191_c0_seq1	1907	ADR64685.1 putative chemosensory ionotropic receptor IR75q.2 [Spodoptera littoralis]	4.00E-133	41%
CcIR8	comp785_c0_seq1	1935	ADR64685.1 putative chemosensory ionotropic receptor IR75q.2 [Spodoptera littoralis]	8.00E-59	29%
CcIR9	comp38814_c0_seq1	1548	NP_649012.2 ionotropic receptor 75a [Drosophila melanogaster]	2.00E-27	27%
CcIR11	comp34038_c0_seq1	1896	AFC91758.1 putative ionotropic receptor IR41a [Cydia pomonella]	1.00E-96	35%
CcIR15	comp11389_c1_seq1	1482	ADR64682.1 putative chemosensory ionotropic receptor IR68a [Spodoptera littoralis]	7.00E-118	45%
CcIR17	comp9541_c0_seq1	1434	ETN63667.1 Ionotropic receptor 76b [Anopheles darlingi]	7.00E-118	46%
CcCCE1	comp205_c0_seq1	1719	XP_008552484.1 PREDICTED: esterase E4-like [Microplitis demolitor]	0.00E+00	64%
CcCCE2	comp14_c0_seq1	1668	NP_001136104.1 carboxylesterase clade B, member 6 precursor [Nasonia vitripennis]	0.00E+00	59%

Table 4.2 Cont.

Gene Name	Illumina EST Name	ORF Length (bp)	NCBI BLASTx Hit	e-value	%AA Identity
CcCCE3	comp504_c0_seq1	1788	EGI65426.1 Esterase FE4 [Acromyrmex echinatior]	0.00E+00	56%
CcCCE4	comp3116_c0_seq1	1575	XP_001604042.1 PREDICTED: esterase FE4 isoform X2 [Nasonia vitripennis]	0.00E+00	50%
CcCCE4b	comp3116_c0_seq1	1053	XP_008216259.1 PREDICTED: esterase FE4 isoform X1 [Nasonia vitripennis]	6.00E-69	40%
CcCCE7	comp2949_c0_seq1	1721	KDR14051.1 Esterase FE4 [Zootermopsis nevadensis]	4.00E-126	40%
CcCCE8	comp960_c1_seq2	1575	XP_001604042.1 PREDICTED: esterase FE4 isoform X2 [Nasonia vitripennis]	6.00E-167	46%
CcCCE9	comp93_c1_seq1	1767	KDR14051.1 Esterase FE4 [Zootermopsis nevadensis]	3.00E-83	35%
CcCCE10a	comp12952_c0_seq1	1697	XP_392698.2 PREDICTED: esterase FE4-like [Apis mellifera]	2.00E-161	46%
CcCCE10b	comp12952_c0_seq1	1662	XP_006610295.1 PREDICTED: carboxylesterase 4A-like isoform X1 [Apis dorsata]	1.00E-159	46%
CcCCE10c	comp12952_c0_seq1	1614	XP_003487272.1 PREDICTED: esterase FE4-like [Bombus impatiens]	5.00E-169	49%
CcCCE10d	comp12952_c0_seq1	1665	XP_003487272.1 PREDICTED: esterase FE4-like [Bombus impatiens]	2.00E-159	46%
CcGST1a	comp111_c0_seq1	660	XP_008554401.1 glutathione S-transferase 1-1-like [Microplitis demolitor]	4.00E-94	61%
CcGST1b	comp111_c0_seq1	651	NP_001165913.1 glutathione S-transferase D1 [Nasonia vitripennis]	2.00E-105	69%
CcGST1c	comp111_c0_seq1	666	XP_006620743.1 glutathione S-transferase 1-1-like [Apis dorsata]	1.00E-98	65%
CcGST2	comp47197_c0_seq1	600	NP_001171499.1 glutathione S-transferase D1 [Apis mellifera]	7.00E-48	53%
CcGST4	comp368_c0_seq1	696	EZA58046.1 Glutathione S-transferase theta-1 [Cerapachys biroi]	1.00E-110	67%
CcGST5	comp3824_c0_seq1	618	XP_003689705.1 glutathione S-transferase-like [Apis florea]	3.00E-73	55%
CcGST6	comp5525_c0_seq1	624	XP_003703955.1 glutathione S-transferase-like [Megachile rotundata]	4.00E-111	82%

Table 4.2 Cont.

Gene Name	Illumina EST Name	ORF Length (bp)	NCBI BLASTx Hit	e-value	%AA Identity
CcGST7	comp8230_c0_seq1	729	NP_001165912.1 glutathione S-transferase O1 [Nasonia vitripennis]	2.00E-128	77%
CcP450_1	comp43439_c0_seq1	1506	XP_001605600.2 PREDICTED: probable cytochrome P450 305a1 [Nasonia vitripennis]	0.00E+00	58%
CcP450_2	comp73896_c0_seq1	1617	XP_393885.1 PREDICTED: cytochrome P450 18a1 [Apis mellifera]	0.00E+00	77%
CcP450_3	comp20087_c0_seq1	1794	EFN61057.1 Probable cytochrome P450 303a1 [Camponotus floridanus]	0.00E+00	59%
CcP450_5	comp259055_c0_seq1	1017	EFN74586.1 Cytochrome P450 306a1 [Camponotus floridanus]	1.00E-156	63%
CcP450_6	comp199929_c0_seq1	1659	XP_001603435.1 PREDICTED: cytochrome P450 307a1 [Nasonia vitripennis]	0.00E+00	71%
CcP450_7	comp84792_c0_seq1	1587	XP_003702139.1 PREDICTED: probable cytochrome P450 4aa1-like [Megachile rotundata]	0.00E+00	64%
CcP450_8	comp115755_c0_seq1	1347	XP_001603476.2 PREDICTED: cytochrome P450 4c3 [Nasonia vitripennis]	0.00E+00	71%
CcP450_9	comp1049_c0_seq1	1545	XP_008550566.1 PREDICTED: cytochrome P450 4C1-like [Microplitis demolitor]	0.00E+00	57%
CcP450_10	comp4386_c0_seq1	1722	NP_001165992.1 cytochrome P450 4G43 [Nasonia vitripennis]	0.00E+00	84%
CcP450_11	comp75639_c0_seq1	1176	XP_006618627.1 PREDICTED: cytochrome P450 4c3-like [Apis dorsata]	3.00E-147	67%
CcP450_12	comp4391_c0_seq1	1554	XP_008217632.1 PREDICTED: cytochrome P450 4C1 [Nasonia vitripennis]	6.00E-170	71%
CcP450_13	comp12159_c0_seq1	1554	XP_001602395.3 PREDICTED: cytochrome P450 4C1 isoform X1 [Nasonia vitripennis]	0.00E+00	55%
CcP450_14	comp6725_c0_seq1	1503	NP_001165999.1 cytochrome P450 6AS30 [Nasonia vitripennis]	0.00E+00	58%
CcP450_15	comp1271_c0_seq2	1596	ACE75339.1 cytochrome P450 [Glyptapanteles indiensis]	4.00E-169	48%
CcP450_16	comp38159_c0_seq1	1580	XP_003703366.1 PREDICTED: probable cytochrome P450 301a1, mitochondrial-like [Megachile rotundata]	0.00E+00	84%

Table 4.2 Cont.

Gene Name	Illumina EST Name	ORF Length (bp)	NCBI BLASTx Hit	e-value	%AA Identity
CcP450_17	comp1595_c0_seq1	1475	EZA62389.1 Cytochrome P450 CYP12A2 [Cerapachys biroi]	6.00E-170	59%
CcP450_18	comp12105_c0_seq1	1519	XP_003393389.1 PREDICTED: cytochrome P450 9e2-like isoform 1 [Bombus terrestris]	2.00E-153	43%
CcP450_19	comp928_c0_seq1	1592	XP_003697790.1 PREDICTED: cytochrome P450 9e2-like [Apis florea]	0.00E+00	58%
CcP450_20	comp14219_c0_seq1	1547	XP_003488941.1 PREDICTED: cytochrome P450 6a2-like [Bombus impatiens]	0.00E+00	50%
CcP450_21	comp3886_c0_seq2	1502	EFN84330.1 Cytochrome P450 6k1 [Harpegnathos saltator]	0.00E+00	74%
CcP450_22	comp62552_c0_seq1	1544	XP_003424479.1 PREDICTED: cytochrome P450 6k1 [Nasonia vitripennis]	0.00E+00	58%
CcP450_23	comp288_c0_seq1	1543	XP_003491922.1 PREDICTED: cytochrome P450 6k1-like [Bombus impatiens]	0.00E+00	58%
CcP450_24	comp2201_c0_seq1	1519	XP_008557951.1 PREDICTED: cytochrome P450 9e2-like [Microplitis demolitor]	7.00E-161	48%
CcP450_25	comp131195_c0_seq1	1268	EFN69019.1 Cytochrome P450 314a1, mitochondrial [Camponotus floridanus]	0.00E+00	77%
CcP450_26	comp97_c0_seq1	1251	XP_008210899.1 PREDICTED: cytochrome P450 4g15 [Nasonia vitripennis]	1.00E-151	54%
CcP450_27	comp73_c0_seq1	1545	XP_001602787.1 PREDICTED: probable cytochrome P450 6a14 [Nasonia vitripennis]	6.00E-180	50%
CcP450_28	comp328011_c0_seq1	1383	XP_003703365.1 PREDICTED: probable cytochrome P450 49a1-like [Megachile rotundata]	0.00E+00	77%
CcP450_30	comp3967_c0_seq1	1701	XP_006616362.1 PREDICTED: cytochrome P450 315a1, mitochondrial-like [Apis dorsata]	0.00E+00	56%
CcP450_31	comp233990_c0_seq1	1317	XP_003487290.1 PREDICTED: probable cytochrome P450 12a5, mitochondrial-like [Bombus impatiens]	6.00E-173	57%

Table 4.3. Olfactory-related gene models annotated from the wheat stem sawfly genome. The position of each gene within the WGS scaffolds is noted along with the number of exons. Additional protein characteristics are also noted, including the length of the predicted protein, the number of predicted transmembrane domains (TMMs), the presence of a signal peptide in the mature protein and the completeness of the coding region. Y denotes a complete coding region; N, an incomplete N-terminus; C, and incomplete C-terminus; and E, a missing exon.

Name	NCBI Scaffold Accession Number	Scaffold position	Exons	Length (AA)	TMMs	Signal Peptide	Complete
CcOBP1	gi|453039718|gb|KB465834.1|	208842-209739	5	142		yes	Y
CcOBP2	gi|453039869|gb|KB465683.1|	2792041-2794452	6	147		yes	Y
CcOBP3	gi|453039108|gb|KB466444.1|	64707-65289	4	150		yes	Y
CcOBP4	gi|453039888|gb|KB465664.1|	677711-679259	6	139		yes	Y
CcOBP5	gi|453039195|gb|KB466357.1|	23879-24565	4	130		yes	Y
CcOBP6	gi|453039940|gb|KB465612.1|	549384-552760	7	173		yes	Y
CcOBP7	gi|453039820|gb|KB465732.1|	173662-174411	6	122		yes	Y
CcOBP9	gi|453039734|gb|KB465818.1|	258044-259619	6	167		yes	Y
CcOBP10	gi|453039195|gb|KB466357.1|	26783-27567	5	135		yes	Y
CcOBP12	gi|453039820|gb|KB465732.1|	170974-171679	5	133		yes	Y
CcOBP13	gi|453039718|gb|KB465834.1|	206991-208466	6	140		yes	Y
CcCSP1	gi|453039188|gb|KB466364.1|	21743-22290	2	117		yes	Y
CcCSP2a	gi|453039999|gb|KB465553.1|	675988-676457	2	133		yes	Y
CcCSP2b	gi|453039999|gb|KB465553.1|	679533-680575	2	128		yes	Y
CcCSP3	gi|453039999|gb|KB465553.1|	930980-931420	2	103		yes	Y
CcCSP4	gi|453038465|gb|KB467087.1|	17708-18309	3	125		yes	Y
CcCSP5	gi|453039742|gb|KB465810.1|	463159-463695	2	125		yes	Y
CcCSP6	gi|453039742|gb|KB465810.1|	465766-467704	3	138		yes	Y

Table 4.3 Cont.

Name	NCBI Scaffold Accession Number	Scaffold position	Exons	Length (AA)	TMMs	Signal Peptide	Complete
CcCSP7	gi|453039742|gb|KB465810.1|	479766-480346	2	124		yes	Y
CcOr1	gi|453039676|gb|KB465876.1|	1033071-1035370	5	395	5		Y
CcOr3	gi|453038779|gb|KB466773.1|	1975-4188	6	383	6		Y
CcOr4	gi|453039993|gb|KB465559.1|	303318-305050	5	399	6		Y
CcOr5	gi|453039945|gb|KB465607.1|	21300-23660	7	415	7		Y
CcOr6	gi|453038272|gb|KB467280.1|	1093-2632	4	407	7		Y
CCOr6b	gi|453039299|gb|KB466253.1|	484231-484752	2	407	7		Y
CcOr7	gi|453039518|gb|KB466034.1|	10722-13240	6	382	6		Y
CcOr8	gi|453039805|gb|KB465747.1|	430248-432281	6	387	7		Y
CcOr9	gi|453038887|gb|KB466665.1|	19204-20814	5	378	6		Y
CcOr11	gi|453038260|gb|KB467292.1|	309545-311601	6	416	6		Y
CcOr12	gi|453039847|gb|KB465705.1|	54263-55993	5	413	7		Y
CcOr13	gi|453039847|gb|KB465705.1|	46530-49408	5	413	7		Y
CcOr14	gi|453039190|gb|KB466362.1|	59295-60131	5	418	6		Y
CcOr15	gi|453038779|gb|KB466773.1|	33595-61943	6	378	6		N
CcOr16	gi|453039299|gb|KB466253.1|	471011-472880	6	411	6		Y
CcOr17	gi|453038887|gb|KB466665.1|	4077-5612	5	375	7		Y
CcOr18	gi|453039935|gb|KB465617.1|	1607875-1609935	6	393	6		Y
CcOr19	gi|453039582|gb|KB465970.1|	146445-148105	6	422	7		Y
CcOr20	gi|453039071|gb|KB466481.1|	406790-408570	5	372	6		Y
CcOr21	gi|453039904|gb|KB465648.1|	2618810-2620773	5	418	7		Y

Table 4.3 Cont.

Name	NCBI Scaffold Accession Number	Scaffold position	Exons	Length (AA)	TMMs	Signal Peptide	Complete
CcOr25	gi|453039299|gb|KB466253.1|	492585-494733	6	407	7		Y
CcOr26	gi|453039299|gb|KB466253.1|	485427-487687	6	407	7		Y
CcOr26b	gi|453039299|gb|KB466253.1|	480008-482271	6	407	8		Y
CcOr27	gi|453039299|gb|KB466253.1|	500167-502497	6	407	7		Y
CcOr29	gi|453038779|gb|KB466773.1|	28742-30816	6	385	6		N
CcOr30	gi|453039904|gb|KB465648.1|	2623388-2624927	6	399	5		Y
CcOr31	gi|453039847|gb|KB465705.1|	35502-36996	4	421	7		Y
CcOr32	gi|453038779|gb|KB466773.1|	10243-11784	6	383	8		Y
CcOr34	gi|453039259|gb|KB466293.1|	306250-308122	7	419	8		Y
CcOr36	gi|453039299|gb|KB466253.1|	506377-508268	5	404	8		Y
CcOr37	gi|453039299|gb|KB466253.1|	498048-499940	6	399	8		Y
CcOr38	gi|453039005|gb|KB466547.1|	102589-103899	7	295	5		N
CcOr39	gi|453039718|gb|KB465834.1|	550750-552363	3	404	7		C
CcOr40	gi|453039847|gb|KB465705.1|	41860-43414	5	410	8		Y
CcOr41	gi|453038868|gb|KB466684.1|	50025-52550	6	354	6		N
CcOr42	gi|453039718|gb|KB465834.1|	562192-563754	6	399	7		N
CcOr43	gi|453039734|gb|KB465818.1|	2316695-2318254	6	381	6		Y
CcOr44	gi|453039847|gb|KB465705.1|	49736-51296	5	415	8		Y
CcOr45	gi|453039149|gb|KB466403.1|	150800-155012	7	396	7		Y
CcOr46	gi|453038887|gb|KB466665.1|	25077-26533	5	382	7		Y
CcOr47	gi|453038779|gb|KB466773.1|	6765-8308	6	404	7		Y
CcOr48	gi|453038779|gb|KB466773.1|	13611-15167	5	409	7		Y

Table 4.3 Cont.

Name	NCBI Scaffold Accession Number	Scaffold position	Exons	Length (AA)	TMMs	Signal Peptide	Complete
CcOr49	gi|453039627|gb|KB465925.1|	242698-247035	7	413	7		Y
CcOr50	gi|453039627|gb|KB465925.1|	209172-209465	6	410	5		Y
CcOr51	gi|453038887|gb|KB466665.1|	9007-10769	5	383	7		Y
CcOr52	gi|453038523|gb|KB467029.1|	2396-3731	4	375	7		Y
CcOr53	gi|453039759|gb|KB465793.1|	1007290-1008995	4	298	6		C
CcOr54	gi|453038779|gb|KB466773.1|	21063-22471	5	313	6		Y
CcOr56	gi|453038779|gb|KB466773.1|	16938-18530	6	387	7		Y
CcOr58	gi|453039718|gb|KB465834.1|	559782-561378	6	407	6		Y
CcOr59	gi|453039847|gb|KB465705.1|	51693-53024	5	338	5		C
CcOr60	gi|453038514|gb|KB467038.1|	4477-5938	5	367	8		C
CcOrco	gi|453039734|gb|KB465818.1|	217215-221196	8	480	7		Y
CcIR1	gi|453039818|gb|KB465734.1|	49774-52788	11	616	3		N
CcIR2	gi|453039896|gb|KB465656.1|	392356-398492	5	935	3		Y
CcIR3	gi|453039470|gb|KB466082.1|	22799-26753	12	835	3		Y
CcIR4	gi|453039869|gb|KB465683.1|	1200711-1203581	9	594	4		C
CcIR7	gi|453039617|gb|KB465935.1|	118577-130804	10	634	3		Y
CcIR8	gi|453039742|gb|KB465810.1	484560-489417	10	645	3		Y
CcIR9	gi|453039875|gb|KB465677.1|	2610902-2611212	7	515	4		N
CcIR11	gi|453039588|gb|KB465964.1|	167733-170202	8	631	4		Y
CcIR15	gi|453039219|gb|KB466333.1|	177846-179738	4	493	4		Y
CcIR17	gi|453039026|gb|KB466526.1|	74008-71679	9	477	4		N

Table 4.3 Cont.

Name	NCBI Scaffold Accession Number	Scaffold position	Exons	Length (AA)	TMMs	Signal Peptide	Complete
CcCCE1	gi|453039535|gb|KB466017.1|	412827-413030	8	572		yes	Y
CcCCE2	gi|453039785|gb|KB465767.1|	384308-386662	9	555		yes	Y
CcCCE3	gi|453039443|gb|KB466109.1|	783651-788296	8	595		yes	Y
CcCCE4	gi|453039143|gb|KB466409.1|	3262-6835	7	524		no	Y
CcCCE4b	gi|453038242|gb|KB467310.1|	166931-168309	5	350		no	E
CcCCE7	gi|453039470|gb|KB466082.1|	147461-151072	10	573		no	Y
CcCCE8	gi|453039143|gb|KB466409.1|	10117-12284	8	524		no	Y
CcCCE9	gi|453039904|gb|KB465648.1|	658683-661391	7	588		no	Y
CcCCE10a	gi|453039904|gb|KB465648.1|	1026518-1031524	9	564		yes	Y
CcCCE10b	gi|453039904|gb|KB465648.1|	1021744-1026648	10	553		no	Y
CcCCE10c	gi|453039904|gb|KB465648.1|	1015073-1017686	11	538		yes	C
CcCCE10d	gi|453039904|gb|KB465648.1|	1010358-1015129	9	554		no	Y
CcGST1a	gi|453038992|gb|KB466560.1|	14982-16775	5	219			Y
CcGST1b	gi|453038992|gb|KB466560.1|	8096-10022	5	216			Y
CcGST1c	gi|453038992|gb|KB466560.1|	3914-5221	5	221			Y
CcGST2	gi|453039665|gb|KB465887.1|	8384-9922	5	199			N
CcGST4	gi|453039500|gb|KB466052.1|	243298-244208	4	231			Y
CcGST5	gi|453039005|gb|KB466547.1|	8587-9530	4	205			Y
CcGST6	gi|453039828|gb|KB465724.1|	894193-895900	4	207			Y
CcGST7	gi|453039880|gb|KB465672.1|	242686-243811	5	242			Y

Table 4.3 Cont.

Name	NCBI Scaffold Accession Number	Scaffold position	Exons	Length (AA)	TMMs	Signal Peptide	Complete
CcP450_1	gi|453039500|gb|KB466052.1|	591619-594706	8	501			Y
CcP450_2	gi|453039089|gb|KB466463.1|	264477-267108	5	538			Y
CcP450_3	gi|453039263|gb|KB466289.1|	230114-232563	8	597			Y
CcP450_5	gi|453039089|gb|KB466463.1|	272275-273445	3	338			Y
CcP450_6	gi|453039389|gb|KB466163.1|	121461-124483	2	552			Y
CcP450_7	gi|453038868|gb|KB466684.1|	96419-100111	7	528			Y
CcP450_8	gi|453039828|gb|KB465724.1|	2182743-2184734	10	448			N
CcP450_9	gi|453039010|gb|KB466542.1|	637879-641122	11	514			Y
CcP450_10	gi|453038929|gb|KB466623.1|	3597-6195	7	573			Y
CcP450_11	gi|453039999|gb|KB465553.1|	1638814-1642050	6	391			N
CcP450_12	gi|453039358|gb|KB466194.1|	689071-692429	12	517			Y
CcP450_13	gi|453038287|gb|KB467265.1|	14750-17385	11	517			Y
CcP450_14	gi|453039490|gb|KB466062.1|	59027-61704	4	500			Y
CcP450_15	gi|453039782|gb|KB465770.1|	92264-93859	1	531			Y
CcP450_16	gi|453039545|gb|KB466007.1|	853767-855947	8	525			Y
CcP450_17	gi|453039935|gb|KB465617.1|	1620344-1623719	7	490			C
CcP450_18	gi|453039439|gb|KB466113.1|	16425-17943	1	505			Y
CcP450_19	gi|453039439|gb|KB466113.1|	3386-4977	1	529			Y
CcP450_20	gi|453039888|gb|KB465664.1|	735807-740023	6	514			Y
CcP450_21	gi|453039458|gb|KB466094.1|	163216-165107	5	499			Y
CcP450_22	gi|453039458|gb|KB466094.1|	159706-161773	5	513			Y
CcP450_23	gi|453039443|gb|KB466109.1|	589620-591629	5	513			Y

Table 4.3 Cont.

Name	NCBI Scaffold Accession Number	Scaffold position	Exons	Length (AA)	TMMs	Signal Peptide	Complete
CcP450_24	gi|453039782|gb|KB465770.1|	97293-98811	1	505			Y
CcP450_25	gi|453040039|gb|KB465513.1|	765-2533	7	421			N
CcP450_26	gi|453039944|gb|KB465608.1|	440059-441742	7	416			N
CcP450_27	gi|453039448|gb|KB466104.1|	1563-3527	6	514			Y
CcP450_28	gi|453039545|gb|KB466007.1|	857163-859889	8	460			C
CcP450_30	gi|453039531|gb|KB466021.1|	71799-73779	5	470			N
CcP450_31	gi|453039071|gb|KB466481.1|	403482-405718	9	447			C

References

Altschul, S.F., Madden, T.L., Schäffer, A.A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D.J. (1997). Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic acids Res* 25: 3389–3402.

Baum, J. A., Bogaert, T., Clinton, W., Heck, G. R., Feldmann, P., Ilagan, O., et al. (2007). Control of coleopteran insect pests through RNA interference. *Nature biotechnology* 25: 1322–1326.

Benton, R., Sachse, S., Michnick, S.W. and Vosshall, L.B. (2006) Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo. *PLoS Biol* 4: e20.

Benton, R., Vannice, K.S., Gomez-Diaz, C. and Vosshall, L.B. (2009) Variant ionotropic glutamate receptors as chemosensory receptors in Drosophila. *Cell* 136: 149–162.

Beres, B.L., Dosdall, L.M., Weaver, D.K., Carcamo, H.A. and Spaner, D.M. (2011) Biology and integrated management of the wheat stem sawfly and the need for continuing research. *Can Entomol* 143: 105–125.

Croset, V., Ritz, R., Cummins, S. F., Budd, A., Brawand, D., Kaessmann, H., Gibson, T. J. and Benton, R. (2010) Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet* 6:e1001064.

Forêt, S., Wanner, K.W. and Maleszka, R. (2007) Chemosensory proteins in the honey bee: insights from the annotated genome, comparative analyses and expressional profiling. *Insect Biochem Mol Biol* 37: 19–28.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology* 29: 644–652.

Gress, J.C., Robertson, H.M., Weaver, D.K., Dlakić, M. and Wanner, K.W. (2013) Odorant receptors of a primitive hymenopteran pest, the wheat stem sawfly. *Insect Mol Biol* 26: 659–667.

Grosjean, Y., Rytz, R., Farine, J.P., Abuin, L., Cortot, J., Jefferis, G.S. and Benton, R. (2011) An olfactory receptor for food-derived odours promotes male courtship in *Drosophila*. *Nature* 478: 236–240.

Hofmann, K. and Stoffel, W. (1993) TMbase-A database of membrane spanning protein segments. *Biol Chem Hoppe-Seyler* 374: 166

Larsson, M.C., Domingos, A.I., Jones, W.D., Chiappe, M.E., Amrein, H. and Vosshall, L.B. (2004) Or83b encodes a broadly expressed odorant receptor essential for Drosophila olfaction. *Neuron* 43: 703–714.

Leal, W.S. (2013) Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol* 58: 373–391.

Morozova, O., and Marra, M.A. (2008). Applications of next-generation sequencing technologies in functional genomics. *Genomics* 92: 255–264.

Nakagawa, T. and Vosshall, L.B. (2009) Controversy and consensus: noncanonical signaling mechanisms in the insect olfactory system. *Curr Opin Neurobiol* 19: 284–292.

Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., et al. (2003). TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics 19: 651–652.

Peterson, T.N., Brunak, S., von Heijne, G. and Nielsen, H. (2011) SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* 8: 785–786

Ronaghi, M., Uhlén, M., and Nyren, P. (1998). A sequencing method based on real-time pyrophosphate. *Science* 281: 363–365.

Robertson, H.M., Gadau, J. and Wanner, K.W. (2010) The insect chemoreceptor superfamily of the parastiod jewel wasp *Nasonia vitripennis*. *Insect Mol Biol* 19: 121–136.

Robertson, H.M. and Wanner, K.W. (2006) The chemoreceptor superfamily in the honey bee, *Apis mellifera*: expansion of the odorant, but not gustatory, receptor family. *Genome Res* 16: 1395–1403.

Rutzler, M. and Zwiebel, L.J. (2005) Molecular biology of insect olfaction: recent progress and conceptual models. *J Com Physiol A Neuroethol Sens Neural Behav Physiol* 191: 777–790.

Silbering, A.F., Rytz, R., Grosjean, Y., Abuin, L., Ramdya, P., Jefferis, G.S. and Benton, R. (2011) Complementary Function and Integrated Wiring of the Evolutionarily Distinct Drosophila Olfactory Subsystems. *J Neurosci* 31: 13357–13375.

Smith, C. D., Zimin, A., Holt, C., Abouheif, E., Benton, R., Cash, E. et al. (2011). Draft genome of the globally widespread and invasive Argentine ant (Linepithema humile). *Proc Natl Acad Sci USA 108*: 5673–5678.

Smith, C. R., Smith, C. D., Robertson, H. M., Helmkampf, M., Zimin, A., Yandell, M., et al (2011b). Draft genome of the red harvester ant Pogonomyrmex barbatus. *). Proc Natl Acad Sci USA* 108: 5667–5672.

Vieira, F.G., Forêt, S., He, X., Rozas, J., Field, L.M. et al. (2012) Unique Features of Odorant-Binding Proteins of the Parasitoid Wasp *Nasonia vitripennis* Revealed by Genome Annotation and Comparative Analyses. PLoS ONE 7(8): e43034.

Vogt, R.G. 2003. Biochemical diversity of odor detection: OBPs, ODEs and SNMPs. In Insect Pheromone Biochemistry and Molecular Biology G. J. Blomquist, R. G. Vogt (eds) pp 391-446. Elsevier Academic Press.

Vosshall, L.B. and Hansson, B.S. (2011) A unified nomenclature system for the insect olfactory coreceptor. *Chem Senses* 36: 497–498.

Wanner, K.W., Anderson, A.R., Trowell, S.C., Theilman, D.A., Robertson, H.M. and Newcomb, R.D. (2007) Female-biased expression of odourant receptor genes in the adult antennae of the silkworm, *Bombyx mori*. *Insect Mol Biol* 16: 107–119.

Wanner, K.W. and Robertson, H.M. (2010) Odorant and gustatory receptor gene families in the Lepidoptera. In: Molecular biology and genetics of Lepidoptera. MR Goldsmith and F Marek, eds (Boca Raton, FL: CRC Press): 153-168.

Wicher, D., Schafer, R., Bauernfeind, R., Stensmyr, M.C., Heller, R., Heinemann, S.H. and Hansson, B.S. (2008) *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature* 452: 1007–1011.

Wurm, Y., Wang, J., Riba-Grognuz, O., Corona, M., Nygaard, S., Hunt, B. G. et al. (2011). The genome of the fire ant *Solenopsis invicta*. *Proc Natl Acad Sci USA* 108: 5679–5684.

CHAPTER FIVE

ANNOTATION OF THE MAJOR CHEMOSENSORY GENE FAMILIES IN THE WHEAT STEM SAWFLY, *CEPHUS CINCTUS*, A MAJOR PEST OF WHEAT IN THE NORTHERN PLAINS

Contribution of Authors and Co-Authors

Manuscript in Chapter 5

Author: Joanna C. Gress

Contributions: Conceived and implemented the study design and data collection. Collected and analyzed data. Wrote the manuscript.

Co-Author: Hugh M. Robertson

Contributions: Laboratory collaborated in sequencing *C. cinctus* genome and transcriptome. Provided feedback on data analysis.

Co-Author: Kim Walden

Contributions: Collaborated in sequencing and assembly of *C. cinctus* genomic data, provided feedback on data analysis.

Co-Author: Kevin W. Wanner

Contributions: Provided feedback on study design, data analysis, and comments on the manuscript.

Manuscript Information Page

Joanna C. Gress, Hugh M. Robertson, Kim Walden and Kevin W. Wanner
Insect Molecular Biology
Status of Manuscript:
__X_ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-review journal
____ Accepted by a peer-reviewed journal
____ Published in a peer-reviewed journal

Published by the Royal Entomological Society

Abstract:

The wheat stem sawfly, *Cephus cinctus* (Hymenoptera: Cephidae), is one of the most important insect pests of wheat in the northern Great Plains region of the United States and Canada, with economic losses exceeding $100 million per year. Traditional pest management strategies including pesticides are generally unsuccessful due to an extended adult flight time and the inaccessible larval stage that feeds within the wheat stem. Research towards integrated pest management strategies based on olfaction has proved promising; however, little is known about the molecular basis of olfaction in this important insect pest. We previously identified candidates respresenting the major olfactory-related gene families including, 11 odorant binding proteins (OBPs), 8 chemosensory proteins (CSPs), 10 ionotropic receptors (IRs), 51 odorant receptors (ORs), 12 carboxylesterases (CCEs), 8 gluthatione S-transferases (GSTs) and 30 cytochrome P450s (P450s) from antennal transcriptome and genome sequences. In this study the sequences are analyzed for conserved protein sequence domains typical of each gene family, to confirm their identity. Phylogenetic relationships within each gene family, with homologous sequences from two other Hymenoptera, the jewel wasp *Nasonia vitripennis* and the honeybee *Apis melifera* were analyzed using Maximum Likelihood trees. This study provides the basis for future gene expression analyses to identify olfactory-related genes that exhibit high antennal and/or sex biased expression, and that might serve as molecular targets for pest control measures.

Introduction

The wheat stem sawfly, *Cephus cinctus* Norton (Hymenoptera: Cephidae) is one of the most economically important pests of wheat (*Triticum aestivum* L. Poaceae) in the Northern Great Plains of North America (Morrill et al. 1994, Morrill 1997). Most economic damage occurs in the Northern Great Plains region encompassing the Canadian prairies and U.S. states of North Dakota and Montana (Fletcher 1904, Ainsle 1920). More recently reports of economic damage to winter wheat have been reported from Colorado and Nebraska (Irell and Pearis 2011, Brandshaw 2013). Economic losses over the past three years have been estimated at more than $100 million annually in the US and Canada with infestation levels reaching 100% in some fields (Morill et al. 1994, Wahl et al 2007). Conventional management strategies have not provided adequate control of this pest due to its biology and life cycle.

The larvae overwinter in the below ground portion of the stem with adults emerging in the late spring during crop growth. Males emerge earlier, and after mating, females begin searching for suitable host plants for oviposition (Ainslie 1920, Weiss et al. 1992). Females usually lay one egg within a wheat stem per visit, in the uppermost developing internode (Holmes and Peterson, 1960). The larva feeds internally on the parenchyma and vascular tissue, damaging the stem and reducing kernel weight and seed set by up to 35% (Holmes 1977). As the wheat crop matures, the larva moves down to the base of the stem where it cuts a notch at ground level, causing the plant to fall over before harvest and creating further yield loss (Beres et al. 2007). The stem is then plugged with

frass and the prepupa forms a cocoon within the stubble in preparation for overwintering (Criddle 1923).

Olfaction is a critical biological process that is essential for the survival of animals. In insects, olfaction allows for the recognition of volatile cues from the environment, for the detection of food, mates, predators and prey. In insects, olfaction is mediated by sensilla (chemosensory hairs), where primary contact with chemical signals takes place and the activation of signaling pathways occurs. The signaling pathway involves several different protein families, and includes: The uptake of odorants from the external environment into the sensory hair; transport of the odorants through the sensillum lymph; interaction with the odorant receptor; activation of a signaling cascade that produces a spike and nerve impulse in the sensory neurons; and finally, integration of the sensory input by the CNS leading to olfactory-mediated behavior. The most important proteins involved in these processes belong to multigene families. These families encode the odorant-binding proteins (OBPs) and chemosensory proteins (CSPs) that are involved in the transport of the odorant through the sensillum lymph. The odorant receptor (OR) family (part of the chemoreceptor superfamily that includes gustatory receptors) constitutes a critical component of the peripheral sensory system, where activation of the olfactory neuron-signaling cascade begins. Chemical signals are terminated by odorant degrading enzymes (ODEs) that degrade the odorant and clear it from the lymph so that the sensitivity of the OR is maintained and new signals can be detected. Additionally, one other gene family may be involved in olfaction, the recently identified ionotropic

receptors (IRs) which are expressed on the dendrites of olfactory neurons that do not contain ORs (Benton et al. 2009).

Odorant binding proteins constitue the largest fraction of proteins found in the sensillar lymph of insect chemosensilla. OBPs are small globular proteins (about 135-200 amino acids in length) that bind to and solubilize hydrophobic odorants such as pheromones. They typically have six conserved cysteine residues that form three disulphide bonds (so called “classic” OBPs) although variant lineages exist. OBPs are synthesized in the accessory cells surrounding the neuron body and are subsequently secreted into the sensillum lymph as a mature protein (the secretory signal peptide at the N-terminus is cleaved with secretion). It has been proposed that they act as molecular carriers that transport and deliver odorants to odorant receptors located on the sensory neuron membrane (Steinbrecht 1998, Leal et al. 2005) as well as possibly removing deleterious compounds from the lymph and deactivating odors following receptor activation (Kaissling 1998, Vogt et al. 1985). OBPs have been identified from many different insect orders including the Orthoptera, but have not been found more broadly in the Arthropoda.

Chemosensory proteins (CSPs) represent another class of small soluble binding proteins (about 130 amino acids in length) that in some cases are also secreted in the lymph of insect chemosensory sensilla (Angeli et al. 1999). The sequences of CSPs are more conserved compared to OBPs with a specific motif of four cysteines that form two disulphide bridges between neighboring residues in the three dimensional structure of the protein. CSPs have been identified more broadly within the Arthropoda, including ticks,

myriapods and crustaceans (Pelosi et al. 2014). While OBP genes typically have several introns, CSP genes typically have a single intron. Although CSPs have been localized in the lymph of some sensilla, they are generally expressed in many different insect organs and there is no direct evidence that they participate in olfaction. Similar to OBPs, all CSPs have an N-terminal signal peptide that is cleaved upon secretion. Several CSPs that are highly expressed in the sensillar lymph are capable of binding to components of several different pheromonal blends *in vitro*. However, not all CSPs are restricted to olfactory organs and it has been postulated that they may also play a role in CO_2 detection, larval development and leg regeneration (Wanner et al. 2007).

ORs are a large diverse gene family primarily responsible for the molecular detection of odors in the insect's environment (Rutzler & Zwiebel, 2005; Vosshall & Stocker, 2007; Leal, 2013). They are believed to have evolved from gustatory receptors (GRs) as arthropods first adapted to a terrestrial environment (Robertson et al. 2010). The insect ORs, first identified from the *Drosophila melanogaster* genome, represent a novel form of chemoreceptor that function as ligand-gated ion channels (Clyne *et al*., 1999; Sato *et al*., 2008; Smart *et al*., 2008; Wicher *et al*., 2008). The OR coreceptor (ORco) is a highly conserved insect OR that acts as a chaperone and dimerization partner for other ORs that impart ligand specificity (Benton *et al*., 2006; Vosshall & Hansson, 2011). Together ORco + ORx form the ligand-gated ion channel. Supporting its critical function as a partner, olfaction in mutant fruit flies lacking ORco is broadly impaired (Larsson *et al*., 2004). ORs have been annotated from the sequenced genomes of several insect species, and range in number from 62 *OR* genes in *D. melanogaster* to 265 in *Tribolium*

castaneum Herbst (Robertson *et al*., 2003; Engsontia *et al*., 2008; Richards *et al*., 2008). In several cases, specific ORs that detect behaviorally important odors have been identified, including the receptor for the ‘queen substance’ 9-oxo-2-decenoic acid in *Apis mellifera* (Wanner et al. 2007b), the bombykol sex pheromone receptor in *Bombyx mori* (Sakurai *et al*., 2004) and the receptor in *Anopheles gambiae* that responds to components of human odor (Carey *et al*., 2010). *A. mellifera* and *Nasonia vitripennis*, also in the order Hymenoptera, have 174 and 225 *OR* genes, respectively (Robertson and Wanner 2006, Robertson et al. 2010). Remarkably, annotations from recently sequenced ant genomes have revealed as many as 407 OR genes (Zhou *et al*., 2012). ORs have been annotated from six hymenopteran species that all belong to the monophyletic ‘true’ wasp suborder Apocrita (Grimaldi and Engel 2005, Zhou et al. 2012). Sawflies, members of the basal, paraphyletic suborder Symphyta (Grimaldi and Engel 2005), are primitively phytophagous and considerably diverged from the more modern Apocrita wasp and bee lineages, suggesting that they have divergent ORs. We previously identified 28 antennal expressed *C. cinctus* ORs, of which one was expressed at higher levels in male antennae and five were expressed at higher levels in female antennae (Gress et al. 2013).

Ionotropic receptors (IRs) are a novel lineage of olfactory receptors that originate from a highly divergent subfamily of ionotropic glutamate recptors (iGluRs) that are found throughout the protosome branch of the animal kingdom (Benton et al. 2009, Croset et al. 2010). IRs were found in the coeloconic sensilla of *D. melanogaster* and are not expressed in the basiconic or trichoid sensilla where the ORs are found, nor were they co-expressed with ORco (Benton et al. 2009). Electrophysiological evidence has

demonstrated the existence of multiple types of coeloconic housing ORNs that respond to acids, ammonia and humidity (Yao et al. 2005). Initial expression and *in situ* hybridization studies in *Drosophila* revealed fifteen *IR* genes expressed in the antenna. Multiple *IR* genes are expressed within the same ORN in complex combinatorial patterns involving 2 to 5 different *IR* genes (Benton et al. 2009). Supporting their novel function as olfactory receptors, IRs are localized in the dendrite of ORNs and not at the nerve synapses, and their miss-expression is sufficient to confer novel odor responsiveness as measured by electrophysiology. Ten antennal-expressed IRs have been identified from *Apis mellifera* and *Nasonia vitripennis* (Croset et al. 2010). Unlike the ORs that do not exhibit detectable levels of sequence conservation between insect orders, six of these IRs were orthologous between *Drosophila* and *A. mellifera*, indicating that expression of this subgroup has been conserved across the 350 million years of evolutionary time that separates the insect orders Diptera and Hymenoptera.

Another class of soluble protein present in the sensillum lymph are the odorant degrading enzymes (ODEs) that include carboxylesterases (CCEs), glutathione-S-transferases (GSTs) and cytochrome P450s (P450s) (Vogt et al. 1985, 2003). These biotransformation enzymes are typically involved in metabolic processes and form larger multi-gene families and are conserved broadly in the animal kingdom. Some members of these large gene families have evolved functions related to olfaction. Once the odorant has reached the OR and triggered the signal transduction into a nerve impulse, it needs to be cleared from the lymph so that the sensitivity of the olfactory system is maintained and new signals can be detected. Continuous activation of the neuron with the same

odorant can cause a loss of sensitivity to that odorant. Additionally, not all odorants that enter a specific class of sensillum are detected by the ORs expressed in that sensillum, and clearing these deleterious odors from the lymph reduces background noise thus improving detection of cognate signals.

Annotating genes is the process of assigning biological information to the DNA sequences (Stein 2001). This process can provide information that links the organism's genetic blueprint to aspects of its biology. The most common annotation is to identify DNA sequences that encode proteins. When working with complementary DNA (cDNA) that has been synthesized from messenger RNA (mRNA), annotating the protein sequences simply involves identifying the open reading frame (ORF) that begins with a start codon and translates (using the standard genetic code) through to the stop codon. While simple in concept, in practice many transcripts are not complete or contain anomalous sequences due to experimental, sequencing or assembly errors. Annotating protein-coding sequences from genomic DNA involves identifying the protein coding exons, the non-coding introns and their splice junctions in the correct order that yields a predicted mature mRNA sequence. *Ab initio* gene prediction was an early bioinformatics development that used computer software and algorithms to predict exons and introns without experimental evidence (Yandell & Ence 2012). While these computational methods are fast they are also prone to errors, particularly with more complex gene structures. For this reason, manual annotation (also known as curation) by human expertise remains a common component of high quality gene annotations. *Ab initio* gene prediction models have now been combined with evidence driven annotation. As DNA

and protein sequence databases have grown, the sequences of homologous genes and proteins from related species can be aligned by sequence similarity and used to predict exon and intron regions. Now, massively parallel sequencing of mRNA using next generation sequencing technology, termed RNA-seq, enables the coding region of many thousands of an organism's genes to be identified quickly and inexpensively. RNA-seq data have the greatest potential to improve the accuracy of gene annotation (Yandell & Ence 2012). The combination of a variety of *ab initio* and evidence driven gene prediction methods into a single software package has been termed annotation pipelines (Stein 2001;Yandell & Ence 2012). A total of 131 olfactory-related genes have been annotated from the WSS (Chapter 4&5).

A second form of annotation is to assign putative functions to proteins based on the presence of conserved sequences and functional domains. Protein sequences from multigene families can be analyzed phylogenetically to identify lineage-specific trends, such as orthology and expansion through duplication. In this study the 131 olfactory-related genes of the WSS are further analyzed and characterized to provide additional functional information.

Materials and Methods

The 131 olfactory-related genes annotated in Chapter 4 were formatted into two FASTA files, one containing the nucleotide sequence of the ORF (Appendix A) and the other containing the translated protein sequence (Appendix B). Based on BLAST search identity (Table 4.2 in Chapter 4), the olfactory-related genes were further subdivided into

separate FASTA files for OBP, CSP, OR, IR, CCE, GST and P450 protein families. The protein sequences were used to create similarity alignments that were in turn used to create phylogenetic trees. Protein alignments were used to identify conserved domains that characterize each different protein family.

WSS ORs were aligned in MAFFT using a BLOSUM62 matrix (Katoh et al. 2005) and manually adjusted to minimize gaps in Jalview (Clamp et al. 2004, Waterhouse et al. 2009). They were then combined with published *N. vitripennis* and *A. mellifera* ORs and again aligned in MAFFT and gaps adjusted in Jalview (Clamp et al. 2004, Waterhouse et al. 2009). The alignment was imported into MEGA6 and a maximum likelihood tree was built using the JTT-matrix model (Tamura et al. 2013). Bootstrap values were calculated based on 1000 replicates and nodes with less than 50% bootstrap support were collapsed. The OR tree was rooted with Orco orthologs as the outgroup.

Phylogenetic analysis of the WSS OBPs was conducted by aligning their sequences with OBPs from *A. mellifera* and *N. vitripennis* using MEGA6 with MUSCLE (Tamura et al. 2013) and manual adjustment to minimize gaps. The alignment was then used to construct a maximum likelihood tree as described for the ORs. The CSPs, ODEs and IRs were then analyzed using the same method.

Results

Odorant Binding Proteins

Eleven genes encoding putative OBPs were identified from the wheat stem sawfly genome (Chapter 4). The translated proteins are all similar in size, ranging from 130 to 173 aa in length (predicted weights of 14 to 20 kDa) (Figure 5.1). The protein alignment illustrates conserved residues and domains typical of the OBP family (Pelosi 1998; Pelosi et al. 2014), including: a predicted signal peptide at the N-terminus; six predicated α-helices; and six conserved cysteine residues. These classical features are conserved in all of the CcOBPs with the exception of CcOBP3 and CcOBP10 (Figure 5.12). CcOBP3 has a substitution of C to S at conserved cysteine position 1 while CcOBP10 has a C to V substitution at cysteine position 6. The single loss of cysteine residues at C1 and C6 has also been reported for *N. vitripennis* OBPs (Vieira et al. 2012).

A phylogenetic analysis of CcOBPs along with known Hymenoptera OBPs (*A. mellifera* and *N. vitripennis)* reveals patterns of gene gain and loss (Figure 5.2). *C. cinctus* has the fewest OBPs between these three species with a ~ 40% reduction compared to *A. mellifera*, possibly due to its basal location in Hymenoptera. In two cases there is strong evidence of conserved 1:1:1 orthologous OBP lineages between the three species. The first being the grouping of CcOPB2:NvOBP2:AmOBP10 with 99% bootstrap support and the second grouping of CcOBP4:NvOBP76:AmOBP1 with 95% support. CcOBP3 + NvitOBP64 and CcOBP10 + NvitOBP26 also form orthologous lineages. In other cases the analysis provides evidence of lineage specific gene expansion. There appears to have been an expansion of OBPs in the Hymenoptera lineage leading to

N. vitripennis, which is predicted to have more than 7 times as many OBP genes compared to the WSS (Figure 5.2). In the analysis, clade members are found in both honeybee and sawfly and *Nasonia* and sawfly: CcOBP12 groups with AmOBP9 with no *Nasonia* ortholgue; and, CcOBP10 grouping with NvOBP26, 18 and 17 but no honeybee orthologues.

Chromosomal clustering of OBP genes is common and was first observed in *D. melanogaster* (Galindo and Smith 2001, Hekmat-Scafe et al. 2002, Vogt et al. 2002). Subsequently clustering of OBP genes has been demonstrated in all 12 *Drosophila* genomes (Vieira et al. 2007), with 69% of OBP genes arranged in 10 clusters of 2-6 genes each. This clustering is also observed in *A. mellifera* where OBPs are arranged in two main clusters of 7-9 genes along with some isolated loci (Forêt and Maleszka 2006). Most of the OBP genes from the wheat stem sawfly are also organized into small clusters within the genome. Only 4 genes, CcOBP2-4 & 6, are represented by isolated loci. Six OBP genes are found as tandem pairs in a head-to-tail orientation. CcOBP1 and CcOBP13 are arranged in tandem head-to-tail on scaffold gi|453039718 approximately 1 Kb apart. However, they only share 27% amino acid identity. Another set of OBPs, CcOBP5 & 10, are in tandem head-to-tail on opposite DNA strands, ~2 Kb apart on scaffold gi| 453039195 but only 16% amino acid identity. Finally, CcOBP 7 & 12 are arranged in tandem head-to-tail about 2.6 Kb apart on scaffold gi| 453039820 and have only 22% amino acid. The low identity of these tandem OBP genes suggests relatively ancient gene duplication events.

Chemosensory Proteins

Eight CSP genes were annotated from the WSS. Their protein alignment (Figure 5.3) illustrates the hallmark characteristics of the family: relatively short sequences (103 to 138 aa long, predicted 12-16 kDa); an N-terminal signal peptide; a "YTTKYDN [VI] [ND] [LV] DEIL" motif near the N-terminus; and 4 conserved cysteine residues. The 4 cysteine motif is strictly conserved with a (CX6CX18CX2C) spacing that in the 3D protein structure forms two disulphide bridges (Campancacci et al. 2003). CcCSP5 showed variation from this spacing with two additional aa inserted between the first two conserved cysteine residues (CX8CX18CX2C) (Fig. 5.3). Similar orthologs have been identified from *L. migratoria* and there are similar variations in this motif in some CSPs reported from honeybees, ants and the jewel wasp (Calvello et al. 2003, Ozaki et al. 2005, Forêt et al. 2007).

The phylogenetic analysis (Figure 5.4) reveals little evidence for recent expansions of CSP genes in the WSS lineage. The median identity between WSS CSP sequences was only 34% and the median similarity was only 46.5%. None of the CcCSPs clustered into a clade with more than 50% bootstrap support. An ML-tree analysis with the addition of *A. mellifera*, *N. vitripennis* and *N. giraulti* CSPs as well as *C. cinctus* CSPs was conducted (Fig. 5.4). NgirCSP4 and CcCSP5 diverge from the typical conserved CSP motifs and they form their own monophyletic group with high bootstrap support and both sequences show the same divergent variation in spacing (CX8CX18CX2C). The other CSPs in the analysis show two distinct paralogous groupings indicating duplication before speciation and may have ancient lineages with CcCSP2, 2b, 4, 6 and 7 in one

group and CcCSP1 and 3 in a second more divergent group. CcCSP6 and NgirCSP3 show a 1:1 orthologous relationship with high bootstrap support within the first paralogous grouping and CcCSP6 is divergent as it contains two introns versus the typical one for this family as does NgirCSP3. CcCSP1 and 3 are contained in a divergent grouping of CSPs, the majority of CSPs share a common motif at their N termini YTTKYDN [VI] [ND] [LV] DEIL. This group contains a truncated C-terminus and is missing the sixth alpha helix in the 3D structure, and has Y to Q substitution that is conserved at the third position before the first conserved cysteine.

CSP genes all have a single intron and are frequently arranged in clusters within insect genomes. In *C. cinctus*, 6 CSPs are arranged in head to tail tandems. One tandem consists of CSP2, CSP2b and CSP3 separated by 251kb on scaffold gi|453039999; all oriented in the same direction with CSP2 and 2b less than 4kb apart. The second tandem consists of CSP5, CSP6 and CSP7 all located within 17kb of one another on scaffold gi|453039742 with csp5 orientated on the plus strand and csp6 and 7 on the minus strand. Only 2 genes are represented by single loci, csp1 and csp4. Six of the eight CSPs have a single intron located 25 bases downstream of the second conserved cysteine. This intron is also conserved in the CSPs of most other species including honeybee, *D. melanogaster, A. gambiae* and *T. castaneum* (Forêt et al. 2007). CSP4 and CSP6 both contain 2 introns, one located 25bp after the second cysteine and the second intron is located 45bp after the first cysteine in CSP4 and 24bp after the last cysteine in CSP6 (which forms a confident orthologue with NgirCSP3).

Odorant Receptors

A total of 53 ORs were annotated from the WSS (Chapter 4). The protein alignment of all WSS ORs (Figure 5.5) demonstrates the features typical of this family, including: lengths ranging between 380 – 420 aa; 6 – 8 predicted transmembrane domains; and a short 3' exon that contains the only conserved domain that can be detected in most insect ORs, D[L|M|V]SLETF[T|G]S[I|V][L|][K|S]T[S|A]FSY[L|F][N|T][L|V]LR (Miller & Tu, 2008).

The WSS ORs were analyzed phylogenetically along with the published ORs from *A. mellifera* and *N. vitripennis* (Robertson et al. 2010). Some ORs can be too divergent and lack sufficient phylogenetic information to form clades with at least 50% bootstrap support, such as CcOr7, 11 and 18. The phylogenetic analysis (Figure 5.6) reveals at least 15 OR clades that exhibit a variety of gene lineage patterns. Gene lineages that are conserved between all species are one common pattern. This includes 8 WSS ORs (CcOr9, 17, 46, 49, 50, 51, 52 & 53) that group with thirty-seven *Nasonia* ORs and seven *A. mellifera* ORs, a multigene lineage conserved in all three species. Other examples of this pattern include a group of 5 WSS ORs (CcOr39, 42, 47 & 58) with 23 *Nasonia* ORs and three honeybee ORs; CcOr10 and 30 that group with 13 *Nasonia* ORs and AmOr141 with high bootstrap support; and, CcOr20 that groups with 11 *Nasonia* and 17 *Apis* Ors.

A second pattern is clades that are conserved in some but not all species lineages. CcOr5 groups with 19 *Nasonia* ORs but *Apis* orthologues are lacking, an example of an apparent expansion in the lineage leading to *Nasonia* but not *Apis*. In contrast, nine WSS ORs (CcOrs6, 16, 24, 25, 26, 26b, 27, 36 & 37) group with twenty-three honeybee ORs without any *Nasonia* orthologues. CcOr4 is a second example, grouping with AmOr168, 169 and 170 without *Nasonia* orthologues.

A less common pattern within multigene families is clades with 1:1 orthology between species. Three ORs form a clade with 1:1:1 species orthology; CcOr8, NvOr296 and AmOr16. In contrast, CcOr34 is in 1:1 orthology with AmOr120 but lacks a *Nasonia* orthologue, suggesting this lineage was lost in *Nasonia*. A second example of this pattern is a 1:1 orthology between CcOr19 and AmOr119 with no representation by *Nasonia*. Similarly, CcOr45 form a highly supported clade with both NvOr78 and 79 but *Apis* orthologues are lacking, suggesting a conserved single gene orthology that was lost in the *Apis* lineage. A 1:1 orthology between CcOr38 and NvOr211 but not *Apis* is a second example of this pattern.

Lineage specific gene expansion results in monophyletic clades. One monophyletic sawfly clade contains 8 Ors (CcOr12, 13, 14, 21, 31, 33, 40 and 44). This group is found within a larger clade that includes CcOr41, 7 *Nasonia* Ors and 52 honeybee Ors, suggesting a large expansion of this gene lineage in *Apis*. A second example of this pattern includes a monophyletic cluster of seven WSS ORs (CcOr3, 32, 48, 54, 55, 56 and 60) that form a larger clade with CcOr1, 15 and 20 along with four *Nasonia* Ors and AmOr160.

Like the *N. vitripennis* and *A. mellifera* ORs (Robertson et al. 2010, Robertson & Wanner 2006), the *C. cinctus* ORs were found clustered in tandem arrays. Four main clusters of ORs are each located on 4 different scaffolds. One cluster on scaffold gi|453038779 contains 8 OR genes (CcOr3, 15, 29, 32, 47, 48, 54, 56) in one main tandem section spanning 14 Kb. Gene clusters often share similar features such as exon number and sequence similarity. With the exception of OR47, all members of this cluster also group together in the same phylogenetic clade (Figure 5.2). All but two of these ORs have six exons; CcOr48 and 54 have 5 exons. CcOr48, 54 and 56 form a tip of the branch, with 48 basal to 54 & 56. Based on their gene structures, it is more likely that CcOR56 is basal to 48 & 54, and the loss of an intron in this clade is a derived trait. Another cluster on scaffold gi|453039299 contains 7 ORs (CcOr16, 25, 26, 26b, 27, 36 & 37) in head-to-tail tandem orientation within a 35 Kb section. All ORs in this cluster form a clade in the phylogenetic tree and all but CcOr36 have 6 exons. CcOrs36 and 37 form a terminal pair in the tree, and the five exons that comprise Or36 suggests the recent loss of an intron in this gene lineage. Four ORs (CcOr9, 17, 46 and 51) form a cluster on scaffold gi|453038887 in the minus strand orientation over a region spanning 21 Kb and 6 ORs (CcOr12, 13, 31, 40, 44 and 59) form a cluster on scaffold gi|453039847 in the plus orientation over a region of 12 Kb. In both cases the clustered ORs also form a clade in the phylogetic tree (Figure 5.2) and the genes have 5 exons with a single exception in each clade. CcOR45 has gained two exons while CcOr31 has lost an exon.

Ionotropic Receptors

Ten IR genes were annotated from the WSS (Chapter 4). An alignment of their protein sequences demonstrates characteristics of this protein family, including a predicted bipartite ligand-binding domain (LBD) at the N-terminus (Figure 5.7). This domain is separated by an ion channel domain and a short cytoplasmic C-terminal region (Mayer, 2006, Benton et al. 2009). An alignment of the sawfly IRs with10 IRs from *N. vitripennis* and *A. mellifera* shows conservation of the pore region, transmembrane segment of the ion channel as well as the two ligand binding domains (Figure 5.7) as reported by Benton et al. (2009). Within *Drosophila* IRs, sequence identity ranged from 10 to 70% (Benton et al. 2009). Similarly, WSS IRs exhibit amino acid sequence identity ranging from 20 to 57% with an average of 35%.

A phylogenetic analysis of the *C. cinctus*, *N. vitripennis* and *A. mellifera* IR sequences was completed (Figure 5.8). Six WSS IRs (CcIR1 – 3, 7, 15 & 16) form 1:1:1 orthologous clades with *Apis* and Nasonia *IRs*, supporting relatively old origins of these lineages within the Hymenoptera. Two clades illustrate lineage specific loss of an IR gene. CcIR11 groups with an *Apis* IR but not a *Nasonia* representative and CcIR4 groups with a *Nasonia* IR but not an *Apis* representative. Some limited lineage specific duplication is observed for *Apis* and *Nasonia* but not the WSS. Consistent with the lack of evidence for recent IR gene duplication, the WSS IR genes are not arranged in clusters or tandem arrays with the genome.

Odorant-degrading enzymes

Forty-nine candidate ODEs representing three protein families were annotated from the WSS genome: 12 CCEs, 8 GSTs and 29 P450s.

Carboxylesterases

CCEs have been characterized as three main functional classes that broadly represent dietary/detoxification, hormone/semiochemical processing and neurological/developmental functions (Oakeshott et al. 2010). All 12 CcCCEs are members of the hormone/semiochemical processing class (termed β-esterases). An alignment of the *C. cinctus* CCEs (Figure 5.9) illustrates the characteristics typical of β-esterases, including: protein lengths ranging from 524 to 588 aa (58.3 to 66.6kDA); conserved RF and DQ residues for the nucleophilic elbow; catalytic triads (S-D/D-H); N-terminal cysteine; and an oxyanion hole (Figure 5.9). CcODE4b was lacking the DQ motif, the N-terminal cysteine residue and part of the oxyanion hole.

A phylogenetic analysis of the β-esterase CCEs from *C. cinctus*, *A. mellifera*, and *N. vitripennis* (Figure 5.10) reveals lineage specific gene expansion and little orthology between the three species. Twelve of the 15 CCEs from *Nasonia* group into a single clade with two *Apis* CCEs. Eight of the 12 WSS CCEs form a monophyletic lineage (CcCCE2, 4, 4b, 8, 10a, 10b, 10c and 10d) with high bootstrap support. CcCCE1 formed an orthologous clade with one *Apis* and two *Nasonia* CCEs (Figure 5.10).

Both *A. mellfera* and *N. vitripennis* genomes displayed CCEs in small clusters of two to four genes (Oakeshott et al. 2005, 2010). Supporting their more recent expansion, five the WSS CCEs in the monophyletic clade displayed in Figure 5.10 were also clustered on the same scaffold. CcCCE9, 10a, 10b, 10c, and 10d were organized in tandem across a 351 Kb region. CcCCE9 was on the plus strand and CCE10a -10d on the minus strand, on scaffold gi|453039904. CcCCE4, 4b and 8 form a branch within this monophyletic CCE clade, and their genes are clustered on scaffold gi|453039143 within a 3.3 Kb region. The CCEs within these clusters have 52-69% amino acid identity.

Glutathione-S-transferases

The alignment of the 8 WSS GSTs with *A. mellifera* and *N. vitripennis* GSTs (Figure 5.11) illustrates typical characteristics of the GST family, including: protein length ranging from 600 to 729 aa (22.8-27.8kDa); and conserved N-terminal and C-terminal domains. The cytosolic GSTs are homo- or heterodimeric proteins that are formed by two subunits. Each subunit folds into two domains, the N-terminal and C-terminal, joined by a variable linker. The N-terminal domain adopts a similar conformation to the thioredoxin domain found in all GST structures (Sheehan et al. 2001) and consists of active G-sites. The larger C-terminal domain consists of a variable number of alpha helices or substrate binding sites. All the WSS GSTs contained the N- and C-terminal domain structures (Figure 5.11).

Phylogenetic analysis of sawfly GSTs along with honeybee and *Nasonia* GSTs revealed six functional classes of GST found in insects (Figure 5.12). Unlike *Nasonia* and *Apis*, Delta class GSTs were the most commonly represented class of WSS GST. Four of

the eight GSTs identified were Delta class, including CcGST1a, 1b, 1c and 2, that all form a clade with NvGSTD1, D3 and D5 and AMGSTD1. Delta class GSTs are one of the largest classes of GSTs in non-hymenotperan insect orders, but are sparsely represented in *N. vitripennis* (Oakeshott et al. 2010). One WSS member of the theta class of GSTs was identified, CcGST4, that formed a strong clade with NvGSTT1, T2 and T3 as well as AmGST theta4. One omega class WSS GST was identified, CcGST7, that grouped strongly with NvGSTO1, O2 and AmGST omega1. Lastly, two sigma class WSS GSTs were identified, CcGST5 and 6 that formed a large clade with four *Nasonia* and 5 *Apis* GSTs. No WSS GSTs belonging to the zeta and epsilon classes of GST were found.

The Delta class is the largest GST class in the sawfly. Four of its 8 GSTs belong to this class. They are located on two scaffolds with three GSTs, CcGST1a-1c located in head to tail tandem on scaffold gi|453038992 over an 11kB region that share 46.8% identity. CcGST2 is located as a single loci on scaffold gi|453039665. The Sigma class GSTs were the next largest group in the sawfly and had 2 members CcGST5 and 6 both of which were represented by single loci on scaffolds gi|453039005 and gi|453039828 respectively and share only 37.8% identity. All remaining GSTs are represented by single loci.

Cytochrome P450s

A total of 29 cytochrome P450 genes were annotated from the WSS (Chapter 4). An alignment of their protein sequences (Figure 5.13) reveals typical characteristics

shared by the P450 family, including: Proteins ranging in size from 338 to 597 aa (38.7-68.5kDa); conserved "ETLR" and "PERF" sequence domains; and, a haem binding domain (consensus PFXXGXXXCXG). All but two of the 29 P450s annotated from the WSS contained the conserved ETLR, PERF and haem domains (Figure 5.13). CcP450_17 and 31 were missing the PERF and haem binding domains. However, both of these annotated genes were incomplete at the C-terminus end of the predicted protein.

Phylogenetic analysis of the *C. cinctus* P450s, along with published P450s from *Apis* and *Nasonia*, revealed four distinct functional classes (Figure 5.14). The mitochondrial, CYP2, CYP3 and CYP4 classes of P450 are named after their prominent vertebrate members (Figure 5.14). The CYP3 class includes two large families named CYP6 and CYP9. Ten of the 29 WSS P450s group into the CYP3 class; CcP450_14, 20, 21, 22, 23 and 27 group with *Nasonia* and *Apis* CYP6 genes while CcP450_15, 18, 19 and 24 group with CYP9 family of P450s. In the CYP6AS clade the *A. mellifera* P450 genes all arose from duplication that occurred after the divergence of WSS from bees and wasps (Figure 5.14).

CYP3 and CYP4 P450 classes are abundant in *Drosophila* and other dipterans, accounting for ~ 40% of the P450 genes (Claudianos et al. 2006). The *A. mellifera* genome shows a reduction in the CYP4 class having only four (AmP450 CYP4AZ1, CYPAA1, CYP4AV1 and CYP4G11) compared to *N. vitripennis* that has 30 CYP4 members. The 29 WSS P450s included 8 CYP4 members (CcP450_7, 8, 9, 10, 11, 12, 13 and 26). The mitochondrial P450s show a high level of conservation and 1:1 orthology with mitochondrial P450s from other insect genomes (Fig. 5.14). CcP450_16, 25, 28, 30

and 31 have 1:1:1 orthology with mitochondrial P450s from *Apis* and *Nasonia*, a phylogenetic pattern typical of housekeeping genes. CcP450_1, 2, 3, 5 and 6 belong to the CYP2 class of P450s, with strong bootstrap support for orthology with their *Apis* and *Nasonia* counterparts.

Six of the 29 WSS P450s are arranged as tandem pairs on a single genomic scaffold, and each pair belongs to the same functional class: CcP450_2 & 5 (CYP2); CcP450_16 & 28 (mitochondrial class); CcP450_18 & 19 (CYP3); and, CcP450_ 21 & 22 (CYP3). CcP450_2 and 5 cluster on scaffold gi|453039500 in the same direction and share 21.8% identity. CcP450_16 and 28 cluster on scaffold gi|453039545 on opposite strands and share 44.2% identity.

Discussion

A total of 131 olfactory-related proteins from the wheat stem sawfly were characterized and identified as members of seven different gene families. Their identity as OBPs, CSPs, ORs, IRs, CCEs, GSTs and P450s was confirmed by analyzing their sequences for conserved domains typical of each family, and by their phylogenetic relatedness to gene family members from other insects. Their accurate identification was facilitated by high quality annotation; most of the sequences were represented by a complete ORF (Chapter 4). The function of some gene families is dedicated to olfaction, such as the OBPs and ORs, and the majority of these genes were expressed in the WSS antennae. The 11 OBPs and 53 ORs are judged to be the complete repertoire encoded in the WSS genome; tBLASTn searches of the genome failed to detect any other unique loci

with homology to these gene families. Similarly, the 8 WSS CSPs appear to represent the complete set of genes in this family. The IRs, CCEs, GSTs and P450s form large gene families that include functions other than olfaction. The IRs, for example, are ligand gated ion channels that are commonly involved in neurotransmitter signaling; a subset of variant IRs expressed in sensory neurons have evolved an olfactory function. The CCE, GST and P450 families are involved in biotransformation reactions, detoxification and metabolism. A smaller fraction of these gene families have evolved olfactory functions, by degrading odors in the sensillum lymph. Ten IR, 12 CCE, 8 GST and 29 P450 genes were identified from the antennal transcriptome. Annotation efforts focused on these olfactory-related transcripts and the complete gene families encoded in the WSS genome were not annotated.

The 11 OBP genes annotated from the WSS constitute one of the smallest numbers observed for this family compared to other insects including Hymenoptera. The *A. mellifera* genome encodes 21 OBP genes and the *N. vitripennis* genome encodes 90, while the mosquito *Anopheles gambiae* has 65. A possible explanation for this reduced number of OBPs may be the basal location of sawflies within the Hymenoptera. Alternatively, odorant transport proteins may be encoded by other genes families, such as the chemosensory protein (CSP) family. It has been suggested that in hymenopterans CSPs are more olfactory-specific than the OBPs (Ishida et al. 2002; Calvello et al. 2005). Interestingly, we did not find any ‘‘plus-C’’ or “minus-C” OBP lineages (Zhou et al. 2004) and all WSS OBPs belonging to the ‘‘classic’’ subgroup (Hekmat-Scafe et al. 2002). The highly conserved pattern of six cysteine residues in insect OBPs is very

important for the formation of three disulphide bonds that stabilize the 3-D protein structure (Leal et al. 1999; Scaloni et al. 1999). Classic OBP structure is composed of 6 cysteine residues forming 3 disulphide bonds, although other forms have also been reported (Lagrade et al. 2011; Lagrade et al 2011b). Interestingly, two WSS OBPs have lost one of the conserved cysteine residues, a trend also reported for some *Nasonia* OBPs (Vieira et al. 2012). The loss of a disulphide bond is thought to make the binding pocket more flexible to larger ligands.

The eight WSS CSP genes correspond in number to the 6 CSPs from *A. mellifera* and 9 CSPs from *N. vitripennis*. Danty et al. (1998) were the first to identity CSPs in the honeybee with partial sequencing of CSP1 and CSP3. Further characterization of these genes was conducted in two additional studies (Briand et al. 2002, Kamikouchi et al. 2004). The 3-D structure of CSPs resembles that of OBPs, and a similar function of transporting hydrophobic odors across the sensillum lymph has been proposed. This hypothesis is based on three lines of evidence: their specific localization to the sensillar lymph (Angeli et al. 1999; Nagnan-Le Meillour et al. 2000; Monteforti et al. 2002), the globular 3-D structure with six α-helices surrounding an internal hydrophobic binding channel, and the ability to bind a variety of hydrophobic odors (Bohbot et al., 1998; Nagnan-Le Meillour et al. 2000; Jacquin-Joly et al. 2001; Briand et al. 2002; Lartigue et al. 2002; Campanacci et al. 2003; Mosbah et al. 2003). Nevertheless, insect CSP genes are also expressed in other tissues including the legs, head, thorax, wings, labial palps, tarsi, proboscis, pheromone glands and ejaculatory ducts. The wide tissue distribution,

even in many tissues that lack sensilla, indicates that they have additional non-olfactory functions.

The insect OR family is characterized by extensive lineage-specific gene duplication and gene loss. With the advent of whole genome sequencing and the analysis of several insect genomes, large regions of clustered duplicated olfactory-related genes have been identified (Robertson & Wanner, 2006; Richards et al., 2008; Robertson et al., 2010; Vieira et al. 2012). We annotated 53 OR genes from the WSS, considerably less compared to other Hymenoptera; including *A. mellifera* (n=164) and *N. vitripennis* (n=225) but similar to *D. melanogaster* gene (n=64). Additionally, the 53 WSS Ors were surprisingly diverse, and formed the basis of many of the more modern and expanded hymenopteran Apocrita lineages represented by *A. mellifera* and *N. vitripennis*. This pattern suggests that the ancestor of sawflies and modern Hymenoptera possessed a diversified repertoire of OR lineages that were not lost in the more modern species of Hymenoptera. The lack of OR gene expansion in the WSS may reflect a relatively stable and conserved life history as a stem mining insect of grasses.

The 10 WSS IR genes are fewer than those identified in *D. melanogaster* (n=66) but similar to the number identified in *Nasonia* (n=9) and *A. mellifera* (n=10) (Benton et al. 2009; Croset et al. 2010). Previous analyses have highlighted IR25a as an atypical member of the IRs, displaying deep phylogenetic conservation and broad expression in many olfactory and gustatory neurons (Croset et al. 2010). In our analysis we see 1:1:1 orthology of IR25a across all three Hymenoptera species as we do with IR8a. It has been proposed that IR25a and IR8a act as co-receptors, analogous to the heteromeric assembly

of iGluR subunits into functional complexes (Gereau & Swanson, 2008) and similar to the pairing of ORs with Orco (Benton et al. 2006; Larrson et al. 2004). In addition to IR25a and IR8a, many of the antennal IRs are highly conserved in insects, both in sequence and expression pattern (Croset et al. 2010). This is in sharp contrast to the insect ORs that have no detectable orthology between the insect orders, with the sole exception of the highly conserved OR co-receptor (Jones et al. 2005).

Fourteen CCEs were identified from the WSS antennal transcriptome, almost three times as many CCEs as found in the genome of *A. mellifera* (n=5) and similar to that found in *N. vitripennis* (n=17) (Oakeshott et al. 2010). Claudianos et al. (2006) hypothesized that the functional significance of odorant degrading enzymes can vary significantly between difference insect species. The reduced number of detoxification enzymes from the honeybee genome is thought to be a consequence of their eusocial structure. The larger number of CCE genes from a basal and phytophagous Hymenoptera species provides some support for this hypothesis.

The WSS has 8 GST genes, comparable to the honeybee (n=10) but fewer than the jewel wasp (n=18). For a phytophagous insect, 8 GST genes represents quite a low number compared to other insect genomes and indicates a modest detoxification capability (*Drosophila* 12 Genome Consortium, 2007; International Silkworm Genome Consortium, 2008; *Tribolium* Genome Sequencing Consortium, 2008). This could relate to the relatively simple life cycle of the WSS as it develops within grass stems feeding on parenchyma cells. Among the six classes of GSTs, the Delta class, along with the Epsilon class, are unique to insects and contain the majority of GSTs associated with

detoxification of insecticides (Tu and Tang 1994; Ranson et al. 2001; Wei et al. 2001). A Delta class GST in *M. sexta* is specifically located in the antennae and involved in odor degradation (Rodgers et al. 1999). The sawfly Delta GSTs CcGST1a, 1b, 1c and 2 may represent specific odor degradation enzymes; expression and functional characterization of these GSTs is needed to verify this prediction. WSS has two GSTs in the Sigma class (CcGST5 and 6); this is the most highly represented class of GSTs in both honeybee and *Nasonia* (Claudianos et al. 2006, Oakeshott et al. 2010). Some members of the insect Sigma class have activity against 4-hydroynonenal, a by-product of lipid peroxidation, and are thought to play a role in protection from oxidative stress (Singh et al. 2001).

Cytochrome P450s can play an important role in the degradation of odorant signals after recognition of the odor molecule at the receptor site (Leal, 2013). However, the primary function of P450s is detoxification and other metabolic functions. Two WSS orthologues (CcP450_29 and 30) may be related to the CYP314A and 315A families that participate in the biosynthesis of 20-hydroecdysone (Rewitz et al. 2007), an important insect hormone. CcP450_17 is in the mitochondrial class of P450s, and has a single orthologue present in *N. vitripennis* but not in *A. mellifera*. Members of this CYP12 clade have been implicated in insecticide resistance in the housefly *M. domestica* and the fruitfly *D. melanogaster* (Guzov et al. 1998; Brandt et al. 2002; Bogwitz et al. 2005). Additionally, members of the CYPAA and CYPAG clades (CcP450_7, 15 and 24) in other species have been found to function in pheromone metabolism and pyrethroid insecticide resistance (Maibeche-Coisne et al. 2000; Pridgeon et al. 2003). Furthermore, dipteran and lepidopteran CYP6s are involved in resistance to a broad range of

insecticides and function to detoxify host plant allelochemicals within the insect's gut (Danielson et al. 1997; Mittapalli et al. 2005; Wen et al. 2005; Li et al. 2007). Some evidence shows that CYP9s may also be involved in detoxification of insecticides and other allelochemicals, as well as semiochemical degradation (Stevens et al. 2000; Poupardin et al. 2008; Malbeche-Coisne et al. 2005).

CcP450_5 and 6 are members of the CYP306 and 307 clades that are involved with 20-hydroxyecdysone biosynthesis (Rewitz et al. 2004). The CYP15 clade, that includes juvenile hormone expoxidase, functions in insect metamorphosis (Helvig et al. 2004). Interestingly, the CYP15 clade is conserved in *N. vitripennis* and *A. mellifera* but not in *C. cinctus*.

The potential of 'reverse chemical ecology', using the proteins that detect odors in the peripheral olfactory system to identify behaviorally active compounds, has been demonstrated by Leal *et al*. (2008). The discovery of the first insect OR sequences from the partially sequenced *D. melanogaster* genome (Clyne et al. 1999; Vosshall et al. 1999) and the demonstration of their preeminent role in detecting and discriminating odors (Hallem et al. 2004) has facilitated progress in non-model pest insect species. Gene expression levels have been used to identify and prioritize candidate olfactory genes, such as the ORs, that detect behaviorally important odors. Typically, ORs that are expressed predominantly in the antennae of one sex (sex-biased expression) respond to odors that mediate behaviors specific to that sex. However, in some cases, insect pheromones are attractive to, and can be detected by, both sexes. Rather than using sex-biased expression, Mitchell et al. (2012) took advantage of the fact that sex pheromone

receptors also tend to be highly expressed relative to other ORs. A total of 57 OR transcripts were identified from the antennae of the cerambycid beetle *Megacyllene caryae*, and by functionally testing only the five most common transcripts the authors identified two pheromone receptors (Mitchell et al. 2012). All protein families involved in odor detection, including ODEs, have been suggested as molecular targets for the disruption of olfactory-mediated behaviors. The next step in this project is to identify which of the candidate WSS olfactory related genes are more highly expressed in the antennae, and/or exhibit sex biased expression, as potential targets.

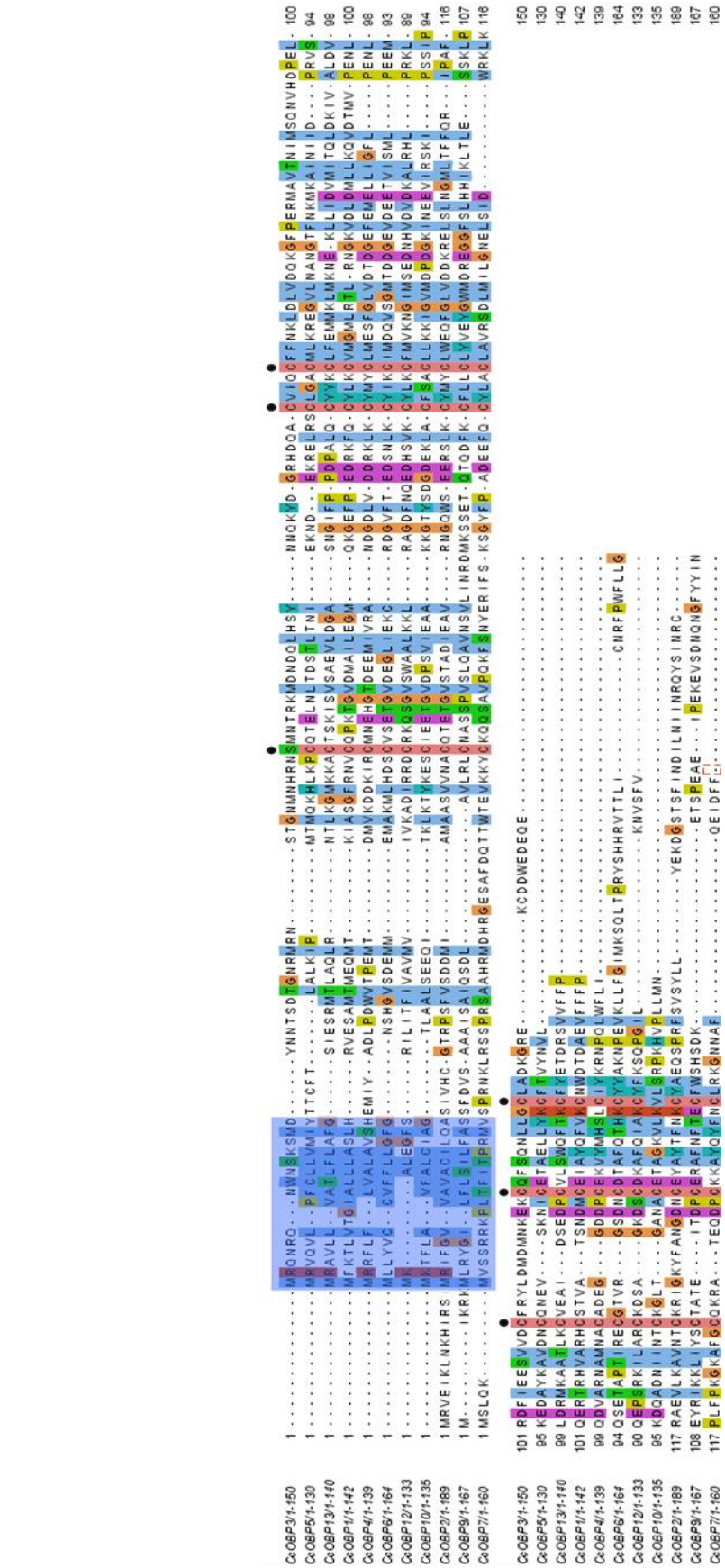
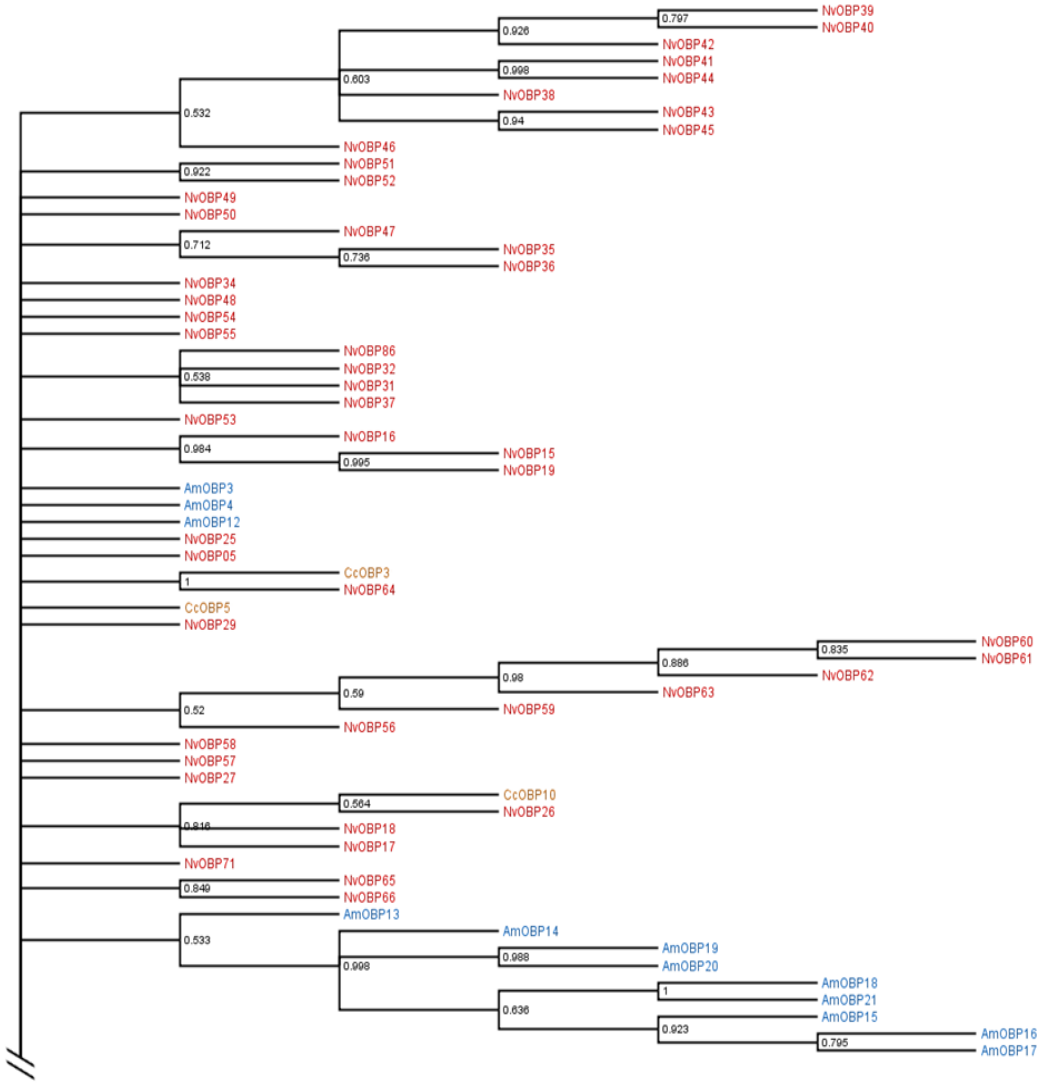


Figure 5.1 Amino acid sequence alignment of wheat stem sawfly (WSS) OBPs annotated from the antennal transcriptome and WGS. A total of 11 WSS OBPs were aligned using CLUSTALX set to default parameters, and the alignment was used to judge the completeness of the WSS OBPs. The signal peptide is highlighted in blue and was determined using SIGNALP 4.1. The 6 conserved cysteine residues characteristic of OBPs are highlighted with a black circle.


NvOBP39
NvOBP40
NvOBP42
NvOBP41
NvOBP44
NvOBP38
NvOBP43
NvOBP45
NvOBP46
NvOBP51
NvOBP52
NvOBP49
NvOBP50
NvOBP47
NvOBP35
NvOBP36
NvOBP34
NvOBP48
NvOBP54
NvOBP55
NvOBP86
NvOBP32
NvOBP31
NvOBP37
NvOBP53
NvOBP16
NvOBP15
NvOBP19
AmOBP3
AmOBP4
AmOBP12
NvOBP25
NvOBP05
CcOBP3
NvOBP64
CcOBP5
NvOBP29
NvOBP60
NvOBP61
NvOBP62
NvOBP63
NvOBP59
NvOBP56
NvOBP58
NvOBP57
NvOBP27
CcOBP10
NvOBP26
NvOBP18
NvOBP17
NvOBP71
NvOBP65
NvOBP66
AmOBP13
AmOBP14
AmOBP19
AmOBP20
AmOBP18
AmOBP21
AmOBP15
AmOBP16
AmOBP17

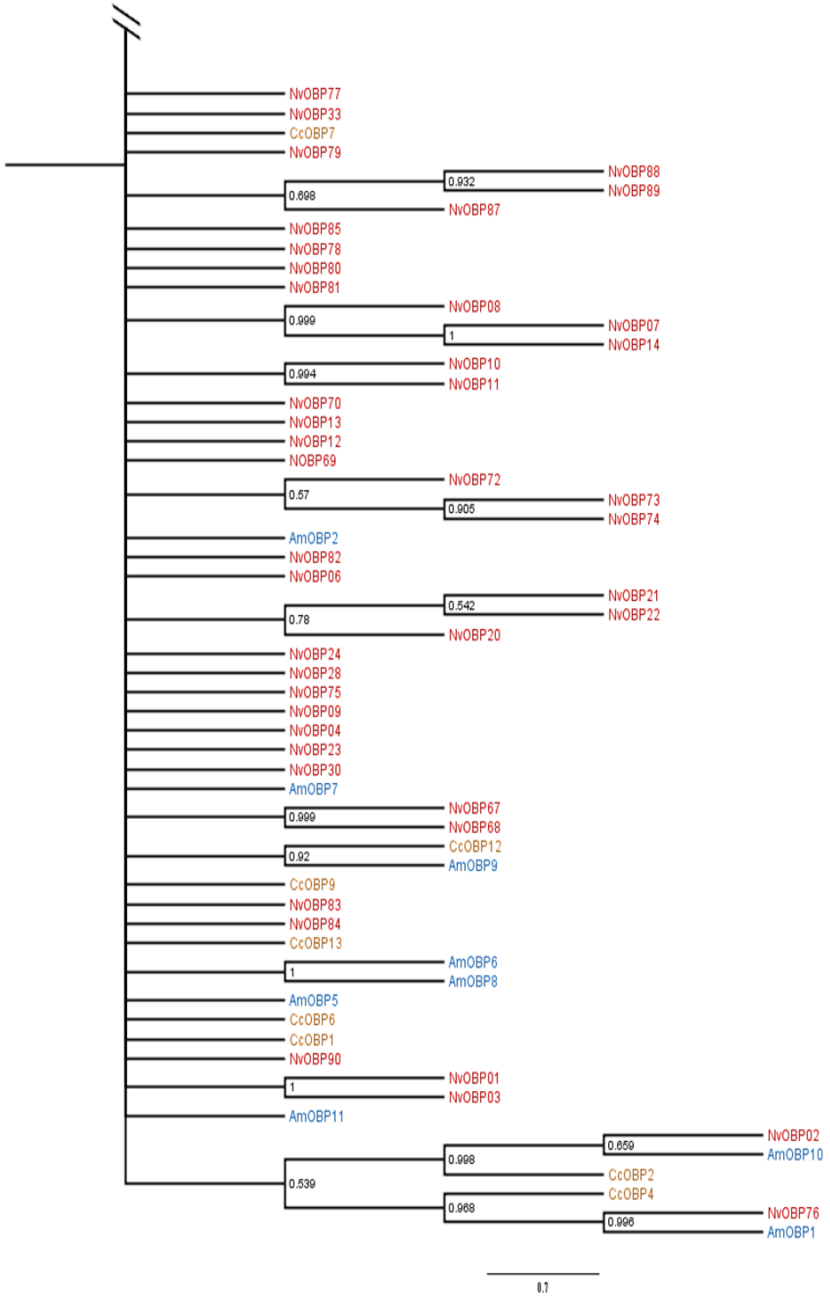


Figure 5.2 Phylogenetic Relationships of 11 wheat stem sawfly (WSS) OBPs with jewel wasp and honeybee OBPs inferred from Maximum Likelihood analysis. Bootstrap values are provided at significant branch points and nodes with less than 50% support are collapsed. Jewel wasp, red font, honey bee, blue font, WSS, brown font.

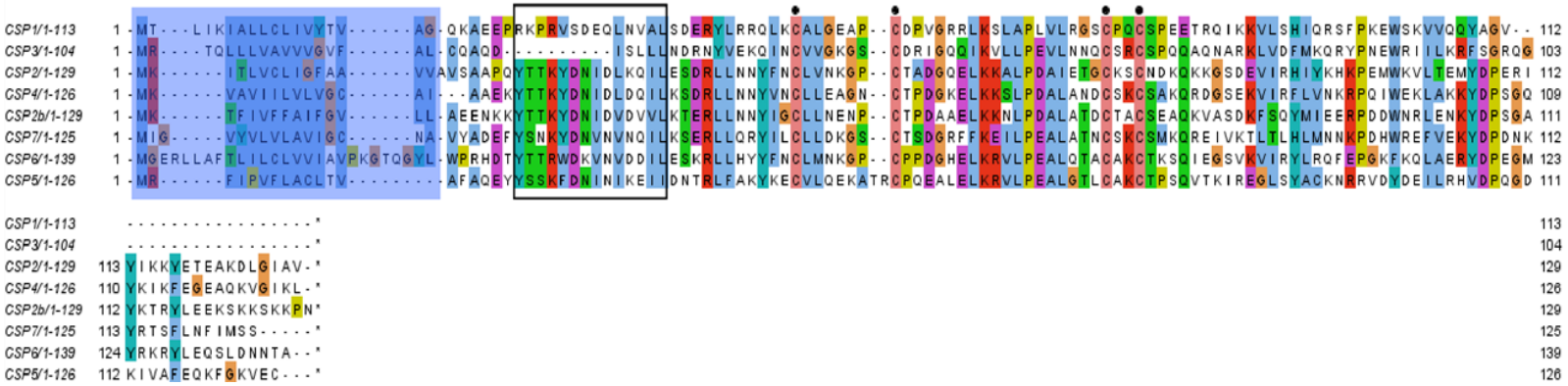


Figure 5.3 Amino acid sequence alignment of wheat stem sawfly (WSS) CSPs annotated from the antennal transcriptome and WGS. A total of 8 WSS CSPs were aligned using CLUSTALX set to default parameters and YYTKDN[VI][ND][LD]DEIL is highlighted with a black box and the 4 conserved cysteine residues are highlighted with a black circle.

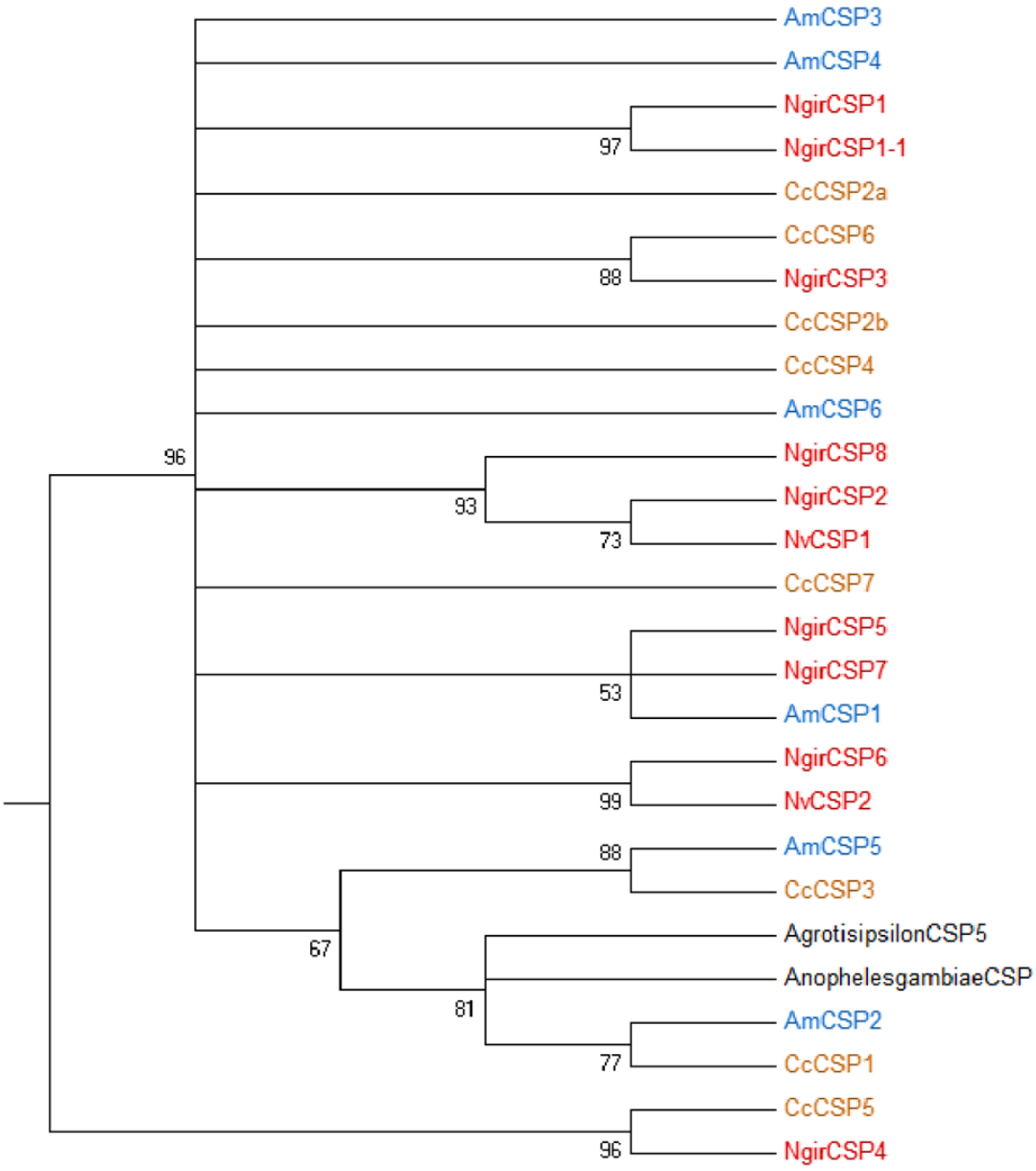
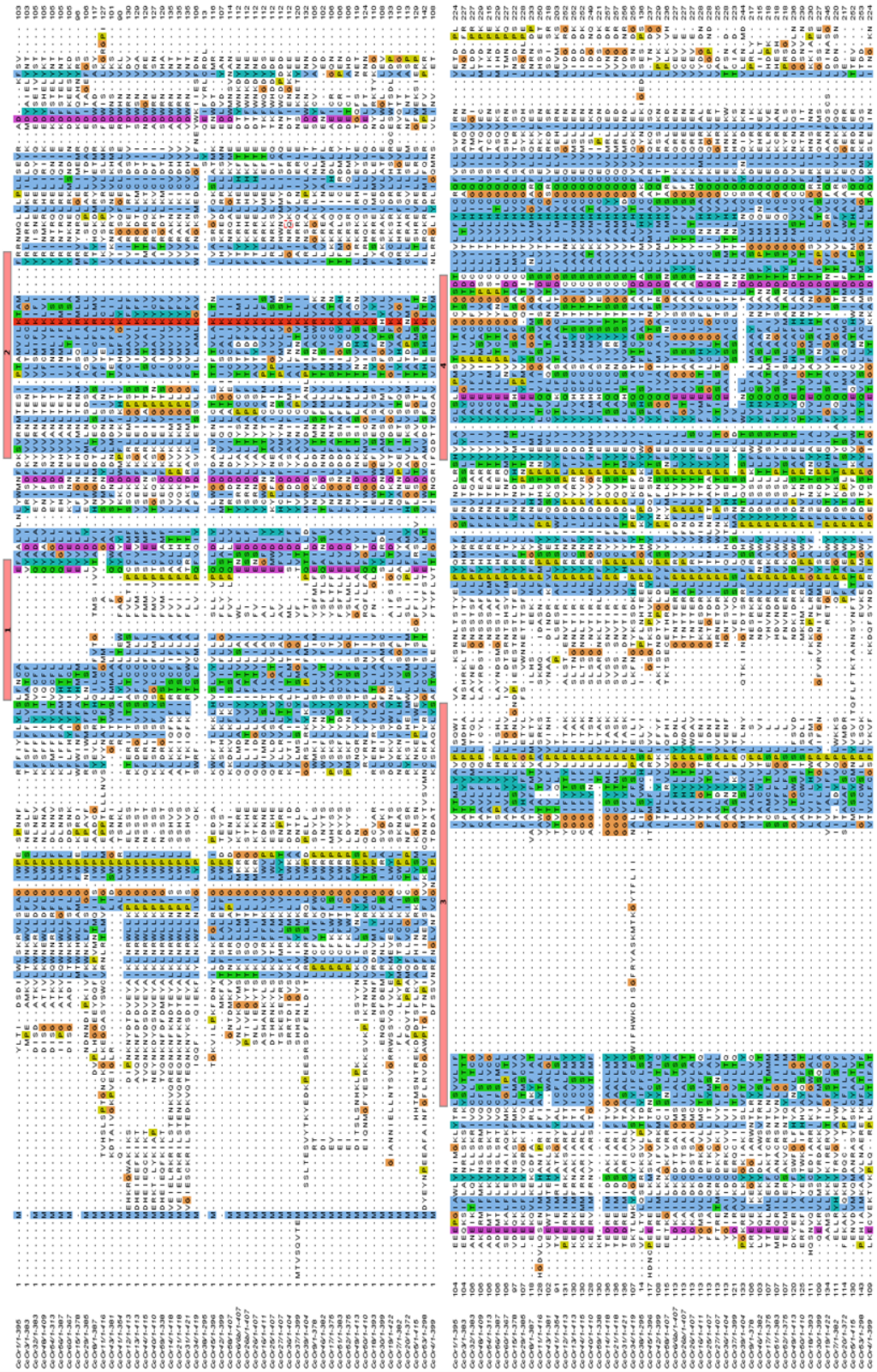


Figure 5.4 Phylogenetic Relationships of 8 WSS CSPs with *N. vitripennis*, *N. giraulti* and *A. mellifera* CSPs inferred from Maximum Likelihood analysis. Bootstrap values are provided at significant branch points and all nodes with less than 50% support were collapsed. Additional CSPs from *A. grotispilon* and *A. gamibae* were included to look at distant relationships among CSPs. Nasonia, red font, honeybee, blue font, WSS, brown font, additional species, black font.



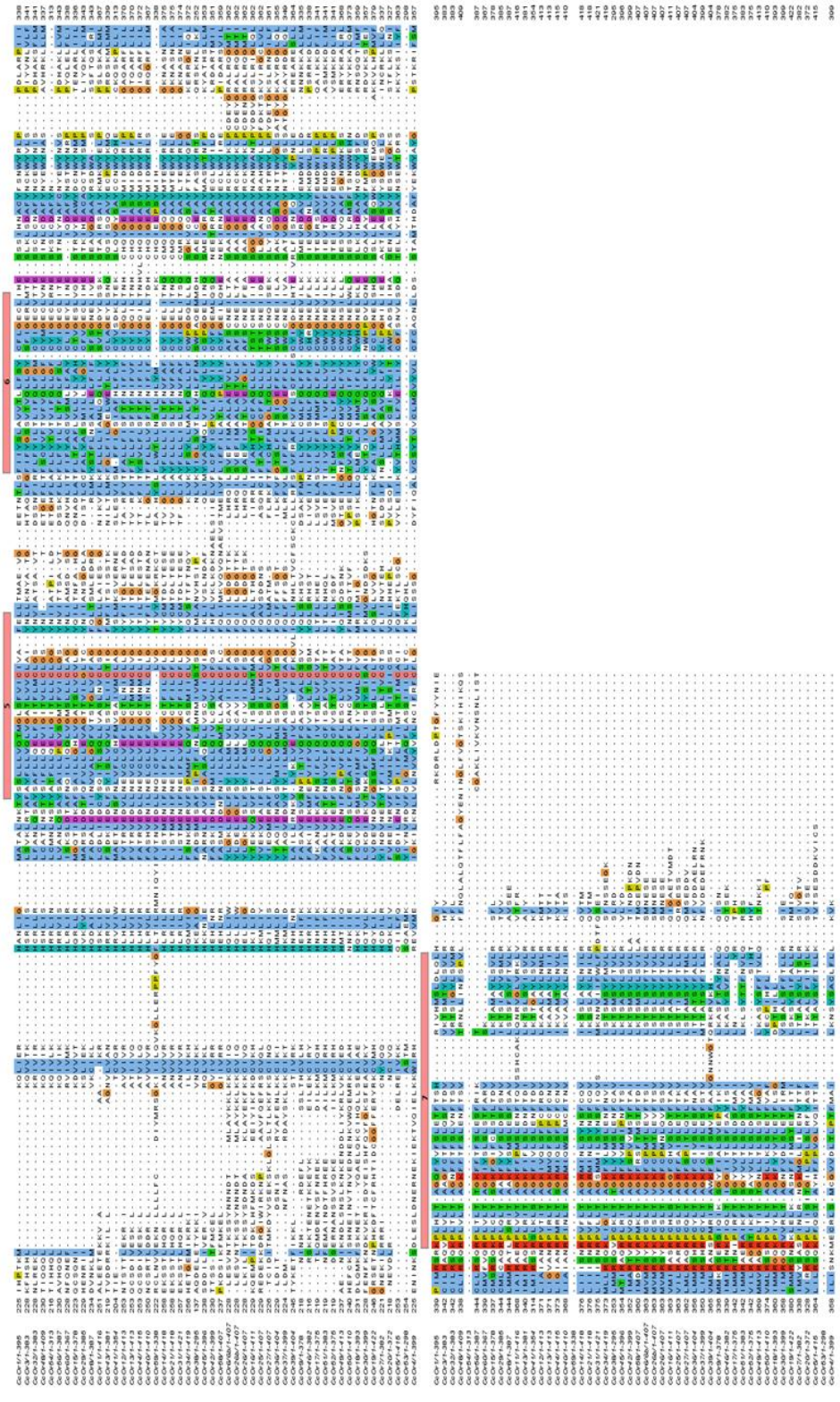
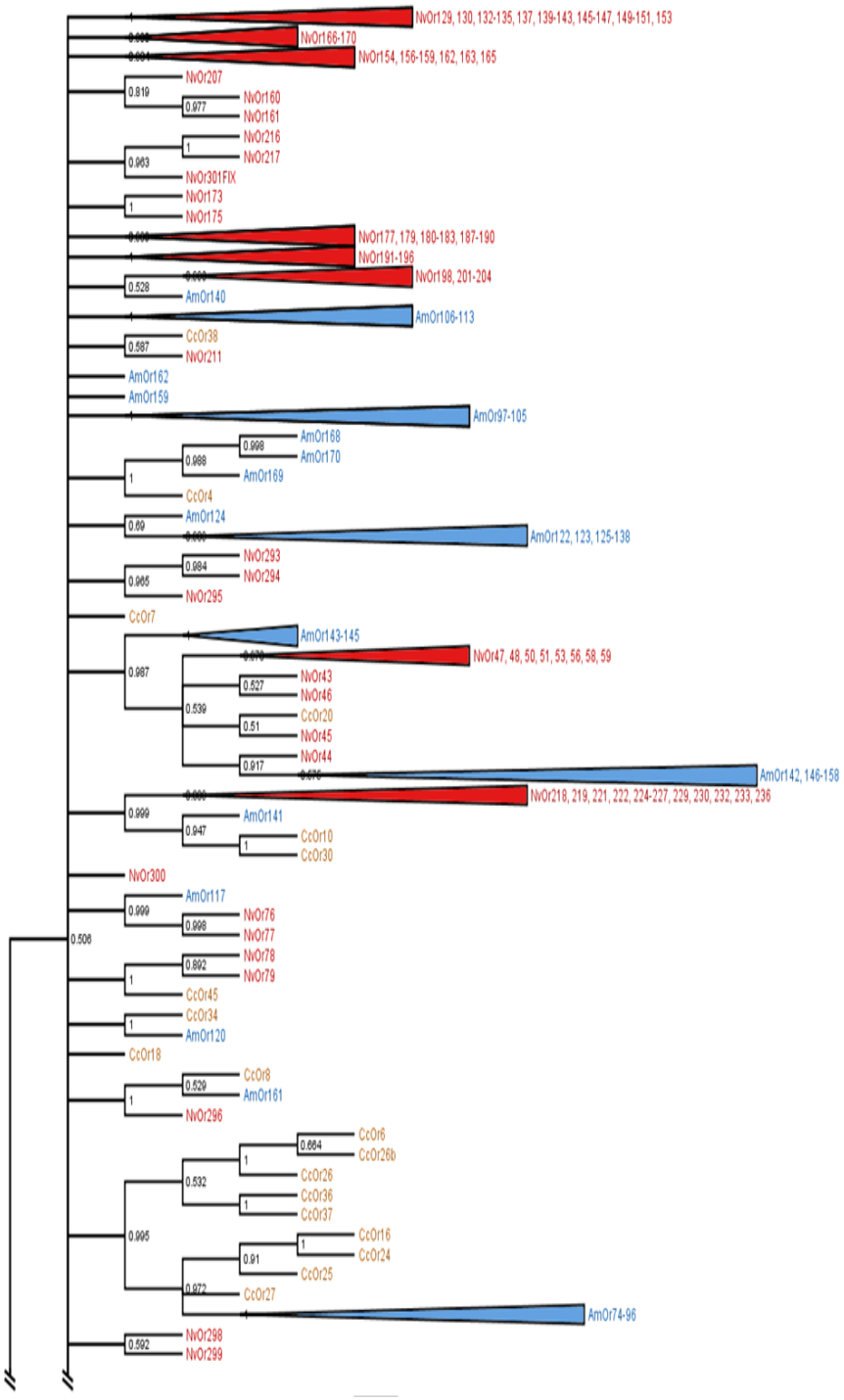
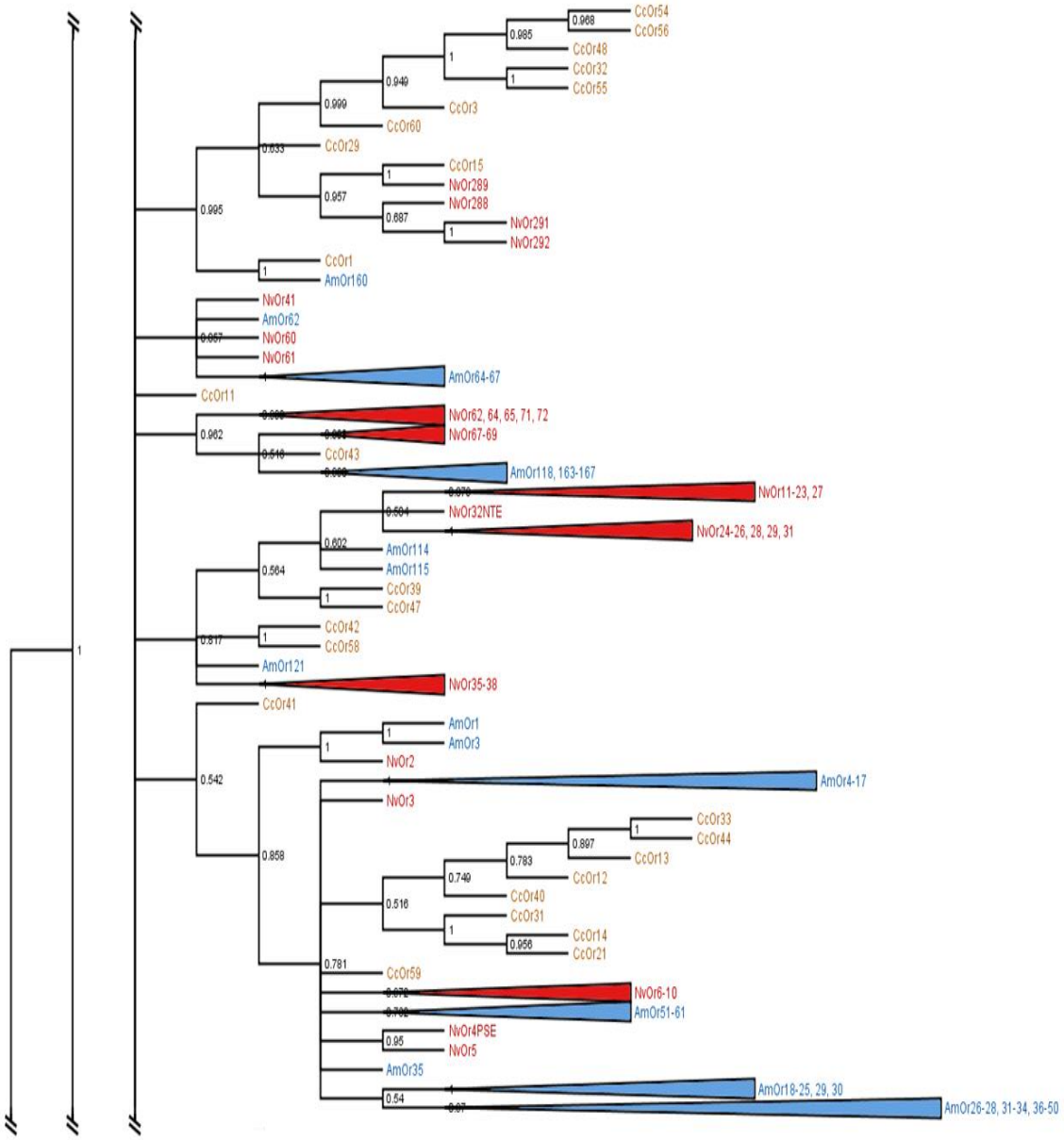


Figure 5.5 Amino acid alignment of WSS ORs annotated from the antennal transcriptome and WGS. A total of 50 WSS ORs were aligned using CLUSTALX set to default parameters and the alignment was used to judge the completeness of the ORs. The 7 transmembrane domains associated with ORs are highlighted as red bars above the alignment and was determined using TMpred.


NvOr129, 130, 132-135, 137, 139-143, 145-147, 149-151, 153
NvOr166-170
NvOr154, 156-159, 162, 163, 165
NvOr207
0.819
NvOr160
0.977
NvOr161
NvOr216
1
0.983
NvOr217
NvOr301FIX
NvOr173
1
NvOr175
NvOr177, 179, 180-183, 187-190
NvOr191-196
NvOr198, 201-204
0.528
AmOr140
AmOr106-113
CcOr38
0.587
NvOr211
AmOr162
AmOr159
AmOr97-105
AmOr168
0.998
0.988
AmOr170
1
AmOr169
CcOr4
AmOr124
0.69
AmOr122, 123, 125-138
NvOr293
0.984
NvOr294
0.965
NvOr295
CcOr7
AmOr143-145
NvOr47, 48, 50, 51, 53, 56, 58, 59
0.987
NvOr43
0.527
NvOr46
0.539
CcOr20
0.51
NvOr45
NvOr44
0.917
AmOr142, 146-158
NvOr218, 219, 221, 222, 224-227, 229, 230, 232, 233, 236
0.999
AmOr141
0.947
CcOr10
1
CcOr30
NvOr300
AmOr117
0.999
NvOr76
0.998
NvOr77
0.506
NvOr78
0.892
NvOr79
1
CcOr45
CcOr34
1
AmOr120
CcOr18
CcOr8
0.529
AmOr161
1
NvOr296
CcOr6
0.664
1
CcOr26b
CcOr26
0.532
CcOr36
1
CcOr37
0.995
CcOr16
1
CcOr24
0.91
CcOr25
0.972
CcOr27
AmOr74-96
NvOr298
0.592
NvOr299


CcOr54
CcOr56
CcOr48
CcOr32
CcOr55
CcOr3
CcOr60
CcOr29
CcOr15
NvOr289
NvOr288
NvOr291
NvOr292
CcOr1
AmOr160
NvOr41
AmOr62
NvOr60
NvOr61
AmOr64-67
CcOr11
NvOr62, 64, 65, 71, 72
NvOr67-69
CcOr43
AmOr118, 163-167
NvOr11-23, 27
NvOr32NTE
NvOr24-26, 28, 29, 31
AmOr114
AmOr115
CcOr39
CcOr47
CcOr42
CcOr58
AmOr121
NvOr35-38
CcOr41
AmOr1
AmOr3
NvOr2
AmOr4-17
NvOr3
CcOr33
CcOr44
CcOr13
CcOr12
CcOr40
CcOr31
CcOr14
CcOr21
CcOr59
NvOr6-10
AmOr51-61
NvOr4PSE
NvOr5
AmOr35
AmOr18-25, 29, 30
AmOr26-28, 31-34, 36-50

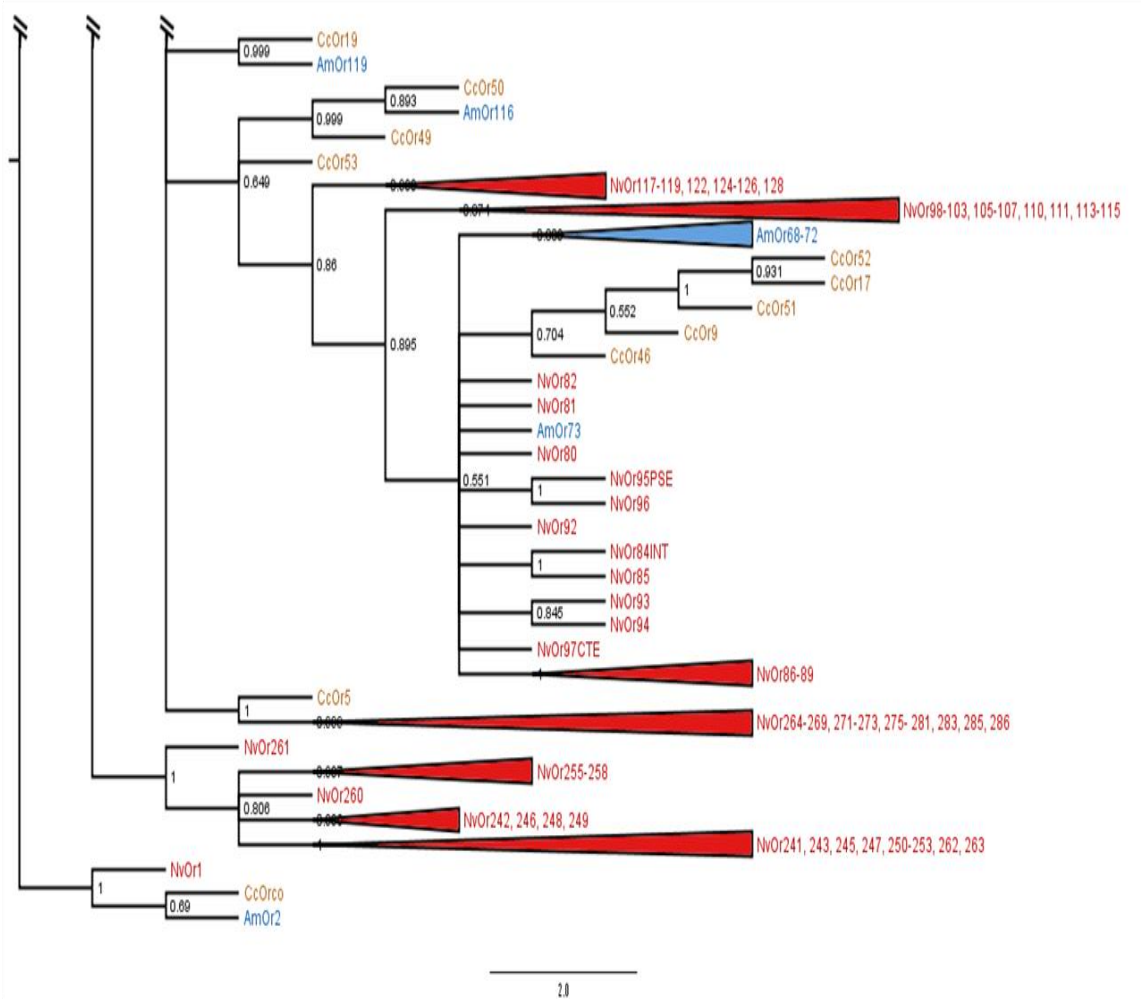
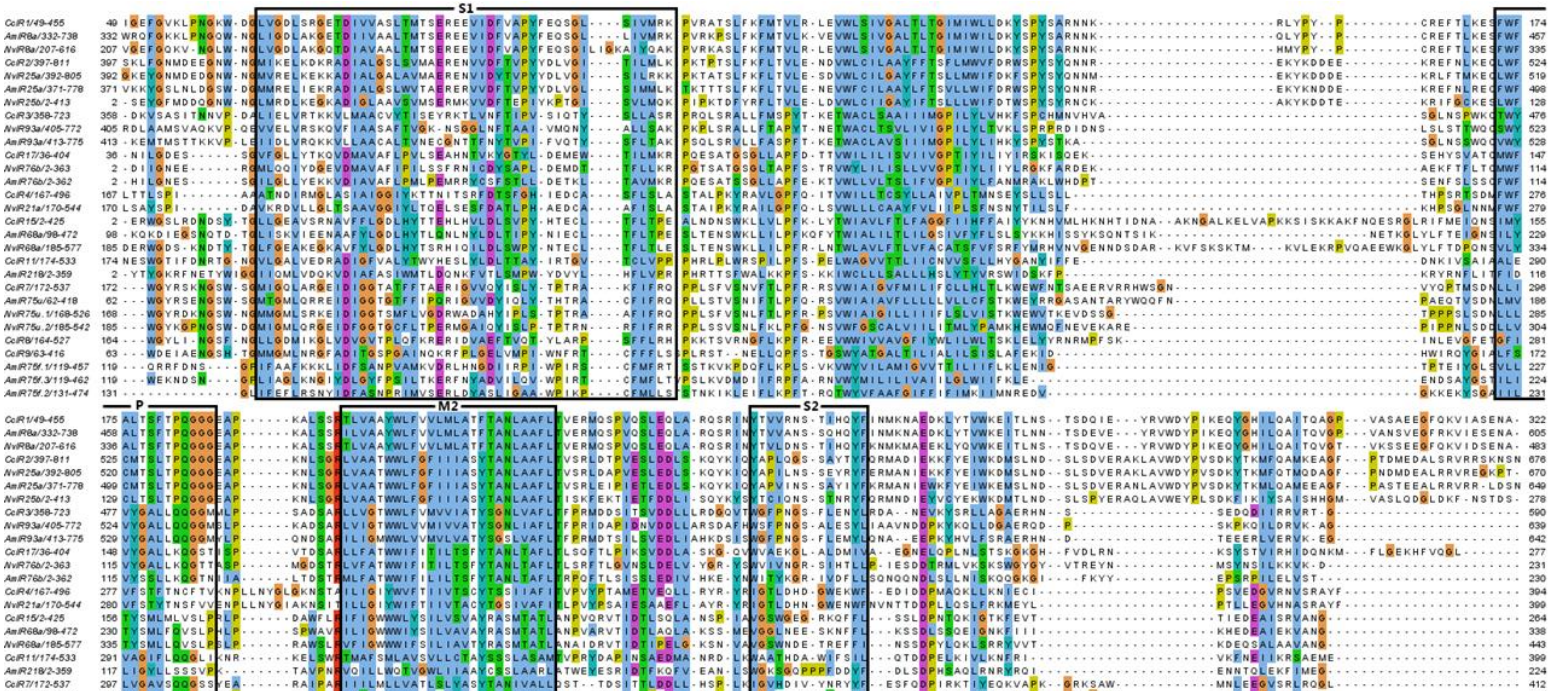


Figure 5.6 Phylogenetic Relationships of 51 WSS Ors with all functional jewel wasp and honeybee Ors inferred from a Maximum Likelihood analysis. Bootstrap values are provided at significant branch points and nodes with less than 50% support are collapsed. The tree was rooted using WSS, jewel wasp and honeybee ORco sequences as the outgroup, based on their bases position in the OR family (Robertson et al. 2003). Jewel wasp, red font, honeybee, blue font, WSS, brown font.


S1
P
M2
S2

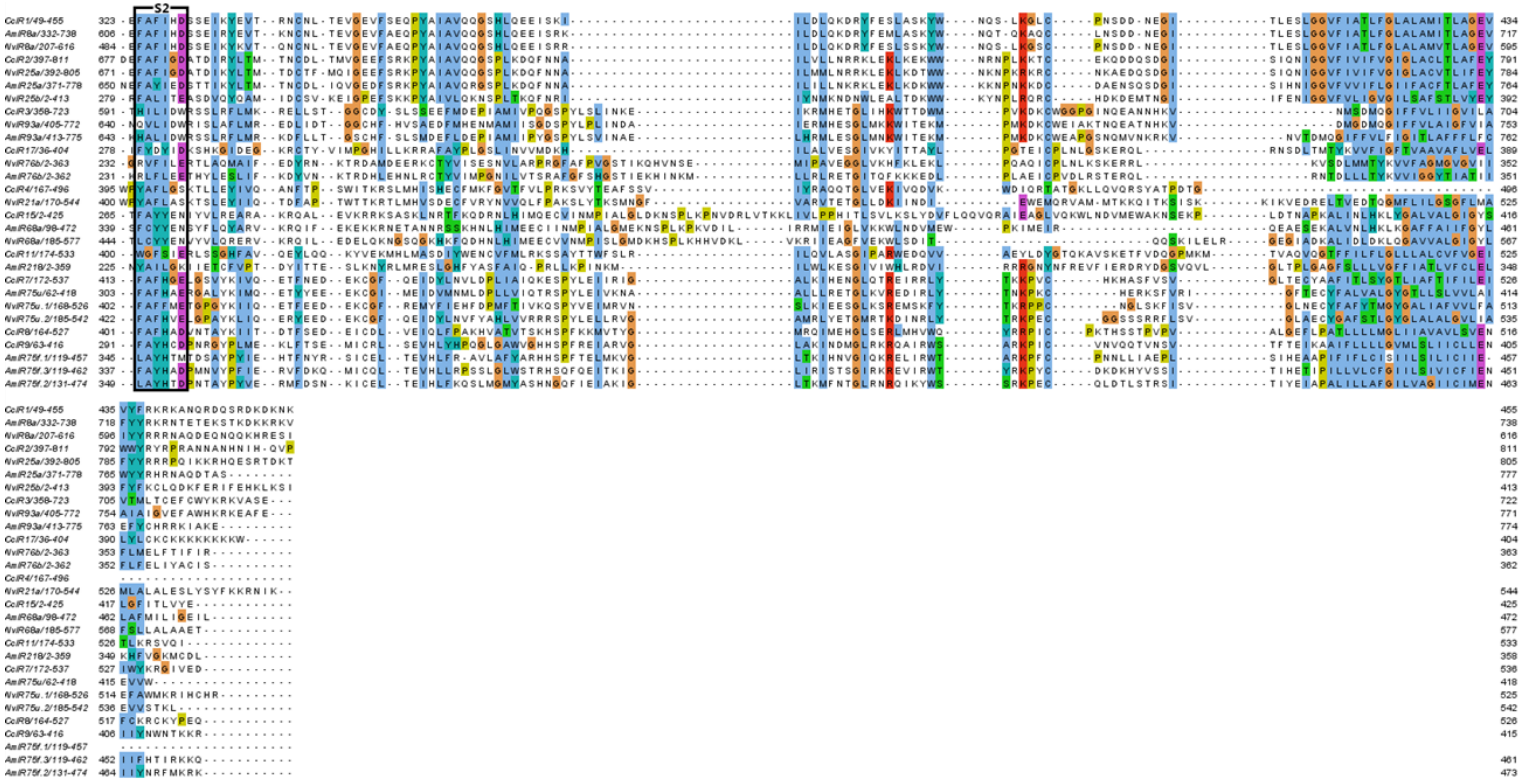


Figure 5.7 Amino acid sequence alignment of WSS IRs annotated from the antennal transcriptome and WGS. A total of 10 IRS were aligned with published sequences of 9 jewel wasp and 10 honeybee IRs using CLUSTALX set to default parameters and the alignment was used to judge the completeness of the IRs. Only the conserved domain section of the IR alignment is shown with the ligand binding domain labeled by a black box marked S1 and S2, the pore region marked P and the conserved ion channel marked M2.

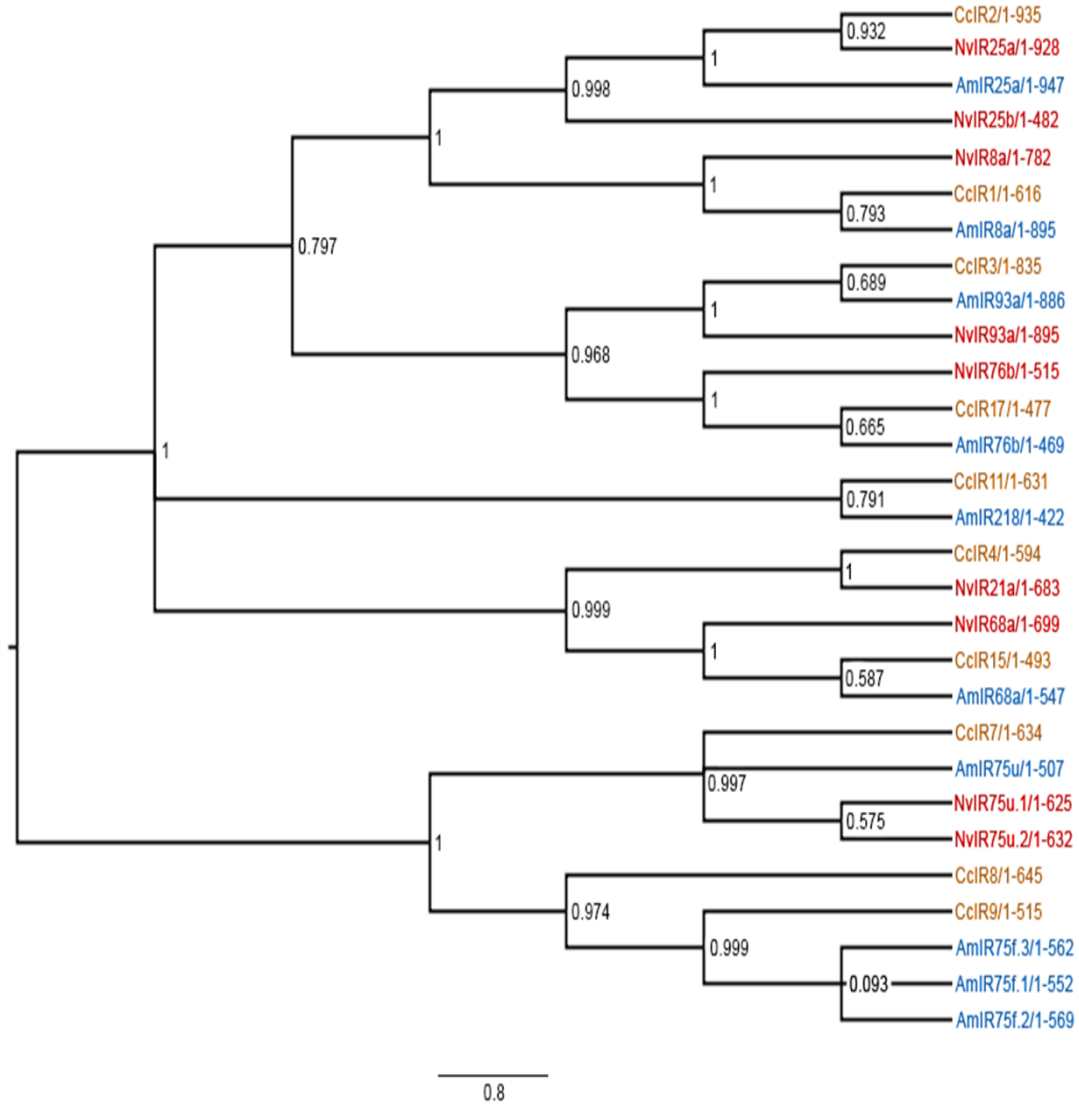
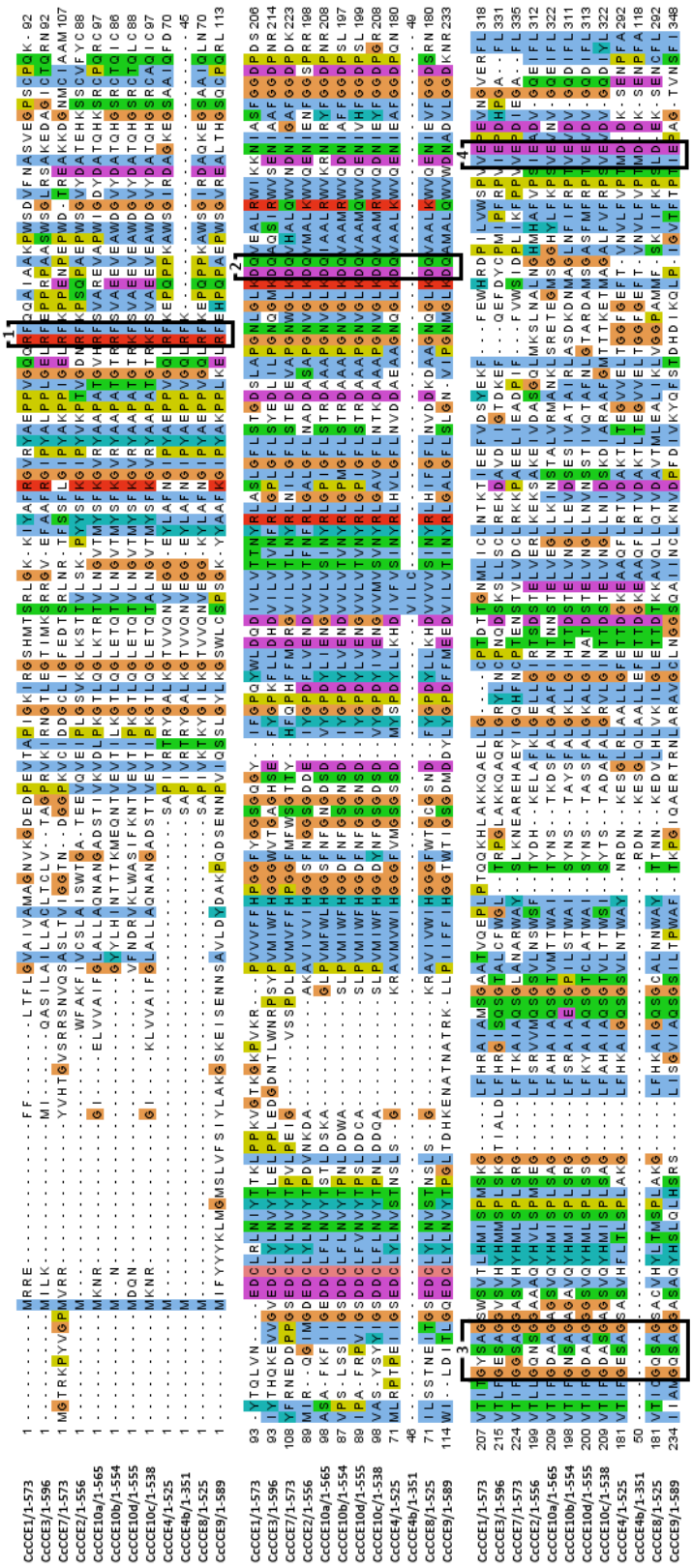


Figure 5.8 Phylogenetic relationships of 10 WSS IRs with jewel wasp and honeybee IRs inferred by Maximum Likelihood analysis. Bootstrap values are provided at significant branch points and nodes with less than 50% support were collapsed. Jewel wasp, red font, honeybee, blue font, WSS, brown font.



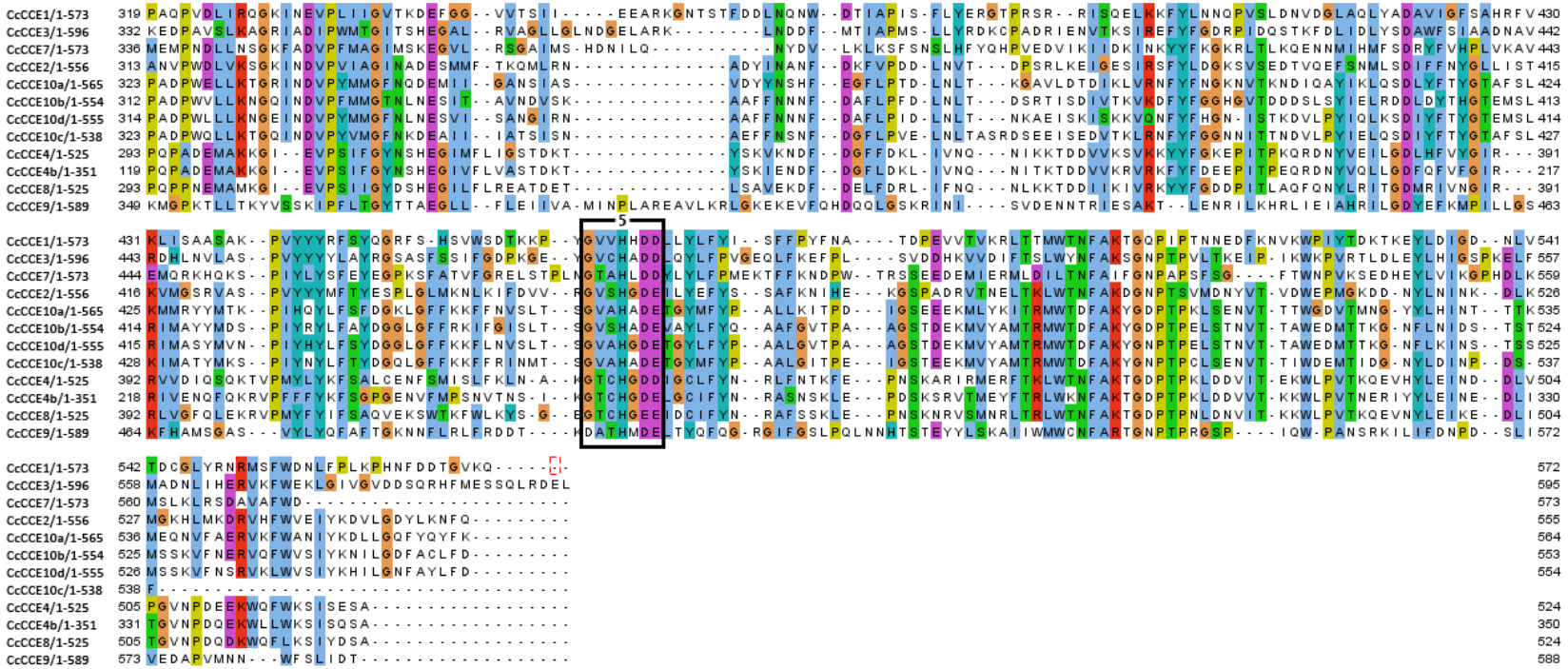


Figure 5.9 Amino acid sequence alignment of WSS CCEs annotated from the antennal transcriptome and WGS. A total of 12 CCEs were aligned using CLUSTALX set to default parameters and the alignment was used to judge the completeness of the WSS CCEs. The 5 conserved domains characteristic of the CCEs are highlighted in black boxes. 1-RF, 2-DQ, 3-GQSAG, 4-VE, 5-GAGHADD.

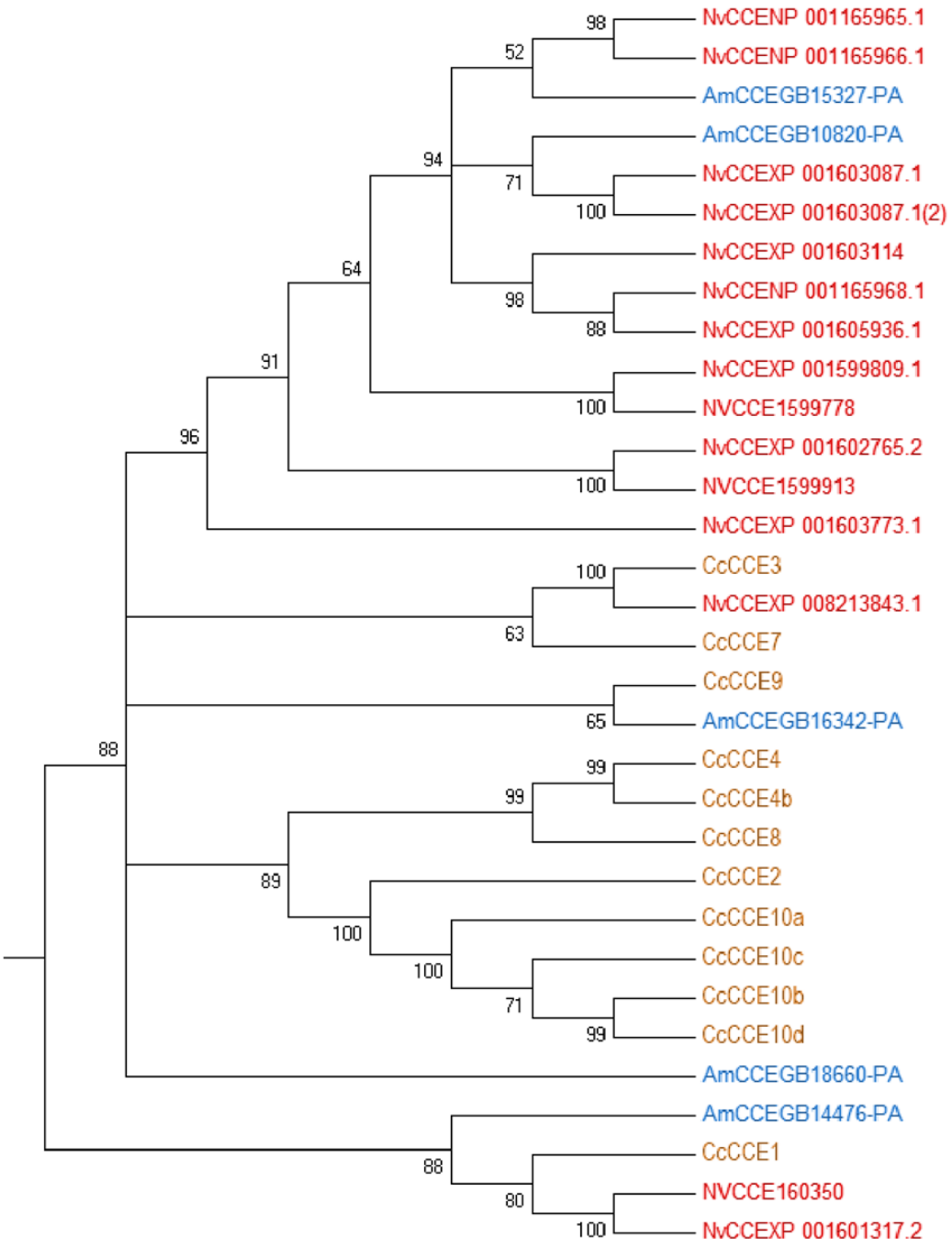


Figure 5.10 Phylogenetic relationships of 12 WSS CCEs with jewel wasp and honeybee CCEs inferred by Maximum Likelihood analysis. Bootstrap values are provided at significant branch points and nodes with less than 50% support are collapsed. Jewel wasp, red font, honeybee, blue font, WSS, brown font.

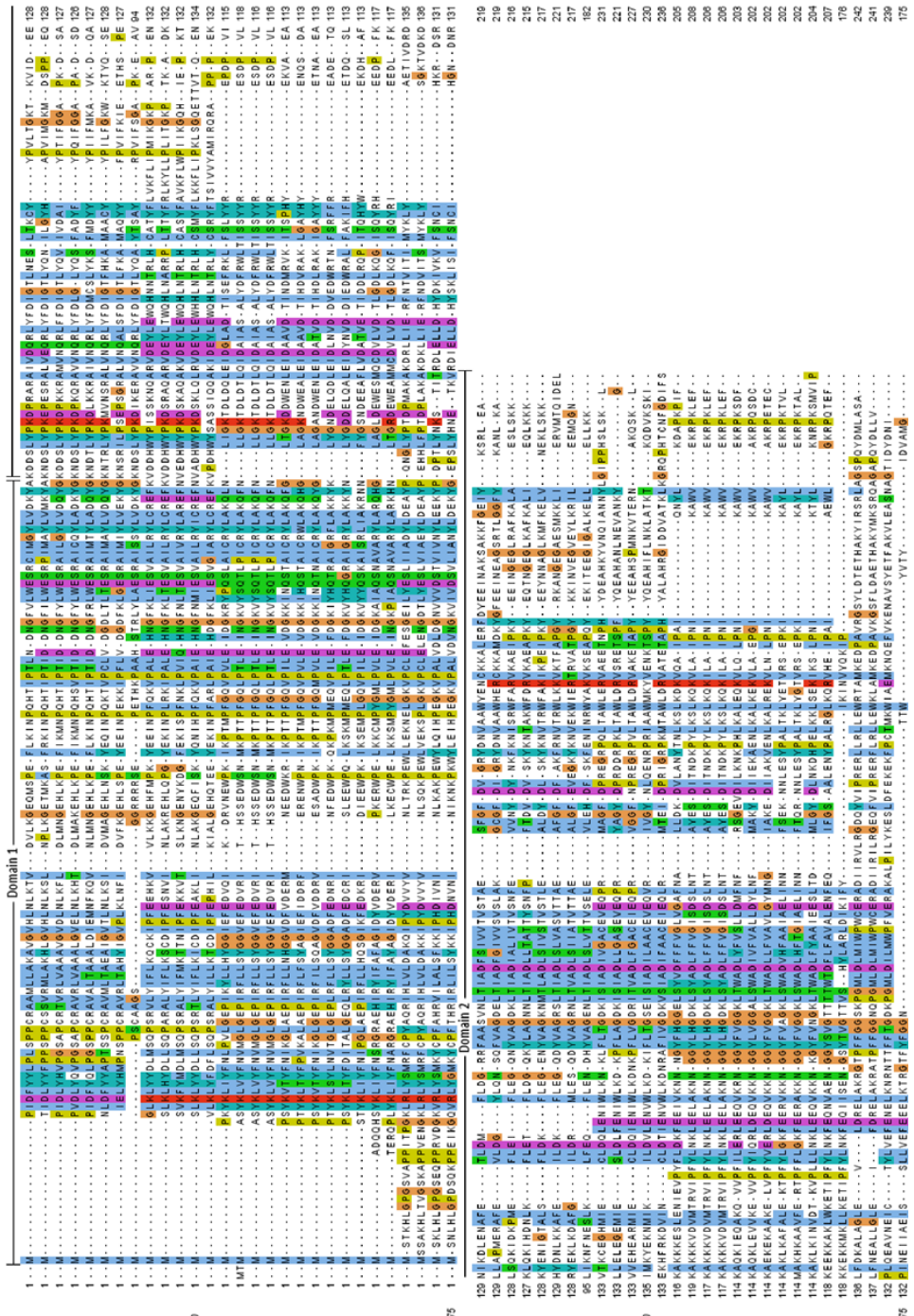


Figure 5.11 Amino acid sequence alignment of WSS GSTs annotated from the antennal transcriptome and WGS. A total of 8 WSS GSTs were aligned with the published sequences of 14 jewel wasp and 11 honeybee GSTs using CLUSTALX set to default parameters, and the alignment was used to judge the completeness of the WSS GSTs. The N terminal and C-terminal domains are labeled.

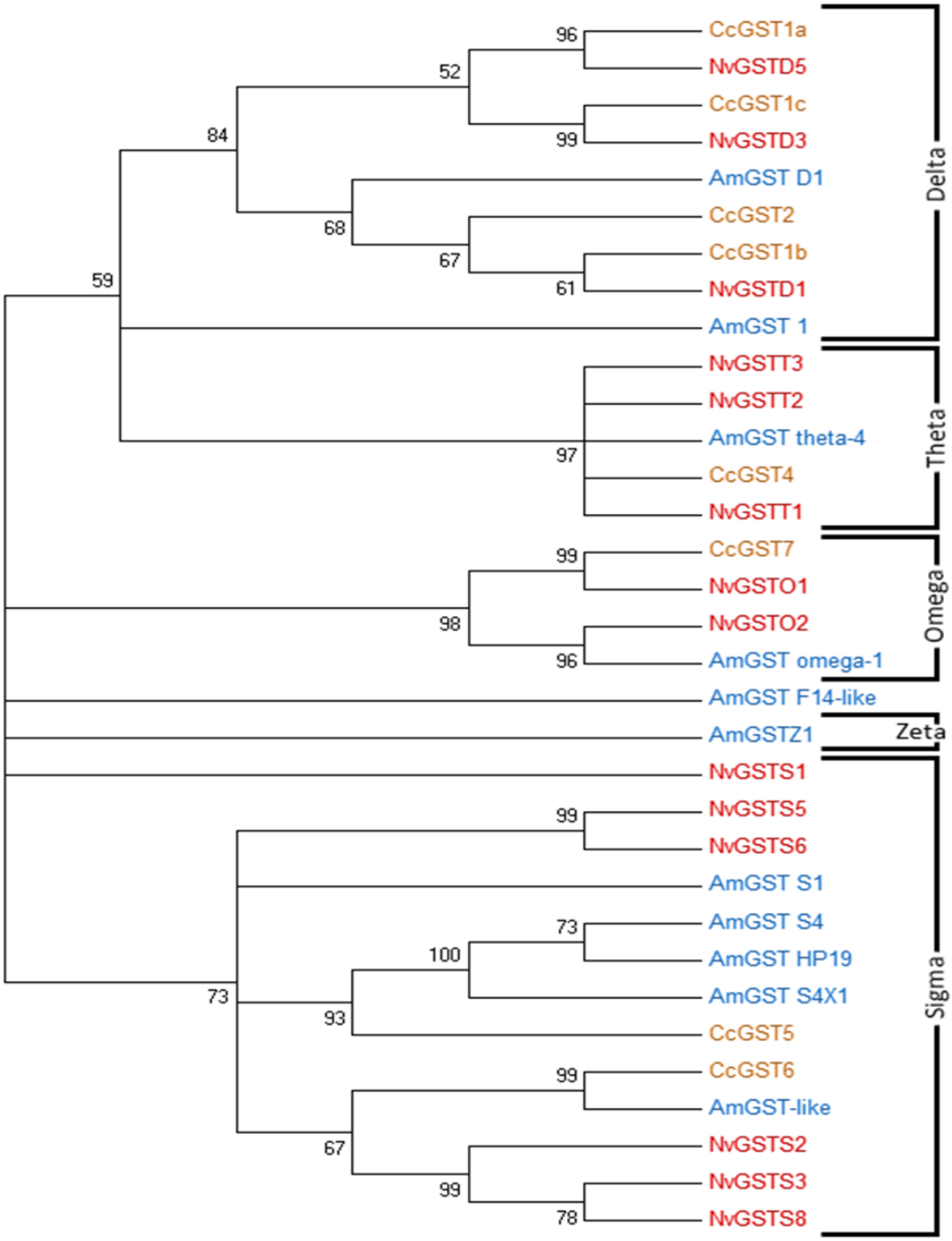
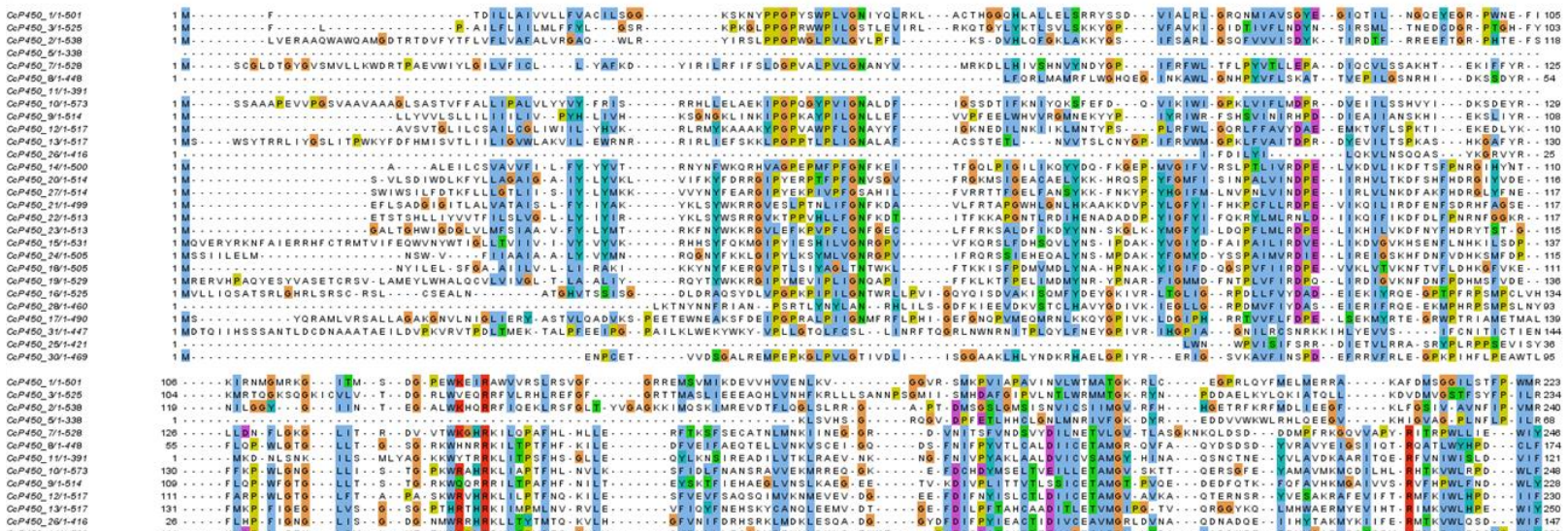


Figure 5.12 Phylogenetic relationships of 8 WSS GSTs with jewel wasp and honeybee GSTs inferred by Maximum Likelihood analysis. Bootstrap values are provided are significant branch points and nodes with support below 50% are collapsed. The main phylogenetic classes of GSTs are identified on the right in black. Jewel wasp, red font, honeybee, blue font, WSS, brown font.



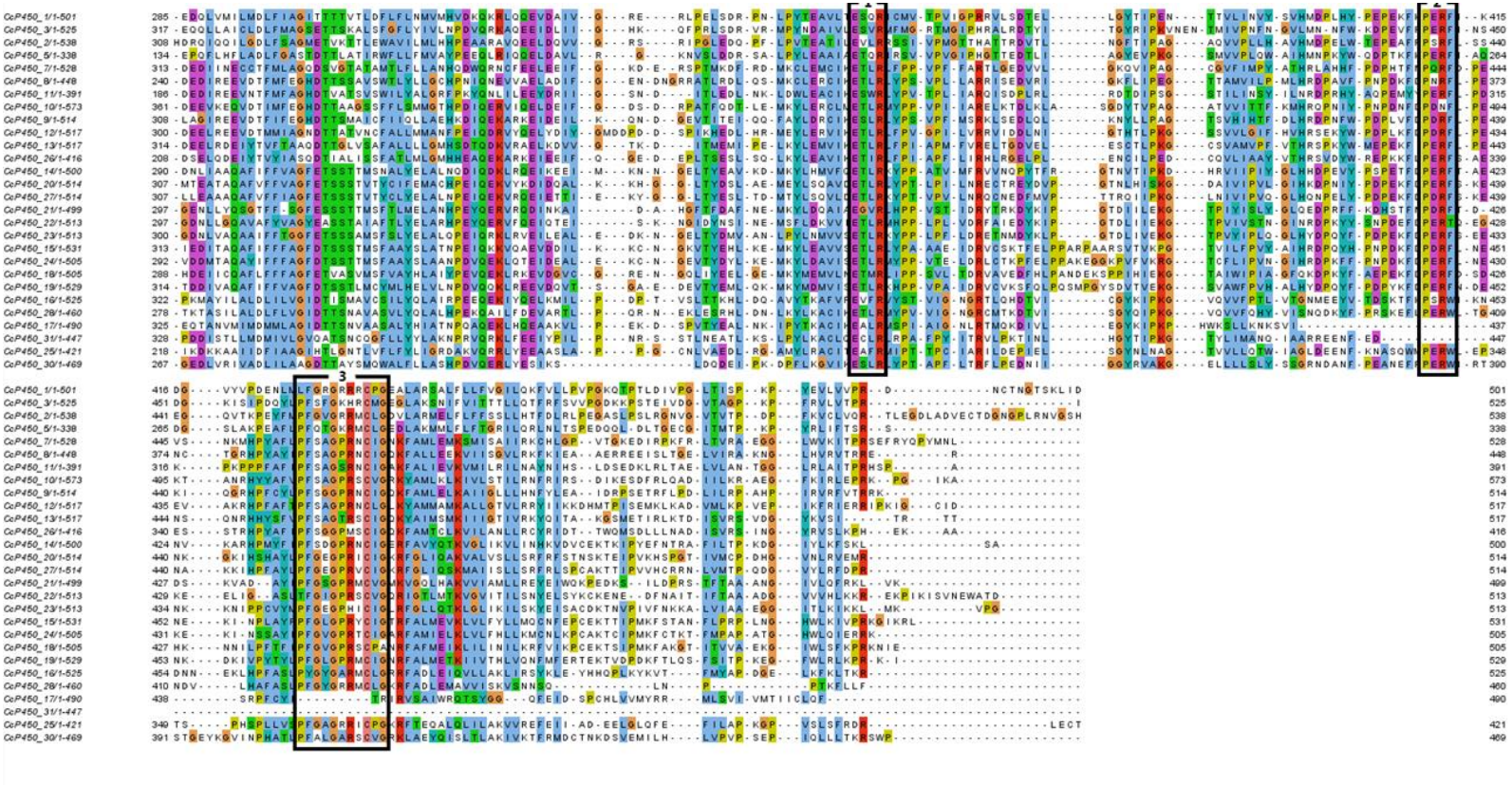
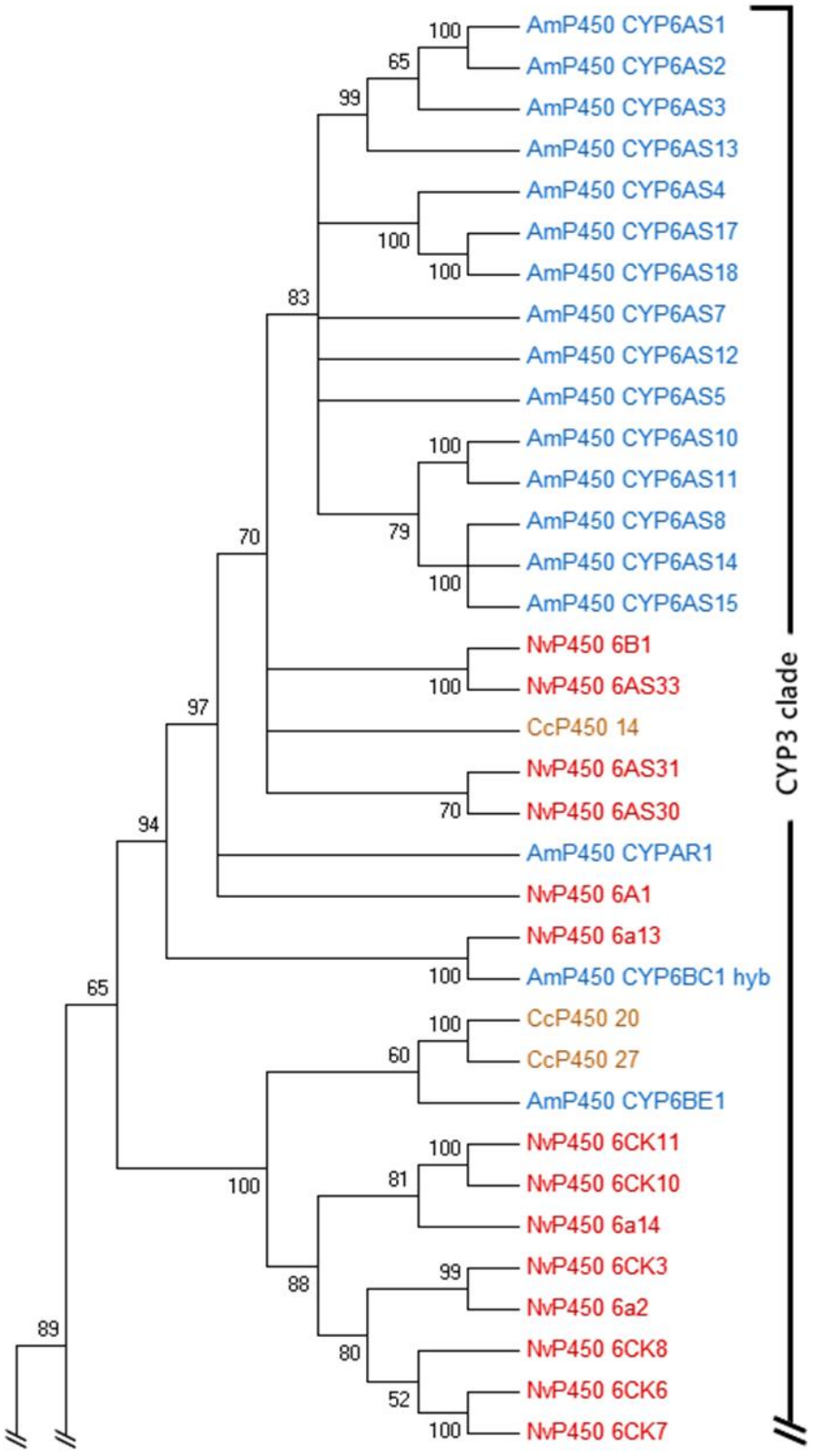
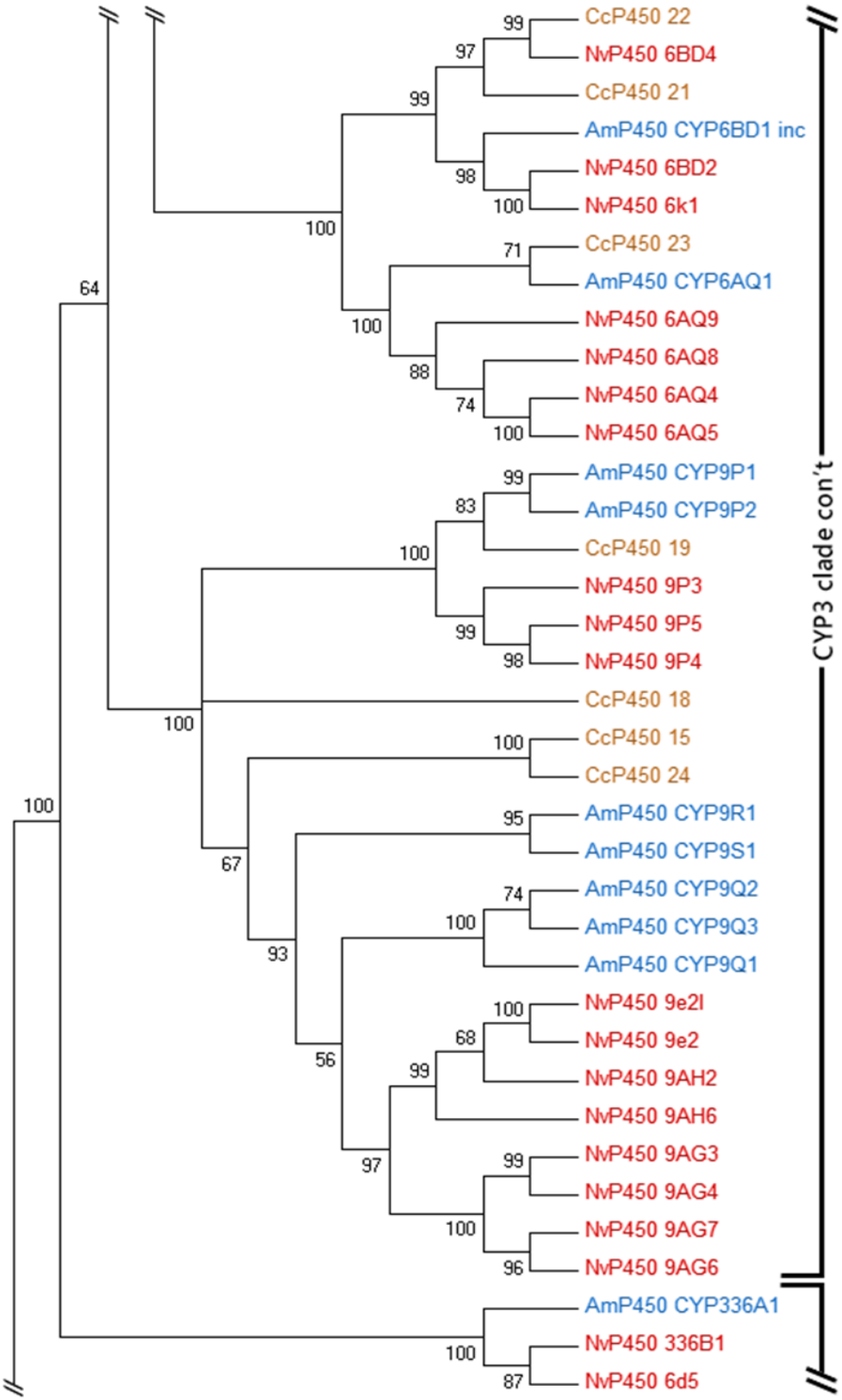
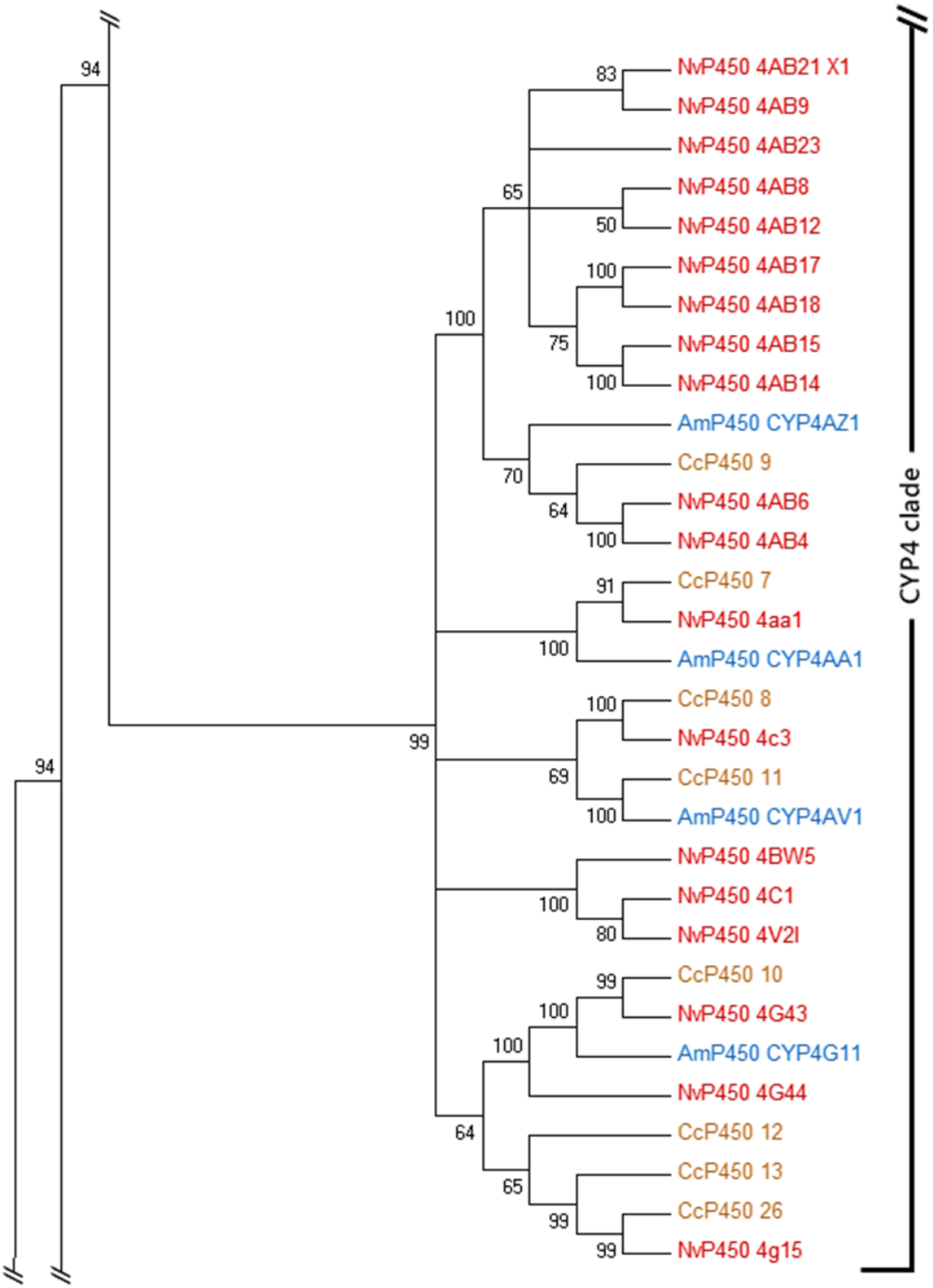


Figure 5.13 Amino acid sequence alignment of WSS P450s annotated from the antennal transcriptome and WGS. A total of 28 WSS P450s were aligned using CLUSTALX set to default parameters and the alignment was used to judge the completeness of the P450s. The 3 conserved motifs are highlighted with numbered boxes, 1-ETRL, 2-PERF, 3-haem binding domain.


AmP450 CYP6AS1
AmP450 CYP6AS2
AmP450 CYP6AS3
AmP450 CYP6AS13
AmP450 CYP6AS4
AmP450 CYP6AS17
AmP450 CYP6AS18
AmP450 CYP6AS7
AmP450 CYP6AS12
AmP450 CYP6AS5
AmP450 CYP6AS10
AmP450 CYP6AS11
AmP450 CYP6AS8
AmP450 CYP6AS14
AmP450 CYP6AS15
NvP450 6B1
NvP450 6AS33
CcP450 14
NvP450 6AS31
NvP450 6AS30
AmP450 CYPAR1
NvP450 6A1
NvP450 6a13
AmP450 CYP6BC1 hyb
CcP450 20
CcP450 27
AmP450 CYP6BE1
NvP450 6CK11
NvP450 6CK10
NvP450 6a14
NvP450 6CK3
NvP450 6a2
NvP450 6CK8
NvP450 6CK6
NvP450 6CK7
CYP3 clade


CcP450 22
NvP450 6BD4
CcP450 21
AmP450 CYP6BD1 inc
NvP450 6BD2
NvP450 6k1
CcP450 23
AmP450 CYP6AQ1
NvP450 6AQ9
NvP450 6AQ8
NvP450 6AQ4
NvP450 6AQ5
AmP450 CYP9P1
AmP450 CYP9P2
CcP450 19
NvP450 9P3
NvP450 9P5
NvP450 9P4
CcP450 18
CcP450 15
CcP450 24
AmP450 CYP9R1
AmP450 CYP9S1
AmP450 CYP9Q2
AmP450 CYP9Q3
AmP450 CYP9Q1
NvP450 9e2l
NvP450 9e2
NvP450 9AH2
NvP450 9AH6
NvP450 9AG3
NvP450 9AG4
NvP450 9AG7
NvP450 9AG6
AmP450 CYP336A1
NvP450 336B1
NvP450 6d5
CYP3 clade con't


NvP450 4AB21 X1
NvP450 4AB9
NvP450 4AB23
NvP450 4AB8
NvP450 4AB12
NvP450 4AB17
NvP450 4AB18
NvP450 4AB15
NvP450 4AB14
AmP450 CYP4AZ1
CcP450 9
NvP450 4AB6
NvP450 4AB4
CcP450 7
NvP450 4aa1
AmP450 CYP4AA1
CcP450 8
NvP450 4c3
CcP450 11
AmP450 CYP4AV1
NvP450 4BW5
NvP450 4C1
NvP450 4V2I
CcP450 10
NvP450 4G43
AmP450 CYP4G11
NvP450 4G44
CcP450 12
CcP450 13
CcP450 26
NvP450 4g15
CYP4 clade

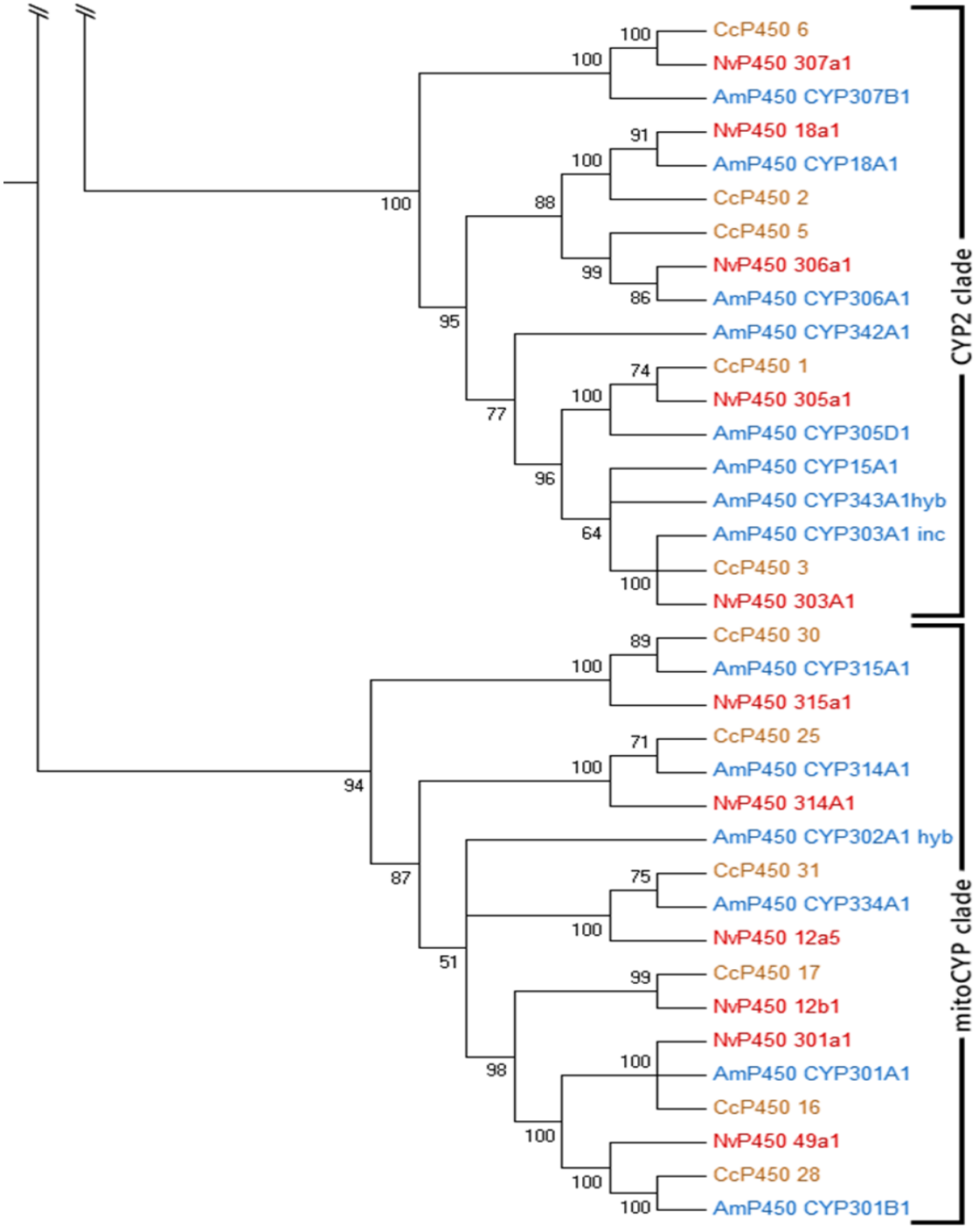


Figure 5.14 Phylogenetic relationships of 28 WSS P450s with jewel wasp and honeybee P450s inferred by Maximum Likelihood analysis. Bootstrap values are provided at significant branch points and nodes with less than 50% support are collapsed. The main phylogenetic classes of P450s are identified on the right of the tree in black, Jewel wasp, red font, honeybee, blue font, WSS, brown font.

References

Abeel, T., Van Parys, T., Saeys, Y., Galagan, J., and Van de Peer, Y. (2012) GenomeView: a next-generation genome browser. *Nucleic Acids Res* **40**(2): e12. doi: 10.1093/nar/gkr995.

Ainslie, C. N. (1920) The western grass-stem sawfly. United States Department of Agriculture. *Bull No* 841.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W., and Lipman, D. J. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* **25**: 3389-3402.

Anderson, A. R., Wanner, K. W., Trowell, S. C., Warr, C. G., Jaquin-Joly, E., Zagatti, P., Robertson, H., and Newcomb, R. D. (2009) Molecular basis of female-specific odorant responses in *Bombyx mori*. *Insect Biochem Mol Biol* **39**: 189–197.

Angeli, S., Ceron, F., Scaloni, A., Monti, M., Monteforti, G., Minnocci, A., Petacchi, R., and Pelosi, P. (1999) Purification, structural characterization, cloning and immunocytochemical localization of chemoreception proteins from *Schistocerca gregaria*. *Eur J Biochem* **262**: 745-754.

Benton, R., Sachse, S., Michnick, S. W., and Vosshall, L. B. (2006) Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo. *PLoS Biol* **4**: e20.

Benton, R., Vannice, K. S., Gomez-Diaz, C., and Vosshall, L. B. (2009) Variant ionotropic glutamate receptors as chemosensory receptors in Drosophila. *Cell* **136**(1): 149-162.

Beres, B. L., Cárcamo, H. A., and Byers, J. R. (2007) Effect of wheat stem sawfly damage on yield and quality of selected Canadian spring wheat. *J Econ Entomol* **100**: 79-87.

Board, P. G., Baker, R. T., Chelvanayagam, G. and Jermiin, L. S. (1997) Zeta, a novel class of glutathione transferases in a range of species from plants to humans. *Biochem J* **328**: 929-935.

Bohbot, J., Sobrio, F., Lucas, P., and Nagnan-Le Meillour, P. (1998) Functional characterization of a new class of Odarant Binding Proteins in the moth *Mamestra bassicae. Biochem Biophys Res Commun **253**:* 489-494.

Bradshaw, J. (2013) Adult Wheat Stem Sawflies Have Begun to Emerge. University of Nebraska Cropwatch. 30 May 2013.

Brandt, A., Scharf, M., Pedra, J. H., Holmes, G., Dean, A., Kreitman, M., and Pittendrigh, B. R. (2002) Differential expression and induction of two *Drosophila* cytochrome P450 genes near the Rst(2)DDT locus. *Insect Mol Biol* **11**: 337–341.

Briand, L., Swasdipan, N., Nespoulous, C., Bezirard, V., Blon, F., Huet, J. C., Ebert, P., and Penollet, J. C. (2002) Characterization of a chemosensory protein (ASP3c) from honeybee (*Apis mellifera L*.) as a brood pheromone carrier. *Eur J Biochem* **269**: 4586–4596.

Bogwitz, M. R., Chung, H., Magoc, L., Rigby, S., Wong, W., O'Keefe, M., et al. (2005) Cyp12a4 confers lufenuron resistance in a natural population of *Drosophila melanogaster*. *Proc Natl Acad Sci USA.* **102**(36): 12807-12812.

Calvello, M., Brandazza, A., Navarrini, A., Dani, F. R., Turillazzi, S., et al. (2005) Expression of odorant-binding proteins and chemosensory proteins in some Hymenoptera. *Insect Biochem Mol Biol* **35**: 297–307.

Calvello, M., Guerra, N., Brandazza, A., D'Ambrosio, C., Scaloni, A., Dani, F.R., Turillazzi, S., and Pelosi, P. (2003) Soluble proteins of chemical communication in the social wasp *Polistes dominulus*. *Cell Mol Life Sci* **60**: 1933–1943.

Campanacci, V., Lartigue, A., Hallberg, B. M., Jones, T. A., Giudici-Orticoni, M. T., Tegoni, M., and Cambillau, C. (2003) Moth chemosensory protein exhibits drastic conformational changes and cooperativity on ligand binding. *Proc Natl Acad Sci USA* **100**: 5069–5074.

Canterel, B. L., Korf, I., Robb, S. M., Parra, G., Ross, E., Moore, B., Holt, C., Sanchez Alvarado, A., and Yandall, M. (2008) MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res* **18**(1): 188-196.

Carey, A. F., Wang, G., Su, C-Y., Zwiebel, L. J., and Carlson, J. R. (2010) Odorant reception in the malaria mosquito *Anopheles gambiae*. *Nature* **464**: 66–71.

Clamp, M., Cuff, J., Searle, S. M., and Barton, G. J. (2004) The Jalview Java alignment editor. *Bioinformatics* **20**: 426-427. doi:10.1093/bioinformatics/btg430.

Claudianos, C., Ranson, H., Johnson, R. M., Biswas, S., Schuler, M. A., Berenbaum, M.R., Feyereisen, R., and Oakeshott, J.G. (2006) A deficit of detoxification enzymes: pesticide sensitivity and environmental response in the honeybee. *Insect Mol Biol* **15**: 615–636.

Clyne, P. J., Warr, C. G., Freeman, M. R., Lessing, D., Kim, J., and Carlson, J. R. (1999) A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. *Neuron* **22**: 327–338.

Criddle, N. (1923) The life habits of *Cephus Cinctus* Norton in Monitoba. *Canadian Entomol* **55**: 1-4.

Croset, V., Ritz, R., Cummins, S. F., Budd, A., Brawand, D., Kaessmann, H., Gibson, T. J. and Benton, R. (2010) Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet* **6**: e1001064. doi: 10.1371/journal.pgen.1001064.

Danielson, P. B., Macintyre, R. J., and Fogleman, J. C. (1997) Molecular cloning of a family of xenobiotic-inducible drosophilid cytochrome P450s: evidence for involvement in host-plant allelochemical resistance. *Proc Natl Acad Sci USA* **94**:10797-10802.

Danty, E., Arnold, G., Huet, J. C., Huet, D., Masson, C., and Pernollet, J. C. (1998) Separation, characterization and sexual heterogeneity of multiple putative odorant-binding proteins in the honeybee *Apis mellifera L*. (Hymenoptera: Apidea). *Chem Senses* **23**: 83–91.

Engsontia, P., Sanderson, A. P., Cobb, M., Walden, K. K. O., Robertson, H., and Brown, S. (2008) The red flour beetle's large nose: an expanded odorant receptor gene family in *Tribolium casteneum*. *Insect Biochem Mol Biol* **38**: 387–397.

Fletcher, J. 1904. Experimental farms reports - Report of the Entomologist and Botanist.Appendix to the Report of the Minister of Agriculture, Sessional Paper No. 16. pp. 172–173.

Forêt, S., and Maleszka, R. (2006) Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (Apis mellifera). *Genome Res* **16**: 1404–1413.

Forêt, S., Wanner, K. W., and Maleszka, R. (2007) Chemosensory proteins in the honey bee: insights from the annotated genome, comparative analyses and expressional profiling. *Insect Biochem Mol Biol* **37**: 19–28.

Galindo, K., and Smith, D. P. (2001) A large family of divergent Drosophila odorant-binding proteins expressed in gustatory and olfactory sensilla. *Genetics* **159**: 1059–1072.

Gereau, R.W., and Swanson, G.T. (2008) The Glutamate Receptors. Humana Press, Totowa, N.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., et al. (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* **29**: 644–652.

Gress, J. C., Robertson, H. M., Weaver, D. K., Dlakić, M., and Wanner, K. W. (2013) Odorant receptors of a primitive hymenopteran pest, the wheat stem sawfly. *Insect Mol Biol* **22**(6): 659-667. DOI: 10.1111/imb.12053.

Grimaldi, D., and Engel, M. S. (2005) Evolution of the Insects. Cambridge University Press, Cambridge,UK. xv+755 pp.

Guzov, V. M., Unnithan, G. C., Chernogolov, A. A., and Feyerseisen, R. (1998) CYP12A1, a mitochondrial cytochrome P450 from the house fly. *Arch Biochem Biophys* **359**: 231–240.

Hallem, E. A., Ho, M. G., and Carlson, J. R. (2004) The molecular basis of odor coding in the Drosophila antenna. *Cell* **117**:965–979.

Hekmat-Scafe, D. S., Scafe, C. R., McKinney, A. J., and Tanouye, M. A. (2002) Genome-wide analysis of the odorant-binding protein gene family in Drosophila melanogaster. *Genome Res* **12**: 1357– 1369.

Helvig, C., Koener, J. F., Unnithan, G. C., and Feyereisen, R. (2004) CYP15A1, the cytochrome P450 that catalyzes epoxidation of methyl farnesoate to juvenile hormone III in cockroach corpora allata. *Proc Natl Acad Sci USA* **101**: 4024-4029.

Holmes, N. D. (1977) The effect of the wheat stem sawfly, *Cephus cinctus* (Hymenoptera: Cephidae), on the yield and quality of wheat. *Can Entomol* **109**: 1591-1598.

Holmes, N. D. and Peterson, L. K. (1960) The influence of the host on oviposition by the wheat stem sawfly, *Cephus cinctus* Nort. (Hymenoptera: Cephidae) II. Resistance to the larva. *Can Entomol* **94**: 348-365.

Holt, C., and Yandell, M. (2011) MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinf* **12**:491. doi:10.1186/1471-2105-12-491.

Irell, B. and Peairs, F. (2011) Wheat Stem Sawfly: A New Pest of Colorado Wheat. Colorado State University Extension. Circular No. 5.612.

Ishida, Y., Chiang, V., and Leal, W. S. (2002) Protein that makes sense in the Argentine ant. *Naturwissenschaften* **89**:505–507.

Jacquin-Joly, E., Vogt, R. G., Francois, M. C., and Nagnan-Le Meillour, P. (2001) Functional and expression pattern analysis of chemosensory proteins expressed in antennae and pheromonal gland of *Mamestra brassicae. Chem Senses* **26**: 833-844.

Jones, W. D., Nguyen T-A ,T., Kloss, B., Lee, K. J., and Vosshall, L. B. (2005) Functional conservation of an insect odorant receptor gene across 250 million years of evolution. *Curr Bio* **15**: R119-R121.

Kaissling, K. E. (1998) Flux detectors versus concentration detectors: Two types of chemoreceptors. *Chem Senses* **23**: 99-111.

Kamikouchi, A., Morioka, M., and Kubo, T. (2004) Identification of honeybee antennal proteins/genes expressed in a sex- and/or caste selective manner. *Zool Sci* **21**(1): 53–62.

Katoh, K., Kuma, K., Toh, H., and Miyata, T. (2005) MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res* **33**(2): 511-518.

Krieger, J., Grosse-Wilde, E., Gohl, T., Dewer, Y., Raming, K., and Breer, H. (2004) Genes encoding candidate pheromone receptors in a moth (*Heliothis virescens*). *Proc Natl Acad Sci USA* **101**: 11845-11850.

Lagarde, A., Spinelli, S., Qiao, H., Tegoni, M., Pelosi, P., et al. (2011) Crystal structure of a novel type of odorant binding protein from Anopheles gambiae, belonging to the C+ class. *Biochem J* **437**: 423–430.

Lagarde, A., Spinelli, S., Tegoni, M., He, X., Zhou, J. J., et al. (2011b) The crystal structure of odorant binding protein 7 from Anopheles gambiae exhibits an outstanding adaptability of its binding site. *J Mol Biol* **414**: 401–12.

Larsson, M. C., Domingos, A. I., Jones, W. D., Chiappe, M. E., Amrein, H., and Vosshall, L. B. (2004) Or83b encodes a broadly expressed odorant receptor essential for Drosophila olfaction. *Neuron* **43**: 703–714.

Lartigue, A., Campanacci, V., Roussel, A., Larsson, A. M., Jones, T. A., Tegoni, M., and Cambillau, C. (2002) X-ray structure and ligand binding study of a moth chemosensory protein. *J Biol Chem* **277**: 32,094–32,098.

Leal, W. S. (2013) Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu Rev Entomol* **58**: 373-391.

Leal, W. S., Barbosa, R. M. R., Xu, W., Ishida, Y., Syed, Z., et al. (2008) Reverse and Conventional Chemical Ecology Approaches for the Development of Oviposition Attractants for Culex Mosquitoes. *PLoS ONE* **3**(8): e3045. doi:10.1371/journal.pone.0003045

Leal, W. S., Chen, A. M., Ishida, Y., Chiang, V. P., Erickson, M. L., Morgan, T. I., and Tsuruda, J. M. (2005) Kinetics and molecular properties of pheromone binding and release. *Proc Natl Acad Sci USA* **102**: 5386–5391.

Leal, W. S., Nikonova, L., and Peng, G. (1999) Disulfide structure of the pheromone binding protein from the silkworm moth, *Bombyx mori*. *FEBS letters* **464**:85–90.

Li, X., Schuler, M. A., and Berenbaum, M. R. (2007) Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. *Annu Rev Entomol* **52**: 231-253.

Maibeche-Coisne, M., Merlin, C., Francois, M. C., Porcheron, P., and Jacquin-Joly, E. (2005) P450 and P450 reductase cDNAs from the moth *Mamestra brassicae*: cloning and expression patterns in male antennae. *Gene* **346**: 195-203.

Maibeche-Coisne, M., Monti-Dedieu, L., Aragon, S., and Dauphin-Villemant, C. (2000) A new cytochrome P450 from *Drosophila melanogaster*, CYP4G15, expressed in the nervous system. *Biochem Biophys Res Commun* **273**: 1132-1137.

Mayer, M. L. (2006) Glutamate receptors at atomic resolution. *Nature* **440**: 456-462.

Mitchell, R. F., Hughes, D. T., Luetje, C. W., et al. (2012) Sequencing and characterizing odorant receptors of the cerambycid beetle *Megacyllene caryae*. *Insect Biochem Mol Biol* **42**(7): 499-505.

Mittapalli, O., Neal, J. J., and Shukle, R. H. (2005) Differential expression of two cytochrome P450 genes in compatible and incompatible Hessian fly/wheat interactions. *Insect Biochem Mol Biol* **35**: 981-989.

Miura, N., Nakagawa, T., Touhara, K., and Ishikawa, Y. (2010) Broadly and narrowly tuned odorant receptors are involved in female sex pheromone reception in Ostrinia moths. *Insect Biochem Mol Biol* **40**: 64-73. doi: 10.1016/j.ibmb.2009.12.011

Monteforti, G., Angeli, S., Petacchi, R., and Minnocci, A. (2002) Ultrastructural characterization of antennal sensilla and immunocytochemical localization of a chemosensory protein in *Carausius morosus* BRÜNNER (Phasmida: Phasmatidae). *Arthropod Struct Dev* **30**: 195–205.

Morrill, W. L., Kushnak, G. D., Bruckner, P. L., and Gabor, J. W. (1994) Wheat stem sawfly damage, rates of parasitism, and overwintering survival in resistant wheat lines. *J Econ Entomol* **87**: 1373-1376.

Morrill, W. L. (1997) The wheat stem sawfly, *Cephus cinctus* Nort. (Hymenoptera: Cephidae) and associated parasitoids in the northern Great Plains of North America. *Trends Entomol* **1**: 171-174.

Mosbah, A., Campanacci, V., Lartigue, A., Tegoni, M., Cambillau, C., Darbon, H. (2003) Solution structure of a chemosensory protein from the moth *Mamestra brassicae*. *Biochem J* **369**: 39–44.

Nagnan-Le Meillour, P., Cain, A.H., Jacquin-Joly, E., Francois, M.C., Ramachandran, S., Maida, R., and Steinbrecht, R.A. (2000) Chemosensory proteins from the proboscis of *Mamestra brassicae*. *Chem Senses* **25**: 541–553.

Nakagawa, T., Sakurai, T., Nishioka, T., and Touhara, K. (2005) Insect sex-pheromone signals mediated by specific combinations of olfactory receptors. *Science* **307**: 1638–1642.

Oakeshott J. G., Devonshire, A. L., Claudianos, C., Ollis, D. L., and Russel, R. J. (2005) Comparing the organophosphorus and carbamate insecticide resistance mutations in cholin- and carboxyl-esterases. *Chemico Biol Interact* 157–158:269–275.

Oakeshott, J. G., Johnson, R. M., Berenbaum, M. R., Ranson, H., Cristino, A. S., and Claudianos, C. (2010) Metabolic enzymes associated with xenobiotic and chemosensory responses in *Nasonia vitripennis*. *Insect Mol Biol* **19**:147–163.

Ozaki, M., Wada-Katsumata, A., Fujikawa, K., Iwasaki, M., Yokohari, F., et al. (2005) Ant nestmate and non-nestmate discrimination by a chemosensory sensillum. *Science* **309**: 311–314.

Pelosi, P. (1998) Odorant-binding proteins: Structural aspects. *Ann NY Acad Sci* **855**: 281–293.

Pelosi, P., Iovinella, I., Felicioli, F and Dani, F.R. (2014) Soluble proteins of chemical communication: an overview across arthopods. *Front Physiol* **5**(320): 1-13.

Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., et al. (2003) TIGR Gene Indices clustering tools (TGICL): A software system for fast clustering of large EST datasets. Bioinformatics 19: 651-652.

Poupardin, R., Reynaud, S., Strode, C., Ranson, H., Vontas, J., and David, J. P. (2008) Cross induction of detoxification genes by environmental xenobiotics and insecticides in the mosquito *Aedes aegypti*: Impact on larval tolerance to chemical insecticides. *Insect Biochem Mol Biol* **38**: 540–551.

Pridgeon, J. W., Zhang, L., and Liu, N. N. (2003) Overexpression of CYP4G19 associated with a pyrethroid-resistant strain of the German cockroach, *Blattella germanica* (L.). *Gene* **314**: 157-163.

Ranson, H., and Hemingway, J. (2005) Mosquito glutathione transferases. *Rev Methods Enzymol* **401**: 226-241.

Ranson, H., Rossiter, L., Ortelli, F., Jensen, B., Wang, X., et al. (2001) Identification of a novel class of insect glutathione S-transferases involved in resistance to DDT in the malaria vector *Anopheles gambiae*. *Biochem J* **359**: 295-304. doi: 10.1042/0264-6021:3590295.

Rewitz, K. F., O'Connor, M. B., and Gilbert, L. I. (2007) Molecular evolution of the insect Halloween family of cytochrome P450s: phylogeny, gene organization and functional conservation. *Insect Biochem Mol Biol* **37**(8): 741-753.

Richards, S., Gibbs, R. A., Weinstock, G. M., Brown, S. J., Denell, R., Beeman, R. W., et al. (2008) The genome of the model beetle and pest *Tribolium castaneum*. *Nature* **452**: 949–955.

Robertson, H. M., Gadau, J., and Wanner, K. W. (2010) The insect chemoreceptor superfamily of the parastiod jewel wasp *Nasonia vitripennis. Insect Mol Biol* **19**: 121–136.

Robertson, H. M. and Wanner, K. W. (2006) The chemoreceptor superfamily in the honey bee, *Apis mellifera*: expansion of the odorant, but not gustatory, receptor family. *Genome Res* **16**: 1395–1403.

Robertson, H. M., Warr, C. G., and Carlson, J. R. (2003) Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* **100**: 14537–14542.

Rogers, M. E., Jani, M. K., and Vogt, R. G. (1999) An olfactory-specific glutathione-S-transferase in the sphinx moth *Manduca sexta*. *J Exp Biol* **202**(12): 1625-1637.

Rutzler, M. and Zwiebel, L. J. (2005) Molecular biology of insect olfaction: recent progress and conceptual models. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* **191**: 777–790.

Sakurai, T., Nakagawa, T., Mitsuno, H., Mori, H., Endo, Y., Tanoue, S., et al. (2004) Identification and functional characterization of a sex pheromone receptor in the silkmoth *Bombyx mori*. *Proc Natl Acad Sci USA* **101**: 16653–16658.

Sato, K., Pellegrino, M., Nakagawa, T., Nakagawa, T., Vosshall, L. B., and Touhara, K. (2008) Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* **452**: 1002–1006.

Scaloni, A., Monti, M., Angeli, S., and Pelosi, P. (1999) Structural analysis and disulfide-bridge pairing of two odorant-binding proteins from *Bombyx mori*. *Biochem Biophys Res Commun* **266**: 386–391.

Sheehan, D., Meade, G., Foley, V. M., and Dowd, C. A. (2001). Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem J* **360**(1): 1-16.

Singh, S. P., Coronella, J. A., Benes, H., Cochrane, B. J., and Zimniak, P. (2001) Catalytic function of *Drosophila melanogaster* glutathione S-transferase DmGSTS1-1 (GST-2) in conjugation of lipid peroxidation end products. *Eur J Biochem* **268**: 2912–2923. doi: 10.1046/j.1432-1327.2001.02179.

Smart, R., Kiely, A., Beale, M., Vargas, E., Carraher, C., Kralicek, A. V., et al. (2008) *Drosophila* odorant receptors are novel seven transmembrane domain proteins that can signal independently of heterotrimeric G proteins. *Insect Biochem Mol Biol* **38**: 770–780.

Stein, L. (2001) Genome annotation: from sequence to biology. *Nature Rev Genet **2**:* 493-503. doi:10.1038/35080529.

Steinbrecht, R. A. (1998) Odorant-binding proteins: expression and function. *Ann NY Acad Sci* **855**: 323–332.

Stevens, J. L., Snyder, M. J., Koener, J. F., and Feyereisen, R. (2000) Inducible P450s of the CYP9 family from larval Manduca sexta midgut. *Insect*

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013) MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Mol Biol Evol* **30**: 2725-2729.

Tu, C. P., and Tang, A. H. (1994) Biochemical characterization of Drosophila glutathione S-transferases D1 and D21. *J Biol Chem* **269**: 27876–27884.

Udomsinprasert, R., Pongjaroenkit, S., Wongsantichon, J., Oakley, A. J., et al. (2005) Identification, characterization and structure of a new Delta class glutathione transferase isoenzyme. *Biochem J* **388**: 763-771.

Vieira, F. G., Forêt, S., He, X., Rozas, J., Field, L. M., et al. (2012) Unique Features of Odorant-Binding Proteins of the Parasitoid Wasp *Nasonia vitripennis* Revealed by Genome Annotation and Comparative Analyses. *PLoS ONE* **7**(8): e43034. doi:10.1371/journal.pone.0043034

Vieira, F.G., Sanchez-Gracia, A., and Rozas, J. (2007) Comparative genomic analysis of the Odorant-binding protein family in 12 *Drosophila* genomes: purifying selection and birth-and-death evolution. *Genome Biol* **8**: R235.

Vogt, R. G. (2003) Biochemical diversity of odor detection: OBPs, ODEs and SNMPs. In: Insect Pheromone Biochemistry and Molecular Biology. G. J. Blomquist, R. G. Vogt (eds) London: Elsevier. Pp 391-446.

Vogt, R. G., Riddiford, L. M., and Prestwich, G. D. (1985) Kinetic properties of pheromone degrading enzyme: the sensillar esterase of *Antheraea polyphemus*. *Proc Natl Acad Sci USA* **82**: 8827–8831.

Vogt, R. G., Rogers, M. E., Franco, M. D., and Sun, M. (2002) A comparative study of odorant binding protein genes: differential expression of the PBP1-GOBP2 gene cluster in *Manduca sexta* (Lepidoptera) and the organization of OBP genes in *Drosophila melanogaster* (Diptera). *J Exp Biol* **205**: 719–744.

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A., and Axel, R. (1999) A spatial map of the olfactory receptor expression in the Drosophila antenna. *Cell* **96**: 725-736.

Vosshall, L. B. and Hansson, B. S. (2011) A unified nomenclature system for the insect olfactory coreceptor. *Chem Senses* **36**: 497–498.

Vosshall, L. B. and Stocker, R. F. (2007) Molecular architecture of smell and taste in Drosophila. *Annu Rev Neurosci* **30**: 505–533.

Wahl, D. V., Shanower, T. G., and Hoelmer, K. A. (2007) A new species of *Collyria* Schiødte (Hymenoptera: Ichneumonidae: Collyriinae), a parasitoid of *Cephus fumipennis* (Hymenoptera: Cephidae) in China, and potential biological control agent for *Cephus cinctus* in North America. *J Kans Entomol Soc* **80**: 43–50.

Wanner, K. W., Anderson, A. R., Trowell, S. C., Theilman, D. A., Robertson, H. M., and Newcomb, R. D. (2007) Female-biased expression of odourant receptor genes in the adult antennae of the silkworm, *Bombyx mori*. *Insect Mol Biol* **16**: 107-119.

Wanner, K. W., Nichols, A. S., Allen, J. E., Bunger, P. L., Garczynski, S. F., Linn, C. E. et al. (2010) Sex pheromone receptor specificity in the European corn borer moth, *Ostrinia nubilalis*. *PLoS ONE* **5**:e8685. doi: 10.1371/journal.pone.0008685

Wanner, K. W., Nichols, A. S., Walden, K. K., Brockmann, A., Luetje, C. W., and Robertson, H. M. (2007b) A honey bee odorant receptor for the queen substance 9-*oxo*-2-decenoic acid. *Proc Natl Acad Sci USA* **104**: 14383–14388.

Wanner, K. W., Willis, L. G., Theilmann, D. A., Isman, M. B., Feng, Q., and Plettner, E. (2004) Analysis of the insect os-d-like gene family. *J Chem Ecol* **30**: 889–911

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. (2009) Jalview Version 2-a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**: 1189-1191. doi:10.1093/bioinformatics/btp033.

Wei, S. H., Clark, A. G., and Syvanen, M. (2001) Identification and cloning of a key insecticide-metabolizing glutathione S-tranferase (MdGST-6A) from a hyper insecticide-resistant strain of the house-fly *Musca domestica. Insect Biochem Mol Biol* **31**: 1145-1153.

Weiss, M. J., Morrill, L. W., and Reitz, L. L. (1992) Wheat stem sawfly (Hymenoptera: Cephidae) Revisited. *Am Entomol* **38**(4): 241-245.

Wen, Z., Baudry, J., Berenbaum, M. R., and Schuler, M. A. (2005) Ile115Leu mutation in the SRS1 region of an insect cytochrome P450 (CYP6B1) compromises substrate turnover via changes in a predicted product release channel. *Protein Eng Des Sel* **18**: 191-199.

Wicher, D., Schafer, R., Bauernfeind, R., Stensmyr, M. C., Heller, R., Heinemann, S. H., and Hansson, B. S. (2008) *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature* **452**: 1007–1011.

Willingham, A. T. and Keil, T. (2004) A tissue specific cytochrome P450 required for the structure and function of Drosophila sensory organs. *Mech Dev* **121**(10): 1289-1297.

Yandell, M., and Ence, D. (2012) A beginner's guide to eukaryotic genome annotation. *Nature Rev Genet* **13**: 329-342. doi:10.1038/nrg3174.

Yao, C. A., Ignell, R., and Carlson, J. R. (2005) Chemosensory coding by neurons in the coeloconic sensilla of the Drosophila antenna. *J Neurosci* **25**(37): 8359-8367.

Zhou, J. J., Zhang, G. A., Huang, W., Birkett, M. A., Field, L. M., Pickett, J. A., and Pelosi, P. (2004) Revisiting the odorant-binding protein LUSH of Drosophila melanogaster: evidence for odour recognition and discrimination. *FEBS Lett* **558**: 23–26.

Zhou, X., Slone, J. D., Rokas, A., Berger, S. L., Liebig, J., et al. (2012) Phylogenetic and Transcriptomic Analysis of Chemosensory Receptors in a Pair of Divergent Ant Species Reveals Sex-Specific Signatures of Odor Coding. *PLoS Genet* **8**(8): e100293

CHAPTER 6

ANTENNAL EXPRESSION OF CANDIDATE OLFACTORY-REALTED GENES FROM THE WHEAT STEM SAWFLY, *CEPHUS CINCTUS*

Contribution of Authors and Co-Authors

Manuscript in Chapter 6
Author: Joanna C. Gress

Contributions: Conceived and implemented the study design and data collection. Collected and analyzed data. Wrote the manuscript.

Co-Author: Charles C. Carey

Contributions: Ran RNAseq data through pipeline analysis. Provided feedback on data analysis.

Co-Author: Aurélien Mazurie

Contributions: Provided feedback on data analysis.

Co-Author: Kevin W. Wanner

Contributions: Provided feedback on study design, data analysis, and comments on the manuscript.

Manuscript Information Page

Joanna C. Gress, Charles C. Carey, Aurélien Mazurie and Kevin W. Wanner
Insect Molecular Biology
Status of Manuscript:
__X_ Prepared for submission to a peer-reviewed journal
____ Officially submitted to a peer-review journal
____ Accepted by a peer-reviewed journal
____ Published in a peer-reviewed journal

Published by the Royal Entomological Society

Abstract:

Olfaction is extremely important for insect survival, mediating key pest behaviors such as host preference, mate choice, and oviposition site selection. Multiple antennal proteins are involved in olfactory signal transduction pathways including odorant binding proteins (OBPs), chemosensory proteins (CSPs), odorant receptors (ORs), ionotropic receptors (IRs) and odorant degrading enzymes (ODEs). The olfactory-related gene repertoire of the economically important agricultural pest, *Cephus cinctus*, a major pest of wheat in the Northern Plains of North America, was previously identified. In this study whole transcriptome RNA sequencing (RNA-seq) was used to compare the expression levels of 131 olfactory-related genes between male and female antennae, male and female whole bodies and whole larva. Enriched gene expression in the antennae and/or sex biased antennal expersion supported the olfactory function of candidate genes. These results advance the molecular knowledge of WSS olfaction and identify candidate genes that may serve as molecular targets in the development of new management strategies.

Introduction

In the peripheral olfactory system of insects, at least seven different protein families are involved in detecting odors. Odorant binding proteins (OBPs) and chemosensory proteins (CSPs) transport odors across the sensillum lymph, odorant receptors (ORs) and ionotropic receptors (IRs) initiate signal transduction across the olfactory neuron membrane after activation by a cognate odor, and odorant degrading enzymes (ODEs) clear the sensillum of chemical stimuli (Korsching, 2002; Leal, 2013). The olfactory sense mediates insect behaviors that are critical for survival, such as recognition of conspecifics, predator and prey relationships, locating food sources and female oviposition (Leal, 2013). Because of their critical role in interfacing with the external environment and mediating important pest behaviors, olfactory-related proteins are molecular targets for new insect pest control techniques. If the proteins that mediate odor detection can be disrupted, then an insect's ability to survive and reproduce is impaired.

The whole genome sequence of several hymenopteran species has now been published, including the honeybee *Apis mellifera* and the jewel wasp *Nasonia vitripennis*, and their olfactory-related genes have been annotated (Robertson & Wanner, 2006; Robertson et al. 2010). Both of these insects are beneficial, the honeybee providing pollination services and the jewel wasp representing a group of insect parasitoids that can control pest insect populations. The wheat stem sawfly (WSS), *Cephus cinctus* (Hymenoptera: Cephidea), is a basal lineage within the Hymenoptera that is primitively

phytophagous. The WSS has been a major pest of wheat in the Great Plains of North America for more than a century. Crop yield losses as a result of lodging and reduced head weight can amount to $100 million per year in the US and Canada (Wahl et al. 2004). Despite the economic importance of this pest, its genetic information has not been widely studied, including knowledge of the genes that mediate odor detection. The rapid advance in DNA sequencing technology was leveraged to rapidly address this deficit in the basic knowledge of the molecular biology of the WSS. Illumina sequencing was used to generate whole genome and transcriptome sequences. These sequences were used to annotate a total of 131 olfactory-related genes (Chapters 4 & 5). In this study, the RNA sequencing technique (RNAseq) was used to quantify the expression levels of the olfactory-related genes in male and female antennae, male and female whole bodies and whole larvae.

RNAseq is an efficient method for studying gene expression, including genes expressed in insect tissues (Bengtsson et al. 2012; Mitchell et al. 2012, Lundberg et al. 2013). During the last decade, microarrays have represented the state of the art technology for simultaneously monitoring the expression levels of thousand of genes (Shalon et al., 2006; Schena et al., 1998). DNA nucleotides complementary to the coding region of genes are spotted onto a glass slide (as many as 10,000 – 20,000 spots per array). mRNA labeled with fluorescent dye is hybridized to the array, and the amount of gene-specific transcript in the pooled RNA is relative to its hybridization to gene specific probes, measured by its location-specific fluorescence. Next generation sequencing (NGS) has opened a new methodology for the massively parallel analysis of gene specific

expression. Next-generation sequencers have platforms that allow for rapid and cost-effective generation of massive amounts of sequence data. RNAseq, the shotgun sequencing of all transcripts in a pool of RNA, provides a measure of transcript abundance (Marguerat & Bahler, 2010).

The use of NGS technologies for the analysis of RNA expression began with researchers working on small regulatory RNAs, whose sequence was typically too short to be captured adequately with the limited resolution provided by microarrays. Sequencing of short regulatory RNAs (Carthew & Sontheimer, 2009; Naqvi et al. 2009) led to whole transcriptome studies by RNAseq, of more than a dozen organisms (He et al. 2008; Core et al. 2008;Wang et al. 2008; Hillier et al. 2009; Hahn et al. 2009; Vera et al. 2008; Trick et al. 2009). Unlike the genome, the transcriptome dynamically changes in response to the tissue or the environment or to intrinsic developmental programs. RNAseq has now been used to analyze transcriptome sequences of several cell types or physiological conditions (Marguerat & Bahler, 2010).

The quantitative analysis of gene-specific transcript number is based on the countable, almost digital nature of the short RNA reads and their ability to be assigned to specific genes based on sequence identity. Early studies assessed the ability of RNA-seq to measure differential gene expression (Marioni et al. 2008; ‘t Hoen et al. 2008; Bloom et al. 2009). These studies all conclude that RNA-seq performs at least as well as microarrays in assessing differential genes expression, provided an adequate sequencing depth. Compared to microarrays, RNA-seq can provide additional genetic information from the transcripts, including assessing alternate splice variants. Another characteristic

of RNA-seq is its high sensitivity, its ability to detect transcripts of low abundance compared to microarrays. In this study we used the RNAseq technique (Xu et al. 2009; Grosse-Wilde et al. 2011; Legeai et al. 2011; Pitts et al. 2011) to quantify the expression levels of 131 olfactory-related WSS genes in five different tissues.

Materials and Methods

RNA Preparation and Illumina Sequencing

Adult male and female wheat stem sawflies were collected from infested wheat fields near Amsterdam MT during July 2012. Antennae were dissected from male and female adults within 24h of collection along with whole bodies. Larvae were collected from wheat stems and smooth brome stems near Conrad and Mocassin, MT in 2011. All tissues were frozen on dry ice and stored at -80°C. For transcriptome sequencing, total RNA was extracted from 500 antennal pairs collected from both male and female insects respectively. Total RNA was extracted from male and female WSS antennae separately as well as whole adults and larvae using a Dounce homogenizer and purified with the RNeasy Mini Kit (Qiagen, Valencia, CA). RNA quality was verified using a 2100 Bioanalyzer RNA Nanochip (Agilent, Santa Clara, CA) and samples had RNA Integrity Number (RIN) values higher than 8.5. The RNA was quantified using a NanoDrop ND-1000 Spectrophotometer (NanoDrop, Wilmington, DE).

Illumina paired end sequencing was performed at the University of Illinois Urbana-Champaign W.M. Keck Center for Comparative and Functional Genomics from 20 ug of male and 20 ug of female antennal RNA and 20ug of whole male and whole

female adults and larvae. First, reads with adaptors or reads containing more than 5% unknown nucleotides (Ns) were removed computationally from the data set. Secondly, low-quality reads containing more than 20% suspect-nucleotides of Phred Quality Score less than 10 were filtered out. Finally, both ends of each individual read were evaluated for three successive suspect-nucleotides and these were trimmed. The WSS genome was sequenced in collaboration with the University of Illinois Urbana-Champaign. Genomic DNA (gDNA) from a single haploid male sawfly collected near Amsterdam, MT was used to sequence 500 bp and 1.5 kb shotgun read libraries. gDNA from additional male sawflies was used to sequence 3 kb and 5 kb mate-pair read libraries.

Genome and GFF annotation sources

The final assembly of the *C. cinctus* genome (NCBI Genbank #s KB465430.1-KB467404.1) was annotated at the University of Utah using the MAKER2 gene annotation pipeline (Holt & Yandell, 2011). This effort produced a list of predicted proteins and their predicted DNA coding sequences from the WSS genome.

The predicted gene models were formatted as GFF files (plain text, 9 column, tab-delimited files). Manually curated olfactory genes (Chapters 4 & 5) were added to the MAKER2 list of predicted genes using bedtools (version 2.20.1) intersect tool, and custom scripts to remove MAKER2 genes redundant with the manual list.

RNA-seq analysis

Reads were inspected for quality using fastqc (version 0.10.1). Trimmomatic (version 0.32) was used to quality trim and filter reads, and to remove reads containing

matches to Illumina sequencing adapters, using default software parameters and the provided TruSeq3-PE.fa adapter sequences. Resulting reads with a minimum length of 80 bp were aligned to the wheat stem sawfly genome using TopHat (version 2.0.12). Parameters for TopHat alignment included max-segment-intron and max-intron-lengths of 10,000 bp, max-multi-hits = 1, no-mixed reads (meaning only reads that aligned as pairs were retained). The modified WSS GFF file was used as a reference for TopHat alignments of the RNA-seq reads. HTSeq-count (HTSeq package, version 0.6.1) software was used to counts reads aligning to CDS regions of each gene as defined in the custom GFF file. EdgeR (version 3.6.8) software was used to generate the normalized reads per kilobase transcript (RPKM). R (R-project.org version 3.1.1) software was used to generate gene expression figures expressed as fragments per kilobase of exon (FPKM). Heatmaps were generated for subsets of data using the R package pheatmap option (version 0.7.7).

Results

The expression of 131 olfactory-related WSS genes in male and female antennae, male and female whole bodies and whole larvae was quantified by RNA-seq. Gene expression levels are quantified by FPKM for each olfactory-related gene. FPKM values for each gene within each tissue are used to calculate tissue-biased values to assess sex and/or tissue biased expression.

Odorant Binding Proteins

Insect OBPs are small soluble proteins, many of which are secreted into the lymph surrounding the chemoreceptor neurons, and that transport odorant molecules (Vogt et al. 1999). CcOBPs exhibited some of the highest levels of transcript expression in antennal tissue of all olfactory-related genes assayed in this study (antennal FPKM as high as 39,822, CcOBP6, Table 6.1), and all but three were enriched in the antennae compared to whole bodies. CcOBPs 1 - 4, 6 and 13 were expressed at antennal FPKM levels ranging from ~ 2,000 to 7,000, and these values were 24 to 228 times higher compared to whole bodies (Table 6.1). The most highly expressed OBP in WSS antennae is CcOPB6, 48 times higher compared to expression in the whole body and 3,000 time higher than expression in whole larvae. OBPS, CcOPB1 and 4, exhibited sex biased expression in WSS antennae. These two OBPs were expressed 3-4 times more abundantly in female compared to male antennae (Table 6.1; Figure 6.1). Three OBPs, CcOBP5, 7 and 9, were not expressed in the antennal tissue at significant levels.

Chemosensory Proteins

The CSPs, like the OBPs, are proposed to bind odors and transport them across the sensillar lymph, although the role of CSPs in olfaction remains unclear. Only two of the CSPs in this study were highly abundant in the antennal tissues, CcCSP2 and 2b. The remaining CSPs were more highly expressed in whole bodies or larvae, suggesting functions other than chemoreception. CcCSP3 and 5 were more highly expressed in whole adult insects while CcCSP 4 and 7 were most abundant in whole larvae (Table 6.1; Figure 6.1).

CcCSP2 was the most highly expressed olfactory related gene in WSS antennae, FPKM score of 84,424 in female antennae. This gene however, was also highly expressed in whole bodies, and was only 6 times more abundant in the antennae. CcCSP2b was expressed at relatively lower levels in the antennae, having an FPKM score of 3,855 in female antennae. This gene however, was 138 times more abundant in the antennae compared to whole bodies (Table 6.1). CcCSP2b was expressed 2.2 times higher in the female compared to male antenna (Figure 6.2).

Odorant Receptors

Insect ORs function as heterodimers to form ligand gated ion channels (Sato et al. 2008; Wicher et al. 2008), with single ORs conferring odor sensitivity to individual odorant receptor neurons (ORNs). OR transcripts exhibited antennal abundances ranging from 0.16 to 720 FPKM (Table 6.1, Figure 6.6), levels far lower than OBPs and CSPs. While their overall expression levels were lower, all the OR genes exhibited significant enrichment in the antennae compared to whole bodies or whole larvae. The ubiquitous co-receptor CcORco was expressed equally in male and female antennae at levels relatively higher than most ORs (FPKM; 324 female antennae, 343 male antennae). ORco was expressed 43 times more abundantly in antennae compared to whole bodies. OR6 and 6B were the only Ors expressed at higher levels than ORco (FPKM values = 711 and 720 in male antennae, respectively) and their antennal enrichment was 46 times greater than whole bodies. CcOr3, 6, 15, 25, and 29 were the next most highly expressed ORs in WSS antennae, all with FPKM values above 100.0 CcOr4, 7, 12, 14, 16, 27 and

42 exhibited moderate FPKM values, ranging from 50 to 82. While moderately expressed, their values continued to demonstrate antennal enrichment. Or4 for example, with an FPKM value of 53.7, was 33 times higher compared to whole bodies. CcOr5, 13, 17, 18, 20, 30, 36, 37, 44 and 58 exhibited moderately low expression levels with FPKM values between 25 and 47. CcOr1, 8, 19, 48 and 53 were expressed at low levels, FPKM values from15 to 24. Finally, the remaining 21 ORs (CcOr11, 21, 31, 32, 34, 38, 39, 40, 41, 43, 45, 46, 47, 49, 50, 51, 52, 54, 56, 59 and 60) were expressed at very low levels, with FPKM values below 14.9 (Table 6.1). While expressed at low or very low levels, these ORs maintain their antennal enrichment. Or1 with an FPKM value of 17.5 was 31 times more abundant compared to whole bodies, and Or31, whose FPKM value was only 8.17 (42 times lower than ORco), was 68 times more abundant compared to whole bodies.

The antennal expression levels of 8 WSS Ors were sex biased with more than 5 fold differences. CcOR14 and 26 were expressed at higher levels in the male antennae while CcOr7, 11, 18, 34, 45 and 48 were expressed at higher levels in the female antennae. CcOr26 had the largest fold change in expression, 18.5 times more abundant in male compared to female antennae. CcOr7 exhibited the next highest fold change, 9.95 times more abundant in female antennae (Figure 6.6). WSS OR expression was mostly restricted to the antennae, and very little expression was noted in whole bodies and larvae (Table 6.1).

Ionotropic Receptors

CcIR transcript levels in the antennae were lower even than the Ors, ranging from 0.96 to 84, the lowest values of any olfactory-related gene family quantified in this study (Table 6.1). Six IRs were expressed in the antennae at FPKM levels above 11 FPKM (Table 6.1), with four above 40 FPKM (CcIr1, 3, 7 and 8). Similar to the Ors, IRs with low expression levels maintained antennal enrichment. IR1, 2 and 9 with FPKM levels of 84.39, 14.41 and 8.05 exhibited antennal enrichment factors of 33, 12 and 67 times, respectively.

DmelIR25a and 8a expression (Benton et al. 2009) and function (Abuin et al. 2011) in ORNs suggest they are co-receptors for other ligand-binding IRs (Benton et al. 2009). We previously identified CcIr2 as the DmelIR25a homolog and CcIr1 as the Dmel IR8a homolog. The function of IR25a and IR8a as co-receptors is consistent with their higher expression levels compared to the other WSS IRs in this study. None of the IRs exhibited significant sex biased expression in the WSS antennae (Figure 6.7).

Odorant Degrading Enzymes

Carboxylesterases

Next to the OBPs and CSPs, the CCEs exhibited the highest levels of transcript abundance in *C. cinctus* (FPKM range 0.2 – 4,847; Table 6.1). Four WSS esterases were abundantly expressed in the antenna (CcCCE2, 3, 8 and 9) while one was abundant in all tissues (CcCCE1) (Table 6.1). Antennal FPKM values of CcCCE2 and 9 were above 4,000 and 1,300, respectively, while the values for CcCCE3 and 8 were 250 and 228,

respectively (Table 6.1). CcCCE2 and 9 were enriched in the antennal tissue 66 and 146 times more than whole bodies, respectively. CcCCE3 and 8 were enriched by factors of 8-10 times. None of the CCEs exhibited significant sex biased expression in WSS antennae. CcCCE8, with a 1.4 fold higher expression level in female compared to male antennae represented the highest value. CcCCE10a and 10b were expressed at much higher levels in the larva compared to whole adults, suggesting a physiological role in larval development. CcCCE10a and 10b form a large monophyletic clade that groups with only one ortholog from *A. mellifera*, AmGB16342PA, an α-esterase type CCE (Oakeshott et al. 2010).

Glutathione S-Transferases

The abundance of CcGST expression ranged from FPKM values of 8 to 2,328, half that of the most abundant CCE gene (Table 6.1). Three CcGSTs were the most abundant in antennal tissue: CcGST1a (FPKM score = 97), CcGST1c (FPKM score = 137) and CcGST1b (FPKM score = 2,238) (Table 6.1). While the most abundant, CcGST1b is enriched only 4 times compared to whole bodies, and was expressed in all sawfly tissues tested, suggesting it is ubiquitously expressed. While CcGST1a and CcGST1c were relatively less abundant, their antennal enrichment factors were higher, 38 and 68 times, respectively. CcGST1b was slightly male biased in the antennae, 2.15 times higher compared to female antennae. CcGST1a and CcGST1c were expressed equally between male and female antennae. All three of these GSTs belong to the Delta class of GSTs.

Cytochrome P450s

The cytochrome P450 transcript abundance was much lower compared to other classes of olfactory-related genes, with FPKM values ranging from 0.03 to 501 (Table 6.1). Only six of 29 P450s were expressed in the antennae, CcP450_9, 15, 23, 24, 27 and 30. Two were expressed 13-17 times more abundantly in antennae compared to whole bodies (CcP450_9 and 23). The remaining four were expressed at levels 5 - 9 times higher in antennae compared to whole bodies. None of the CcP450s exhibited significant sex biased expression.

There are four main phylogenetically distinct groupings P450 genes. The six P450s genes expressed in the WSS antennae belong to three of the phylogenetic classes. CcP450_9 is a member of CYP4, CcP450_15, 23, 24, 27 all belong to the CYP3 group and P450_30 is a member of the mitochondrial CYP group.

Discussion

Profiling tissue specific gene expression is a first step towards characterizing gene function. Tissues and organs dedicated to particular functions express a combination of general house keeping genes and genes that have evolved to carry out the particular function. Olfactory neurons express genes involved in cellular metabolism, but also express olfactory-related genes that function only in the detection of volatile odors. Practical experimental constrains often limit the number and specificity of tissues dissected and analyzed for gene expression, and the first preliminary studies of large gene families often utilize more general tissue collections. In this study the expression of 131

WSS olfactory-related genes was analyzed in male and female antennae, male and female whole bodies and whole larvae. The interpretation of gene expression levels in whole insect bodies is limited since it represents a sum of all the different tissues. The goal of this study was to identify genes that function in the peripheral olfactory system; antennae are the primary organ housing olfactory neurons and their supporting cells. Three patterns of gene expression support a putative role in olfaction: 1) high expression levels in the antennae, 2) enrichment in antennae compared to whole bodies and whole larvae (supporting antenna-specific expression and function), and 3) sex-biased expression in antennae (supporting sex specific olfactory functions). OR, IR and OBP gene families function almost exclusively to detect odors, and the expression of almost all of the WSS members of these genes families were enriched in the antennae. CSP, CCE, GST and P450 gene families have functions broadly related to physiology, and typically only a few members of these genes families have evolved olfactory functions. In this study, only a subset of these gene families exhibited higher expression levels in the antennae.

During the last decade microarrays represented the dominant technology for profiling the tissue specific expression of large numbers of genes simultaneously. RNAseq has quickly augmented or replaced microarrays. In either case, real time quantitative RT- PCR (qRT-PCR) is used to validate the expression of specific genes, from the global expression data. The expression of 28 WSS Or genes was assayed previously by qPCR (Gress et al. 2013). The expression results for these Ors obtained by RNAseq in this chapter agreedvery well with the previously published data, validating the present approach. In both cases CcORco was the second most highly expressed OR

gene, second only to CcOR6. Gress et al. (2013) identified CcOr 26 as male-biased, with expression at levels 14.7 - 16.2 times higher in male compared to female antennae. In this study RNAseq identified OR26 as having an 18.5 fold higher expression in male antennae. qPCR analysis identified five ORs as female-biased (Gress et al. 2013) and RNAseq analysis provided supporting results in each case. However, results from the two different techniques do not agree perfectly. RNAseq in this study found CcOR7 to be expressed at levels 9.95 times higher in male compared to female antennae, results contrary those reported in Gress et al. (2013). When such discrepancies arise, it will be important to conduct additional qPCR experiments to resolve the expression results.

OBPs and CSPs transport odors across the sensillar lymph and are the first proteins to interact with odorants and ORs. Previous studies using other insects including *B. mori* have shown that that some OBPs and most CSPs are expressed throughout insect development (including pupae) in non olfactory tissues (Gong et al. 2007, 2009). Similar to these published results, the expression of 6 of 9 WSS OBPs were enriched in antennae and only two of 7 CSPs were expressed at abundant levels in the antennae. Based on very high expression levels in the antennae, CcOBP6 and CcCSP2 are interesting gene candidates for future functional studies. Since OBPs are proposed to participate in odor discrimination by binding a defined group of molecular structures (Rutzler & Zwiebel 2005), the OBPs specifically expressed in larvae (CcOBP9) may function in larval olfaction. However, more detailed expression analysis would be required to identify potential physiological functions. CSPs were first defined as chemosensory proteins, but several subsequent expression studies have revealed that they are expressed throughout

many different tissues (Pelosi et al. 2006; Jacquin-Joly et al. 2001) and likely function in other physiological processes such as development (Kitabayashi et al. 1998). Thus, it is possible that the larvae-specific CSPs (CcCSP4), in addition to the other CSPs expressed in larvae (CcCSP2 and 7), participate in larval development.

Insect ORs are expressed in specific subsets of olfactory neurons, one OR gene for each olfactory neuron type. The response spectra of the OR that forms a ligand gated ion channel dictates the physiological response of the olfactory neuron to odors. Insect antennae will have different proportions of olfactory neurons tuned to specific odor classes. For example, male moths will have a high proportion of olfactory neurons tuned to female produced sex pheromones and female moths will have higher proportions of olfactory neurons tuned to host plant volatiles. Because OR gene expression levels in the antennae correlate roughly with abundance of olfactory neurons expressing the OR gene, highly abundant and/or sex biased OR expression have proved a reliable indicator of ORs that detect behaviorally important odors (Wanner et al. 2007; Jordan et al. 2009; Legeai et al. 2011). CcORs 3, 6, 6B, 15, 25 and 29 are interesting candidates because of their high overall expression levels. OR26, identified by Gress (et al. 2013) is a male-biased receptor that is a candidate for the detection of female produced pheromone. This study also identified CcOr14 as expressed at levels 10 times higher in male compared to female antennae. CcORs 7, 11, 18, 34, 45 and 48 are candidate host plant volatiles receptors that might mediate host seeking behavior, because of their higher levels of expression in female antennae. Alternatively, female biased ORs may detect male produced pheromones used in courtship behavior. To date only one OR in Hymenoptera has been

functionally characterized, AmOr11 which responds to 9-ODA, the main component of the ‘queen substance’ in *A. mellifera* (Wanner et al. 2007). No WSS ORs identified in this study have orthology with AmelOr11.

Insect IR olfactory receptors have evolved from ionotropic glutamate receptors (Croset et al. 2010), a family of conserved ligand-gated ion channels associated with synaptic communication in both eukaryotes and prokaryotes. Since their discovery (Benton et al. 2009), IRs have been primarily described as olfactory receptors in *D. melanogaster* (Abuin et al. 2011; Ai et al. 2013), although other functional roles have been assigned to IRs expressing in non-antennal tissues (Senthilan et al. 2012; Zhang et al. 2013). Like the ORs, the WSS IRs were expressed at low levels, but exhibited significant enrichment in the antennae, consistent with expression in olfactory neurons. None of the IRs were sex biased in their expression, consistent with their proposed function as receptors of essential nutrients. It has been proposed that the ancestral chemosensory function of IRs is likely to be in the detection of water-soluble, non-volatile compounds, and that antennal IRs gained olfactory function (Croset et al. 2011). Interestingly, CcIR7, 8 and 9 are homologous to *Drosophila* IRs that respond to diverse aldehydes and acids. CcIr17 is homologous to the *Drosophila* phenylethyl amine receptor (DmelIR76b) (Ai et al. 2010). The detection of these compounds by *C. cinctus* IRs remains to be verified.

In contrast to ORs and OBPs, whose entire gene families have evolved primarily to function in olfaction, ODEs have evolved sporadically from three large gene families that function in biotransformation pathways. Individual members of the CCE, GST and

P450 gene families have evolved to be expressed in the sensillum lymph, where these enzymes participate in the rapid deactivation of insect odor and pheromone signals (Ishida & Leal, 2002). Collectively, the CCE, GST and P450 enzymes carry out the majority of metabolic transformations responsible for disarming toxic xenobiotics (Ranson et al. 2002; Feyereisen, 2005; Oakeshott et al. 2005; Ranson & Hemingway, 2005; Li et al. 2007). The P450s and CCEs in particular function to clear signals related to the reception of kairomones and pheromones (Rogers et al. 1999; Ishida & Leal, 2002, 2005).

A total of 49 WSS ODE transcripts were identified from the antennal transcriptome. Candidate ODE genes involved in olfaction were identified by their high expression levels in antennae, antennal enrichment compared to whole bodies, and sex biased expression. Four WSS CCE genes (CcCCE2, 3, 8 and 9) met these criteria, but CcCCE2 and 9 are the most interesting candidates due to their high levels of expression and specific enrichment in antennae. CcCCE2, 3, 8 and 9 are orthologous to the α-esterase group with dietary/detoxification functions (Claudianos et al. 2006). α-esteraes have been consistently implicated in organophosphate insecticide resistance (Oakeshott et al. 2005; Hartley et al. 2006). In another Hymenoptera species, *N. vitripennis*, an α-esterase CCE EST was identified from the wasp's venom gland, although its function in venom metabolism was not clear (de Grafet al. 2010; Oakeshott et al. 2010). Beckage and Gelman (2004) hypothesized that the CCE might disrupt juvenile hormone metabolism in the host that receives the venom. CcGST 1a, 1b and 1c, members of the Delta class of GSTs, were expressed in WSS antennae. The Delta class of GSTs is unique to insects and

contains the majority of the GSTs associated with detoxification of insecticides. A small subset of the Delta class insect GSTs can catalyse the dehydrochlorination of the organochlorine insecticide DDT (Tang & Tu, 1994; Ranson et al. 2001; Lumjuan et al. 2005) and other members of the insect specific GST class metabolize organophosphate insecticides (Huang et al. 1998; Wei et al. 2001). Whereas other classes of insect GSTs have been proposed to play a secondary role in protection against insecticides, for example by ameliorating the effects of oxidative stress, all those GSTs directly implicated in insecticide metabolism belong to the Delta or Epsilon classes (Claudianos et al. 2006). No epsilon class GSTs have been identified in *Nasonia*, *Apis* (Claudianos et al. 2006; Oakeshott et al. 2010) or *C. cinctus* (this chapter). While CcGST1b was the most abundant, CcGST1a and 1c maybe be more interesting targets for functional characterization since their antennal enrichment factors were much higher than CcGST1b.

Several members of dipteran and lepidopteran P450s class CYP6 radiations are involved in resistance to a broad range of insecticides (OPs, synthetic pyrethroids, DDT and neonicotinoids; Carino et al. 1994; Liu & Scott, 1996; Daborn et al. 2002; Li et al., 2007; Muller et al. 2008) and/or the detoxification of host plant allelochemicals in the gut (Danielson et al. 1997; Mittapalli et al. 2005;Wen et al. 2005; Mao et al. 2006, 2007; Li et al. 2007; Rupasinghe et al. 2007). Less is known of the functions of the P450 CYP9 class but the limited evidence available also implicates them in the detoxification of insecticides and other allelochemicals (Stevens et al. 2000; Poupardin et al. 2008), with some evidence also for involvement in semiochemical metabolism (Maibeche-Coisne et

al. 2005). The sawfly P450 class CYP6 CcP450_23 and 27 and to a lesser extent P450 class CYP9s CcP450_9, 15 and 24 are thus good candidates for involvement in its detoxification of insecticides and host metabolites.

Cytochrome P450s are large families of genes that function in detoxification and other metabolic pathways. Many gene lineages are highly conserved because their functions are important as house keeping genes. Members of the P450 CYP4J+ clade, that include CcP450_9, have been linked to lipid metabolism (Simpson, 1997; Feyereisen, 2005) and caste determination in the fire ant *Solenopsis invicta* (Liu & Zhang, 2004). Related P450 clades have been implicated in pheromone metabolism (Maibeche-Coisne et al. 2002, 2005), juvenile hormone and metabolism and DDT, pyrethroid, and carbamate insecticide resistance (Sutherlandet al.1998; Scharf et al. 2001; Maibeche-Coisne et al. 2002). CcP450_30 is an orthologue of the *D. melanogaster* Halloween gene shd, which encodes the steroid 20-hydroxylases (Gilbert, 2004). These mitochondrial P450s are 1:1:1:1:1 orthologues in the honeybee, jewel wasp, sawfly, mosquito and fruit fly, and CcP450_30 likely has a similar function. Only 6 CcP450s were expressed in the antennae of WSS, CcP450_9, 15, 23, 24, 27 and 30. None exhibited high levels of antennal enrichment. , CcP450_9 and 23 may be the most promising olfactory-related candidates since they were expressed 13-17 times more abundantly in the antennae compared to whole bodies.

In this chapter RNA-seq based differential analysis of WSS olfactory-related gene expression supported the putative functional roles of a large number of OBP, CSP, OR and IR genes in olfaction. Based on overall expression levels, antennal enrichment and

sex biased expression, several of these candidate olfactory genes were prioritized for future functional studies. ODE genes in the CCE, GST and P450 genes families were more commonly expressed in non olfactory tissues and were not significantly enriched in the antennae. However, 4 CCE, 3 GST and 6 P450 genes were identified as candidates for future functional studies. Expression levels of these candidate genes need to be validated by qPCR prior to functional studies. The aim of future functional studies is to identify olfactory proteins that mediate import WSS behaviors and to develop these genes as new targets for insect control.

Table 6.1 Expression levels of olfactory-related genes in five wheat stem sawfly tissues.

FPKM						
Olfaction Gene	antenna female	antenna male	whole female	whole male	larva	Fold Difference Fa/Ma
OBP1	3662.83	1075.78	15.74	40.15	2.05	3.40
OBP2	1822.57	3132.74	13.09	128.37	22.92	0.58
OBP3	189.25	122.99	6.99	10.32	15.00	1.54
OBP4	6945.76	1915.51	41.85	29.41	1.86	3.63
OBP5	6.08	2.65	0.64	0.84	20.40	2.30
OBP6	24481.85	39822.45	250.23	823.74	13.22	0.61
OBP7	1.30	0.61	0.10	292.88	0.40	2.13
OBP9	1.65	0.76	1.75	11.80	11.45	2.16
OBP13	4716.64	7288.16	31.01	158.82	1.91	0.65
CSP2	85424.12	67675.69	14059.63	24438.91	488.53	1.26
CSP2b	3855.29	1788.35	27.72	93.18	1.44	2.16
CSP3	0.12	0.48	584.12	262.72	9.33	-3.08
CSP4	55.11	15.78	342.96	561.94	4889.74	3.49
CSP5	3.80	2.29	53.12	116.24	38.20	1.66
CSP6	5.37	4.04	0.45	0.18	7.31	1.33
CSP7	37.95	43.49	188.90	270.87	634.37	0.87
Or01	17.46	17.67	0.56	0.39	0.00	0.99
Or03	172.44	36.03	0.74	0.76	0.09	4.79
Or04	53.72	16.02	1.65	1.54	0.33	3.35
Or05	47.31	33.53	0.93	1.41	0.67	1.41
Or06	173.93	711.27	1.47	14.27	0.17	-4.09
Or06b	206.10	720.12	3.04	15.62	0.14	-3.49
Or07	71.19	7.15	0.78	1.30	0.04	9.95
Or08	19.30	12.60	0.19	0.69	0.21	1.53
Or09	23.58	11.18	0.49	1.08	0.17	2.11
Or11	12.25	2.19	0.06	0.15	0.00	5.59
Or12	51.00	19.32	2.79	1.20	0.24	2.64
Or13	46.72	34.79	0.39	0.68	0.00	1.34
Or14	15.18	82.02	0.15	1.89	0.00	-5.40
Or15	98.30	100.90	0.79	1.85	1.04	0.97
Or16	67.90	35.82	0.45	0.74	0.08	1.90
Or17	31.36	6.59	1.33	0.37	0.70	4.76

Table 6.1 Cont

FPKM						
Olfaction Gene	antenna female	antenna male	whole female	whole male	larva	Fold Difference Fa/Ma
Or18	32.97	5.75	1.11	0.94	0.42	5.73
Or19	17.66	10.47	1.15	1.39	0.12	1.69
Or20	27.47	7.98	0.31	0.36	0.20	3.44
Or21	12.52	5.78	0.11	0.73	0.00	2.17
Or25	123.21	231.92	1.22	5.60	0.00	0.53
Or26	9.28	171.94	0.18	3.84	0.04	-18.54
Or26b	46.78	123.55	0.43	2.75	0.08	-2.30
Or27	45.77	51.89	0.30	0.88	0.80	0.88
Or29	74.91	172.18	0.80	4.43	0.89	0.44
Or30	44.45	35.33	8.36	14.83	11.90	1.26
Or31	1.79	8.17	0.00	0.12	0.04	-4.55
Or32	5.13	1.33	0.06	0.03	0.13	3.85
Or34	8.50	3.38	1.30	1.18	2.15	2.51
Or36	28.44	13.23	0.03	0.21	0.00	2.15
Or37	23.09	31.39	0.19	1.41	0.49	0.74
Or38	0.16	0.21	0.25	0.00	0.11	0.78
Or39	1.66	0.73	0.21	0.74	0.05	2.29
Or40	6.94	5.59	0.18	0.19	0.00	1.24
Or41	1.44	1.43	2.20	1.84	1.00	1.01
Or42	56.35	15.57	11.52	2.45	0.61	3.62
Or43	4.15	5.96	0.03	0.17	0.04	0.70
Or44	36.13	10.55	0.54	0.34	0.04	3.42
Or45	11.34	1.63	0.12	0.13	0.08	6.96
Or46	11.07	4.08	0.49	1.10	0.47	2.72
Or47	11.24	2.52	0.74	0.30	0.25	4.46
Or48	19.61	3.81	0.15	0.16	0.28	5.15
Or49	4.36	4.66	0.75	0.90	0.36	0.94
Or50	14.31	7.55	0.75	0.31	0.16	1.90
Or51	14.97	4.31	0.36	0.60	0.51	3.47
Or52	7.94	1.52	0.13	0.27	0.04	5.22
Or53	21.77	11.18	0.41	1.02	0.00	1.95
Or54	12.09	2.59	0.08	0.16	1.15	4.66
Or56	8.38	1.77	0.16	0.00	0.08	4.73

Table 6.1 Cont

FPKM						
Olfaction Gene	antenna female	antenna male	whole female	whole male	larva	Fold Difference Fa/Ma
Or58	27.37	9.37	1.46	2.41	0.76	2.92
Or59	9.68	5.29	0.07	0.15	0.10	1.83
Or60	10.98	10.81	0.14	0.21	0.04	1.01
Orco	324.79	343.89	2.18	8.10	0.58	0.94
Ir1	75.09	51.09	2.29	2.56	0.50	1.47
Ir2	14.41	12.75	1.25	5.65	3.48	1.13
Ir3	43.15	17.18	2.04	4.37	0.59	2.51
Ir4	2.45	1.15	0.36	0.15	0.19	2.12
Ir5	1.39	0.96	1.48	6.63	1.98	1.45
Ir7	49.49	33.69	3.94	3.79	0.36	1.47
Ir8	84.39	26.01	0.75	1.32	0.25	3.24
Ir9	8.05	4.47	0.12	0.32	0.22	1.80
Ir11	11.75	5.65	2.35	6.07	1.59	2.08
Ir15	2.68	2.45	0.28	0.24	0.37	1.09
CCE1	427.30	844.76	400.36	1327.38	33.11	-1.98
CCE2	4126.24	4847.59	33.29	77.21	2.04	0.85
CCE3	249.99	252.55	29.86	123.98	29.54	0.99
CCE4	43.29	38.75	30.21	80.12	14.73	1.12
CCE4b	16.23	14.60	7.10	10.81	10.58	1.11
CCE7	22.28	20.48	1.74	2.84	25.38	1.09
CCE8	228.74	161.02	21.73	31.50	0.44	1.42
CCE9	1400.32	1326.77	8.54	33.86	3.12	1.06
CCE10a	1.75	2.78	132.91	10.02	289.52	0.63
CCE10b	2.01	1.98	31.09	5.52	844.06	1.01
CCE10c	0.16	0.04	21.02	0.32	59.68	3.92
CCE10d	0.02	0.02	7.93	0.36	110.91	0.98
GST1a	97.65	73.72	2.44	10.71	50.21	1.32
GST1b	1080.47	2328.04	94.23	560.75	842.87	-2.15
GST1c	137.16	137.29	2.02	7.66	1.78	1.00

Table 6.1 Cont

FPKM						
Olfaction Gene	antenna female	antenna male	whole female	whole male	larva	Fold Difference Fa/Ma
GST2	8.32	4.30	0.19	1.79	0.00	1.93
GST4	42.74	46.58	45.83	52.91	34.95	0.92
GST5	106.62	77.41	146.90	77.00	251.21	1.38
GST6	41.25	42.65	11.84	40.34	22.25	0.97
GST7	47.78	61.10	148.37	120.79	109.14	0.78
P450_1	9.34	4.45	1.13	3.34	3.84	2.10
P450_2	5.76	4.04	4.26	10.91	14.47	1.43
P450_3	0.71	0.75	0.12	0.28	0.66	0.95
P450_5	0.54	0.18	1.47	0.94	1.21	2.94
P450_6	0.93	0.79	0.23	0.42	0.12	1.18
P450_7	3.11	3.02	13.64	26.71	17.06	1.03
P450_8	2.33	1.57	9.09	1.86	0.98	1.48
P450_9	62.25	119.42	4.59	73.76	0.38	0.52
P450_10	60.61	24.02	3707.90	7621.96	742.30	2.52
P450_11	3.94	2.17	13.98	12.73	5.62	1.82
P450_12	18.41	9.06	8.56	10.69	22.62	2.03
P450_13	0.96	1.53	3.86	5.65	27.19	0.63
P450_14	30.62	33.31	29.07	40.69	38.18	0.92
P450_15	182.81	146.90	24.62	44.77	0.06	1.24
P450_16	6.88	5.02	1.59	1.27	32.33	1.37
P450_17	93.42	63.87	16.69	44.90	2.31	1.46
P450_18	33.21	19.88	33.42	38.53	7.34	1.67
P450_19	185.28	116.45	77.99	180.27	13.67	1.59
P450_20	4.72	19.11	3.90	8.70	257.37	-4.05
P450_21	45.68	50.83	314.47	375.40	104.37	0.90
P450_22	4.83	3.09	24.16	34.34	16.13	1.56
P450_23	501.03	410.35	29.07	88.50	1.09	1.22
P450_24	154.85	118.59	25.28	44.14	0.00	1.31
P450_25	0.69	0.64	6.75	3.45	23.00	1.08
P450_26	13.44	6.69	4.75	6.85	6.59	2.01
P450_27	189.57	214.52	7.38	24.15	84.24	0.88

Table 6.1 Cont

FPKM						
Olfaction Gene	antenna female	antenna male	whole female	whole male	larva	Fold Difference Fa/Ma
P450_28	0.43	0.16	0.05	0.00	12.11	2.62
P450_30	92.87	57.89	19.04	35.90	50.85	1.60
P450_31	0.03	0.06	0.00	0.03	0.22	-2.04

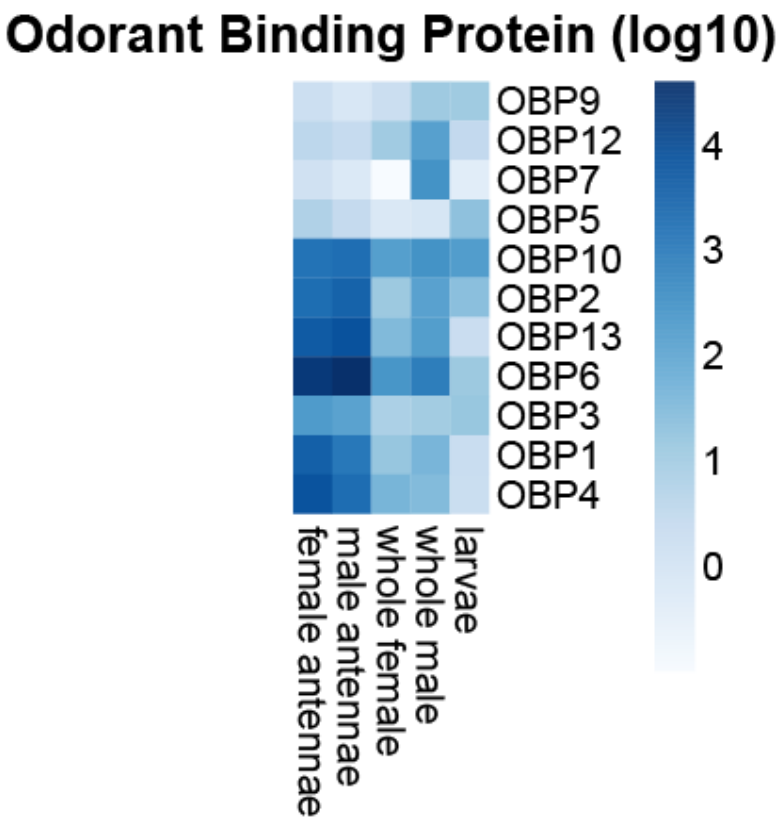


Figure 6.1. Expression levels of odorant binding protein genes in five wheat stem sawfly tissues, quantified by RNAseq. Expression levels in the heat map plot are reported as log10 values of the FPKM scores listed in Table 6.1. Increasing values are represented by darker shades of blue.

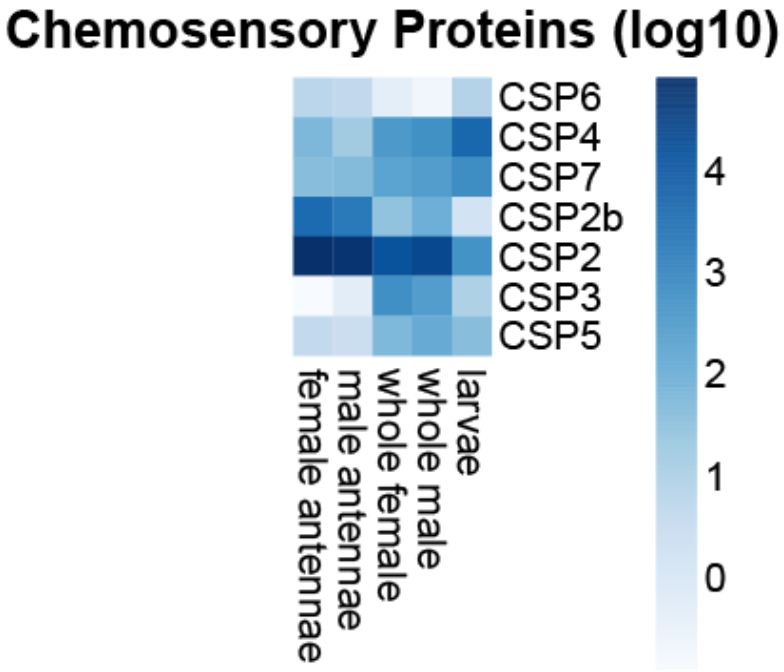


Figure 6.2. Expression levels of chemosensory protein genes in five wheat stem sawfly tissues, quantified by RNAseq. Expression levels in the heat map plot are reported as log10 values of the FPKM scores listed in Table 6.1. Increasing values are represented by darker shades of blue.

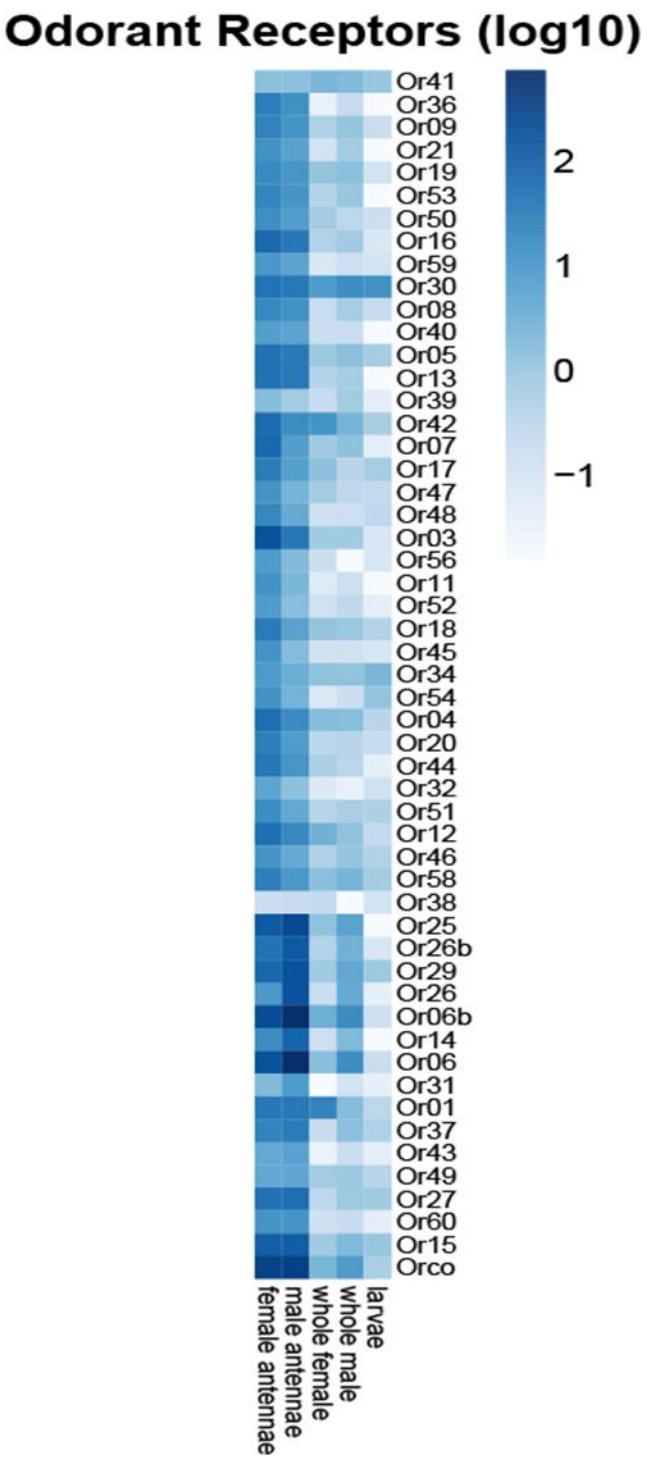


Figure 6.3. Expression levels of odorant receptor genes in five wheat stem sawfly tissues, quantified by RNAseq. Expression levels in the heat map plot are reported as log10 values of the FPKM scores listed in Table 6.1. Increasing values are represented by darker shades of blue.

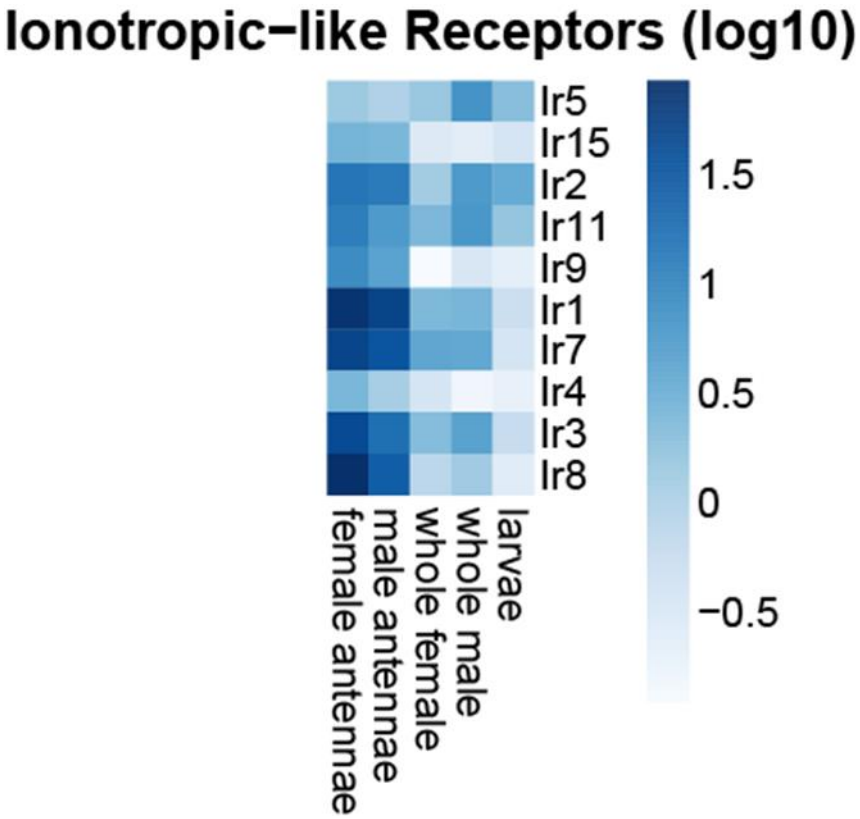


Figure 6.4. Expression levels of ionotropic receptor genes in five wheat stem sawfly tissues, quantified by RNAseq. Expression levels in the heat map plot are reported as log10 values of the FPKM scores listed in Table 6.1. Increasing values are represented by darker shades of blue.

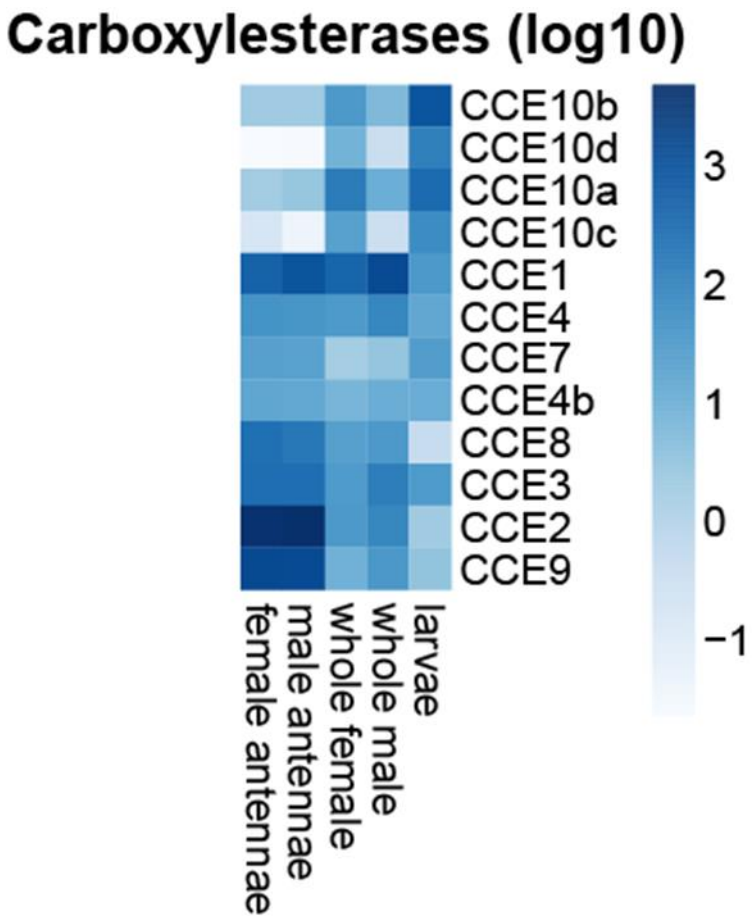


Figure 6.5. Expression levels of carboxylesterase genes in five wheat stem sawfly tissues, quantified by RNAseq. Expression levels in the heat map plot are reported as log10 values of the FPKM scores listed in Table 6.1. Increasing values are represented by darker shades of blue.

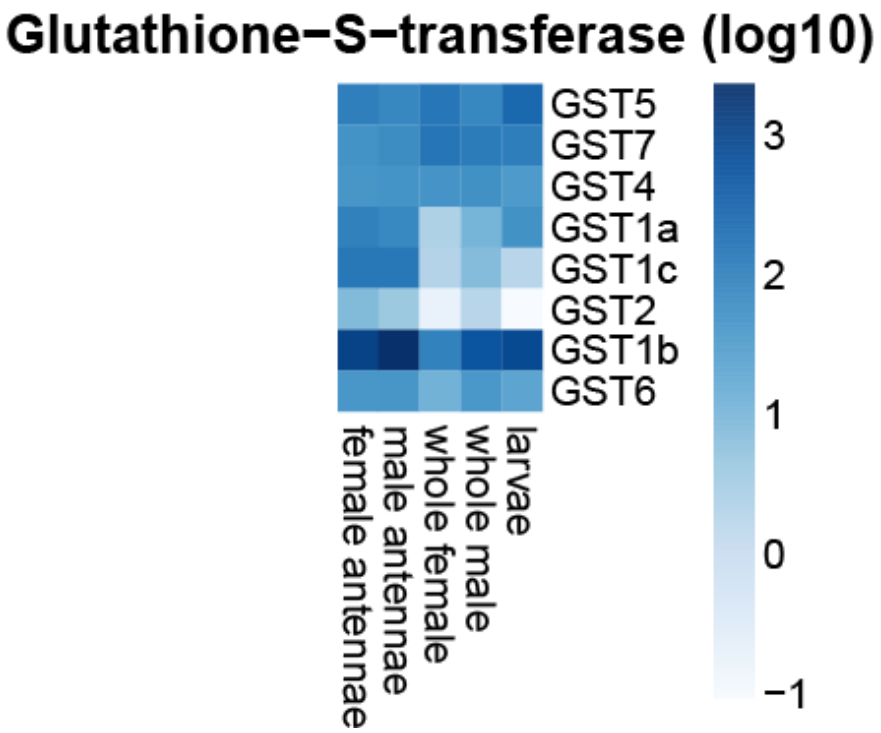


Figure 6.6. Expression levels of glutathione S-transferase genes in five wheat stem sawfly tissues, quantified by RNAseq. Expression levels in the heat map plot are reported as log10 values of the FPKM scores listed in Table 6.1. Increasing values are represented by darker shades of blue.

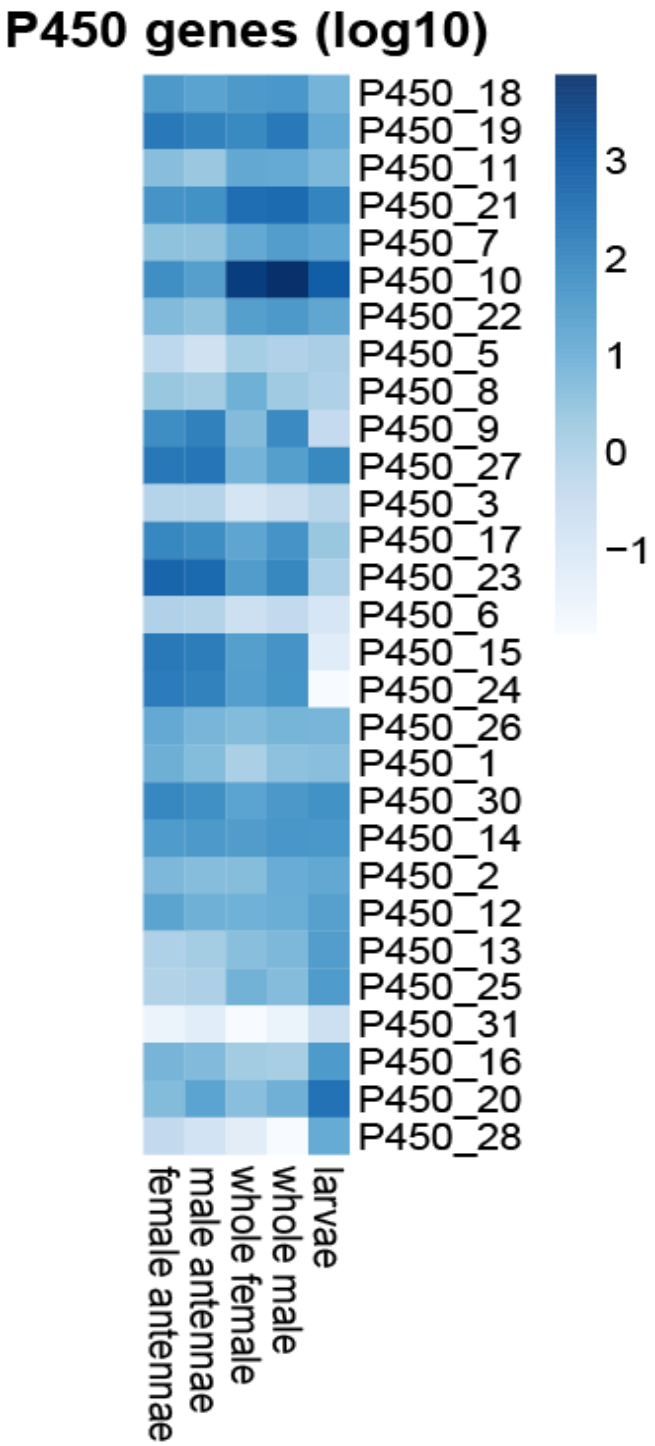


Figure 6.7. Expression levels of cytochrome p450 genes in five wheat stem sawfly tissues, quantified by RNAseq. Expression levels in the heat map plot are reported as log10 values of the FPKM scores listed in Table 6.1. Increasing values are represented by darker shades of blue.

References

Abuin, L., Bargeton, B., Ulbrich, M. H., Isacoff, E. Y., Kellenberger, S., & Benton, R. (2011). Functional architecture of olfactory ionotropic glutamate receptors. *Neuron*, *69*(1), 44-60.

Ai, M., Blais, S., Park, J. Y., Min, S., Neubert, T. A., & Suh, G. S. (2013). Ionotropic glutamate receptors IR64a and IR8a form a functional odorant receptor complex in vivo in *Drosophila*. *The Journal of Neuroscience*, *33*(26), 10741-10749

Beckage, N. E., & Gelman, D. B. (2004). WASP PARASITOID DISRUPTION OF HOST DEVELOPMENT: Implications for New Biologically Based Strategies for Insect Control*. *Annual Reviews in Entomology*, *49*(1), 299-330.

Bengtsson JM, Trona F, Montagne N, Anfora G, Ignell R, et al. (2012) Putative chemosensory receptors of the codling moth, *Cydia pomonella*, identified by antennal transcriptome analysis. PLoS One 7: e31620

Benton, R., Vannice, K.S., Gomez-Diaz, C., Vosshall, L.B. (2009) Variant ionotropic glutamate receptors as chemosensory receptors in *Drosophila*. *Cell* **136**(1): 149–162.

Bloom J, Khan Z, Kruglyak L, Singh M, Caudy A (2009) Measuring differential gene expression by short read sequencing: quantitative comparison to 2-channel gene expression microarrays. BMC Genomics 10:221

Carino, F. A., Koener, J. F., Plapp Jr, F. W., & Feyereisen, R. (1994). Constitutive overexpression of the cytochrome P450 gene CYP6A1 in a house fly strain with metabolic resistance to insecticides. *Insect biochemistry and molecular biology*, *24*(4), 411-418.

Carthew, R. W., & Sontheimer, E. J. (2009). Origins and mechanisms of miRNAs and siRNAs. *Cell*, *136*(4), 642-655.

Claudianos, C., Ranson, H., Johnson, R. M., Biswas, S., Schuler, M. A., Berenbaum, M.R., Feyereisen, R., and Oakeshott, J.G. (2006) A deficit of detoxification enzymes: pesticide sensitivity and environmental response in the honeybee. *Insect Mol Biol* **15**: 615–636.

Core L.J., Waterfall J.J., Lis J.T. (2008) Nascent RNA sequencing reveals widespread pausing and divergent initiation at human promoters. Science 322:1845–1848

Croset, V., Ritz, R., Cummins, S. F., Budd, A., Brawand, D., Kaessmann, H., Gibson, T. J. and Benton, R. (2010) Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet* **6**:e1001064.

Daborn, P. J., Yen, J. L., Bogwitz, M. R., Le Goff, G., Feil, E., Jeffers, S., et al. (2002). A single P450 allele associated with insecticide resistance in *Drosophila*. *Science*, *297*(5590), 2253-2256.

Danielson, P. B., MacIntyre, R. J., & Fogleman, J. C. (1997). Molecular cloning of a family of xenobiotic-inducible drosophilid cytochrome P450s: evidence for involvement in host-plant allelochemical resistance. *Proceedings of the National Academy of Sciences*, *94*(20), 10797-10802.

De Graaf, D. C., Aerts, M., Brunain, M., Desjardins, C. A., Jacobs, F. J., Werren, J. H., & Devreese, B. (2010). Insights into the venom composition of the ectoparasitoid wasp *Nasonia vitripennis* from bioinformatic and proteomic studies. *Insect molecular biology*, *19*(s1), 11-26.

Feyereisen, R. (2005). Insect cytochrome P450. *Comprehensive molecular insect science*, *4*, 1-77.

Gilbert, L. I. (2004). Halloween genes encode P450 enzymes that mediate steroid hormone biosynthesis in *Drosophila melanogaster*. *Molecular and cellular endocrinology*, *215*(1), 1-10.

Gong, D. P., Zhang, H. J., Zhao, P., Lin, Y., Xia, Q. Y., & Xiang, Z. H. (2007). Identification and expression pattern of the chemosensory protein gene family in the silkworm, *Bombyx mori*. *Insect biochemistry and molecular biology*, *37*(3), 266-277.

Gong, Z. J., Zhou, W. W., Yu, H. Z., Mao, C. G., Zhang, C. X., Cheng, J. A., & Zhu, Z. R. (2009). Cloning, expression and functional analysis of a general odorant-binding protein 2 gene of the rice striped stem borer, *Chilo suppressalis* (Walker) (Lepidoptera: Pyralidae). *Insect molecular biology*, *18*(3), 405-417.

Gress, J. C., Robertson, H. M., Weaver, D. K., Dlakić, M., and Wanner, K. W. 2013. Odorant receptors of a primitive hymenopteran pest, the wheat stem sawfly. *Insect Molecular Biology* 22(6): 659-667.

Grosse-Wilde, E., Kuebler, L. S., Bucks, S., Vogel, H., Wicher, D., & Hansson, B. S. (2011). Antennal transcriptome of *Manduca sexta*. *Proceedings of the National Academy of Sciences*, *108*(18), 7449-7454.

Hahn DA, Ragland GJ, Shoemaker DD, Denlinger DL (2009) Gene discovery using massively parallel pyrosequencing to develop ESTs for the flesh fly *Sarcophaga crassipalpis*. *BMC Genomics* 10:234

Hartley, C. J., Newcomb, R. D., Russell, R. J., Yong, C. G., Stevens, J. R., Yeates, D. K., et al. (2006). Amplification of DNA from preserved specimens shows blowflies were preadapted for the rapid evolution of insecticide resistance. *Proc. Natl. Acad. Sci. USA* 103(23): 8757-8762.

He, Y., Vogelstein, B., Velculescu, V.E., Papadopoulos, N., Kinzler, K.W. (2008) The antisense transcriptomes of human cells. *Science* 322:1855–1857

Hillier, L.W., Reinke, V., Green, P., Hirst, M., Marra, M.A., Waterston, R.H. (2009) Massively parallel sequencing of the polyadenylated transcriptome of *C. elegans*. *Genome Res* 19:657–666

Holt, C., & Yandell, M. (2011). MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC bioinformatics* 12(1): 491.

Huang, H. S., Hu, N. T., Yao, Y. E., Wu, C. Y., Chiang, S. W., & Sun, C. N. (1998). Molecular cloning and heterologous expression of a glutathione S-transferase involved in insecticide resistance from the diamondback moth, *Plutella xylostella*. *Insect Biochemistry and Molecular Biology*, 28(9): 651-658.

Ishida, Y., & Leal, W. S. (2002). Cloning of putative odorant-degrading enzyme and integumental esterase cDNAs from the wild silkmoth, *Antheraea polyphemus*. *Insect biochemistry and molecular biology*, *32*(12): 1775-1780.

Ishida, Y. and Leal, W. S. 2005. Rapid inactivation of a moth pheromone. *Proc. Natl. Acad. Sci. USA* 102: 14075–14079.

Jacquin-Joly, E., Vogt, R. G., Francois, M. C., and Nagnan-Le Meillour, P. (2001) Functional and expression pattern analysis of chemosensory proteins expressed in antennae and pheromonal gland of *Mamestra brassicae. Chem Senses* 26: 833-844.

Jordan, M.D., Anderson, A., Begum, D., Carraher, C., Authier, A., Marshall, S.D. et al. (2009) Odorant receptors from the light brown apple moth (*Epiphyas postvittana*) recognize important volatile compounds produced by plants. *Chem Senses* 34: 383–394.

Korsching, S. (2002). Olfactory maps and odor images. *Current opinion in neurobiology*, *12*(4), 387-392.

Leal, W. S. (2013) Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol*. 58: 373-391.

Legeai, F., Malpel, S., Montagné, N., Monsempes, C., Cousserans, F., Merlin, C., et al. (2011). An Expressed Sequence Tag collection from the male antennae of the Noctuid moth *Spodoptera littoralis*: a resource for olfactory and pheromone detection research. *BMC genomics*, 12(1): 86.

Li, X., Schuler, M. A., & Berenbaum, M. R. (2007). Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. *Annu. Rev. Entomol.*, 52: 231-253.

Liu, N., & Scott, J. G. (1998). Increased transcription of CYP6D1 causes cytochrome P450-mediated insecticide resistance in house fly. *Insect biochemistry and molecular biology*, *28*(8): 531-535.

Liu, N. and Zhang, L. (2004) CYP4AB1, CYP4AB2, and Gp-9 gene overexpression associated with workers of the red imported fire ant, *Solenopsis invicta* Buren. *Gene* 327: 81–87.

Lumjuan, N., McCarroll, L., Prapanthadara, L. A., Hemingway, J., & Ranson, H. (2005). Elevated activity of an Epsilon class glutathione transferase confers DDT resistance in the dengue vector, *Aedes aegypti*. *Insect biochemistry and molecular biology*, *35*(8), 861-871.

Lundberg, M., Boss, J., Canbäck, B., Liedvogel, M., Larson, K. W., Grahn, M., et al. (2013). Characterisation of a transcriptome to find sequence differences between two differentially migrating subspecies of the willow warbler *Phylloscopus trochilus*. *BMC genomics*, 14(1): 330.

Maibeche-Coisne, M., Jacquin-Joly, E., Francois, M.C. and Nagnan-Le Meillour, P. (2002) cDNA cloning of biotranformation enzymes belonging to the cytochrome P450 family in the antennae of the noctuid moth *Mamestra brassicae*. *Insect Mol Biol* 11: 273–281.

Maibeche-Coisne, M., Merlin, C., Francois, M.C., Porcheron, P. and Jacquin-Joly, E. (2005) P450 and P450 reductase cDNAs from the moth *Mamestra brassicae*: cloning and expression patterns in male antennae. *Gene* 346: 195–203.

Mao, W., Rupasinghe, S., Zangerl, A.R., Schuler, M.A. *and* Berenbaum, M.R. (2006*)* Remarkable substrate specificity ofCYP6AB3 in *Depressaria pastinacella*, a highly specialised caterpillar. *Insect Mol Biol* 15*:* 169–179.

Mao, W., Schuler, M.A. *and* Berenbaum, M.R. *(*2007*)* Cytochrome P450s in *Papilio multicaudatus* and the transition from oligophagy in the Papilionidae. *Insect Mol Biol* 16*:* 481–490.

Marguerat, S., & Bähler, J. (2010). RNA-seq: from technology to biology. *Cell Mol life Sciences*, *67*(4), 569-579.

Marioni, J.C., Mason, C.E., Mane, S.M., Stephens, M., Gilad, Y. (2008) RNA-seq: an assessment of technical reproducibility and comparison with gene expression arrays. *Genome Res* 18:1509–1517

Mittapalli, O., Neal, J. J., & Shukle, R. H. (2005). Differential expression of two cytochrome P450 genes in compatible and incompatible Hessian fly/wheat interactions. *Insect Mol Biol* 35(9): 981-989.

Mitchell, S. N., Stevenson, B. J., Müller, P., Wilding, C. S., Egyir-Yawson, A., Field, S. G. et al. (2012). Identification and validation of a gene causing cross-resistance between insecticide classes in *Anopheles gambiae* from Ghana. *Proc. Natl. Acad. Sci. USA* 109(16):, 6147-6152.

Müller, P., Warr, E., Stevenson, B. J., Pignatelli, P. M., Morgan, J. C., Steven, A., et al. (2008). Field-caught permethrin-resistant Anopheles gambiae overexpress CYP6P3, a P450 that metabolises pyrethroids. *PLoS Genetics*, 4(11): e1000286.

Naqvi, A.R., Islam, M.N., Choudhury, N.R., Haq, Q.M.R. (2009) The fascinating world of RNA interference. *Int J Biol Sci* 5:97–117

Nomura Kitabayashi, A., Arai, T., Kubo, T., & Natori, S. (1998). Molecular cloning of cDNA for p10, a novel protein that increases in the regenerating legs of *Periplaneta Americana* (American cockroach). *Insect Mol Biol* 28(10): 785-790.

Oakeshott J. G., Devonshire, A. L., Claudianos, C., Ollis, D. L., and Russel, R. J. (2005) Comparing the organophosphorus and carbamate insecticide resistance mutations in cholin- and carboxyl-esterases. *Chemico Biol Interact* 157–158:269–275.

Oakeshott, J. G., Johnson, R. M., Berenbaum, M. R., Ranson, H., Cristino, A. S., and Claudianos, C. (2010) Metabolic enzymes associated with xenobiotic and chemosensory responses in *Nasonia vitripennis*. *Insect Mol Biol* 19:147–163.

Pelosi, P., Zhou, J. J., Ban, L. P., & Calvello, M. (2006). Soluble proteins in insect chemical communication. *Cellular and Molecular Life Sciences CMLS*, 63(14): 1658-1676.

Pitts, R. J., Rinker, D. C., Jones, P. L., Rokas, A., & Zwiebel, L. J. (2011). Transcriptome profiling of chemosensory appendages in the malaria vector *Anopheles gambiae* reveals tissue-and sex-specific signatures of odor coding. *BMC genomics*, 12(1): 271.

Poupardin, R., Reynaud, S., Strode, C., Ranson, H., Vontas, J., & David, J. P. (2008). Cross-induction of detoxification genes by environmental xenobiotics and insecticides in the mosquito *Aedes aegypti*: Impact on larval tolerance to chemical insecticides. *Insect Biochem Mol Bio*, 38(5):, 540-551.

Ranson, H., Claudianos, C., Ortelli, F., Abgrall, C., Hemingway, J., Sharakhova, M. V. et al. (2002). Evolution of supergene families associated with insecticide resistance. *Science*, *298*(5591), 179-181.

Ranson, H., and Hemingway, J. (2005) Mosquito glutathione transferases. *Rev Methods Enzymol* 401: 226-241.

Ranson, H., Rossiter, L., Ortelli, F., Jensen, B., Wang, X., Roth, C., et al. (2001). Identification of a novel class of insect glutathione s-transferases involved in resistance to DDT in the malaria vector *Anopheles gambiae*. *Biochem. J 359*: 295-304.

Robertson, H. M., Gadau, J., and Wanner, K. W. 2010. The insect chemoreceptor superfamily of the parastiod jewel wasp *Nasonia vitripennis*. *Insect Mol. Biol*. 19: 121–136.

Robertson, H. M. and Wanner, K. W. 2006. The chemoreceptor superfamily in the honey bee, *Apis mellifera*: expansion of the odorant, but not gustatory, receptor family. *Genome Res* 16: 1395–1403.

Rogers, M. E., Jani, M. K., and Vogt, R. G. (1999) An olfactory-specific glutathione-S-transferase in the sphinx moth *Manduca sexta*. *J Exp Biol* 202(12): 1625-1637.

Rupasinghe, S. G., Wen, Z., Chiu, T. L., & Schuler, M. A. (2007). *Helicoverpa zea* CYP6B8 and CYP321A1: different molecular solutions to the problem of metabolizing plant toxins and insecticides. *Protein Engineering Design and Selection*, *20*(12), 615-624.

Rutzler, M. and Zwiebel, L.J. (2005) Molecular biology of insect olfaction: recent progress and conceptual models. *J Com Physiol A Neuroethol Sens Neural Behav Physiol* 191: 777–790.

Sato, K., Pellegrino, M., Nakagawa, T., Nakagawa, T., Vosshall, L. B., and Touhara, K. (2008) Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* 452: 1002–1006.

Scharf, M.E., Parimi, S., Meinke, L.J., Chandler, L.D. and Siegfried, B.D. (2001) Expression and induction of three family 4 cytochrome P450 (CYP4) genes identified from insecticide-resistant and susceptible western corn rootworms, *Diabrotica virgifera* virgifera. *Insect Mol Biol* 10: 139–146.

Schena, M., Heller, R. A., Theriault, T. P., Konrad, K., Lachenmeier, E., & Davis, R. W. (1998). Microarrays: biotechnology's discovery platform for functional genomics. *Trends in biotechnology*, 16(7): 301-306.

Senthilan, P. R., Piepenbrock, D., Ovezmyradov, G., Nadrowski, B., Bechstedt, S., Pauls, et al. (2012). *Drosophila* Auditory Organ Genes and Genetic Hearing Defects. *Cell*, 150(5): 1042-1054.

Shalon, D., Smith, S. J., & Brown, P. O. (1996). A DNA microarray system for analyzing complex DNA samples using two-color fluorescent probe hybridization. *Genome Res* 6(7): 639-645.

Simpson, A. E. C. M. (1997) The cytochrome P450 4 (CYP4) family. *Gen Pharmacol* 28: 351–359.

Stevens, J. L., Snyder, M. J., Koener, J. F., & Feyereisen, R. (2000). Inducible P450s of the CYP9 family from larval *Manduca sexta* midgut. *Insect Mol. Biol* 30(7): 559-568

Sutherland, T. D., Unnithan, G. C., Andersen, J. F., Evans, P. H., Muralaliev, M. B. and Szabo, L. Z. (1998) A cytochrome P450 terpenoid hydroxylase linked to the suppression of juvenile hormone synthesis. *Proc Natl Acad Sci USA* 95: 12884–12889.

‘t Hoen, P. A. C., Ariyurek ,Y., Thygesen, H. H., Vreugdenhil, E., Vossen, R. H. A. M., de Menezes, R. X., Boer, J. M., van Ommen, G. B., den Dunnen, J. T. (2008) Deep sequencing-based expression analysis shows major advances in robustness, resolution and inter-lab portability over five microarray platforms. *Nucleic Acids Res* 36:e141

Tang, A. H., & Tu, C. P. (1994). Biochemical characterization of *Drosophila* glutathione S-transferases D1 and D21. *J of Biolog Chem* 269(45): 27876-27884.

Trick, M., Long, Y., Meng, J., Bancroft, I. (2009) Single nucleotide polymorphism (SNP) discovery in the polyploid *Brassica napus* using Solexa transcriptome sequencing. *Plant Biotechnol J* 7:334–346

Vera, J. C., Wheat, C. W., Fescemyer, H. W., Frilander, M. J., Crawford, D. L., Hanski, I., Marden, J. H. (2008) Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. *Mol Ecol* 17:1636–1647

Wahl, D. V., Shanower, T. G., and Hoelmer, K. A. (2007) A new species of *Collyria* Schiødte (Hymenoptera: Ichneumonidae: Collyriinae), a parasitoid of *Cephus fumipennis* (Hymenoptera: Cephidae) in China, and potential biological control agent for *Cephus cinctus* in North America. *J. Kans. Entomol. Soc*. 80: 43–50.

Wang, E. T., Sandberg, R., Luo, S., Khrebtukova, I., Zhang, L., Mayr, C., Kingsmore, S. F., Schroth, G. P., Burge, C. B. (2008) Alternative isoform regulation in human tissue transcriptomes. *Nature* 456:470–476

Wanner, K.W., Anderson, A.R., Trowell, S.C., Theilman, D.A., Robertson, H.M. and Newcomb, R.D. (2007) Female-biased expression of odourant receptor genes in the adult antennae of the silkworm, *Bombyx mori*. *Insect Mol Biol* 16: 107–119.

Wei, S. H., Clark, A. G., & Syvanen, M. (2001). Identification and cloning of a key insecticide-metabolizing glutathione S-transferase (MdGST-6A) from a hyper insecticide-resistant strain of the housefly *Musca domestica*. *Insect biochem mol biol* 31(12): 1145-1153.

Wen, Z., Baudry, J., Berenbaum, M.R. *and* Schuler, M.A. (2005) Ile115Leu mutation in the SRS1 region of an insect cytochrome P450 (CYP6B1) compromises substrate turnover via changes in a predicted product release channel. *Protein Eng Des Sel* 18*:* 191–199.

Wicher, D., Schafer, R., Bauernfeind, R., Stensmyr, M. C., Heller, R., Heinemann, S. H., and Hansson, B. S. (2008) *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. *Nature* 452: 1007–1011.

Xu, Y. L., He, P., Zhang, L., Fang, S. Q., Dong, S. L., Zhang, Y. J., & Li, F. (2009). Large-scale identification of odorant-binding proteins and chemosensory proteins from expressed sequence tags in insects. *BMC genomics*, 10(1): 632.

Zhang, Y. V., Raghuwanshi, R. P., Shen, W. L., & Montell, C. (2013). Food experience-induced taste desensitization modulated by the *Drosophila* TRPL channel. *Nature neuroscience* 16: 1468-1476

CHAPTER SEVEN

GENERAL CONCLUSIONS

The wheat stem sawfly (WSS) has been a major pest of wheat in Montana and the Northern Great Plains of North America for more than a century. Economic impacts of the wheat stem sawfly date back to the early settlement of the Prairies where severe infestation and damage was recorded as early as 1922 (Criddle, 1922). The cryptic location of the larval stage that lives within the wheat stem has rendered control by insecticides ineffective. Similarly, the emergence and flight period of the adult stage is too protracted for effective control using contact insecticide sprays. As a result, extensive crop damage continues to persist and monetary losses of more than $100 million per year in the U.S. and Canada continue to occur (Wahl et al., 2007). Tolerant wheat varieties with solid stems represent the most common pest management strategy that is applied to control damage, but control and yield are variable. Producers in Montana and the region need new pest-control strategies to manage damage caused by WSS. Modern insect genomics provide new genetic targets for novel pest control strategies.

Recent advances in the molecular understanding of insect olfaction have prompted olfactory-related genes as novel targets to disrupt pest behaviors mediated by olfaction. An original objective of this thesis was to identifty candidate odorant receptors (ORs) from the antennal transcriptome using “454” pyrosequencing technology (Chapter Three). At the time, 454 represented the state-of-the-art DNA sequencing technology. However, the pace of DNA sequencing technology has advanced so quickly during the

last five years that Illumina sequencing technology took over as state-of-the-art technology while this thesis research was ongoing. Illumina sequencing technology was applied to the WSS genome and antennal transcriptome and enabled the objectives of this dissertation to be expanded to include all olfactory-related gene families and their expression in the antennae.

Four protein families are involved in detecting odors in the peripheral olfactory system of insects: odorant-binding proteins (OBPs), chemosensory protein (CSPs), odorant receptors (Ors) and ionotropic receptors (IRs). Three protein families are involved in degrading odor signals: carboxyl esterases (CCEs), glutathione S-transferases (GSTs), and cytochrome P450s (P450s). A total of 131 candidate transcripts were identified from the antennal transcriptome and used along with the whole geneome sequence to generate high quality gene model annotations (Chapter Four). The predicted protein sequencs were analyzed bioinformatically to confirm their protein family identity (Chapter Five). Finally, RNA-seq. data from five different WSS tissues were mapped to the high quality gene models to assay tissue-specific gene expression based on number of sequence reads per gene length. Collectively, these data were used to prioritize gene candiates based on expression levels in the male and female antennae providing a list for future functional studies.

Future research needs to link olfactory genes with specific odors that mediate important insect behaviors. One approach is to identify the ligands of olfactory proteins to screen for potential attractants or repellents that can be tested in behavioral assays, an approach that has been termed “reverse chemical ecology.” For example, CcOr26 is

highly expressed and significantly biased towards expression in male antennae. This expression pattern is typical of an OR that detects odors important to male, but not female, insects, such as female-produced sex pheromones for mating. The *Xenopus* oocyte system can be used to functionally characterize ORs, identifying odors that specifically activate the receptor (Stortkuhl & Kettler, 2001; Wetzel et al., 2001). If previously published WSS pheromone components do not activate CcOr26, then deorphanizing this receptor may lead to the identification of new attractants.

REFERENCES CITED

Abeel, T., Van Parys, T., Saeys, Y., Galagan, J., and Van de Peer, Y. 2012. GenomeView: a next-generation genome browser. *Nucleic Acids Res* 40(2): e12. doi: 10.1093/nar/gkr995.

Ai, M., Min, S., Grosjean, Y., Leblanc, C., Bell, R., Benton, R., and Suh, G.S. 2010. Acid sensing by the *Drosophila* olfactory system. *Nature* 468: 691–695.

Ainslie, C.N. 1920. The western grass-stem sawfly. *USDA. Bull. No.* 841.

Ainslie, C.N. 1929. The Western Grass-Stem Sawfly: A Pest of Small Grains. *U.S. Dept of Agric Techn Bul No* 57: 1–24.

Ainslie, C.N. 1929. The western grass-stem sawfly - a pest of small grains. *USDA. Bull. No*. 157.

Altner, H. 1977. Insect sensillum specificity and structure: an approach to new typology. *Olfaction and Taste* 6: 295–303.

Altner, I., Hatt, H. and Altner, H. 1983. Structural properties of bimodal chemo- and mechanosensitive setae on the pereiopod chelae of the crayfish, *Austropotamobius torrentium*. *Cell Tissue Res* 242: 357–374.

Altner, H. and. Prillinger, L. 1980. Ultrastructure of invertebrate chemo-, thermo-, and hygroreceptors and its functional significance. *Intl Rev Cytol* 67: 69–139.

Altschul, S. F., Madden, T. L., Schaffer, A. A., Zhang, J., Zhang, Z., Miller, W. and Lipman. D. J. 1997. Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25: 3389–3402.

Anderson, A. R., Wanner, K. W., Trowell, S. C., Warr, C. G., Jaquin-Joly, E., Zagatti, P., Robertson, H. and Newcomb, R. D. 2009. Molecular basis of female-specific odorant responses in *Bombyx mori*. *Insect Biochem. Mol. Biol*. 39: 189–197.

Angeli, S., Ceron, F., Scaloni, A., Monti, M., Monteforti, G., Minnocci, A., Petacchi, R. and Pelosi, P. 1999. Purification, structural characterization, cloning and immunocytochemical localization of chemoreception proteins from *Schistocerca gregaria*. *Eur. J. Biochem* 262: 745–754.

Baldwin, I. T. and Schultz, J. C. 1983. Rapid changes in tree leaf chemistry induced by damage: evidence for communication between plants. *Science* 221: 277–279

Bartlet, R. J., Cossé, A. A., Petroski, R. J. and Weaver, D. K. 2002. Cuticular hydrocarbons and novel alkenediol diacetates from wheat stem sawfly (*Cephus cinctus):* natural oxidation to pheromone components. *J Chem Ecol* 28: 385-405.

Baum, J. A., Bogaert, T., Clinton, W., Heck, G. R., Feldmann, P., Ilagan, O., et al. 2007. Control of coleopteran insect pests through RNA interference. *Nature biotechnology* 25: 1322–1326.

Benton, R., Sachse, S., Michnick, S. W. and. Vosshall, L. B. 2006. Atypical membrane topology and heteromeric function of *Drosophila* odorant receptors in vivo. *PLoS Biol* 4: e20.

Benton, R., Vannice, K. S., Gomez-Diaz, C., Vosshall, L. B. 2009. Variant ionotropic glutamate receptors as chemosensory receptors in Drosophila. *Cell* 136(1): 149–162.

Beres, B. L., Cárcamo, H. A. and Byers, J. R. 2007. Effect of wheat stem sawfly damage on yield and quality of selected Canadian spring wheat. *J Econ Entomol* 100: 79–87.

Beres, B. L., Dosdall, L. M., Weaver, D. K. Cárcamo, H. A. and Spaner, D. M. 2011. Biology and integrated management of wheat stem sawfly, *Cephus cinctus* (Hymenoptera: Cephidae), and the need for continuing research. *Can Entomol* 143: 105–125.

Berg, J. E., Bruckner, P. L., Carlson, G. R., Dyer, A., Eckoff, J., Kushnak, G. D., et al. 2006. Winter wheat variety performance summary in Montana. *MAES*

Blodgett, S. L., Goosey, H. B., Waters, D. K., Tharp, C. I. and Johnson, G. D. 1996. Wheat stem sawfly control on winter wheat. *Arthropod Manag Tests* 22: 331-332.

Bogwitz, M. R., Chung, H., Magoc, L., Rigby, S., Wong, W., O'Keefe, M., et al. 2005. Cyp12a4 confers lufenuron resistance in a natural population of *Drosophila melanogaster*. *Proc Natl Acad Sci U.S.A.* 102(36): 12807-12812.

Bohbot, J., Sobrio, F., Lucas, P., and Nagnan-Le Meillour, P. 1998. Functional characterization of a new class of Odarant Binding Proteins in the moth *Mamestra bassicae. Biochem Biophys Res Commun* 253*:* 489-494.

Board, P. G., Baker, R. T., Chelvanayagam, G. and Jermiin, L. S. 1997. Zeta, a novel class of glutathione transferases in a range of species from plants to humans. *Biochem J* 328: 929-935.

Bradshaw, J. 2013. Adult Wheat Stem Sawflies Have Begun to Emerge. University of Nebraska *Cropwatch*. 30 May 2013.

Brandt, A., Scharf, M., Pedra, J. H., Holmes, G., Dean, A., Kreitman, M., and Pittendrigh, B. R. 2002. Differential expression and induction of two *Drosophila* cytochrome P450 genes near the Rst(2)DDT locus. *Insect Mol Biol* 11: 337–341.

Breer, H. 1997. Molecular mechanisms of pheromone reception in insect antennae. In Insect Pheromone Research New Directions (ed. R. T. Cardé and A. K. Minks), pp. 115–130. New York: Chapman & Hall.

Briand, L., Swasdipan, N., Nespoulous, C., Bezirard, V., Blon, F., Huet, J. C., Ebert, P. and Penollet, J. C. 2002. Characterization of a chemosensory protein (ASP3c) from honeybee (*Apis mellifera L*.) as a brood pheromone carrier. *Eur J Biochem* 269: 4586–4596.

Bruce, T. J. A. and Pickett, J. A. 2011. Perception of plant volatile blends by herbivorous insects - finding the right mix. *Phytochemistry* 72: 1605–1611.

Bruce, T. J. A., Wadhams, L. J. and Woodcock, C. M. 2005. Insect host location: a volatile situation. *Trends Plant Sci* 10: 269–274.

Buteler, M., Weaver, D. K. and Peterson, R. K. D. 2009. Exploring the oviposition behavior of the wheat stem sawfly when encountering plants infested with cryptic conspecifics. *Environ Entomol* 38: 1707–1715.

Buttery, R. G., Xu, C. and Ling. L. C. 1985. Volatile components of wheat leaves (and stems): possible insect attractants. *J Agric Food Chem* 33: 115–117.

Callenbach, J. A. and. Hansmeier, M. P. 1944. Wheat stem sawfly control in severely infested areas. *Montana Extension Service in Agriculture and Home* Economics Circular No. 156.

Calvello, M., Brandazza, A., Navarrini, A., Dani, F. R., Turillazzi, S., et al. 2005. Expression of odorant-binding proteins and chemosensory proteins in some Hymenoptera. *Insect Biochem Mol Biol* 35: 297–307.

Calvello, M., Guerra, N., Brandazza, A., D'Ambrosio, C., Scaloni, A., Dani, F. R., Turillazzi, S., and Pelosi, P. 2003. Soluble proteins of chemical communication in the social wasp *Polistes dominulus*. *Cell Mol Life Sci* 60: 1933–1943.

Campanacci, V., Lartigue, A., Hallberg, B. M., Jones, T. A., Giudici-Orticoni, M. T., Tegoni, M., and Cambillau, C. 2003. Moth chemosensory protein exhibits drastic conformational changes and cooperativity on ligand binding. *Proc Natl Acad Sci USA* 100: 5069–5074.

Canterel, B. L., Korf, I., Robb, S. M., Parra, G., Ross, E., Moore, B., Holt, C., Sanchez Alvarado, A., and Yandall, M. 2008. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res* 18(1): 188-196.

Carey, A. F., Wang, G., Su, C-Y., Zwiebel, L. J. and Carlson, J. R. 2010. Odorant reception in the malaria mosquito *Anopheles gambiae*. *Nature* 464: 66–71.

Chevereux, B., Pfisterer, T., Drescher, B., Driesel, A. J., Müller, W. E. G., Wetter, T. and Suhai, S. 2004. Using the miraEST Assembler for Reliable and Automated mRNA Transcript Assembly and SNP Detection in Sequenced ESTs. *Genome Res* 14: 1147–1159.

Clamp, M., Cuff, J., Searle, S. M. and Barton, G. J. 2004. The Jalview Java alignment editor. *Bioinformatics* 20: 426-427. doi:10.1093/bioinformatics/btg430.

Claudianos, C., Ranson, H., Johnson, R. M., Biswas, S., Schuler, M. A., Berenbaum, M. R., Feyereisen, R., and Oakeshott, J. G. 2006. A deficit of detoxification enzymes: pesticide sensitivity and environmental response in the honeybee. *Insect Mol Biol* 15: 615–636.

Clyne, P. J., Warr, C. G., Freeman, M. R., Lessing, D., Kim, J. and Carlson, J. R. 1999. A novel family of divergent seven-transmembrane proteins: candidate odorant receptors in *Drosophila*. *Neuron* 22: 327–338.

Cochran, V., Danielson, J., Kolberg, R., and Miller, P. 2006. Dryland cropping in the Canadian prairies and the US Great Plains. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America. Dryland Agriculture, 2nd ed., Agronomy Monograph 23.

Cossé, A. A., Bartelt, R. J., Weaver, D. K., and Zilkowski, B. W. 2002. Pheromone Components of the Wheat Stem Sawfly: Identification, Electrophysiology, and Field Bioassay. J. Chem. Ecol. 28(2): 407-423.

Criddle, N. 1915. The Hessian fly and the western wheat stem sawfly in Manitoba, Saskatchewan and Alberta. Canadian Department of Agriculture, Entomology Branch Bulletin, 11: 1-23.

Criddle, N. 1922. The western wheat-stem sawfly and its Control. Dominion of Canada Department of Agriculture Pamphlet 6- New Series 3-8.

Criddle, N. 1923. The life habits of *Cephus Cinctus* Norton in Monitoba. *Canadian Entomol*. 55: 1-4.

Croset, V., Ritz, R., Cummins, S. F., Budd, A., Brawand, D., Kaessmann, H., Gibson, T. J., and Benton, R. 2010. Ancient protostome origin of chemosensory ionotropic glutamate receptors and the evolution of insect taste and olfaction. *PLoS Genet*. 6, e1001064. doi: 10.1371/journal.pgen.1001064.

Danielson, P. B., Macintyre, R. J., and Fogleman, J. C. 1997. Molecular cloning of a family of xenobiotic-inducible drosophilid cytochrome P450s: evidence for involvement in host-plant allelochemical resistance. *Proc Natl Acad Sci USA* 94:10797-10802.

Danty, E., Arnold, G., Huet, J. C., Huet, D., Masson, C., and Pernollet, J. C. 1998. Separation, characterization and sexual heterogeneity of multiple putative odorant-binding proteins in the honeybee *Apis mellifera L*. (Hymenoptera: Apidea). *Chem Senses* 23: 83–91.

Delaney, K. J., Weaver, D. K. and Peterson, R. K. D. 2010. Wheat photosynthesis and yield reductions: wheat stem sawfly (Hymenoptera: Cephidae) herbivory interacts with abiotic conditions and host plant resistance. *J Econ Entomol* 103: 516–524.

DeMoraes, C. M., Lewis, W. J., Paré, P. W., and Tumlinson, J. H. 1998. Herbivore infested plants selectively attract parasitoids. Nature 393: 570-574.

DePauw, R. M., Clarke, J. M., Knox, R. E., Fernandez, M. R., McCraig, T. N., and McLeod, J. G. 2000. AC Abbey hard red spring wheat. Canadian Journal of Plant Science 80: 123-127.

DePauw, R. M., McLeod, J. G., Clarke, J. M., McCaig, T. N., Fernandez, M. R., and Knox, R. E. 1994. AC Eatonia hard red spring wheat. Can. J. Plant Sci 74: 821–823.

DePauw, R. M., Townley-Smith, T. F., Humphreys, G., Knox, R. E., Clarke, F. R., and Clarke, J. M. 2005. Lillian hard red spring wheat. Canadian Journal of Plant Science 85: 397–401.

Dicke, M., Sabelis, M. W., Takabayashi, J., Bruin, J., and Posthumus, M. A. 1990. Plant strategies of manipulating predator-prey interactions through allelochemicals: prospects for application in pest control. J. Chem. Ecol. 16: 3091-3118

Dicke, M. and van Loon, J. J. A. 2000. Multitrophic effects of herbivore-induced plant volatiles in an evolutionary context. Entomol. Exp. Appl. 97: 237-249.

Drummond, A., Ho, S., Phillips, M., and Rambaut, A. 2006. Relaxed phylogenetics and dating with confidence. PLoS Biol, 4(5):e88.

Drummond, A. J., and Rambaut, A. 2007. Beast: Bayesian evolutionary analysis by sampling trees. BMC Evol Biol. 7:214.

Durand, N., Carot-Sans, G., Bozzolan, F., Rosell, G., Siaussat, D., Debernard, S., et al. 2011. Degradation of Pheromone and Plant Volatile Components by a Same Odorant-Degrading Enzyme in the Cotton Leafworm, Spodoptera littoralis. PLoS ONE 6(12): e29147.

Engsontia, P., Sanderson, A. P., Cobb, M., Walden, K. K. O., Robertson, H., and Brown, S. 2008. The red flour beetle's large nose: an expanded odorant receptor gene family in *Tribolium casteneum*. *Insect Biochem. Mol. Biol.* 38: 387–397.

Farmer, E. E. and Ryan, C. A. 1990. Interplant communication: airborne methyl jasmonate induces synthesis of proteinase inhibitors in plant leaves. Proceedings of the National Academy of Sciences of the USA. 87: 7713-7716.

Farstad, C. 1940. The development of western wheat stem sawfly (*Cephus cinctus* Nort.) in various host plants as an index of resistance [PhD dissertation]. Iowa State College, Ames, IA. P 83.

Farstad, C. W., Platt, A. W., and McGinnis, A. J. 1949. Influence of wheat varieties on the sex ratio of the wheat stem sawfly, *Cephus cinctus* Nort. (Hymenoptera: Cephidae). 80th Annual Report of the Entomological Society of Ontario.

Feng, H. L., Zhang, D. X., Chang, T. R., Guo, W. Y., Zhang, R. C., and Xin, Y. 1992. Scanning electron microscopy studies of antennal sensilla and ommateum eye of ant. Entomological Knowledge 29: 292–294.

Fletcher, J. 1904. Experimental farms reports - Report of the Entomologist and Botanist.Appendix to the Report of the Minister of Agriculture,Sessional Paper No. 16. pp. 172–173.

Forêt, S., and Maleszka, R. 2006. Function and evolution of a gene family encoding odorant binding-like proteins in a social insect, the honey bee (Apis mellifera). *Genome Res* 16: 1404–1413.

Forêt, S., Wanner, K. W., and Maleszka, R. 2007. Chemosensory proteins in the honey bee: insights from the annotated genome, comparative analyses and expressional profiling. *Insect Biochem. Mol. Biol*. 37: 19–28.

Fritzsche-Hoballah, M. E. and Turlings, T. C. J. 2001. Experimental evidence that plants under caterpillar attack may benefit from attracting parasitoids. Evolutionary Ecology Research 3: 553–565.

Fuentes-Contreras, E. and Niemeyer, H. M. 1998. Dimboa glucoside, a wheat chemical defense, affects host acceptance and suitability of *Sitobian avanae* to the cereal aphid parasitoid *Aphidius rhopalosiphi*. Journal of Chemical Ecology 24: 371-381.

Galindo, K., and Smith, D. P. 2001. A large family of divergent Drosophila odorant-binding proteins expressed in gustatory and olfactory sensilla. *Genetics* 159: 1059–1072.

Gao, Y., Luo, L. Z., and Hammond, A. 2007. Antennal morphology, structure and sensilla distribution in Microplitis pallidipes (Hymenoptera:Braconidae). Micron 38: 684–693

Gereau, R.W., and Swanson, G.T. 2008. The Glutamate Receptors. Humana Press, Totowa, N.

Gianoli, E. and Niemeyer, H. M. 1998. Allocation of herbivory-induced hydroxamic acids in the wild wheat Triticum uniaristatum. Chemoecology 8: 19-23.

Goosey, H. B. 1999. In field distributions of the wheat stem sawfly, (Hymenoptera: Cephidae), and evaluation of selected tactics for an integrated management program. [Master's Thesis]. Montana State University, Bozeman, Montana.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature biotechnology* 29: 644–652.

Gress, J. C., Robertson, H. M., Weaver, D. K., Dlakić, M., and Wanner, K. W. 2013. Odorant receptors of a primitive hymenopteran pest, the wheat stem sawfly. Insect Molecular Biology 22(6): 659-667.

Grimaldi, D., and Engel, M. S. 2005. Evolution of the Insects. Cambridge University Press, Cambridge,UK. xv+755 pp.

Grosjean, Y., Rytz, R., Farine, J. P., Abuin, L., Cortot, J., Jefferis, G. S., and Benton, R. 2011. An olfactory receptor for food-derived odours promotes male courtship in *Drosophila*. *Nature* 478: 236–240.

Guzov, V. M., Unnithan, G. C., Chernogolov, A. A., and Feyerseisen, R. 1998. CYP12A1, a mitochondrial cytochrome P450 from the house fly. *Arch Biochem Biophys* 359: 231–240.

Hallem, E. A., Ho, M. G., and Carlson, J. R. 2004. The molecular basis of odor coding in the Drosophila antenna. *Cell* 117:965–979.

Hamilton-Kemp, T. R. and Anderson, R. A. 1984. Volatile compounds from *Triticum aestivum*. Phytochemistry 23: 1176–1177 .

Hamilton-Kemp, T. R. and Anderson, R. A. 1986. Volatiles from winter wheat: identification of additional compounds and effects of tissue source. Phytochemistry 25**:** 241–243.

Hartel, K. D., Frederick, B. A., and Shanower, T. G. 2003. Isolation and characterization of microsatellite loci in wheat stem sawfly Cephus cinctus and cross-species amplification in related species. Mol. Ecol. Notes 3(1): 85-87.

Hartlieb, E**.** and Anderson, P. 1999. *Olfactory-released behaviours. In Hansson, B.S. (ed.),* Insect Olfaction. *Springer, Berlin, pp.*315*– 349.*

Hekmat-Scafe, D. S., Scafe, C. R., McKinney, A. J., and Tanouye, M. A. 2002. Genome-wide analysis of the odorant-binding protein gene family in Drosophila melanogaster. *Genome Res* 12: 1357– 1369.

Helvig, C., Koener, J. F., Unnithan, G. C., and Feyereisen, R. 2004. CYP15A1, the cytochrome P450 that catalyzes epoxidation of methyl farnesoate to juvenile hormone III in cockroach corpora allata. *Proc Natl Acad Sci USA* 101: 4024-4029.

Hilker, M. and Meiners, T. 2002. Induction of plant responses towards oviposition and feeding of herbivorous arthropods: a comparison. Entomol. Exp. Appl. 104: 181-192.

Hill, E. S., Iwano, M., Gatellier, L., and Kanzaki, R. 2002. Morphology and physiology of the serotonin-immunoreactive putative antennal lobe feedback neuron in the male silk moth Bombyx mori. Chem. Senses 27: 475-483.

Hofmann, K. and Stoffel, W. 1993. TMbase-A database of membrane spanning protein segments. *Biol Chem Hoppe-Seyler* 374: 166

Holmes, N. D. 1977. The effect of the wheat stem sawfly, *Cephus cinctus* (Hymenoptera: Cephidae), on the yield and quality of wheat. *Can. Entomol*. 109: 1591-1598.

Holmes, N. D. 1978. The wheat stem sawfly. Proc. 26th Annu. Mtg. Entomol. Soc. Alberta, pp. 2-13.

Holmes, N. D. 1982. Population dynamics of the wheat stem sawfly, *Cephus cinctus* (Hymenoptera Cephidae), in wheat. Can. Entomol. 114: 775-788.

Holmes, N.D. and Peterson, L.K. 1960. The influence of the host on oviposition by the wheat stem sawfly *Cephus cinctus* Nort. (Hymenoptera:Cephidae). *Can J Plant Sci* 40: 29–46.

Holmes, N. D. and Peterson, L. K. 1960. The influence of the host on oviposition by the wheat stem sawfly, *Cephus cinctus* Nort. (Hymenoptera: Cephidae) II. Resistance to the larva. *Can. Entomol*. 94: 348-365.

Holt, C., and Yandell, M. 2011. MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12:491. doi:10.1186/1471-2105-12-491.

Howse, P. E. 1998. Pheromones and Behaviour. Insect Pheromones and their Use in Pest Management. P. E. Howse, I. D. R. Stevens and O. T. Jones. London, Chapman & Hall.

Irell, B. and Peairs, F. 2011. Wheat Stem Sawfly: A New Pest of Colorado Wheat. Colorado State University Extension. Circular No. 5.612.

Ishida, Y., Chiang, V., and Leal, W. S. 2002. Protein that makes sense in the Argentine ant. *Naturwissenschaften* 89:505–507.

Ishida, Y. and Leal, W. S. 2005. Rapid inactivation of a moth pheromone. Proc. Natl. Acad. Sci. USA 102: 14075–14079.

Isidoro, N., Bin, F., Colazza, S., and Vinson, S. B. 1996. Morphology of antennal gustatory sensilla and glands in some parasitoids Hymenoptera with hypothesis on their role in sex and host recognition. Journal of Hym. Research 5: 206–239.

Ivie, M. A. 2001. On the Geographic Origin of the Wheat Stem Sawfly (Hymenoptera: Cephidae): A New Hypothesis of Introduction from Northeastern Asia. Am. Entomol. 47: 84-97.

Jacquin-Joly, E., Vogt, R. G., Francois, M. C., and Nagnan-Le Meillour, P. 2001. Functional and expression pattern analysis of chemosensory proteins expressed in antennae and pheromonal gland of *Mamestra brassicae. Chem Senses* 26: 833-844.

Jiménez-Martínez, E., Bosque-Pérez, S. N. A., Berger, P. H., Zemetra, R. S., Ding, H., and Eigenbrode, S. D. 2004. Volatile cues influence the response of *Rhopalosiphum padi* (Homoptera: Apididae) to barley yellow dwarf virus-infected transgenic and untransformed wheat. Environ. Entomol. 33: 1207-1216.

Jin, X., Brandazza, A., Navarrini, A., Ban, L., Zhang, S., Steinbrecht, R. A., Zhang, L., and Pelosi, P. 2005. Expression and immunolocalisation of odorant-binding and chemosensory proteins in locusts. Cell. Mol. Life Sci. 62:1156–1166.

Jones, O. T. 1998. Practical Applications of Pheromones and Other Semiochemicals. Insect Pheromones and their Use in Pest Management. P. E. Howse, I. D. R. Stevens and O. T. Jones. London, Chapman & Hall.

Jones, W. D., Nguyen T-A ,T., Kloss, B., Lee, K. J., and Vosshall, L. B. 2005. Functional conservation of an insect odorant receptor gene across 250 million years of evolution. *Curr Bio* 15: R119-R121.

Kaissling K.E. 1998. Flux detectors versus concentration detectors: Two types of chemoreceptors. Chem. Senses. 23: 99-111.

Kamikouchi, A., Morioka, M., and Kubo, T. 2004. Identification of honeybee antennal proteins/genes expressed in a sex- and/or caste selective manner. *Zool Sci* 21(1): 53–62.

Katoh, K., Kuma, K., Toh, H., and Miyata, T. 2005. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res*. 33(2): 511-518.

Keil T. A. 1999. Morphology and development of the peripheral olfactory organs. In: Insect Olfaction, edited by Hansson BS, editor. Berlin: Springer. p. 6–47.

Kessler, A. and Baldwin, T. 2001. Defensive function of herbivore-induced plant volatile emissions in nature. Science 291: 2141-2144.

Knodel, J. J., Beuzay, P. B., Eriksmoen, E. D., and Pederson, J. D. 2009. Pest management of wheat stem maggot (Diptera: Chloropidae) and wheat stem sawfly (Hymenoptera: Cephidae) using insecticides in spring wheat. Journal of Agricultural and Urban Entomology 26: 183-197.

Krieger, J., Ganssle, H., Raming, K., and Breer, H. 1993. Odorant binding proteins of Heliothis virescens. Insect Biochem. Mol. Biol. 23: 449-456.

Krieger, J., Grosse-Wilde, E., Gohl, T., Dewer, Y., Raming, K., and Breer, H. 2004. Genes encoding candidate pheromone receptors in a moth (*Heliothis virescens*). *Proc Natl Acad Sci USA* 101: 11845-11850.

Krieger, J., von Nickisch-Rosengk, E., Mameli, M., Pelosi, P., and Breer, H. 1996. Binding proteins from the antennae of Bombyx mori. Insect Biochem. Mol. Biol. 26: 297-307.

Lagarde, A., Spinelli, S., Qiao, H., Tegoni, M., Pelosi, P., et al. 2011. Crystal structure of a novel type of odorant binding protein from Anopheles gambiae, belonging to the C+ class. *Biochem J* 437: 423–430.

Lagarde, A., Spinelli, S., Tegoni, M., He, X., Zhou, J. J., et al. 2011b. The crystal structure of odorant binding protein 7 from Anopheles gambiae exhibits an outstanding adaptability of its binding site. *J Mol Biol* 414: 401–12.

Landolt, P. J. and Phillips, T. W. 1997. Host plant influences on sex pheromone behavior of phytophagous insects. Annual Review of Entomology. 42**:** 371-391.

Lanning, S. P., Carlson, G. R., Eckhoff, J., Kushnak, G. D., Kephart, K. D., Stougaard, R. N., et al. 2006. Spring wheat variety performance summary in Montana. Montana State University Plant Sciences: Crops Bulletin.

Larkin, M.A, Blackshields, G., Brown, N.P., Chenna, R. McGettigan, P.A., McWilliam, H., et al. 2007. ClustalW and ClustalX version 2. Bioinformatics 23: 2947–2948.

Larsson, M. C., Domingos, A. I., Jones, W. D., Chiappe, M. E., Amrein, H., and Vosshall, L. B. 2004. Or83b encodes a broadly expressed odorant receptor essential for Drosophila olfaction. *Neuron* 43: 703–714.

Lartigue, A., Campanacci, V., Roussel, A., Larsson, A. M., Jones, T. A., Tegoni, M., and Cambillau, C. 2002. X-ray structure and ligand binding study of a moth chemosensory protein. *J. Biol. Chem*. 277: 32,094–32,098.

Laughlin, J., Ha, T., Jones, D., and Smith, D. 2008. Activation of pheromone sensitive neurons is mediated by conformational activation of pheromone binding protein. Cell 133: 1255–1265.

Leal, W. S. 2012. Odorant Reception in Insects: Roles of Receptors, Binding Proteins, and Degrading Enzymes. Annu Rev Entomol Sep 27.

Leal, W. S. 2013. Odorant reception in insects: roles of receptors, binding proteins, and degrading enzymes. *Annu. Rev. Entomol*. 58: 373-391.

Leal, W. S., Barbosa, R. M. R., Xu, W., Ishida, Y., Syed, Z., et al. 2008. Reverse and Conventional Chemical Ecology Approaches for the Development of Oviposition Attractants for Culex Mosquitoes. *PLoS ONE* **3**(8): e3045. doi:10.1371/journal.pone.0003045

Leal, W. S., Nikonova, L., and Peng, G. 1999. Disulfide structure of the pheromone binding protein from the silkworm moth, *Bombyx mori*. *FEBS letters* 464:85–90.

Leal, W. S., Chen, A. M., Ishida, Y., Chiang, V. P., Erickson, M. L., Morgan, T. I., Tsuruda, J. M. 2005. Kinetics and molecular properties of pheromone binding and release. *Proc. Natl. Acad. Sci. USA* 102: 5386–5391.

Lenardis, A. E., vand Baren, C., Di Leo Lira, P., and Ghersa, C. M. 2007. Plant-soil interactions in wheat and coriander crops driving arthropod assemblies through volatile compounds. Europ. J. Agronomy 26: 410-417.

Li, X., Schuler, M. A., and Berenbaum, M. R. 2007. Molecular mechanisms of metabolic resistance to synthetic and natural xenobiotics. *Annu Rev Entomol* 52: 231-253.

Livak. K.J. and Schmittgen, T.D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25: 402–408.

Lou, K. F., Weiss, M. J., Bruckner, P. L., Morrill, W. L., Talbert, L. E., and Martin, J. M. 1998. RAPD Variation Within and Among Geographic Populations of Wheat Stem Sawfly (Cephus cinctus Norton). J. of Heredity 89(4): 329-335.

Maibeche-Coisne, M., Merlin, C., Francois, M. C., Porcheron, P., and Jacquin-Joly, E. 2005. P450 and P450 reductase cDNAs from the moth *Mamestra brassicae*: cloning and expression patterns in male antennae. *Gene* 346: 195-203.

Maibeche-Coisne, M., Monti-Dedieu, L., Aragon, S., and Dauphin-Villemant, C. 2000. A new cytochrome P450 from *Drosophila melanogaster*, CYP4G15, expressed in the nervous system. *Biochem Biophys Res Commun* 273: 1132-1137.

Mayer, M. L. 2006. Glutamate receptors at atomic resolution. *Nature* 440: 456-462.
Mitchell, R. F., Hughes, D. T., Luetje, C. W., et al. (2012) Sequencing and characterizing odorant receptors of the cerambycid beetle *Megacyllene caryae*. *Insect Biochem Mol Biol* 42(7): 499-505.

McBride, D. K, Koop, D. D., and Nyegaard, C. W. 1989. Wheat stem insect pests and management practices. North Dakota State University Extension Service, Fargo, ND. E-680.

McKenna, M. P., Hekmat-Scafe, D. S., Gaines, P., Carlson, J. R., and Putative, D. 1994. Melanogaster pheromone-binding proteins expressed in a subregion of the olfactory system. J. Biol. Chem. 269: 16340-16347.

Meers, S. B. 2005. Impact of harvest operations on parasitism of the wheat stem sawfly, *Cephus cinctus* Norton (Hymenoptera: Cephidae). [Master's Thesis]. Entomology, Montana State University, Bozeman, Montana. 129 p.

Mitchell, R. F., Hughes, D. T., Luetje, C. W., Millar, J. G., Soriano-Agatón, F., Hanks, L. M., *et al*. 2012. Sequencing and characterizing odorant receptors of the cerambycid beetle *Megacyllene caryae*. *Insect Biochem Mol Biol* 42: 499–505.

Mittapalli, O., Neal, J. J., and Shukle, R. H. 2005. Differential expression of two cytochrome P450 genes in compatible and incompatible Hessian fly/wheat interactions. *Insect Biochem Mol Biol* 35: 981-989.

Miura, N., Nakagawa, T., Touhara, K., and Ishikawa, Y. 2010. Broadly and narrowly tuned odorant receptors are involved in female sex pheromone reception in Ostrinia moths. *Insect Biochem Mol Biol* 40: 64-73. doi: 10.1016/j.ibmb.2009.12.011

Monteforti, G., Angeli, S., Petacchi, R., and Minnocci, A. 2002. Ultrastructural characterization of antennal sensilla and immunocytochemical localization of a chemosensory protein in *Carausius morosus* BRÜNNER (Phasmida: Phasmatidae). *Arthropod Struct. Dev*. 30: 195–205.

Morozova, O., and Marra, M.A. 2008. Applications of next-generation sequencing technologies in functional genomics. *Genomics* 92: 255–264.

Morrill, W. L. 1983. Early history of cereal grain insect pests in Montana. Bull. Entomol. Soc. Amer. 29: 24–28.

Morrill, W. L. 1997. The wheat stem sawfly, *Cephus cinctus* Nort. (Hymenoptera: Cephidae) and associated parasitoids in the northern Great Plains of North America. *Trends Entomol*. 1: 171-174.

Morrill, W. L., Gabor, J. W., and Kushnak, G. D. 1992. Wheat stem sawfly (Hymenoptera: Cephidae) at low temperatures. Environ. Entomol. 22: 1358-1361.

Morrill, W. L. and Kushnak, G. D. 1996. Wheat stem sawfly (Hymenoptera: Cephidae) adaptation to winter wheat. Environ. Entomol. 25: 1128-1132.

Morrill, W. L., and Kushnak, G. D. 1999. Planting date influence on the wheat stem sawfly (Hymenoptera: Cephidae) in spring wheat. J. Agric. Entomol. 16: 123-128.

Morrill, W. L., Kushnak, G. D., Bruckner, P. L., and Gabor, J. W. 1994. Wheat stem sawfly damage, rates of parasitism, and overwintering survival in resistant wheat lines. *J. Econ. Entomol*. 87: 1373-1376.

Morrill, W. L., Kushnak, G. D., and Gabor, J. W. 1998. Parasitism of the wheat stem sawfly (Hymenoptera: Cephidae) in Montana. Biol. Control 12:159-163.

Morrill, W. L., Weaver, D. K., and Johnson, G. D. 2001. Trap strip and field modification for management of the wheat stem sawfly (Hymenoptera: Cephidae). J. Entomol. Sci. 36: 34-45.

Mosbah, A., Campanacci, V., Lartigue, A., Tegoni, M., Cambillau, C., Darbon, H. 2003. Solution structure of a chemosensory protein from the moth *Mamestra brassicae*. *Biochem*. J. 369: 39–44.

Munro, J. A. 1945. The Wheat Stem Sawfly and Harvest Loss. North Dakota Agricultural Experimental Station. Bimonthly Bulletin. 7: 12-16.

Munro, J. A. 1947. Wheat sawfly is on the wing. Co-Op Grain Quarterly, National Confederation of Grain Cooperatives. 5: 39-42.

Nagnan-Le Meillour, P., Cain, A.H., Jacquin-Joly, E., Francois, M.C., Ramachandran, S., Maida, R., and Steinbrecht, R.A. 2000. Chemosensory proteins from the proboscis of *Mamestra brassicae*. *Chem Senses* 25: 541–553.

Nakagawa, T., Sakurai, T., Nishioka, T., and Touhara, K. 2005. Insect sex-pheromone signals mediated by specific combinations of olfactory receptors. *Science* 307: 1638–1642.

Nakagawa, T. and Vosshall, L.B. 2009. Controversy and consensus: noncanonical signaling mechanisms in the insect olfactory system. *Curr Opin Neurobiol* 19: 284–292.

Nansen, C., Macedo, T. B., Weaver, D. K., Peterson, R. K. D. 2005a. Spatiotemporal distributions of wheat stem sawfly eggs and larvae in dryland wheat fields. Can. Entomol. 137: 428-440.

Nansen, C., Weaver, D. K., Sing, S. E., Runyon, J. B., Morrill, W. L., Grieshop, M. J., Shannon, C. L., and Johnson, M. L. 2005b. Within-field spatial distribution of *Cephus cinctus* (Hymenoptera: Cephidae) larvae in Montana wheat fields. Can. Entomol. 137: 202-214.

Nelson, W. A. and Farstad, C. W. 1953. Biology of *Bracon cephi* (Gahan) (Hymenoptera: Braconidae), an important native parasite of the wheat stem sawfly, *Cephus cinctus* Nort. (Hymenoptera: Cephidae), in Western Canada. Can. Entomol. 85: 103-107.

Neuhaus, E. M., Gisselmann, G., Zhang, W., Dooley, R., Stortkuhl, K., and Hatt, H. 2005. Odorant receptor heterodimerization in the olfactory system of Drosophila melanogaster. Nat. Neurosci. 8: 15–17.

Oakeshott J. G., Devonshire, A. L., Claudianos, C., Ollis, D. L., and Russel, R. J. 2005. Comparing the organophosphorus and carbamate insecticide resistance mutations in cholin- and carboxyl-esterases. *Chemico Biol Interact* 157–158:269–275.

Oakeshott, J. G., Johnson, R. M., Berenbaum, M. R., Ranson, H., Cristino, A. S., and Claudianos, C. 2010. Metabolic enzymes associated with xenobiotic and chemosensory responses in *Nasonia vitripennis*. *Insect Mol Biol* 19:147–163.

Ochieng, S. A., Park, K. C., Zhu, J. W., and Baker, T. C. 2000. Functional morphology of antennal chemoreceptors of the parasitoid Microplitis croceipes (Hymenoptera: Braconidae). Arthropod Structure & Development 29: 231–240.

Ozaki, M., Wada-Katsumata, A., Fujikawa, K., Iwasaki, M., Yokohari, F., et al. 2005. Ant nestmate and non-nestmate discrimination by a chemosensory sensillum. *Science* 309: 311–314.

Party, V., Hanot, C., Said, I., Rochat, D., and Renou, M. 2009. Plant terpenes affect intensity and temporal parameters of pheromone detection in a moth. Chem. Senses 34: 763-774

Peck, G. E. 2004. Potential semiochemicals of wheat (*Triticum aestivum* L.) induced by oviposition and feeding of the wheat stem sawfly, *Cephus cinctus* Norton (Hymenoptera: Cephidae) [Master's Thesis]. Montana State University, Bozeman, Montana. 65 p.

Pelosi, P. 1998. Odorant-binding proteins: Structural aspects. *Ann NY Acad Sci* 855: 281–293.

Pelosi, P., Iovinella, I., Felicioli, F and Dani, F. R. 2014. Soluble proteins of chemical communication: an overview across arthopods. *Front Physiol* **5**(320): 1-13.

Perez, J., Weaver, D. K., and Morrill, W. L. 2006. Infestation of wheat and downy brome grass by wheat stem sawfly and subsequent larval performance. Environ. Entomol. 35: 1279-1285.

Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., et al. 2003. TIGR Gene Indices clustering tools (TGICL): A software system for fast clustering of large EST datasets. Bioinformatics 19: 651-652.

Pertea, G., Huang, X., Liang, F., Antonescu, V., Sultana, R., Karamycheva, S., et al. 2003. TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. *Bioinformatics* 19: 651–652.

Peterson, T.N., Brunak, S., von Heijne, G. and Nielsen, H. 2011. SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods* **8**: 785–786

Pichersky, E. and Gershenzon, J. 2002. The formation and function of plant volatiles: perfumes for pollinator attraction and defense. Curr. Opin. Plant Biol. 5: 237-243.

Pichersky, E., Noel, J. P., and Dudareva, N. 2006. Biosynthesis of plant volatiles: nature's diversity and ingenuity. Science 311: 808-811

Piesik, D., Weaver, D. K., Peck, G. E., and Morrill, W. L. 2006. Mechanically-injured wheat plants release greater amounts of linalool and linalool oxide. Journal of Plant Protection Research 46*:* 29–39.

Piesik, D., Weaver, D. K., Runyon, J. B., Buteler, M., Peck, G. E., and Morrill, W. L. 2008. Behavioural responses of wheat stem sawflies to wheat volatiles. Agric. For. Entomol. 10: 245-253.

Poupardin, R., Reynaud, S., Strode, C., Ranson, H., Vontas, J., and David, J. P. 2008. Cross induction of detoxification genes by environmental xenobiotics and insecticides in the mosquito *Aedes aegypti*: Impact on larval tolerance to chemical insecticides. *Insect Biochem Mol Biol* 38: 540–551.

Preston, C. A., Laue, G., and Baldwin, I. T. 2004. Plant-plant signaling: application of trans- or cis-methyl jasmonate equivalent to sagebrush releases does not elicit direct defenses in native tobacco. Journal of Chemical Ecology 30: 2193–2214.

Pridgeon, J. W., Zhang, L., and Liu, N. N. 2003. Overexpression of CYP4G19 associated with a pyrethroid-resistant strain of the German cockroach, *Blattella germanica* (L.). *Gene* 314: 157-163.

Quiroz, A., Pettersson, J., Pickett, J. A., Wadhams, L. J., Niemeyer, H. M. 1997. Semiochemicals mediating spacing behavior of bird cherry-oat aphid, Rhopalosiphum padi, feeding on cereals. Journal of Chemical Ecology 23: 2599–2607.

Quisenberry, S. S., Schotzko, D. J., Lamb, P. F., and Young, F. L. 2000. Insect distribution in spring pea-winter wheat-spring barley crop rotation system. J. Entomol. Sci. 35: 327-333.

Raguso, R. A. 2008. Start making scents: the challenge of integrating chemistry into pollination ecology. Entomol. Exp. Appl. 128(1): 196–207.

Ranson, H., and Hemingway, J. 2005. Mosquito glutathione transferases. *Rev Methods Enzymol* 401: 226-241.

Ranson, H., Rossiter, L., Ortelli, F., Jensen, B., Wang, X., et al. 2001. Identification of a novel class of insect glutathione S-transferases involved in resistance to DDT in the malaria vector *Anopheles gambiae*. *Biochem J* 359: 295-304. doi: 10.1042/0264-6021:3590295.

Reddy, G. V. P. and Guerrero, A. 2004. Interactions of insect pheromones and plant semiochemicals. Trends Plant Sci. 9: 253-261.

Reinecke, A., Ruther, J., Tolasch, T., Francke, W., and Hilker, M. 2002. Alcoholism in cockchafers: orientation of male *Melolontha melolontah* towards green leaf alcohols. Naturwissenschaften 89: 265-269.

Reitz, S. R., and Trumble, J. T. 1997. Effects of linear furanocoumarins on the hervibore *Spodoptera exiqua* and the parasitoid *Archytas marmoratus:* host quality and the parasitoid success. Entomol. Exp. Appl. 84: 9-16.

Rewitz, K. F., O'Connor, M. B., and Gilbert, L. I. 2007. Molecular evolution of the insect Halloween family of cytochrome P450s: phylogeny, gene organization and functional conservation. *Insect Biochem Mol Biol* 37(8): 741-753.

Richards, S., Gibbs, R. A., Weinstock, G. M., Brown, S. J., Denell, R., Beeman, R. W., et al. 2008. The genome of the model beetle and pest *Tribolium castaneum*. *Nature* 452: 949–955.

Robertson, H. M., Gadau, J., and Wanner, K. W. 2010. The insect chemoreceptor superfamily of the parastiod jewel wasp *Nasonia vitripennis*. *Insect Mol. Biol*. 19: 121–136.

Robertson, H. M. and Wanner, K. W. 2006. The chemoreceptor superfamily in the honey bee, *Apis mellifera*: expansion of the odorant, but not gustatory, receptor family. *Genome Res* 16: 1395–1403.

Robertson, H. M., Warr, C. G., and Carlson, J. R. 2003. Molecular evolution of the insect chemoreceptor gene superfamily in *Drosophila melanogaster*. *Proc. Natl. Acad. Sci. USA* 100: 14537–14542.

Roemhild, G. R. 1954. Morphological resistance of some of the Gramineae to the wheat stem sawfly *(Cephus cinctus* Norton) [Master's Thesis]. Montana State College, Bozeman, MT. 42 p.

Rogers, M. E., Jani, M. K., and Vogt, R. G. 1999. An olfactory-specific glutathione-S-transferase in the sphinx moth *Manduca sexta*. *J Exp Biol* 202(12): 1625-1637.

Ronaghi, M., Uhlén, M., and Nyren, P. 1998. A sequencing method based on real-time pyrophosphate. *Science* 281: 363–365.

Roux, O., Van Baaren, J., Gers, C., Arvanitakis, L., and Legal, L. 2005. Antennal structure and oviposition behavior of the Plutella xylostella specialist parasitoid: Cotesia plutellae. Microscopy Research and Technique 68: 36–44.

Rouyar, A., Party, V., Presern, J., Blejec, A., Renou, M. 2011. A general odorant background affects the coding of pheromone stimulus intermittency in specialist olfactory receptor neurons. PLoS ONE 6:e26443 10.1371/journal.pone.0026443.

Runyon, J. B., Hurley, R. L., Morrill, W. L., and Weaver, D. K. 2001 .Distinguishing adults of *Bracon cephi* and *Bracon lissogaster*(Hymenoptera: Braconidae), parasitoids of the wheat stem sawfly (Hymenoptera: Cephidae). Canadian Entomol. 133: 215–217

Runyon, J. B., Morrill, W. L., Weaver, D. K., and Miller, P. R. 2002. Parasitism of the wheat stem sawfly (Hymenoptera: Cephidae) by *Bracon cephi* and *B, lissogaster* (Hymenoptera: Braconidae) in wheat fields bordering tilled and untilled fallow in Montana. J. Econ. Entomol. 95: 1130-1134.

Ruther, J., Reinecke, A., and Hilker, M. 2002. Plant volatiles in the sexual communication of *Melolontha hippocastani*: response towards time-dependent bouquets and novel function of (Z)-3-hexen-1-ol as a sexual kairomone. Ecol. Entomol. 27.

Ruther, J., Reinecke, A., Tolasch, T., and Hilker, M. 2001. Make love not war: a common arthropod defense compound as sex pheromone in the forest cockchafer *Melolontha hippocastani*. Oecologia 128: 44-47.

Rutzler, M. and Zwiebel, L. J. 2005. Molecular biology of insect olfaction: recent progress and conceptual models. *J. Comp. Physiol. A Neuroethol. Sens. Neural Behav. Physiol*. 191: 777–790.

Rytz, R., Croset, V., and Benton, R. 2013. Ionotropic receptors (IRs): chemosensory ionotropic glutamate receptors in Drosophila and beyond. Insect Biochem. Mol. Biol. 43:888-897.

Sakurai, T., Nakagawa, T., Mitsuno, H., Mori, H., Endo, Y., Tanoue, S., et al. 2004. Identification and functional characterization of a sex pheromone receptor in the silkmoth *Bombyx mori*. *Proc Natl. Acad. Sci. USA* 101: 16653–16658.

Sanchez-Gracia, A., Vieira, F. G. and Rozas, J. 2009. Molecular evolution of the major chemosensory gene families in insects. *Heredity* 103: 208–216.

Sato, K., Pellegrino, M., Nakagawa, T., Nakagawa, T., Vosshall, L. B., and Touhara, K. 2008. Insect olfactory receptors are heteromeric ligand-gated ion channels. *Nature* 452: 1002–1006.

Scaloni, A., Monti, M., Angeli, S., and Pelosi, P. 1999. Structural analysis and disulfide-bridge pairing of two odorant-binding proteins from *Bombyx mori*. *Biochem Biophys Res Commun* 266: 386–391.

Schiestl, F. 2010. The evolution of floral scent and insect chemical communication. Ecol. Lett. 13: 643–656.

Schneider, D. 1964. Insect Antennae. Annual Review of Entomology 9: 103-122.

Schröder, R. and Hilker, M. 2008. The relevance of background odor in resource location by insects: a behavioral approach. BioScience 58: 308-316.

Seamans, H. L., Manson, G. F., and Farstad, C. W. 1944. The effect of the wheat stem sawfly (*Cephus cinctus* Nort.) on the heads and grains of infested stems. Ontario Entomol. Soc. Annu. Rpt. 75: 10-15.

Sheehan, D., Meade, G., Foley, V. M., and Dowd, C. A. 2001. Structure, function and evolution of glutathione transferases: implications for classification of non-mammalian members of an ancient enzyme superfamily. *Biochem J* 360(1): 1-16.

Silbering, A. F., Rytz, R., Grosjean, Y., Abuin, L., Ramdya, P., Jefferis, G. S. and Benton, R. 2011. Complementary Function and Integrated Wiring of the Evolutionarily Distinct Drosophila Olfactory Subsystems. *J Neurosci* 31: 13357–13375.

Sing, S. E. 2002. Spatial and Biotic Interactions of the wheat stem sawfly with wild oat and Montana dryland spring wheat. Ph.D. dissertation, Montana State University, Bozeman, Montana.

Singh, S. P., Coronella, J. A., Benes, H., Cochrane, B. J., and Zimniak, P. 2001. Catalytic function of *Drosophila melanogaster* glutathione S-transferase DmGSTS1-1 (GST-2) in conjugation of lipid peroxidation end products. *Eur J Biochem* 268: 2912–2923. doi: 10.1046/j.1432-1327.2001.02179.

Skoog, F. E. and Wallace, L. E. 1963. Application of systemic insecticides as seed treatment to protect wheat plants against grasshoppers and wheat stem sawfly. J. Econ. Entomol. 57: 199-205.

Smart, R., Kiely, A., Beale, M., Vargas, E., Carraher, C., Kralicek, A. V., et al. 2008. *Drosophila* odorant receptors are novel seven transmembrane domain proteins that can signal independently of heterotrimeric G proteins. *Insect Biochem. Mol. Biol*. 38: 770–780.

Smith, C. R., Smith, C. D., Robertson, H. M., Helmkampf, M., Zimin, A., Yandell, M., et al. 2011b. Draft genome of the red harvester ant Pogonomyrmex barbatus. *). Proc Natl Acad Sci USA* 108: 5667–5672.

Smith, C. D., Zimin, A., Holt, C., Abouheif, E., Benton, R., Cash, E. et al. 2011. Draft genome of the globally widespread and invasive Argentine ant (Linepithema humile). *Proc Natl Acad Sci USA 108*: 5673–5678.

Somsen, H. W. and Luginbill, P. 1956. Bracon lissogaster Mues. a parasite of the wheat stem sawfly. USDA Technical Bulletin No. 1153. pp. 1-7.

Stein, L. 2001. Genome annotation: from sequence to biology. *Nature Rev Genet* 2*:* 493-503. doi:10.1038/35080529.

Steinbrecht, R. A. 1997. Pore structures in insect olfactory sensilla: a review of data and concepts. Int. J. Insect Morphol. Embryol. 26: 229–245.

Steinbrecht RA. 1998. Odorant-binding proteins: expression and function. *Ann NY Acad Sci* 855: 323–332.

Steinbrecht, R. A., Ozaki, M., and Ziegelberger, G. 1992. Immunocytochemical localization of pheromone-binding protein in moth antennae. *Cell Tissue Res.* 270: 287–302.

Stevens, J. L., Snyder, M. J., Koener, J. F., and Feyereisen, R. 2000. Inducible P450s of the CYP9 family from larval Manduca sexta midgut. *Insect Mol Biol* 7(3): 559–568

Stortkuhl K. and Kettler, R. 2001. Functional analysis of an olfactory receptor in *Drosophila melanogaster*. *Proc Natl Acad Sci USA* 98: 9381–85

Tamura ,K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. 2011. MEGA5: Molecular Evolutionary Genetics Analysis using Maximum Likelihood, Evolutionary Distance, and Maximum Parsimony Methods. *Mol Biol Evol* 28: 2731–2739.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. 2013. MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0. *Molecular Biology and Evolution* 30: 2725-2729.

Tegoni, M., Campanacci, V., and Cambillau, C. 2004. Structural aspects of sexual attraction and chemical communication in insects. Trends Biochem. Sci. 29: 257–264.

Tscharntke, T., Thiessen, S., Dolch, R., Boland, W. 2001. Herbivory, induced resistance, and interplant signal transfer in Alnus glutinosa. Biochem. Syst. Ecol. 29: 1025–1047.

Tu, C. P., and Tang, A. H. 1994. Biochemical characterization of Drosophila glutathione S-transferases D1 and D21. *J Biol Chem* 269: 27876–27884.

Turlings, T. C., Tumlinson, J. H., and Lewis, W.J. 1990. Exploitation of herviboreinduced plant odors by host-seeking parasitic wasps. Science 250: 1251-1253.

Udomsinprasert, R., Pongjaroenkit, S., Wongsantichon, J., Oakley, A. J., et al. 2005. Identification, characterization and structure of a new Delta class glutathione transferase isoenzyme. *Biochem J* 388: 763-771.

Van Baaren, J., Boivin, G., Le Lannic, J., and Nénon, J. P. 1999. Comparison of antennal sensilla of Anaphes victus and A. listronoti (Hymenoptera, Mymaridae), egg parasitoids of Curculionidae. Zoomorphology 119: 1–8.

Van Loon, J. J. A., de Boer, J. G., and Dicke, M. 2000. Parasitoid-plant mutualism: parasitoid attack of herbivore increases plant reproduction. Entomologia Experimentalis et Applicata 97: 219–227.

Vieira, F.G., Forêt, S., He, X., Rozas, J., Field, L.M. et al. 2012. Unique Features of Odorant-Binding Proteins of the Parasitoid Wasp *Nasonia vitripennis* Revealed by Genome Annotation and Comparative Analyses. *PLoS ONE* 7(8): e43034. doi:10.1371/journal.pone.0043034

Vieira, F.G., Sanchez-Gracia, A., and Rozas, J. 2007. Comparative genomic analysis of the Odorant-binding protein family in 12 *Drosophila* genomes: purifying selection and birth-and-death evolution. *Genome Biol* 8: R235.

Visser, J. H. 1986. Host odor perception in phytophagous insects. Ann. Rev. Entomol. 31: 121-144.

Vogt, R. G. 2003. Biochemical diversity of odor detection: OBPs, ODEs and SNMPs. In Insect Pheromone Biochemistry and Molecular Biology G. J. Blomquist, R. G. Vogt (eds) pp391-446. Elsevier Academic Press.

Vogt, R. G. 2005. Molecular basis of pheromone detection in insects. In: L. Gilbert, K. Iatrou, S. Gill, editors. Comprehensive Insect Physiology, Biochemistry, Pharmacology and Molecular Biology. London: Elsevier. pp. 753–804.

Vogt R. G., Prestwich, G. D., and Lerner, M. R. 1991. Odorant-binding-protein subfamilies associate with distinct classes of olfactory receptor neurons in insects. J. Neurobiol. 22: 74-84.

Vogt, R. G. and Riddiford, L. M. 1981. Pheromone binding and inactivation by moth antennae. Nature 293: 161–163.

Vogt, R. G., Riddiford, L. M., and Prestwich, G. D. 1985. Kinetic properties of pheromone degrading enzyme: the sensillar esterase of *Antheraea polyphemus*. *Proc. Natl Acad. Sci. USA* 82: 8827–8831.

Vogt, R. G., Rogers, M. E., Franco, M. D., and Sun, M. 2002. A comparative study of odorant binding protein genes: differential expression of the PBP1-GOBP2 gene cluster in *Manduca sexta* (Lepidoptera) and the organization of OBP genes in *Drosophila melanogaster* (Diptera). *J Exp Biol* 205: 719–744.

Vosshall, L. B., Amrein, H., Morozov, P. S., Rzhetsky, A., and Axel, R. 1999. A spatial map of the olfactory receptor expression in the Drosophila antenna. Cell 96: 725-736.

Vosshall, L. B. and Hansson, B. S. 2011. A unified nomenclature system for the insect olfactory coreceptor. Chem Senses 36: 497–498.

Vosshall, L. B. and Stocker, R. F. 2007. Molecular architecture of smell and taste in Drosophila. Annu. Rev. Neurosci. 30: 505–533.

Vosshall, L. B., Wong, A. M., and Axel, R. 2000. An olfactory sensory map in the fly brain. Cell 102(2):147-59

Wahl, D. V., Shanower, T. G., and Hoelmer, K. A., 2007. A new species of *Collyria* Schiødte (Hymenoptera: Ichneumonidae: Collyriinae), a parasitoid of *Cephus fumipennis* (Hymenoptera: Cephidae) in China, and potential biological control agent for *Cephus cinctus* in North America. J. Kans. Entomol. Soc. 80: 43–50.

Wallace, L. 1962. Field-plot tests of chemicals for wheat stem sawfly control. J. Econ. Entomol. 55: 909-912.

Wallace, L. E. and McNeal, F. H. 1966. Stem sawflies of economic importance in grain crops in the United States. Agricultural Research Service United States Department of Agriculture (in Cooperation with Montana Agricultural Experimental Station), Technical Bulletin 1350.

Wanner, K.W., Anderson, A.R., Trowell, S.C., Theilman, D.A., Robertson, H.M. and Newcomb, R.D. 2007. Female-biased expression of odourant receptor genes in the adult antennae of the silkworm, *Bombyx mori*. *Insect Mol Biol* 16: 107–119.

Wanner, K. W., Nichols, A. S., Allen, J. E., Bunger, P. L., Garczynski, S. F., Linn, C. E. et al. 2010. Sex pheromone receptor specificity in the European corn borer moth, *Ostrinia nubilalis*. *PLoS ONE* 5:e8685. doi: 10.1371/journal.pone.0008685

Wanner, K. W., Nichols, A. S., Walden, K. K., Brockmann, A., Luetje, C. W., and Robertson, H. M. 2007. A honey bee odorant receptor for the queen substance 9-*oxo*-2-decenoic acid. Proc. Natl. Acad. Sci. USA 104: 14383–14388.

Wanner, K.W. and Robertson, H.M. 2010. Odorant and gustatory receptor gene families in the Lepidoptera. In: Molecular biology and genetics of Lepidoptera. MR Goldsmith and F Marek, eds (Boca Raton, FL: CRC Press): 153-168.

Wanner, K. W., Willis, L. G., Theilmann, D. A., Isman, M. B., Feng, Q., and Plettner, E. 2004. Analysis of the insect os-d-like gene family. *J Chem Ecol* 30: 889–911

Waterhouse, A. M., Procter, J. B., Martin, D. M. A., Clamp, M., and Barton, G. J. 2009. Jalview Version 2-a multiple sequence alignment editor and analysis workbench. Bioinformatics 25: 1189-1191. doi:10.1093/bioinformatics/btp033.

Weaver D. K., Buteler, M., Hofland, M. L., Runyon, J. B., Nansen, C., et al. 2009. Cultivar preferences of ovipositing wheat stem sawflies as influenced by the amount of volatile attractant. Journal of Economic Entomology 102*:* 1009–1017.

Weaver, D. K., Sing, S. E., Runyon, J. B., and Morrill, W. L. 2004. Potential Impact of Cultural Practices on Wheat Stem Sawfly (Hymenoptera: Cephidae) and Associated Parasitoids. Montana State University Department of Entomology. J. Agric. Urban Entomol. 21(4): 271-287.

Wei, S. H., Clark, A. G., and Syvanen, M. 2001. Identification and cloning of a key insecticide-metabolizing glutathione S-tranferase (MdGST-6A) from a hyper insecticide-resistant strain of the house-fly *Musca domestica. Insect Biochem Mol Biol* 31: 1145-1153.

Weiss, M. J. and Morrill, W. L. 1987. Influence of planting date and spring tillage on the wheat stem sawfly. Mont. Agr. Res. 4: 2-5.

Weiss, M. J., Morrill, L. W., and Reitz, L. L. 1992. Wheat stem sawfly (Hymenoptera: Cephidae) Revisited. Am. Entomol. 241-245.

Wen, Z., Baudry, J., Berenbaum, M. R., and Schuler, M. A. 2005. Ile115Leu mutation in the SRS1 region of an insect cytochrome P450 (CYP6B1) compromises substrate turnover via changes in a predicted product release channel. *Protein Eng Des Sel* 18: 191-199.

Wetzel, C., Behrendt, H., Gisselmann, G., Stortkuhl, K., Hovemann, B. and Hatt, H. 2001. Functional expression and characterization of a Drosophila odorant receptor in a heterologous cell system. *Proc Natl Acad Sci USA* 98: 9377–80

Wicher, D., R. Schafer, R. Bauernfeind, M. C. Stensmyr, R. Heller, S. H. Heinemann, and B. S. Hansson. 2008. *Drosophila* odorant receptors are both ligand-gated and cyclic-nucleotide-activated cation channels. Nature 452: 1007–1011.

Willingham, A. T. and Keil, T. 2004. A tissue specific cytochrome P450 required for the structure and function of Drosophila sensory organs. *Mech Dev* 121(10): 1289-1297.

Wojtasek, H. and Leal, W. S. 1999. Conformational change in the pheromone-binding protein from Bombyx mori induced by pH and by interaction with membranes. J Biol Chem 274: 30950–30956.

Wurm, Y., Wang, J., Riba-Grognuz, O., Corona, M., Nygaard, S., Hunt, B. G., et al. 2011. The genome of the fire ant Solenopsis invicta. *Proc Natl Acad Sci USA* 108: 5679–5684.

Xu, P., Atkinson, R., Jones, D. N., and Smith, D. P. 2005. Drosophila OBP LUSH is required for activity of pheromone-sensitive neurons. Neuron 45: 193-200.

Yandell, M., and Ence, D. 2012. A beginner's guide to eukaryotic genome annotation. *Nature Rev Genet* 13: 329-342. doi:10.1038/nrg3174.

Yao, C. A., Ignell, R., Carlson, J. R. 2005. Chemosensory coding by neurons in the coeloconic sensilla of the Drosophila antenna. J. Neurosci. 25(37): 8359-8367.

Yuan, X., Zhang, S., Zhang, Z., Kong, X., Wang, H., Shen, G., and Zhang, H. 2013. Antennal morphology and sensilla ultrastructure of the web-spinning sawfly Acantholyda posticalis Matsumura (Hymenoptera: Pamphiliidae). *Micron* 50: 20-28.

Zacharuk, R. Y. 1985. Antennae and sensilla. G.A. Kerkut, L.I. Gilbert (Eds.), Comprehensive Insect Physiology, Biochemistry and Pharmacology, Pergamon Press, London, pp. 1–70.

Zhang, G. J. and Xiao, G. X. 1992. Scanning electron microscopy studies of antennal sensilla of Spalangia endius and Nasonia vitripennis. Entomological Knowledge 35: 154–159.

Zhou, J. J., Zhang, G. A., Huang, W., Birkett, M. A., Field, L. M., Pickett, J. A., and Pelosi, P. 2004. Revisiting the odorant-binding protein LUSH of Drosophila melanogaster: evidence for odour recognition and discrimination. FEBS Lett. 558: 23–26.

Zhou, X., Slone, J. D., Rokas, A., Berger, S. L., Liebig, J., et al. 2012. Phylogenetic and Transcriptomic Analysis of Chemosensory Receptors in a Pair of Divergent Ant Species Reveals Sex-Specific Signatures of Odor Coding. PLoS Genet 8(8): e100293

APPENDICES

APPENDIX A

cDNA SEQUENCES OF CEPHUS CINCTUS OLFACTORY-RELATED GENES

>CcCSP1
ATGACTCTAATCAAGATCGCCCTCCTGTGTCTGATAGTTTATACCGTAGCAGG
ACAAAAGGCAGAGGAGCCAAGAAAACCACGCGTTTCCGACGAACAGCTTAA
TGTGGCACTGAGCGACGAGCGATACTTGAGAAGACAGCTTAAGTGTGCCTTG
GGAGAAGCACCGTGCGATCCTGTTGGTCGCCGTTTGAAAAGTCTGGCTCCAC
TGGTACTACGAGGTTCCTGTCCGCAGTGTAGTCCCGAGGAAACAAGACAAAT
TAAAAAGGTTCTTTCCCACATTCAAAGAAGCTTCCCGAAGGAATGGTCAAAG
GTAGTGCAACAGTATGCCGGTGTATAA

>CcCSP2
ATGAAAATCACCCTGGTTTGTCTTATCGGCTTTGCAGCTGTAGTAGCTGTCTC
GGCTGCTCCCCAATATACTACGAAGTATGATAATATTGATCTGAAACAAATTC
TTGAGAGTGATCGACTCTTGAATAATTATTTCAATTGTCTGGTCAACAAAGGA
CCCTGCACTGCTGACGGTCAAGAACTCAAAAAAGCTCTTCCTGATGCCATCG
AGACCGGATGCAAGAGTTGCAACGATAAACAGAAGAAAGGCTCGGATGAGG
TCATAAGGCACATCTACAAACACAAACCTGAAATGTGGAAAGTTCTCACCGA
AATGTACGACCCTGAACGCATTTACATTAAGAAATACGAAACTGAGGCCAAG
GATCTTGGCATTGCTGTTTAA

>CcCSP2b
ATGAAGACCTTTATCGTCTTTTTTGCTATTTTCGGTGTTCTCCTTGCCGAGGAG
AATAAAAAGTACACCACGAAGTACGACAACATAGACGTTGATGTTGTCCTTA
AGACCGAGAGATTGCTCAATAATTACATAGGATGTCTCTTGAATGAAAATCC
TTGTACTCCTGACGCAGCTGAATTAAAAAAAAATTTACCAGATGCTTTGGCCA
CGGATTGTACAGCGTGTAGCGAGGCTCAGAAAGTGGCTTCTGACAAGTTTTC
TCAGTATATGATTGAAGAGAGACCGGATGATTGGAACCGATTAGAGAACAAG
TACGATCCTAGCGGAGCATACAAGACTCGTTATCTTGAGGAAAAATCTAAAA
AATCAAAGAAGCCAAATTAA

>CcCSP3
ATGAGAACTCAACTGTTATTGGTTGCTGTGGTCGTAGGAGTTTTTGCTCTGTG
CCAGGCACAGGATATTTCTCTTTTACTGAACGACAGAAATTATGTCGAGAAA
CAGATTAATTGTGTAGTCGGTAAAGGATCGTGTGATAGAATAGGACAGCAGA
TCAAAGTTTTGCTACCGGAAGTACTGAATAATCAGTGCTCACGATGCTCTCCC
CAACAGGCTCAGAATGCTCGGAAACTGGTCGATTTTATGAAACAACGTTATC
CCAACGAATGGCGCATCATTCTTAAACGTTTTTCTGGACGTCAGGGTTGA

>CcCSP4
ATGAAGGTTGCCGTAATCATCCTCGTCCTCGTAGGATGCGCCATTGCCGCCGA
AAAGTATACGACCAAGTATGATAACATTGACCTAGATCAGATCTTGAAAAGC
GATCGTCTTCTTAACAACTACGTAAATTGCCTGTTGGAAGCTGGAAATTGCAC
TCCCGACGGCAAGGAACTGAAAAAATCTCTGCCCGATGCTCTCGCCAACGAT
TGCAGCAAGTGCAGCGCCAAGCAGAGGGATGGCAGTGAGAAGGTGATTCGT
TTTCTGGTAAACAAGCGTCCCCAGATTTGGGAGAAACTTGCTAAGAAGTATG

ATCCAAGCGGCCAATACAAGATCAAGTTCGAGGGAGAGGCACAGAAGGTTG
GAATCAAGCTGTAA

>CcCSP5
ATGCGATTTATACCTGTATTTCTTGCCTGCCTGACGGTTGCTTTTGCACAAGA
ATATTATTCTAGTAAATTTGACAATATTAATATCAAAGAAATTATCGACAATA
CACGATTGTTTGCAAAGTACAAAGAATGCGTTCTACAAGAAAAGGCAACTCG
CTGTCCTCAGGAAGCATTGGAACTAAAAAGAGTGTTACCGGAAGCATTGGGT
ACTCTGTGCGCTAAGTGTACTCCCTCACAAGTTACTAAAATACGGGAGGGTCT
GTCTTACGCCTGTAAAAATCGAAGAGTGGATTATGATGAAATACTGAGACAT
GTTGATCCCCAGGGAGATAAAATTGTTGCTTTTGAACAAAAATTTGGAAAAG
TTGAATGTTGA

>CcCSP6
ATGGGAGAGCGTTTGCTCGCGTTTACGTTAATTTTGTGCTTAGTGGTGATCGC
CGTCCCAAAGGGAACACAGGGTTACCTTTGGCCAAGACATGACACGTATACG
ACACGCTGGGACAAAGTCAACGTGGATGATATCCTCGAGAGCAAGAGACTCC
TTCACTATTATTTCAATTGCCTGATGAATAAGGGTCCTTGCCCGCCGGATGGT
CATGAACTCAAGAGAGTTTTGCCAGAAGCTCTGCAAACAGCCTGTGCAAAAT
GTACAAAGAGCCAAATCGAAGGATCTGTCAAAGTCATAAGATATCTCAGGCA
ATTTGAGCCTGGGAAATTCAAGCAACTTGCTGAGAGATACGATCCCGAGGGC
ATGTACAGAAAACGATACCTCGAACAATCATTAGATAATAATACGGCATAA

>CcCSP7
ATGATTGGCGTTTATGTCCTCGTGTTGGCTGTCATTGGCTGTAACGCAGTTTA
CGCTGATGAATTTTACAGCAATAAATATGACAATGTAAATGTGAATCAAATT
TTAAAAAGCGAACGGCTTTTGCAACGTTATATTCTCTGTTTACTTGACAAAGG
ATCTTGTACATCTGACGGCCGCTTCTTTAAAGAAATATTACCGGAGGCTTTAG
CAACGAATTGTTCAAAATGCAGCATGAAGCAAAGAGAAATAGTGAAAACCTT
GACTTTGCATCTAATGAATAATAAGCCTGACCACTGGCGCGAATTTGTGGAA
AAaTACGATCCGGATAACAAATATAGAACTTCCTTCCTAAATTTCATAATGTC
AAGTTGA

>GST1a
ATGCCTATCGATTTCTACTATTTACCACTGAGCCCACCGTGTCGAGCAGCGAT
GCTGCTGGCCAAGGCCGTCGGTGTCCATTTGAACTTAAAAACAGTTGACGTA
TTAAAGGGAGAGCAAATGTCACCGGAATTTTTAAAGATAAATCCTCAACATA
CAATACCAACCCTTAATGACAATGGTTTCGTATTATGGGAAAGCCGTTGCATT
ATGGGTTATCTCGTGGACAAATATGCCAAGGATGATTCGTTATATCCAAAAG
ATCCACGCGCACGCGCAATCGTAGATCAACGATTGTACTTCGACATTGGAAC
TCTAAACGAATCACTCACGAAATGTTACTATCCAGTTTTAACGGGGAAAACC
AAAGTTATCGACGAGGAAAATATAAAGGATTTGGAGAATGCCTTCGAGACTC
TCGACATGTTTCTGGATGGCAGACGGTTTGCCGCCTCTGTCAATTTGACGATC
GCTGACTTCAGCATCGTGGTAACAGTCAGCACAGCGGAAAGTTTTGGTTTTG

ACGTCGGTCGATACGACAACGTTGCGGCCTGGTATGAAAATTGTAAAAAAGC
TCTCGAGCGATTCGACTACGAGGAGATCAATGCGAAGAGTGCGAAAAAGTTT
GGCGAGATCTACAAGAGTCGGCTGGAGGCTTAA

>GST1b
ATGCCCATCGACCTATATCACGTACCCGGAAGTGCACCATGTCGGACTGTAA
GACTGGTGGCTGCTGCTGTGGGCGTCGACTTGAACCTGAAGTTCCTGGATTTG
ATGAACGGCGAGCACCTAAAGCCTGAATTTATCAAGATGAATCCGCAGCACA
CAATTCCTACGATAGACGACAATGGCTTTTATCTTTGGGAAAGCCGAGCGATT
TTGGGCTATTTGGTCGACCAATATGGCAAGGATGACTCACTCTATCCAAAGG
ACCCTAAGAAGCGTGCAATGGTAAATCAAAGACTGTTCTTTGATATAGGCAC
CCTGTACCAGGTTATCGTCGATGCTATTTACCCCACGATTTTCGGTGGTGCAC
CAAAGGATTCTGCTCTGTCTCAAAAGATCGACAAACCCATGGAATTCCTGGA
AATATTTTTGGAAGGCCAAAATTATGTTGCTGGAGACAAATTGACTATTGCA
GATATTGCGCTGATTGTCACTGTTTCAAATTTCGAAGTAGTCAATTACGATTA
CAACAAATTCAAAAACATCAGCCGTTGGTTCGCCAGAGTCAAGGCTGAGATT
CCAAAGTATGAAGAGATCAATGGCGAAGGCTTGAGAGCATTCAAAGCACTTG
CCGAGAGTTTATCCAAGAAGTGA

>GST1c
ATGAATTTGGATTTTTATTACGCCCCTACCAGCTCACCTTGCCGTGCCGTTAT
GCTGACTGCTGAAGCCATCGGAGTGACTTTGAATTTAAAGAGCATTGATGTG
ATGGCTGGGGAACACCTTAATTCGAAATACGAGCAAATAAATCCACAAAAAA
CTGTTCCATGTTTGGTGGACGGTGATTTGACATTGACAGAAAGTCGTGCTATA
ATGGCATATTTGGTCGACCAGTATGGAAAAAATACCAGACTCTATCCGAAAA
TGGTAAACAGTCGTGCGTTAGTCAATCAAAGACTGTATTTTGATATTGGTACG
TTTCACAAAGCAATGGCAGCTTGTTACTATCCGATATTATTTGGAAAATGGAA
AACGTATCAGTCCGAGCATTACGACAATTTGAAAAAGGCTTTTGAAATTTTG
GATAAATTTCTCGAGGGTCAGGATTACGTAGCAGGACGCAGCTTAACGATAG
CTGATCTTTCTTTAGTGGCATCCGTTACCACTGCAGAGGCTTTCGGTTTTGATT
TTGCAAAATTCAGAAACGTCACAAGATGGTTGAAGAAGGTGAAGACCTTTGC
TCCTGGATACCGCAAGGCTAATGGAGAAGGAGCCGAATCAATGAAGAAGCT
GATCGAACGAGTCATGACACAAATAGACGAGTTATAG

>GST2
GTCTATCGATCTTTATTTGATGCCCGTAAGTGCCCCATGCCGAGCTGTGCGGC
TGGTAGCGGCGGCCGTCGGCGTCGATCTGAACCTGAAACTCATAAACCCGCA
GCACACAGTACCCGTAATCGACGACAATGGCTTTATCTCGCTGAAAGCCGAG
CGATTTCGAGCTACTTGGTCGACAAATACGGGAAGAACGACTCCCTCTACCC
GAAGGACATCAAAGAACGTGCAGTCGTCAATCAGCGATTGTATTTCGACATT
GGTACCCTGTATCAGGCATATACGAGTGCTTACAGACCTGTAATATTCAGCG
GAGCCCCTAAGGAAGCTGTGTTGATTAAAAATTTCAACGAAAGTCTGAAACT
CTTTGAACAATTTTTGGAGAACGATCATTACGTGGCTGGAGAAAATCTTACG
ATTGCTGATCTTTCGTTGGTTACCACCGTATCCGAGCTTGAGGTCTTGGAACA

TGACTTCAGCAAGTTTGAGAACATTAATCGCTGGTACGCCAAAGTTAAATCC
GAAGCACCGCAATACGAGAAGATCACCGAAGAAGGAATTGGGGCATTGAAA
GAATTGCTCGAATTGCTTAAGAAATAA

>GST4
ATGGGTTTAAAACTGTATTATGACTTGATGTCACAACCTTCCAGGGTCCTCTA
TATATTCCTCAAACAATGTAAAATCCCTTTCGAGGAACATAAAGTAGTTTTAA
AAAAGGGTGAACACTTTATGCCAAAATATGAAGAGATCAATCCATTTCAAAA
GGTTCCAGCTCTTGAGCATAATGGATTTAAACTGATAGAAAGCGTGGCAATT
CTGCGATATCTTTGTCGGGAATTCAAGGTAGATGATCACTGGTATCCCATATC
ATCCAAAAATCAAGCAAGGGTTGATGAGTACCTTGAATGGCAGCATAATAAT
ACTAGATTGCATTGCGCAACATATTTCTTAGTAAAGTTTTTGATACCAATGAT
CAAAGGGAAGCCTGCAAGACCAGAAAATGTTACAAAGTGTGAGGGACATAT
GATTGAATGTCTAGATCAACTTGAAAATATTTGGTTGAAAAACAAACTATTCC
TCACAGGTGACAAGATAAGTATTGCTGATATTGTAGGTGCCTGTGAGGTTGA
ACAACCAAGAATGGCAGGCTTTGATCCTCGTGAAGGTAGACCACAGCTCACT
GCGTGGCTTGAAAGAGTAGCTGAAGAGACAAATCCTCATTATGATGAAGCGC
ATAAATATGTTAATCAGATTGCAAACAACTACTTGGGAATACCACCACACAG
CTTGAGTAAGCTTTAA

>GST5
ATGCCGAAATATAAGCTAATATATTTCAATTTGCCTGTTCTCGGCGAACCGAT
AAAGTATCTTTTACATTATGGCGGAATTGAATTCGAAGACGTACAGATAAAG
GATCGCGTGAACGAATGGCCGAAGATAAAGCCGACAATGCCCTTCGGCCAA
GTACCTGTTCTTGAAATTGATGGGAAACGGTATCCTCAGAGTTTAGCGATATG
TCGTTACCTGGCTAAGCAATTGAACCTAATCGGTAAAACAGATCTGGATCAG
TTGGAGATAGACGGGTTAGCCGATACGATAAGCGAATTTCGAAAATTGTTCA
GCCTGTACTACAGGGAACCCGATCCAGTTATAAAGGCTAAGAAGAAGGAATC
GCTGGAGAATATTGAGGTTCCATATTTCTTGGATAAATTTGAGGAAATCGTGA
AGAATAACAATGGATATTTTCATGGCGGAGAGCTGAGCTACGTGGACTTCTT
TTTCGTTGGATTAGTCGGTGCCTTTAATGCTCTGTTGGACAAAGATGTCTTGG
CAAACTATACAAATTTGAAATCCCTCAAGGACAAGGTCCAAGCTATACCGGC
AATACAAAATTATTTGAAAGACGCACCTAAGCCGATATTTTAA

>GST6
ATGGCGGACCAGCAGCACAGCTATAAACTGATCTATTTCAACTCTCGTGGCC
GTGCCGAGCACATTCGTTACATATTCGCGTTCGCCGGTATCGATTACGTCGAC
GAGAGGGTACCGAAGGAACGCTGGCCGGAACTCAAGAAATGTATGCCCTAC
GGGATGCTTCCGGTCTTGGAAATCGATGGGAAAGCCATTGCTCAGAGCAACG
CAGTTGCGAGGTATCTGGCGCGGCAGCATGGCCTCGCAGGCCTCGACGAATG
GGAATCCATGCAGTGCGATGTTCTTGTCGACACTCTGGGCGACCTCAAGCAA
GGGATCTCGCAGTATCGCCACGAGGAGGATCCGTTCAAGAAGGAGGAAAAG
AAAGCGAAACTTTGGAAGGAGACGATACCATTTTATTTGAACAAATTCGAGC
AGAACGTTGCTGAGAATGGCGGATACTCCGTTGGCACTACTACAACGTGGAC

GGATTTTCTGTTCGCCGTGGCGTTGGAGAATTTCGAGCAAATCTTCGGATCCG
CTGCTCTTGAGAACTACCCTGCTCTCAGGGGGTTGAAGCAAAGGGTTCATGA
AATACCGGCTATAGCCGAATGGCTGGGTAAACGACCCCAGACGGAATTCTAA

>GST7
ATGAGTACCAAGCATTTAGGACCTGGTTCTGTCGCCCCACCCATTACACCTGG
TAAGCTACGGTTGTACAGTATGAGATTTTGTCCTTACGCACAAAGAATACATC
TTGTACTTGATGCTAAACAAATACCGTATGATGTGGTTTATGTAAATCTTACT
CGAAAACCAGAATGGTTACTGGAGAAAAATGAATTGGGGAAAGTACCTTGTT
TAGAATTTGAGAGTGGAGAAATTCTATATGAAAGTCTAATTATAGCAGATTA
TTTGGATGAAGCTTATCCACAAAATGGTCTTTATCCAACTGATCCTATGGCTA
AGGCTAAGGATAGACTACTTATTGAGAGATTTAATACAGTTATAACAATTAT
GTACAAGTTGTATGCAGAAACAATAGTTGACCGGGATTTATTTGACAAGGCT
TTAGCTGGCTTAGAAGTTCTTGATAGAGAACTTGCAAAACGAGGGACACCAT
TCTTTGGAGGAAGTAAACCAGGAATGCTGGATTTAATGATATGGCCATGGTG
CGAGAGAGCTGATATTATTAGAGTTTTACGCGGTGATCAATATGTTATACCTC
GTGAGCGGCTTTTAAGATTGCTGGAATGGCGAACTGCAATGAAAGAGGATCC
TGCTGTGCGGGGTAGTTATCTAGACACAGAAACACATGCTAAATATATTAGA
AGCCGTCTAGCTGGTTCACCACAATATGACATGCTCGCTTCTGCATAG

>CcIR1
ATGCAATTACTGGGGTCGAACGTGGAACCAAAATCCGGACACTGTGAAAATA
AAAATTCAACTATCACACGGAATAACTCTACTATGGAGTTATTCAATAAAAA
TGTAATAGCTAAACTTAATGGGAACGGCACGCTATTTTATATTCCGGATATTC
AATTAATAACCTATAAAGCAGAAATGGAATTGCAAACTTATGGTAACGGTAG
CTTAGAGAAATTTGCCGATTGGAATATCGAAAAGAAAGTCACGACGTTACCA
GGGAAGACCATAGAACCGGCAAAACGATTTTTCCGCGTTGGAACCACTCAGT
CGATCCCTTGGTCGAATCTGAGGAGAGACCCGGTTACAGGTGAATCGATAAA
AGACAGTAATGGTAAAGAAGTCTGGGAAGGTTATTGCATAGATTTTATAAGC
GAACTTGCAAAAGGAATGAATTTCGATTACGAAATAGTCATTCCTAAGTCAG
GAGAATTTGGGGTGAAATTACCAAATGGAAAATGGGACGGACTAGTTGGTGA
TTTGTCACGTGGGGAAACGGACATTGTGGTGGCTTCACTGACGATGACTTCTG
AACGCGAAGAAGTCATTGATTTCGTTGCACCATATTTTGAACAATCTGGACTC
TCAATTGTCATGAGGAAACCTGTACGTGCGACGTCACTTTTTAAGTTCATGAC
TGTATTGAGGCTCGAGGTCTGGTTGAGTATTGTTGGAGCACTTACTTTGACTG
GAATAATGATTTGGCTATTAGATAAATATTCACCATACAGTGCAAGAAATAA
TAAGCGTCTCTATCCGTATCCATGCAGAGAATTCACGCTAAAGGAGAGCTTCT
GGTTTGCTCTAACTTCCTTCACACCACAAGGCGGTGGCGAAGCACCCAAAGC
ATTATCTTCCAGAACCCTGGTGGCTGCTTATTGGCTTTTCGTTGTATTAATGCT
GGCTACTTTTACTGCCAATTTGGCAGCTTTCCTTACAGTGGAGAGAATGCAGT
CACCTGTACAATCACTCGAACAACTGGCTCGACAATCGCGAATAAATTATAC
AGTTGTTCGAAATTCCACTATTCATCAATATTTTATTAATATGAAAAATGCGG
AGGATAAGCTCTATACCGTTTGGAAGGAAATCACATTAAACAGTACTAGCGA
CCAAATTGAATATCGCGTGTGGGACTATCCGATAAAGGAACAATATGGTCAC

ATACTTCAAGCCATCACACAAGCTGGTCCCGTAGCATCGGCGGAGGAAGGTT
TTCAAAAAGTAATCGCCAGTGAAAATGCCGAGTTCGCTTTCATTCACGATTCT
TCGGAAATAAAGTACGAGGTGACAAGAAATTGTAACTTGACTGAAGTCGGCG
AAGTTTTCTCCGAGCAGCCATATGCAATTGCTGTTCAACAAGGAAGCCATCTT
CAAGAAGAGATCAGTAAAATTATTCTTGATCTTCAAAAAGATCGATATTTCG
AATCACTCGCATCAAAATATTGGAATCAATCACTCAAAGGATTGTGTCCAAA
TTCTGATGACAATGAGGGCATCACTTTGGAGAGTTTAGGTGGTGTTTTCATAG
CGACACTTTTTGGTTTGGCATTAGCAATGATCACTCTAGCTGGAGAAGTCGTA
TACTTCAGGAAGCGTAAAGCAAATCAGCGCGATCAAAGCAGAGACAAGGAC
AAGAATAAAAAACAGGACGAAAGATTTTACAGGACGTCTGACATACATAAG
AGTCTTGGAAAATTAGAAATTAAACCGGCTCCAACTGTTTCCTTTATTGGGAA
ACCAACTGGTCCAAGGACACGTGTCTCTCATATATCGGTGTATCCTAGACACC
TTCCTTTTAAGGATTAA

>CcIR2
ATGAAACTCCTGCTCCTGGTGATCGCGATCCACTGCGCGACTTTCTCTCGACG
TGGGTACGAAGTTGCCGAAGCACAACAGGATCCAGGGGGAACGTCGGCCAA
TAGGGATGTCAGACCCGTTAATCTCTTTATCGTGAACGACGTGACTAACGAC
ATAGCCAACAAGAGCGTGATAAATGCCCTGAAGACTCTGAGAGATAAGAATC
CAGGCTGGATAGGAGAGGTCTCCACTGTGCTGATCAACGGATCCGATCCGAG
AGATACATTGGACAGACTCTGCGCCGCTTGGGATTTGGCGCTACGAGCAGGA
AGCACCAAAGTCCCAGACTTGGTCCTGGACACGACCAAGTCAGGTTTGGGCG
CTGAGACTGTCTGCTCCTTCACGGCAGCTCTGGGCTTGCCCACTCTGTCCGCT
CAGTACGGTCAGCAAGGTGATCTACAGCAATGGGAGAATTTGACACCGGATC
AAACAAAGTATCTGGTGCAAGTGATGCCACCGCCCGACCTTGTTCCTGAGGT
CATCAGACAGCTGTGCCTGTCCATGAATATCAGCAACGCCGCCATTCTCTTCG
ACGAGGGTTACGTAATGGATCACAAGTACAAGAGTCTTCTGCAAAACGTGCC
GACGAGACACGTGATCATCGAGGCGAAGCAAACGACCACCGAGTTGCAGTC
GCAGCTGTCGCGGTTGCGTGATCTTGACATTGTCAACTTTTTCGTCCTCGGCG
ACCAGGAGACCCTTAGTGTCGCTCTGGACGCAGGCGAGTCCTTGAACTTCAC
AGGTAGAAAGTACGCCTGGTACGGACTCACACTCGACGACTTCAGTCTACAA
TGCGCCTGCAGGAACATCTCCGTTCTCTTTTTGAAACCGAAGCCCGCCGCAAG
TCAGCAAGTCTTGAGCGAGCTGACCAGTCGTGGCTTATTGCCAAAACCACTG
CTGCTCTCGGCATTCTACTACGACTTGGTTCGTCTTGGTATCTCGGCCATGAG
GTCGGCGTTGGAGTCCCGTCAGTGGCCCAGAGAACCCAGTCACATCAGCTGC
GACGACTACAACGTCAACAACACTCCCGTGAGGAACTTTGACTTTTTGTCCAC
GTTGAAAGTCACGACGCGCAACAGCGACTTTGTCCCGACTTACGGAGGATTC
GCTTGGGGCAAGTCGAATGGCGACCATCGCGCCAACTTCAGCATGACGATCA
ATCTGGTGATAATTAACAACGGGAATCCCATATCAACGGAGGAAATAGGCGC
GTGGCAAGCCGACGTCGACGGCGACTTGCAGGTCTTCGAGGAGGATCTATTA
ACGAATCACACGGCGCTAACGAGCTATCGCGTAGTAACGGTTCAGCAGGCTC
CATTCATCGTGTACGATGCTGAAAACGACACGTGGTCAGGTTACTGCATCGA
CTTACTCAACGAGATTCGTGAGATTCTCATGTTCGAGTATGAGATCCGTGAGG
TCGATGACAAGCTCTTCGGTAACATGGACGAAGAAGGGAATTGGAACGGGAT

GATAAAGGAGCTGAAAGACAAGAGAGCGGACATAGCGCTGGGCTCATTGTC
CGTGATGGCCGAGAGGGAGAACGTAGTGGACTTCACAGTGCCCTATTACGAC
CTAGTGGGCATAACGATTCTCATGCTAAAACCCAAAACACCCACCTCCCTCTT
CAAGTTCCTCACTGTTCTCGAGAATGACGTCTGGCTCTGCATCCTGGCCGCAT
ATTTTTTCACGAGCTTCCTCATGTGGGTCTTCGATCGCTGGTCGCCGTACAGC
TATCAGAATAACCGCGAGAAGTACAAGGACGACGAGGAGAAGCGCGAGTTC
AACCTGAAGGAGTGCCTTTGGTTCTGCATGACTTCCTTGACGCCTCAGGGTGG
AGGCGAAGCGCCGAAAAATCTTTCGGGCAGACTAGTCGCGGCAACCTGGTGG
CTCTTTGGGTTCATCATCATCGCTTCCTACACTGCGAATCTTGCTGCTTTTTTG
ACGGTCTCTAGGTTGGATACTCCCGTCGAGTCTTTGGATGACCTCAGCAAGCA
GTACAAGATTCAATATGCCCCGCTCCAGGGTTCGTCCGCCTACACCTACTTCC
AGAGGATGGCCGACATCGAGAAGAAATTTTACGAGATCTGGAAGGATATGA
GTCTGAACGACAGTCTCTCTGACGTGGAAAGGGCGAAGCTGGCCGTTTGGGA
TTATCCGGTGAGTGACAAATACACGAAGATGTTTCAAGCCATGAAGGAGGCC
GGCTTCCCCACGGACATGGAGGACGCCCTGAGTAGAGTAAGACGCTCTAAGA
ATTCCAACGACGAATTTGCATTCATCGGCGATGCAACCGACATTAGGTATCTT
ACCATGACCAATTGCGACCTCACTATGGTGGGCGAGGAGTTCTCTAGGAAGC
CGTACGCCATTGCTGTCCAGCAAGGATCGCCTCTCAAGGATCAGTTCAATAA
CGCAATCCTGGTCCTACTGAACAGACGAAAGTTGGAGAAGCTCAAAGAGAA
GTGGTGGAATCGAAATCCGTTGAAGAAGACGTGCGAGAAGCAAGACGACCA
GAGTGACGGGATTAGCATTCAAAATATCGGTGGCGTTTTCATTGTCATTTTCG
TGGGAATTGGCTTGGCTTGCCTGACACTGGCCTTTGAATACTGGTGGTATAGA
TATCGACCCAGAGCCAATAACGCGAATCATAATATTCACCAAGTGCCGAAGG
GTGTACCAAATGTTAAGCCGATAAGATTTAATTTGCAACCAGCAGCGACACG
CGCGTTTGAGTCACAGCCAAAGTTTCGGGCGCGTTTTTAA

>CcIR3
ATGGCGGTTATCATCGAAGAGAATTTCTTCAAGGATAAATCCTATTATGATTG
GACATCAAAAGAAGTTTATAATTTTATATTTAACGGTGTTAAGGATAATATGA
AATTTAGTAGTATCGATTTTCATATCTTCCATAATGTCGATATTAGTCTCAAA
AGAGATTATACTGTGTTGCTATCCATTGTTACGTGTGATGAAACATGGCGGTT
ATTTAAAATCGCGCAGAAGGAGCATCTTGTTCATTTAGCAATAACTGATTCGG
ATTGCCCACGACTTCCGGAAAATGGTGGAATAAGTATACCTATGATGCAACC
GGGCGAGGAGTTGCCCCAAATTCTCTTGGATTTAAGGATGAGCAAGGCGTTG
AATTGGAAAAAAATAAATGTCTTACATGATCAAATCCTTGGAAAAGATACGA
TCAGCAGAATTCTCATGGTCTTTTCGAGAGAGTTTGAAAAAAATTTGGAGTTA
GCATCTAGATCAGTTTTTGCAATTCAATACAAAGAAACTGATTGGGCTCGTAG
GCAGCAGCTAGTGAATATACTTTCCGGATTTGATGTAGATCAACTTGGAAATT
GCTTTTTGGTCTTCACCACAGCAGACATGGTTGCGGTTATCATGGAAGCACGA
TCCTTGAAAATGGTTAATACATCTAGTCAATGGCTTTATGTCATAAATGATGC
TAGCGAAAGAGATGTTAATATAACATCATTCAGCGAATTATTAGAAGAAGGG
GAAAACATAGCTTTCCTTTATAATATAACAAACAAAAACAAAACGTGCAGTA
TTAATTTAGACTGTTATGGCAAAGAAATTATTAGAGCATTGGCAATCACTTTG
GACAATGCTTATGTCAAAGAAATTGAATTGTACGAGCAAGTTACTGACGAAG

AATTTGAGATCGTTCGCTTGACCAAGGCTGAGAGACGGCGTGAGATCATCGA
AAATATGATTAAGGAACTTTCTTTGGAAAGAACAATATATGGCGGTAGCTGC
GGTCAGTGTTTATCATGGAAAATCACAGCAACAGTTACCTGGGGCACGTCAT
TCATAAGTGAGAATGAAACGATCGGACAGCTTATTGATTCTGGATTTTGGTCG
CCTGGAATTGGACTGAATATGAATGATGTTATATTTCCACATGTGCAACACGG
TTTTCGTGGAAAGAAATTGCCAATGGTATCCTATCATAACCCACCTTGGCAAA
TAGTTAGTCATACTGAAAGTGGTGAAGTTCAATATACTGGATTAATCTTTAAT
ATTATCAAGCATCTGAGTGTAAAATTAAATTTTACATACACCGTGGTAACGAC
GACGAACGCACGAGATACCAAAATAGCCAATCACACTCGTCATGTTAAAATG
TTACAAAAAACCATAGACAAAGTTTCCGCATCGATAACCAATAACGTTCCAG
ATGCTTTGATAGAGTTAGTGCGAACAAAAAAGGTTTTAATGGCTGCATGTGT
ATATACTATTAGCGAATACCGGAAGACTCTCGTCAACTTTACAATACCTGTAT
CTATCCAGACGTACAGTCTTCTAGCTTCCAGACCAAGACAGTTGTCCAGGGC
ATTGCTCTTCATGTCCCCATATACCAAGGAGACCTGGGCCTGCCTGAGTGCTG
CTATTATCATAATGGGACCAATTTTGTACCTTGTGCACAAGTTCAGTCCTTGC
CATATGAATGTGCATGTGGCTTCAGGCTTGAATTCACCTTGGAAATGTACGTG
GTATGTCTACGGTGCCCTTCTTCAGCAAGGTGGGATGATGCTACCTTCTGCAG
ATAGTGCTCGCCTACTTGTGGGAACCTGGTGGCTTTTCGTCATGGTCGTCATA
GCAACGTACTCCGGAAACTTGGTTGCCTTCTTAACATTCCCAAGGATGGATGA
TTCGATTACCTCTGTGGATGATCTCCTTTTGCGTGACGGTCAGGTTACTTGGG
GCTTTCCCAATGGCAGTTTCCTTGAGAATTATCTACGTGATGCCAACGAAGTA
AAGTACAGTAGATTATTGGCTGGTGCTGAAAGACATAATTCATCCGAGGATC
AAGATATTATCAGAAGAGTTCGCACAGGGACACATATTTTAATTGACTGGAG
GAGTTCGTTAAGATTTTTAATGAAACGCGAACTCCTGTCTACCGGCGGGTGCG
ATTATTCACTTAGTTCAGAGGAGTTCATGGATGAGCCAATAGCCATGATCGTG
CCTCAAGGAAGTCCATACTTGTCCCTTATTAATAAGGAAATCAAAAGAATGC
ACGAGACTGGTCTCATTCATAAATGGACAACAGACTGGATGCCTGTCAAGGA
TAAATGCTGGGGAGGTCCTGGCATCAATCAAGAAGCTAATAATCACAAAGTC
AACATGAGCGACATGCAGGGTATTTTTTTCGTCCTGATTATAGGAGTCATACT
AGCAGTGACAATGTTAACCTGCGAATTCTGTTGGTACAAACGGAAGGTTGCC
AGTGAACGGAAGTTGATACGCCCTTTTGTCCCTTGA

>CcIR4
ATGAAAATATTAATTTTATTAGCCATACTTGTCAATATCTTATATTACGTAAA
TTGTGGAACTATTACGAAGAATATACTAAGAAATAATTCCGAATCAAAAGCT
CTCGCTGGATTATTGGTACATATTATTAATAAATATTTCGATGATTGCGCAAT
TATAATTATTTACGACGCTCACTTCGAAAATAGCGAATTTCTTCCACATTTCC
TTCCAAAAATCACACAAACATTCGTTCACAAAAGTCTCAACTTTTCACTCGTC
GATAAAACACAAATTGCCTTTAATAATAAAAATTACAATTATATACTCTTTCT
TGAAAATTTATCAACAATAAAGTACGCTATCGATCGCGAGACTACCAATAAA
ATAGTTATCGTCACTACAGCAACTTCTTGGTTAGTTAATGATTTCTTGAGAAG
TCACACATCCAAATCATATATTAATTTACTAATCATCACACAAAGTGCCCGTA
AGAAGAGAAATGAAAATGCATATTTGCTGTATACACATAAATTATATCTGGA
TGGCGTTGGAACGAGTAAAGCAATCTTACTGACATCTTGGATCAATAATTCAT

TGACAAAGCCACACGTTCTACTATTTCCGGATAAATTACAAGATGGCTTCAA
AGGTCACAGGATTTTAATATCAACCGCACACCAACCGCCATTCGCTATAAGG
AAAGCTATTGCCGGTACCGATGAGATAGCGTGGGATGGAATCGACGTACGTT
TAATCAAATTACTCGGTGACATATTAAACTTTACAGCAGATTTTCGTCCGCCA
ACATCAACGACACTTAGTCCAATCGCTGCAGCAACAAATGACATCAGAATGG
GCTTGGCTTCTATAGCTATTGGTGGAATTTACAAAACGACGAACATAACCTCT
CGTTTCGACACTTCTTTTGGTCACATTGAAGACTGCGCGTCATTTTTATCACTT
GCATCTACTGCATTACCCAAATATCGCGCCGTTCTTGGTCCTTTTCAAATAAC
TGTCTGGATTCTATTGACATGTTCATATTTACTTGCAATTGTTCCGTTGACAAT
GAACAGCGAATACTCTCTACTATCTCTGCTGACCCATCCGAGCAGGACAAGC
GACATGTTCTGGTTTGTGTTCAGTACATTTACGAACTGCTTCACTGTGAAAAA
TCCTTTGTTGAACTACGGACTAGGAAAGAATTCTACGGCGATTCTAATTGGTA
TTTATTGGGTATTTACTATCATTGTAACGTCTTGCTATACCAGCTCTATTATCG
CCTTCATCACTGTGCCCGTCTATCCAACCGCAATGGAAACGGTTGAGCAACTT
CTCAGATACAGATATCGAATTGGAACTTTGGATCATGATGGATGGGAAAAGT
GGTTCGAGGATATCGACGATCCAATGGCTCAGAAATTATTAAAAAATATCGA
ATGTATTCCTTCAGTCGAAGATGGCGTCAGGAATGTCAGTCGAGCTTATTTCT
GGCCATATGCATTTTTAGGGTCTAAAACATTATTGGAATACATAGTCCAGGCT
AATTTTACTCCAAGCTGGATAACCAAGAGATCTCTTATGCATATCAGTCACGA
GTGTTTCATGAAATTCGGCGTGACTTTCGTGCTACCACGTAAATCGGTTTACA
CGGAGGCTTTCAGTAGCGTGATTTATCGAGCACAACAGACCGGCCTCGTTGA
GAAGATCGTTCAGGACGTTAAGTGGGACATACAGCGTACAGCAACGGGAAA
GTTACTACAAGTCCAAAGGTCCTACGCTACGCCTGACACCGGTTGA

>CcIR7
ATGATTTCTGTGCTTCTTGGAGTTTTAATAATTAACGCGTGTCTCCCTGTTTCT
TCAGGAATGGACGGAACTCTCCTGAAATTTCTTGTGGACGCGAGCATACTCTT
TGCTCCTACGAGGATGACCGCCTTCTTGTGCATCGATCAAAATGATGCCTTGC
ATTTATTAAGAGCAATGTCAAATGCTCATTTATCATATGATATTAGAATATCT
AATCACAATCTCGATATGAGAGTTTCCCGCGAGAATCGACATTACGTCTCGTT
CTTACTCGACCTCAACTGCCCCGAGGCCATCACAATATTGCGGCAGGCGAAC
GAGAGTGCATTATTTGTAGGGCCGCTCAAATGGCTCATTCTACAAGATTTAAA
GGATAGGAAGAAACAAAACGATTCATATCTAGATAGTGTCTTTGGAAACCTC
GCAATCTTCCCGGACAGTGAGTTAGTATTGGCCCAAAGACTTGATAAAAATT
CAACGAGATTACTCTCTCCATATCGACCGAGTCCTTATCATAAATTAATTATT
GAGGAAATAGGATTATGGGATTTACAAAATGGAGTAACTCTTCTGAATCATG
ACGTGTCATCGAGAAGAAGAAGGAATCTTGCAAACACTCCGCTAAAATCATG
TTTGGTGGTGACGGATCCTGATACAATAAATCATTTAACAGATTATCAGGAC
AAGCATTCAGATACTATCACAAAAACTAATTACCCATGGATAATGCACCTCG
TCCGAATGATGAATGCTACAGTGACTTTCTCCATCGCAAATACTTGGGGGTAC
CGAAGTAAAAACGGCTCTTGGAGCGGCATGATCGGTCAGTTAGCCAGGGGTG
AAATCGATATAGGTGGAACGGCGACATTCTTTACAGCTGAGCGTATTGGGGT
AGTGCAATATATTTCACTCTACACTCCTACTAGAGCTAAGTTTATTTTTCGAC
AACCACCATTGTCTTTCGTCAGTAATGTCTTCACACTGCCATTCCGACGATCC

GTTTGGATTGCTATTGGAGTATTTATGATCTTGATATTTTGCTTACTCCATCTG
ACATTGAAATGGGAATGGTTCAACACCTCGGCAGAAGAAAGAGTAAGACGT
CATTGGTCAGGGAATGTCTATCAGCCAACTATGAGCGACAATCTCCTCATTCT
TGTCGGCGCTGTATCTCAACAAGGTTCCTCTTATGAGGCCAGAGCAATACCA
GCTCGTATTATTATTCTCATGTTACTTGTGGCTACCCTAAGCCTTTACGCTTCT
TACACGGCTAACATAGTCGCATTGCTACAGTCCACTACTGATTCCATTACCAC
TCTCGATGACTTACTACACAGCCCATTGAAAATTGGAGTTCATGACATTGTAT
ACAATCGCTACTACTTTGAGTCCTTTCAAGATCCTATAAGGAAGACAATTTAC
GAGCAGAAAGTGGCTCCAAAAGGACGAAAGTCTGCGTGGATGAACTTGGAG
GAAGGCGTAAGTCGCCTGCGACAAGGACTATTTGCATTTCATGGAGAACTCG
GCTCTGTGTACAAAATAGTTCAAGAAACTTTCAATGAAGACGAGAAATGTGG
TTTTCAAGAAATTGACTATTTAAACGTTCTCGACCCATTGATAGCCATTCAAA
AAGAATCTCCTTATCTTGAAATTATACGTATAGGAGCCCTTAAGATACATGAA
AATGGCTTGCAGACCCGAGAGATAAGACGATTGTACACGAAAAAACCTGTAT
GTCATAAACATGCAAGTTTTGTAAGTGTAGGTTTGACTGAATGTTACGCCGCG
TTTATCACACTTAGTTACGGTACTCTCATCGCTTTTACTATTTTCATCCTAGAA
ATCATATGGTATAAAAGAGGAATAGTCGAGGACGTAGTACAGCCGTTTGATT
ACGTAAAATAA

>CcIR8
ATGGATTTACGGTTCTTGCTGTTTGTATTTGTCTGTGCGATACGATGCAGCGA
CGCAAATGGATTGTCTGCACAATTTTTACAAGATTATTTTGCTGACAAATTTA
TACATCAAATTGTTGTATTTGCATGTTACAATCATCTTGAAAGAGTGAGTCTA
TCTAGACAATTAATGTTGTCCGACACAAAGCTGTCTTACATTTCAATTTCGGA
AAATATTAACATGAAGACAACTTTAGCAGTGAATTATTATAAGCTTGGTATA
GTTCTCGATTTGGATTGTACAAAGAGTGAGATAGTGCTTGACAAGTTTCAAA
AGGAACGTCTGCCCTATAACGAATCTTACTTTTGGaTGGTGATATCCAAAGAT
TCTATGCCCCCGATCGAAATTCTTCGGAAACTTCCTCTGACTGTGGAGACGGA
GTTTACTTTGGCACTTCATCACGATGAAATCTATGTACTTTATGACGTTTATA
ATCCAAGTTATCGTCACGGTGGTCGTTTAAATGTTACGTATATGGGCGTTTGG
AATCCTCTTGATGGCCTCTCTATAAAGCTTACCCAATACAAATACAAAAGAC
GTGCTGATTTTCAAGGCTTACATCTCAACTTTTCGATCGCGCTATCGAATCCG
CCTCTACCGGATTTAATGACTTATATCTCCAATCCCATTAATCGCCATCTTGAT
ACAATGCATCGATTTAATTATGCTCTTCTGCTACAGCTACGAGATTACTATAA
TTATACAATGATTTTGACGCAATCAGAAACTTGGGGATATTTAATAAATGGC
AGTTTTAACGGGTTGCTGGGAGATATGATTAAAGGACTTGTTGACGTGGGTG
TAACTCCATTGCAATTTAAACGCGAGAGAATCGACGTTGCTGAATTTACAGT
ACAGACTTATTTGGCCAGGCCTAGTTTCTTTCTCCGTCATCCGAAGAAGACCA
GCGTACGCAATGGATTTCTAAAGCCATTCAGGGAAGAAGTTTGGTGGATCGT
TGTAGCAGTAGGATTTATATACTGGTTAATACTATGGTTAACCAGTAAATTGG
AATTATATTACAGGAATAGAATGCCTTTTAGTAAAATAAATTTAGAAGTGGG
TTTCGAAACTGGTTTTATTTGTATGGCAGCCATTTCTCAGCAAGGCTTAACCG
ATAGTCCAAGATTGTATTCAGGCCGAATAGTGTTTTTGTCGTTATTTGTTTGG
GCATTATTGCTTTATCAATTTTATTCAGCAAGCATAGTTGGTTCACTACTCGC

AGAGCCACCTCGTTATATAAATGACTTGAAAGATCTTCTGGAAAGCAATTTA
GAAATTGGCATTGACGATATTGCTTACAATTATGACTTCTTTGCGACAACAAG
TGACCCCGTCGCCTTAGAACTTTATAAGCGGAAAGTGGCGCCTAATAAGTAT
AGAAAAAAGCCGGCTTTTTATAATCATACAGAGGGATTGCGGAAAGTGAAGA
AAGGCGGTTTTGCTTTTCACGCAGACGTGAATACTGCGTATAAGATCATTACG
GATACATTCTCGGAAGATGAAATTTGTGATTTGGTGGAAATTCAGTTATTTCC
AGCTAAGCACGTTGCAACTGTGACTAGCAAGCATTCTCCGTTTAAGAAGATG
GTGACTTATGGAATGAGACAAATTATGGAACATGGACTTTCAGAACGTTTGA
TGCATGTGTGGCAATATAGACGACCAATTTGCCCGAAAACTCATAGTTCTAC
ACCTGTTCCTGTAGCATTGGGAGAATTTCTTCCAGCTACTTTGCTTCTGTTGAT
GGGTTTAATTATTGCTGTTGCTGTTTTGAGCGTGGAGAATTTTTGTAAACGTT
GCAAGTATCCTGAGCAACTTTCGTGTAATACATATGTGATACCAGGATTCAA
AAATCAGGCAGATATCAATATGTATATTCAAGGGCACTCCTAA

>CcIR9
ATGTCTGAGACTGCAGGACATTATACCTATACAAACTTATACAATTGGCTGAT
TCTTGGAGAAAGTCTTGACAGTGTGTTGCCTCGTTTAAATGACAGTACCTACA
GTGTGGTAACTGATCTTGTGATAGCAACTCCAATACTAGACGGTTACGTACTC
CATGACGTTTACAATCCTTCGAAATACCACGGTGGATCATTGAACGTCACGG
AGCTTGGCGACTGGCATAGAAGTTCAGGCTTGAATATTACACTGACACACAG
TAAATTTAATAGACGTGCAAACTATCATGGAATGATTCTGAGAGTTCTTACTG
TGACGCAGTTTAAACCGAACAATACTCGACTGGAGGATTATTTAGAAGATTA
TTCAAACAAGTCAATCGATCTTATGTCCAAATTTGTTTATCCCCTGATGCAAC
ATTTATGTGATATTTTTAATTTCACGATGGACCTCTCTACAGACTTTTATTGGG
ACGAAATCGCAGAAAATGGATCTCATACAGGAATGATGGGCATGTTAAATCG
TGGATTTGCCGACATAACAGGAAGTCCTGGGGCCATTAATCAAAAGAGATTT
CCATTGGGTGAATTAGTCATGCCAATTTGGAATTTTCGGACATGCTTCTTCTT
CTTATCTTCACCGTTGAGATCAACAAATGAATTGCTACAACCATTTTCCACTG
GTTCTTGGTATGCTACTGGAGCCCTAACCATACTGATCGCTTTGATTCTTTCTA
TATCATTGGCTTTCGAGAAAATTGATCATTGGATTCGGCAGTACGGAATCGCT
TTATTTTCTACCATATGTATACTCTGTCAACAAGGATCCTCGGTGATTCCGAA
CTTCTTGACTGGTCGCTTGGTGTTCCTGCAATTGTTGTTCTTTAGTCTACTAAT
TTTCAATTATTATTCCGCTTGTATTGTTTCGGTTCGCTTAAATCAACCATTGGA
TAAGATAAACGATTCGCTTAATGAACTTGCCGAAAGTAACATGGACATTGCA
GCCGAGCGTATTGTTTACTATGATTATCTACTGGAGACATCCACAGATTGGGA
AGTGCAAATGCTGAAGGAGAAACGTTGGTCAAAGATGTCGGATGCTAAAAA
ATATTTACCACTAGAAGAGGGTTTAAAACGTATTACTCGTGGTGGTTTTGCTT
ACCATTGCGACCCCAATAGAGGATATCCTCTTATGGAAAAATTATTCACTTCG
GAAATGATATGTAGACTGTCAGAAGTGCATCTGTATCACCCACAAGGACTGG
GAGCATGGGTGGGTCACCACAGTCCGTTTCGCGAAATCGCAAGAGTCGGATT
AGCAAAGATCAATGATATGGGACTCAGAAAGCGACAGGCAATTCGTTGGAGT
GCAAGAAAACCAATCTGTGTAAATGTACAGCAAACCGTGAACAGTGTGACAT
TCACTGAAATAAAGGCGGCAATTTTTTTACTATTGCTGGGAGTTATGCTGTCC

CTCATAATATGCCTGCTAGAGAATATCATATACAACTGGAATACGAAGAAAC
GGTACGAGAGAGATAATTAA

>CcIR11
ATGTGCCATTTCAAAATGGGAGGAATGCGGATGCGTTTGTTAATTTTAAAATT
TTACGTTTTCGTATGTTGCGCGTGTTTCGAGACTCAAGAAATTTGTAAAATGG
TTGATCTGATAGCACAACGGTTTTTAAACGACACGTGTCTTGTCGTATTTTCA
AATAATGGTTGCGCATTCAAGTCTGTGATACCTGCGCTTTATTTTGACCCGTA
CGACACTCTTTTGGAAAATGAAAATATGGATAGCGAGACTTACAAAACAACT
TATCTAACGGAAATTTTGGAAATTGGTTGCAATGGGTTTGTTATTCAAACTAA
AAATCCGGCACGGCAATTTGTACTTTTGGAAGAAACTTTTGTTCTTGTCAATA
ACAGGGGCAATCGAAAATTAATTTTCTTGCCGCTACTGGAAAAATCCTCGTTC
AAGAATCTTGATTCTCATTCAATAGTTGAGACGAGATCTGCCAATTATGTTCC
AAATTTAATAATGGTGGAACCTTCAGTCACCAAAAATATTTTTAATTTGAAGA
CTCATTATTTTGGCGGGACTAGTAATCAGTCAAAAGAAATTATCGTCTTGGAT
ACGTATAACATTACCTCAGGGAAATTTTTATTAGACGCAAACTTGTTTCCTGA
TAAGATAAAAGATTTGCAAGGAAAAATTTTGCGTGCTGCTGTTTTTCACTATA
TGCCATACGTCGTTCTTACTGGGAATGACACTCAAAATAATTCACAGATCGAC
GGCACTGAATGGCGGATGGCATTGCAATTTGCAAAATTTCATAATTGCACTTT
TGAGCGTTTTGTCAGCGATGAAGAATCGTGGGGAACAATTTTTGATAATCGA
ACAGGAAACGGCGTCTTAGGTGCCCTCGTGGAGGACAGAGCTGACATTGGAT
TTGTGGCACTATATACGTGGTATCATGAATCCTTGTATTTGGATTTAACAACG
GCGTACATAAGAACTGGAGTTACTTGTTTAGTTCCACCGCCACATCGTTTGCC
ACTGTGGCGCTCGCCAATTTTACCGTTTTCCCCAGAACTATGGGCGGGAGTTG
TCACAACCTTAATTATTTGTAACGTAGTCAGCTTTTTACTTCACTATGGAGCT
AATTATATTTTTTTCGAAGATAATAAAATAGTCTCCGCCATTGCGGCTTTAGA
AGTTGCCGGAATATTTCTTCAACAAGGATTAATAAAAAATCGGAAAGAACTT
TCGTGGCGCACAATGGCATTTAGCATGTTAGCAGTGTCGGTATTGCTGTGTAC
CGCTTACAGCAGTAGCTTGGCATCAGCCATGACAGTTCCCAGGTACGACGCC
CCTATAAATAGCGCAGAAGACATGGCGAATCGTGATCTAAAGTGGGCAGCCA
CTCACGATGCTTGGATATTTTCAATTTTGCAAACCGACGATCCTGAATTGAAA
ATCGTATTGAAGAATTTTCGAATTGTAAAATTCAATGAGATTATAAAACGCA
GCGCAGAAATGGAATGGGGTTTCAGTATAGAGAGACTATCGTCTGGACATTT
TGCGGTTCAAGAATACCTGCAGCAGAAGTACGTGGAGAAAATGCACTTGATG
GCCAGTGACATCTATTGGGAAAATTGCGTATTCATGTTACGGAAGAGTTCAG
CGTACACGACATGGTTTAGCCTGAGGATTTTGCAAGTTCTGGCTTCCGGTATA
CCAGCCCGTTGGGAAGACCAAGTTGTAGCAGAGTATTTGGATTATGGGACCC
AAAAAGCCGTGTCGAAAGAGACCTTCGTCGATCAAGGACCGATGAAGATGA
CAGTGGCTCAAGTTCAAGGCACATTCTTTATCTTATTTCTGGGACTGTTACTC
GCTCTCTGTGTCTTCGTTGGTGAGATCACTTTGAAAAGATCTGTACAAATATA
A

>CcIR15

ATACCTGGTACAAAGGAAGATTTAGATACGGAGTTAAGCTCTTTGATGAAAA
AACAAAAAATCTTCATGGCAAAAATTGCGGAAAAGTCAATAAATTTTAAGGC
ATTGGTATACGTGCCACCGGATGTAAAAACAGAACGTTGGGGTTCACTAAGA
GATAATGATTCGTATACTGGATTATTGGGAGAAGCAGTTTCAAGAAATGCTG
TATTTTTTCTTGGTGATCTTCATTATACAACAGAACATCTTCATGTGCTCGACT
TATCTGTTCCGTATCATACAGAATGCTTGACCTTTCTTACACCAGAAGCTTTG
AATGATAATTCATGGAAGCTGCTTGTTCTACCTTTCAAGTTGTACACATGGAT
TGCAGTACTATTTACTTTATTTGCTGGTGGATTTATTATACACTTCTTTGCTAT
TTATTATAAAAATCATGTAATGTTACATAAGAACCACACAATCGACAATGCT
GCAAAGAATGGAGCTTTAAAGGAACTCGTTGCACCAAAAAAAAGTATTAGTA
AGAAAGCGAAATTTAATCAGGAATCTCGAGGCCTACGTATATTTATGGAAAT
ACAAAACAGCATTATGTATACATATAGTATGTTAATGCTAGTATCTTTGCCAC
GTTTACCTGATGCCTGGTTTCTTCGTATATTTATTGGATGGTGGTGGCTGTACA
GTATTTTAGTATCTGTTGCTTATCGTGCAAGTATGACTGCAACTCTAGCTAAT
CCTGTACAAAGGGTAACTATCGACACGTTAAGCCAATTGGCTAATAGTCCAA
TAGCCGTTGGTAGTTGGGGAGAGGGCCGAAAGCAATTTTTCTTATCATCTTTG
GATCCGAATACACAGAAAATCGGAACTAAATTCGAGGTGACCACTATTGAGG
ATGAGGCTATATCACGAGTAGCAAATGGAACATTTGCTTATTATGAAAATAT
CTACGTTTTGCGCGAAGCTCGTGCTAAACGACAAGCTTTGGAGGTGAAACGA
AGAAAAAGTGCATCTAAGCTGAATCGTACATTTAAGCAAGATCGAAATTTGC
ATATAATGCAAGAATGTGTTATAAATATGCCGATTGCCCTTGGTCTGGATAAA
AATTCTCCACTGAAACCAAACGTGGACCGGCTTGTAACTAAAAAACTAATTG
TATTACCGCCACATATAACTTTATCAGTATTGAAATCATTGTATGATGTATTC
CTGCAACAGGTACAACGAGCTATAGAAGCCGGTTTAGTTCAGAAATGGTTGA
ACGACGTGATGGAATGGGCCAAGAATTCCGAAAAACCTTTAGATACTAATGC
TCCAAAAGCTCTTATAAATCTTCACAAACTTTACGGAGCATTAGTTGCTCTCG
GAATTGGATATTCTCTTGGCTTTATAACGTTAGTATATGAAATACTACATTGG
AAATACATTGTTCTCAAAGATTCGAAATACGACAAATATTATCTGGATGTGTT
CTATAAGTAA

>CcIR17
ATGACGGGCTTTGAATTGATGCTCATCGAGATTTGCTCGATTCTGCAAAATGT
TACGAGTAATAACAGCTGTCTGCCGAGCCAAGTCAAAAAAGAAATTCCATCC
CGCTATGACGATCAAGTATTTTCTGCCGTTGCGCCATATTTGTCTGAGACAAG
AAATAATTTTCCTACGGAACTTCGAATAACGACGTTACAGGATCCGCCATTGT
CGGGCACTAAGCTCGTAAACGGTAGTCTGGTCGGCGTTGGACATGCATTTTA
CATCTTCAATCTTATGCAAGCAAAATTAAATTTCACCTATAAAATAATATTGC
CAGAAAGAAATATATTGGGCGATGAAAGTTCCGGTGTATTTGGATTACTTTAT
ACAAAGCAAGTCGATATGGCCGTGGCGTTTTTACCAGTTCTCTCAGAAGCTCA
TAACACCGTTAAATACGGTACTTATTTAGATGAAATGGAATGGACCATACTC
ATGAAACGTCCTCAAGAATCAGCAACAGGTTCCGGGCTCTTGGCACCTTTTG
ACACTACAGTCTGGATTCTGATTTTGATATCAGTCATAATAGTGGGACCTACA
ATTTATATTTTAATTTACATAAGATCGAAAATCTCACAGGAAAAATCAGAAC
ATTACTCCGTAGCGACTTGTATGTGGTTTGTTTATGGTGCATTGCTTAAGCAA

GGATCAACAATCTCCCCAGTAACAGACTCAGCCAGATTGCTGTTTGCCACCT
GGTGGATATTCATAACAATTTTAACGTCATTCTACACTGCGAATCTCACAGCC
TTCTTGACATTGTCGCAGTTTACGTTACCGATCAAGTCGGTCGATGATTTAGC
CAGTAAAGGTCAAGTATGGGTAGCAGAAAAAGGTCTCGCATTGGATATGATA
GTGGCAGAGGGTAACGAGTTGCAACCACTGAATCTTTCGACCAGCAAAGGAA
AAGGTCATTTCGTGGATCTACGAAATAAATCATATTCAACCGTCATTCGTCAC
ATTGATCAAAATAAAATGTTCCTGGGTGAAAAACATTTTGTTCAAGGATTGAT
TTTTTATGATTACATAGACAAAAGTCATAAGGGGATCGATGAAGGTAAACGC
TGCACATATGTAATAATGCCAGGACACATTTTATTAAAGAGAAGAGCCTTTG
CGTATCCCTTGGGCAGCCTAATAAATGTCGTCATGGACAAACATATCTTAGCG
TTGGTGGAGTCCGGTATAGTGAAATATATAACAACGGCTTATTTACCTGGGA
CAGAAATATGCCCTTTGAATTTGGGATCCAAGGAACGTCAATTGAGAAACTC
GGATCTGACTATGACTTATAAAGTCGTCTTCATCGGGTTCACCGTGGCGGCAG
TCGCCTTCCTTGTCGAGCTCTTGTATCTTTGCAAATGTAAAAAGAAGAAGAAG
AAGAAGAAATGGTAA

>CcOBP1
ATGTTTAAGACACTCGTGACGGGCATCGCGCTCCTCCTCGCGTCTCTTCATCG
CGTGGAATCGGCTATGACGATGGAGCAAATGACGAAAATAGCTTCGGGCTTC
AGGAATGTCTGTCAGCCAAAAACCGGTGTTGACATGGCCATTCTCGAGGGGA
TGCAGAAAGGGGAATTCCCGGAGGATAGGAAATTTCAGTGCTACCTCAAATG
CGTTATGGGTATGTTGCGAACGCTGCGAAATGGCAAAGTCGACCTCGATATG
CTCCTGAAACAAGTGGACACCATGGTGCCCGAGAACTTACAAGAGCGCACGA
GACACGTTGCCAGACATTGCTCCACCGTGGCCACCAGCAACGACATGTGCGA
GATCGCGTATCAGTTTGTCAAATGCAACTGGGATACGGATGCCGAAGTTTTCT
TCTTTCCCTGA

>CcOBP2
ATGAGAGTTGAAATCAAATTAAATAAGCATATACGATCTACAATGAGAATTT
TCGGTGTGGTAGCTGTTGCCTGTATTCTTCAAGCATCGATCGTCCATTGCGGG
ACGCGGCCCAGCTTCGTTTCTGATGACATGATAGCTATGGCTGCAAGTGTCGT
TAACGCATGCCAGACTGAGACAGGAGTATCTACTGCCGACATAGAAGCCGTA
CGAAATGGACAATGGTCCGAGGAACGTTCCCTTAAGTGTTACATGTACTGCC
TTTGGGAACAATTCGGCCTTGTGGATGATAAGAGGGAATTGAGTCTCAATGG
GATGCTGACATTTTTCCAACGTATACCAGCATTCCGAGCTGAAGTATTGAAAG
CAGTAAACACTTGCAAGAGGATCGGTAAATATTTCGCTAATGGTGACAACTG
CGAATACGCATATACATTCAACAAGTGTTACGCGGAACAGTCACCAAGGTTT
TCAGTCAGTTATTTATTATATGAAAAAGATGGATCGACTAGTTTCATTAACGA
CATTCTTAATATTATTAATCGTCAATACTCCATTAATCGATGCTAA

>CcOBP3
ATGCGGCAGAACAGACAAAATTGGAATAGCAAATCAATGGATTATAATAATA
CAAGCGATACTGGTAATAGAATGCGAAATTCTACTGGCAATATGAATCATAG
AAATTCAATGAATACTAGAAAAATGGACAATGATCAGTTGCATTCTTATAAT

AATCAAAAGTACGATGATCGGCATGATCAAGCGTGCGTTATACAGTGCTTCT
TCAATAAACTCGATCTGGTCGACCAGAAAGGATTTCCAGAACGTATGGCAGT
TACCAATATAATGTCACAGAATGTACATGACCCGGAATTACGTGATTTCATTG
AAGAGTCTGTAGTGGATTGTTTCCGTTATCTTGATATGGATATGAATAAAGAG
AAGTGTCAATTCTCTCAGAATCTCTTGGGTTGTTTGGCTGACAAGGGACGAGA
AAAATGCGACGATTGGGAAGATGAGCAAGAATAA

>CcOBP4
ATGCGGAGATTTCTGTTCCTGGTGGCTCTTGCCGTGTCGCACGAGATGATCTA
CGCAGACCTACCGGATTGGGTAACACCGGAAATGACAGATATGGTGAAAGA
CGACAAGATACGCTGTATGAACGAACACGGAACGGACGAAGAGATGATCGT
GAGGGCAAATGATGGTGATCTGGTGGACGATCGGAAGCTCAAGTGTTACATG
TACTGCCTGATGGAATCGTTCGGTTTGGTCGATACCGATGGTGAGTTCGAGAT
GGAACTTTTGATTGGCTTCCTGCCGGAGAACCTTCAAGACGTAGCCCGGAAC
GCGATGAACGCCTGTGCCGACGAAGGCGGTGATGATCCGTGCGAAAAAGTCT
ACATGCACTCGCTCTGCATCTACAAAAGAAATCCTCAGCTATGGTTCCTGATA
TAA

>CcOBP5
ATGAGGGTCCAGGTATTACCGTTTTGTCTACTTGTGATGATTTACACAACTTG
TTTCACTTTGGCATTGAAAATACCTATGACAATGCAGAAACATCTGAAGCCCT
GTCAAACAGAATTGAACCTAACTGATAGCACACTTACGAATATCGAAAAAAA
TGACGAGAAAAGAGAATTGAGATCGTGTTTGGGTGCTTGTATGTTAAAGAGG
GAAGGAGTTCTTAATGCTAACGGGACGTTTAACAAAATGAAAGCCATTAATA
TAATTGATCCTCGTGTATCAAAAGAGGATGCCTATAAGGCTGTTGATAATTGC
CAGAATGAAGTATCCAAGAATATATGTGAAACGGCCGAACTGCTTTATAAAT
GTTTTACGGTATACAACGTGTTATAA

>CcOBP6
ATGTTGCTTTACGTTTGTTGCGTATTTTTTCTTCTCGGCTTTGGCAACTCGCAC
GGAGTATCTGATGAAATGATGGAAATGGCCAAAATGTTACATGATTCATGTG
TCAGTGAAACTGGTGTTGACGAAGGTCTGATCGAAAAATGTCGCGATGGAGT
ATTTACAGAGGACAGTAATTTGAAGTGTTACATAAAGTGCATCATGGATCAA
GTATCAGGAATGACTGATGATGGAGAAGTTGACGAAGAAACTGTGATCAGTA
TGTTACCTGAAGAGATGCAATCAGAGACAGCTCCAACTATTCGAGAATGTGG
AACAGTTAGGGGATCAGACAACTGCGATACCGCTTTTCAAACTCATAAATGT
TATTACGCTAAAAATCCAGAAGTAAAATTACTTTTTGGTATAATGAAATCGCA
ATTGACGCCACGTTATAGTCATCATAGGGTTACCACATTAATATGTAATCGAT
TTCCTTGGTTTCTATTGGGTTAA

>CcOBP7
ATGGTGAGTTCAAGAAGGAAACCTTTGACATTTATTACTCCAAGGATGGTAT
CACCAAGGAACAAGCTGAGAAGTTCGCCAAGGAGTGCAGCGCACAGAAAGC

TGACGATCCTTGCATTCGACCAAACCACGTGGACGGAAGTAAAGAAGTATTG
CAAGCAGCAGTCAGCTGTTCCTCAAAAGTTTTCCAACTACGAGAGGATATTCT
CAAAGTCTGGATACTTTCCTGCGGATGAGGAGTTCCAGTGTTACTTGGCTTGC
CTCGCCGTAAGATCGGACCTGATGATCCTGGGAAACGAGTTGAGCATCGACT
GGCGCAAACTCAAACCACTTTTCCCCAAGGGCAAAGCGTTCGGGTGTCAGAA
ACGTGCAACCGAACAGGATCCCTGCAAGAAAGCGTATCAATATTTCAACTGT
TTACGAAAGGGAAACAACGCGTTCCAGGAGATTGATTTCTTCTGA

>CcOBP9
ATGATAAAAAGAAAAATGTTGCGATACGGTATATTATTCTTGAGCATTTTATT
CGCCAGCTCATTTGACGTTAGTGCTGCTGCCATAAGCGCCATTCAGAGTGACC
TTGCTGTTCTTCGATTATGCAATGCATCAAGTCCTGTGTCACTTCAGGCTGTC
AATTCAGTGCTAATAAATCGTGATATGAAAAGTTCTGAGACACAGACGCAAG
ACTTCAAATGCTTCCTGCTTTGCCTTTACGTCGAATACGGTTGGATGGATCGC
GAAGGAGGTTTCAGTTTGCATCACATTAAGTTGACACTGGAATCGTCAAAGC
TGCCGGAATATCGTATCAAAAAGTTAATTTACAGTTGTACCGCAACAGAAAT
AACAGATCCTTGTGAGAGAGCATTCAACTTTACTGAATGCTTCTGGTCGCACA
GTGACAAGGAAACGAGCCCAGAAGCTGAGATCCCTGAGAAAGAAGTCTCGG
ATAACCAGAATGGATTTTACTATATAAACTAG

>CcOBP10
ATGAAGACCTTCCTGGCTGTATTTGCTTTGTGCATTGCTGGTACTTTGGCTGC
ACTGTCCGAAGAACAAATCACCAAGTTGAAGACTTATAAAGAATCCTGTATT
GAGGAAACTGGCGTTGATCCTTCTGTTATCGAGGCTGCCAAGAAGGGTACCT
ACTCCGATGGTGATGAAAAGTTAGCATGTTTCTCTGCTTGCTTGTTAAAGAAA
ATTGGAGTTATGGATCCCGATGGTAAGATCAACGAAGAAGTGATTCGTTCCA
AAATTCCTTCCAGTATCCCGAAGGACCAGGCAGATAACATCATAAATACTTG
CAAGGGACTAACTGGTGCTAACGCTTGCGAAACTGCAGGAAAAGTCCTAAAA
GTGTTATCTCGACCAAAACATGTTCCTCTTCTCATGAATTAA

>CcOBP12
ATGAAAGCCCTCGAGGGATTCAGCCGCATTCTCATCACCTTTATAGTTGCCGT
TATGGTGATCGTCAAGGCCGATATTAGGCGAGATTGTCGAAAGCAAAGCGGA
GTATCATGGGCTGCGCTGAAGAAGCTCAGGGCTGGCGACTTCAATCAGGAAG
ACCACAGTGTCAAGTGTTACCTCAAGTGTTTCATGGTGAAGAATGGCATTATG
AGCGAGGACAATCACGTCGATGTGGATAAGGCCTTAAGGCATTTACCAAGAA
AGCTACAGGAGCCTTCACGAAAAATTCTCGCACGCTGTAAGGACTCGGCCGG
CAAAGATTCTTGTGACAAAGCTTTTCAGATTGCCAAATGCTACTTCAAGTCGC
AGCCCGGGATCCTCAAAAATGTCTCCTTCGTGTAA

>CcOBP13
ATGAGAGCTGTCCTGCTGGTAGCGACGCTCTTCCTGGCGTTTGGATCGATCGA
GAGCAGAATGACATTGGCGCAATTGAGAAATACGCTAAAGGGTATGAAGAA
AGCCTGCACATCGAAAATCAGTGTTTCGGCAGAGGTACTGGACGGTGCTTCT

AACGGCATCTTCCCACCCGATCCGGCATTGCAGTGTTACTACAAATGCCTCTT
CGAGATGATGAAGCTGATGAAAAACGAAAAGCTCCTCATTGACGTAATGATA
ACGCAATTAGATAAGATTGTGGCTCTGGACGTACTTGATAGAATGAAGGCGG
CTACGCTAAAGTGCGTCGAAGCAATCGACAGTGAAGATCCGTGCGTGCTGTC
CTGGCAATTCACCAAGTGCTTCTACGAAACGGACCGTTCGGTCGTCTTTTTCC
CCTGA

>CcCCE1
ATGCGTCGTGAATTCTTTTTGACTTTCCTAGGTGTCGCCATTGTTGCGATGGC
GGGAAACGTCAAAGGAGATGAAGATCCAGAAGTGACAGCTCCTATTGGGAA
AATTCGTGGTTCCCATATGACGAGTAGACTTGGTAAAAAAATTTATGCATTTC
GTGGTGTTAGGTATGCCGAACCTCCCGTAGGACAACAACGATTCCAGCAAGC
CATAGCAGCTAAGCCTTGGTCTGATGTCTTCAATGCTTCCGTCGAAGGACCCT
CGTGTCCTCAAAAATACACACAATTGGTCAATGAGGATTGTCTTCGATTAAAC
ATTTACACGACAAAGTTGCCGCCGAAAGTTGGAACAAAGGGAAAACCTGTGA
AGAGACCAGTTGTCGTATTTTTTCATCCAGGAGGATTTTACGGGGGCTCCGGA
CAAGGTTATATCTTTGGACCACAATATTGGTTGGATCAAGATATTGTATTGGT
TACAACGAATTACCGTCTTGCGTCCTTGGGTTTTCTAAGCACCGGTGATTCCC
TGGCTCCTGGCAATCTTGGTCTTAAAGATCAAGTCGAGGCTTTGCGTTGGATT
AAGAAAAATATTGCCAGTTTTGGAGGCGATCCTGATTCCGTGACTATAACCG
GTTACAGCGCTGGCTCTTGGAGTGTTACGCTGCATATGATTTCACCAATGTCT
AAGGGCCTATTTCACAGAGCAATTGCAATGAGCGGTGCAGCCACTGTACAGG
AACCACTTCCAACGCAGCAGAAACATTTGGCCAAAAAACAGGCGGAACTTCT
TGGTTGTCCAACTGACACCACTGGAAATATGTTGATCTGTTTGAACACTAAGA
CGATTGAAGAGTTTGTTGATAGTTATGAAAAATTCTTTGAATGGCATCGCGAT
CCTATTTTAGTTTGGTCTCCGGTCGTAGAACCTGAAGTAAATGGAGTAGAAA
GATTTCTTCCAGCTCAACCTGTTGATTTAATTCGACAGGGAAAGATAAACGA
AGTGCCATTGATTATCGGTGTTACAAAGGATGAGTTTGGTGGTGTAGTTACTT
CAATAATTGAAGAAGCACGCAAAGGAAATACTTCTACATTTGATGATCTTAA
TCAGAATTGGGATACTATTGCACCCATTAGTTTTCTTTATGAACGTGGTACAC
CTCGATCCAGAAGGATAAGTCAGGAACTGAAAAAGTTCTATTTGAATAATCA
ACCAGTTAGTCTAGACAATGTTGATGGACTTGCTCAACTCTATGCTGACGCTG
TCATTGGATTTTCTGCTCATCGTTTTGTCAAATTGATCAGTGCTGCTTCGGCGA
AACCAGTTTATTATTATAGATTTTCTTATCAAGGACGATTCAGTCACAGTGTA
TGGAGTGACACTAAAAAACCTTACGGTGTGGTCCATCATGATGATCTATTGTA
TCTCTTCTACATAAGCTTTTTCCCCTATTTCAACGCAACGGATCCTGAAGTCGT
CACTGTTAAACGATTAACCACAATGTGGACCAATTTTGCGAAAACTGGACAA
CCCATACCGACTAACAACGAAGACTTTAAGAACGTCAAGTGGCCGATATATA
CCGATAAGACTAAAGAGTATCTGGATATCGGAGATAATCTCGTCACAGATTG
CGGCCTCTACCGAAACAGAATGAGCTTCTGGGATAACTTATTTCCGCTTAAAC
CTCATAATTTTGATGATACAGGCGTTAAGCAATAA

>CcCCE2

ATGTGGTTCGCAAAGTTCATAGTTTGCTCTCTGGCTATAAGCTGGACGGGTGC
CACCGAGGAAGTCCAAGTGGAGATTCCTCTGGGTTCAGTGAAGGGTCTCAAA
AGTACTACAGTGCTGAGCAAACCATATTACAGTTTCAAAGGGATACCATACG
TGAAACCGACTGTCGGCGACAACAGGTTTAAGCCATCGCAACCGGCAGAGCC
TTGGTCAGGGGTGTACGACGCCACGGAGCATAAATCAAGCTGCGTATTTTAT
TGCATGATAAGACAAGGTATAATGGGCGACGAAGACTGTCTCTATTTGAACG
TCTACACGCCGGACGTAAACAAAGACGCTGCTAAGGCTGTCATAGTGGTTAT
TCATGGTGGCAGTTTTAACGGGGGTTCCGGTGACGATGAAATCTACGGTCCG
GACTTTTTGGTAGAGAACGACGTGGTCCTTGTCACGTTTAACTTCAGACTCGG
TGCATTGGGATTTTTGAATGCGGATGACGCCAGTGCACCGGGTAACGTTGGG
CTGAAGGATCAAGTAATGGCCCTGAAATGGGTTCAGGAAAATATTGAAAATT
TCGGAGGATCCCCTCGGAGAGTGACCCTCCTTGGACAGAACTCCGGAGGTGC
TGCCGCTCAGTACCACGTGTTGTCCCCTATGTCCGAAGGATTATTTTCACGAG
TCGTCATGCAAAGTGGATCAGTCCTGAACAGCTGGTCGTTTACTTACGACCAC
AAGGAATTGGCTTTCAAATTGGGAGAGCTTCTTGGCATACGTACGAGCGATA
GCACCGAACTGGTTGAGAGACTAAAGGAAAAATCAGCCAAAGAGATAGTAG
ATGCCAGCGGTCAACTTATGAAGTCTCTGAATGCCTTGAACGGACATATGCA
CGCTTTTGTACCGTCGATCGAGGCTGACGTGGGTCAGGAGATTTTCCTAGCTA
ACGTACCCTGGGATTTGGTTAAATCCGGAAAAATCAATGACGTTCCAGTGAT
CGCCGGTATTAACGCTGACGAGTCTATGATGTTCACGAAACAAATGTTACGT
AACGCGGATTACATAAATGCGAATTTCGATAAGTTCGTACCCGACGACTTGA
ACGTCACCGATCCGTCGCGCCTCAAGGAAATCGGCGAATCCATCAGATCGTT
TTATCTCGATGGCAAAAGCGTCTCGGAGGACACCGTTCAAGAATTCTCAAAC
ATGTTGAGTGACATCTTTTTCAACTACGGCCTTCTAATTTCCACGAAGGTTAT
GGGCTCCCGTGTGGCTTCGCCTGTTTACTATTATATGTTTACGTACGAGTCGC
CACTTGGGCTCATGAAGAATTTAAAGATATTCGATGTGGTGAGAGGAGTATC
GCACGGGGACGAAATACTGTATGAGTTCTATTCGTCAGCTTTTAAAAATATTC
ATGAAAAGGGATCACCGGCTGATCGCGTTACGAATGAATTGACTAAGCTATG
GACCAACTTCGCCAAGGACGGCAATCCAACATCGGTCATGGATAACTATGTG
ACGGTCGATTGGGAGCCTATGGGTAAAGATGACAATTATCTGAACATTAATA
AGGATTTAAAGATGGGAAAGCATCTGATGAAAGACAGAGTCCATTTCTGGGT
TGAAATATACAAGGATGTGCTTGGTGATTATTTGAAAAATTTCCAATAG

>CcCCE3
ATGATCTTGAAAATGATCCAAGCGTCGATCCTCGCGATTCTTCTCGCTTGTCT
AACGTGTCTCGTCACGGCTGGTCCAAGAGTCAAAATTCGAAACGGCACCTTG
GAAGGTACCATCATGAAGAGCAGACGCGGCGTGGAGTTCGCTGCCTTCAGGG
GCATCCCTTACGCGCTACCACCTCTTGGAGAACTGAGATTCGAGCCACCCAG
ACCAGCCGCATCTTGGTCCGGTCTACGTTCGGCTAAGGAAGATGCTGGGATC
TGTACCCAACGCAATATTTACACGCATCAGAAAGAAGTCGTTGGAGTAGAGG
ATTGTCTTTATCTCAACGTCTACACCCTTGAACTACCACCCCTCGAGGATGGT
GATAACACTCTTTGGAACCGGCCTTCCTATCCGGTGATGATCTGGTTCCACGG
AGGCGGATGGGTCACCGGCGCTGGGCACTCTGAATTTTACGGTCCCAAATTT
CTCCTGGACCATGACGTCATCCTCGTCACCGTAAACTTCAGGCTGGGTCCATT

GGGCTTCCTCAGCTACGAGGACCTGATCTTGCCTGGAAATCAGGGGATGAAG
GATCAGGCCCAGTCCATACGTTGGGTCAGTGAGAACATTGCTGCCTTCGGGG
GTGACCCCAACAGAGTAACTCTCTTCGGGGAAAGTGCAGGTGGCGTTAGTGT
TCACTACCATATGATGAGTCCACTTTCAAAAGGTACGATAGCTCTTGACCTGT
TTCATCGAGGAATTTCACAGAGTGGCACCGCTCTGTGCTTTTGGGGACTGACC
AGGCCCGGATTGGCTAAGAAACAGGCCCAACGTCTTGGTCGATATTTGAATT
GTCCTAATCAAGATTCGAAGTCATTGCTATCCTGTCTTCGTGAAAAAGACGCT
GTTGACATTATTGGCACCGACAAAGAATTTCAGGAATTCGATTACTGCCCAAT
GATCCCATTCAGACCGGTCATCGAGCCAGATCATCCTGGCGCATTCCTCAAG
GAAGATCCAGCAGTCTCATTGAAGGCCGGACGTATTGCTGATATTCCATGGA
TGACTGGTATCACGTCCCACGAAGGAGCCCTCAGAGTTGCTGGTCTTTTGGGT
CTCAACGATGGCGAATTGGCCAGAAAGCTGAACGACGACTTCATGACTATTG
CTCCCATGAGTCTACTCTACAGGGATAAATGCCCTGCGGATCGTATTGAGAAT
GTGACGAAGTCTATTCGCGAGTTCTACTTTGGTGATCGGCCAATCGATCAATC
GACCAAGTTCGATCTTATCGACCTTTACTCGGACGCCTGGTTCTCCATTGCTG
CTGATAATGCAGTCAGGGATCATCTGAATGTCCTTGCTAGTCCAGTTTATTAT
TATTATCTTGCGTACAGAGGAAGCGCTTCCTTTAGCTCTATTTTTGGAGATCC
AAAAGGGGAGTATGGCGTTTGTCATGCGGATGACTTGCAGTATCTTTTCCCTG
TTGGGGAACAATTGTTCAAGGAGTTTCCATTGAGCGTGGATGATCACAAAGT
GGTCGATATATTTACCAGTCTTTGGTATAATTTTGCCAAATCAGGAAATCCAA
CACCTGTGCTCACGAAAGAGATTCCAATCAAATGGAAACCAGTGAGGACTCT
GGATCTTGAATATCTACACATTGGAAGTCCAAAGGAATTGTTTATGGCTGAC
AATCTTATACACGAGAGAGTTAAGTTCTGGGAGAAATTAGGAATCGTGGGAG
TTGATGATTCCCAACGTCATTTTATGGAATCATCTCAACTCAGAGACGAGTTA
TAA

>CcCCE4
ATGAGCGCACCGATTGTGAGAACGAGATACGGAGCCTTGAAAGGTACAGTTG
TACAAAATGTAGAAGGCGGAGAGTACTTGGCCTTCAATGGTATTCCCTATGC
TGAGCCACCAGTGGGCCAACTAAGGTTCAAGGAACCCCAACCTCCAAAAGCA
TGGTCAGGCATCAGGGATGCTGGGAAAGAAGGTTCTGCAGCAATACAATTCG
ATATGTTGAGACCGACTCCTGAAATTATTGGGAGTGAAGATTGTCTCTATTTG
AATGTTTCTACTAATTCATTGTCTGGCAAAAGGGCTGTTATGGTATGGATACA
TGGAGGAGGATTTGTGATGGGCTCAGGAAGCAGTGACATGTACAGTCCGGAC
TATCTACTTAAACACGACATAGTTTTCGTCTCCATTAATTACAGACTTCATGT
ACTCGGCTTCCTAAACGTGGATGATGCGGAGGCCGCAGGAAATCAAGGTTTG
AAAGATCAAGTAGCAGCATTGAAATGGGTCCAGGAGAACATAGAAGCATTT
GGCGGCGATCCGCAGAACGTGACGATCTTTGGAGAAAGTGCCGGTTCTGCTT
CTGTCCATTTCTTGACACTCTCACCATTGGCCAAAGGACTATTTCACAAGGCA
ATTGGCCAGAGTGGTTCCGTTCTCAACACTTGGGCCTATAACAGGGATAATA
AGGAATCAGGACTACGATTAGCTGCATTATTAGGATTTGAAACTACAGATGG
TAAAGAAGCTGCACAGTTTCTACGGACAGTCGATGCTAAAACATTGACAGAA
GGAGTGGTAGAGCTTACTGGTGGTTTCGGAGAGTTCACGGTTAACGTTTTATT
TGTACCAACAATGGATGACAAGAGTGAAAATCCCTTTGCACCTCAGCCAGCA

GATGAAATGGCCAAAAAAGGAATTGAGGTTCCAAGTATATTCGGTTACAATA
GTCACGAAGGGATTATGTTCCTTATAGGATCTACGGATAAGACATACAGTAA
AGTCAAAAATGACTTCGATGGCTTTTTTGATAAACTCATTGTTAATCAAAATA
TTAAAAAAACTGATGACGTCGTAAAAAGTGTCAAGAAATACTATTTTGGGAA
GGAGCCAATCACACCAAAGCAACGCGACAATTACGTAGAGATTCTAGGTGAC
CTGCATTTTGTCTATGGAATTCGACGAGTAGTAGATATTCAATCACAGAAAAC
AGTACCGATGTATCTTTATAAATTCTCTGCCCTATGTGAAAATTTTAGTATGA
TATCACTCTTCAAACTTAACGCAAAAGGTACATGTCATGGAGATGATATAGG
CTGTCTATTCTACAATCGTTTATTCAATACCAAGTTCGAACCGAATTCTAAGG
CTCGTATAAGGATGGAGCGTTTCACCAAGTTGTGGACGAACTTTGCCAAAAC
CGGGGACCCAACGCCAAAGCTCGACGACGTGATAACTGAGAAATGGCTACC
AGTAACTAAGCAAGAAGTTCATTATCTAGAAATTAATGATGATCTTGTACCC
GGCGTCAATCCGGATGAAGAAAAATGGCAATTCTGGAAAAGCATATCTGAAT
CCGCGTAA

>CcCCE4b
ATGAGCGCACCGATTGTGAGAACGAGATACGGAGCCTTGAAAGGTACCGTTG
TACAAAATGTAGAAGGCGGAGAGTACTTGGCCTTCAATGGTATTCCCTACGC
TGAGCCACCAGTGGGCCAACTAAGGTTCAAGGTAATTTTATGTCGGGACAAT
AAGGAATCAGGACTACAATTAGCTGCATTATTAGAATTTGAAACTACAGATG
GTAAAGAAGCTGCACAGTTTCTACGGACAGTCGATGCTAAAACATTGACAGA
AGGAGTGGTAGAGCTTACTGGTGGTTTCGGAGAGTTCACGGTTAACGTTTTAT
TTGTACCAACAATGGATGACAAGAGTGAAAATCCCTTTGCACCTCAGCCGGC
AGATGAAATGGCCAAAAAAGGAATTGAGGTTCCAAGTATATTCGGTTACAAT
AGTCACGAAGGGATTGTGTTCCTTGTAGCATCTACGGATAAGACATACAGTA
AAATCGAAAATGACTTCGATGGCTTTCTTGATAAACTGATTGTTAATCAAAAT
ATTACAAAAACTGATGACGTCGTAAAAAGGGTCAGAAAATTTTATTTTGATG
AGGAACCGATAACACCTGAGCAACGCGACAATTACGTACAGCTATTAGGTGA
CTTTCAATTTGTCTTTGGAATTCGACGAATTGTAGAAAATCAATTCCAAAAAA
GGGTTCCATTTTTTTTCTACAAATTCTCTGGACCAGGTGAAAACGTGTTTATG
CCATCCAATGTCACAAACAGCATCAAAGGTACGTGTCATGGAGATGAACTCG
GTTGTATATTCTACAACCGTGCATCTAATTCCAAGTTGGAACCAGATTCTAAG
AGTCGTGTAACGATGGAGTATTTCACCAGGCTCTGGAAGAACTTTGCCAAAA
CCGGGGATCCAACGCCAAAGCTCGACGATGTGGTAACAAAGAAATGGCTAC
CAGTGACAAACGAAAGAATCTATTACCTAGAAATCAATGAAGATCTTATAAC
CGGAGTCAATCCGGATCAAGAGAAATGGCTGTTATGGAAAAGCATATCCCAA
TCCGCTTAA

>CcCCE7
ATGGGAACAAGAAAGCCATACGTTGGACCGATGGTACGACGATACGTTCATA
CCGGGGTTTCGAGACGTTCTAACGTACAGTCTGCTAGTCTAACGGTTATTGGT
GGTACGAATGATGGAGGACCAAAAGTTTGTATCGACGATGGTTGTCTGATTG
GTTTTGAAGATACTAGCAGATTGAACAGGACATTTAGTTCATTCCTTGGAATT
CCATACGCTAAACCCCCTATTGGGGAATTACGATTCAAGCCTCCGGAAAATC

CAGAACCCTGGGACACGCGTGAAGCTAAAAAAGATGGAAACATGTGTATCG
CAGCCATGTATTTCCGTAATGAAGATGACCCACCAGGCAGTGAAGACTGCCT
TTACTTGAATGTTTATACACCAGTGTTACCTGAAATAGGCGTTAGTTCACCGG
ATCTTCCAGTAATGGTCTTTTTTCATCCTGGTGGTTTTATGTTCTGGTCAGGTA
CCACCTATCATTTTCAACCACACTTTTTCATGGACCGTGGTGTCATCCTTGTCA
CTTTAAATTACCGTCTCAATATACTAGGATTTTTAAGTACCGAGGATGAGGTA
GCTCCTGGAAATTGGGGACTTAAAGATCAAGTACATGCTCTGCAGTGGGTTA
ACGATAATATTGGGGCCTTTGGTGGTGATCCTGATAAGGTTACTATTTTTGGC
GGTAGTGCTGGCGGAGCTTCAGTACATTATCACATGCTAAGTCCACTTTCCAG
AGGATTGTTCACGAAGGCCATCGCCCAATCAGGTCTTGCCAATGCAAGATGG
GCTTATTCACTGAAGAATGAAGCAAGAGAGCATGCAGAATATATAGGTCAAC
ATTTTAATTGTCCTACTGAAAATTCGACTGTTTTGGTTGACTGCCTTCGTAAG
AAACCAGCAGAGGAACTGATTGAAGCTGATCCAATCTTTTTTGTTTGGTCTAT
TGATCCATTAATTAAATTCCCACCCGTAGTAGAGCCGGATATCGAAGGAGCG
TTTCTTATGGAGATGCCAAATGATCTATTGAATTCTGGAAAATTTGCGGATGT
ACCGTTCATGGCTGGAATTATGTCAAAAGAAGGAGTTTTGAGAAGCGGCGCT
ATCATGTCTCATGACAATATACTTCAGAACTATGATGTTCTCAAGTTAAAGTC
GTTCTCTAATAGTTTACACTTTTATCAGCATCCGGTAGAGGATGTAATCAAAA
TCATTGACAAAATTAACAAATACTATTTTAAGGGAAAACGCCTGACTCTTAA
ACAAGAGAACAATATGATACACATGTTTTCTGACAGATATTTTGTCCATCCGT
TGGTGAAAGCTGTTGAGATGCAGCGTAAACATCAGAAATCACCAATTTATCT
GTACAGTTTTGAATATGAGGGACCAAAAAGTTTTGCTACAGTATTTGGTAGA
GAACTTAGTACTCCCTTAAATGGTACAGCGCATTTAGATGATTACTTGTATCT
ATTTCCTATGGAGAAAACCTTCTTCAAGAATGATCCTTGGACCAGAAGTAGT
GAAGAGGATGAAATGATAGAAAGAATGTTAGACATTCTAACAAATTTTGCAA
TTTTTGGTAATCCTGCCCCTAGTTTCTCTGGGTTTACGTGGAATCCAGTTAAAT
CTGAGGACCATGAATATCTAGTCATAAAGGGACCTCACGATCTAAAAATGTC
ATTAAAATTGAGATCAGATGCTGTCGCCTTTTGGGATAA

>CcCCE8
ATGAGTGCACCGATTGTAAAAACGAAATATGGGATTTTGAAAGGTACCGTTG
TACAAAATGTAGAAGGTGGAAAGTACTTGGCTTTCAATGGTATTCCCTATGCT
GAGCCACCAGTTGGCCAACTTAGGTTCAAGGAACCCCAACCTCCAAAAGCAT
GGTCAGGCATCAGGGATGCTCAGAAAGAAGGTTCTGCAGCAATACAATTAAA
TATACTGAGTTCTACTAATGAAATAACTGGGAGTGAAGATTGTCTCTATTTGA
ATGTTTCTACTAATTCACTGTCTGGTAAACGGGCTGTTATAGTATGGATACAT
GGAGGCGGTTTCTGGACAGGCTGTGGAAGTAATGACTTCTACGGTCCGGACT
ATCTACTTAAACATGATGTAGTTGTCGTATCCATTAACTATAGACTTCATATA
TTTGGCTTCCTGAACGTGGACGACAAGGATGCTGCAGGTAATCAAGGTTTAA
AAGATCAAGTTGCTGCGTTGAAATGGGTCCAGGAAAATATAGAAGTATTTGG
GGGCGATTCGCGAAACGTAACTATCTGTGGACAAAGTGCTGGTTCTGCTTGT
GTCCATTACTTGACAATGTCGCCATTAGCCAAAGGATTATTTCACAAGGCTAT
TGGTCAGAGTGGTTGCGCTCTGAACAACTGGGCCTATACCACGAACAATAAA
GAGGATGTCTTACATTTAGTTAAAATATTAGGACTTGAAACTACAGATACTA

AGAAAGCTGTGCAGCTTCTGCAAACAGTCGATGCAGTAACAATGTTGGAGGC
ATTGATTAAGCTTGTCGGAGGACCAGCAATGATGTTCTCCAAGATTATATTTA
AACCATCGTTGGATGAAAAAAGTGAAAATTGCTTTCTGCCTCAGCCACCTAA
TGAAATGGCGATGAAAGGAATTGAAGTTCCTAGTATCATCGGATATGATAGC
CACGAAGGAATTCTATTTCTTCGAGAAGCTACGGATGAGACATTAAGTGCAG
TCGAAAAGGATTTTGATGAATTGTTCGATAGACTTATTTTCAATCAGAATCTT
AAAAAAACGGATGATATCATAAAAATCGTTAGAAAATACTACTTTGGTGACG
ACCCAATTACATTGGCGCAATTTCAAAATTACTTGCGCATTACAGGCGACATG
CGCATTGTGAATGGAATTCGAAGATTAGTAGGTTTTCAATTGGAGAAACGAG
TACCGATGTATTTCTATATTTTTTCTGCTCAAGTTGAAAAATCGTGGACAAAA
TTTTGGCTCAAATACAGTGGAGAAGGTACATGTCATGGAGAAGAAATCGACT
GTATATTCTACAATCGTGCATTCAGTTCCAAATTGGAACCGAATTCTAAAAAT
CGTGTATCGATGAATCGTCTCACCAGGCTGTGGACGAACTTTGCCAAAACTG
GAGATCCAACGCCAAATCTCGACAACGTGATAACTAAGAAATGGCTACCAGT
AACTAAGCAAGAAGTCAATTACCTAGAAATTAAAGAAGATCTTATAACCGGA
GTCAATCCGGATCAAGATAAATGGCAGTTTTTGAAAAGTATATATGACTCTG
CTTAA

>CcCCE9
ATGATATTTTACTATTACAAACTTATGGGAATGTCTTTGGTGTTTTCTATTTAT
TTGGCTAAGGGTTCTAAAGAAATTAGTGAGAATAATAGTGCAGTCTTAGACT
ATGATGCAAAACCTCAGGATTCAGAGAATAATCCTGTTATTCAAAGTTCCTTA
GGATATCTTAAGGGCAGCTGGCTTTGTTCACCTAGTGGAAAATATTATGCGGC
TTTTAAGGGGATACCTTATGCGAAACCACCCCTTAAAGAATTGCGCTTTGAGC
ATCCACAACCCGCTGAGCCTTGGAGTGGTGTAAGAGAAGCTCTTACACATGG
GAGTCAATGTCCTCAACGTTTATGGATACTGGATATCACCCTGGGACAAGAA
GACTGTCTTTATTTGAATGTTTATACACCAGGGCTGACCGACCATAAGGAAA
ATGCTACCAACGCGACAAGGAAATTATTACCAATAATCTTCTTTATACATGGT
GGCACTTGGACAATAGGATCTGGAGATATGGATGATTACCTATATGGTCCCG
ATTTCTTCATGGAGGAGGATGTTATACTGGTTACAATTAATTATCGTCTTGGT
GCTCTAGGGTTCCTCTCCCTGGGTAACGTTATACCTGGGAACATGGGGTTGAA
GGATCAGGCGATGGCTCTTCAATGGGTTTGGGATAATGCCGATGTCTTGGGT
GGAGATAAAAATAGAATTATAGCTATGGGTCAAAGTGCTGGTGGTGCTTCTG
CCCAATATCATTCTCTTCAACTTCACTCTAGGTCTTTAATCTCAGGTGTTATCG
CCCAGAGCGGTTCAATTCTAACACCCTGGGCCTTTACGAAACCCGGGATACA
GGCTGAGAGGACCAGAAATCTTGCAAGGGCCGTTGGATGTCTGAATGGAGGT
TCCCAAGCGATTATTAATTGCCTTAAAAATGTTGATCCTTTTGACATAGTGAA
ATATCAATTTTCTACAGATCATGACATCAAACAACTTCCTATTGGTGTGACTT
TTACTCCGACAATAGAACCAGCTGGTACAGTTAATTCTTTTATTAAAATGGGT
CCCAAGACACTTTTGACTAAATACGTATCAAGTAAAATACCTTTTCTCACTGG
ATATACAACCGCCGAGGGACTTTTATTTTTAGAAATTATTGTGGCCATGATAA
ATCCTCTAGCTCGAGAAGCTGTACTTAAACGATTGGGAAAAGAAAAAGAAGT
TTTTCAACACGATCAACAACTTGGGAGTAAAAGAATCAATATATCTGTGGAT
GAAAATAATACTCGAATAGAGAGTGCAAAAACTTTGGAAAACAGGATCTTGA

AGCACCGTCTAATAGAAATAGCCCATCGTATTCTTGGTGACTACGAATTCAA
AATGCCAATATTATTAGGATCTAAGTTCCATGCTATGTCTGGGGCTAGCGTTT
ACCTTTATCAATTTGCTTTTACTGGAAAAAATAACTTCTTACGACTTTTCCGCG
ATGACACTAAAGATGCTACACATATGGACGAACTTACATACCAGTTTCAAGG
AAGAGGTATTTTTGGTAGTTTACCACAATTGAACAATCATACATCTACCGAAT
ACTATCTATCCAAGGCAATTATATGGATGTGGTGCAATTTTGCTAGAACAGGT
AACCCAACACCAAGAGGAAGTCCTATTCAGTGGCCAGCAAATTCAAGGAAG
ATACTTATATTTGATAATCCAGACTCGTTAATTGTTGAAGATGCACCAGTCAT
GAACAACTGGTTCTCATTAATAGACACCTAG

>CcCCE10a
ATGAAGAACAGAGGTATCGAATTAGTGGTGGCCATTTTTGGTTTGGCTCTCCT
GGCACAAAATGCCAACGGGGCGGATTCGACTATTGTCAAGGTCGACCTACCC
AAAGGTACCTTGCAAGGGTTAAAAACGCGGACTGTGTTAAATGGCGTCACTA
TGTACAGCTTCAAGGGCGTAAGATACGCAGCCCCCGCAACTGGTGTACATAG
ATTCTCTGTTGCTCGAGAGGTCGCGCCCTGGATTGGTGATTACGATGCAACCC
AACACAAATCAAGATGCCCTCAGAGATGCGCTTCAGCTTTTAAATTCATCATT
GGAGAAGATGATTGTTTATTCTTGAACGTCTACACTTCCACTCTTGATTCTAA
AGCTGGTCTACCAGTGATGTTTTGGTTGCATGGTGGATCCTTTAATTTTGGTA
ACGGCGACAGTGACGTGTATGGACCCGATTATCTTATCGAAAATGGCGTCGT
ATTGGTCTCGATAAATTACAGACTGGGTCCAATAGGATTCCTCAGCACAAGA
GATGCTGCTGCTCCAGGCAACGTGGGTCTTAAGGATCAAATTGCTGCTTTGAG
GTGGGTCAAGCGCAATATTCGTTACTTTGGTGGAGATCCAAACCGGGTAACC
ATCTTTGGTGATGCTGCTGGTGCTGGTTCTGTACAGTATCATATGATCTCACC
ACTCTCTGCTGGTCTCTTCGCCCATGCCATTGCTCAGAGCGGTACCGTCATGA
CTACCTGGGCCATTACCTACAACTCTACCAAAGATTCCTTTGCTCTAGGAGCA
GCATTCGGAATCAATACCAATAATTCAACCGAACTTGTCGAAGGATTATTGA
AAATTAATAGTACCGCCCTTGTTCGAATGGCAAACAAATTGTCTAGAGAGAC
GGAGGGTATGTCAGGAGGTCATTACCTTTTCAAACCCTCAGTAGAAGTTAAC
GTGGGTCAGGAAATATTCCTGCCGGCTGATCCTTGGGAACTTCTAAAGACTG
GTCGAATTAATGACGTTCCATACATGATGGGTTTCAATCAGGACGAGATGAT
CATAGGCGCTAACAGTATCGCAAGTGTGGATTACTATAATAGTCACTTTGAA
GGATTCCTACCAACTGACCTCAATCTTACAAAAGGCGCAGTACTTGACACTG
ACATTAAACTAGTGAGAAATTTCTATTTCAATGGAAAGAATGTAACTAAGAA
TGACATCCAAGCTTATATTAAGCTGCAAAGTGATCTGTACTTCACCTACGGAA
CAGCATTCTCTTTGAAGATGATGAGATATTACATGACCAAACCAATCCACCA
GTATCTGTTCTCCTTCGATGGAAAACTAGGATTCTTCAAAAAATTCTTTAACG
TATCTTTGACATCCGGTGTAGCCCATGCAGATGAGACTGGATATATGTTCTAC
CCAGCCCTTCTTAAAATTACGCCGGATATAGGCTCAGAAGAAGAAAAGATGT
TATACAAAATAACAAGAATGTGGACCGACTTTGCTAAATATGGTGATCCAAC
TCCGAAGTTATCCGAAAACGTTACGACAACCTGGGGCGATGTAACAATGAAC
GGATACTATTTGCACATCAACACAACCACAAAAATGGAGCAGAATGTATTTG
CGGAAAGAGTCAAATTCTGGGCGAACATATACAAAGATCTACTTGGCCAATT
TTATCAATACTTCAAATAA

>CcCCE10b
ATGAACGGATACTATTTGCACATCAACACAACCACAAAAATGGAGCAGAATA
CTGTCGAGGTCACCATACTAAAAGGTACCTTGCAAGGGTTGGAAACACAGAC
TGTATTAAATGGCGTGACTATGTACAGCTTTAAGGGTGTAAGATATGCAGCC
CCAGCAACTGGTACACGAAGATTCTCTGTTGCCGAAGAGGTCGAGGCCTGGG
ATGGTGTTTACGATGCAACTCAACACGGATCAAGATGCACTCAGATATGTGT
TCCATCTCTTTCATCTATCATTGGCTCAGATGACTGTCTATTCCTCAACGTCTA
CACTCCAAACCTTGATGATTGGGCTTCTTTACCAGTGATGATTTGGTTCCACG
GTGGAGACTTCAATTTTGGCTCCGGCAACAGTGACATCTATGGGCCTGATTAT
CTCGTCGAAAATGACGTCGTACTTGTCACAGTAAATTACCGACTGGGTCCAAT
GGGATTCCTCAGCACGAGAGATGCTGCTGCTCCAGGCAACGTCGGTCTTAAG
GATCAAGTTGCTGCCATGAGGTGGGTCCAGGACAATATTGCGTTCTTCGGTG
GAGATCCATCTCTGGTTACCATCTTTGGTAATTCTGCTGGTGCTGGAGCTGTA
CAGTATCATATGATCTCACCACTCTCTAGAGGTCTATTCAGCCGTGCCATTGC
TGAAAGTGGTCCAATTTTATCTACTTGGGCAATTTCTTATAATTCTACTGCGT
ATTCGTTTGCCTTGGGGAAACTACTAGGTATCCATACCACCGATTCCACTGAA
CTTGTCAACGGATTATTGGAAGTCGATAGTGAGTCCATTGTTGCAACAGCCAT
CAGGCTAGCTAGTGACAAGGATAATATGGCAGGTCTCAACTTCATTTTCAGA
CCCACAGTGGAAGTTGACGTTGGTCAGGATATTTTCCTTCCGGCTGACCCTTG
GGTACTTTTAAAGAATGGTCAAATCAACGATGTTCCGTTCATGATGGGTACCA
ACTTAAACGAGTCAATTACTGCAGTGAATGATGTCAGCAAGGCAGCGTTCTT
CAATAACAATTTTGATGCTTTTCTACCATTCGATCTTAATCTCACAGACAGTA
GAACAATTAGTGATATTGTTACTAAAGTCAAAGACTTTTATTTCGGTGGCCAT
GGTGTCACCGATGATGACTCACTTTCTTATATCGAGCTGAGAGATGATCTAGA
TTACACGCATGGGACGGAAATGTCTTTAAGGATAATGGCATACTACATGGAC
AGTCCGATTTATCGTTACCTATTTGCATACGATGGAGGACTTGGATTTTTCAG
AAAGATTTTCGGCATATCACTGACCTCTGGAGTATCTCATGCGGATGAAGTTG
CCTATCTCTTCTATCAAGCTGCTTTTGGCGTCACTCCGGCTGCTGGATCGACA
GATGAAAAAATGGTTTACGCAATGACAAGAATGTGGACCGATTTCGCAAAAT
ACGGTGACCCAACTCCTGAGCTTTCCACGAATGTCACCACAGCTTGGGAAGA
CATGACAACGAAAGGAAATTTCTTGAACATCGATTCAACTTCAACCATGTCG
TCGAAAGTTTTCAACGAAAGAGTCCAATTTTGGGTATCCATATACAAAAATA
TTCTTGGAGATTTTGCATGTCTCTTTGACTAA

>CcCCE10c
ATGAAGAACAGAGGTATCAAATTAGTGGTGGCCATTTTTGGTTTGGCTCTCCT
GGCACAAAATGCCAACGGGGCGGATTCGACTACTGTCGAGGTCACCATACCA
AAAGGTACCTTGCAAGGGTTGGAAACACAGACTGCATTAAATGGCGTGACAA
TGTACAGCTTTAAGGGTGTAAGATATGCAGCTCCAGCAACTGGTACACGAAA
ATTCTCTGTTGCCGAAGAGGTCGAGGCCTGGGATGGTGTTTACGATGCAACTC
AACACGGATCAAGATGCGCACAAATTTGTGTAGCATCTTATTCATATTACATT
GGCTCAGATGACTGTCTATTCCTCAACGTCTACACTCCAAACCTTGATGATCA
GGCTTCTTTACCAGTGATGATTTGGTTTCACGGTGGAGACTACAATTTTGGTT

CCGGCGACAGTGACGTTTATGGACCCGATTATATTGTCGAAAATGGCGTCGT
AATGGTGTCGGTCAATTACAGATTGGGTGCAGTAGGATTCCTCAACACGAAA
GATGCTGCTGCTCCAGGCAACGTGGGTCTTAAGGATCAAGTTGCTGCCATGA
GGTGGGTCCAGGACAATATTGAATACTTTGGTGGAGATCCAGGACGCGTAAC
CATCTTTGGTGATGCTTCTGGTGCTGGATCTGTACAGTACCACATGATCTCAC
CACTCTCTGCTGGTCTCTTCGCCCATGCCATTGCTCAGAGCGGTACCGTCCTA
ACTACCTGGTCCATTTCCTATACCTCAACCGCAGATGCCTTCGCTTTAGGAAG
AGTATTAGGAATCAATACCACCGATTCCACTGAACTCGTTAATGGATTACTGA
ATATTGATTCTAAGGACATTGCTAGGGCAGCTTTTGGGATGACTACAACTAA
GGAGACAATGGCAGGAGCTCTATTCGTGTTCAGACCTTCAGTCGAGGTTGAC
GTTGGTCAGGATATATATCTTCCTGCTGATCCTTGGCAACTTTTAAAAACCGG
CCAGATCAACGACGTTCCTTATGTAATGGGTTTCAATAAAGACGAAGCGATT
ATAATCGCTACTAGTATCAGTAACGCAGAATTCTTCAATAGTAACTTCGACGG
TTTCCTGCCTGTGGAACTGAATCTCACAGCCAGTAGAGACAGTGAAGAGATC
AGCGAAGACGTTACAAAACTGAGAAACTTCTATTTTGGTGGAAATAATATCA
CCACCAACGACGTGCTGCCATATATCGAGCTTCAGAGCGACATATACTTCAC
GTACGGGACAGCCTTCTCTTTAAAAATAATGGCGACCTACATGAAAAGTCCT
ATCTACAATTATCTTTTTACGTATGACGGACAACTTGGATTCTTCAAAAAATT
CTTCCGCATTAATATGACGTCGGGCGTAGCCCATGCAGATGAAACTGGCTAT
ATGTTCTATCCGGCTGCTCTTGGTATTACGCCAGAAATTGGATCGACAGAAGA
AAAAATGGTTTATGCAATGACTAGAATGTGGACAGATTTTGCAAAATATGGC
AATCCAACTCCATGCCTTTCTGAGAACGTTACCACAATTTGGGACGAAATGA
CGATAGATGGAAATTATTTGGACATCAATCCAGATTCATTC

>CcCCE10d
ATGGACCAGAATGTCTTTAATGATAGAGTCAAATTGTGGGCGTCTATATTCAA
AAATACTGTCGAGGTCACCATACCAAAAGGTACCTTGCAAGGGTTGGAAACA
CAGACTGTATTAAATGGCGTGACTATGTACAGCTTCAAGGGTGTAAGATATG
CAGCTCCAGCAACTGGTACACGAAAATTCTCTGTTGCCGAAGAGGTCGAGGC
CTGGGATGGTGTTTACGATGCAACTCAACACGGATCAAGATGCACTCAGTTA
TGTATTCCAGCTTTTCGACCTGTGATTGGATCAGATGACTGTCTATTTGTCAA
CGTCTACACTCCAAGCCTTGATGATTGCGCTTCTTTACCAGTAATGATTTGGT
TCCACGGTGGAGATTTCAATTTTGGGTCCGGCAACAGTGACATCTATGGACCT
GATTATCTTGTCGAAAATGGCGTTGTACTTGTCACAGTAAATTATCGACTAGG
TCCAATAGGATTCCTCAGCACAAGAGATGCTGCCGCTCCAGGCAACGTGGGT
CTTAAGGATCAAGTTGCTGCCATGCAGTGGGTCCAGGAAAATATTGTACACT
TTGGTGGAGATCCATCTCTCGTAACCATCTTTGGCGATGCTGCTGGTGGTGGA
TCTGTACAGTATCATATGCTGTCACCACTTTCTAGTGGTCTTTTCAAATATGCC
ATTGCTCAAAGTGGTACCTGTTTGGCTACTTGGGCAATTTCATATAATTCCAC
CGCATCTTCCTTCGCCCTGGGAAAAGCACTGGGTATCAATGCCACTGATTCCA
CTGAACTTGTCAATGGATTATTGAATGTCAATAGTAGCACCATTGTTCAAACA
GCCTTCAATCTGGGTACTGCTAGGGATGCCATGTCTGGTGCCTCCTTTATGTT
CAGACCCACAGTAGAAGTTGACGTCGGTCAGGATATTTTCCTTCCGGCGGAT
CCCTGGCTGCTTCTGAAAAACGGTGAAATTAATGATGTTCCTTACATGATGGG

CTTCAATCTCAACGAATCAGTTATATCTGCCAATGGTATCAGAAATGCGGCAT
TCTTCAACAATAATTTTGATGCTTTTCTACCAATCGACCTCAATCTCACAAAT
AAAGCAGAAATTAGTAAAATTAGTAAAAAAGTACAAAATTTTTATTTCCATG
GTAATATCAGTACTAAAGATGTACTTCCTTATATTCAGCTGAAGAGTGATATC
TACTTCACATATGGGACGGAAATGTCTTTAAGGATAATGGCATCCTACATGGT
GAATCCGATTTATCATTACCTATTTTCATACGACGGAGGACTGGGATTTTTCA
AAAAGTTTTTAAACGTATCGTTGACCTCTGGTGTAGCACATGGGGATGAAAC
TGGCTATCTCTTCTATCCAGCTGCTCTTGGCGTCACTCCGGCTGCTGGATCGA
CAGATGAAAAAATGGTTTACGCAATGACAAGAATGTGGACCGATTTCGCAAA
ATACGGTAACCCAACTCCTGAGCTTTCCACGAATGTCACCACAGCTTGGGAA
GATATGACAACGAAAGGAAATTTCTTAAAGATCAATTCAACTTCATCCATGT
CGTCGAAAGTTTTCAACTCAAGAGTCAAATTATGGGTTTCCATATACAAACAT
ATTCTTGGAAATTTTGCATATCTTTTTGACTAA

>CcOR1
ATGTATTTAACCATAGATTCTGACATATTGTGGCTGAGTAAGCGAGTCTTGTC
ATTGGCTGGTATTTGGCCAGAGAGCCCGAATAATTTTCGATTTTTCATTTACC
TTCTGTACTTGAGTTTATTCAACTGCGCCGAATTTGCCGCGCTCGTACTGAAT
TTGTATTGGATGAATTTTGACAAGTCCGTCAGGAATATGACAGAGTCTATACC
CACGGCCATGGTGATTCTGAAGACGGCGATGTTCCGTAGAAATATGCAGCTG
TTGTTGCCGTTGTTGAGTGAAGTCAGAGCAGACAAATTCAGCGTAGAGGAAG
AGCCAGGAATTGCATGGTTGTATAATATCATGGGAAAGTTGTATACACGTCT
CTCCGTCGTCCTGATATTCATCGTCACGACGATGCTATACGCAGTACCGTTAT
CCCAGTGGATCGTTGCCAAATCAAATAATTTAACCTCAACCTACGAGCTACCT
TATCAAATGTATTTTGGATTTGAGATTAACGACTTGCGTAGTCACGTACTGGC
GTGTCTCAGTCTACTGCCAATGAGCACAGTTTTGACAATCGGCTGTACGGGAT
CGGATACATTGTTGGTTGTCCTGATATTCTACCTATGCCGACAGTTTGTATTG
CTTTCCGTGCGAATTCGAAATGTCGAGACTGATCCGTTGATTCACCCAACGAA
AATGAAACAGCTGATCGAGAGACACGCCAATCTCATTGGAATGGCCACAGCT
TTAAATAAGACTTACAGCTCCCTCTTGCTGGTGCAAACTATGGGTCTTAGTTT
TGTCATTTGCATAGTGGCTTTCGAGTTACTCACGATGGCAGAAGTTGGAGAG
GAAACGAATACCTTGAGCTTCATTATTTATTCACTGGCGGTTGTAACGCTACT
TTTCAGCTACTGCTTTCTCGGCGAGTGTTTAATTCACGAGAGTTCCTCGATTC
ACAACGCTTGCTATTTTTCAAACTGGTACCGATTGCCGCCAGACCTCGCCAGG
CCAATAATCATCCCGATCATGAGATCACGAAAACCCTTGCACTTGACAGCAG
GACAATTTTACGTaTTTTCCCTAGAAACATTTACATCTCATAGGATTGTCCACA
TGTCATTGGATCTGCAGCTACATGGAATCTATAGGAAGGATAGGCTCGATCC
CACTGGTTTTTATTACAACATTGAGTAA

>CcOR3
ATGCCGGAAGCTATGAAAGTGTTAACGTGGAACAAGTGGGTGCTAGAGTTTC
TAGGAATTTGGCCATCCAACGAAAGTCTTTTTACGTTTTCCTTCTTCTTCTTTC
TTGTGTCTATGGCAACTTGTTTTGTTTATGCGGACTTAATTTATCGTATCTCCG
ACTTTAAATACGTCGTTGAAAATTTAACGGAAAATATCGTATTAACCTTGCTG

TGCTGTAAAATTGGTCTCTACAGATTGAATAGACGATTAATGAAGGAAATAT
TACTTGATATCAAAATGGATTACGCAATAGAACTGTACAACACAGAGGAACA
AAAGTCTATCTTTTTAGCATACAACAGATTGTCCAAGTCATTCATTAAATACT
CAGTAACAACAACGACCGTTGCAACGGTACTTTATTATATTCAACCGTTAATG
GATCATGCAAACTCCCACAGAAAATTAACAGAAAATTCGTCGATTACCTACA
TTTTGCCTTATCACATGCGTATGTCTTTCAACATTACGGAATCGTCGTTATATT
ATTATGTATACGCTTACGAAGGAGTGCTCGTCCCGATTATCGCATGCGGATAT
TCCGGCACCGATTGTCTATTGGTGACACTGACGCTACATTTGTGCGCGCAGAT
TTCCGTTTTAGCAAATCAAGTGGAAAATTTTAATGGCGATTTTCGTAAATTTC
ATAGTCATCTAAAACAAGTTGTCACTAAGCATTCTCGAATAATAAGTCTCTCG
GTAAAGTTACGTACCGCGTTTGCTTACTTTCTGCTCGTACAATTAGTGGGTGC
AACTTTAGTAGTATGCCTGGCTATTTACAATCTTCTAAAGAATTATGCCACAG
GTCACACTGCTCAATTATTCGGGTTTATCTTCTATGGTTCTTCTGTGACTGTGC
AGCTTCTCGGTTATTCTTTTATCGGAGAACGCCTTATGACTGAGAGCTTGAAG
TTGAGCGATGCATTTTACAATTGCAAATGGTACGTTTTGTCACCGATATACGT
CAATCTTGTTTTATTTTGTATACTACGCGCGCAAGTGCCACTTGTATTAACAG
CTGCAGGATTTTGTACATTTTCCCTGCAAAGTTTCACAGACGTGGTCAAAACG
TCAATGACTTACTTGTCGGTATTGCGACATTTCGTGTAA

>CcOR4
ATGGATTTTTCTAGTGTCAATAGATTCAACGGTTTGGTTAATTTTGTCTGTGG
AAATTTGTTACCGTTGACAGACGACGCTACGAAATTTTCTAAAGCTCAGAAA
CTTTACTCGGCTTTCACCTGGATCCTTGAATTCACTTATTTTCTCGTTACGACT
TTTGGAATTTTTTTTACCACTAACCAACGCATTTTTCAAGACACTACCGTCAA
CCAGGCTGTAGTTATTGAGATTCTTATTTTGGGTTTTTATATGAACTTACGAC
GGGGTTTGATTTATCGATTGATTGGCCAGATGAACAGCGTTTTAATCAATGTC
GAAACCTTAAAGGAATGTGTGGAGAAGACTGTGAAACCTTTGCAGAGACCCT
TGAAGCTCTACACGATATTCGCTACTGTGACAGTAGTTTTGTTTTGCGGTTCG
CCAATCTACAAAGTGTTTAAGAAGGATCAATTTAGCTACAACGACTTTCGTAT
ACCTGCTTATATACCTGGTGAACCTTATTCGACTGGATTATTTATTGCTGGCA
TTCTTTTCGAGACACTGGGTGGATTCTATACAATTTTGAAAAAAGCTAGTATC
GATATTTACTTGATACACGTCATTACCCTTCTTACTGCTCAGTACAAGTATTTA
AGTTTGGAACTAATTAATATAATTAATGGAAAGAACGAAAATATTAATAAAA
GTGATCTCGAATCGTTGGACAATGAGAGAAATGAAAAAATCGAAAAGACTGT
ACAGATTGAATTAAAAAAGTGGATTCGACATCGTGAAGTTGTGATGGAAATT
GGAAAAATCCTGAAGAATTTACTGGTCTTGAATGTCGTCTACGTCTACCTGAA
TTGCATATTTCGTTTCTGTTTTTTGGGATTTTTACTGATTTCGAGTTCTGGTGA
TTATTTTATACAAGCTTTGGTATGTTCGTACACTGTCGTTTGCTTGATGCAGGT
ATACGTGTTATGCTTTTGCGCTCAAAATTTATTGGACTCGAGTACAGCCATGA
CGCACGATGCCTTCTACGAAAAATGGTACGCTTACGGTCCATCGACTAAACG
TATCTTCTCAATGATATTGATCTCGAATAAAATGGAATGTCGCCTCTCCATGT
GTGGAGTCGTTGACTTGACTCTGCCAACCTTCATGGCGATACTGAAGAATTCT
TACTCAGCTTGCTTGTTCTTACTGAAGGTGAAGTAA

>CcOR5
ATGCATCACACGATGTCAAATACTCGCGAAAAGGACCCGGATACTTCTTTCCT
CAAGTATGCAGATTTTCATATAAACCTGTTACGAAAATCAGGATTTTATTCGA
TGAAAGGGATATCAAATAAAATCAACAAGGAGCCCACTATATGGGAAGTGCT
GCTGGTTCTGACAATTTCCACTTGCGGATTTTTCATTATCATTTTGGAGTTTCG
CAGCGTTGCAGTATCACTAGGATCGGATACGGGATTTGTAATTGCTGTTTTAT
CAGGCACTCTCACGGCTACTTTAAGCATGTCAAAGGGCTTAACGATACTCAC
GTCCCATCGAGAAGTCCGTGAGCTTCTTCTGCGTTTGTCTGGCTTTTGGGAAA
AGTCTATAGAGCGACCAGAAAACGTGGACGTTATGGTGCAAATGGCTAATCG
AGCGTCTTACCTGAGCAAATGCTATGCGGCCACAGTAGTGATAATGTGCAGC
AGCTACTGCATGAATCCCTACGTGAGCGTTATTACACAGTTTTTATTTACCAA
GACAGCAAACAATTCGTACAACTTTACTGCTACTACTTTTCCAACCGTCTATC
CTTTTGATCTGTCATACTTTCCGAAATACGTAGTCTGGATTTTATTCGAACAA
GCCGTCTGTCTCCTAATGACACTCCATTGGATTGCCTGCGACACTCTCTTTCC
CATGTGCGCTACGCATCTTGCTATACAATTTCAAATTTTACGACGTGACCTGG
AGAGGACTACGGAAGTAGATGAATTACGTGAAATCGTTAAAAAACAAATCAT
ACTCTTTCAATCTTGCGACATTCTGGAAAATATATTCAGTCCAATAATATTTC
TCACTATAATCATGACTTCCACTATCATGTGTGCCTGTATATTCCAATTTGAA
AAAACACTGTCATGCGGGGTTTATTTGGAAATTATTAAATACGTAACGCATAT
GATGTCGCTTTTCGTGGAAATTCTGATTTACTGTGGATTCTCGAATGTTCTAA
GTGATCAGACGGAACTTCTATATCATGCAGCTTATAATAGCGAATGGACAGA
TCGTAGCAAGAAATACAAATCCATAATATATTTTTTAATATTACGATCGCAAA
AACCATTCCAGTGTACAGCTTACCACTTTTTTCCAGTTGGTCTAGTTCAAATC
ACGACGATATTGACGACAGCCGTTTCTTATTTCACGCTTTTGAAAACCGTAAC
AAGCGAATCAGATGACAAAGTTATATGCTCGTAA

>CcOR6/6b
ATGGTTAATTTGGTGAAAGGTTACATGTCGACTAAGATATCGCAATTGCTAAT
GACACTTATTGGTATGAAACGCGGTAAAACTAAACGTGAACAGTTGTTAATG
GACGCGTTATTTGTTTATATACTTGCGACGGTTTTATCGGCAATATGGCTAGA
AAATTCGGAATTATTTTATTCACGGAACGATCTGTATGCCCTCACTTATAGTG
CACCATGTTGCTTCACCGTGACATTCGATTTTGTAAAATTAATGATATTCACA
TACAAGCGGCATGAACTGCATGAACTGCATAAATTTACTGAAGATACATATT
GGAATAAGGATTACAATGAACTGGATAAAGCAATACTAGACAAGTGCGATA
CTACAAGTGCCATTGGAATGAGTATATTAGCTCTAACATCAGCAATCCTTGCA
TTTCATTATTTGACAGGACCATACTTGGATAACTTAGGAACAAATACCACTGA
GAGAACGTTACCATTTCGAGTAGTATTCGACTTCCCAATTACAGTGACCCCAT
TATATCAGATATTGTACTTTATTGAGGTGATTGGAACAATTAGTATAGGTATT
TGCAGCGTAGCGTTCGCCTCGTATCTATTCTACACGTGCATCTTTGTTTCCGGC
TTCTTTAAAATCTTGCAGCGAGAGTTGGAGAACGTGTGCGAGGTAGAACTAG
AAAGTGTAAACACAAAATCATCATATAACAATAATGATACTATGCTTGCATA
CAAAAAACTCAAAAAGTGCGTCATCCAACATCAACTACTAATTTGGTATTTG
GGTAAACTGGAAGGACTATTCTCCTACATATTGTTGATGCTCGTATTATGTGC
CGTCATAATACTCTGCTTTGCTGGGTTTCAAATAATTTTGGGAGATGGAACCA

CGAAACTGCATCGTCAAATTTTAAGCGTAGAATTTATCATGGCTGCTCTGGCT
GAAACAGTCCTCTTCGCTTTCTCGTGCAATGAAATATTGACAGCCAGCGCCGC
TATTGGGGAAGCTGCGTATCGATGCAAATGGTACAAGTTACCGTGCGATGAA
TATGGTCGAGCACTGAGACAGGGAATGACAATAATGGTCATGAGGTCCTACA
AGCCGTGCTCATTGACAGTTGGCAAATTTTGTCCCATGACTCTGGAAGTGTTC
ACTTCCATACTGAGTACGTCATTATCCTATTTTACGGTGTTACGGTCAATGAA
TGAAAGCGAGTGA

>CcOR7
ATGTTTCTGCTAAAATACCCTATGCAATACACAAGTTTTTGTTTGAAAATCAT
TTGCTGTTGGCCCATATCAAAGAACGCTTCAAATTTGAGAAAAAATTTCGATT
TTCTTCACTTTCTCTTCGTCATCAGTGGATTTATTCTTATCTCCATTCAAGCAT
TAGTGTACGCAATAATACAGGCGAACAATCCTCAAGAAATAATTGAAACCCT
TAGTACCCAGGTGATATATTTGCATGCAGCGGGTAAAGTGGGAATCATGCGG
TTGCATCGTCATAAATTTAGTCGACTCCTCCACGGAGTTGAGAGAGTTCAAAC
TACAGCCTCTGGAGAACTTCTTCGGTATCATAAAGAGTATACTCGCGTGGGTT
ATGTCGCTTATTGGATCTATACTATCAGCGTGTTTATTATTACATTAGATTATA
TAGTTCAACCTTTGTGGAAAAAATCACAATTCCTACCAACTGGGGCCTGGTAT
CCCTTTGATTACAAGAAATCGACGTTTTTCTACACCCTGGCATATTTTCAACA
AATTATATGCATTACGTTCAGTGGCTGCGCTAGTACTACTGAAATCACCTTCG
GTGTTTTTATATTCTTTGCCTGCGCGAGACTCAAAGTTTTACAGCGAAAATTC
CAGCAGCTTTCTGATAATTCAAACGGTAATGAGAAGATTTTAAGAAGAAGGA
TCTGCAACTATGTTCAACAACACTGCGACATATTACGCTATATCAATGACGTC
AACGAAACTTATACCTTCATTATACTTGTGCTATTCTTGTCAATTCTTCTAACG
GTTTGCTGTACAAGTTTTTTGATAATTAATGTCACCGAACATTCATTGGACTC
CATTTTTATAAATTCCTTACTAATGGTAGCTTCGGCTGTTCAACTGTTATTCTA
CTACCTTCCAGGGCATATACTTATTGAAGAGGCTAAAACAATTGCAGAATCT
GTTTATTACAGCGGTTGGGAATCATTATCAATAAATTGCCGTAAATTACTTTT
ACAAATCATGGTGTATAGTGCATCTCCTATCAATCTGCGGAACGGAAAAATG
GGACTTTTAATCTTGGAGAATTATACCACATTCCTCACCACTGCTGCATCTTA
CTTGACGTCTTTGCGTTCAATTGTGGGAACAGTGTAG

>CcOR8
ATGGATGTCCCTCTACATGGCGGGGAGGAATACGATCAATTTATAAAGCCAG
TGATGAATACTATGCAAATAATATCAATATGGCCTCTAGCAGCAGACTGTGG
ATCTTACGAGTATCTCCTCAGAATCTGCCATCAGATTTTAATGTTTTTCGTAAC
GGGTACTATGTCTATCGTAGTGACAGCTGATGTTATTCACAATTGGGGAAAC
ATGGACACGGCTACCGAATGTTCGCTCATAGCATCTGCCTTCGTTTTGTGCTT
TTTACGATTAATGGTCTACACGTATCATCAGAAGGATATGAGATACGTGGTC
GAAACAATGAGATCtGATTGGGCAGACGCATCCTACGAGGAAAAACAGGTTC
TTAAGGAAAAATGTGATTTCGCCTTTAGGTTGGCAAAATATTTCATTGCTACT
GTTGCAATTACTATAGCATTCTTTATGACCGTTCCTATGCTAGAGACGTATAT
TCTTCACTCCGAGGAAAAAATATTACCATTTCGTGGTTACTTCTTTCTTAACC
ATACTTTGTCTCCGAATTACGAGATGATTTATATTTTTGAAATAATTGCCGGC

TCTTTCGGAGGCAGCATGATAGCTGGCGTCACCAGTTTCAATCTCGTAGTAAT
AATGCATGGAGCAGCGAGGTTCAGTCTTTTACAAAAGAAATTAGAGTCTCTG
AATAGGAATGATCCAGACGTCAACAAGCTTATGGTGAAGTGCATTAAGCTAC
ATCAAGATGCTATCAAATTTGCCGACGCTCTCGAGGACATCATAAACGTAGT
GGCGCTGGGACAATTTGTCACCAGCACGGGCTTGGTGTGCTTCGCAGGATTTC
AGCTGACTTCGATGCTCGAGGACAGAGGACGACTGATGAAGTACTCAACCTT
TTTGAACTCCGCAATCCTCGAGTTGTTCATATTCAGTTTTAGCGGAAACGAGT
TGATAGTCGAGAGCGAAGCTGTCGGCGATGCTGCCTACCGCAGTGATTGGGC
CAGTAGTTCCTTTACTCAGAGTCTGCGAATATTGATGATGCGAGCTACGTTAC
CTAGCAGAATCACTGCAGCTAAATTCTACAGTATGTCCCTTGAGAGTTTCTCA
GCGGTTCTCAGCACGTCATTCTCGTACTTCACTGTTCTGAAAGCTGTAAGCGA
GGAGTAA

>CcOR9
ATGCGTACTCTGCCGGTGTGTTTTGTAATTTTCAAATTCTGGGGTCTCTGGCG
ACCAAGTGATTACTTAAGCTGGATGAAAATTCTCTACAATATCTACAGCTTGC
TAATCGTCGTTATAATGTATAGTTTCATGTTATTCGAGTTTATCGACGTGTTCA
TCAATATAAAAAGTATTGATGATTTCACGATGAATTCATTTATGCTCGTGACG
TTCACTAACGCGTGCTGGCGAGTGGCAAAAATTTTGGCGGGACGTGGGAAAA
TTTTAAAAGTGATAAACTTGTTGACGTCTGATTATTGTGTAGCAGTTGATAAA
CGGGAAGTGGAAATAAAAGAAGGATACGATGGGATAGCCAGGTGGAATACA
CTTCGTTATGTAATATTGGTTCAGATAACAGTTGCACTTATGCTAATTGTACC
CCTTTCCAATGAAAGTCGGAAAAGGATTTTACCATTTAGATCATGGCTACCTT
ATGATTTATCTTCTTTAAAATTATTTTGGTTATCCTATGTGCATCAGTGTATCG
CCATTATGGCTGCAGCTTACATAAATGTGGCTACTGATAGTTTCATTTCCGGA
CTTATGATACAAATATGCTCTCAATTGGACATTTTGAAATATCGTTTGATCAA
GCTGCCACGTTTGTACAATATCAATCATCGAGATGAATTTTTAAGTTCCTTAA
CTCACTGCGTATTACACCACGAGCATATATTCAAATTTGCTAGCGCGGTAGAG
AGTGTTTCTAATCCGATTACAATCGTGCAATTTTGTGCCAGTGCCGTTGCATT
ATGCTCCAGTGTTTATCAGTTATCGAAACACAGTGTTGATAGTGCGAAATTTA
TGCCGTTGATACTTTATTTACTTTGTATGCTGTTCCAGCTATTTTTTTATTGTTG
GTACGGGAACGAGGTCATTCTTAAGAGCATGGAAGTTCGAGATGCGGTTTAC
GAAATGGATTGGATATTATTGGATAACAATAATAAAAAGGCTCTTTTAATGA
TGATGAAACGAAGTGAAAGATCAATTCAAATCAAAAGTGGATATTTTATTGC
CTTATCTGTCGAGCCGTATATTAAGATATTAAAAGCATCATTTACCGCATATA
ATGCGCTGCAACAAATTTCAAATTAA

>CcOR11
ATGACGGTACATAGTCTGTCACCAACTGGAAATTGCAAAGGCCTTAAGGAAG
GTCAAGCCTCTTACTCCTGGTGCGTTCGCAATTTGCGCACATTAATGGTCTTC
ACGGGAATCTGGCCGATGGAACCGCCAACGCTTCTTTTGAACGTGTCTTATTA
TTACAACGCAACAACGTTCACCCTTGTCATTTGCGGAATGATGGCCGGCGCTA
TCAGTGTGATGGACAACTACGATCTGCTCGTGGACAATCTCAGTATAAATCTG
ATATTTACCGAAATCTTCATAAAGTGCATTCTGATAAAGATTTATTCGAAACC

ACTCTCAAGAGTACTGTCGCTCATGAAATTCGATTGGATAAGTCTAAAGGGA
AGGGGTCCCATTCACGGAGACGTACTTCAAAGCGAGAATATTATGCTCTTGC
ATGCCAATATTCCAAGGATATTTTTCATCGCCTACACCGCCCTCGCACTTGTG
GCTTGGACCGCTACAGTGGTGGCGGCTGTTTCCAGAAAAAGTTCAAAAATGC
AAATAGATGCGAGTAATGCATTTCCGATGCCCTCTTGGTATCCGTTCGAGATG
CACTCTACTCCAAATTACGAAATGCTACTTACGTTTCAAGTGATAATAGGATG
TTCAATAGCAGTGAGTAGTGCAGCAGTGGACAGTCTACTCGTGACTGCAGTT
TTCCACGTTTGCGGTCAATTGGAGATTTTACGAAAATATTTTGAAAATTTACA
TTCATCGGAAACGGCGCTGGAGGAAACTGAGAAGAAGGTGGCTGCTGCGATC
AAGAGACATTCGAAACTGATCGATTTGTGCGACCTGATCGAGGATTGTTATA
GCCAAATAACTCTTAGTCAACTCCTAGTAGCCAGCTTGAACGTTTGTCTATCA
GGATTTGGATTACTATTAGCTATAGAATCGGGTAATATTAAAGTATTTCTCAA
GTTTCTACTCCTCCTAATTGCGATGCTTCAACAGATACTCATATATTCTGTTAC
TGGGGATTATTTGTCATCTAAGAGTACAGCCATTCGAAGTGCCATATACAAA
ATGAAATGGTACGAGTTGCCACCTTCGCTAAGTAAAGCCCTAATGTTAATCG
CGATTAGGGCAGAACGGCCGTTGGTCGTTACAGCTGGAAAATTTTTTCCTATG
TCTTTGGAGAACTTCACGCAAGTAAGTAGCCACTGTGCTAAACTGCAAAATC
GAGTAGGTCTGGTACGAAAACTTGTTGTAACTTTCAGATAA

>CcOR12
ATGGAGCATAAAAGAGGGTGGGCCAAAATCAAAAGTGATATTCCAAATAAG
AACTACGATACGGACGTGGAATACGCGATAAAATTAAATCGATGGTTATTAA
AACCAATTGGCGTTTGGCCGTTGGAGTTATCGTCATCAAGAACCGAGAGAAT
TGTCACGATAACCAGCGCTGTTACGTGTTGCCTTTTAATGAGTTTCGTTTTAA
CCATTCCATGTTGCATCGAAATGTTTTCCTCCCAAAAGGATTTTAAATCTAGA
CTGGAAATGCTTGGACCTTCAAGTTTTTGTGTCATGGCTGTTATTAAATTTTTC
TTTTTCGTAATTCGCGGTAAAGAAATTAGATTTTGTATCGATTCCGTCGTTAC
CGATTGGCGAAACGTCGACGTACCAGAGGAACGTAATATTATGTTACGGAAG
GCAAAGTCTGCGCGATTTTTAACTACGGTCTGTGCACTCTTTATGTATTATGG
GGGCATTTTTTACACTACTTATTTACCATTAACGACAGCCAAGGCATTATCTA
CTGAAAATGTTACCATTAGAATTCTACCTTATCGTTGTAATTTTATACTGTTTG
ATCCGTACGCGCCACTGTTTTTCGACATCATTTATTTTCTGCAATGTTTATCCG
CTGCTTTTATGTTCACCTTGACGTCGGGAGTGTGCAGCTTGGCTGTGAATTTT
ATAATCCATGCATGTGGACAGTGCCAAATTATTGGATTACTTTTGGAAAATCT
CGTTGACGGTAGAGGAAATACTTCAACCACCCTTGAAAAGAGAATAGCTGTT
GTCATCGTGCGACACTTACATTTGCTAAGGTTTGTCACACGCGTGGAAGATGT
ATTGAATGAAGTGTGTCTCGTGGAATTTTTAGGTTGCACCCTGAATATATGCC
TTCTTGGGTATTACTTTATAACGGGATTGGAGACAGCTGACACTGCTAGATTC
GTTACTTTTGCTCTACTGTTTATATCCTTCACGTTTAATATTTTTATATTTTGCT
ATATCGGCCAGTTATTGACAAATCACTGTCATCAAATTGGAGAAATATCCTAT
ATGATCGATTGGTATCGCATCCCGGGAGCCCAAGCCCGATTTTTAATATTGCT
AATTGCTATTGCTAATCGACCAGTAACAATCACAGCTGGTAAAATCGTCCAA
CTGTCTTTTCCGTGTTTTCGTGACGTATTAAAAGCTGCATTAGCTTATCTGAAT
ATGCTTCGAAAGATGACCACCTAA

>CcOR13
ATGGACCACGAAATAGAGGAATTTAAAATCAAGACTGCCGTTCAGAATAAAA
ACTTCGACTTCGACGTGGAATACGCGATTAAATTAAATCGCTGGTTATTGAAG
CCATTCGGGATTTGGCCGTTAAATTCGTCATCCACAAAATTTGACAGAGGAAT
ATCAGTAATTTCTTCATTCATCTGTTGCCTCCTGCTGTTTTTCGTGATGATTCC
GTCCTTCGTCGTAATGTTCGTCAGTGAAAAAGATTTTAAAGGAAGACTTGAG
ATTGTGGGACCCACGAGTTTTATTATTATGGTTGTTCTAAAATATTTTTTCCTT
ATCACTCGTGGCGACACTCTTAAAATGTGTATCGATACTATCATTAGTGACTG
GAGTAATGTTCAAGCAAAGGAAGAACGTAAAATAATGTTCAGAAATGCAAA
GATTGCTCGATTATTTACAATAATCTGCGTCTCCTTCATGTATTGCGGTGGAA
TTTTTTACAGTATCTTTTTACCTCTGATAACGGCCAAATCGTTGACAAATGGC
AATAACTTGACTATTAGAATTCTACCATATCGATGCAATTTTATAATCTTTGA
TCCATATAGGCGACCAATTTTTGACATCGTGTATGTTGTGCATTGCTTCTGTA
GTGTCATTATGTATTCTATCACAACGGGCATCTGCAGCCTAGCTGCAAAGTTT
GTGATGCATGCCTGTGGACAGTGCGAAATCGTTATGTCACTTTTGGAAAATCT
TATTGACGACGACAAACAATGTTCCGATATTGTGGAAAGCAAGTTGGCTACT
ATTATTGTGCAACATTTGCACGTGATCAGATTTGTGACGCGCGTGGAGGATCT
TTTAAACGAAGTGTGTCTTGTGGAATTTTTGGGCTGCACTATGAACATGTGTC
TTGTCGGATATTACATTATAACGGGATTCGAATCAGCTGACACCGTCAGATTC
ATTACTTTCACCCTACTATTTATATCTTTCACGTTTAATATTTTCGTATTTTGTT
ATATTGGTCAAATATTGACAAATCATTGCCATCAAATTGGAGAAGCATCATA
CATGATCGATTGGTATCGCTTCCCAGGAACACAAGCACGATTTCTAATATTGC
TAATTGGCGTCGCAAATCGACCAATAAGACTCACAGCTGGTAAAATGGTCCA
ATTCTCTTTTCCTTGTTTCTGCAACGTGATAAAAGCAGCAATGGCATACCTAA
ATATTATCAGAACAGTCACCATATGA

>CcOR14
ATGGTGGAAATAGAATTGCGCAAAAGGATTTTAAGCACAGAAAATAAAGTA
CAGAGGGAACAAAATAAGAATTTCAAGAACGACACAGAGTATGCCTTAAAA
TTGAATCGCTGGCTGTTGAACCCAATTTCATTATGGCCACTATCGTCACATGT
ATCAGCACTCAAGAAAATCCAATTTAAATTTATACGAACAAGCTGCTTCATCT
TCCTGGCATTTGTAATAATCCCAGCCTGTCTGCATACGATCCTCGTACAGAAG
GATCCCAAAGTGAAGTTAAAAATGATCGGACCTCTCGGTTTCTGCGTAATGG
TTATCGTGAAATATTTCGTCTTCGTCGTTCGTGCTAAAAATATAAAAATATGC
GTCGATCATGTTGTCGCAGACTGGCGCAACGTTAATACGACTGAGGATCGTG
AGATCATGATTGACAGTGCAAAGATTGCTCGGATATTCACTGTTGGTGCTGCA
CTCCTCATGTACGGTGGTGGTTGTTCTTATACCATGCTTCTGCCACTGACAGC
ATCCAAATCTGTATCGAGTTCTAACGTTACGATAAGGGTGCTCCCGTACCCCT
GTTATTTCTTTTTCTTCGATCCCCAGGTCAGTCCTACATACGAAATCGTATTTA
GTGCTCAGATCTTGTGCAACATTGTCAGATATTCCACGACATCCGGAGTCTGC
AGCTTGGCTGTCGTATTTGCAATGCATGTTTATGGTCAGTGTCAAGTTTTGAT
GAGACTTTTGGAAGACTTCGTAAACGGCGACAGAGAGAAATCATCAACTTTA
CACCAGAGACTGGCCAATGTTGTTGTGCGTCACTTGCGGATTCTTAGGCTGAT

ATCTACAATGGAAAATATTTTTAATGAGATTTTTCTTATAGAGGTGGTTGGTT
GTACATTCATTTTGTGCTTTCTTGGATACTATTGCATGACGGATCTGACAGAG
AGCGAAACAGTTGGCCTAATTGCGTATTTTCTTTTATTAATATCACTGACGTT
CAATGTCTTCGCGTTTTGTTACATCGGAGAGCTTATAACGAATCAGTGCATGC
AAATCGGACAGGCGGCTTATATGACAGAGTGGTACAGATTAGAAGGGAAAA
ACGCGAGTAATTTAATTGCACTAATTATCATCTCAAATCGTCCAGTATTTCTG
ACAGCTGGAAGAATGATAAACTTGTCTTACAACAGTTTTTGTCAAGTGATAA
AATCATCTTTAGCGTACCTGAATATTCTTCGTCAAGTAACGATGTAA

>CcOR15
ATGACCTGGAATCATTGGCTTCTCGCGATGCTCGGATTTTGGCCGGAAAAGCC
CAGAGATATTTGGTTCTGGATAAACTTCGGTTATTTTGCGTATCACATGACAA
TGGAATACGTAGATCTATTTCTCTTTATCGGAAATCTCGAGCACGTAATCATG
AACCTAACTGAGAACATGGCGTTCTCACAAATTTTTATCCGCATGCTCATGAT
GCGAGTATACAATCGTCAATTAGGAGAATTAATAATGGAAATGAGAAAGGAT
TTTCAAGCTCACAATTACAGATCCGTCGACGAGCAAAAAATATTTTTATCGTA
TAATTCAAAGTCTAAGACGTTTATGAAACTTTTAATGGCGTTTGTTGCACTAA
CCGCATCCTCGTATTACTTGAAGCCTATACTTGGAAATTTGGGAAATGATCCA
ATCGAAAGCGAAAGTACCAATTCAACGTTAACTTTTGAATTACCGTACCGTTT
TTACTTACTGTACAATGTCAATGACACGCATACTTATATGATAACTTATCTTT
CTCATCTTCCATTTGTCTTTGTGAGTGGTTTTGGTCAGAGCGCAGCAGATTGT
CTGATGGTCACGCTGGTTTTTCATGTTTGCGGACAATTGTCTGTGCTCACGTT
ACGAATTAGCAGTATCAATAGCGACCCGTCACAATGTAGTCAGAATATAAAG
GACGTCGTGGTGACTCATCAGAGACTTTTGCGAATGGGACAAACCATAGACA
AAGCTTTCAGTGCCATTTTATTAGGGCACCTTGTAGGTGCAACGTCTTTGGTT
TGTGTGCTCGGATATCAAATTTTAACGAATTTTGCACACGGACAAAATGCCG
ATCTTGCCACATTTTTAACTTTTGCATTTTTGGTTCTTCTTGTGCTTTATGCACA
TTGTACTGTTGGCGAGAGTCTCGTTCAAGAGAGTACCCGAGTCTACGAAGCTT
GGTATGATTGTAATTGGTACAATATGCCAACGGAAAATGCGAGACTAATTAT
TCTTTGTATGTCACGATCGCAAAAGCCGCTTTGTTTGACATCTGGCAAATTTG
GAATTTTTTGCTTAAGTACATTAACTGACGTTTTAAAAACGGCAATGGCATAT
TTATCCGTTTTGCGATCTTTCCTGTAG

>CcOR16
ATGGCGTCACACGCGAACAAATATCTTTCGATTAAGTTGGTACGATTTTTCAT
GAAGCTAGTTGGCATTTGGGTGCCGAAAGATAATAATGAACAATGGATAATG
AACTGTGCGTTGTCATACACGATACTAGTCATACTTCTTGCAGTGGCTATAGA
AGGATTCGATATTTATTATTGTTGGGGAAATTTTTATGCTACGACCTACACAG
CGTGCGCTACTATGCCAGTGATTATTGTATTAGCCAAAATATTTTCCTTTTtGC
TACGACGTAAACTCGTGATGGAGCTAATTGAATTTACTGAGACAAAATTTTG
GCAGGGAGATTATGATTCTTATGGGATGAGTGTACTgGATGATATTGATAGAA
AAGGTGTTCTgCTGATGTGCACGTTTATATTTTTTGTCCAAGGAACCGTCATTG
GTTACGTGCTAACACCAATAATTGAGAATATTGGAAAAAaTGAATCAGACAG
AGTCCTACCTGTCACGCTATGGATAAATATACCGGTAACAACAACGCCTTATT

TTGAAATTTGTTTTGTCATGGAGAGTCTAACAATCGTTCATATCGGAATTTGT
TTCTTTTGCTTTGATATCTTTCTCTGCATACTTAACATACACGCTGCTGGCCAA
TTCAAAATGTTACAACACAGATTCGCACTAGTCTACGACGGAGATAACAAGC
AGGAAATGCCAATCACAGATAAATTACACTTCATGAAAGATTCCGAGGAAAT
TTATATAAAATTCAAAGACTGTGTAAAGCATCACAAGATTCTCATCGATTACA
CAGAAAAAGTGCAGAGTGTCTTTACCTTCATTATACTATGTCAAATTTTAATA
TCAAGTCTAATGATGACCATGGCAGGATTTCAAGCACTTCTCACTCATGGTTC
CATAATCAGACGTCTTATATTTATCGCCCACACTTCGGGTTGCTTCGCCCAAT
TGCTGTTATTCACGTCAACGTGCCACGAGATCATCATTGAAAGTGGTGGTATT
GCCGATGCTGCTTACAACGCTAATTGGGCAAATACCCCTTACGATGACGTCG
GAAGGAGCTTGCGAATTGGTTTGCAGCTGGTAATGATAAGAGCCAGTCGTCC
GTGTCATTTAAGTGCCGGAGGCTTCTGTCACGTGTCCCTCGACACATTCACTG
CTGTATTGAGTACAGCAGCGTCATACTTCACTCTGCTACGACAAATCGGTGCC
GAAACTGTAATGGATACATGA

>CcOR17
ATGGAAGTTCTGCCGTTGTGCTTTAAGTTGTGGACATTAAGCGGTGTTTGGAG
ACCAATGCATTATTCTTCTCCAGTGTCAAAATCCTTGTACACTGTTTACAGTCT
TGCAGTACTTATTACACTGTACAGTTTGACGTTTTTCGAATTTCTAGATATTGT
ATTTAATTTCGGTTCTCTCGATGATTTCGCGAACACCGCCTTTTTACTTTTGTC
GATGATAATAATCTGTTGTAAAGCGACAAATACTCTGAAAAAACGTGCCCAA
ATTATTGAGATACATGATTTGCTGAGGGCTGAGATATGCAGAGGTCAAAATA
CGACGGAGAATTTAATGCTAGAGAAGTTTGCCAAAACCTGTAGATCCAATAC
ACTAAATCTTTTCACAATGATGATGTCATGCGCAATGTGCTTAACTTTGGAGT
CGTTGCTATATCACGTGAATGATCGTCTACTGCCAATGAAAATTTGGTTGCCA
TATAGTCTTTCGTCTCTGACGTTATTTTCTCTGTCGTTTCTGTACCAAGTAGTG
ACTATGACTTTGGCTAATGCAATTACCGTAGCAAATGACACGTACATAACTG
GCTTAATGATTGAGATTTGCGCCCAACTGGAAATATTAAAACTTCGTTTGGTC
GAATTACATGACCCCAAGACTTTACAAATGGATGAGAATTATTCCTTCAATCG
AGAGAAAGATATTCTCAAAATGTGCATTCAGCATCACAATCATATATTCAAA
GTGAAGGCAAATGTTGAAAAAGCATTTAATTCGATATTTTTAATTCAGATTGT
TACAAGCACATTGGTTTTCTGCGTCACAGCACTAACCTTCTTAAAGCACGAGA
TATTAAGTGTTGAATTGGCCAGTGTGATTTTGTATTTTTTTACAATGATGTTCC
AACTGTTTTTATTTTGTTGGTACGGAAACAAAGTTATTTTAAAGAGTACCGAA
GTAAGGAGTGCTGTATTTGAAATGGATTGGATTCCTTTACCTCAGGCAATTAA
AAAGGATTTTATATTTCTTATGATGAGAACCAACATACCAATCCGATTTACCA
GTGGATACGTTGTCACTCTGTCTTTGGATTCGTTTATGGCTATTTTGAATTATC
GTACACCACATTAA

>CcOR18
ATGAACAGAAACAATTTCGTAAGGGACAACGTGATGATTTACAAGTGTCTGG
GTTCCTGGCCACTCGACTGCGTTGCCAGACGTCTCGAGAACACCCGATATATC
CTCGGCTTGACCCTCATCGCTGTAATCACTGTTTTCAATGGTCTACAGTATCTC
GACCTCTTCATTGAGTGGGGAGACTGGGGATTTATATCCGAAAATATCAGTG

TTTCCTTTATCTACAGTATATTCGTGTCCAAGATTTACGTTTTTTATTCACGTC
GTCGAGAAATTATGGAAATGGTCCTCGAAATAGATAATTATGTGCGAAAAAC
TTACAAAGATGGTCATCAAAGTAACGTCCAGATTCTACAATCCTGTGAAGAT
ATTGCCAGACGAATAAAAATTGTCTTCGGGTGGTCTGCTCTCCTGACTGTCAT
TATCTTGCATACCTGGCCAATAGCTTCTATGATATTTAAAAAGGACCCCAATC
TGAGAATGCTTGGAGTTCCTGCATATTTTCCATTTTCATTAAATTCGACAAAT
AATTATGCCATTGCTTATGTTGTTGAGATCGTTACAGCATCTGTCATGTCTTTA
CACACAGTTAATTTTGACACCTTCTTTATATCTTGCATCCTTTTTGCCATAGGA
CGCTTAAGAATTCTACATAACAGCATAAACAATATAAAAAAGATAGCACCAT
CTGATTTACAAATGAAATTTAGTAAGAATGAGACATACCAGGCTGAACTTCA
AAAATgCATTCACCAATTATTGTCTGAGGCCATTGCGGAGCATCAAAAAATTA
TTTTATTCTGCCAAGTGATTGATTCTATGCTCAGTTGGATTATGTTTATTGGCC
TTTTCaTATATTCTGTCATGTTGTGCTTTGTTGGGATGCGAATAATTATGATCG
GACCATCAATAAAGTTATTACAACTAATGGAATATTTGGCAATAATGATGAC
TCAGATATTTCTATATTATTGGCATGGAAACGAGTTGAAATTGGAGAGTTTAA
AAATTCATGATGCTGCATACAATTGTGATTGGTATAGTTTCGATCGCAATTCT
CAACAGACTATAATGTTTATAATGATGCGTGGTCAAAAACCGATTTATTTGAC
TGCAGGAAAATTTTACTATGTCACTCTTGAGACATTTTTATCGGGGGATCCTA
CTCATATTTTACTGTAA

>CcOR19
ATGGGTATTGCTAATAATATAGAATTATTGAATACCAGTGTTGGAAGAAGAT
GGAGCAGTTATCAAACGGTACTAGAATATAAAATGGAGGTCGAGTGTCTCAA
ATTGGCAGGACTTTGGTATCTGTCATCGGACACTTCTGATTGCGTCAAGGTGA
TTTACTGGATTTATAATAAAATTGTGTTTGCTGCCATCGCCATTTTTAGTATTG
GCCTTCTTACGGATATTATCATTAATTACAAGGATTTAGTGACTTTCACGGAC
AGTGGTTGTATTTTTGTTGGCATATCGGTTGTAAGTTTCAAAGCTGTCGTTTTC
CAGTGTAAAAAGTCAAGGATTGACAGGCTTGTCCATTCGGTCAGACAGTGTG
AGTCTTGGACGAGTGATTTGGTTGATCCTGCCGCAATGTCGCTGCTGAAGAA
ATACAAAATTCGAGAAAGGGTGACAATCTTGGGATTCTCATGTTTGGGATGT
GTGCTGGTGATAGCTCTGATATTTTTTGTACCAAGAGAAACAGGAGAACTGC
CAATACGTTGTTGGTATCCTTTTGACACAACGGTGACACCGATGCACCAGATT
GTATTTGGGATACAGAGTTTTGCTGTTGCTGTAGGAATGATCGCCATTATTGG
AATGGACAATACCGTCTTCGTTCTTTGTGGTCGTGTGCTGTTTCAGCTGGAAA
TATTAGCTGCGAATTTTCAAAGTTGTTCGCTGGATATTAACGCCTCTGGTGGA
AGATCTGAAACTAATCCTCATAAGGATTTCACCTGTTTCCGGCACACAATAGA
CTGCGGAGGCTTTTTTGAGAGATACCGGAAGTGCGTCATGCATCATCAATATC
TTATTTTATTAGTTGATGAAGTGAATGATGTATTTGGCTCATCCATGTTTTCCC
AATTACTGAGCAGCAGCTTGATAATATGCTTGACAGGATTTCAATCATTATTG
GTGGTTGGGCACGGTACGAATTTTTTGAAATTTGCTATTTACCTAGGAGCAGC
TTTCAGCCAATTGCTTTATTGGTGTTGCTTAGGAAATGAACTGAGTCATCAAA
GTTCCCTTTTACTGGAATCCCAATGGAAATCAGGCTGGGAAATGCAGCCGGC
AAAGAAAGTACTTCATCCAATGGTATTTTCAATGATGAGATCCAATAAGAAG
CTTGAGCTGAAAGCTAGCAATTTTTTTATAATGTCCACAGAGACATTCATAAC

GATTCTTAGCAAATCCTATTCCATATTTGCTTTATTAAATAACATGATTCAGT
AA

>CcOR20
ATGGCGTTTCAAGTGACTCTACCGCGAGCTATGCAATTTTTGCTAGTCTGTGG
AAAAATATCATGCACGTTACCACCAAATTTTGAAAGCCGAAGATACAAATCA
ATTATCCACGAAATCATGTGGTTGATATCTATGTCTGTGACAATGAGTTTAAT
AATTCCTCTTTTTCTGGCAATGTATCACTTTGGTGAAGATTTGGCAGTATTAA
CAAAATCAGTGTCAGAATCAATGACCGCGACAGATCCACTTTTGAATATGAT
TTTCTGCAAATTAGAGAGAGCTCGAATTCAGGTTATTCTCAGTGAGTTAATGA
ACTTCCTAAATACTGCATCTCCGTTTGAAAAAGCTAAACTTCAAAAACACGTT
GACCAATGCACACCGTTCTATCTGTTCATACTTATTTTATTTTCACTGACTGCT
GTTGCCTTTTCATGCGGGCCCATCGTAATGGACAGGCCATTTCCGGCTGAAGC
GTGGTATCCTTTTTCAACGGATCCACTTCCTATCAGTTGTTCCATATACGTGCT
TCAAATTATGGTCATAGCACAAGCCGCGATGTGTATACATATGGATTTTATGA
TTGCATTTCTTTTGTGGTACGCCGTGGCGAGTTTTGAAATATTGGGAGAGAAA
TTTCGTCGTGTCGAAAACGAAGTTGATCTTCGAAATTGTATCGTTCAACATCA
GAAATTGATAGTATTTGTTCGTGAAATTAATCGTGTCTTTTATGTAATGATTCT
AAAAACACCTCTATCGATGACAATTAGCATAATCTGTTCTAGCGTTCAGCTGA
TTCACCACGAACCTCTGCCAGTATTGTCCCAGTTTATTTTGCTGCTAGTAGTTT
GCTCTATAAAGCTATACATCACTGCTTGGCCAGCGGACAGTTTGATCCACGCT
AGCGAAAACATCGCTTCGTCAGTTTACGAATCATCTTGGATTGGTAAATCTTC
AACGTTTCTGAAGAGTCTGAACATTGTCATACGAAGAAGTCAAAAACCTCTC
ATTATTTCAATCACTGGCATACTTCCTCCACTGTCTCTACGATTTTACGCAAGC
TTTATCACCAAAGCACTGTCGTTCATCACGACATTGAAGAGTGTTATCAGCGA
GTGA

>CcOR21
ATGGTGGAAATAGAATTGCGCAAAAGGATTTTAAGCACAGAAAATAAAGTA
CAGAGGGAACAAAATAAGAATTTCAAGAACGACACAGAGTATGCCTTAAAA
TTGAATCGCTGGCTGTTGAACCCAATTTCATTATGGCCACTATCGTCACATGT
ATCAGCACTCAAGAAAATCCAATTTAAATTTATACGAACAAGCTGCTTCATCT
TCCTGGCATTTGTAATAATCCCAGCCTGTCTGCATACGATCCTCGTACAGAAG
GATCCCAAAGTGAAGTTAAAAATGATCGGACCTCTCGGTTTCTGCGTAATGG
TTATCGTGAAATATTTCGTCTTCGTCGTTCGTGCTAAAAATATAAAAATATGC
GTCGATCATGTTGTCGCAGACTGGCGCAACGTTAATACGACTGAGGATCGTG
AGATCATGATTGACAGTGCAAAGATTGCTCGGATATTCACTGTTGGTGCTGCA
CTCCTCATGTACGGTGGTGGTTGTTCTTATACCATGCTTCTGCCACTGACAGC
ATCCAAATCTGTATCGAGTTCTAACGTTACGATAAGGGTGCTCCCGTACCCCT
GTTATTTCTTTTTCTTCGATCCCCAGGTCAGTCCTACATACGAAATCGTATTTA
GTGCTCAGATCTTGTGCAACATTGTCAGATATTCCACGACATCCGGAGTCTGC
AGCTTGGCTGTCGTATTTGCAATGCATGTTTATGGTCAGTGTCAAGTTTTGAT
GAGACTTTTGGAAGACTTCGTAAACGGCGACAGAGAGAAATCATCAACTTTA
CACCAGAGACTGGCCAATGTTGTTGTGCGTCACTTGCGGATTCTTAGGCTGAT

ATCTACAATGGAAAATATTTTTAATGAGATTTTTCTTATAGAGGTGGTTGGTT
GTACATTCATTTTGTGCTTTCTTGGATACTATTGCATGACGGATCTGACAGAG
AGCGAAACAGTTGGCCTAATTGCGTATTTTCTTTTATTAATATCACTGACGTT
CAATGTCTTCGCGTTTTGTTACATCGGAGAGCTTATAACGAATCAGTGCATGC
AAATCGGACAGGCGGCTTATATGACAGAGTGGTACAGATTAGAAGGGAAAA
ACGCGAGTAATTTAATTGCACTAATTATCATCTCAAATCGTCCAGTATTTCTG
ACAGCTGGAAGAATGATAAACTTGTCTTACAACAGTTTTTGTCAAGTGATAA
AATCATCTTTAGCGTACCTGAATATTCTTCGTCAAGTAACGATGTAA

>CcOR25
ATGGACACGCATAGAAACAAATACTTGTCGATAAAGGTGACGAAATTTTTTA
TGAAAATCGTTGGCATATGGCTCCCAGAAAGCAAGCATGAACAATTTGTCTT
GGACGTGTCTCTGTTCGTTACCATCGCTGGCACTGTTCTTTCTATCCTGTTTGA
GATGTGGGATATTTACAAATATCCATTTAACTTCAACGAAGCAGTTTATATAA
TATGCAATATACTTACACCAGGCATTGTCTTATTCAAGCTATCAATGATTCGA
CTTAATCGAAAGTCTTTATACGAACTTATCGATATATGTCAAACCAAATTTTG
GCACGACGATTACGATGAGTTTGGTATTGCCATTTTGCAAAATTGCGAGACG
AAATGTGTCCTTTTGATTACGTCATATATGTCTTTTGCACTATTCACAGCTATT
ACCTATACCGTACGATCGATCATAGATAATATAGGTAAAACTGGAACAGACA
AAATATTACCTTTTACGATGTGGCTCAATGAGACTATGGCAAGGGCACCGTA
TTTTCAATTGCTATTCATTTTTGAGGGAATTATACTGTGTTACCTGGGAGTCG
GCTTTTTTTGCATCGACAATTTCTTTTGTATAATCAATATTCACGTTGCTGGTC
AATTCAAAATACTCCAAGGCAAATTGGAAAGATTGTGCGGACCGAATGATCG
TGAAGACGAAAAAAAGGATCGTGGAATTTGGATAAGAAAGAATCCAGCGGC
CGTCTTTCAAGAATTTCGAAGTTGCGTTCAACTGCATAAAATGCTTATTTACT
ACGTGGAAAAGGTCGAAGCCATCTTTAGCCTCATAATACTCTGTCAGGTGAT
ATTGTCAAGTATACTCATGTGTCTAGCCGGTTTTCAAGCAGTTTCCGACGACA
ATTCGGCAAGCCAACGTTGTATTTTTACTGCGTATACAATTGGTTGTTTTTTTC
AACTGTTACTCTATACGTCAACGAGCAACGAAATAATAGATGAAAGTCTTGG
CGTAGCCAATGCGGCTTATCGGGCTCATTGGTATCTTTTGCCTTTCGATAAAA
CAAGTAAAGTTATCAGGGGTTGCTTAGTGCTTGTCATATTACGAGCAAGACG
CCCATGTTCTTTAACGGCAGGAAGATTCTTTGCCATATCCTTAGAAACATTTA
CAAAAGTGATCAGTACCACCATATCGTACTTTACTTTACTAAGACAACGTGG
AGAGTCTTCATAG

>CcOR26
ATGTCGAATTTGATAGAAGGTTACACGTCGACGAAGATATCGCAAATTCTAA
TAACACTTATTGGTATGAAACGCGGTAAAACTAAACGTGAACAGTTGTTAAT
AGACGGGTTATTGGTTTATATATTTGCGACGGTTTTAATCGCAATATTTGTAG
AAAATTCGGATTTAATTTATTCACGGAACGATCTCTATGCACTCACTTATAAT
GCACCATGTTCCTTCGCCGTGACATATGATTTCGTAAAATTACTGATATTCAC
GTACAAGCGTCGTGAACTGTATGAACTGCACAAATTTACTGAAGATACATTTT
GGAATAAGGATTACAATGAACTGGATAAAGCAATATTGGACAAGTGCGATA

GTACAAGTGCCATTGGAATGTGTGTACTATCTGTACTAGCAACCATAGTTGCA
ATTCATTATTTGACGGGACCATACTGGGATGCCGTAGGAACAAATACCACTG
AGAGAACGTTACCGTTTCGAGTAGTATTCGACTTGCCACTTACAGTGACCCCA
TATTATGAGATATCATATGTTATTGAGGTGATTGGAGCATTTAGTGTAGGTCT
GTGCAGCGTAGCATTCGCCTCGTATCTATTCTACACGTGCACCTTCGTTTCCG
GCCACTTTAAAATCTTGCAACGAGAGTTAGAGAATGTGTGCGAGGTAGAACT
AAAAATTTTAATCACAAAATCATCATATAGTGATAATGATGCTAAGCTTGCAT
ACGAAAAATTCAAAAAGTGCATCGTGCAACATGAACTACTAATTGGGTATTT
GGGTAAACTGGAAAGTTTATTCTCCTACATATTCTTAATGCTCGTATTATGTA
TCGTCATAATACTCTGCTTTTCTGGATTTCAATTCATTTTGGGAGATGGAACTT
CGAAACTCCATCGTCAAATTCTAAGTGCAGAATATATTGTGACTACTTTGGTT
GAAACAGGCCTCTTTGCTTTCTCGTGCAATGAAATATTTGAAGCCAGCGCCGC
TATTGGGGAAGCTGCGTATCGATGCAAATGGTACAAGTTACCGTGCGATGAA
AATGGTCGAGCACTGAGACAGGGAATGACCATAATGGTCATGAGGTCCTACA
AGCCGTGCTCATTGACAGTTGGCAAATTTTGTCCCATGAATCTACAAGTGTTC
AGTTCCGTACTGAGTACGTCATTATCCTATTTTACGGTGTTACGGTCAATGAA
TGAAAGCGAGTGA

>CcOR26b
ATGCCAACTATTGTGGAAGGTTACACATCGACGAAGATATCGCAATTGCTAA
TGACACTTATTGGTATGAAACGCGGTAGCACTAAACATGAACAGTTGTTAAT
AAACACGTTATTGGTCTACATATTTACGACGGTTTTAATTGCAATAAGTATCG
AAAGTTCAGATTTATTTTATTCACGGAACGATCTCTATGCCCTCACTTATAAT
GCACCATGTTCCTTCACCGTGACATTCGATTTTGTAAAATTACTGATATTCAC
ATACAAGCGGCATGAACTGCATGAACTGCATAAATTTACTGAAGATACATTT
TGGAATAAGGATTACAATGAACTGGATAAAGCAATACTAGACAAGTGCGATA
CTACAAGTGCCATTGGAATGAGTATATTAGCTCTAGTATCAACTATCCTTGCA
TTTCATTATTTAACAGGACCATACTGGGATGCCTTAGGAACAAATACCACTGA
GAGAGCGTTACCATTTCCAGTAGTGTTCAACTTGCCACTCACAGTGACCCCAT
ATTATGAGATATTGTATGGTACTGAGGCGATTGGAGCAATTAGTATAGGTATT
TGCAGCGTAGCGTTCGCCTCGTATCTATTCTACACGTGCATCTTTGTTTCCGGC
TTCTTTAAAATCTTGCAGCGAGAGTTGGAGAACGTGTGCGAGGTAGAACTTG
AAAGTGTAAACACAAAATCATCATATAACAATAATGATACTATGCTTGCATA
CAAAAAACTCAAAAAGTGCGTCATCCAACATCAACTACTAATTTGGTATTTG
GATGAACTGGAAGGACTATTCTCCTACATATTGTTGATGCTCGTATTATGTGC
CGTCATAATACTCTGCTTTTCTGGATTTCAAATAATTTTGGGAGATGGAACCA
CGAAACTGCATCGTCAAATTTTAAGCGTAGAATTTATCATGGCTGCTCTGGCT
GAAACAGTCCTCTTCGCTTTCTCGTGCAATGAAATATTGACAGCCAGCGCCGC
TATTGGGGAAGCTGCGTATCGATGCAAATGGTACAAGTTACCGTGCGATGAA
TATGGTCGAGCACTGAGACAGGGAATGACCATAATGGTCATGAGGTCCTACA
AGCCGTGCTCATTGACAGTTGGCAAATTTTGTCCCATGACTCTGGAAGTGTTC
ACTTCCGTACTGAGTACGTCATTATCCTATTTTACGGTGTTACGGTCAATGAA
TGAAAGCGAGTGA

>CcOR27
ATGACGTCAAAAGAGAGTGAATATCGGTCGGTAAAATTTATGCGAATTTTAA
TGAAAATAGTTGGCATGTATTATACCGAAAATCCTCGTGAACGATTGTTGTTA
CGTGTCGCTTTGACATACGCAATTATTGCAATCTTATTTGCACTCGCCGTAGA
GTTCGTTGACCTGTATCACTGTTTAGGAGATTTTTCGGCTGTCATGTACAATTT
ATGCTCAACAATGCCCCTTGTTATGGTTTTGGTGAAGATTTCGAATTTCTTATT
TCACTGGAACGTCATGATGTACCTAATCTCATTTGCGCAAAATAATTTCTGGT
GTGATCCTGACGACGACTTCACCCGCGAGACAATGAAGCGTTGCGACAAATA
TGGAAAAATATTTGTCTACCTCTTCACAAATTTGGTATTGTTCGCAGTATTGG
ATTACATCTTTGCACCAGTTGTGGAAAATTTTCATCGAAACGAAACCGACAG
AATTCTTCCATTCACACTGTGGGTCAATTTACCAGTCACTGTCACGCCTTATT
ATGAAATCACTTATACTATACAGTCTTTGTCAACACTTTACACAGGAATTTGT
ACGTGCTTCTTCGACAACTTCATATCAGTCCTAAATATTTACGTAGCTGGTCA
GCTGGAAATTTTGGGACATCGTGTCGAAACTGTAGCAGACACGTGTATAGAT
ATCACCATGAAGGATTACGTGTCGGAGAAATCAAAATTGGGACTTTCGTTAA
CCTTGAAAAAATTTAAAAGCTGTATAAATCAACATCAAATCCTGATATCTTAT
ATAGAACAAATGGAACGTGCTTTCACTTTGATTCTACTTGGACAATTAATACT
TTCCAGTTTTGTtATCTGCGTCGGTGGTTTTCAGCTCATGGCCACGATGAGCTT
CCTCCGGAAGTTTACATTTATCTGTCACTTCATGGCTGGTCTCACGCAACTCC
TACTGTACACTTGGAGCTGCAATCAAATAAGTGAAAAAAGTCTTTACATCTC
GCAAGCAGCTTACAACACGCGTTGGTACCTTCTGCCATACGATGAAACTGGA
AAATCGTTGAGAAACGGAATATTATTTCTTATGCTGAGGGCGCAGCGGCCGT
GTCAGCTTACAGCAGGAAAATTTTCACCAATAACACTGCAAACATTGACTGC
GATATTAAGTACGGCTATGTCTTATTTTACAATGCTACGACAAATGAGCGAA
GACGATGTGTGA

>CcOR29
ATGAATGATAATAACGATATACCTAAAATCGTTTTTAAATGGAACAAACTAT
TTCTGTCTATTGGTGGAAACTGGCCCCTAGAACCAGCCTATTTCAGATTCTCC
ATATGGATGTCATTCTTTACCTTGAATATCAGCCTCCTGTATGTGGATCTGTAT
GAAGTTTTCGGCAATTTGGAACAAATGCTGTTAAATCTCTCAGACAGCGTGAT
ACAAAGCTTGATACTTGCGAAGCTGCTGCTCTTTCGATTCAGCGAACCACTGG
CACGTTTGATCACTGACGCCCAAGAAGATATCACAGCCGGAGAGTTTGGGAG
TCTGGAAGAGAAAAAATGCTTTCTGGAATATTATCAGCGAGGCAAAATTTTC
TATCAAATAACAATGTCATGTGTGGTGTTCGCTGTTCACGTATACTTCTTCAA
AGGATTAGAAACATATCTATTCTCAGTGTGGAACAACGAAACGACAACTTCT
TTTGTTCTTCCGTACAGAACAAAATTATTTTTCAATTTGACAAGTCCTAATACT
TACATAATTCTGTATATACTTGAATATCCAATGGTTTATATTATTGCATGTCA
GGGGGCAACTATTTGCTTACTGGTCACTCTTATATTTCACGTTTGTGGTCAGTT
AGCTATTGTATCATATAAAATCCAACATTTAAGAGGAAATTTACCTGAGCAT
GAAAATCGAATTATTAAAAGTTTAATCGAGAAACATGTTAAATACGTTAGGA
TGGCTAGATCTTTGGACGATACTTTTCAATTCGTTTTGCTTTTAGAAGTAGCA
GGAACTACAGTTATTCTCGGCCTCACTTGCTATAACATTATAGCGAATAGTGG
TGATTTAGCTGACATATCAACATTATGCTGCTTTGCGCTGTATGCATCTTCTAT

GATCTTATTGCTTTACGGATATTGCTTCGTTGGCGAATGTCTGATACACGAGA
GTACAAAGATACACGAAGCGTGTGCTCAATGCACATGGTATAGTATGCCACT
GATCTACCAAAAAGCTTTAATAATGTGCATGTTATGTGCACAAAGGCCTTTAC
AATTGACAGCTGCGAAATTTTACGTATTTTCTTTAGACAGCTTCAGCAATGTC
ATTAAAACATCCATCGCTTATGTATCCATGTTACGTACTGTCGTATAA

>CcOR30
ATGGAAAATCAGGAGTCATTCGATGAAATGATACGGGTGAATCTTTTTCTTCT
GAAATTCTCAGGAATAATTCGTGCGTCCGTAGGCAAAATATCCCTGACAGAA
ATCTTGTCAGTTTTGGCATTTGGACTTTTGGCAGCAATGTCTTGCTCCTATTTA
AGGGATCTGCTGCTAACTGATAATATGGAAGTGGCGATGCAATCTTTGGCAT
TTATGATAACTGGCTTCGGAAATACCTTACATTATATACTGATAGCTAAAAGT
CGTGCAAAGCTTTGTGACGTATTGGCAACGTTCGAGGATCTTTGGGGGCTCTT
GGAAGTTCAAGAGAAACGAGTGCTAATGTCTTACGTCAGGGACGCCAAGAA
ACTCACGTATTTCTTTATGTCACAGTGTGCAGCCACCGTTTTTTTATACGTTGG
AGCACCGGTTATTTTTGGAAATGGTTTTGTCAGGGTCAATGGAAATATTACGG
AAAGGATGTTACCATATTCATTGATCTTTGAATGCAAAGACTCACCCTGCTAT
GAAATCCTTTATATTCTACAGATACTAACGGTTATCAACATAGCTATCACGTA
TATCGGTGTTGATACCATTGGTCCTGTTTTAATTTTGACTGTCAGTGGCCATAT
GAAGATTATCCAAAATCGAATAATGTCATTGGGATCGTCCGAGGAATTTGAA
AATTTCAAAAACGACAAAGTGAAGCGAATATCTGATTACGAAATTTCGCATT
TTCAGTATTTTGGGAAATCTTTCGAAGCCTGTGTCAAGTATCATCAGACTGTT
CTAAAACTTTGCAAGGACATCGAAAAAGTCACGAACAAAGCTTTTCTCGTGC
AACTAATAACGAGCACGTACAGCATTTCGGCGATTGGTTTCAAGATGGTTGG
GAACGATTCCGACAAATCGAAGTACGTTACGCAAATTGTACTTTCGTTGGTTC
AGCTGTTTCTTTGCAATTGGCCTCCCGATGTTCTTCAGAACGAAAGCCAGGCA
GTGGCGTATGCTGCATATTTTATGCCCTGGTATCGTTGCAGCAGAGACGTCAA
AAAATCCACGGAAATCATAATAATGAGAGGGCAACGTGTTGTTCGTTTAACA
GCTGGTAATTTCGTCGACCTGTCATTGGAGACATTTATTAGGATGGTTTCGTC
GGCATTGTCTTTTTTCACACTCCTCAGAAGTATCGAGTGA

>CcOR31
ATGGTGGGAATAGAATCGTGCAAAAGGATTTTAAGCACAGAAGATAAAGTA
CAGACGGAACAAAATAAGAATTACAAGAGCGACATAGAGTATGCCTTAAAA
TTGAATCGCTGGCTGCTAAACCCAATTTCAATATGGCCACTATCGTCACATGT
ATCAACACTCACGAAAATCCAATTTAAATTTATACGAACATGTTGCTGGGTCT
TTCTGGCGTTTTTAATAATCCCAACCAGACTACACACGATCTTTGCGAAAACG
GATCGCGCCGTGAAATTAAAAATGATCGGACCTCTCGGTTTCTGCGTAATGGT
TATCGTGAAATATTTCGTCTTCGTCGTTCGTGCTAAAAATATAAAAATATGCG
TCGATCATGTTGTCGTAGACTGGCGCAACGTTAATACGACTGAGGATCGTGA
GATCATGATTGATAGTGCAAGGATTGCTCGGTTATTCACTGCTGCTTCTGCAC
TCTTCATGTACGGCGGTGGTGTTTTTTATACGTCCGTTCTGCCACTAACAGCA
TCCAAATCTTTATCAAATGATAATGTTACGATAAGAGTGCTGCACCCCTTCTA
TTTCTTTACCTACGATCCGCAGACCACTCCTATGTACGAACTTGTATTTCTTGT

TCATACCTTGAGTGACGTTGTCATGTATTCCACGACTTCCGGAGTCTGCAGCT
TGGCTGTCGTATTTGCAATGCATGTTGATGGTCAGTGTCAAGTTTTGATGAGA
CTTTTGAATGACTTCGTTGACGGCAACAGAGAGAAATCATCAACCTTACACC
AGAGATTGGCCAATGTTGTTGTGCGTCACTTGCGGATTCTCAGGCTGATATCT
ACAATGGAAAATATTTTTAATGAGATTTTTCTTGTAGAGGTGGTTGGTTGTAC
ATTCATTTTGTGCTTTCTTGGATACTATTGCATGACGGATCTGACAGAGAGCG
AAACATTTGGCCTAATTGCGTATTTTCTTTTATTAATATCACTGACGTTCAATG
TCTTCGCGTTTTGTTACATCGGAGAGCTTATAACGAGTCAGTGCATGCGAATC
GGACAAGCGGCATATTTGACAGAATGGTACAGATTAGGGGGGAAAAATGCC
AGTAATTTAATTATATTAATTATCATCTCAAATCGTCCAGTACTGCTGACAGC
TGGAAGCATGATGATTCTGTCTTACAACAGTTTTATTCAAGTGAGCATGAAAA
ACAATGTAATAATATTCTGGCTATTCCCTGATACGTTTCAATCAGAAATTTAA

>CcOR32
ATGGACATATCCGACGCCACGAAAGTGTTAAAATGGAACAAACGATTGCTCG
ACGTTCTAGGGCTGTGGCCTTTGAACTTAAATGAAGTCAAATTCTCTTTTTTCT
TTATTTACATAACAGTTCAGTGTTTTTTACAGTATGCCGATTTAGCAGAATAC
ATTTACGACTTTAACTATGTTGTCAGAAATTTAACAGAGACCATCGTTTTGAG
TATGATTTTTCTCAAAATTTTTATATACAGAATTATTTCAAATGAGTTGCGCG
AATTAATACAGTACATTCAAGAGGATTATTCGGAAATAGTTTACAGTACGGC
TAATGAAATAAAAACTTTTCTGCAATATACCTTATTGTCAAAGAGAATTGTTC
AGTGCTTATTAATAACATGTGGATTTACTACGGTGCTCTTTTACATGCAGCCA
TTAACTACTCAATTATTGGCATACAATGAGGGAAATTCATCGACTTCCTTCAT
ATTGCCATTTCACATTCGGTTATTCTTCGATGTAACGCAAGCGCGAATATATT
ACATTATATATGCCTGTGAGATATTCGTCGTCCCATTGGTTTTGTGTGGATTTA
TTGGTACCGATTGTCTATTGATAACGCTGGTGCTACACATTTGTGGGCAAATT
TCGATTTTGACAATGCAAGTGGGAATTTTGTTGGACGATCCACGAAATCTGCG
TGAAAAGCTCAAACGGATCGTCATAAGGCATTGTCGTCTGCTAAGACTCTTTG
CAAATTTGCAGTCCGCGTATTCTGCATTCCTACTGCAAGAACTTTTTGGTATA
ACTTTTCTAATGTGCTTAGGAAGTTACAATGTTATCGCGACTTCTGCCGTGAC
AGACAGCTCAAATTTTATAAGATTCTTATTTTATATTCTTACTCTAATATTTCA
ACTTTTTGGTATGTGCTATATAGGGGAATGCGTTACGACTGAGAGCTCATGCT
TGTGCAACGCATTTTACAATTGTGAATGGTACAACATATCACCAGATCATGCA
AAATCTTTCCTGATGTGCATATTACGATCGCAAAAGCCTCTCACATTAACTGC
TGGAAATTTCTTTACATTCTCCTTGGTAAATTTCACAAGTGTCGTAAGAACGT
CCATCGGTTATTTATCAGTGATGCGAAAATTTTTATAG

>CcOR34
ATGATTCAACAGTTTCAAATAGAAAAATTTATTTATCTCAACAAATGGTTTCT
AAACATTGTTGGCCTATATCCACAGAAATCGTGGAGATTTGTTACATCATTAT
TTTGCATGTTGCTTCTTGTGATACCGCAGTTTGTGCAGATTTATTTATTCTGTA
CTGATCTGTCAGTGATACTTGAGACGAGTTCGGTTTTATTCACAATTCTGTTA
GCGATGCTAAAAGGCGCCGTGTGGATATTCAATGGTAAATCTATGGAGGACT
TAGTGGGATTTTTATTTAATGAGTATTGGAAGATAATAGAGTCTTTTGACAAT

TCTAAGACTCTTATGAAATATGCTGGGTACTATATCTGTGTATATTATTCTTAT
GCTATATATATTTTTTGTTTCTCTTGGATTTTTCATTGGAAAGATATTTCTGGT
TTTAGATACGCTAGTAAGATGACCAAGGGTTATACGTTTTTAATAATAAACGC
TTTACTATTTTTCTTCAGTTTACCACCCATTGAGATACTCATACTAAAATTTAA
TGGCACTTATAAGAGCAGTGATAAGCATTTTCCATTCCCTGCGACTTATCCAG
AATTTCTTAAGGAGTTTCCATATTTTCAGATTGTTTATTTCTCACAAATTATTG
CTACAATGACTTGTGCTTTGGTGATACTTGCCACTGACACTTTAATAGCAACC
GCGTTGTTTCACACTTGCGGTCAGTTTGTTGTTATACAACAAAAACTAGAACT
TCTCAGCAGTAACGACGGACATGAGACGGGTGAAATGATCAAGAGGAAAAT
AATATTAATAGTGAAGCATCATCAAATGGTGATAGGGTTTTCGAATAAAATG
GAAAGGGTATTTAGTCCCATGATGTTTCTTCAAGTATTTGCAAGCAGCATGAT
AATTTGTCTTGTCGGACTACAAGTTAGCACAACATTTACCAACCAGTACAAAT
TAGTCAAGTATTTTTCGTACTTGCTAATGGCGTTGTTCCAGTTGCTACTTTTTT
GCTGGCCAGGTGATCAATTGTTATTGCAAAGTGGCAAAGTATGTCAGTCCGC
ATACTTTACAAAGTGGTACCAGTTTGGTAAGGAACGGAGAGGTGAAATACAG
TTGATGTTGTTACGTAGTCAGAAGCTAATTGGTATAACAGCTGGCAAGTTTTA
CTTGATGAGCTTAGAAAATTTTAATGTAATATTGAGTACCTCCATGTCGTACT
TTATGGTGCTCAGGAGTTTCAACTCATCGGAGGGTAAATAA

>CcOR36
ATGTCGCGCCGTACAGATATCGGTGTGTCGGTAAAAATAATAAAATTTTCTAT
GAAATTAATTGGGTTATGGAAAGCTGATAATACAATTGATAAAGTTGTCACG
ATTTTAATTTATACATGTCTATTAACTCTGATGACCATCGGCGTGATGTTATC
GTTAATAGACTTTTTTTACACTTTGGATGACATTTATAGTGCTGTGAATATATT
TTGCGCAATGATTGCTATTCTTAATTTAATCGGGAAATTAACTATCTTTGGTA
TTAATCGACAGAAAGTATTTGACATTATTGACCGTATAGAAGATACTATTGA
AAATGGAAAGAAAGAGGAATACGGTATAAATGCGATTATAGATTGTGAGCG
TAAGTGCGTCGTGTTGGTAATAATCCTTGGAATTCTAACTCAAGGGGCAACTG
CGAATTACGTGATTCTTCCTTTCGTGGAAAATTTCAATGGAAACACGAGTGTA
AAGAGTTTACCAATCCAGATCCATGTTACTAAGCTCTCTCACAGCGAATCACC
GTATTACGAAATTGGATTTGTTCTCGAGGCCGCGATAGGCGTCAGCGCTGCA
ATTTGTTCAGTTACGATCAACATATTTTTGATGACAACAAATCTTCATCTTGT
ATGTCAATTCGAAATATTACACAACAAATTGAAATGGACTTTTTCAAACGACT
TGGATATAACGGATTCCAATATCTCACGTTACGCTTTTGAAAATTTAAAAAAG
TGCATTAAAATTCACATGATGCTCATCGACTACACTGCACAAATAGAAAACG
CTTGTTCGTATATGATACTTGTGCAAATGTTATCGTCTGGCTTACTGATGTGC
ACATCTGGTTTTCAAATATTTTTCTTTTCTGGAACGATACTAAAGTTCGTATTT
GGAACCTCTTTATTCCTGACAAGTATCGGAGAATTTTTTCTCGTTTCTTGGTCA
TGCAACGAGATCATATTGGCGAGTGCAAAAGTTGGCGATAGCGCCTACAATA
CAACTTGGTACGGTCTTTCAGCAACTGGTTATGGAAAAGCATATCGAGATGG
ATTACAATTATTAATGATGCGAAGTCATCGTCCGTGTTACCTGACAGCTGGAA
AATTTTGCATCGTATCGTTGGAATCCTTCAATGCGGTGATAACAACGGCAACG
TCATATTTCGCGCTGTTGCGTAAATTCGTGGACGATGCTGAACTACGTAATTA
G

>CcOR37
ATGACAGTGTCACAAGTCACTGAAATGTCCCACCATTCAAATATAGGTGTGT
CCCTTGCGGTAATAAAATACTCTATGAAATTAATTGGATTATGGAAAGCTGAT
AATACACTTGATATATTTCTCATGAGTTCAATTTTCATTTATACAATCAGCAT
GGGGATTGCCGGTATTGTATTCACAGTAACTGATCTTTTTTACGTTTTCGATG
ACATTTATGCAGCCGTAAATATTATTTGTCCAACGATTGCTATTTTTAATAAT
ACTGTGAAATTAATAATCTTTGCGATTAATCGACGGCAAGTTTTGAATATTAT
AGAACGTTTGGAAAATTCGATTAAAAACGAAACGTACGAGGAATACGATATA
AGAGCCGTCAAAGATTGTGAACGTCAATGTATTATTTTGGTAATAACTTTTGT
AATTCTGACTCAGGGGGCAGCTTCTAATTATAAATTTTTTCCGGCAACGGAAG
TAGAAATTTACCAGTCACTATCCATTTTGGTGGATTCGCTTATGGCGAATCAC
CTTATTTCGAAATTGGATTTTGTATCGAGTTCCGCACTGTGTTTAGTTACGCTT
AATAATTTTCTTACGACAACAAATCTTCATCTTGCGTGTCAATTCAAAATTTT
ACACAACAAACTAAAATCAACTTGTATCGCCGATTTGGATATGAATTTCAAT
GCCTCACGTGATGCTTACAGCAAGTTAAAAAAAATGTATAAAAACTCACAAAA
TGCTCATCGACTACACTGCTCAAGTAGAAGATGTTTATACATACATAATACTA
ATGCAAATGTTTGCATCCAGTGTAGTAATATGCGCAGCTGGTTTTCAATTATT
TTTCTTTTCGGGGTCGATGCTGCGCTTCGTTTTGAGTATTTTTTTCTTTCTGACT
TCCGTCGGCGAATTTTTTCTCTTTTCCTGGTCATGTAACGAAATCATCGTGGC
GAGTGCGACAGTTGGTGATGGCGGCTACAATACGACCTGGTACAGTCTTTCC
GCAACTGGTTATGGAAAAGCATATCGAGACGGATTGCAATTATTAATTATGC
GAAGTCATCGTCCGTGTTATCTCACAGCTGGAAAATTTTGTATAATATCACTA
GAATCCTTTTCCACAGTGATGACAACAGCAACGTCATACTTCACGCTGTTACG
AAATTTCGTTGACGAGGATGAATTTCGTAATAAATAA

>CcOR38
TCGTACCAGGAAATTTTCTATCGGTTATGCAGAGATCTAGTATTTTTAACGTA
CCTACAGAGTGAACGTAAAAAATCTGTTTTACCGACGACAGATTACATCTTTT
TCCTTTCTTCACTTATCTTTTCAGTCTGGTGTCACGCTTCATTTTCACTATACAT
TATTGGACCTTTCCTTAATCACACAGAGGAACGACCATTTCCAGGGCATACA
AAATATCCCTTAGACGAAGACAAATATTATTGGTTCATCGTATCTCATCAGTC
GATATGTACTTTTACCGTTTCCGCCATCACTGCTGCCGTCGATGCACTTTTTGT
CATGCTCGTTCATCACATATGCGCCAAATTTGCAGTACTTGGATATAACCTCG
AGAAAATCGGCGAGGATCTATCAGAAAGTCCATCGGCGCAAGACGAGCTGG
AAACGTATAAAAAAATTGTGAACTGCGTAAAAAAACACAAGGAAGCTATAG
GTTTCGCTGATTCAATGGAGTCAGCATTTTCTATACCAACGTTCATCAATCTT
ACGATGAACATGATAGTAATGAGTACTTCCGGTTTTATGCTAGCCAACGTCCA
TTCGATACCAAATCTAGTAAAATACATAGGATTGTCAGTCACCCTAACCGTTC
ATCTTTTCTTCATGAGTTGGCCAGCCCAGGAGATGATGGACCACAGTGCAGA
AATTTGCGAATTCGCGTATTTCGCATCATGGTATCGCACATCGCAAAGATCTA
AAAATTTATTGAAGTTCCTGATGATGAGGAGTCAAGTTCCTTGCAAGTTGACA
ACAGGAAAGCTCTACGTCATGTCCTTGGAAAATTTTTGCGCGGTCCTGAAGA
CGTCCATGTCCTATTTTACAGTATTAAAGTCCTTGAGAGACTAG

>CcOR39
ATGTCTTCCTTAACAGAAAGCGTATATACTAAATACGAGGATAAGCCAGAAG
AATCAAGATCTGATTTTGAGAATCTAGACATTACCTTGAGATGGAATCGCTTT
AGCCTACGTCAGCTTGGGATTTGGCCAGATCCGGAATTGTTTGGAGGCAGAT
CATTGAGGAAATACAAGTTTCTGTTTCCTGCATTTTTTATGCTTGCATTCACCA
TCATCCCGCAGACGACGTTGCTTGATATGGTCTGGGGAGATCTGGACCTCGTC
GTCGATAATTTATGTGTGGCTAACCTACCAATGATTTCCGCTCTGATGAAGCT
TTTAATACTCCGATTCAACAGTAAAGCTCTACGACCAGTTCTCGCCTGCATTA
CTGAGGATTGGAAAAATGTAAATAACCCAGGGAAACGCGAAGTCTTACTGAA
GTATGGAAAAATTGCAAAAATTATATCCCTCGTCTTTACCTTTCTGATATACA
TGACGGTCATCAGCTACGTTGTTCTTCAAATATATTTGAATTATCAGACAAAG
ATAATAAACGGAACGGATACATCGCGACGTTTCCTTTTTCAATCTCACTTTCC
TTATAACACGCAAGTCAGTCCTAGTTACGAATTGACTTGGCTTATCAACCTAT
TGACCCTTGACAACATATTGGCTATCCTCGTACTTAACGTATGTAGCCAATAC
AAGAATCTGCAAATCGACTTGAAAAATGTATTAGAAAACATGGATACGTACA
AGGAAACGTTTATAAAAAAACTTTCAACCATTGTACGCAAGCACAACCACTT
GAACAGGTTGGCAATTTTATTGCGAAACAAAGTTTTATATAAATTGGTCGAA
ATAGTTTTATCTTTTTTCTGTTTCAAAGGTGTGCTGATACAATTGATAAATCAT
TTAACCGTATGCTTCTCGTGCAAATGTTTGCTTCTACGGTTCAGCTTTGTTTTC
AGGGCTTTCAAAGTTTTTCTCTCTTGTTTCCTCTTCTCCTGTTATATCGGTGAT
AAACTTCATACCGAGGTTAGACTGTTTATTTCAAATTTTCAATATTATCCTAA
CTTAGAAGAGAGAGAAGCAAGGGAGTCGTTGCTTGTGATGTTAAGAGCAAAC
AGTCCTATGCGAATAACAGCCTGGAAATTCTCTTCGTTTTCGATGGTCACATT
CGCCGCGGTAGGCAATAACTGGGGAACTCAAAGAAAACGCGTTAAACATTG
A

>CcOR40
ATGGAACGTAAAACTTTGTATAAGATCGCGCCAAATGAATACAATAAGGATT
ACCAGTCAAACGTGGAATACGCTATTAAATTGAATCGCTGGTTATTAAAACC
GATTGGAATTTGGCCCATTAGGTCATCATCGAAATTGTCCAATATCATATCAT
TTATTTACTCAGGTTTATGTTGCTTTATAATGATTTTTATGGTGGTTCCGACGT
GCATCGAGATATTCACTGGTGAACAGAGTCTTAAATCAAAGCTAGAAATTTT
TGGACCTACTAATTTTTGTGCCATGGCTATTATAAAGTATTGTGTCTTTGCGTT
TCGGAGCAGAAAACTTAAAGTTTGTATAGATTTCATTATTGCCGATTGGCGTG
AAGTTAGTAACAAAGAGGAACGCGTCATTATGTTCAGAAATGTATATATAGC
TAGATCTTTGACAGGAGTTTGTGTTACTTTCATGTTTGCTGGTGGTATTTTTTA
TAATATACTCTTACCTCTGCTCACATCCAATTCTCTCTCAAGTGGTAACGTGA
CCATTAGAATGTTGCCATACCGTGGTAATTACATATTGTTTGATGCATATGTG
GGACCGTTTTACCACATGGTATACCTCATGCATTGTTTGTGCGCCATAATCAT
GTACTCCGTCACAACTGTTGTGTGCAGTTTAGCTACAAAGTTTGTAATGCACA
CGTGTGGACAGTGCCAAATTGTGATATCACTCTTGAAAAATCTTATTGACGGC
AATCAAAATTGTTCAAGAACCTTGGATGACAGATTGGCTGTTGTTGTTGTACG
ACATCTCCGTGTACTCAGATTTGCATCACACGTGGAAAACATTTTAAATGAAC

TATGCCTTCTGGAATTTGTGGGTTGTACTATAAATTTGTGTCTCCTTGGATATT
ATTTTATTACAGAATTTGAGAATGCCAATACACTCGGACTCATTACTTTCTTT
CTCTTACTCGTCTCCTTCACTTTTAATATTTTTATATTTTGTTATGTCGGCGAA
CTGTTAACGGATCACTGTCATCAAATAGGAGAAGCATCTTACATGATCGATT
GGTATCGTCTATCAGGGAGGCAAGGACGATTTTTGATAATGATAATTGCAAT
TGCCAATCGACCCATAAATCTCACAGCTGGTAAAATGATACAATGGTCTCTG
ATGTGCTTCACCAACGTAATAAAAGTGGCAGTAGCATACCTGAATATTATCC
GTAAAATAACTTCGTGA

>CcOR41
ATGCAAGCATTAGGACTTTGGCCACGAACATCCAACAAGCTTCGTTTGACGA
TCATTGCAATTTATTTGCTTGCATTAACCGAATGGTTTGCTGGGCAGTATGCA
TATTACATCACTAAAAAACTTGATTGGATGCCTTTTATTGTAAAAATAAAACA
TACAGTAGTAGCCCATATGTTCATTGGGAAATATCTATTGGCCATTATTGCTG
GAAGTCAGTTTGGAGAATGTTTATCCGCCCTGGAAAGAGATTGGCATAATGT
CAAATTGACGTATGAAATAGAGATCATGCGAATTTATGCCACCAAAGGCAGA
TATTACGCTCTAATCTGTTTCATCTTGTCTTTCACAACCGTCACGTTACAATTA
ATAGTCCCATTGACTCCTTTATCGGAATCTCGCGCCTTTCCTTATCCATCGCAC
TATTGGTTCGATTACCAAAGCAGCCCATATTACGAAATTGTTTATACTTTAAA
AATATTTTCAAGTACCATTCTTGCCGCATGTAACGTTACTGTTGATGGATTTA
TTATTATTCTAATTATGCATGTCATTGGACAATTGGAGCTCGTTAGTGAGCAT
ATAAGAAACATGGAGGTTATGAAATCAACGATAGAAACTTGTGTACAGCGAC
ATAGAGTTGTTATATGGATGGCAGAGACGTTGGAGAATTCATTTAATTTGGTT
TTTTTATGCCATTTCATGGTATCAGCCATTGTGATGTGTTTGGCAGGATATTCA
TTGATGAAGAGCGTGGAAAGAAACGAGAGCCTTGAATCTGTTATTTCTATGT
TCTTTATTGGATCAATTGCTTTTCATCTCTTTCTTTATTGTTACGTCAGCGATA
AATTAACGGAAAAGAGTCTGCAAGTTGGTTACTCGGCTTACTGTTGCAACTG
GTATCAGCATGAATGTAAACAATCGAAACCTTTACTTATTTTAATGATTCGTT
CATCGAGATCGTTAAAATTGTCAGCAGGGAAATTCATCGTATTGTCACTTGAA
AATTTCACAGACGTACTAAAAACGACTGGTGCATATTTATCAGTATTGAGAA
AATTTTAA

>CcOR42
ATGAAATTTGCAACGGACTTCAGTCGTTTGGCTTTAGAATTTTTTGGCTTTTG
GCCTGACACGGAGTATAGTCAAGCTAGTAAACATTTAAAACTGAAGTGTATC
ATCACAATATTATTCCTTCTGGGATTTTTGTATTTACCCCAAAGTGCTTTGCTC
TTAATGAACATTGGGGATTTGAATTTGGTTATACAGATATTAGCTACTGGAAA
AATTATAATCTCCTGTGCCATAATAAAGATAATAATTCTACATATTAAAAGAG
AAGATTTGAGAAGACTTCTCAAAAGTATGAATGAAGATTTGGTGGATTACGC
CAACGAGGAAATAAGGGAGATCATGTTGAAGAACGCAATATTCGTTAGAAA
AATGTGTATCTCATACTCCGTTACAATCTGCATAACAATGTTACTTTATATAG
TAGTCAAAATATTTTTTTACTTCGCGAAAATCTCGGATAACGACACATATCCT
GGTCTTGATTTATTCGTTCCTAGTTATTTGCCAAATTATCTTTTCAATAGTCCA
AGTTTCGAATTAATATATGCTGGGCAAATTGTGGCCATTATGATTGCAATCAG

CGCGTATTCCAGAAGTGACAGTTTACTTGTGATGCTGATAATGCATGTTTCTA
CTCAATTTACAATATTGCGATCATCCCTGAAGAATCTACCTGTCAAAGTTGGT
ACTACGAGTTCCACATTTATGTGCGAACTTGGAGTAATCGTCCGACGTCACGA
GCACCTGAACAGGTTCGCTTCGAAAATTGATAACATCTTCAATATGATTTTAC
TATTACAAATCGTTCTGTGCTCCGTATTGCTGTGCTTTCAAAGTTTTCACTTTC
TTGTGAAATTGAAAGATAAGAACGCGGAATTGTCTATCATCGAATTGGTATTT
CTTTGTTTCTACATTTGTCCCGTATTAATTCAACTGTATATTTATTGTTATTTCG
GTGAAATGCTTCAATCCCAGAATGAGGAGACGAGAAGAACCGCGTACGAGT
GTCTTTGGTATATCTTGGATCTAAGAGATGCAAGAACAATGATTATAATAATG
GCGCGTACAGTTCGTCCAGTGCAGTTAACTGCTGGGAAATTCAGCCCCGTCGT
AATGTCGACTTTCACGGCTATCTTAAAAACATCAATGAGTTATCTCTCAGTGC
TATTAGCTACAATCAATGATCCGAAGGACAATTAG

>CcOR43
ATGAAGGACACTGCGATAGTTGGTAAACCAGTAGAAATTGGTCTTCGATTAA
TTGACTCCTGGCCTGGTGCTTCTAATAGAATTGGGCGGCTGACAGTTTGGTCT
CTCATGATGGCAGCACTGATTTTTCAGTACTGGGATGCTGTGAGCGTTTTCAA
CGATTTGGACAATCTCATGGACAACTTCAGTGTCACCATCACCGAGACGTTGT
TCTTTGTCAAGCTGATCATCGTCTACAAGAATCGTCGGTATGTCAATGAAGTT
TTGAAGCATATGTCAGAAGACTGGAACAGCGTTAAGAGCGTCGAGGAATGG
ACCGTGATGACGGAACATGCTAAACTGTCCAGGATATTTTATATCTGGGCTTT
AGGACTTTACGTAGGGACTGTTGTCCTTTTTCTGCCTGTTGTAATAAATCACT
ACAATGCTGCTGATATAAACGATAGGAAATTCGTTTTGCCCTCTGAATATCCT
TTTCAAAGTAAAGTGTCGCCCGTCTATGAAGTCCTTTGTTGCGTTCAGTTCCTT
CAGGCTGTACTAACTGCCGCCGGGAACGCTCTTACCGAAAGCCTCCTTGTAA
CGTTGGTGTTACACGCAGGATCTAGATTGGTCCTATTGCGCAAGGATATATCG
CGATTTTCTGATATTAGCCGGACTGTTGACGACAGGCGGAAGATATTATTGGC
TGGCAATGTCCTTGTAGCCAATCATCGAAGAGTCATCGAATTTTCAGACAAG
ATTGAGGACCTCTTCTCTTATATATCATTGGTTCAGGTATTATCGAGTACTTTG
ATAATCTGTACCATTGGTTTCATGTTTATTACGTCCATTTCATCCACCAAGAAT
ATCCTTACGCTCATGAAATTTGGATTGTTCATTCTTGGTGAACTTTGGGAGAC
ATTGGCGTACTGTTTGGCTGGAGAGTATTTGAGTAATCAGTCTCAATCGATCA
GTCAGGCTGTTTATGAATGTCCCTGGTACAAGATGCAGCCTAGGGACAGTAA
GATGCTAATGATGATAATGATAAGAGCTCAGAAACCTCTTAGAATTACTGCT
GGAAAATTCATCTTTTTATCACTGGACAACTTCACTGACATTTTAAAAACATC
TCTCTCCTACATTTCGGTCCTGCGTGCTATATACTAG

>CcOR44
ATGGATCACGAAATAGAGCAGTGTAAAATCAAGACtACCGTTCAGAATAAAA
ACGTAGACTCGGACGTGGAATATGCGATTAAATTAAATCGCTGGTTATTGAA
GCCATTCGGGATTTGGCCGTTAAATTCATCGTCGACTCAATTCGAAAGACTAG
TATCAGTAATCTCACAATTCGTTTGCTGCTTCCTACTGCTCTTCATGATGATTC
CGTCCATTATAGAAATGTTTGTCAGTGAAAAAAATTTTAAAGCAAGACTCGA
TATTTTGGCACCCACGAGTTTTTCTGTTACGGTTGCTATAAAATATATTGTATT

TATGACTCGTGGTAGACAGCTTAAAACGTGTATCGATACTATCATTAATGATT
GGGGTAACGTCCGTGAAAAGCAAGAACGTGAAATGATGATACGTAACGCAA
GAATTGCTCGACTTTTCGCAATAATTTGTGTCTCGTTTATGTATTGCGGTGGA
ATTTTTTACAGTATCTTTTTACCTCTGATAACGGCAAAGTCATTGACAAGTGG
CAATAACTTGACTATTAGAATTCTACCATATCGATCCAATTATATACTGTTTG
ATCCATACGTAAGACCAGTTTTTGACATTGTGTATATTGCGCACTGCTTCTGT
AGTTTCGTTATGTGTTCCATCACAACGGGCATCTGTAGCCTAGCTGCAAAGTT
TGTGATGCATGCCTGTGGACAGTGCGAAATCGTTATGTCACTTTTGGAAAATC
TTATTGACGACGACAAACAATGTTCTGATATTGTGGAAAGCAAGTTGGCTAC
TGTTATTTTACAACATTTGCACGTGATCAGATTTGCGACGCGCGTGGAGGATC
TTTTAAACGAAGTGTGTCTTGTGGAATTTTTGGGCTGCACTATGAACATGTGT
CTTATTGGATATACCATTCTGACGGGATTGGAATCAACTGACACTACTAAATT
CATTACTTACTCAGTACTATTTCTATCCTTCACGTTTAATATTTTCATATTGTG
CTATATCGGTCAAATTTTGACAAATCATGTACTTTGCCATCAAATTGGAGAAG
CTTCATACATGATCGATTGGTATCGCATCCGAGGAATCCAAGCACGATTTTTG
ATATTACTAATTGGTATAGCAAATCGACCAATGAGACTCACAAGTGGTAGAA
TGATCCAATTGTCTTTCCCGTGTTTCTGCAACGTGATAAAAGTAGCAATGACA
TACCTAAATGTTCTCAGAAAAGTCACAGCGTAA

>CcOR45
ATGACTGGGAAAGTCATTTTACCAAAATTTGATAATTACATTTTACTTAATAA
ATTTGGACTTCGTCTACTTGGCCTTTGGCCCATACCCGAGGACTCGGCTAAGT
GGAAGTTCACCCTGAAGAAGCTTCACGTCTCGCTCGTGTACGTCTTGCTGCTC
AGTCTACTGGTGCCCCAGCTCTTGGATCTTTATATTCTGTGGGGCGATGTCGA
TGCAAATGTGGAAAATCTTTGCACCAGCCTCACGACTTTCACTGTTCTTGCCA
AATTGACGAATATCGTTGCCTCGCGAGGAGTATTTCAGAAATCCTTGGCCACC
ATGAAAGAAAATTGGGACACCATCATGCGGCACGACAATTGTCCCGAGGAG
AGGGAAATTCTTTTAAAGATGTCAAAAGTGGGATTCGTCTTTACAAGGAATT
ACTGTATAATTATGTACATCACTGCAGGAATGTATTTTTTGAGGCCTCTCGTT
GTTGGCTCAGGGACAAAAAGTCCTCACGTTAAGGAGTATCCATTTTGTGCCTG
GTACTACTACGATCAATTTTCAAATCTCACCTATGGTATATTTTATTTCTCTCA
GGTAATAATAGGTTTCTTCTGTGGCACCGGAAATTTTACATTGGACAGTCTGT
GCTTGGTAATGGTCTACCACGCGTGTGCACAACTTCGTATACTTCAAAAGCAA
ATAAGTCAGCTGACCAATGATGGCTCGGATGATATCTTGATCGTCGAAAGGG
TTCGTCAGTTGGTGAAGCTTCATAAGAAAAATATAGAGAATGCCCGCAACCT
GGAGGCCGTCTTCAGTGGTGCGAGTGCCCAGCAGCTGCTCGTAAGCTGCGTC
ATCATCTGCGTGATTGGACTCAAGCTGATCGTGTCTCTGAATGATGCCTTCCA
GTTAATAATGTACGTGGCATACATGCAGCTGGTGATCTTCCAAATTTACTTAT
ACTGCAGTCCCGGCGATGAACTGATCAATCAGAGTATGGAAATTGGCAGAGC
CGCCTATATGGCGTCTTGGACAAATTTTCCAAAATATGCCACGCACTCTTTAC
TCATGATGACTATCAGAGCACAGAGGCCACTCAGAATCACTGCTGGAAGATT
TTACGTCATGTCCATACCAAATTTTACTTCTATTCTCAAAACTTCTGCCTCGTA
CTTATCCGTTCTACGTGTATTGTACGATTGA

>CcOR46
ATGGATATACTTCCATTGTGCTTTACAATTTTTCAATTTTGTGGTTTTTGGCGA
CCAATTAATTATACGTCATGGAGCAAATTTTTGTACAATTGCTATACAGTATT
CGTGACATTTTGTATATATAGCTTCACCCTGTCTCAAGCGATAAATAGCATAA
CAGTGCTTGATGACATTAACAATTTGACTAACAATACATTTATGCTTGTGACA
ATGATTAGTGTTTGCTGTAAAATTGCGAATATATTAGTAAAGCGCGAGGCAG
TAATTAACGCGATTAATATGTTCTTGGAAAATTTTGAGGATCTTGTAGAGCAG
AAGATAAAAAAGAAATATGATTTAATAGCTTGGTCAGTTACACGGAATTATT
CAATACTGGTAACGATCACGGCAATATGTTACATTGTGACACCTTTGGTAATT
AATGATATTGAAAAACGAGTTTTACCAAACCGAGCATGGTATCCTTATGATTT
ATCATCATTAACATTATTTTGGTTGAGTTACGTGCATCAAAGTTTTGCCCTGG
CGTCAGCAGCCTTCATCAATGTTGCTAATGACACATTTATTCCGGGGTTGATG
ATTCAAAACTGTGCTCAGCTGGAAATACTGGAGCATCGTGTGGAGAAGTTGA
TAATATTAACAAGAAGTAAAAAAACTGAAAATGAGACAAAGAAGGAGTATA
AAGAAATTTTGGAAAAATGCATTCGTTATCATAATCGAATAATAAAATTTGC
CAAAATTATTGAAAAAATGTTTGAACCGGTTATTTTCGTGCAGTTTTTTGTAA
CGGTTTTGACATTGTGCTCAAGTATGTATCAACTTTCAAGGCATAGCACATTG
ATCGTTGAAATTATTAATTTAATATTTTATTTGATCTCCATGTTATTCCAATTG
TTCTTTTACTGCCATCGCGGTAACGAATTATTTCTAAAGAGTATAGAGGTTGG
CGACAATTTATACAAAATGGACTGGTTATCATTGAGACCTTCTCGAAGAAAG
GATTTATTAATAATCATGATCAGAACCAATAAACCAATACAATTAACAATTG
GTGGCATAATAACTTTGTCCTTCAGTGCTTATATAAGTATATTAAAAGCATCA
TATTCAGTATTCAACTTGTTGCAACAAACATCTGAAAAATGA

>CcOR47
ATGGATATTTCTGAAGCCGAGAAAGTTTTAACATGGAACAAGTGGTTGCTCG
ATTTTCTTGGAATCTGGCCTACCAACAAAAATAATACTAAATTCTATTTTTTCT
TTACTTTCATAACAGTTCAATGCTGCTTACAGTACGCTGATTTAATAGAATAT
ATCGATGACTTCGAATATGTTCTCGTAAATTTAACTGAAAGTATTGTGTTTAC
CATGATACTTCTCAAGATTGTTCAATATAAAATTAATACGGGACGACTGATTG
AATTAATACGTGACATTAGGACGGATTATTCAAAAGAACTGTACCGCACAGA
AGAAGAAATGGCTATTTTCATACGGTATAATTCATTTTCAAAAATAATTCTAT
ATTGCTTATCAATATCAGGAGTGATTACCAGTAGCCTCTACTATATTCGGCCA
TTGTCCAGCTACCTGTTGGCTCACAATGATCCCACAGGAAACTCATCCAGTTC
CTTCATCTTGCCATATCACGTCCGGTTATTCTTCGATTTGACGGAAGTGCCAA
CGTATTACGCCATATATGCCTGTGAAATATGGGTCATCCCGATGGTTGTGTGC
GGATTTATGGGCACCGATTGTTTACTGATAACTCTGGTGCTGCATGTTTGTGG
ACAATTTTCAATTTTAACAATCCAAGTGGAAAATTTAACGCACGATCCACGA
GAACTGCATGATGAGCTCAAACGTATCGTTATAAAACATCGTCGCTTGATAA
GACTCTTTGCTAATTTGCATTCTGCTTATTCTACATTTCTACTGCAAGAACTCG
TTGGTATCACTTTACTTACGTGCTTAGGCAGCTACAACGTTATCGTGACTCCT
ATTCTGAAAGAAACTACATATTTTTTGATATTTGTATTTTACATTATTAGCATA
ATTTCTCAGCTTTATGGTTTCTGCTATATAGGAGAATGTCTAATAATTGAGAG
TGAATGTTTGTGTCACGCATTTTACAATTCTGAATGGTACAATGCGCCACCAC

ATCATGGAAAATTTATTCTGATGTGCATATTACGCTCCCAAAGGCCTCTCGCA
TTGACTGCAGGAAACATATTTACATTCTCTTTGGAAAGTTTCACGCATGTTTT
AAAAACATCAATGGCTTACCTATCAGTTATACGAAACTGCTTATAA

>CcOR48
ATGGACATATCAGGTGCCACGATAGTTTTGAAGTGGAACGAATGGCTGCTTG
ACTTTTTGGGACTGTGGCCTTTAAATTTGAATAATGCCAAATTCTCTTTCTTTT
TTATTTACATAATAATACAATGCTTCTTGCAGTATGCCGCCTTAGTAGACAAT
ATTTTTGACTTGAGCTACGTTGTCGCGAATTTAACAGAAACCGTCGTCTTCTG
CATGATAGTTTTAAAACTTGTTATATACAGAATCAATATGAAAAGATTGCATG
AATTAATACGAATTATTAAAGAAGATTATTCACATGAGCTTTATAAAACAGC
AAAGGAACGAATGATTTTTATGAAATATAATTCATTGTCAAGAATGATTGTTC
AGTGCTTTTCGATACTGTGTGTGTGTGCTGCGGTGCTCTTCTATATTCAGCCAT
TGATTTGTTATTTATTAGCATACAGAGACTCCACAGGAAATTCATCGAGTGCC
TTCGTATTGCCATATCATATTCGATTATTTTTTAATCTAACTGAAGCGAGAAC
GTATTATATTATATATGCTTGTGAGATATTGATTATTCCGATGTCTGCATGTG
GATACGTTGGGCCTAGCTGTCTACTGATAACATTGGTGCTACATATTTGTGGA
CAATTATCCATTTTGGCTACTCAAGTGGAATGTATGACGTACGATCCAAAAAC
TATACAGCAGCAGCTCAAACAAATCGTCATAAAGCATAGTCACTTAATAAGT
CTCTGTGCTACTTTGAATTCCACGTATTCTATATTTTTATTACAAGAAGTCATT
GGTATTACTGTTCTTCTGTGTCTAGGCAGTTACAATATTACTGAAACTGGCGA
ATTTTTGACATTTTCTTGTTATGTTTTTACTGTGTTTGTTCAACTTCTTGGTTTC
TGTTATATGGGAGAATGCCTTGTAAATGAGAGCATAAACTTGTGCGATGCAT
TTTACAATTATGAATGGTATAATGCATCGGCAGTGCATAGAAAATTGCTTCTT
ATGTGTCTTATACGATCTCAGAGGCCTCTCGTATTAACAGCTGGAAAATTCTT
CACGTTCTCCTTGGAAAATTTCACAAGCGTAACTCGAAATTTATTAATTAATT
TTTCTCCGGTATTATTTTTTAATCAATTAGCTCTTCAAACTTTCTTATTCGCAG
GTTATGAAAACATCAATGGTTTATTTGTCGGTACTTCAAAAATTCATATAAAA
CAGTCATGA

>CcOR49
ATGGATATTACATCTTTGTCCAATCACAAGCTACCGAAAATATCGTCCTATTA
CAATGTCAAGCTTACGTTACTCGTTAATAAATATTTTGGATATTGGCCGTCGC
TTAATGTAAATCAACGATTTTATGCTATCTATACGTTCTTTAGTTTTTCACTGA
CAATCGGCGCTATTTTACTTGCCGAACTTGTTTACGTGATCGTTTATATACGT
GACTTTAACAAATTGACATCGATTTCGCTTCTTTTATTCACAAACGTCGTTCA
CGTTTCTAAGATACAGAATTTAATTCGTAAGCGTAAGCAGATACGCCATCTGT
TGGAATATGTAGAAACTGGTCAATTCTCGAACGAAACGGATAAATATGAAAG
GATTGTAACGTATTTCTCTTGGTTTGGACTATTTCATTATGCGATTAATGTCTT
GTTTGGAATGGGCGCTATTACTGGATGGTCAATTATATCCTTGTTTAGCGTTG
ACAATGATAAAATTGACAGAAGATTGCCTAGCGATGGATGGTTTCCATATGA
TCCATTAAAATCACCTGCCTTTGAACTCACATGTGTCTATGAAATAATTATAC
TTCACGTATGTGCCTTTCATAACTTAGCTCTGGATCATTTGACGGTCGGTTTCA
TGTCTGTCGTATGTTGTCAATTTGTAATTTTAAATTGTAATCTACAAAGTATTG

GAGACGATGTACTCAATGCTGAAAATATAAAAAATGATTTATCGAATTCATT
GAAAAATAAGGAAAATGACGACTTGATTTATAAACGACTGGTTAAGTGCGTC
GAAGATCATATGACAATTATTCAATTTGCCAAAGACGTCGAGAACATATTTA
ATTTACCAGTTTTTCTTCAATTTCTGGAAAGTTGCCTGGTCATTTGCGTAACCG
CCTATCAGATTAGTCAAGATTCCAATAAAGGTACTAGTGAAATTTTGGGAAA
TTTTAGTTATTTAATGTGCATGGTGTACCAATTGTTTGTCTACTGTTGGTACGG
AAACTCCATATTATTGCTGAGTGAAAGCGTCGTTCAGGCTGCTTTTTCTGGCA
ACTGGTGGAAATCGAGCGAACGATATAAACGGGCTTTGCGTCTGATGATGGT
GAGAGCAGGCAGACCACTTACTATCACTACTGGAAATCTATTGACGCTATCA
TTAAGTACGTTCATGGGAATTTTGAAAGCATCTTACTCGTTCTTCACGGTATT
GCAAAGTACCAACAAGAAAATCTAG

>CcOR50
ATGGAAATACAAAATATGGGTTTCTATGAGTCGAGAAAAAAATCTGTGAAAC
CCATAAAAACGAATGTTCACTTGCAAGTAAGTCTCAGCATAATTCGCTACAT
GGGAACATGGCCGCCACAGGGTCGCTACCGAAATTTATACTTCGTCTATAGC
TTCCTTGTCTTGAACATAATAATTAGGCAGTTTCTTTTTGCTCAGTTTGGTACG
ATAATTATGATTTGGGGTGATATTGAAAAAATGGTAAGCTGCGGTGCCCTCCT
GATGACGAATATTGCCCACGCCTTTAAGGTTGTCGTCATGCTTCGACATCAGA
AGTCAATTCAGATTCTTCTGGATAGTCTACAGGACGAGTTATTTACCAGAAAT
AACGAGAGATTTAAATTTATCGCATATGAATATACTTGGAAAGGCATTTTTCA
TCATGTCGCATATCAAAGCTTTGGAACCGTAGCTGTTCTCTGTTGGGTTTTTA
CTTCAATATCAGATCTAATTATGAAACATATGAAACGTTTACCAATTCCACTC
TGGTATCCTTTCAACGTAACAAATACGCCAGCATTCGAATTGATTACTCTTCA
ACAAACTGTTGGAGTCATAAGCGGATGTTTCCACAACGTGGCGATGGATACT
TTGATCACGGGTCTGATAACAGTTGCTTGTTGTCAATTTGAATTGTTGAAGAC
CAATATCATTACGATTGGGAGTAACGTTAATCATGAATTTTGTAAAATTTCCC
ATAATGAAATTAACGTCACGAACGTTGAAGATAACGAGAATAAAGTCTGGCA
AGAGATGAGAAAGTGCATCGCGCATAACAATAAGATTATTGAGTTTTCGAAA
GAAATTCAATCGCTTTTCGGTACAGCGATTTTCTTGCAATTTCTTGCAAACTG
TGTTATCATTTGCTTAACGACTTTTAACATGTCTCAGACAACTGTATTTGTTCC
ATCGGAAATACTTGGAACAGTCGCATACACTTGTTGTATGGTATATCAAATTT
TTATCTATTGCTGGCATGGAAATGAATTGTGGCTTCAAAGCGAAAGTGTCCTG
CGGATGTCTTTCACGTCGAATTGGTGGGAACATAATCGACGTTATAAAAAAG
CTTTACAAATGGTAATGCTAAGATCAAGTCGACCAGTCATCCTGACTGCTGG
AAAATTACTAAAGTTATCTTTAGAAACGTTTGTTGCATTTTACGAGTGTCCTA
TTCACTTTTTACCGTTTTAA

>CcOR51
ATGGAAATTCTGCCTTTGTGCTTCAAGTTCTGGACTATATGTGGTCTTTGGAG
ACCAGTTCGTTATTATTCTTCGGTGTCAAAATTCTTTTACGACTCCTACAGTTG
CGTGACGGTTCTTGTGTTGTTTAGTTTCACGCTTCTAGAATTCGTAGACATCG
CCTTGAATTTTGGAAACGTCGATGACTTTACGAGTAACGCTTTTCTCTTTTTAA
CGATGATTGGCGTGTGTTACAAAGTGGCGCATACCTTTATGAAACGTTTGCAA

ATTATTTGCATAAGAGACATGCTGTACGACGAGATCTGCAGACCTGAAAATA
TGGAGGAACAGAGAATACTTGATAATTGTGCTAATGCCTGTAGATCCAATAC
GGTAAATCTTGGGATAATAATTATGAGTTCCATAGTGTTTTTTCTTGTGGGAC
CACTATTGCATGATGTCAATGACCGCGTACTGCCGTTTAGAGCATGGCTGCCA
TACAGTCTGTCATCGTTGCCAGTGTTTTACCTGACTTATCTCCATCAAAGCTTG
GCGATAATATCTGCTGGTTTGGTGTCCGTGGCAAGCGATACATTTATATCCGG
ATTAATGATTCAGGCTTGTGGCCAGTTGGATATTCTAAAATGTCGTTTAGCCA
ATTTGCATGAATCGCACAATTTCGAAATGGCTATAAATGATACATTCCATCGG
GAGGAGGCTATTCTCAAAATGTGTATTAGGCATCACAATCATATATTCAAATT
CGCAGCTCTCGTGGAAAAGACTTTTAATCATGTAGTACTTTTTCAATGTTTTA
TAAGCGCCTTCGTTGTTTGCACCACTACACTTACATTATTGAATCATCAGTTA
CTGAGTATTGAATTCATTACTATAATTGTGTATTTTCTTTCAATAGTCTTTCAA
TTGCTTGTATATTGTTGGTATGGGAATGAAGTAATTTTGAAGAGCATCGAAGT
CACAGATGCTGTGTTTGAAATGGATTGGATTCGTCTACCTGCATCAATGAAAA
AGGATTTAATAATGATTATGATGAGGGCTGAGAGAGCAATCAAGTTCACCAG
TGGATACGTTCTCACTCTGTCTTTGGATTCGTATGTGGCTATTTTAAAGTTATC
ATACACTACATTTAATGTATTGCAACGATCATCTTAG

>CcOR52
ATGGAAATTCTGCCGTTGTGCTTTAAGTTGTGGACGATAGGCGGTTTTTGGAG
ACCACTCGATTACTATTCTCCAATGTCAAAATTCTTGTACAACTGTTACAGTT
TTGTGAGTTTTCTTGGATTGTACAGTCTGATGTTATTCGAATTCCTGGATATTG
TCTTTAATTTCGGAAACTTAGACGATTTTACAAGTATCGCCTTCATGTTTATG
ACGGTATTAAGTGATTGTTTAAAAGCAGCAAATACCCTAGGAAAACGTTCAC
GGATAATTGAGATACAAGACATGCTAACCGATGAGACATGCATAGCTCACGA
TACGGAGGAACAAATGATACTAGAAAGGTCTGCCAAAGTGTGCAGATCGAAT
ACACAATATCTTGGGATATTAATAATGAGTTTTGTTGTATTCCTATTTGTGGG
ATCATTGCTGTTTAACGTGAATGAGCGCATACTGCCGGTCAAAGTTTGGCTGC
CTTACAGTCTCTCATCTTTGACATCATTTTCTCTGTCTTATTTATACCAAATTT
GGGTTTCGATGCTGGCCGCATTTATGGCCACGGCAAATGACACGTATATAAC
TGGATTAATGATCCAGATTTGCGCCCAGCTGGATATCTTGAAATATCGTTTGC
TCAGTTTACCAAGATCGGGAGATTCCGAAAGGAATGCAAATTCTTCCGTCAG
CCAAAAAGAAATTATTCTCAAAATGTGTATTCGGCATCACAACCATATAATC
AAAGTTGCAGCTACTGTGGAGGAAACCTTCAGTTCCATTATATTTTTTCAATG
TTTCGTAAGCACCACCGTCATTTGCGTTGCTACGTTTATACTTTTGAAATCTGA
CTTCATGAGTGTTGAATTCATCATCACCCTCATGTATTTTCCTCCATTACTATT
CGAATTGTTAATATACTGCTGGTATGGGAATGAAGTGACGTTAAAGAGTACC
GAAATCAGGGATGCTGTGTTCGAAATGGATTGGATTCCTCTGCCTGTGTCAAT
GAAAAAGGATTTAATACTTGTTATACTCAGGGCTGGCAGAACAATCAAATTC
ACCAGTGCATACGTTCTCACTCTGTCTTTGGATTCGTTTATGGCTGTCAGTATT
CACACCCATTATTTTTGATGA

>CcOR53

ATGGACTACGAGTACAATCCTGAAGAAGCTTTTGCTATAAATTTCGGATTTCT
TCGAGTGGATGGTGCATGGCCAACTGGGATAACGAATCCAATACTAAGGTTC
ATTATCAATGAAGTATCTTTTATCGTTGTCAAAAGTGTAATCTGTCAGAATAA
TTATACAGTTTCTGTAATGAATTGTTGTAGATTTTTGTACAGAATTTATGGATT
TTGTGTTCTATCCATCTTCACAGTGTTATACTTTTCAACTGAAATTATTTTGAT
GACAGTGATATGGGGTGATGTCCAGGCGATGGTTGATTGCTCCTTTCTGTTTC
TCACGCATCTCAGCCATTCGGTGAAGCTAACGAATTTCTTACTCCGTCAAAAA
CGTCTAGACAGGATGATTTTTCAATTGAGGGAGCCAATGTTTGTACCGAAAA
GACCTGAACATATCGCTATCATGAAGGCAGCAGTGCTCAATGCTGAACGGGA
GACCAAGATGTTCTTAACAATTACATTCGGCACGATCATCTGTTGGTCCATAT
TACCGTTGCTTAGTCAGAAAGAGGTTAAAGAACTTCCTTTGATGGGATGGTTC
CCGTTTGATACAACTCAAAGCCCGGCGTATGGTTTTACGTATACTTATCAAGT
GATATCGGTCCTTTTAAATGCTTGCGTAAACGCAATGATGGACACGATGACTT
CTGGTCTGCTCATTCTCATGGGTGCTCAATTGGACATGTTGAAAGAAAATTTA
CAATCTCTGACCGATGACGCTTCAACTATGTCACAAGCTGAAATGAAGTATA
TTTATAACGTGAGAGATGTATTTCAAATTGGTATCATGGCTCAGTTTTTAGCC
AGTAGTATTATTATATGTCTGACGTGTTTTAAGTATCACCAGCATCTGTAG

>CcOR54
ATGGATATTTCCGGGGCCACGAAAGTTTTACAGTGGAATAATCGGTTACTCC
GTTTCCTAGGACTTTGGCCCTTTGACCTAAATAACGTAAAATTCATGTTTTTCT
TTACCTACGTGACAGTTCAGTGTTTCTTGCAGTATGCCGGCTTACTGGAATAT
ATTTCCGACCTGAATTATGTTGTCGCAAATTTAACAGAAACCATCATCATCAA
TATGCTTATTTTCAAAATTTCTATATACAGAATTAATACAAGACAACTGCGTG
AATTAATCCAGAATATTGAGAAAGATTTTTCGACAGAACTTTATAATACAGC
AGACGAAATGACAATTTTTTTGAAATATAACTCATTGTCAAAGACTATTGTTC
AGTGTATATCCATAGTGTGCGGAGCTACTGCTGTACTCTTTTACATTCGAAAT
TCTTCAAGTGCCTTCTTGATGCCATATCATATCCGATTGTTCTTTAACCTGACT
GAAGCAAGAACATATTACATAGTATATGCTTGTGAAATATTTATGATTCCGCT
GGTCACATGCGGATATGCTGGTCCTAGTTGTCTATTGATAACATTGGTGCTAC
ATATCTGTGGACAAGTTTCAATTTTGACATGTCAAGTGGAAAACTTGATAAA
AAATCCAGAAACTATACATCACCAGCTGAAACAAATTGTTTTAAAGCATCGT
CGTTTGATAAATCTCTGCATGAATTTAAATTCCACGTATGCTACATTCTTACT
GCAGGAACTCATTGGTATTACTTTGCTTCTGTGCTTAGGCGGTTATAACATTA
TTGCGACTCCTATCTTAGATGAAACTGGACATTTTTTGGCATTCTTATTATACA
CTGTTACAGTGATCTTTCAACTTCTTGGTTTTTGTTACATCGGCGAATGTCTTA
GGAATGAGAGCAAAAGCTTGTGCGATGCATTTTACAATTAA

>CcOR56
ATGATAATACCCGGGGCCACAAAAGTTTTGCAGTGGAATCATTGGTTGCTCG
GTTTCATAGGACTTTGGCCCTTTGATTTAAATAACAGCAAATTCATATTTTTCT
TTGCTTACGCAATGGTTCACACTTTCTTACAGTATGGTGATTTGATAGAACAC
ATTTCTGACTTGAACCATGTTGTCGCGAATTTAACAGAAACCATCATCATTAA
TATGCTTATTTTCAAAATGTCTATATACAGAATTAATACAAGACAACTGCGTG

AATTAATCCAGAATATTGAGAAAGATTTTTCGACAGAACTTTATAATACAGC
AGACGAAATGACCATTTTTTTGAAATATAACTCATTGTCAAGAACTATTGTTC
AGTGCTTTTCAATAATGTGTTTGATTAGTCCGATCCTCTTCTACATTCATCCAC
TGCTAAGTCATTTATTAGCATACAACGATTCCATGGGAAATTCATCAATTGCC
TTCGTATTCCCAATTCATTTTCGATTATTCTTCAACCTGACTGAGGAGAGAAC
ATATTACATTATATATGCTTGTGAAATATTGCTTGTTCCTACGTGTACATGCG
GATATAATGGTCCTATTTGTCTAATGATAACACTGGTGCTACATACTTGTGGA
CAGATTTCTATTTTGGCAAGTCAAGTGAAAAGTATGATACACGATCCAAAAG
CTGTACACCAACAGCTTAAACAAATCGTCATAAAGCATCGTCGTGTGATAAG
TCTCGTTGCAAATTTGCAGTCCGCGTATTCTGCAATCCTACTGCCAGAAGTTT
CTGGTATGACTTTTGTAATTTGCTTAGGCAGTTACAACGTTATCACGACTTCT
GCCGTGACAGACAGCTCAAAATTCTTGAAATTCTTATTTTATATTCTTACCCT
AACTTTTCAACTTTTTAGCCTGTGCTACATAGGAGAATGTCTTATAACTGAGA
GCACAAACTTATACAACGCGTTTTGCAATTACGAATGGTACAACGTATCACC
AGATCATGCAAAACTTCTCGTAATGTGCCTGTTACGATCGCAAAGGCCTCTCA
CACTAACTACTGGAAAATTCTTTACGTTTTCTTTAGAAAGTTTTAGAATTAAA
ACTTGTGGCGCAAAATTGATTGTAAAAGTAAATTCAAACCTTATTAGCACTTA
A

>CcOR58
ATGGGCAATACAGATATGAAATTCGTTACGAATATCCATCGTTTGGTTTTAGC
ACCAATTGGTCTCTGGCCTGACGTTGAAAACATTAAAAAAGATAAAGTGTTA
AAATTCAAGATTATCTCATCAATATTATTAGTTCTCGTCTTTGTCTATTTACCG
CAAACTATATTACTCGTAACGAGAATCGATGATTTTGATTTAGTCATACAGAT
ATTAGCCACTGGAGAAATTGTATGCTCGTGTGCTTTAATAAAAATAATAATTC
TCTATTTTAATAGACAAGCTTTGGGAAAACTCCTCAAGTACATGGAAGAAGA
TTGGATGAATTCTGTTAATGCCAACGAGGAAACAAAAGGGATCTTGCTGAAA
AAAGGGAAATTTGTCAGAAGATTTTGCATTTCAAATATACTTTTCTCATACTT
GACATTTTTATTGCGTGTACTCGTCAAAATTCAATTTCACATCACAAAAACTT
CAGAATGCGAAACGCATCGTGGTCTCGAATTTTTTTTCCCCAGCTACTTTCCA
AAATACATGCTGAAAAGTCCAAATTTTGAATTAATATTTATTGGGCAAATTGT
TGCTACTGTGATTGTGATCAATATTTATATAGGATGTGACAATTTTCTTGTGTT
ACTGATAATGCACGTTTCTGCCCGATCGACGATACTGCGAGCGTTGCTAAGG
GATTTACCACTGAAAGTCCATCCAAAGGATTCCATTAAATTTATGAAGGAAC
TTGGCATCATTGTTCGACGACATGAGCACCTGAACAGGTTTGCTTCTATGATC
GATGATAATTTCAATATAATATTACTCGCGCAAATGACTTTACTCGCGGCATT
ACTGTGCTTTCTATGTTTTCAATTAATTATGAAAGTGCAAGTTCAGAATGCAG
AAGTTTCTACCATGGAAATAGTATTTTTAATTTTCTTTATTCTTGCCACATTAA
TTCCGTTATACATTTATTGTTATTTCGGTGAAATACTTCAACACGAGAACGAA
AAGATAAGAAATGCAGCATACGAGTGCCTGTGGTATATTCGCGAACCGATAG
ATGCAAGAAGTGTGATATTAATAATGGATCGTACAGTTCGTCCAATGAAATT
AACTGCTGGAAAATTCAGGTCTTTCACCATGTCAACTTTTACTGATATTCTAA
AAACATCAATGAGTTATCTCTCAGTGATATTAGCTACAATGCAAGAGCCGGT
GGACAATTAA

>CcOR59
ATGGACCACGAAATAGAGCAATTTAAAATCAAGACTACCGTTCAGAATAAAA
ACTTCGACTTCGACATGGAATACGCGATTAAATTAAATCGCTGGTTATTGAAG
CCATTCGGGATTTGGCCGTTAAATTCGTCATCCACAAAATTTGACAGGGGAAT
ATCAGTAATTTCTCCATCCATCTGTTGCCTCCTGCTGTTCTTCGTGATGATTCC
GTCCTTCGTCGCAATGTTCGTAAGTGAAAAAGATTTTAAAGGAAAACTTGAG
ATTGTGGGACCCACGAGTTTCATTATTATGGTTGTTTTAAAATATTTTTTCCTT
ATCACTCGTGGCGACACTCTTAAAATGTGTATCGATACTCTCATTAGTGACTG
GCGTAATGTTCATGCAAAGCATACCTTTTTACCTGTGTTAACAGCCAAATCGT
TAGCCCGTGGCAATAAGTTGACTATTAGACTAGTACTATTTTTTTCGTGTAAT
TTTATACTTTTTGATCCATACGTAAGACCAATTTTTGAAATTGTCTACATTGCG
CATTGCTTCTGTAGTGTCATTGTGTATTCTATCACAACCGAAATCTCCAGCCT
AGCTGCAAAGTTTGTGATACATGCCTGTGGACAGTGCGAAATCGTAATACCA
CTGCAAAATATTATTGATAGCGATAAGCGCTGTTTGAATGTTGTGGAAAAAA
ATTTACTGTTGTTATTTTGCGATATTTACATGAGATCAGGTAAAATAATTATC
TTACAAGTAAAAGGACTTTTGGAGAGGCCACCATTCTATGGATTTTTAACAAT
TCTGCGAATGAATATTCAATATATATTTGTCACAAATGTGGAAGATCTTTTAA
ATCAAATATGTTTTGTGGAATTTTTGGTCGGTACTACTGTGTATATGGGTAAA
CGAAAGTGCACTGAAGCTCATTACTCTCTTTTTCTCTGGACACTATTAAATTC
AATATTACATTTATATATGAAAAAGTGCCATCAGATTGAAGAACCATCATAC
ATGATCCATTGA

>CcOR60
ATGGATATATCAGGAGCTGCTGATATATTAACTTGGAACAAATGGGTGTTTTC
GTTTTTGGGTATCTGGCCATTAGATGATAGCAATAAACTATTTTTATTTCATTA
CATATATTTTATGATTCACACATGTTTGGAGTATGTCGAACTGTTTAAGTATTT
AAACAATTTAGAATATGTAGTCGCAAATCTGACTGAGAATTTAGTGATCACT
CTTACCTTTTTAAAGATTAGTGCTTACAGAATTAATAAAAATGAATTACGTAA
AATGTCAAATGACATAAATGACGACTTTAAAGAAGAGTCTTATAAGGATAGT
GAAGAGAAAGCATTATTTTTGGAATACAATGCAATAGCTCAAAAATTTATGA
AAGTTGGGATCCCGATAACATTAATTGCGGCAGTACTGTATTATTTGAGACCT
TTAACTGGACATGTAATCGTGGATACATCACGAAATACATCACACAGTTATA
TTATGCCGTATCGCATTACATTATTCTTTGAAATTACAAATTTTCAAATGTAC
GTCTTCATGTACGCTTTTGAAATTCTAAACGTCCCAATCATAGCCTTTGGATTT
ATTGGAACTGATTGTTTATTAATTACGTTGGTATTACACTTGTGCGGTCAATT
AGCTGTTTTGTCACACCAAGTGCGAAATTTAACAATGAATCCGCAAAATTTTC
AAAATGAATTCCGAGTGATAGTTATGAAGCATCTTCGACTCATAAGAATTTC
GAAAAGTTTAGATACGGCATTCAATCAGATTTTAATTCAACAACTCGTCGGA
ATGAGTTTGCTTTTATGTTTGGCTGGATACAATCTTATAGCGATGTCTGATAG
TGGACAAAATGTGCATCTAATAACGTTTATCTTCTATGCCTGTTCGGTATCCA
TGTTACTCTTTGCTTATTGTTTAATCGGGGAATATCTCATTAACGAAATCCAA
GATGCTTTTTACAATTCTACATGGTACAACAGGCCACCTTATCAGCTGGAACT

ATTTTTGATTTGTATGGTATTTGTTCAACGTCCGCTGGTACTAACTGCTGGAA
AAATCTATTCATTGTCTTTAAGTACTTTTGCGCGTGTAAGTAAGTAA

>Orco
ATGATGAAATTTAAGCAACAGGGCCTGGTCGCTGACCTGATGCCCAATATTC
GTCACATGCAATTCAGCGGACATTTTATGTTCAATTATTACAATGATACCGGT
GGTTCCACGAAGCTCTTTCACACGATTTATTGTTCCATTCACTTGTTCCTGATT
CTATTGCAATTTGGATTATGCTGTGTCAATCTGACCCTGGAAAGGGCTGACGT
GGATGACCTTACCGCTAACACGATCACTGTCCTCTTCTTCGCTCACAGTATTA
TTAAATTAGCTTACTTTGCCGTACGGAGCAAGCTCTTCTATCGCACCCTCGGC
ATCTGGAACAATCCAAACAGTCATCCTCTCTTTGCTGAGAGCAATGCACGTTA
TCATGCGATCGCATTGACGAAGATGAGACGGTTGCTTGCCGCTGTGGGTGCA
GCGACGATTCTGACAGTTTGCGCCTGGACCGGGATCACCTTCGTTGGGGATTC
CGTGAAGAAGGTTACCGATCCTGTGACTAATGAAACGATGACGGTGGAGATA
CCAAGACTGATGCTTCGCTCTTGGTATCCATACGATGCAAGTCACGGGATGG
CCCACGTCCTAACGCTCATCTATCAATTCTACTTTCTCCTAATCACTACGATG
GACGCCAATTCTCTTGATGTCCTCTTTTGTTCCTGGCTGCTATTTGCCTGCGAG
CAATTGCAACATCTTAAGCAGATTATGAAACCCCTTATGGAACTGAGTGCTA
CCCTTGATACAGTGGTGCCACATACTAACGAGCTCTTTAAGGCTGGTAGTACC
GATCATCTGAGGGACACGCAAGGTACTCAACCGATGGCACCACCTCCTAATG
AAAACATGCTTGATATGGATCTCAGAGGAATTTACAGTAACAGACAGGATTT
TACTGCCACCTTTAGGACAGCTGCTGGGATGAATTTCAATGGTGGGGTTGGTC
CTAATGGGCTGACAAAGAAGCAGGAGATGTTAGTTAGAAGTGCCATTAAATA
TTGGGTTGAGAGACACAAGCACATCGTCAGATTAGTCACCGCTATTGGTGAC
GCTTACGGAGTTGCCCTTCTTTTCCATATGTTGATAACAACCGTATCTTTGACT
CTGTTGGCATATCAAGCGACAAAGGTGAACACCGTTGACGTGTACGCTGCTA
CTGTGATCGGTTATGTCCTCTACACACTGGGACAAGTCTTTCTCTTCTGCATAT
TTGGAAATCGTCTAATCGAGGAGAGCTCATCCGTCATGGAAGCTGCATATTC
CTGTCACTGGTACGACGGTTCCGAGGAAGCCAAAACTTTTGTACAAATTGTCT
GTCAGCAGTGTCAAAAGGCCATGTCAATTTCCGGCGCGAAATTCTTCACCGTT
TCTCTTGATCTCTTCGCTTCCGTGCTGGGAGCTGTGGTTACTTATTTTATGGTG
CTGGTACAGCTCAAGTAA

>CcP450_1
ATGTTTACAGACATACTGCTGGCCATAGTCGTTCTTTTATTCGTTGCCTGCATC
TTATCCGGCGGGAAATCTAAAAACTACCCACCAGGACCTTATTCTTGGCCACT
GGTTGGGAATATATATCAGCTCCGAAAATTAGCATGCACTCATGGTGGTCAG
CATCTTGCACTTTTGGAGCTTTCACGAAGGTATTCCAGTGACGTGATAGCACT
TCGTTTAGGAAGGCAGAATATGATAGCTGTCTCGGGTTATGAAGGAATTCAA
ACGATTCTGAATGGTCAGGAGTACGAAGGACGTCCGTGGAACGAATTCATAA
AGATCAGGAATATGGGAATGCGTAAAGGAATCACAATGTCGGATGGACCCG
AATGGAAAGAAATACGAGCCTGGGTGGTAAGGTCCTTGAGAAGCGTTGGGTT
TGGACGTAGAGAAATGTCTGTCATGATCAAGGACGAGGTCGTCCACGTTGTG
GAGAATCTTAAGGTTGGCGGAGTACGCTCAATGAAACCAGTTATCGCCCCGG

CAGTAATAAATGTTCTGTGGACGATGGCAACGGGAAAGCGACTGTGCGAGGG
ACCCAGGTTACAGTACTTTATGGAATTAATGGAACGTCGTGCCAAAGCCTTC
GATATGTCTGGTGGAATCCTCTCGACTTTCCCTTGGATGCGATATTTAGCTCC
GGAATTCTCAGGGTACAACGTACTTATGGTTTTTAATGAGGAGATCAAGAAC
TTTCTTATGGATATTATAAACGAGCATAAGGAAAATTATACTCCAGGAAACG
AAGATGACTTAATTGATATGTTCCTAGGTGAAATGTACAGTGGAAAAGGAGC
AAAGGCTGGTTTTACAGAGGATCAGCTGGTGATGATATTGATGGATCTCTTCA
TTGCGGGTATCACGACCACAACAGTGACTTTGGATTTCCTCTTTCTTAACATG
GTGATGCACGTGGACAAGCAGAAACGACTGCAGCAGGAAGTAGACGCCATA
GTCGGGCGTGAGAGGCTTCCGGAATTGAGCGACAGACCTAATCTACCATACA
CCGAAGCCGTATTGACCGAGTCCCAGAGAATTTGCATGGTTACGCCAGTCAT
CGGACCGAGACGAGTTCTCTCCGACACGGAACTACTTGGTTACACAATACCT
GAAAACACCACCGTCCTGATTAACGTCTACAGCGTTCATATGGATCCACTGC
ACTACCCTGAACCAGAGAAATTCAAACCGGAACGGTTTATAAAAGATGGCGT
CTACGTGCCAGATGAGAATCTAATGCTCTTCGGTCGTGGAAGACGACGTTGT
CCTGGAGAAGCCTTAGCGAGGTCTGCACTCTTTTTACTGTTTGTTGGTATTCT
ACAGAAATTCGTACTGCTACCAGTTCCTGGTAAACAAACTCCTACTTTGGACA
TCGTGCCAGGATTGACCATCTCGCCTAAGCCTTATGAAGTATTAGTTGTTCCT
CGTGATAATTGCACCAATGGCACATCGAAATTAATAGACTGA

>CcP450_2
ATGCTCGTTGAACGCGCAGCGCAGTGGGCCTGGCAGGCCATGGGTGACACGA
GGACCGACGTCTTCTACACCTTCCTGGTCTTCCTCGTGGCTTTTGCCCTTGTGC
GGGGCGCACAGTGGCTGCGATACATTCGCTCGTTGCCACCTGGCCCATGGGG
ACTTCCGGTCTTGGGCTACCTTCCCTTTCTCAAGAGCGACGTACATCTGCAGT
TTGGGAAACTCGCCAAGAAGTACGGATCCATATTCAGCGCGAGACTCGGGAG
CCAGTTCGTAGTCGTTATCAGCGACTACAAGACAATTCGCGATACGTTTCGTC
GGGAGGAATTCACCGGGAGGCCGCACACAGAGTTCAGCAACATCCTTGGAG
GATACGGTATCATTAATACCGAGGGCGCATTGTGGAAGCATCAGCGCCGGTT
CATTCAAGAGAAGCTTCGCAGTTTCGGTTTGACGTATGTGGGTGCCGGCAAG
AAGATTATGCAGTCCAAGATAATGCGGGAGGTGGACACGTTCTTGCAAGGAC
TCTCTCTGCGTCGCGGTGCACCCACGGATATGTCGGGTTCCTTAGGCATGTCC
ATCAGCAACGTTATTTGCTCGATAATAATGGGCGTGCGCTTTCACCACGGCGA
GACCAGATTCAAGAGATTCATGGACCTGATCGAAGAGGGCTTCAAACTCTTC
GGCAGTATCGTCGCCGTCAATTTCATTCCTGTGATGCGATACTTGCCCTTCCA
CAAGACGATTTGCAACAAGATTTCCGCGAATCACAAAGAGATGGCCAACTTC
TTTCAACAGGCGGTGAACCAGCATCGCGCTACCTTTGACGCCGGTAACGTCA
GAGACCTGGTGGACACGTATCTTCTAGAGATCGAGAAAGCGAAGGTTGAAGG
TCGCGAGGCTACGCTCTTCGAAGGAATGAATCATGACCGTCAGATACAACAA
ATTCTTGGGGACCTCTTCTCCGCTGGAATGGAAACCGTGAAGACGACTCTTGA
GTGGGCAGTCATCCTCATGCTTCATCATCCAGAGGCTGCACGTGCTGTTCAGG
AGGAGCTTGATCAAGTCGTGGGTCGATCAAGGATTCCAGGATTGGAGGATCA
ACCCTTCTTACCTGTTACCGAAGCCACTATTTTGGAAGTGTTACGCCGCTCAA
GCATCGTGCCGATGGGGACGACGCACGCCACTACTCGCGATGTCACGTTGAA

CGGTTTCACGATCCCAGCTGGAGCACAGGTGGTACCGCTCCTGCACGCCGTT
CACATGGACCCCGAGCTCTGGACAGAGCCCGAGGCGTTCCGACCAAGCAGGT
TCCTGAGTTCAGAGGGCCAAGTCACGAAGCCCGAATACTTCATGCCCTTTGG
CGTCGGAAGACGCATGTGCTTGGGTGACGTCCTGGCACGCATGGAGCTCTTC
CTCTTCTTCAGCTCCCTTCTTCACACCTTCGACTTGCGGCTACCCGAGGGCGC
CTCCTTGCCAAGTCTACGTGGTAACGTCGGGGTCACTGTCACCCCGGATCCAT
TCAAGGTCTGTCTCGTTCAAAGGACTTTAGAGGGTGACCTCGCGGACGTGGA
GTGCACCGACGGCAATGGACCACTTCGCAACGTTGGTAGTCATTGA

>CcP450_3
ATGTTTCTGCCAGCTATATTATTCCTTATTATACTGATGTTATTCTTCTACTTG
GGATCAAGAAAACCAAAAGGTCTTCCGCCAGGACCACGTTGGTGGCCTATTC
TTGGCTCAACCCTAGAAGTAATTCGGCTTCGAAAACAGACCGGCTATTTATAC
AAGACTCTCTCGGTTCTCTCCAAAAAATATGGTCCAGTTTTCGCTGTAAAAAT
TGGCATTGATACTATTGTATTCTTGAACGATTATAATAGCATAAGATCAATGC
TGACCAATGAGGATTGTGACGGGAGGCCTACTGGACACTTTTACAAAATGAG
AACACAAGGAAAGTCACAGGGTAAAATTTGTGTTTTGGTCACTGATGGTCGC
TTATGGGTTGAACAACGTAGATTTGTTCTTCGGCATTTAAGAGAATTTGGCTT
TGGCCGCACCACTATGGCCTCATTAATAGAAGAAGAAGCTCAGCACTTGGTC
AATCACTTTAAAAGACTTTTGTTAAGTGCCAATAATCCGGTGTCGGAAAGCA
GACAACAAATCAAATCTTGTTCAAATAACAATGGGCAGATTTATCAACTGAT
CTCTGATACGAATATTAATAATACAATGATGACACACGAAAAAGACAACATA
GAATCGGACACTAAGAAGAGAAAAAATGAAGGAATGACTATTGAGGATGCT
TATGTGAAAGCTGACGATTATGATGAAGTGAGAAAAATCTCACAGTCCTCCG
GAATGATAATCAGCATGCATGATGCCTTTGGCATTCCTGTACTCAATACATTA
TGGCGAATGATGACTGGGAAAAGATATAATCCGGATGATGCTGAGCTGAAAT
ATCTTCAAAAAATTGCAACTCAACTTTTAAAGGATGTGGATATGGTGGGCAG
TACATTCAGTTACTTTCCGATACTGAGATACTTGGCGCCTGAAATGTCTGGTT
ACAAATCATTTGTTGCAACGCACGAGCGAATCTGGAGATTTCTTAAGAAAGA
ATTAGAAAATCACAAGACTAACATCAATCCTGAGAAGCCTAAAGATCTGATG
GACGCCTATATAAATATTTTACAGACAGAATCTTACAGTAAGACGTTCTCAG
GTATGCTTATTTCAAAAATTAATATTCCAATAAAAGATGCCATTACTAGTATA
ATATTTATTTCAGAACAACAATTACTAGCTATCTGCTTGGATCTCTTCATGGC
TGGATCAGAAACAACTTCCAAAGCTCTCAGTTTTGGTTTTCTCTACATCGTAC
TCAATCCAGACGTCCAGAGAAAAGCTCAGGAGGAAATTGATTTGATAATTGG
TCATAAACAGTTTCCACGCCTGAGTGATAGAGTAAGAATGCCCTATAACGAC
GCGATTGTTCTGGAATCGGTACGGATGTTCATGGGACGTACCATGGGTATAC
CACACAGAGCCTTAAGGGATACATATATCACAGGGTATCGGATACCAAAGGT
AAATGAAAATACCATGATAGTGCCTAATTTCAATGGGGTTCTTATGAATAACT
TCTGGAAAGATCCAGAAGTATTTCGTCCAGAGAGATTTATCAACTCTGATGG
AAAGATCTCAATACCCGATCAATATCTGCCATTTAGCTTTGGTAAACACCGTT
GCATGGGAGAAGGTCTGGCCAAGAGCAATATATTTGTAATAACAACGACCCT
CTTGCAAACCTTTAGGTTTTCTGTTGTGCCAGGTGACAAGAAACCGAGCACA

GAAATAGTGGACGGCGTCACAGCAGGACCAAAACCGTTTCGTGTCCTTGTTA
CACCACGTATATAA

>CcP450_5
AAATTATCGGTGCATTCCGGTAGACAAGGTGTGGATCCCTTCGAGACTTTACA
TCACTGCCTGGGTAATCTGATGAATCGCATTGTCTTTGGCAAGGACTACCGGG
AGGACGACGTCGTGTGGAAGTGGTTGCGTCATCTTCAGGAAGAAGGAGTCAA
GCATATCGGCGTGGCAGGGCCGCTTAACTTTCTGCCGATTCTCAGGCTGCTTC
CACGATTTAGTAAAGTGTTTGACTATCTGATCGATGGGAAGCGGAAGACTCA
TCAAGTGTACAAGCAGATCATCGCCGAGCACAGGAACCAGGTTGACAATGGA
GAAGAGACGATGGACAGTGTCTTGGCTGCGTTCGACGATGAGATGAGACGTC
GGATGAACAGTGATGATCCTGGTCACTTTACCGAACCTCAGTTTCTTCATTTT
CTGGCGGATCTCTTTGGCGCCAGTACAGACACCACCTTGGCGACGATTCGCT
GGTTCCTCCTCTTCATGGTTGCATACCCCGAGGAGCAGCTCAGGATTCAGCAG
GAGTTGGATGCTGTTCTTCGAGGGAAGAATGTGTCTCTCGATGACAGAAGTG
CCTTGCCGTACCTCGAAGCTGCCATAGCGGAAACTCAACGTATCAGGAGCGT
CGTCCCCGTCGGCATACCTCACGGAACCACTGAGGACACATTGATCGCCGGT
TACGAAGTTCCAAAGGGAAGCATGGTGGTTCCCTTACAGTGGGCTATTCACA
TGAATCCGAAATACTGGCAGGACCCAACGAAATTCAAACCGGAAAGATTCAT
CGCCCAGGACGGAAGTCTCGCCAAGCCAGAAGCTTTTCTTCCCTTTCAGACTG
GAAAGCGGATGTGCCTCGGCGAAGATCTAGCCAAGATGATGCTCTTCCTTTTT
ACTGGTCGAATCCTCCAGAGGCTTAACCTAACCTCACCTGAAGATCAACAGC
TGGACTTGACTGGAGAGTGTGGCATAACTATGACACCGAAACCATATCGATT
AATATTCACATCTCGTTCCTAA

>CcP450_6
ATGGATGTTGGAGCATTTTGGATCCTCTTCACCTGTGCCACCATCCTGGCCAT
GATCGCCAGGAATCGTTTACGGTCTGGACCCACGGAAGAGGAGGGTGACACG
GTAGCCTGGGAGAAGGGCAAGCTCCGCAAGGAGAAGGATCAAGCGGATGAC
ATCAAGAAGGATAAGCGAGCTGGAAGAATGAAGAAGCTTCTGAATGGTCCT
AGGGCACTGCCCCTGTTTGGGAGTCTCCATCTCCTGGCGAGCCCAGGTGGAC
CCTTCGAAGCGTTCACCAGGTTGGCACGGAAGTACGGAGACATCTATGAGAT
TCGTCTAGGAGTAGCTACCTGCGTGGTGGTCAGCAGCTACAGTCTCGTCAAG
GAAGTTCTCATCACGAAGGGTAGTCACTTTGGGGGTCGTCCTGATTTTCTACG
TTTCCACCAGCTCTTCGGGGGTGATCGAAATAACTCTCTCGCGCTTTGCGATT
GGTCCGATCTGCAGATGAAGAGACGATCATTGGCCAGGAGTTTCTGTTCGCC
ACGAAGAGGAAGTCCTCAGCAGGAGGAACTGTCGAGGGTGGCTACCTTGGA
GGCATCGCAATTCATGTCGGCGCTGACGAACGAGGAATCTCAGGACGTTCTG
CGCGGCGAACGTTCCTTGAAGCCACTTCTTCTGGGAGCCGTGGCGAACATGTT
CACCAGGTACATGTGCTCGACGCGGTTCAGTTACGCCGATCAGGAATTCTGC
AAGATCGTCAGGACCTTCGATGAAATCTTCTGGGACATCAATCAAGGCTACG
CTGTGGACTTTTTACCCTGGCTCAGACCGTTCTATGCTGGTCATATGAAGAGG
TTGAATAATTGGGCTACGGGGATAAGAACCTTCATCCTGAAGAGGATCATCG
ATCGGCATCGCGCCAGTCTCGACTCCGTGAACGGGGTCCCTAGGGATTTCAC

TGATGCCTTGCTTCTTCATTTAGAGAGTCCCGGGTCCGGTTTATCCTGGGAAC
ACATTATATTTGAGCTGGAGGATTTTCTAGGAGGTCATTCTGCGATTGGGAAC
CTGGTGATGCTCATACTAGCCAATTCAGTGGTGCATTCTGAGGTCCAGAGTAA
GATCCAGGCAGAGTGTGATACCGTCCTAGCCCAACCTGGTCGATCACGAGCC
CTGGTTACTCTAGAAGACAGAGCCGACATGCCCTACACCGAGTCCGTCATCT
GGGAAACTCTGCGTATATCCAGCTCACCGATAGTTCCTCACGTAGCGACCCA
GGACACTGAAATCGGCGGCTACAACGTGGCAAAAGATACAGTCGTTTTCGTT
AACAATTACGAACTGAATCTAGGGGAGTCTTACTGGGGTCCTGGATCCCGTG
AGTTTCGTCCTGAACGTTTCCTTAAGTACTCCGTCCCAGAGGCCGCTGGAGAC
TCCGGCGAGAAGCCCAAGGTCAGGAGGCCTGAACACTTCGTCCCATTTTCAA
CTGGCAAAAGGACCTGTATTGGTCAGCGACTGGTCCAGGGTTTTACATTCGTC
CTGGTAACGGCTATCTTTTCACGCTTCGACATCTCTCAGGCCTCCGAGGATCT
GGCTTCTCGTCTGTTACCAGGATGCGTCGCGGTACCTCCGGATCCTTTCCACC
TGGTGCTCGTTCCTAGAGCAGGCAAATCCGATGACTGA

>CcP450_7
ATGTCTTGTGGCTTGGATACTGGTTATGGCGTCTCTATGGTATTACTCAAGTG
GGATCGAACTCCTGCCGAGGTCTGGATTTACCTCGGTATTCTTGTTTTCATTTG
TCTCCTGTACGCCTTCAAGGATTACATACGAATTTTGAGATTTATATTCAGCC
TTGACGGACCTGTGGCTCTACCAGTGTTGGGAAATGCTAATTATGTAATGCGC
AAAGATTTACTACATATAGTATCGCACAATGTTTATAACGATTATGGACCAAT
TTTTCGATTCTGGCTGACCTTCTTACCCTACGTCACTCTCTTGGAGCCGGCGG
ATATACAATGTGTACTTAGTAGTGCGAAACACACCGAAAAGATATTTTTTTAT
AGACTACTGGACAATTTCCTTGGGAAGGGATTGATCACCAGGGACGTTGTCA
CTTGGAAGGGACATAGGAAGATACTTCAACCGGCATTTCATCTTCATCTACTC
GAACGGTTCACCAAGTCGTTCTCAGAGTGCGCTACAAATTTAATGAACAAAA
TTATTAATGAAGGTGGACGGGATGTGAATATCACGAGTTTCGTTAATGACTCT
GTCTACGACATTTTGAATGAAACTGTATTGGGAGTGACTTTAGCCAGTGGAA
AGAACAAGCAGCTCGACAGCGATGATGACATGCCGTTTAGAAAGGGGCAAG
TGGTGGCACCTTATCGTATAACCAGGCCCTGGCTGCTCATCGAGTGGATTTAC
CAGTTAACGAGCATTGGTCGAGCCGAGCAGAAGCAGCGCGATGAGCTGTTCA
ATGCCTGCCACAAGATGATGGAAAGGACCCGGCAAAATATTATGAAAAACG
GTTCAACCATAACAGATAGTTCGCAGAATGCAGTCAAAAAAATATCCCTCTT
AGAGTTCATGGTCGAAATCAGTGGAAAGCATCCTGAATTTACGGACGAGGAT
ATTATTAACGAATGCTGCACCTTCATGCTGGCTGGACAGGATTCCGTTGGTAC
TGCAACGGCCATGACACTTTTCCTTTTAGCCAATCACCAGGACTGGCAGAGG
AATTGCTTCGAGGAGCTCGAGGAGATTTTTGGAAAGGACGAGAGGTCCCCGA
CTATGAAAGACTTTCGGGATATGAAATGCCTCGAGATGTGTATTAAGGAGAC
GCTCAGACTATTTCCTCCAGTTCCTTTTTTCGCGAGGACGTTAGGCGAAGATG
TTGTCCTCGGGAAACAAGTAATTCCAGCAGGTTGTGGAGTCTTCATCATGCCT
TACGCGACTCATCGACTAGCTCATCACTTTCCAGATCCTCATACGTTCAATCC
TCAAAGGTTTGATCCTGAAGTTTCTAATAAAATGCATCCTTATGCGTTTCTGC
CATTTAGCGCAGGGCCCAGAAATTGCATTGGAAACAAATTTGCAATGCTTGA
AATGAAGTCGATGATCAGTGCCATTATACGCAAGTGTCACTTGGGTCCAGTG

ACCGGAAAGGAGGATATACGACCCAAATTTCGACTGACTGTCAGAGCCGAGG
GTGGACTCTGGGTGAAGATCACACCACGTAGTGAATTTAGGTACCAACCCTA
CATGAACCTGTGA

>CcP450_8
TTGTTTCAACGACTTATGGCTATGCGATTTTTATGGGGTCATCAAGAGGGGAT
CAATAAAGCATGGCTTGGTAACCATCCTTACGTATTTCTCAGCAAAGCAACTA
CCGTAGAACCCATCTTAGGAAGTAATCGCCATATCGACAAGAGCTCTGATTA
TCGATTCCTTCAGCCGTGGTTAGGAACGGGTCTACTGACAGGTTCTGGAAGA
AAATGGCATAATCGAAGGAAGATACTTACACCCACCTTCCATTTCAAAATCC
TAGAGGATTTTGTTGAGATTTTCGCCGAGCAAACTGAACTGCTCGTTAACAAA
GTTTCGTGTGAAATTGGCCAAGATTCGTTTAATATTTTTCCCTATGTGACTCTC
TGCGCCCTGGATATAATTTGCGAAACCGCCATGGGTCGTCAAGTTTTTGCCCA
ATACGATAGTGACAGTGACTATGTGCGCGCTGTGTATGAAATTGGAAGTATC
ATTCAAACACGTCAGGCTACTCTCTGGTATCATCCGGATTGTTTGTTTCGATT
AACGTCGCTCTATAAGAAGCATCAAGAATGTATTCGTATACTTCATGGATTTT
CGAATAAGGTAATAGCTGAAAGACGTCGTGAAATCAAACGAACACGGAATG
AGACTGTTTCAGAGGATGAAGGACGCAAGAGACTGGCTTTTCTAGATTTACT
TATAGAGGCCTCACAAAATGGAACTTCTCTTTCTGATGAAGATATACGTGAA
GAGGTTGACACCTTTATGTTTGAGGGACATGATACAACTTCCTCCGCCGTTAG
CTGGACGCTTTATCTTCTTGGATGCCATCCAAATATTCAAAATGAAGTGGTAG
CCGAATTGGCTGATATTTTTGGTGAGAATGATAATGGACGCCGAGCGACTTT
GCGTGATCTGCAAAGTATGAAGTGCCTTGAAAGATGTATTAAAGAGACACTC
CGATTATATCCTAGTGTTCCTTTATTAGCGCGACGAATCTCAGAGGATGTTCG
CATAGGAAAGTTTCTTATTCCTGAAGGAACCACAGCCATGGTGATCCTGCCA
ATGTTACATCGCGACCCTGCAGTATTTCCTAATCCAGATAAATTTGACCCAAA
TCGTTTTTTACCAGAAAATTGCACTGGTCGACATCCATACGCTTATATACCTT
TCAGCGCAGGACCAAGAAATTGTATTGGACAAAAGTTTGCACTTTTAGAAGA
GAAAGTTATAATCTCAGGTGTATTGCGAAAATTTAAAATTGAGGCTGCAGAA
CGACGGGAGGAAATAAGCTTGACTGGCGAACTCGTTATCAGAGCGAAAAAT
GGTTTGCATGTGAGAGTTACACGACGTGAAAGATAA

>CcP450_9
ATGTTGCTCTACGTAGTACTCTCTTTATTGATTTTAATAATTATTCTTATAGTT
CCATATCACTTAATAGTACATAAAAGTGGAAACGGTAAACTCATTAACAAGA
TTCCAGGCCCTAAGGCATATCCAATTTTGGGAAATCTGTTAGAATTTGTAGTA
CCTTTTGAAGAACTCTGGCATGTTGTACGTGGCATGAACGAAAAATATTACCC
GATTTATAGAATTTGGAGATTCTCGCATTCAGTGATAAATATTCGTCATCCCG
ATGATATTGAGGCCATCATCGCCAACTCAAAGCACATAGAAAAAAGTTTGAT
ATATAGGTTCCTTCAGCCATGGTTCGGCACTGGTCTCCTTACAAGCACAGGAC
GTAAATGGCAACAGCGTCGAAGAATTCTTACACCAGCCTTTCACTTCAACATT
TTGACCGAATATTCTAAAACATTTATCGAGCATGCAGAAGGACTTGTAAATTC
GCTGAAAGCGGAAGGTGAAGAAACTGTGAAAGACATCGTCCCGCTTATCACC
ACTGTTACACTGAGTTCAATATGCGAAACCGCAATGGGTACGCCCGTACAAG

AGGATGAAGATTTTCAAACAAAATTTCAATTTGCTGTTCATAAAATGGGCGC
AATTGTTGTGTCAAGAGTTTTTCATCCGTGGCTTTTTAATGACTGGTTGTACA
AATTCACACCGACTTTCCGCCTGCAAACAAAAGTTCTTGATATACTGCATGGA
TTTTCATTAAAAATAATAAACATACGACAGCAATATCACGATGCAACCGGCG
GGAAGTATTTAAATCAGTTTTCTACTGACAAGAAGAATATCGACGGGAACGA
AGTGAAAGACAGTCAACGTAAAAGATTGTCAATGCTTGATCTTCTTATTGGA
GTATCCAGAAATACTGGTCAGATAGACCTCGCTGGTATCAGAGAAGAAGTGG
ACACCTTCATATTCGAAGGACATGACACCACGTCGATGGCTATATGTTTCATT
ATACAACTCCTTGCGGAGCATAAGGATATTCAGGAAAAGGCAAGGAAAGAG
ATCGATGAAATATTGAAACAAAATGATGGAGAAGTGACGATAACGGAAATC
CAACAGTTCGCCTATCTTGATCGTTGCATCAAGGAGTCACTGCGTCTCTATCC
CAGCGTGCCATTCATGTCCCGAAAACTGTCCGAAGATCTTCAATTAAAGAATT
ATCTGCTACCTGCTGGAACATCGGTGCATATACATACCTTTGATCTCCATAGG
GATCCCAATTTTTGGCCTGATCCTTTGGTGTTTGATCCTGATAGATTTCTGCCA
GAGAAGATACAGGGTCGACATCCGTTTTGCTATTTGCCGTTCAGCGGGGGAC
CTCGTAATTGTATTGGCCAGAAGTTTGCGATGCTGGAACTGAAGGCTATTATA
GGTCTGCTGCTGCATAATTTCTACTTAGAAGCAATCGATCGCCCTAGCGAGAC
ACGCTTTCTACCAGATCTGATATTGAGACCGGCTCATCCTATTCGCGTAAGGT
TTGTTACCAGACGAAAGTGA

>CcP450_10
ATGAGTAGTGCTGCTGCTCCTGAAGTGGTACCGGGATCGGTGGCTGCAGTGG
CTGCAGCTGGCCTTTCGGCATCCACAGTTTTCTTTGCTCTCCTCATCCCAGCCC
TGGTACTGTACTACGTCTACTTCAGGATCTCGCGTCGTCATCTCCTGGAACTC
GCTGAGAAGATTCCGGGCCCTCAAGGATACCCCGTGATTGGAAATGCTCTGG
ATTTCATTGGAAGTTCTGACACAATCTTCAAGAACATTTACCAAAAGTCCTTC
GAATTCGATCAAGTGATCAAGATATGGATTGGACCTAAGCTTGTCATCTTCCT
GATGGACCCACGTGACGTGGAGATCATCCTCTCGAGTCACGTTTACATCGAC
AAATCCGACGAGTACAGATTCTTCAAGCCTTGGCTTGGAAATGGTCTCCTCAT
CTCCACGGGACCAAAATGGCGTGCTCACCGTAAACTGATTGCGCCCACTTTC
CACTTGAACGTCCTGAAGAGCTTCATCGACCTCTTCAATGCCAACTCTCGTGC
CGTCGTTGAGAAGATGCGTCGGGAACAAGGCAAGGAGTTTGATTGCCACGAT
TACATGTCTGAACTGACTGTTGAGATCCTCTTGGAAACTGCAATGGGTGTCTC
TAAGACCACCCAGGAACGAAGTGGATTCGAATATGCCATGGCCGTCATGAAG
ATGTGCGACATTCTTCACTTGCGTCACACCAAAGTCTGGCTCCGACCTGACTG
GCTGTTCAACCTGACCAAATACGGTAAGGATCAAATTCAGCTTCTTGAAATC
ATCCACGGGCTCACCAAGAGAGTCATCAAGCGTAAGAAGGAAGAGTACAAG
ACTGGCAAGCGCAGTGTGATCAATCCTGCTGTGGCTACTATCCAATCAGAAA
TGAAGGCTGCTGCGAAGGGCAAGGAGTCCAAGAACTCCAGTACAACTGTAGT
CGAAGGACTGTCATTCGGTCAGTCTGCTGGTCTCAAGGATGACTTGGACGTC
GATGACAATGATGTCGGTGAGAAGAAGAGGCTTGCTTTCCTGGATCTCCTGA
TCGAAGCTGGTCAAAACGGAGTAGTCCTCACCGACGAGGAAGTTAAGGAAC
AGGTCGACACCATCATGTTCGAGGGACACGATACTACGGCTGCTGGATCCAG
CTTCTTCCTTTCAATGATGGGAACTCACCCTGACATCCAAGAGAGAGTCATCC

AGGAGTTGGATGAGATCTTCGGCGACAGTGACAGACCCGCCACTTTCCAGGA
CACCCTGGAGATGAAGTATCTTGAACGTTGCCTCATGGAGACTCTTCGCATGT
ATCCACCTGTACCAATAATTGCTCGTGAATTGAAGACAGATCTCAAGCTAGC
ATCCGGAGACTACACCGTACCCGCTGGAGCAACCGTAGTCATCACGACCTTC
AAGATGCACCGACAGCCCAATATTTATCCTAACCCGGATAATTTCGATCCGG
ACAACTTCCTTCCCGAGAAGACTGCAAACCGCCATTATTATGCATTTGTACCA
TTCTCTGCTGGACCAAGATCTTGCGTCGGCCGTAAATACGCCATGCTCAAGCT
GAAGATTGTCCTCTCGACCATCCTGAGGAACTTCCGTATCAGATCTGACATCA
AGGAGTCTGACTTCAGGCTACAGGCTGATATCATCCTGAAGAGAGCAGAGGG
CTTCAAGATCAGGCTGGAACCAAGGAAACCAGGTATCAAGGCCTGA

>CcP450_11
ATGAAAGACAATTTAAGTAATAAAATTCTGTCTATGCTTTATGCAGGTAAGA
AATGGTACACTCGAAGGAAGCTCATCACACCCAGTTTCCACAGTGGACTATT
GGAACAATATCTAAAAAATTCAATTCGCGAAGCTGACATACTTGTCACTAAA
CTTCGTGCAGAAGTTAACAAGAATGGATTCAATATTGTTCCATATGCGAAGCT
TGCAGCTCTTGACGTTATTTGCGTTAGTGCTATGGGTTATCATATCAATGCGC
AAAGTAATTGCACGAACGAGTACGTTCTCGCTGTCGATAAAGCCGCCAGAAT
CACTCAGGAAAGATTCGTTAATATCTGGATATCACTGGACGTGATATTCAAG
CGAACTACTGCAGGAAAGGAATTCAAAAAAGCATTGGAAATAATGGATAGG
TTCACGGAGAAAGTAATCACTGGCAGAAAAGCAGAATGGAGATCGAAGAGA
GATGGAAACTTTAATGAATTCCCAAAGAAACGAAAAGCTTTGTTAGACATGT
TACTGGAAATGTCCGACGACGGCAAAATCTTATCGGACGAGGATATCAGGGA
GGAAGTAAATACTTTTATGTTTGCCGGGCATGATACAGTAGCTACAAGTGTTT
CCTGGATTCTTTATGCACTTGGAAGATTTCCAAAATATCAGAACCTCATTCTT
GAGGAATACGACAGAATTATCGGATCGAATGATATCACTTTGGAAGATTTAA
ATAAATTGGATTGGCTGGAAGCATGCATCAAGGAATCTTGGCGACTTTATCC
AGTAACGCCTCTCATAGCAAGACAAATATCTGATCCATTACGTTTGAGAGAC
ACAGATATACCATCGGGATCGACAATTCTGATAAATTCGTATATCCTCAATCG
GGATCCTCGACATTACGCCCAACCGGAAATGTATTATCCAGAACGCTTCCTG
CCGGACAAGCCGAAACCGCCGCCTTTTGCTTTCATACCGTTTAGTGCTGGTTC
GCGTAACTGCATCGGTGCAAAGTTTGCACTGATAGAAGTGAAGGTCATGATT
TTACGAATCCTTAATGCGTATAATATACATTCGTTAGACAGTGAAGATAAGCT
GCGACTCACGGCCGAATTGGTGCTTGCGAACACTGGTGGACTTCGTTTGGCG
ATAACTCCACGACATTCGCCAGCTTGA

>CcP450_12
ATGGCGGTCTCCGTGACTGGACTGATCTTGTGCAGTGCAATTCTATGCGGTTT
GATTTGGATTATATTGTATCATGTAAAAAGATTAAGAATGTACAAAGCAGCT
GCAAAGTATCCGGGGCCTGTGGCCTGGCCTTTCCTTGGAAATGCATATTATTT
CATAGGGAAGAACGAAGATATTCTCAACAAAATCATCAAGCTAATGAATACC
TATCCTTCGCCATTGCGTTTCTGGTTGGGACAACGTTTATTCTTTGCAGTCTAC
GATGCAGAAGAAATGAAGACAGTTTTCCTCAGTCCTAAAACCATCGAAAAGG
AAGATCTATATAAGTTTGCACGGCCTTGGCTCGGAACGGGATTATTTACAGC

ACCAGCATCGAAATGGCGGGTACATAGGAAATTGATACTGCCAACGTTTAAT
CAGAAAATACTTGAGTCTTTTGTGGAAGTATTCTCGGCACAATCTCAAATAAT
GGTTAAAAATATGGAGGTGGAAGTGGACGGAGAAGAATTTGACATATTTAAC
TACATATCACTTTGTACACTGGACATTATATGCGAAACCGCTATGGGCGTCGC
CGTAAAAGCTCAGACCGAACGGAATTCTCGTTATGTCGAATCTGCTAAAAGA
GCTTTCGAAGTAATATTTACCAGAATGTTTAAAATATGGTTGCATCCCGATAT
TATTTTCAATCATACTCAACTTGGAGCGGATCAACGTACCTGTGTGGAATATT
TGCACAGTTTGACGAACGATGTAATACAGAAGAAGAAACAAACATTTCATCC
AGTCAATGGCAAAATAGCTAATAGTGAGGTAGTAGAGAATCAACCGAATCA
ACGAAAAGCCTTCTTAGATTTATTGATGGAATTATCTAACAACGGTACAAAA
TTTACCGACGAGGAACTAAGAGAAGAAGTCGACACAATGATGATTGCAGGA
AACGATACCACTGCTACTGTAAACTGCTTTGCGCTGCTGATGATGGCGAATTT
CCCAGAAATACAGGACCGAGTTTATCAGGAGCTTTACGATATTTATGGTATG
GACGATCCGGACGACTCACCAATAAAACATGAGGATTTACATCGCATGGAGT
ATTTAGAACGTGTCATCAAAGAAACTCTGCGACTTTTTCCAGTCGGACCGATT
CTTGTACGACGTGTAATCGATGATTTGAATATAGGAACACACACGTTACCAA
AGGGTAGTTCCGTGGTTCTTGGAATATTTCACGTTCACAGAAGCGAAAAGTA
TTGGCCTGATCCACTGAAGTTTGATCCGGACAGATTTCTTCCTGAAGAAGTTG
CAAAACGTCATCCTTTTGCTTTCACTCCATTCAGTGCAGGTCCGAGAAACTGT
TTAGGACTGAAGTATGCTATGATGGCTATGAAGGCATTATTAGGAACGGTAC
TCCGAAGGTATATCATTAAGAAGGATCACATGACACCTATATCAGAAATGAA
ATTAAAGGCAGACGTTATGTTGAAACCAGTTGAGCCTATTAAATTCAGAATT
GAAAGAAGAATCCCAAAAATTGGTTGTATCGACTAA

>CcP450_13
ATGTCTTGGTCATACACGCGACGTCTCATATACGGTTCCTTGATTACGCCATG
GAAATATTTTGATTTTCATATGATTTCCGTCACGTTAATCATTTTGATCGGGGT
TTGGCTGGCAAAAGTAATTTTGGAATGGCGAAATCGAAGAATCAGACTCATA
GAATTTTCGAAAAAACTACCAGGACCGCCAACCTTGCCATTAATCGGGAATG
CATTGGCATTTGCATGTAGCTCGACCGAAACATTAAATGTAGTAACAAGTTTG
TGCAATTACGGTCCCATATTTCGAGTTTGGATGGGACCGAAACTCTTCGTTGC
TATAACGGATCCTCGAGATTACGAGGTAATTCTTACGAGTCCGAAAGCATCG
CACAAGGGTGCCTTTTATCGTTTTATGAAACCCTTCATCGGCGAAGGTCTTGT
GAGCGGTTCAGGGCCAACGCACAGAACTCATCGTAAAATCATCATGCCTATG
TTAAACGTAAGGGTTCTGGAAGTGTTCATACAGTATTTTAACGAACACAGCA
AATACTGTGCAAATCAGCTCGAAGAGATGGTAGATACCGGAGAATTTGATAT
TTTGCCATTTACCGCACACTGTGCAGCCGACATAACACTGGAAACCGTAATG
GGTATACCTGGTACAGTACAAAGGGGTGGATACAAACAATTGATGCATTGGG
CCGAAAGAATGTACGAAGTCATTCATACGAGAATAATGAAAGTCTGGCTGCA
TCCCGAATGGATCTATTCATGGACAGAGTACAGAGCGCAAGAGAAAATAGG
ACAAAATGTAATTCAAGGTTTTACAGAATCTGCGATAGAACGTAAGAAAAAG
GAGCACTATGCTTTGGAGCGTGGCGCATTGATATCCAATAGGCCCAGAACGA
TGATCTTGGAACAGCTGATAACCTACGTGGAAAATACCCACGTGATGAACGA
CGAGGAACTGCGAGACGAGATATACACCGTATTTACTGCTGCACAAGACACG

ACGGGCTTGGTCAGCGCGTTTGCACTACTCTTACTGGGCATGCACAGCGACA
CACAGGACAAAGTGCGAGCAGAGTTGAAGGACGTTGTGGGAACGAAGGATA
TTACTATGGAAATGATTCCAGAATTGAAGTATCTCGAAATGGTAATAAAAGA
GACGCTACGTTTGTTCCCTATAGCTCCAATGTTCGTGAGAGAACTCACGGGCG
ACGTTGAACTTGAATCGTGCACGCTTCCGAAAGGCTGTTCCGTGGCGATGGTC
CCGTTCGTGACGCATCGAAGTCCAAAATATTGGATGGAACCGGAAAAATTTA
TACCGGAACGATTTTTGCCCGAGAATTCGCAAAATCGGCATCACTATAGCTTC
GTGCCGTTCAGTGCCGGTACCAGAAGCTGCATCGGTCAAAAATACGCAATTA
TGAGTATGAAAATTATTATTGGCACTATTGTCAGAAAATATCAAATAACAGC
GAAAGGCTCGATGGAGACGATCCGATTGAAAACTGATATATCGGTTCGATCG
GTCGATGGATATAAAGTATCGATAACTCGTACCACCTGA

>CcP450_14
ATGGCAGCACTGGAGATACTGTGCAGCGTTGCTGTTGTATTTATATTGTTTTA
CTACTATGTTACGAGAAATTATAATTTTTGGAAACAACGTCATGTCGCCGGGC
CTGAGCCTATGTTTCCTTTTGGAAATTTTAAAGAAATAACGTTTGGGCAGTTA
CCAATTGGCATTCTAATAAAACAATACTACGATCAATTCAAAGGAGAGCCCA
TGGTCGGTATATTTGTCAGAAGCTTACCTACTCTGATAGTGCGTGATCCTGAA
TTGGTTAAGGACGTATTAATAAAGGACTTTACGAGTTTCCCAAATAGAGGCA
TACACTATAACACAAAAGTCGATCCGTTAAGCGAAAATTTGTTTATGCTCAA
AGAAACGAAATGGCGACCACTTCGTCGAAAACTCACTCCGGTATTCACGTCC
GGAAAATTGAAGCAAATGTTCTATCTATTGTTGGAATGCGGAGATGTATTCG
AAAAGTATCTTGATCATCTTTCGGAAAAGAATGAACCAATAGAATGTCGAGA
CATTACAGCAAAGTTTACCACAGAGGTCATCGGTACATGTGCATTTGGTTTGA
AGACTGATGCGATTGCTGATGACAACAGCGAGTTTCGAAAAATGGGTATCAG
AGTATTTTCTACAAGTTTTATGAAGACGTTGCAGTTTATTCTGAGACAGTCTA
TGCCCTGGCTTTATAATATTTTGGGAAAATGTATGGCTGACTGGGTCGTCATT
GATTTCTTTACAAAATCCATAAAGGACACAATAGAACATAGAAAACAGAATA
ACATTGTCAGACATGATTTTATTGACCTTCTTATAGATATGAAAGAACATTCT
GATAACATGAGTGACGTCGATGTCACGGATAATTTAATTGCTGCCCAAGCTTT
TATTTTCTTTGTTGCTGGATTTGAAACTTCCTCTACGACCATGAGCAACGCCCT
CTACGAATTAGCATTGAATCAAGATATACAGGATAAGCTTCGCCAAGAAATA
AAGGAAGAGATTATGAAGAACAATGGCGAGTTAACCTATGAAGCTGTAAAG
GACATGAAATACCTTCATATGGTATTTCAAGAAACTTTAAGAAAATATCCAC
CGGCAACTGTAATGTTTAGAGTTGTAAATCAGCCGTACACGTTTCGTGGGACT
AATGTCACGATTCCTAAAGATCATAGAGTAATCATTCCTATTTATGGCCTTCA
TCATGATCCAGAAGTATATCCTTCACCGGAAACGTTCGATCCTGAAAGGTTCA
CTGCAGAAAACGTTAAAGCTCGACATCCAATGTACTTTATACCGTTTAGCGAC
GGTCCTAGGAACTGCATAGGTGAAAGATTTGCGGTATATCAGACAAAAGTTG
GTCTGATTAAAGTTCTAATCAATCATAAAGTAGACGTATGTGAGAAAACAAA
AATCCCATACGAATTCAATACAAGAGCTTTCATTCTGACACCGAAAGATGGT
ATCTACTTGAAATTTTCTAAGTTAAGCGCATAA

>CcP450_15

ATGCAGGTGGAACGCTATCGAAAGAATTTCGCTATCGAACGTAGACATTTTT
GCACAAGGATGACTGTGATTTTTGAGCAATGGGTGAATTACTGGACCATTGG
ACTTTTGACGGTGATTATTGTAATCGTATACGTTTACGTGAAACGACATCACA
GTTACTTCCAAAAAATGGGAATTCCATATATTGAGTCACATATTCTCGTAGGA
AATCGTGGACCCGTAGTTTTCAAACAGAGATCACTTTTCGATCATTCTCAGGT
GCTCTACAATTCCATACCAGACGCAAAGTACGTTGGCATATATGATTTTGCTA
TACCAGCAATTCTGATACGGGACGTCGAACTCATTAAGGATGTCGGTGTTAA
ACATTCTGAAAATTTTTTGAATCACAAAATCCTCTCCGATCCGCAACTGGATC
CGCTCTTTGCAAACAATCTCTTCAATTTACAAAACGACAAATGGCGAAAAAT
GCGAGCTCTGCTGACTCCAGCTTTCACATCGAGCAAAATGAAGATGATGTTT
GATATGATAACGAAGTGCGCTACAGATTTTGCGGATTTTATAGAAAAACAAT
CTCCCATTGTTGTGGACGTGAAGGATGGCTTTGCAAGATACACAAACGACGT
GATCGCCGGTGCAGCTTTTGGAATTACCGTGAATTCTATGAAGGAGCGAAAT
AATGAATTCTACTTAATGGGCAAAGAGGCGACTACTAGTGGGTTTGTAGAAT
TAATGAAAATGATCCTCTATCAAGCTTCTCCGACATTAGCGAAAAAGTTGAA
GGCTAACGTGGTCAGTCACCGGGCAGTCAATTACTTCTGTAACATTATCGAAT
CGACCATTGCAACGAGGAAAACGCAAGGGATTCAAAGATCTGACATGATTCA
ACTTCTAATGGAGGCGAAAGATTCTGGAGATAAGTTGGATTTGACTATCGAA
GATATAACAGCGCAAGCATTCATATTTTTCTTCGCGGGTTTCGATACGTCATC
CTCAACTATGAGCTTTGCTGCTTATTCGCTGGCTACTAATCCGGAAATTCAAA
AAAAGGTTCAAGCGGAGGTCGATGATATTCTTAAGAAATGTAATGGCAAAGT
GACGTACGAACATCTCAAAGAAATGAAATATCTCGAAGCAGTTGTCAGCGAA
ACTCTAAGACTGTACCCGGCAGCCGCCGAAATAGATCGAGTTTGCAGTAAAA
CATTCGAACTTCCTCCAGCTAGACCAGCTGCAAGATCCGTTACTGTGAAACCT
GGCACCGTTATATTATTTCCCGTGTATGCAATTCATCGCGACCCTCAATATCA
TCCCAATCCCGACAAATTTGATCCTGACAGATTCTTGAATGAGAATGAGAAG
ATCAACCCGTTAGCATATTTTCCTTTTGGCCTTGGACCGAGATACTGTATTGG
CACTAGATTTGCTCTAATGGAAGTTAAATTAGTTCTCTTTTATCTACTGATGC
AGTGTAACTTTGAACCATGCGAAAAAACAACTATTCCTATGAAATTCAGTAC
TGCAAATTTCCTCCCAAGACCTTTAAATGGTCATTGGTTGAAAATTGTACCAA
GAAAGGGAATAAAGCGTTTGTAA

>CcP450_16
ATGGTATTATTAATTCAATCAGCGACGTCACGTCTGGGACATCGCTTGTCACG
ATCATGCCGATCGTTATGTAGCGAAGCACTGAATGCAACGGGTCATGTGACG
TCGTCTATTTCCGGAGACCTTGATAGGGCCCAATCGTATGATCTTGTACCAGG
CCCAAAACCTATACCGATTTTGGGCAACACGTGGCGCCTTCTACCGGTCATA
GGTCAATATCAGATATCGGACGTGGCTAAAATATCTCAAATGTTTTATGACG
AATATGGAAAAATCGTTAGACTGACTGGTCTGATTGGTAGGCCTGATTTATTA
TTTGTCTACGATGCCGATGAAATTGAAAAAATATATAGACAAGAGGGACCCA
CGCCGTTTAGGCCATCTATGCCGTGTCTTGTTCACTACAAGAGTGAGGTTAGG
AAGGACTTTTTCGGAAAATTGCCTGGGGTTGTTGGAGTACACGGGGAACCGT
GGAGGGAATTTCGTACGCGGGTCCAAAAACCAGTTTTACAACCTCAAACTGT
ACGGAAATATATTGGGCCTATCGAGGTGGTCACTGATGATTTTATTCAGAGA

ATTGAAAAAATTAAAAGGGACGATTGTGAACTTCCGGCTGATTTTGATAATG
AAATACATAAATGGGCTTTGGAGTGTATTGGCAGAGTTGCGCTCGATGTCAG
ACTTGGATGCCTTGGCGATAAGTTACCGCTAGATTCCGAACCGCAGAAAATC
ATTGATGCGGCAAAGTATGCTTTGAGGAATGTCGCCGTTTTGGAATTAAAGG
CGCCATTTTGGAGATACATTCCTACGCCACTTTGGACGAGATACGTGAAAAA
TATGAATTATTTTATTGAAATTTGCATGAAGCATATAGACGTCGCGATGGAAA
GATTGAAAACCAAAAAAGCTTTGAGCGAGTTGGATCTGTCGTTAATCGAAAG
AATTTTGGCAAAAGAATCTGATCCAAAAATGGCATATATTTTGGCTCTTGATC
TAATACTAGTCGGAATAGATACGATTTCAATGGCAGTATGTTCCATTTTGTAT
CAATTAGCAATCAGACCGGAAGAACAAGAAAAGATTTATCAAGAATTAAAA
ATGATACTTCCGGATCCCACCGTGTCGTTAACAACAAAACATTTGGATCAAG
CAGTATATACAAAAGCATTCGTCCGTGAAGTGTTCCGAGTATACTCCACTGTC
ATTGGCAATGGTCGTACCTTGCAACACGACACTGTCATCTGCGGTTACAAAAT
ACCAAAAGGTGTCCAAGTAGTTTTCCCCACATTGGTAACCGGAAATATGGAA
GAATACGTAACGGATTCAAAAACCTTCAAACCTTCTCGTTGGATAAAAAACG
ACAATAATGAAAAATTGCACCCATTCGCATCATTGCCTTATGGTTATGGCGCC
CGAATGTGTCTGGGCCGAAGATTCGCCGATTTAGAAATACAAGTCTTACTAG
CAAAGCTCATCCGCTCTTACAAATTGGAATATCATCACCAGCCTCTAAAGTAC
AAAGTGACTTTCATGTACGCACCTGACGGTGAACTCAAGTTCAAATTAACTA
AGCGATAA

>CcP450_17
ATGTCATACCAAAGAGCAATGTTGGTTCGCAGTGCTTTGCTCGCTGGTGCGAA
AGGTAACGTGCTTAATATTGGCTTGATTGAACGTTATGCCAGCACAGTGTTAC
AAGCGGATGTAAAGAGTCCTGAGGAAACGGAATGGAACGAGGCCAAGTCAT
TCGATGAAATACCAGGACCAAGGGCACTGCCGATCATTGGGAACATGTTCAG
ATTTCTTCCTCACATAGGTGAATTCGGAAATCAGCCGGTCATGGAACAAATG
AGAAATTTGAAAAAGCAGTATGGCCCGATTGTTAAACTGGACGGTATACCGC
ATCGTCGAACAGTCGTATTTCTGTTCGATCCGGAACTCAGCGAGAAAATGTA
CAGAACGGAAGGTCGATGGCCGACCAGAATTGCTATGGAGACAATGGCACT
GTATCGCGAGGAGAGGCCCGAATTGTATAAAGGAAATTATGGATTAGCTTCG
AGCCAAGGCAAGGATTGGCACAATTTTCGTTCCAAAGTAAATCAACACATGA
TGCAACCTCGAACGATAAAACCTCACGTGGGACAAATTAATGAGGTGGCTCA
GGACTTTATTGAAAAAATGCGCAAACTTCGAGATCCGGAAACTCTGGAACTT
CCGAAAAGCTTTAACAATGAAATGAACAAATGGGCGCTTGAATCTATTTGCG
CAATCGCACTGGACCAAAGACTCGGATGCCTCGAATTTGACCTGCCGACAGA
TTCGGAGCCTCAAATAATGATAAATGCTGTTCACGACATGTTCGATCTCTTCT
ATAAATTGGAAATTCTTCCCTCCTTGTGGAAAGTCTATAGGACACGAAATTTG
AAAAAATTATTTCATTGCTTCGACATCATGAGTGGAATTAGCAAACGATACA
TTAGCCTGGCTAGAGAGAAACTACGTGAGTCAACATCAATGGGCAACGAAGA
ACTAAGTATGCTGCAACGACTTGTAGCTATCGACGAGCAGACAGCAAATGTC
ATGATAATGGACATGATGCTTGCAGGAATAGATACTACCAGCAATGTCGCCG
CAAGCGCTCTCTATCATATTGCGACTAATCCTCAAGCCCAAGAGAAACTCCA
CCAAGAAGCAGCCAAGGTGTTGCCAGAGAAAGATTCGCCAGTAACTTACGAG

GCGCTTAATAAAATTCCCTATACGAAAGCCTGCATCAAGGAAGCCCTGAGAA
TGTCACCCATCGCCATTGGTAATTTAAGAACTATGCAAAAGGACATTGTACTT
GAAGGCTATAAAATTCCTAAGCCGCACTGGAAATCCTTGCTAAAAAATAAAT
CCGTGATTTCCAGGCCATTTTGTTACATTACCAGAATCCGAGTATCGGCAATA
TGGCGTCAGACAAGTTACGGCGGGCAATTTGAAATCGATAGTCCGTGCCATT
TAGTGGTTATGTATCGACGGATGCTTTCTGTTATTGTTATGACTATTATCTGTT
TGCAGTTCTAG

>CcP450_18
ATGAATTATATACTCGAATTGAGTTTCGGAGCGGCGATAATTTTGGTGCTCTT
AATACGGGCTAAAATAAAAAAATATAATTATTTCAAAGAACGCGGTGTGCCC
ACTCTAAGCATTTATGCGGGTCTTACTAATACGTGGAAGCTGTTTACAAAGAA
AATATCTTTTCCCGACATGGTGATGGACCTTTATAACGCGCATCCAAATGCCA
AGTACATCGGCATCTTCGATCAAGGATCGCCGGTTTTCATCATTCGGGATCCT
GAAGTGGTCAAGTTGGTAACAGTAAAAAACTTCACGGTCTTTCTCGATCACA
AAGGATTCGTCAAGGAAGAGGTCGACCCTCTTCTAGGTGGTGGACTATTCAA
TATGAAGGGTGGCAAGTGGAAGGAAATGAGAGCGGTCATCAGTCCAACGTTC
ACAACCAGTAAAATGAAATGCATGTTCGAGCTAATGTCAAAATGTGGCCAGG
ATTTTGTTGAATTTTTGGAAAAGGAAAAACAGCCGCTTGTTCTCGAAATGAA
AGACGTTTTCACGCGTTACGCCAATGACGTAATAGCGACAACAGCATTTGGC
GTGACAGTGAACAGCATTCACGATAAGGAAAATGAATTTTATATGATGGGCG
TAGAAGCGACGAGCTTCACTCCACTTCGTAGCTTCAAGACAATCCTATTTCAA
TGGTATCCAAAGATAATGAAGTTCCTTAAGTTTAAAATAGTGTCCGATAAGG
TGTCCAACTTTTTTCGTGATATTGTGATAACGACCATAAATACGCGCAAGGAA
CAAAATATACATAGACCGGATCTCTTGAACATCCTGATGCACGCGAGAATGA
AGGATCCGGACAGAGAGTTGACGCATGACGAAATAATATGTCAAGCTTTCCT
CTTCTTTTTCGCGGGTTTCGAAACCGTCGCCTCGGTGATGTCATTCGTCGCGT
ATCATTTAGCTATATATCCGGAAGTGCAGGAGAAACTGCGCAAAGAGGTGGA
TGGGGTTTGCGGACGAGAAAATGGTCAACTGATTTACGAGGAACTTGGTGAA
ATGAAATATATGGAAATGGTTCTGAATGAAACAATGAGGCTTATTCCACCGA
GCGTTTTGACCGATCGCGTTGCTGTGGAAGATTTTCATTTACCGGCCAATGAT
GAAAAATCACCACCGATACACATCGAAAAAGGCACTGCTATTTGGATTCCTA
TAGCGGGATTTCAAAAGGATCCGAAATACTTTGCAGAACCCGAAAAGTTCGA
TCCGGAACGCTTCGATTCTGATCATAAGAATAATATTTTGCCATTTACTTTCA
TCCCCTTTGGTGTCGGTCCCAGAAGCTGTCCTGCTAATCGTTTCGCTTTTATGG
AGATCAAACTTATTTTGATAAACATATTGAAAAGATTTGTTATAAAACCTTGC
GAAAAAACATCTATACCTATGAAGTTTGCTAAGGGTACGATAACTGTCGTAG
CGGAAAAAGGTATTTGGTTATCTTTCAAACCTAGGAAAAACATTGAATAA

>CcP450_19
ATGCGTGAACGTGTCCATCCAGCCCAGTATGAAAGTTACGTCGCGAGTGAAA
CGTGTCGTTCAGTTCTCGCGATGGAGTATCTGTGGCATGCATTGCAATGTGTC
CTCGTAATTGTCGGCCTCACTCTGGCCGCTCTAATTTACCGTCAGTACACTTA
TTGGAAGAAACGTGGGATCCCATACATGGAAGTGATACCACTTATTGGAAAT

CAGGCGCCGATCTTCTTCAAACTGAAAACTTTTCCCGAGCTTATAATGGACAT
GTATAACCGATACCCCAATGCAAGATTCATCGGTATGATGGATTTTAATACGC
CCCTCGTCTTCATTAGGGACCCCCAATTGATTCGCGACATTGGCGTGAAGAAT
TTCGACAACTTTCCGGATCATATGAGCTTCGTCGATGAGAAGGTGGACAAGC
TTTTTGGAAAGAACATATTTTCTCTTCGCGGCGAACGCTGGAAGGAAATGAG
GGCGACGCTGAGTCCCAGCTTTACCGCCAGCAAGATGAAGTTCATGTTCGAA
TTGATTTCCAAGACTTCCAAGGATTTCGTCACATATCTACTCGAGCATCCGGA
AGAGACAAAGTCCATCGAGACGAAAGATGCCTTCACGCGATACACGAACGA
CGTGATCGCCAGTGCAGCTTTTGGCATAAGCGTTAATTCGATGCGTGAAAAA
GATAATGACTTCTATATCCACGGAAAGGAGGCAACAAATTTTGAGGGATTCT
GGAAAATTATGAAGTTTGTTATGTTTCGCGTATGTCCGAGTATTTTGCGATTG
ATGGATCAAGGCTTTGTAACGCGCGACACGGCAACATTTTTTAGGACCCTTAT
CCAAGATGCTGTTAAGGCTAGAGACGAAAAGGGAATCGTCAGACCGGATAT
GATTCACTTGCTGATGCAGGCCAGAGATAACGAACGTGTATCATTGTCCACC
GATGACATCGTGGCTCAAGCTTTTATATTCTTTGTGGCTGGTTTCGATACTTCC
TCGACTTTGATGTGTTACATGCTTCATGAGTTGGTGCTCAATCCTGACGTCCA
GCAAAAGCTTCGCGAAGAGGTTGATCAGGTGACTTCTGGTGCCGAGGACGAG
GTCACCTATGAGATGTTACAGAAGATGAAGTACATGGACATGGTCATCTCTG
AAACTCTACGAAAACATCCTCCAGTTCCTGCGATCGATCGTGTCTGTGTTAAG
AGTTTTCAGCTTCCGCAGTCGATGCCAGGGTACAGTGATGTCACAGTTGAGA
AGGGCTCTGTCGCGTGGTTTCCAGTGCACGCTCTTCACTATGATCCACAATAT
TTTCCGGATCCATATAAATTTGATCCTGAAAGATTTAATGACGAGAATAAGG
ACAAAATTGTGCCTTACACTTATTTGCCCTTTGGACTGGGTCCTCGAATGTGT
ATTGGAAATAGATTCGCGTTGATGGAGACGAAAATTATCGTTACACATTTGG
TGCAAAACTTTATGTTCGAACGAACGGAAAAGACTGTTGATCCGGATAAGTT
TACTCTACAAAGTTTTAGTATTACACCGAAAGAAGGTTTTTGGCTTCGGCTAA
AGCCACGAAAAATTTAG

>CcP450_20
ATGTCGGTCTTATCCGACATTTGGGACTTGAAGTTCTACCTGCTAGCTGGGGC
GATTGGCGCAATTTACTTGTACGTCAAGTTGGTGATTTTCAAGTATTTTGATA
GGAGAGGCATTCCGTATGAAAGACCCACGTTTCCTTTTGGAAATGTCAGTGG
AGTATTTCGAGGAAAAATGTCAATAGGGGAAGCGTGCGCGGAGTTGTACAAG
AAGCATCGACAATCCCCGTACTTTGGCATGTTTATATCAATTAATCCAGCCTT
GGTGATCAATGATCCTGAGATCATTCGTCACGTCCTGACGAAGGACTTTTCCC
ATTTCCACGACAGAGGAATATATGTGGATGAGAAAGTTGATCCACTTTCCGG
ACATTTGTTTAGCCTTCCCGGTGAGAAATGGAAGAATTTGCGCACCAAATTG
ACACCGACTTTTACGTCTGGCAAAATCAAGCAAATGTTTCCCATTGTGAACCA
ATGCGGCGTGGAACTTGGCGATTGCGTCGAGAAGGTTTTGAAAATTGACGAA
ACTGTGGAAATTAAGGATATCGTGGCGAGATATACCACTGACGTCATATCGA
GCGTGGCTTTCGGTTTGGAAACCAACAGCTTGAAGGAACCGAACAACGAATT
CCGAAGATGTGGGCAATTGGTATTCTCATCAAGACCACTGACGACTTTGTTAT
CCCTTTTTATGCCGAGTCTTTTGCGCATATTGAAGCTGCCAAATACAGACACA
CGAGTTACCAATTTTTTCACAAAAGTATTTAGTGATACAGTCAATTACAGGAA

GCAAAATAACATTAAGCGGCACGACATGTTGAACATGCTGATGCAGTTGATG
GACAAAGGATACGTTGAAACTGACGACGCGAACGAAATTCCCAAAGGGAAA
GCAGTAAACGATAGCAAGAAGCTGGAAATGACGGAAGCTACGGCCCAGGCC
TTTGTCTTCTTTGTAGCTGGTTTTGAAACTTCCTCGTCTACCGTGACGTACTGC
ATTTTCGAAATGGCCTGTCACCCAGAGATTCAAGAGAAGGTTTACAAAGACA
TAGACCAAGCATTGAAGAAACATGGCGGCCTGACGTACGACAGTTTGGCAGA
GATGGAATATCTTTCCCAGGCAGTAGACGAAACTTTGCGGTTATATCCCACAC
TTCCGATACTGAATCGAGAATGCACCAGGGAATACGATGTGCCTGGAACCAA
CTTGCACATTTCAAAAGGAGATGCAATAGTCATTCCAGTTTTGGGTATTCATA
AGGATCCTAATATATATCCTGATCCCGAGAAATTCGATCCTGAACGTTTTAGT
AAGGAAAATAAAGGAAAGATTCACTCACACGCTTATCTTCCGTTTGGCGAAG
GACCAAGGATATGCATTGGCAAAAGATTCGGCTTGATTCAAGCGAAAGTAGC
CTTGGTCAGTTTACTGTCTCGATTTCGATTTTCCACAAACTCAAAAACGGAAA
TTCCAGTGAAGCACAGTCCTGGAACTATTGTAATGTGTCCAGATCACGGAGT
AAATCTGCGGGTCGAAATGCGATAA

>CcP450_21
ATGGAGTTCCTAAGTGCGGATGGTATCGGCATCACTCTGGCTTTGGTGGCTAC
AGCAATATCGTTGTTTTACATTTACGCCAAGTATAAACTCTCTTATTGGAAAC
GACGTGGTGTAGAGTCTCTGCCAACGAATTTGATATTTGGCAATTTTAAGGAT
GCCGTACTATTCCGTACAGCACCAGGGTGGCATCTTGGAAATCTTCATAAAG
CCGCTAAAAAAGACGTTCCGTACCTTGGATTCTATATATTTCACAAGCCCTGC
TTCCTACTTCGCGATCCTGAAGTCATCAAGCAGATCTTGATAAGAGACTTCGA
GAACTTTTCCGATCGTCATTTTGCTGGTTCCGAACAGAAAGACAGTATTGGAA
TGAAGAATCTTTTCGGTATTAAAAATCCCAACTGGAAATACCTGAGGGCCAA
GATCACTCCGACCCTTACACGTGGCAAATTGAAGCAGATGTTTCCCCTTATGC
TGGAGACTGGTGAACCCATGATGAAGTATTTGGAGAATCAACCTGAAAATAA
AAATGGTGTAAAAGTGATTGACGCTCAAGAACTTAGTTATAAATACACCACC
GATTTAATCGCATCAATCGCACTCGGCACCAAAATGGATTCCTTCCATTATCC
CAACGAAGAATTTTCCAAAGCAGTGATGGAATTCTTCCATGGATTCAAACGT
ATGGTCGCCCTTGTCACCGTGTTCTTCATGCCTGAATTGGTGGAGATGATTGG
CACCCGCATGTTATTCAATTCCAGTTTCGTAAAGAAGGTCTTCTGGGAGGCCA
TGGAGTCCAGAGAACGCACTGGAAATAAGAGAGGAGATTTTATTGACTCCCT
TCTGCAACTTAAAAATGGCGAACAGAACTCCGACTACAAATTCGAGGGTGAG
AATCTTCTTTACCAATCCGGCACCTTCTTCTCTGGCTTCGAGTCCAGCTCCAC
GACAATGTCTTTCACTCTTATGGAACTTGCCAATCATCCGGAGTACCAGGAAC
GTGTTCGTCAGGACATCAACAAAGCTATTGACGCTCATGGCTTCACCTTTGAT
GCCTTCAATGAAATGAAGTATTTAGATCAAGCAATTGCCGAAGGGGTTAGGT
TGCATCCTCCTGTATCTACCATCGACAGATACACACGTAAAGACTATAAGATT
CCGGGCACCGACATTATTCTCGAGAAGGGTACACCCATATACATTTCCCTTTA
TGGTCTTCAGGAAGATCCTCGTTTCTTCAAAGACCATTCAACATTTAACCCAG
ACAGATTCACAGACGACAGTAAAGTCGCCGATGCTTATATTCCTTTTGGCAGC
GGACCAAGGATGTGCGTTGGAATGAAAGTTGGTCAGCTTCATGCTAAGGTTG
TCATCGCGATGCTTTTACGCGAATACGAGATTTGGCAGAAGCCTGAAGACAA

GAGCATTCTGGATCCTCGTTCGACCTTCACTGCAGCAGCTAACGGCATTGTAT
TGCAATTTAGAAAATTGGTTAAATGA

>CcP450_22
ATGGAGACGTCAACATCGCACTTGCTGATTTACGTAGTGACGTTTATACTGTC
GTTAGTCGGTCTCCTCTATATTTATATCAGATATAAGTTGTCATATTGGAGCA
GACGAGGTGTCAAGACGCCTCCTGTCCACTTGCTATTTGGTAACTTCAAAGAT
ACTATTACATTCAAGAAAGCACCTGGCAATACGTTACGGGATATACACGAGA
ATGCTGACGCAGACGATCCCTATATCGGCTTCTATATATTCCAAAAGAGATAC
TTGATGCTGAGAAATCTTGACATTATTAAGCAGATATTCATAAAGGATTTTGA
TTTGTTTCCAAATCGTAATTTTGGTGGAAAACGAGAAATAGACTCCGTAGGCC
TTATTAATCTCTTGGGAATAAGACAACCGGGATGGAAATACCTGAGACAAAA
GATAACCCCATTTATCACTGGACAAAAACTAAGAAGAATGATACCTCTTATG
GTGGATTCCGGATATCCGATGACGAATTTTATCGATGATAAATCCACTGACA
GCGAAGGATGGAAGGAAGTTGAATTGAAGGATATCTCGTCTAGATATAGTAC
CGATGTTATATCTTCTTTGGCTTTTGGCGTTACCACCAATTCTTTTAACAAGAA
TGATACTGCCTTTTGGAAAGCAGGACTGAAGATACTCTCTGGTACTAAGAGA
AATATTATACTTATGATTGTATTCCTCTTACCTGAGCTTATTGATATTATTTAT
CCAATAATGACTGGTCCTGCACAATTTTTTAGAGAAGTTTTTTGGGACTCGAT
GAACACTCGCGAGAAAGCCGGTTATAAGCGTGGCGACTTGATTGACTCTCTT
CTGCAATTAAAAAATGCCCAACAGAATCCGATATTCAAATTTGAAGGCGATA
ATTTGTTGGGTCAAGCTGTCGCTTTTTATGTTGCTGGATACGAAGCATCTTCC
ACTGCAATCGCTTTTACACTGTACGAATTGGCACGACACCCGGAATATCAAG
AACGCGTCTTTCAGGAAATTCAAACGGAAATATCGAAAAATGGTATCGATTA
TAATTCCATAAACGAGATGAGTTTCTTGGATAAAGTAGTTATTGAGACGTTGA
GGATGCATCCGCCTTTACCTTTGGTTGATAGATTTGCTATTGAAGATTATAAG
ATTCCAGGGACTGATTTGATAATAGAAAAAGGTACACCTGTTATTGTGTCGA
CGAACGGTATAAATCGTGATCCCAAATATTATTCAAATCCGGATGAATTTGAT
CCTGAAAGAACACAGGAAGGTAAAGAAGAATTGATTGGGGCTTCGTTGACTT
TTGGAATTGGACCTCGGTCGTGCGTAGGACAAAGGATAGGGACTCTTATGAC
AAAAGTCGGTGTTATAACGATACTATCAAATTATGAGCTATCTTACAAGTGTA
AGGAAAACGAAGATTTTAATGCTATAACAATATTCACAGCTGCTGCAGATGG
TGTTGTTGTTCATTTAAAAAAGCGGGAGAAACCTATCAAAATCAGTGTCAAT
GAATGGGCGACTGATTGA
>CcP450_23
ATGGGGGCGTTAACGGGACATTGGATCGGCGATGGTCTCGTGCTGATGTTTTC
AATAGCCGCTGTCTTTTACCTTTACATGACGAGGAAATTTAATTATTGGAAAA
AACGCGGAGTTCTGGAATTCAAGCCGGTTCCGTTCCTCGGTAATTTCGGGGA
GTGCCTCTTCTTCCGTAAATCGGCCTTGGACTTTATCAAGGATTACTACAATA
ATTCCAAAGGACTTAAATACATGGGTTTCTACATACTGGACCAGCCTTTCCTG
ATGCTGCGCGATCCTGAGCTAATCAAGCACATTCTGGTCAAGGACTTCAACT
ACTTTCACGACAGATACACAAGCACTGGGAAAAAAGACATCCTGGGTAATGC
GAATTTATTTACCATCAAGAATCCAGCCTGGAGATTCCTTAGGGGTAAGCTCA
CGCCGTTCTTCACATCTGGCAAGATGAAAAAAATGTTTGAGCTTATGCTCGCC

GTCAACAAGGATCTTGACACTTATTTGGAATCTCTTGATATTGACAGCAAAGG
TCGTGAAATGGAGATGAAGGAACTGTGTGCAATGTACACTACTGACCTTATT
GGAGTAACGGCTTACGGGCTCAGAGTAAATTCCCTGAATAATCCAGACGCCG
AATTTCGAGCTTGTGGACGTCAAATATTTAACTTTAATTGGAGACGCGGTATG
GAGTTGGCCTGCACCTTTTTTGCGCCAAAAATTGTCGATCTATTTGGATTTAA
GTTCTTTAGTAAGGAGAGCACCGAGTTTCTTCGAAGAGCATTCTGGGATACG
ATCGACGAGCGAGTCAAATCTGGCGTCAAAAGGAATGACCTTATAGACCTTC
TTATTGATCTCAGGGAGCAGACGAAAAATGGCGAGAACAAGGAATTCAGATT
TGACGGTGACAATTTGGTCGCTCAGGCAGCAATTTTCTTTACCGGCGGATTTG
AAACTTCTTCGTCGGCAATGTCGTTCTCGCTTTACGAACTTGCACTGCAGCCT
GAAATACAGAGGAAGCTACGAGTTGAAATTCTGGAAGCCTTGGAAGATAAA
AATGGCGAACTTACTTACGATATGGTTGCAAATCTTCCTTACCTGAATATGGT
GGTGTCGGAAACTTTAAGGAAATATCCGCCGCTGCCGTTTTTGGATCGTGAA
ACCAACATGGATTACAAATTGCCTGGAACTGACTTGATCGTGGAGAAAGGAA
CACCCGTTTATATACCCCTTCAAGGCCTTCATTACGATCCACAATATTTTCCA
GACCCAGAGAAATTTGATCCCGAACGATTTTCCGAAGAGAACAAGAAAAATA
TACCTCCTTGTGTTTATATGCCTTTTGGGGAGGGTCCTCATATTTGCATTGGTC
TTCGATTTGGACTTTTGCAAACGAAATTGGGTCTTATAAAGATACTTTCCAAA
TACGAGATATCAGCCTGTGACAAAACTAACGTTCCGATAGTATTTAATAAGA
AAGCGTTGGTCATTGCTGCTGAAGGTGGGATAACCTTGAAAATTAAGAAATT
AATGAAGGTACCTGGTTAA

>CcP450_24
ATGTCTTCGATAATTTTGGAATTGATGAATTCGTGGGTTTTTATAATAGCGGC
AATAGCCGCATTATATGTCTACATGAACAGGCAGGGCAATTATTTCAAAAAA
TTGGGAATTCCTTACTTGAAATCGTATATGTTAGTGGGAAATCGTGGACCAGT
AATCTTTAGACAGAGATCGTCCATAGAACACGAACAGGCGCTTTACAATTCT
ATGCCAGACGCCAAATATTTTGGCATGTACGATTATTCTTCGCCGGCAATCAT
GGTTCGTGATATCGAGCTGATCAGAGAGATCGGCAGTAAGCACTTTGACAAT
TTTGTGGATCACAAAAGTATGTTCGATCCGGAGCTCGATTCTCTGTTATCAAG
CACTGTCTTTTTCTTACGGGAGGACAAATGGCGACGAATGCGATCTCTTCTAA
GTCCGGCATTTACGTCAAGCAAAATAAAGACAATGTTTGGGTTGATATCGAA
TTGCGCAACAAATTTGGCAGATTTTGTGGTGAAACAACAAGAACCGATTATT
TTGGACATGAAGGACCTTTTCTCAAGATGCACGAACGACGTGATCGCCAGCA
CCGCTTTTGGGATTAATTTGAATTCAATGGTGGATCGCCAAAATGAATTTTAT
AAAATGGGTAGAGAAGTGACGGATTTTGGATTTTTACGATCGTCAAAAATAT
TTTGTATGCAAATGTGGCCGAAATTGGCGAGGAAATTGAAGGCTAGTTTACT
TACTGATCGAATGGTTGATTTTTTCCATCAAATTATTGCTACAACCGTTGCTA
CGAGAAGAGCACAAGGAATTCAAAGACCCGACTTGATTCATCTTCTTATGCA
GGCCAGCGAATCTGAAACATCTTCAAAATTGACTGTTGATGACATGACTGCA
CAAGCATATATATTTTTCTTCGCCGGTTTTGATACGTCATCTACAACCATGTCT
TTTGCTGCTTATTCGCTTGCTGCTAATCCTGACGTTCAAGAAAAACTTCAAAC
GGAAATTGACGAAGCTCTCGAAAAATGCAATGGCGAAGTTACGTATGACTAT
CTGAAAGAAATGAAATATCTTGATGCAGTTATCAGTGAAACCTTGAGGATGT

ATCCACCAGTGACAGAACTAGACCGATTGTGCACAAAACCATTTGAGCTTCC
TCCAGCTAAAGAAGGTGGGAAACCCGTATTTGTGAAACGTGGTACCTGTTTC
TTAATTCCAGTTAACGGAATTCATCGAGATCCCAAATTCTTTCCCAATCCAGA
TAAATTTGATCCTGAAAGATTTTTGAATGAAAAGGAAAAGATCAATTCTTCA
GCATATATACCATTTGGAGTCGGACCAAGAACGTGTATTGGCGCTAGATTCG
CCATGATAGAATTAAAATTGGTTCTTTTTCACCTATTAAAAATGTGTAATTTG
AAGCCATGTGCAAAAACATGCATACCGATGAAATTCTGTACCAAAACCTTTA
TGCCAGCACCTGCAACTGGTCATTGGTTGCAAATCGAAAGAAGAAAATGA

>CcP450_25
CTTTGGAATTGGCCGGTGATTTCTATTTTCTCTCGTCGTGACATCGAGACGGT
GCTAAGACGTGCTTCGAGATATCCTCTGCGTCCACCTTCAGAAGTCATTTCGT
ACTACCGACAATCTCGGCCAGACCGCTACACCAACCTTGGCCTCGTCAACGA
ACAAGGCATCACGTGGCACAAGCTACGATCAGCTCTGACGCCGGAACTGATG
GGAGCCAATACAGTACTGGGATTTCTTCCTGCATTAAATGACGTCACTGATGA
CTTCGTGGACTTAATACGTTCGCAACGAAGTGGCTCGAAGGTTGTTGGATTCG
AGGAGCTTGCATATCGTATGGGCCTTGAAAGCACGTGCACCCTCATTCTGGG
ACGTCGCTTGGACTTTCTTAAACCTGATTCAGGCTCGTTAACAGCCACGTTGG
CCGAGGCTGTAAGAGTTAATTTCACAGCGTCCAGGGACGCCTTCTACGGGCT
GCCATTTTGGAAATTGTTACCCACCAAAGCTTACAAGCAGCTTGTTGAGAGTG
AAGATGTTATTTACGACACGATATCCGAATTGATGGAGACGACTATCTGGGA
GAAACAGGATGACGCTAAGGATGAAATCGTCGAGGCTGTTTTTCAATCTATT
CTCAGGGAAAAGAATCTTGATATAAAAGATAAAAAAGCTGCGATTATTGATT
TTATAGCTGCTGGTATACATACGCTTGGTAACACACTCGTCTTTCTTTTTTATC
TTATTGGTCGCGACGCCAAAGTGCAGAGACGTCTATACGAGGAGGCTGCATC
TCTTGCACCGCCAGGGTGCAATCTTGTCGCAGAGGATCTTCGTGGTGCCATGT
ATCTACGTGCTTGCATTACTGAGGCATTCAGAATGATCCCCACGACACCATGT
ATTGCTAGAATTTTGGACGAACCCATTGAACTTTCGGGTTATAATCTAAATGC
TGGGACCGTGGTTCTGCTTCAAACGTGGATAGCAGGACTGGATGAAGAAAAT
TTCAAAAATGCATCTCAGTGGATGCCGGAAAGGTGGTTGGAACCCACTTCAC
CCCATTCACCACTTTTAGTATCACCCTTTGGTGCTGGCCGACGTATTTGTCCA
GGAAAACGTTTCACGGAACAGGCATTGCAGCTTATATTGGCTAAGGTCGTTC
GAGAATTCGAAATCATCGCCGACGAAGAACTCGGTCTACAATTTGAATTTAT
TTTGGCGCCAAAGGGTCCTGTATCCTTGTCGTTTCGTGATCGTCTCGAATGTA
CGTAA

>CcP450_26
ATATTTGACATTTTATACATATTGCAGAAAGTTTTGAATAGTCAGCAAGCCAG
TTACAAAGGACGGGTGTATCGATTTCTTCATCCATTTATTGGAAATGGCTTAA
TAAGCGGAGATGGTAATATGTGGCGTCGACATAGGAAATTATTAACGTACAC
CATGACCCAAAAGGTTCTTCATGGATTTGTCAACATTTTTGATAGACACAGTC
GGAAGTTAATGGATAAGTTAGAATCGCAAGCAGACGGTGGTTATGATTTTGA
TATTTTTCCATATATTGAAGCTTGCACTATTGATATTGTTTGCGAAGCTGTTAT
GGGTCGCTTAGATGTTAATGCACAGGATAATGCAGACCAAGAGATTATTCAT

TATACTGCCAAAATGTACAAAATTATATTCGAAAGAATGACCAAAGTATGGC
TGCAATCAGATTGGATATTCAATCAAACGCGCTATTATACAGAACAACAAAG
AGGACGAGAAGTTATACAAAATTTTGTACATACCTGCGTCTCAGCAAAACGA
AATGACCAGAACTATACACCCTTGTTAAGCAGCAATCAAAAATCAATTTTGG
AACATCTCCTAGAATCATTGAAAAGTTCATCAAATTCTCTGACTGATTCAGAA
TTACAGGATGAAATTTACACGGTTTACATCGCATCCCAAGACACCATAGCTCT
CATAAGCAGCTTCGCAACTCTTATGCTAGGCATGCACCACGAAGCACAGGAA
AAGGCAAGAAAGGAGATAGAGGAAATATTCCAAGGAGAAGATGAACCCCTC
ACGTCTGAATCTTTATCTCAATTGAAATACCTGGAAGCAGTTATTAAGGAGAC
CATTCGTCTATTTCCAATAGCACCGTTTTTAATAAGACATCTGCGTGGAGAAC
TTCCTCTGGAGAATTGCATCTTGCCAGAAGATTGTCAAGTACTAATAGCGGCA
TATGTTACTCATAGAAGTGTCGATTATTGGAGAGAACCAAAGAAATTTCTAC
CGGAGAGATTTTCTGCAGAGGAAAGTAGTACACGGCATCCATATGCATTCAT
ACCATTCAGCGGAGGTCCCATGTCATGCATAGGACAAAAATTTGCAATGACA
TGTTTGAAAGTTATACTCGCGAATCTCTTACGATGCTACAGAATCGATACCAC
TTGGCAAATGTCAGACTTGTTACTCAATGCAGACATTTCTGTTCGCTCCATTA
ATGGCTATCGTGTCTCGTTAAAACCGCACGAGAAAGCAGCTTGA

>CcP450_27
ATGTCGTGGATCTGGTCGATTTTGTTTGACACGAAGTTTCTTCTATTGGGGAC
ATTGATTATTTCGATTTACTTGTATATGAAAAAAGTGGTTTATAATTATTTCG
AAGCAAGAGGAATTCCATACGAAAAACCGATCGTTCCCTTCGGCAGTGCTCA
CATACTTTTTGTTCGACGTACAACATTCGGTGAATTATTTGCTAATTCTTACAA
AAAGTTCAATAAATATCCTTATCATGGTATTTTCATGTTGAATGTACCGAACT
TGGTTATTAATGATCCTGAGATCATTCGTCTCGTGCTCAATAAGGATTTTGCT
AAGTTTCATGATCGCGGATTATATTTCAATGAGAAAGCCGATCCGCTTTCCGG
TCACTTGTTCGTCCTCGGTGGTGAGAAATGGAAGAGGCTGCGCGCCAAATTG
TCGCCAACTTTTACATCTGGAAAAATCAAACAAATGTTTCCTGTGGTCAGCGA
GTGCAGCCAATTGCTCGCAGAACACATTAATGAAAGCATGAAGACGGATAGT
GACTTTGAGATAAAAGATTTGCTAGGAAGATATTCCACTGACGTCATCGCGA
GCATAGCTTTTGGTTTGGATAGCAGATGCTTAAAAGATCCAAACAGTGATTTT
CGAAAATATGGCAAAAAAATTTTCACCGGACAGACAATCCGAACGTTCCTGT
TCATATTCACTCCGCGTCTTGTTGATCGGTTGGGAATTAGGCTTTCGCAGTCG
GATGTCAATGAATTCTTCATCAAGACGTTCACAGAAATGGTCCAGTACCGAA
AAGACAATAATATAGAACGAAGCGACATGCTGAATTTACTGATACAGTTGAT
GGACAAAGGTTACGTTGAATCTGATAACGTAAAGACTGATAAGTCAAATGGA
GAGGTCGACACGCGAAAACTCGGTTTACTCGAAGCGGCAGCACAGGCCTTTG
TATTTTTCATAGCTGGATTCGAAACATCATCTTCCACTGTTACATACTGTCTTT
ATGAATTAGCACTTAATCCAGAAATTCAAGAGAAAGTTAGACAAGAAATTGA
GACTACGATAGAAAAATACGGTGGCCTCACATACGAGTCTTTGACCGATATG
ACGTACCTGAGTCAAGTACTTGATGAAACTTTGAGGAAATATCCGACGGTTC
CTGTACTCAATCGACAATGCAACGAGGATTTCATGGTGCCCACGACAAGGCA
GATAATTCCGAAAGGATTGAATATAGTTATTCCAGTTCAGGGTCTTCATCAGA
ATCCGGAATTATACCCTGATCCAGAAAAGTTCGATCCTGATCGTTTTAGTAAG

GAAAATGCCAAAAAAATTCATCCGTTCGCCTACTTGCCTTTCGGAGAAGGAC
CCAGAGTGTGCATTGGAAAGAGATTTGGACTTATACAGTCGAAGATGGCCAT
AATTTCCCTATTGTCCCGCTTCCGATTGTCACCGTGTGCAAAAACAACAATTC
CTGTTGTACATTGTCGTAGAAATTTGGTGATGACTCCGCAAGATGGAGTCTAT
TTGCGATTCGATCCGCGATAA

>CcP450_28
TTGAAAACTAATTATAATAATTTTCGAATAGCAAATCCATCTCGAACCTTATA
TAATTATTTAGCCAATCGTCACCTAATACTTTCTGGCGATTTCAAAATCGAAG
AAGTAGACAAAGTCTCCACGTGCCTTCACGCGGTCTATGGTGACATAGTGAA
AATTGAAGGTCTTTTAGGAAGACCGGATATGGTATTTATTTATGACGCAAGC
GAAATTGAGAGAATTTTTCGTCAGGAGGAAAAAATGCCTCATCGTCCTTCCA
TGCCATCTTTGAATTACTACAAGCACGTATTGCGCAAAGAACATTTTGGTAAC
AATGCAGGAGTCATTGGCGTTCACGGTGAAAGCTGGTATAATTTTCGAAGTA
AAGTGCAACAAGTGATGCTGCAGCCACGATCAGCTAGGATGTACATTGGTGC
TATTGAGGAGGCCAGCCAATCGTTTCTTCGCAGGATAGACAGAATAAAAGAT
GAAAATGACGAAACGCCAGATGACTTTCTCAATGAAATTCATAAATGGTCCT
TAGAATCAATTGCAAATGTAGCGCTGGATGTCAGACTCGGTTGCCTAGACAA
CGACGCACCGTTGGAGACGCAAAAATTAATAGATGCCGTAAATGTATTCTTC
AAAAATGTGGGTGTGTTGGAATTGAAAATACCATTTTGGAAAATATTCAATA
CACCCACTTGGCAAAAATTTGTTAACGCTCTGGACGTCATAGTGGATATTACC
AGGAAGTATACGGACGCTGCGCTGACGAGAGAAAAAACTAGTGGGGATAGT
GAAAAAGAATTATCAATTTTGGAAAAAATTTTGGCAGTTGAGAACGACACGA
AAACGGCCTCTATCCTTGCTCTGGATTTATTTTTAGTTGGGATCGACACGACG
TCTAATGCTGTTGCATCTGTACTTTATCAATTGGCACTGCATCCCGAAAAACA
GGCAATTCTTTTCGACGAGGTGGCACGTACCCTTCCGCAACGGAATGAAAAA
TTGGAATCAAGACACTTGGACAATTTGAAGTACTTGAAAGCCTGCATAAAGG
AAACTTTACGAATGTATCCAGTAGTAATTGGAAATGGACGTTGTATGACAAA
AGATACGGTAATCAGTGGTTATCAAATTCCAAAAGGGGTGCAAGTAGTATTC
CAGCATTACGTTATCAGTAATCAGGACAAGTATTTCCCAAGGAGTAAAGAAT
TCCTTCCAGAACGTTGGTTGACCGGAAATGACGTTCTCCACGCGTTTGCCTCA
CTTCCGTTTGGCTACGGAAGACGAATGTGTCTTGGCAAACGTTTCGCGGATTT
AGAAATGGCTGTAGTAATTAGTAAGGTGAGTAACAATAGTCAATTAAACCCC
CCAACAAAATTCCTACTTTTCTGA

>CcP450_30
ATGGAAAACCCATGCGAGACTGTCGTGGATTCAGGAGCTTTGAGAGAAATGC
CTGAACCTAAAGGACTTCCGGTTCTCGGGACTATTGTAGATCTGATCATTTCC
GGTGGAGCTGCAAAACTTCACCTATATAACGACAAAAGACATGCAGAACTGG
GCCCAATCTATAGAGAACGTATCGGTTCCGTTAAGGCTGTGTTCATAAATTCT
CCAGACGAATTTCGCCGGGTCTTCCGCCTAGAAGGACCTAAACCGATCCACT
TTTTACCTGAAGCCTGGACACTGTATAATGAGATCCGAAAGTGTAAACGCGG
TCTATTCTTTATGGATGGAGACGAATGGCTGTACTTCCGCAGAATAGCAAAC
AAACTTCTCTTGAGACCGAAATCCGAGGAATTTATGTCAAGACCATGTCAGA

ATGCTGCGGATAGTCTTGCCAGCAGATGGGAAATTTATAGTGAAAATGGAAA
ATTGATCCCGGAATTGGAAGCGCAGCTTTATCAATGGTCAATTGAAGTGATG
CTGGCAACCCTTATGGGAAGAAATAAGTGGGAGGAGTATGGTCCAAATATTA
TGAGGAAGAGTGAAAATCTGGCGAAGAATTTACATAGGATATTTGAGTACTC
AGCAACGTTATCATTGATCCCAGCGAAACTTGCCATGAGAATTAAACTTCCG
ATATGGAAGAAATTCGTGGGATCCATAGACGAAGTCATGGAGATTGTAAGGA
TCTTGGTTCCTGAGATTATTCGACTCGATGGAGACGGTCTTACTAAAATGATA
TTGGACGAAGGAGTTCAGGGTGAAGATCTCGTGCGAATTGTGGCAGACTTGA
TTTTGGCTGCTGGTGATACGACGGCCTATTCCATGCAGTGGGCACTATTCCTC
CTAGCTAGTCACCCAGATGTACAAGAGAGACTGTACGAAAGCATAAAATCTC
TGGACCAGGATGAGATACCAAAAGATCCATTTCTAAAGGGTGTCATAAAAGA
ATCTCTCAGGCTGTACCCCACTGCTCCATTCCTCACTAGATTTTTACCAGAAG
ATAACATCATTGGCGGTTATCGCGTAGCTAAAGGGGAACTTCTGCTGCTTTCG
CTCTACTCCAGTGGACGCAACGATGCCAATTTTCCAGAAGCCAATGAATTCC
GCCCGGAAAGATGGATTAGAACGTCTACAGGAGAATACAAAGGTGTCATTAA
TCCTCATGCTACTTTACCATTTGCACTTGGTGCAAGATCCTGCGTTGGACGAA
AGCTAGCTGAATATCAAATATCTCTAACATTGGCGAAGATTGTTAAAACTTTT
CGTATGGATTGTACGAACAAGGACAGTGTCGAGATGATCTTGCATCTTGTTCC
TGTGCCTTCAGAACCAATCCAGTTGTTGCTTACAAAGAGAAGTTGGCCGTGA
AAAGAGATAAGGATCGTCACGATCTATTTACGTATATTCACGGATATAAATT
CTACATTGCGAGTGTTCATCGTTAAAATATCAAATACTATTAATTTCATCCTT
GACTCGATTAAAGTATTTAATTGGAAATCTTATGCATTTACAGTGCTATTATG
CAATCAAGAACCATTATCTGAGGGACAATAA

>CcP450_31
ATGGATACACAAATTATTCATTCTTCATCAGCAAACACGTTGGATTGCGATAA
CGCAGCTGCAACTGCAGAAATACTGGACGTGCCAAAAGTGCGAGTGACGCCT
GATTTGACAATGGAGAAGACTGCACTTCCGTTCGAAGAAATTCCAGGACCCG
CTATACTGAAGCTCTGGGAGAAATATTGGAAATACGTGCCTCTCCTTGGTACG
CAACTGTTTTGCAGTCTGCTAATCAACAGATTTACCCAGGGTCGATTAAACTG
GAATAGAAATATTACGCCATTGCAATATTTATTCAACGAGTATGGTCCTATTG
TTCGAATTCATGGACCAATTGCCGGAAATATTCTGAGGTGTTCAAACAGGAA
GAAGATTCACCTATACGAAGTGGTATCGATATTTTGCAACATTACCATTTGCA
CTATAGAAAATACAGACTGCCCGGACCATTCTCTATGTTAGCAAGGCGCAGA
ATGGACGGACGTAAAGGTAAAATTAGAAAAGCCATTCACTCAACAACTCTCG
CAATACTTCGACCGTCTCGAACTTACTAGCGACGAACTGGTACCTGGTCAGTT
TATTGAAGAATTGACGCGTTGGAGTATGGAGTGTTTTTCCATATTGATGTTCA
ACAGACGTCTTGGATTCCTTGAATCAACCGGATTAAACGTCGCTTCCGAACCT
GCAAGAATAATTGAAGCACTAACGACGGCCCACGTGTATCTGAGTCGTTGCG
AAACTGGTTTTCAAGTATGGCGATTCTTTGAAACTCCGTTTGCCAAGAAGCTC
TTTGGCGCGTGTGACGTTATCGATAGTGTTATTGGGAAGTATATTCGACAAGC
GCAAAATAAAGTTCGCCATCGTACCCTTACACCCCTCGATGATAAGGATACG
ATGGCTGAGAAGAGTTCACCGATTCTTGAAAAAATGTTATTCAACGAACGAA
TTCATCCTGATGACATTTCTACTCTTCTCATGGACATGATCGTTCTCGGAGTTC

AAGCTACTAGTAATTGTCAAGGTTTTCTCCTGTATTATCTCGCAAAAAATCCC
CGTGTACAAAGGAAATTATTCGAAGAGATATATCCCATTTTACCCAATAGAA
GTTCCACCTTGAATGAGGCAACTCTTAAGAGCTTGCCTTACCTTAAAGCTTGT
TTACAAGAGTGTCTCAGATTACGACCAGCCTTCCCTTACATAACGAGAGTTCT
TCCGAAAACTATCAATTTACATGGATATACAATTCCTAAAGGGACGTACCTG
ATAATGGCTAATCAAATAGCTGCGCGACGCGAAGAGAATTTTGAAGATCGAT
AA

APPENDIX B

CEPHUS CINCTUS OLFACTORY PROTEINS MASTER LIST

>CcCSP1
MTLIKIALLCLIVYTVAGQKAEEPRKPRVSDEQLNVALSDERYLRRQLKCALGEA
PCDPVGRRLKSLAPLVLRGSCPQCSPEETRQIKKVLSHIQRSFPKEWSKVVQQYA
GV

>CcCSP2
MKITLVCLIGFAAVVAVSAAPQYTTKYDNIDLKQILESDRLLNNYFNCLVNKGPC
TADGQELKKALPDAIETGCKSCNDKQKKGSDEVIRHIYKHKPEMWKVLTEMYD
PERIYIKKYETEAKDLGIAV

>CcCSP2b
MKTFIVFFAIFGVLLAEENKKYTTKYDNIDVDVVLKTERLLNNYIGCLLNENPCT
PDAAELKKNLPDALATDCTACSEAQKVASDKFSQYMIEERPDDWNRLENKYDP
SGAYKTRYLEEKSKKSKKPN

>CcCSP3
MRTQLLLVAVVVGVFALCQAQDISLLLNDRNYVEKQINCVVGKGSCDRIGQQIK
VLLPEVLNNQCSRCSPQQAQNARKLVDFMKQRYPNEWRIILKRFSGRQG

>CcCSP4
MKVAVIILVLVGCAIAAEKYTTKYDNIDLDQILKSDRLLNNYVNCLLEAGNCTPD
GKELKKSLPDALANDCSKCSAKQRDGSEKVIRFLVNKRPQIWEKLAKKYDPSGQ
YKIKFEGEAQKVGIKL

>CcCSP5
MRFIPVFLACLTVAFAQEYYSSKFDNINIKEIIDNTRLFAKYKECVLQEKATRCPQ
EALELKRVLPEALGTLCAKCTPSQVTKIREGLSYACKNRRVDYDEILRHVDPQGD
KIVAFEQKFGKVEC

>CcCSP6
MGERLLAFTLILCLVVIAVPKGTQGYLWPRHDTYTTRWDKVNVDDILESKRLLH
YYFNCLMNKGPCPPDGHELKRVLPEALQTACAKCTKSQIEGSVKVIRYLRQFEP
GKFKQLAERYDPEGMYRKRYLEQSLDNNTA

>CcCSP7
MIGVYVLVLAVIGCNAVYADEFYSNKYDNVNVNQILKSERLLQRYILCLLDKGS
CTSDGRFFKEILPEALATNCSKCSMKQREIVKTLTLHLMNNKPDHWREFVEKYD
PDNKYRTSFLNFIMSS

>CcGST1a
MPIDFYYLPLSPPCRAAMLLAKAVGVHLNLKTVDVLKGEQMSPEFLKINPQHTIP
TLNDNGFVLWESRCIMGYLVDKYAKDDSLYPKDPRARAIVDQRLYFDIGTLNES
LTKCYYPVLTGKTKVIDEENIKDLENAFETLDMFLDGRRFAASVNLTIADFSIVVT
VSTAESFGFDVGRYDNVAAWYENCKKALERFDYEEINAKSAKKFGEIYKSRLEA

>CcGST1b
MPIDLYHVPGSAPCRTVRLVAAAVGVDLNLKFLDLMNGEHLKPEFIKMNPQHTI
PTIDDNGFYLWESRAILGYLVDQYGKDDSLYPKDPKKRAMVNQRLFFDIGTLYQ
VIVDAIYPTIFGGAPKDSALSQKIDKPMEFLEIFLEGQNYVAGDKLTIADIALIVTV
SNFEVVNYDYNKFKNISRWFARVKAEIPKYEEINGEGLRAFKALAESLSKK

>CcGST1c
MNLDFYYAPTSSPCRAVMLTAEAIGVTLNLKSIDVMAGEHLNSKYEQINPQKTV
PCLVDGDLTLTESRAIMAYLVDQYGKNTRLYPKMVNSRALVNQRLYFDIGTFHK
AMAACYYPILFGKWKTYQSEHYDNLKKAFEILDKFLEGQDYVAGRSLTIADLSL
VASVTTAEAFGFDFAKFRNVTRWLKKVKTFAPGYRKANGEGAESMKKLIERVM
TQIDEL

>CcGST2
VYRSLFDARKCPMPSCAAGSGGRRRRSEPETHKPAAHSTRNRRQWLYLAESRAI
SSYLVDKYGKNDSLYPKDIKERAVVNQRLYFDIGTLYQAYTSAYRPVIFSGAPKE
AVLIKNFNESLKLFEQFLENDHYVAGENLTIADLSLVTTVSELEVLEHDFSKFENI
NRWYAKVKSEAPQYEKITEEGIGALKELLELLKK

>CcGST4
MGLKLYYDLMSQPSRVLYIFLKQCKIPFEEHKVVLKKGEHFMPKYEEINPFQKVP
ALEHNGFKLIESVAILRYLCREFKVDDHWYPISSKNQARVDEYLEWQHNNTRLH
CATYFLVKFLIPMIKGKPARPENVTKCEGHMIECLDQLENIWLKNKLFLTGDKISI
ADIVGACEVEQPRMAGFDPREGRPQLTAWLERVAEETNPHYDEAHKYVNQIAN
NYLGIPPHSLSKL

>CcGST5
MPKYKLIYFNLPVLGEPIKYLLHYGGIEFEDVQIKDRVNEWPKIKPTMPFGQVPV
LEIDGKRYPQSLAICRYLAKQLNLIGKTDLDQLEIDGLADTISEFRKLFSLYYREPD
PVIKAKKKESLENIEVPYFLDKFEEIVKNNNGYFHGGELSYVDFFFVGLVGAFNA
LLDKDVLANYTNLKSLKDKVQAIPAIQNYLKDAPKPIF

>CcGST6
MADQQHSYKLIYFNSRGRAEHIRYIFAFAGIDYVDERVPKERWPELKKCMPYGM
LPVLEIDGKAIAQSNAVARYLARQHGLAGLDEWESMQCDVLVDTLGDLKQGIS
QYRHEEDPFKKEEKKAKLWKETIPFYLNKFEQNVAENGGYSVGTTTTWTDFLFA
VALENFEQIFGSAALENYPALRGLKQRVHEIPAIAEWLGKRPQTEF

>CcGST7
MSTKHLGPGSVAPPITPGKLRLYSMRFCPYAQRIHLVLDAKQIPYDVVYVNLTRK
PEWLLEKNELGKVPCLEFESGEILYESLIIADYLDEAYPQNGLYPTDPMAKAKDR
LLIERFNTVITIMYKLYAETIVDRDLFDKALAGLEVLDRELAKRGTPFFGGSKPG

MLDLMIWPWCERADIIRVLRGDQYVIPRERLLRLLEWRTAMKEDPAVRGSYLDT
ETHAKYIRSRLAGSPQYDMLASA

>CcIR1
MQLLGSNVEPKSGHCENKNSTITRNNSTMELFNKNVIAKLNGNGTLFYIPDIQLIT
YKAEMELQTYGNGSLEKFADWNIEKKVTTLPGKTIEPAKRFFRVGTTQSIPWSNL
RRDPVTGESIKDSNGKEVWEGYCIDFISELAKGMNFDYEIVIPKSGEFGVKLPNG
KWDGLVGDLSRGETDIVVASLTMTSEREEVIDFVAPYFEQSGLSIVMRKPVRATS
LFKFMTVLRLEVWLSIVGALTLTGIMIWLLDKYSPYSARNNKRLYPYPCREFTLK
ESFWFALTSFTPQGGGEAPKALSSRTLVAAYWLFVVLMLATFTANLAAFLTVER
MQSPVQSLEQLARQSRINYTVVRNSTIHQYFINMKNAEDKLYTVWKEITLNSTSD
QIEYRVWDYPIKEQYGHILQAITQAGPVASAEEGFQKVIASENAEFAFIHDSSEIK
YEVTRNCNLTEVGEVFSEQPYAIAVQQGSHLQEEISKIILDLQKDRYFESLASKY
WNQSLKGLCPNSDDNEGITLESLGGVFIATLFGLALAMITLAGEVVYFRKRKAN
QRDQSRDKDKNKKQDERFYRTSDIHKSLGKLEIKPAPTVSFIGKPTGPRTRVSHIS
VYPRHLPFKD

>CcIR2
MKLLLLVIAIHCATFSRRGYEVAEAQQDPGGTSANRDVRPVNLFIVNDVTNDIAN
KSVINALKTLRDKNPGWIGEVSTVLINGSDPRDTLDRLCAAWDLALRAGSTKVP
DLVLDTTKSGLGAETVCSFTAALGLPTLSAQYGQQGDLQQWENLTPDQTKYLV
QVMPPPDLVPEVIRQLCLSMNISNAAILFDEGYVMDHKYKSLLQNVPTRHVIIEA
KQTTTELQSQLSRLRDLDIVNFFVLGDQETLSVALDAGESLNFTGRKYAWYGLT
LDDFSLQCACRNISVLFLKPKPAASQQVLSELTSRGLLPKPLLLSAFYYDLVRLGI
SAMRSALESRQWPREPSHISCDDYNVNNTPVRNFDFLSTLKVTTRNSDFVPTYGG
FAWGKSNGDHRANFSMTINLVIINNGNPISTEEIGAWQADVDGDLQVFEEDLLTN
HTALTSYRVVTVQQAPFIVYDAENDTWSGYCIDLLNEIREILMFEYEIREVDDKLF
GNMDEEGNWNGMIKELKDKRADIALGSLSVMAERENVVDFTVPYYDLVGITIL
MLKPKTPTSLFKFLTVLENDVWLCILAAYFFTSFLMWVFDRWSPYSYQNNREKY
KDDEEKREFNLKECLWFCMTSLTPQGGGEAPKNLSGRLVAATWWLFGFIIIASYT
ANLAAFLTVSRLDTPVESLDDLSKQYKIQYAPLQGSSAYTYFQRMADIEKKFYEI
WKDMSLNDSLSDVERAKLAVWDYPVSDKYTKMFQAMKEAGFPTDMEDALSR
VRRSKNSNDEFAFIGDATDIRYLTMTNCDLTMVGEEFSRKPYAIAVQQGSPLKD
QFNNAILVLLNRRKLEKLKEKWWNRNPLKKTCEKQDDQSDGISIQNIGGVFIVIF
VGIGLACLTLAFEYWWYRYRPRANNANHNIHQVPKGVPNVKPIRFNLQPAATR
AFESQPKFRARF

>CcIR3
MAVIIEENFFKDKSYYDWTSKEVYNFIFNGVKDNMKFSSIDFHIFHNVDISLKRD
YTVLLSIVTCDETWRLFKIAQKEHLVHLAITDSDCPRLPENGGISIPMMQPGEELP
QILLDLRMSKALNWKKINVLHDQILGKDTISRILMVFSREFEKNLELASRSVFAIQ
YKETDWARRQQLVNILSGFDVDQLGNCFLVFTTADMVAVIMEARSLKMVNTSS
QWLYVINDASERDVNITSFSELLEEGENIAFLYNITNKNKTCSINLDCYGKEIIRAL
AITLDNAYVKEIELYEQVTDEEFEIVRLTKAERRREIIENMIKELSLERTIYGGSCG

QCLSWKITATVTWGTSFISENETIGQLIDSGFWSPGIGLNMNDVIFPHVQHGFRGK
KLPMVSYHNPPWQIVSHTESGEVQYTGLIFNIIKHLSVKLNFTYTVVTTTNARDT
KIANHTRHVKMLQKTIDKVSASITNNVPDALIELVRTKKVLMAACVYTISEYRKT
LVNFTIPVSIQTYSLLASRPRQLSRALLFMSPYTKETWACLSAAIIIMGPILYLVHK
FSPCHMNVHVASGLNSPWKCTWYVYGALLQQGGMMLPSADSARLLVGTWWL
FVMVVIATYSGNLVAFLTFPRMDDSITSVDDLLLRDGQVTWGFPNGSFLENYLR
DANEVKYSRLLAGAERHNSSEDQDIIRRVRTGTHILIDWRSSLRFLMKRELLSTG
GCDYSLSSEEFMDEPIAMIVPQGSPYLSLINKEIKRMHETGLIHKWTTDWMPVKD
KCWGGPGINQEANNHKVNMSDMQGIFFVLIIGVILAVTMLTCEFCWYKRKVASE
RKLIRPFVP

>CcIR4
MKILILLAILVNILYYVNCGTITKNILRNNSESKALAGLLVHIINKYFDDCAIIIIYD
AHFENSEFLPHFLPKITQTFVHKSLNFSLVDKTQIAFNNKNYNYILFLENLSTIKYA
IDRETTNKIVIVTTATSWLVNDFLRSHTSKSYINLLIITQSARKKRNENAYLLYTH
KLYLDGVGTSKAILLTSWINNSLTKPHVLLFPDKLQDGFKGHRILISTAHQPPFAI
RKAIAGTDEIAWDGIDVRLIKLLGDILNFTADFRPPTSTTLSPIAAATNDIRMGLAS
IAIGGIYKTTNITSRFDTSFGHIEDCASFLSLASTALPKYRAVLGPFQITVWILLTCS
YLLAIVPLTMNSEYSLLSLLTHPSRTSDMFWFVFSTFTNCFTVKNPLLNYGLGKN
STAILIGIYWVFTIIVTSCYTSSIIAFITVPVYPTAMETVEQLLRYRYRIGTLDHDGW
EKWFEDIDDPMAQKLLKNIECIPSVEDGVRNVSRAYFWPYAFLGSKTLLEYIVQA
NFTPSWITKRSLMHISHECFMKFGVTFVLPRKSVYTEAFSSVIYRAQQTGLVEKIV
QDVKWDIQRTATGKLLQVQRSYATPDTG

>CcIR7
MISVLLGVLIINACLPVSSGMDGTLLKFLVDASILFAPTRMTAFLCIDQNDALHLL
RAMSNAHLSYDIRISNHNLDMRVSRENRHYVSFLLDLNCPEAITILRQANESALF
VGPLKWLILQDLKDRKKQNDSYLDSVFGNLAIFPDSELVLAQRLDKNSTRLLSPY
RPSPYHKLIIEEIGLWDLQNGVTLLNHDVSSRRRRNLANTPLKSCLVVTDPDTINH
LTDYQDKHSDTITKTNYPWIMHLVRMMNATVTFSIANTWGYRSKNGSWSGMIG
QLARGEIDIGGTATFFTAERIGVVQYISLYTPTRAKFIFRQPPLSFVSNVFTLPFRRS
VWIAIGVFMILIFCLLHLTLKWEWFNTSAEERVRRHWSGNVYQPTMSDNLLILV
GAVSQQGSSYEARAIPARIIILMLLVATLSLYASYTANIVALLQSTTDSITTLDDLL
HSPLKIGVHDIVYNRYYFESFQDPIRKTIYEQKVAPKGRKSAWMNLEEGVSRLRQ
GLFAFHGELGSVYKIVQETFNEDEKCGFQEIDYLNVLDPLIAIQKESPYLEIIRIGA
LKIHENGLQTREIRRLYTKKPVCHKHASFVSVGLTECYAAFITLSYGTLIAFTIFIL
EIIWYKRGIVEDVVQPFDYVK

>CcIR8
MDLRFLLFVFVCAIRCSDANGLSAQFLQDYFADKFIHQIVVFACYNHLERVSLSR
QLMLSDTKLSYISISENINMKTTLAVNYYKLGIVLDLDCTKSEIV
LDKFQKERLPYNESYFWMVISKDSMPPIEILRKLPLTVETEFTLALHHDEIYVLYD
VYNPSYRHGGRLNVTYMGVWNPLDGLSIKLTQYKYKRRADFQGL

HLNFSIALSNPPLPDLMTYISNPINRHLDTMHRFNYALLLQLRDYYNYTMILTQSE
TWGYLINGSFNGLLGDMIKGLVDVGVTPLQFKRERIDVAEFTVQ
TYLARPSFFLRHPKKTSVRNGFLKPFREEVWWIVVAVGFIYWLILWLTSKLELYY
RNRMPFSKINLEVGFETGFICMAAISQQGLTDSPRLYSGRIVFLS
LFVWALLLYQFYSASIVGSLLAEPPRYINDLKDLLESNLEIGIDDIAYNYDFFATTS
DPVALELYKRKVAPNKYRKKPAFYNHTEGLRKVKKGGFAFHAD
VNTAYKIITDTFSEDEICDLVEIQLFPAKHVATVTSKHSPFKKMVTYGMRQIMEH
GLSERLMHVWQYRRPICPKTHSSTPVPVALGEFLPATLLLLMGLI
IAVAVLSVENFCKRCKYPEQLSCNTYVIPGFKNQADINMYIQGHS

>CcIR9
MSETAGHYTYTNLYNWLILGESLDSVLPRLNDSTYSVVTDLVIATPILDGYVLHD
VYNPSKYHGGSLNVTELGDWHRSSGLNITLTHSKFNRRANYHGMILRVLTVTQF
KPNNTRLEDYLEDYSNKSIDLMSKFVYPLMQHLCDIFNFTMDLSTDFYWDEIAE
NGSHTGMMGMLNRGFADITGSPGAINQKRFPLGELVMPIWNFRTCFFFLSSPLRS
TNELLQPFSTGSWYATGALTILIALILSISLAFEKIDHWIRQYGIALFSTICILCQQGS
SVIPNFLTGRLVFLQLLFFSLLIFNYYSACIVSVRLNQPLDKINDSLNELAESNMDI
AAERIVYYDYLLETSTDWEVQMLKEKRWSKMSDAKKYLPLEEGLKRITRGGFA
YHCDPNRGYPLMEKLFTSEMICRLSEVHLYHPQGLGAWVGHHSPFREIARVGLA
KINDMGLRKRQAIRWSARKPICVNVQQTVNSVTFTEIKAAIFLLLLGVMLSLIICL
LENIIYNWNTKKRYERDN

>CcIR11
MCHFKMGGMRMRLLILKFYVFVCCACFETQEICKMVDLIAQRFLNDTCLVVFSN
NGCAFKSVIPALYFDPYDTLLENENMDSETYKTTYLTEILEIGCNGFVIQTKNPAR
QFVLLEETFVLVNNRGNRKLIFLPLLEKSSFKNLDSHSIVETRSANYVPNLIMVEP
SVTKNIFNLKTHYFGGTSNQSKEIIVLDTYNITSGKFLLDANLFPDKIKDLQGKILR
AAVFHYMPYVVLTGNDTQNNSQIDGTEWRMALQFAKFHNCTFERFVSDEESWG
TIFDNRTGNGVLGALVEDRADIGFVALYTWYHESLYLDLTTAYIRTGVTCLVPPP
HRLPLWRSPILPFSPELWAGVVTTLIICNVVSFLLHYGANYIFFEDNKIVSAIAALE
VAGIFLQQGLIKNRKELSWRTMAFSMLAVSVLLCTAYSSSLASAMTVPRYDAPI
NSAEDMANRDLKWAATHDAWIFSILQTDDPELKIVLKNFRIVKFNEIIKRSAEME
WGFSIERLSSGHFAVQEYLQQKYVEKMHLMASDIYWENCVFMLRKSSAYTTWF
SLRILQVLASGIPARWEDQVVAEYLDYGTQKAVSKETFVDQGPMKMTVAQVQG
TFFILFLGLLLALCVFVGEITLKRSVQI

>CcIR15
IPGTKEDLDTELSSLMKKQKIFMAKIAEKSINFKALVYVPPDVKTERWGSLRDND
SYTGLLGEAVSRNAVFFLGDLHYTTEHLHVLDLSVPYHTECLTFLTPEALNDNS
WKLLVLPFKLYTWIAVLFTLFAGGFIIHFFAIYYKNHVMLHKNHTIDNAAKNGAL
KELVAPKKSISKKAKFNQESRGLRIFMEIQNSIMYTYSMLMLVSLPRLPDAWFLRI
FIGWWWLYSILVSVAYRASMTATLANPVQRVTIDTLSQLANSPIAVGSWGEGRK
QFFLSSLDPNTQKIGTKFEVTTIEDEAISRVANGTFAYYENIYVLREARAKRQALE
VKRRKSASKLNRTFKQDRNLHIMQECVINMPIALGLDKNSPLKPNVDRLVTKKLI

VLPPHITLSVLKSLYDVFLQQVQRAIEAGLVQKWLNDVMEWAKNSEKPLDTNAP
KALINLHKLYGALVALGIGYSLGFITLVYEILHWKYIVLKDSKYDKYYLDVFYK

>CcIR17
MTGFELMLIEICSILQNVTSNNSCLPSQVKKEIPSRYDDQVFSAVAPYLSETRNNF
PTELRITTLQDPPLSGTKLVNGSLVGVGHAFYIFNLMQAKLNFTYKIILPERNILGD
ESSGVFGLLYTKQVDMAVAFLPVLSEAHNTVKYGTYLDEMEWTILMKRPQESA
TGSGLLAPFDTTVWILILISVIIVGPTIYILIYIRSKISQEKSEHYSVATCMWFVYGA
LLKQGSTISPVTDSARLLFATWWIFITILTSFYTANLTAFLTLSQFTLPIKSVDDLAS
KGQVWVAEKGLALDMIVAEGNELQPLNLSTSKGKGHFVDLRNKSYSTVIRHIDQ
NKMFLGEKHFVQGLIFYDYIDKSHKGIDEGKRCTYVIMPGHILLKRRAFAYPLGS
LINVVMDKHILALVESGIVKYITTAYLPGTEICPLNLGSKERQLRNSDLTMTYKV
VFIGFTVAAVAFLVELLYLCKCKKKKKKKW

>OBP1
MFKTLVTGIALLLASLHRVESAMTMEQMTKIASGFRNVCQPKTGVDMAILEGM
QKGEFPEDRKFQCYLKCVMGMLRTLRNGKVDLDMLLKQVDTMVPENLQERTR
HVARHCSTVATSNDMCEIAYQFVKCNWDTDAEVFFFP*

>OBP2
MRVEIKLNKHIRSTMRIFGVVAVACILQASIVHCGTRPSFVSDDMIAMAASVVNA
CQTETGVSTADIEAVRNGQWSEERSLKCYMYCLWEQFGLVDDKRELSLNGMLT
FFQRIPAFRAEVLKAVNTCKRIGKYFANGDNCEYAYTFNKCYAEQSPRFSVSYLL
YEKDGSTSFINDILNIINRQYSINRC*

>OBP3
MDYNNTSDTGNRMRNSTGNMNHRNSMNTRKMDNDQLHSYNNQKYDDRHDQ
ACVIQCFFNKLDLVDQKGFPERMAVTNIMSQNVHDPELRDFIEESVVDCFRYLD
MDMNKEKCQFSQNLLGCLADKGREKCDDWEDEQE

>OBP4
MRRFLFLVALAVSHEMIYADLPDWVTPEMTDMVKDDKIRCMNEHGTDEEMIVR
ANDGDLVDDRKLKCYMYCLMESFGLVDTDGEFEMELLIGFLPENLQDVARNAM
NACADEGGDDPCEKVYMHSLCIYKRNPQLWFLI*

>OBP5
MRVQVLPFCLLVMIYTTCFTLALKIPMTMQKHLKPCQTELNLTDSTLTNIEKNDE
KRELRSCLGACMLKREGVLNANGTFNKMKAINIIDPRVSKEDAYKAVDNCQNE
VSKNICETAELLYKCFTVYNVL

>OBP6
MLLYVCCVFFLLGFGNSHGVSDEMMEMAKMLHDSCVSETGVDEGLIEKCRDGV
FTEDSNLKCYIKCIMDQVSGMTDDGEVDEETVISMLPEEMQSETAPTIRECGTVR

GSDNCDTAFQTHKCYYAKNPEVKLLFGIMKSQLTPRYSHHRVTTLICNRFPWFL
LG*

>OBP7
MSLQKMVSSRRKPLTFITPRMVSPRNKLRSSPRSAAHRMDHRGESAFDQTTWTE
VKKYCKQQSAVPQKFSNYERIFSKSGYFPADEEFQCYLACLAVRSDLMILGNELS
IDWRKLKPLFPKGKAFGCQKRATEQDPCKKAYQYFNCLRKGNNAFQEIDFF

>OBP9
MIKRKMLRYGILFLSILFASSFDVSAAAISAIQSDLAVLRLCNASSPVSLQAVNSV
LINRDMKSSETQTQDFKCFLLCLYVEYGWMDREGGFSLHHIKLTLESSKLPEYRI
KKLIYSCTATEITDPCERAFNFTECFWSHSDKETSPEAEIPEKEVSDNQNGFYYIN*

>OBP10
MKTFLAVFALCIAGTLAALSEEQITKLKTYKESCIEETGVDPSVIEAAKKGTYSDG
DEKLACFSACLLKKIGVMDPDGKINEEVIRSKIPSSIPKDQADNIINTCKGLTGAN
ACETAGKVLKVLSRPKHVPLLMN*

>OBP12
MKALEGFSRILITFIVAVMVIVKADIRRDCRKQSGVSWAALKKLRAGDFNQEDH
SVKCYLKCFMVKNGIMSEDNHVDVDKALRHLPRKLQEPSRKILARCKDSAGKD
SCDKAFQIAKCYFKSQPGILKNVSFV*

>OBP13
MRAVLLVATLFLAFGSIESRMTLAQLRNTLKGMKKACTSKISVSAEVLDGASNGI
FPPDPALQCYYKCLFEMMKLMKNEKLLIDVMITQLDKIVALDVLDRMKAATLK
CVEAIDSEDPCVLSWQFTKCFYETDRSVVFFP

>CcCCE1
MRREFFLTFLGVAIVAMAGNVKGDEDPEVTAPIGKIRGSHMTSRLGKKIYAFRG
VRYAEPPVGQQRFQQAIAAKPWSDVFNASVEGPSCPQKYTQLVNEDCLRLNIYT
TKLPPKVGTKGKPVKRPVVVFFHPGGFYGGSGQGYIFGPQYWLDQDIVLVTTNY
RLASLGFLSTGDSLAPGNLGLKDQVEALRWIKKNIASFGGDPDSVTITGYSAGSW
SVTLHMISPMSKGLFHRAIAMSGAATVQEPLPTQQKHLAKKQAELLGCPTDTTG
NMLICLNTKTIEEFVDSYEKFFEWHRDPILVWSPVVEPEVNGVERFLPAQPVDLIR
QGKINEVPLIIGVTKDEFGGVVTSIIEEARKGNTSTFDDLNQNWDTIAPISFLYERG
TPRSRRISQELKKFYLNNQPVSLDNVDGLAQLYADAVIGFSAHRFVKLISAASAK
PVYYYRFSYQGRFSHSVWSDTKKPYGVVHHDDLLYLFYISFFPYFNATDPEVVT
VKRLTTMWTNFAKTGQPIPTNNEDFKNVKWPIYTDKTKEYLDIGDNLVTDCGLY
RNRMSFWDNLFPLKPHNFDDTGVKQ

>CcCCE2
MWFAKFIVCSLAISWTGATEEVQVEIPLGSVKGLKSTTVLSKPYYSFKGIPYVKPT
VGDNRFKPSQPAEPWSGVYDATEHKSSCVFYCMIRQGIMGDEDCLYLNVYTPD

VNKDAAKAVIVVIHGGSFNGGSGDDEIYGPDFLVENDVVLVTFNFRLGALGFLN
ADDASAPGNVGLKDQVMALKWVQENIENFGGSPRRVTLLGQNSGGAAAQYHV
LSPMSEGLFSRVVMQSGSVLNSWSFTYDHKELAFKLGELLGIRTSDSTELVERLK
EKSAKEIVDASGQLMKSLNALNGHMHAFVPSIEADVGQEIFLANVPWDLVKSGK
INDVPVIAGINADESMMFTKQMLRNADYINANFDKFVPDDLNVTDPSRLKEIGES
IRSFYLDGKSVSEDTVQEFSNMLSDIFFNYGLLISTKVMGSRVASPVYYYMFTYE
SPLGLMKNLKIFDVVRGVSHGDEILYEFYSSAFKNIHEKGSPADRVTNELTKLWT
NFAKDGNPTSVMDNYVTVDWEPMGKDDNYLNINKDLKMGKHLMKDRVHFWV
EIYKDVLGDYLKNFQ

>CcCCE3
MILKMIQASILAILLACLTCLVTAGPRVKIRNGTLEGTIMKSRRGVEFAAFRGIPY
ALPPLGELRFEPPRPAASWSGLRSAKEDAGICTQRNIYTHQKEVVGVEDCLYLNV
YTLELPPLEDGDNTLWNRPSYPVMIWFHGGGWVTGAGHSEFYGPKFLLDHDVIL
VTVNFRLGPLGFLSYEDLILPGNQGMKDQAQSIRWVSENIAAFGGDPNRVTLFGE
SAGGVSVHYHMMSPLSKGTIALDLFHRGISQSGTALCFWGLTRPGLAKKQAQRL
GRYLNCPNQDSKSLLSCLREKDAVDIIGTDKEFQEFDYCPMIPFRPVIEPDHPGAF
LKEDPAVSLKAGRIADIPWMTGITSHEGALRVAGLLGLNDGELARKLNDDFMTI
APMSLLYRDKCPADRIENVTKSIREFYFGDRPIDQSTKFDLIDLYSDAWFSIAADN
AVRDHLNVLASPVYYYYLAYRGSASFSSIFGDPKGEYGVCHADDLQYLFPVGEQ
LFKEFPLSVDDHKVVDIFTSLWYNFAKSGNPTPVLTKEIPIKWKPVRTLDLEYLHI
GSPKELFMADNLIHERVKFWEKLGIVGVDDSQRHFMESSQLRDEL

>CcODE4
MSAPIVRTRYGALKGTVVQNVEGGEYLAFNGIPYAEPPVGQLRFKEPQPPKAWS
GIRDAGKEGSAAIQFDMLRPTPEIIGSEDCLYLNVSTNSLSGKRAVMVWIHGGGF
VMGSGSSDMYSPDYLLKHDIVFVSINYRLHVLGFLNVDDAEAAGNQGLKDQVA
ALKWVQENIEAFGGDPQNVTIFGESAGSASVHFLTLSPLAKGLFHKAIGQSGSVL
NTWAYNRDNKESGLRLAALLGFETTDGKEAAQFLRTVDAKTLTEGVVELTGGF
GEFTVNVLFVPTMDDKSENPFAPQPADEMAKKGIEVPSIFGYNSHEGIMFLIGSTD
KTYSKVKNDFDGFFDKLIVNQNIKKTDDVVKSVKKYYFGKEPITPKQRDNYVEIL
GDLHFVYGIRRVVDIQSQKTVPMYLYKFSALCENFSMISLFKLNAKGTCHGDDIG
CLFYNRLFNTKFEPNSKARIRMERFTKLWTNFAKTGDPTPKLDDVITEKWLPVTK
QEVHYLEINDDLVPGVNPDEEKWQFWKSISESA

>CcCCE4b
MSAPIVRTRYGALKGTVVQNVEGGEYLAFNGIPYAEPPVGQLRFKVILCRDNKE
SGLQLAALLEFETTDGKEAAQFLRTVDAKTLTEGVVELTGGFGEFTVNVLFVPT
MDDKSENPFAPQPADEMAKKGIEVPSIFGYNSHEGIVFLVASTDKTYSKIENDFD
GFLDKLIVNQNITKTDDVVKRVRKFYFDEEPITPEQRDNYVQLLGDFQFVFGIRRI
VENQFQKRVPFFFYKFSGPGENVFMPSNVTNSIKGTCHGDELGCIFYNRASNSKL
EPDSKSRVTMEYFTRLWKNFAKTGDPTPKLDDVVTKKWLPVTNERIYYLEINED
LITGVNPDQEKWLLWKSISQSA

>CcCCE7
MGTRKPYVGPMVRRYVHTGVSRRSNVQSASLTVIGGTNDGGPKVCIDDGCLIGF
EDTSRLNRTFSSFLGIPYAKPPIGELRFKPPENPEPWDTREAKKDGNMCIAAMYFR
NEDDPPGSEDCLYLNVYTPVLPEIGVSSPDLPVMVFFHPGGFMFWSGTTYHFQPH
FFMDRGVILVTLNYRLNILGFLSTEDEVAPGNWGLKDQVHALQWVNDNIGAFG
GDPDKVTIFGGSAGGASVHYHMLSPLSRGLFTKAIAQSGLANARWAYSLKNEAR
EHAEYIGQHFNCPTENSTVLVDCLRKKPAEELIEADPIFFVWSIDPLIKFPPVVEPD
IEGAFLMEMPNDLLNSGKFADVPFMAGIMSKEGVLRSGAIMSHDNILQNYDVLK
LKSFSNSLHFYQHPVEDVIKIIDKINKYYFKGKRLTLKQENNMIHMFSDRYFVHP
LVKAVEMQRKHQKSPIYLYSFEYEGPKSFATVFGRELSTPLNGTAHLDDYLYLFP
MEKTFFKNDPWTRSSEEDEMIERMLDILTNFAIFGNPAPSFSGFTWNPVKSEDHE
YLVIKGPHDLKMSLKLRSDAVAFWD

>CcCCE8
MSAPIVKTKYGILKGTVVQNVEGGKYLAFNGIPYAEPPVGQLRFKEPQPPKAWS
GIRDAQKEGSAAIQLNILSSTNEITGSEDCLYLNVSTNSLSGKRAVIVWIHGGGFW
TGCGSNDFYGPDYLLKHDVVVVSINYRLHIFGFLNVDDKDAAGNQGLKDQVAA
LKWVQENIEVFGGDSRNVTICGQSAGSACVHYLTMSPLAKGLFHKAIGQSGCAL
NNWAYTTNNKEDVLHLVKILGLETTDTKKAVQLLQTVDAVTMLEALIKLVGGP
AMMFSKIIFKPSLDEKSENCFLPQPPNEMAMKGIEVPSIIGYDSHEGILFLREATDE
TLSAVEKDFDELFDRLIFNQNLKKTDDIIKIVRKYYFGDDPITLAQFQNYLRITGD
MRIVNGIRRLVGFQLEKRVPMYFYIFSAQVEKSWTKFWLKYSGEGTCHGEEIDCI
FYNRAFSSKLEPNSKNRVSMNRLTRLWTNFAKTGDPTPNLDNVITKKWLPVTKQ
EVNYLEIKEDLITGVNPDQDKWQFLKSIYDSA

>CcCCE9
MIFYYYKLMGMSLVFSIYLAKGSKEISENNSAVLDYDAKPQDSENNPVIQSSLGY
LKGSWLCSPSGKYYAAFKGIPYAKPPLKELRFEHPQPAEPWSGVREALTHGSQCP
QRLWILDITLGQEDCLYLNVYTPGLTDHKENATNATRKLLPIIFFIHGGTWTIGSG
DMDDYLYGPDFFMEEDVILVTINYRLGALGFLSLGNVIPGNMGLKDQAMALQW
VWDNADVLGGDKNRIIAMGQSAGGASAQYHSLQLHSRSLISGVIAQSGSILTPW
AFTKPGIQAERTRNLARAVGCLNGGSQAIINCLKNVDPFDIVKYQFSTDHDIKQL
PIGVTFTPTIEPAGTVNSFIKMGPKTLLTKYVSSKIPFLTGYTTAEGLLFLEIIVAMI
NPLAREAVLKRLGKEKEVFQHDQQLGSKRINISVDENNTRIESAKTLENRILKHR
LIEIAHRILGDYEFKMPILLGSKFHAMSGASVYLYQFAFTGKNNFLRLFRDDTKD
ATHMDELTYQFQGRGIFGSLPQLNNHTSTEYYLSKAIIWMWCNFARTGNPTPRG
SPIQWPANSRKILIFDNPDSLIVEDAPVMNNWFSLIDT

>CcCCE10a
MKNRGIELVVAIFGLALLAQNANGADSTIVKVDLPKGTLQGLKTRTVLNGVTMY
SFKGVRYAAPATGVHRFSVAREVAPWIGDYDATQHKSRCPQRCASAFKFIIGED
DCLFLNVYTSTLDSKAGLPVMFWLHGGSFNFGNGDSDVYGPDYLIENGVVLVSI
NYRLGPIGFLSTRDAAAPGNVGLKDQIAALRWVKRNIRYFGGDPNRVTIFGDAA
GAGSVQYHMISPLSAGLFAHAIAQSGTVMTTWAITYNSTKDSFALGAAFGINTN

NSTELVEGLLKINSTALVRMANKLSRETEGMSGGHYLFKPSVEVNVGQEIFLPAD
PWELLKTGRINDVPYMMGFNQDEMIIGANSIASVDYYNSHFEGFLPTDLNLTKG
AVLDTDIKLVRNFYFNGKNVTKNDIQAYIKLQSDLYFTYGTAFSLKMMRYYMT
KPIHQYLFSFDGKLGFFKKFFNVSLTSGVAHADETGYMFYPALLKITPDIGSEEEK
MLYKITRMWTDFAKYGDPTPKLSENVTTTWGDVTMNGYYLHINTTTKMEQNV
FAERVKFWANIYKDLLGQFYQYFK

>CcCCE10b
MNGYYLHINTTTKMEQNTVEVTILKGTLQGLETQTVLNGVTMYSFKGVRYAAP
ATGTRRFSVAEEVEAWDGVYDATQHGSRCTQICVPSLSSIIGSDDCLFLNVYTPN
LDDWASLPVMIWFHGGDFNFGSGNSDIYGPDYLVENDVVLVTVNYRLGPMGFL
STRDAAAPGNVGLKDQVAAMRWVQDNIAFFGGDPSLVTIFGNSAGAGAVQYH
MISPLSRGLFSRAIAESGPILSTWAISYNSTAYSFALGKLLGIHTTDSTELVNGLLE
VDSESIVATAIRLASDKDNMAGLNFIFRPTVEVDVGQDIFLPADPWVLLKNGQIN
DVPFMMGTNLNESITAVNDVSKAAFFNNNFDAFLPFDLNLTDSRTISDIVTKVKD
FYFGGHGVTDDDSLSYIELRDDLDYTHGTEMSLRIMAYYMDSPIYRYLFAYDGG
LGFFRKIFGISLTSGVSHADEVAYLFYQAAFGVTPAAGSTDEKMVYAMTRMWT
DFAKYGDPTPELSTNVTTAWEDMTTKGNFLNIDSTSTMSSKVFNERVQFWVSIY
KNILGDFACLFD

>CcCCE10c
MKNRGIKLVVAIFGLALLAQNANGADSTTVEVTIPKGTLQGLETQTALNGVTMY
SFKGVRYAAPATGTRKFSVAEEVEAWDGVYDATQHGSRCAQICVASYSYYIGS
DDCLFLNVYTPNLDDQASLPVMIWFHGGDYNFGSGDSDVYGPDYIVENGVVMV
SVNYRLGAVGFLNTKDAAAPGNVGLKDQVAAMRWVQDNIEYFGGDPGRVTIF
GDASGAGSVQYHMISPLSAGLFAHAIAQSGTVLTTWSISYTSTADAFALGRVLGI
NTTDSTELVNGLLNIDSKDIARAAFGMTTTKETMAGALFVFRPSVEVDVGQDIYL
PADPWQLLKTGQINDVPYVMGFNKDEAIIIATSISNAEFFNSNFDGFLPVELNLTA
SRDSEEISEDVTKLRNFYFGGNNITTNDVLPYIELQSDIYFTYGTAFSLKIMATYM
KSPIYNYLFTYDGQLGFFKKFFRINMTSGVAHADETGYMFYPAALGITPEIGSTEE
KMVYAMTRMWTDFAKYGNPTPCLSENVTTIWDEMTIDGNYLDINPDSF

>CcCCE10d
MDQNVFNDRVKLWASIFKNTVEVTIPKGTLQGLETQTVLNGVTMYSFKGVRYA
APATGTRKFSVAEEVEAWDGVYDATQHGSRCTQLCIPAFRPVIGSDDCLFVNVY
TPSLDDCASLPVMIWFHGGDFNFGSGNSDIYGPDYLVENGVVLVTVNYRLGPIGF
LSTRDAAAPGNVGLKDQVAAMQWVQENIVHFGGDPSLVTIFGDAAGGGSVQYH
MLSPLSSGLFKYAIAQSGTCLATWAISYNSTASSFALGKALGINATDSTELVNGLL
NVNSSTIVQTAFNLGTARDAMSGASFMFRPTVEVDVGQDIFLPADPWLLLKNGEI
NDVPYMMGFNLNESVISANGIRNAAFFNNNFDAFLPIDLNLTNKAEISKISKKVQ
NFYFHGNISTKDVLPYIQLKSDIYFTYGTEMSLRIMASYMVNPIYHYLFSYDGGL
GFFKKFLNVSLTSGVAHGDETGYLFYPAALGVTPAAGSTDEKMVYAMTRMWT
DFAKYGNPTPELSTNVTTAWEDMTTKGNFLKINSTSSMSSKVFNSRVKLWVSIY
KHILGNFAYLFD*

>CcOr1
MYLTIDSDILWLSKRVLSLAGIWPESPNNFRFFIYLLYLSLFNCAEFAALVLNLYW
MNFDKSVRNMTESIPTAMVILKTAMFRRNMQLLLPLLSEVRADKFSVEEEPGIA
WLYNIMGKLYTRLSVVLIFIVTTMLYAVPLSQWIVAKSNNLTSTYELPYQMYFGF
EINDLRSHVLACLSLLPMSTVLTIGCTGSDTLLVVLIFYLCRQFVLLSVRIRNVETD
PLIHPTKMKQLIERHANLIGMATALNKTYSSLLLVQTMGLSFVICIVAFELLTMAE
VGEETNTLSFIIYSLAVVTLLFSYCFLGECLIHESSSIHNACYFSNWYRLPPDLARPI
IIPIMRSRKPLHLTAGQFYVFSLETFTSHRIVHMSLDLQLHGIYRKDRLDPTGFYY
NIE

>CcOr3
MPEAMKVLTWNKWVLEFLGIWPSNESLFTFSFFFFLVSMATCFVYADLIYRISDF
KYVVENLTENIVLTLLCCKIGLYRLNRRLMKEILLDIKMDYAIELYNTEEQKSIFL
AYNRLSKSFIKYSVTTTTVATVLYYIQPLMDHANSHRKLTENSSITYILPYHMRM
SFNITESSLYYYVYAYEGVLVPIIACGYSGTDCLLVTLTLHLCAQISVLANQVENF
NGDFRKFHSHLKQVVTKHSRIISLSVKLRTAFAYFLLVQLVGATLVVCLAIYNLL
KNYATGHTAQLFGFIFYGSSVTVQLLGYSFIGERLMTESLKLSDAFYNCKWYVLS
PIYVNLVLFCILRAQVPLVLTAAGFCTFSLQSFTDVVKTSMTYLSVLRHFV

>CcOr4
MDFSSVNRFNGLVNFVCGNLLPLTDDATKFSKAQKLYSAFTWILEFTYFLVTTFG
IFFTTNQRIFQDTTVNQAVVIEILILGFYMNLRRGLIYRLIGQMNSVLINVETLKEC
VEKTVKPLQRPLKLYTIFATVTVVLFCGSPIYKVFKKDQFSYNDFRIPAYIPGEPY
STGLFIAGILFETLGGFYTILKKASIDIYLIHVITLLTAQYKYLSLELINIINGKNENI
NKSDLESLDNERNEKIEKTVQIELKKWIRHREVVMEIGKILKNLLVLNVVYVYLN
CIFRFCFLGFLLISSSGDYFIQALVCSYTVVCLMQVYVLCFCAQNLLDSSTAMTH
DAFYEKWYAYGPSTKRIFSMILISNKMECRLSMCGVVDLTLPTFMAILKNSYSAC
LFLLKVK

>CcOr5
MHHTMSNTREKDPDTSFLKYADFHINLLRKSGFYSMKGISNKINKEPTIWEVLLV
LTISTCGFFIIILEFRSVAVSLGSDTGFVIAVLSGTLTATLSMSKGLTILTSHREVRE
LLLRLSGFWEKSIERPENVDVMVQMANRASYLSKCYAATVVIMCSSYCMNPYV
SVITQFLFTKTANNSYNFTATTFPTVYPFDLSYFPKYVVWILFEQAVCLLMTLHW
IACDTLFPMCATHLAIQFQILRRDLERTTEVDELREIVKKQIILFQSCDILENIFSPIIF
LTIIMTSTIMCACIFQFEKTLSCGVYLEIIKYVTHMMSLFVEILIYCGFSNVLSDQTE
LLYHAAYNSEWTDRSKKYKSIIYFLILRSQKPFQCTAYHFFPVGLVQITTILTTAV
SYFTLLKTVTSESDDKVICS

>CcOr6/6b
MVNLVKGYMSTKISQLLMTLIGMKRGKTKREQLLMDALFVYILATVLSAIWLEN
SELFYSRNDLYALTYSAPCCFTVTFDFVKLMIFTYKRHELHELHKFTEDTYWNK
DYNELDKAILDKCDTTSAIGMSILALTSAILAFHYLTGPYLDNLGTNTTERTLPFR

VVFDFPITVTPLYQILYFIEVIGTISIGICSVAFASYLFYTCIFVSGFFKILQRELENVC
EVELESVNTKSSYNNNDTMLAYKKLKKCVIQHQLLIWYLGKLEGLFSYILLMLV
LCAVIILCFAGFQIILGDGTTKLHRQILSVEFIMAALAETVLFAFSCNEILTASAAIG
EAAYRCKWYKLPCDEYGRALRQGMTIMVMRSYKPCSLTVGKFCPMTLEVFTSI
LSTSLSYFTVLRSMNESE

>CcOr7
MFLLKYPMQYTSFCLKIICCWPISKNASNLRKNFDFLHFLFVISGFILISIQALVYAI
IQANNPQEIIETLSTQVIYLHAAGKVGIMRLHRHKFSRLLHGVERVQTTASGELLR
YHKEYTRVGYVAYWIYTISVFIITLDYIVQPLWKKSQFLPTGAWYPFDYKKSTFF
YTLAYFQQIICITFSGCASTTEITFGVFIFFACARLKVLQRKFQQLSDNSNGNEKIL
RRRICNYVQQHCDILRYINDVNETYTFIILVLFLSILLTVCCTSFLIINVTEHSLDSIFI
NSLLMVASAVQLLFYYLPGHILIEEAKTIAESVYYSGWESLSINCRKLLLQIMVYS
ASPINLRNGKMGLLILENYTTFLTTAASYLTSLRSIVGTV

>CcOr8
MDVPLHGGEEYDQFIKPVMNTMQIISIWPLAADCGSYEYLLRICHQILMFFVTGT
MSIVVTADVIHNWGNMDTATECSLIASAFVLCFLRLMVYTYHQKDMRYVVETM
RSDWADASYEEKQVLKEKCDFAFRLAKYFIATVAITIAFFMTVPMLETYILHSEE
KILPFRGYFFLNHTLSPNYEMIYIFEIIAGSFGGSMIAGVTSFNLVVIMHGAARFSL
LQKKLESLNRNDPDVNKLMVKCIKLHQDAIKFADALEDIINVVALGQFVTSTGL
VCFAGFQLTSMLEDRGRLMKYSTFLNSAILELFIFSFSGNELIVESEAVGDAAYRS
DWASSSFTQSLRILMMRATLPSRITAAKFYSMSLESFSAVLSTSFSYFTVLKAVSE
E

>CcOr9
MRTLPVCFVIFKFWGLWRPSDYLSWMKILYNIYSLLIVVIMYSFMLFEFIDVFINI
KSIDDFTMNSFMLVTFTNACWRVAKILAGRGKILKVINLLTSDYCVAVDKREVEI
KEGYDGIARWNTLRYVILVQITVALMLIVPLSNESRKRILPFRSWLPYDLSSLKLF
WLSYVHQCIAIMAAAYINVATDSFISGLMIQICSQLDILKYRLIKLPRLYNINHRDE
FLSSLTHCVLHHEHIFKFASAVESVSNPITIVQFCASAVALCSSVYQLSKHSVDSA
KFMPLILYLLCMLFQLFFYCWYGNEVILKSMEVRDAVYEMDWILLDNNNKKAL
LMMMKRSERSIQIKSGYFIALSVEPYIKILKASFTAYNALQQISN

>CcOr11
MTVHSLSPTGNCKGLKEGQASYSWCVRNLRTLMVFTGIWPMEPPTLLLNVSYY
YNATTFTLVICGMMAGAISVMDNYDLLVDNLSINLIFTEIFIKCILIKIYSKPLSRVL
SLMKFDWISLKGRGPIHGDVLQSENIMLLHANIPRIFFIAYTALALVAWTATVVA
AVSRKSSKMQIDASNAFPMPSWYPFEMHSTPNYEMLLTFQVIIGCSIAVSSAAVD
SLLVTAVFHVCGQLEILRKYFENLHSSETALEETEKKVAAAIKRHSKLIDLCDLIE
DCYSQITLSQLLVASLNVCLSGFGLLLAIESGNIKVFLKFLLLLIAMLQQILIYSVT
GDYLSSKSTAIRSAIYKMKWYELPPSLSKALMLIAIRAERPLVVTAGKFFPMSLEN
FTQVSSHCAKLQNRVGLVRKLVVTFR

>CcOr12
MEHKRGWAKIKSDIPNKNYDTDVEYAIKLNRWLLKPIGVWPLELSSSRTERIVTI
TSAVTCCLLMSFVLTIPCCIEMFSSQKDFKSRLEMLGPSSFCVMAVIKFFFFVIRGK
EIRFCIDSVVTDWRNVDVPEERNIMLRKAKSARFLTTVCALFMYYGGIFYTTYLP
LTTAKALSTENVTIRILPYRCNFILFDPYAPLFFDIIYFLQCLSAAFMFTLTSGVCSL
AVNFIIHACGQCQIIGLLLENLVDGRGNTSTTLEKRIAVVIVRHLHLLRFVTRVED
VLNEVCLVEFLGCTLNICLLGYYFITGLETADTARFVTFALLFISFTFNIFIFCYIGQ
LLTNHCHQIGEISYMIDWYRIPGAQARFLILLIAIANRPVTITAGKIVQLSFPCFRDV
LKAALAYLNMLRKMTT

>CcOr13
MDHEIEEFKIKTAVQNKNFDFDVEYAIKLNRWLLKPFGIWPLNSSSTKFDRGISVI
SSFICCLLLFFVMIPSFVVMFVSEKDFKGRLEIVGPTSFIIMVVLKYFFLITRGDTLK
MCIDTIISDWSNVQAKEERKIMFRNAKIARLFTIICVSFMYCGGIFYSIFLPLITAKS
LTNGNNLTIRILPYRCNFIIFDPYRRPIFDIVYVVHCFCSVIMYSITTGICSLAAKFV
MHACGQCEIVMSLLENLIDDDKQCSDIVESKLATIIVQHLHVIRFVTRVEDLLNEV
CLVEFLGCTMNMCLVGYYIITGFESADTVRFITFTLLFISFTFNIFVFCYIGQILTNH
CHQIGEASYMIDWYRFPGTQARFLILLIGVANRPIRLTAGKMVQFSFPCFCNVIKA
AMAYLNIIRTVTI

>CcOr14
MVEIELRKRILSTENKVQREQNKNFKNDTEYALKLNRWLLNPISLWPLSSHVSAL
KKIQFKFIRTSCFIFLAFVIIPACLHTILVQKDPKVKLKMIGPLGFCVMVIVKYFVF
VVRAKNIKICVDHVVADWRNVNTTEDREIMIDSAKIARIFTVGAALLMYGGGCS
YTMLLPLTASKSVSSSNVTIRVLPYPCYFFFFDPQVSPTYEIVFSAQILCNIVRYSTT
SGVCSLAVVFAMHVYGQCQVLMRLLEDFVNGDREKSSTLHQRLANVVVRHLRI
LRLISTMENIFNEIFLIEVVGCTFILCFLGYYCMTDLTESETVGLIAYFLLLISLTFN
VFAFCYIGELITNQCMQIGQAAYMTEWYRLEGKNASNLIALIIISNRPVFLTAGRM
INLSYNSFCQVIKSSLAYLNILRQVTM

>CcOr15
MTWNHWLLAMLGFWPEKPRDIWFWINFGYFAYHMTMEYVDLFLFIGNLEHVI
MNLTENMAFSQIFIRMLMMRVYNRQLGELIMEMRKDFQAHNYRSVDEQKIFLS
YNSKSKTFMKLLMAFVALTASSYYLKPILGNLGNDPIESESTNSTLTFELPYRFYL
LYNVNDTHTYMITYLSHLPFVFVSGFGQSAADCLMVTLVFHVCGQLSVLTLRISS
INSDPSQCSQNIKDVVVTHQRLLRMGQTIDKAFSAILLGHLVGATSLVCVLGYQI
LTNFAHGQNADLATFLTFAFLVLLVLYAHCTVGESLVQESTRVYEAWYDCNWY
NMPTENARLIILCMSRSQKPLCLTSGKFGIFCLSTLTDVLKTAMAYLSVLRSFL

>CcOr16
MASHANKYLSIKLVRFFMKLVGIWVPKDNNEQWIMNCALSYTILVILLAVAIEGF
DIYYCWGNFYATTYTACATMPVIIVLAKIFSFLLRRKLVMELIEFTETKFWQGDY
DSYGMSVLDDIDRKGVLLMCTFIFFVQGTVIGYVLTPIIENIGKNESDRVLPVTLW
INIPVTTTPYFEICFVMESLTIVHIGICFFCFDIFLCILNIHAAGQFKMLQHRFALVY

DGDNKQEMPITDKLHFMKDSEEIYIKFKDCVKHHKILIDYTEKVQSVFTFIILCQIL
ISSLMMTMAGFQALLTHGSIIRRLIFIAHTSGCFAQLLLFTSTCHEIIIESGGIADAA
YNANWANTPYDDVGRSLRIGLQLVMIRASRPCHLSAGGFCHVSLDTFTAVLSTA
ASYFTLLRQIGAETVMDT

>CcOr17
MEVLPLCFKLWTLSGVWRPMHYSSPVSKSLYTVYSLAVLITLYSLTFFEFLDIVF
NFGSLDDFANTAFLLLSMIIICCKATNTLKKRAQIIEIHDLLRAEICRGQNTTENLM
LEKFAKTCRSNTLNLFTMMMSCAMCLTLESLLYHVNDRLLPMKIWLPYSLSSLT
LFSLSFLYQVVTMTLANAITVANDTYITGLMIEICAQLEILKLRLVELHDPKTLQM
DENYSFNREKDILKMCIQHHNHIFKVKANVEKAFNSIFLIQIVTSTLVFCVTALTFL
KHEILSVELASVILYFFTMMFQLFLFCWYGNKVILKSTEVRSAVFEMDWIPLPQAI
KKDFIFLMMRTNIPIRFTSGYVVTLSLDSFMAILNYRTPH

>CcOr18
MNRNNFVRDNVMIYKCLGSWPLDCVARRLENTRYILGLTLIAVITVFNGLQYLD
LFIEWGDWGFISENISVSFIYSIFVSKIYVFYSRRREIMEMVLEIDNYVRKTYKDGH
QSNVQILQSCEDIARRIKIVFGWSALLTVIILHTWPIASMIFKKDPNLRMLGVPAYF
PFSLNSTNNYAIAYVVEIVTASVMSLHTVNFDTFFISCILFAIGRLRILHNSINNIKK
IAPSDLQMKFSKNETYQAELQKCIHQLLSEAIAEHQKIILFCQVIDSMLSWIMFIGL
FIYSVMLCFVGMRIIMIGPSIKLLQLMEYLAIMMTQIFLYYWHGNELKLESLKIHD
AAYNCDWYSFDRNSQQTIMFIMMRGQKPIYLTAGKFYYVTLETFLSGDPTHILL

>CcOr19
MGIANNIELLNTSVGRRWSSYQTVLEYKMEVECLKLAGLWYLSSDTSDCVKVIY
WIYNKIVFAAIAIFSIGLLTDIIINYKDLVTFTDSGCIFVGISVVSFKAVVFQCKKSRI
DRLVHSVRQCESWTSDLVDPAAMSLLKKYKIRERVTILGFSCLGCVLVIALIFFVP
RETGELPIRCWYPFDTTVTPMHQIVFGIQSFAVAVGMIAIIGMDNTVFVLCGRVLF
QLEILAANFQSCSLDINASGGRSETNPHKDFTCFRHTIDCGGFFERYRKCVMHHQ
YLILLVDEVNDVFGSSMFSQLLSSSLIICLTGFQSLLVVGHGTNFLKFAIYLGAAFS
QLLYWCCLGNELSHQSSLLLESQWKSGWEMQPAKKVLHPMVFSMMRSNKKLE
LKASNFFIMSTETFITILSKSYSIFALLNNMIQ

>CcOr20
MAFQVTLPRAMQFLLVCGKISCTLPPNFESRRYKSIIHEIMWLISMSVTMSLIIPLF
LAMYHFGEDLAVLTKSVSESMTATDPLLNMIFCKLERARIQVILSELMNFLNTAS
PFEKAKLQKHVDQCTPFYLFILILFSLTAVAFSCGPIVMDRPFPAEAWYPFSTDPL
PISCSIYVLQIMVIAQAAMCIHMDFMIAFLLWYAVASFEILGEKFRRVENEVDLR
NCIVQHQKLIVFVREINRVFYVMILKTPLSMTISIICSSVQLIHHEPLPVLSQFILLLV
VCSIKLYITAWPADSLIHASENIASSVYESSWIGKSSTFLKSLNIVIRRSQKPLIISIT
GILPPLSLRFYASFITKALSFITTLKSVISE

>CcOr21

MVEIELRKRILSTENKVQREQNKNFKNDTEYALKLNRWLLNPISLWPLSSHVSAL
KKIQFKFIRTSCFIFLAFVIIPACLHTILVQKDPKVKLKMIGPLGFCVMVIVKYFVF
VVRAKNIKICVDHVVADWRNVNTTEDREIMIDSAKIARIFTVGAALLMYGGGCS
YTMLLPLTASKSVSSSNVTIRVLPYPCYFFFFDPQVSPTYEIVFSAQILCNIVRYSTT
SGVCSLAVVFAMHVYGQCQVLMRLLEDFVNGDREKSSTLHQRLANVVVRHLRI
LRLISTMENIFNEIFLIEVVGCTFILCFLGYYCMTDLTESETVGLIAYFLLLISLTFN
VFAFCYIGELITNQCMQIGQAAYMTEWYRLEGKNASNLIALIIISNRPVFLTAGRM
INLSYNSFCQVIKSSLAYLNILRQVTM

>CcOr25
MDTHRNKYLSIKVTKFFMKIVGIWLPESKHEQFVLDVSLFVTIAGTVLSILFEMW
DIYKYPFNFNEAVYIICNILTPGIVLFKLSMIRLNRKSLYELIDICQTKFWHDDYDE
FGIAILQNCETKCVLLITSYMSFALFTAITYTVRSIIDNIGKTGTDKILPFTMWLNET
MARAPYFQLLFIFEGIILCYLGVGFFCIDNFFCIINIHVAGQFKILQGKLERLCGPND
REDEKKDRGIWIRKNPAAVFQEFRSCVQLHKMLIYYVEKVEAIFSLIILCQVILSSI
LMCLAGFQAVSDDNSASQRCIFTAYTIGCFFQLLLYTSTSNEIIDESLGVANAAYR
AHWYLLPFDKTSKVIRGCLVLVILRARRPCSLTAGRFFAISLETFTKVISTTISYFT
LLRQRGESS

>CcOr26
MSNLIEGYTSTKISQILITLIGMKRGKTKREQLLIDGLLVYIFATVLIAIFVENSDLI
YSRNDLYALTYNAPCSFAVTYDFVKLLIFTYKRRELYELHKFTEDTFWNKDYNE
LDKAILDKCDSTSAIGMCVLSVLATIVAIHYLTGPYWDAVGTNTTERTLPFRVVF
DLPLTVTPYYEISYVIEVIGAFSVGLCSVAFASYLFYTCTFVSGHFKILQRELENVC
EVELKILITKSSYSDNDAKLAYEKFKKCIVQHELLIGYLGKLESLFSYIFLMLVLCI
VIILCFSGFQFILGDGTSKLHRQILSAEYIVTTLVETGLFAFSCNEIFEASAAIGEAA
YRCKWYKLPCDENGRALRQGMTIMVMRSYKPCSLTVGKFCPMNLQVFSSVLST
SLSYFTVLRSMNESE

>CcOr26b
MPTIVEGYTSTKISQLLMTLIGMKRGSTKHEQLLINTLLVYIFTTVLIAISIESSDLF
YSRNDLYALTYNAPCSFTVTFDFVKLLIFTYKRHELHELHKFTEDTFWNKDYNEL
DKAILDKCDTTSAIGMSILALVSTILAFHYLTGPYWDALGTNTTERALPFPVVFNL
PLTVTPYYEILYGTEAIGAISIGICSVAFASYLFYTCIFVSGFFKILQRELENVCEVE
LESVNTKSSYNNNDTMLAYKKLKKCVIQHQLLIWYLDELEGLFSYILLMLVLCA
VIILCFSGFQIILGDGTTKLHRQILSVEFIMAALAETVLFAFSCNEILTASAAIGEAA
YRCKWYKLPCDEYGRALRQGMTIMVMRSYKPCSLTVGKFCPMTLEVFTSVLST
SLSYFTVLRSMNESE

>CcOr27
MTSKESEYRSVKFMRILMKIVGMYYTENPRERLLLRVALTYAIIAILFALAVEFV
DLYHCLGDFSAVMYNLCSTMPLVMVLVKISNFLFHWNVMMYLISFAQNNFWC
DPDDDFTRETMKRCDKYGKIFVYLFTNLVLFAVLDYIFAPVVENFHRNETDRILP
FTLWVNLPVTVTPYYEITYTIQSLSTLYTGICTCFFDNFISVLNIYVAGQLEILGHR

VETVADTCIDITMKDYVSEKSKLGLSLTLKKFKSCINQHQILISYIEQMERAFTLIL
LGQLILSSFVICVGGFQLMATMSFLRKFTFICHFMAGLTQLLLYTWSCNQISEKSL
YISQAAYNTRWYLLPYDETGKSLRNGILFLMLRAQRPCQLTAGKFSPITLQTLTAI
LSTAMSYFTMLRQMSEDDV

>CcOr29
MNDNNDIPKIVFKWNKLFLSIGGNWPLEPAYFRFSIWMSFFTLNISLLYVDLYEV
FGNLEQMLLNLSDSVIQSLILAKLLLFRFSEPLARLITDAQEDITAGEFGSLEEKKC
FLEYYQRGKIFYQITMSCVVFAVHVYFFKGLETYLFSVWNNETTTSFVLPYRTKL
FFNLTSPNTYIILYILEYPMVYIIACQGATICLLVTLIFHVCGQLAIVSYKIQHLRGN
LPEHENRIIKSLIEKHVKYVRMARSLDDTFQFVLLLEVAGTTVILGLTCYNIIANSG
DLADISTLCCFALYASSMILLLYGYCFVGECLIHESTKIHEACAQCTWYSMPLIYQ
KALIMCMLCAQRPLQLTAAKFYVFSLDSFSNVIKTSIAYVSMLRTVV

>CcOr30
MENQESFDEMIRVNLFLLKFSGIIRASVGKISLTEILSVLAFGLLAAMSCSYLRDLL
LTDNMEVAMQSLAFMITGFGNTLHYILIAKSRAKLCDVLATFEDLWGLLEVQEK
RVLMSYVRDAKKLTYFFMSQCAATVFLYVGAPVIFGNGFVRVNGNITERMLPYS
LIFECKDSPCYEILYILQILTVINIAITYIGVDTIGPVLILTVSGHMKIIQNRIMSLGSS
EEFENFKNDKVKRISDYEISHFQYFGKSFEACVKYHQTVLKLCKDIEKVTNKAFL
VQLITSTYSISAIGFKMVGNDSDKSKYVTQIVLSLVQLFLCNWPPDVLQNESQAV
AYAAYFMPWYRCSRDVKKSTEIIIMRGQRVVRLTAGNFVDLSLETFIRMVSSALS
FFTLLRSIE

>CcOr31
MVGIESCKRILSTEDKVQTEQNKNYKSDIEYALKLNRWLLNPISIWPLSSHVSTLT
KIQFKFIRTCCWVFLAFLIIPTRLHTIFAKTDRAVKLKMIGPLGFCVMVIVKYFVF
VVRAKNIKICVDHVVVDWRNVNTTEDREIMIDSARIARLFTAASALFMYGGGVF
YTSVLPLTASKSLSNDNVTIRVLHPFYFFTYDPQTTPMYELVFLVHTLSDVVMYS
TTSGVCSLAVVFAMHVDGQCQVLMRLLNDFVDGNREKSSTLHQRLANVVVRH
LRILRLISTMENIFNEIFLVEVVGCTFILCFLGYYCMTDLTESETFGLIAYFLLLISLT
FNVFAFCYIGELITSQCMRIGQAAYLTEWYRLGGKNASNLIILIIISNRPVLLTAGS
MMILSYNSFIQVSMKNNVIIFWLFPDTFQSEI

>CcOr32
MDISDATKVLKWNKRLLDVLGLWPLNLNEVKFSFFFIYITVQCFLQYADLAEYIY
DFNYVVRNLTETIVLSMIFLKIFIYRIISNELRELIQYIQEDYSEIVYSTANEIKTFLQ
YTLLSKRIVQCLLITCGFTTVLFYMQPLTTQLLAYNEGNSSTSFILPFHIRLFFDVT
QARIYYIIYACEIFVVPLVLCGFIGTDCLLITLVLHICGQISILTMQVGILLDDPRNL
REKLKRIVIRHCRLLRLFANLQSAYSAFLLQELFGITFLMCLGSYNVIATSAVTDS
SNFIRFLFYILTLIFQLFGMCYIGECVTTESSCLCNAFYNCEWYNISPDHAKSFLMC
ILRSQKPLTLTAGNFFTFSLVNFTSVVRTSIGYLSVMRKFL

>CcOr34

MIQQFQIEKFIYLNKWFLNIVGLYPQKSWRFVTSLFCMLLLVIPQFVQIYLFCTDL
SVILETSSVLFTILLAMLKGAVWIFNGKSMEDLVGFLFNEYWKIIESFDNSKTLMK
YAGYYICVYYSYAIYIFCFSWIFHWKDISGFRYASKMTKGYTFLIINALLFFFSLPP
IEILILKFNGTYKSSDKHFPFPATYPEFLKEFPYFQIVYFSQIIATMTCALVILATDT
LIATALFHTCGQFVVIQQKLELLSSNDGHETGEMIKRKIILIVKHHQMVIGFSNKM
ERVFSPMMFLQVFASSMIICLVGLQVSTTFTNQYKLVKYFSYLLMALFQLLLFCW
PGDQLLLQSGKVCQSAYFTKWYQFGKERRGEIQLMLLRSQKLIGITAGKFYLMS
LENFNVILSTSMSYFMVLRSFNSSEGK

>CcOr36
MSRRTDIGVSVKIIKFSMKLIGLWKADNTIDKVVTILIYTCLLTLMTIGVMLSLIDF
FYTLDDIYSAVNIFCAMIAILNLIGKLTIFGINRQKVFDIIDRIEDTIENGKKEEYGIN
AIIDCERKCVVLVIILGILTQGATANYVILPFVENFNGNTSVKSLPIQIHVTKLSHSE
SPYYEIGFVLEAAIGVSAAICSVTINIFLMTTNLHLVCQFEILHNKLKWTFSNDLDI
TDSNISRYAFENLKKCIKIHMMLIDYTAQIENACSYMILVQMLSSGLLMCTSGFQI
FFFSGTILKFVFGTSLFLTSIGEFFLVSWSCNEIILASAKVGDSAYNTTWYGLSATG
YGKAYRDGLQLLMMRSHRPCYLTAGKFCIVSLESFNAVITTATSYFALLRKFVD
DAELRN

>CcOr37
MTVSQVTEMSHHSNIGVSLAVIKYSMKLIGLWKADNTLDIFLMSSIFIYTISMGIA
GIVFTVTDLFYVFDDIYAAVNIICPTIAIFNNTVKLIIFAINRRQVLNIIERLENSIKN
ETYEEYDIRAVKDCERQCIILVITFVILTQGAASNYKFFPATEVEIYQSLSILVDSL
MANHLISKLDFVSSSALCLVTLNNFLTTTNLHLACQFKILHNKLKSTCIADLDMN
FNASRDAYSKLKKCIKTHKMLIDYTAQVEDVYTYIILMQMFASSVVICAAGFQLF
FFSGSMLRFVLSIFFFLTSVGEFFLFSWSCNEIIVASATVGDGGYNTTWYSLSATG
YGKAYRDGLQLLIMRSHRPCYLTAGKFCIISLESFSTVMTTATSYFTLLRNFVDED
EFRNK

>CcOr38
SYQEIFYRLCRDLVFLTYLQSERKKSVLPTTDYIFFLSSLIFSVWCHASFSLYIIGPF
LNHTEERPFPGHTKYPLDEDKYYWFIVSHQSICTFTVSAITAAVDALFVMLVHHI
CAKFAVLGYNLEKIGEDLSESPSAQDELETYKKIVNCVKKHKEAIGFADSMESAF
SIPTFINLTMNMIVMSTSGFMLANVHSIPNLVKYIGLSVTLTVHLFFMSWPAQEM
MDHSAEICEFAYFASWYRTSQRSKNLLKFLMMRSQVPCKLTTGKLYVMSLENF
CAVLKTSMSYFTVLKSLRD

>CcOr39
MSSLTESVYTKYEDKPEESRSDFENLDITLRWNRFSLRQLGIWPDPELFGGRSLR
KYKFLFPAFFMLAFTIIPQTTLLDMVWGDLDLVVDNLCVANLPMISALMKLLILR
FNSKALRPVLACITEDWKNVNNPGKREVLLKYGKIAKIISLVFTFLIYMTVISYVV
LQIYLNYQTKIINGTDTSRRFLFQSHFPYNTQVSPSYELTWLINLLTLDNILAILVL
NVCSQYKNLQIDLKNVLENMDTYKETFIKKLSTIVRKHNHLNRLAILLRNKVLY
KLVEIVLSFFCFKGVLIQLINHLTVCFSCKCLLLRFSFVFRAFKVFLSCFLFSCYIGD

KLHTEVRLFISNFQYYPNLEEREARESLLVMLRANSPMRITAWKFSSFSMVTFAA
VGNNWGTQRKRVKH

>CcOr40
MERKTLYKIAPNEYNKDYQSNVEYAIKLNRWLLKPIGIWPIRSSSKLSNIISFIYSG
LCCFIMIFMVVPTCIEIFTGEQSLKSKLEIFGPTNFCAMAIIKYCVFAFRSRKLKVCI
DFIIADWREVSNKEERVIMFRNVYIARSLTGVCVTFMFAGGIFYNILLPLLTSNSLS
SGNVTIRMLPYRGNYILFDAYVGPFYHMVYLMHCLCAIIMYSVTTVVCSLATKF
VMHTCGQCQIVISLLKNLIDGNQNCSRTLDDRLAVVVVRHLRVLRFASHVENILN
ELCLLEFVGCTINLCLLGYYFITEFENANTLGLITFFLLLVSFTFNIFIFCYVGELLT
DHCHQIGEASYMIDWYRLSGRQGRFLIMIIAIANRPINLTAGKMIQWSLMCFTNVI
KVAVAYLNIIRKITS

>CcOr41
MQALGLWPRTSNKLRLTIIAIYLLALTEWFAGQYAYYITKKLDWMPFIVKIKHTV
VAHMFIGKYLLAIIAGSQFGECLSALERDWHNVKLTYEIEIMRIYATKGRYYALIC
FILSFTTVTLQLIVPLTPLSESRAFPYPSHYWFDYQSSPYYEIVYTLKIFSSTILAAC
NVTVDGFIIILIMHVIGQLELVSEHIRNMEVMKSTIETCVQRHRVVIWMAETLENS
FNLVFLCHFMVSAIVMCLAGYSLMKSVERNESLESVISMFFIGSIAFHLFLYCYVS
DKLTEKSLQVGYSAYCCNWYQHECKQSKPLLILMIRSSRSLKLSAGKFIVLSLEN
FTDVLKTTGAYLSVLRKF

>CcOr42
MKFATDFSRLALEFFGFWPDTEYSQASKHLKLKCIITILFLLGFLYLPQSALLLMN
IGDLNLVIQILATGKIIISCAIIKIIILHIKREDLRRLLKSMNEDLVDYANEEIREIMLK
NAIFVRKMCISYSVTICITMLLYIVVKIFFYFAKISDNDTYPGLDLFVPSYLPNYLF
NSPSFELIYAGQIVAIMIAISAYSRSDSLLVMLIMHVSTQFTILRSSLKNLPVKVGT
TSSTFMCELGVIVRRHEHLNRFASKIDNIFNMILLLQIVLCSVLLCFQSFHFLVKLK
DKNAELSIIELVFLCFYICPVLIQLYIYCYFGEMLQSQNEETRRTAYECLWYILDLR
DARTMIIIMARTVRPVQLTAGKFSPVVMSTFTAILKTSMSYLSVLLATINDPKDN

>CcOr43
MKDTAIVGKPVEIGLRLIDSWPGASNRIGRLTVWSLMMAALIFQYWDAVSVFND
LDNLMDNFSVTITETLFFVKLIIVYKNRRYVNEVLKHMSEDWNSVKSVEEWTVM
TEHAKLSRIFYIWALGLYVGTVVLFLPVVINHYNAADINDRKFVLPSEYPFQSKV
SPVYEVLCCVQFLQAVLTAAGNALTESLLVTLVLHAGSRLVLLRKDISRFSDISRT
VDDRRKILLAGNVLVANHRRVIEFSDKIEDLFSYISLVQVLSSTLIICTIGFMFITSIS
STKNILTLMKFGLFILGELWETLAYCLAGEYLSNQSQSISQAVYECPWYKMQPRD
SKMLMMIMIRAQKPLRITAGKFIFLSLDNFTDILKTSLSYISVLRAIY

>CcOr44
MDHEIEQCKIKTTVQNKNVDSDVEYAIKLNRWLLKPFGIWPLNSSSTQFERLVSV
ISQFVCCFLLLFMMIPSIIEMFVSEKNFKARLDILAPTSFSVTVAIKYIVFMTRGRQ
LKTCIDTIINDWGNVREKQEREMMIRNARIARLFAIICVSFMYCGGIFYSIFLPLITA

KSLTSGNNLTIRILPYRSNYILFDPYVRPVFDIVYIAHCFCSFVMCSITTGICSLAAK
FVMHACGQCEIVMSLLENLIDDDKQCSDIVESKLATVILQHLHVIRFATRVEDLL
NEVCLVEFLGCTMNMCLIGYTILTGLESTDTTKFITYSVLFLSFTFNIFILCYIGQIL
TNHVLCHQIGEASYMIDWYRIRGIQARFLILLIGIANRPMRLTSGRMIQLSFPCFCN
VIKVAMTYLNVLRKVTA

>CcOr45
MTGKVILPKFDNYILLNKFGLRLLGLWPIPEDSAKWKFTLKKLHVSLVYVLLLSL
LVPQLLDLYILWGDVDANVENLCTSLTTFTVLAKLTNIVASRGVFQKSLATMKE
NWDTIMRHDNCPEEREILLKMSKVGFVFTRNYCIIMYITAGMYFLRPLVVGSGTK
SPHVKEYPFCAWYYYDQFSNLTYGIFYFSQVIIGFFCGTGNFTLDSLCLVMVYHA
CAQLRILQKQISQLTNDGSDDILIVERVRQLVKLHKKNIENARNLEAVFSGASAQ
QLLVSCVIICVIGLKLIVSLNDAFQLIMYVAYMQLVIFQIYLYCSPGDELINQSMEI
GRAAYMASWTNFPKYATHSLLMMTIRAQRPLRITAGRFYVMSIPNFTSILKTSAS
YLSVLRVLYD

>CcOr46
MDILPLCFTIFQFCGFWRPINYTSWSKFLYNCYTVFVTFCIYSFTLSQAINSITVLD
DINNLTNNTFMLVTMISVCCKIANILVKREAVINAINMFLENFEDLVEQKIKKKY
DLIAWSVTRNYSILVTITAICYIVTPLVINDIEKRVLPNRAWYPYDLSSLTLFWLSY
VHQSFALASAAFINVANDTFIPGLMIQNCAQLEILEHRVEKLIILTRSKKTENETKK
EYKEILEKCIRYHNRIIKFAKIIEKMFEPVIFVQFFVTVLTLCSSMYQLSRHSTLIVEI
INLIFYLISMLFQLFFYCHRGNELFLKSIEVGDNLYKMDWLSLRPSRRKDLLIIMIR
TNKPIQLTIGGIITLSFSAYISILKASYSVFNLLQQTSEK

>CcOr47
MSSLTESVYTKYEDKPEESRSDFENLDITLRWNRFSLRQLGIWPDPELFGGRSLR
KYKFLFPAFFMLAFTIIPQTTLLDMVWGDLDLVVDNLCVANLPMISALMKLLILR
FNSKALRPVLACITEDWKNVNNPGKREVLLKYGKIAKIISLVFTFLIYMTVISYVV
LQIYLNYQTKIINGTDTSRRFLFQSHFPYNTQVSPSYELTWLINLLTLDNILAILVL
NVCSQYKNLQIDLKNVLENMDTYKETFIKKLSTIVRKHNHLNRLAILLRNKVLY
KLVEIVLSFFCFKGVLIQLINHLTVCFSCKCLLLRFSFVFRAFKVFLSCFLFSCYIGD
KLHTEVRLFISNFQYYPNLEEREARESLLVMLRANSPMRITAWKFSSFSMVTFAA
VGNNWGTQRKRVKH

>CcOr48
MDISGATIVLKWNEWLLDFLGLWPLNLNNAKFSFFFIYIIIQCFLQYAALVDNIFD
LSYVVANLTETVVFCMIVLKLVIYRINMKRLHELIRIIKEDYSHELYKTAKERMIF
MKYNSLSRMIVQCFSILCVCAAVLFYIQPLICYLLAYRDSTGNSSSAFVLPYHIRL
FFNLTEARTYYIIYACEILIIPMSACGYVGPSCLLITLVLHICGQLSILATQVECMTY
DPKTIQQQLKQIVIKHSHLISLCATLNSTYSIFLLQEVIGITVLLCLGSYNITETGEFL
TFSCYVFTVFVQLLGFCYMGECLVNESINLCDAFYNYEWYNASAVHRKLLLMC
LIRSQRPLVLTAGKFFTFSLENFTSVTRNLLINFSPVLFFNQLALQTFLFAGYENIN
GLFVGTSKIHIKQS

>CcOr49
MDITSLSNHKLPKISSYYNVKLTLLVNKYFGYWPSLNVNQRFYAIYTFFSFSLTIG
AILLAELVYVIVYIRDFNKLTSISLLLFTNVVHVSKIQNLIRKRKQIRHLLEYVETG
QFSNETDKYERIVTYFSWFGLFHYAINVLFGMGAITGWSIISLFSVDNDKIDRRLP
SDGWFPYDPLKSPAFELTCVYEIIILHVCAFHNLALDHLTVGFMSVVCCQFVILNC
NLQSIGDDVLNAENIKNDLSNSLKNKENDDLIYKRLVKCVEDHMTIIQFAKDVEN
IFNLPVFLQFLESCLVICVTAYQISQDSNKGTSEILGNFSYLMCMVYQLFVYCWY
GNSILLLSESVVQAAFSGNWWKSSERYKRALRLMMVRAGRPLTITTGNLLTLSL
STFMGILKASYSFFTVLQSTNKKI

>CcOr50
MEIQNMGFYESRKKSVKPIKTNVHLQVSLSIIRYMGTWPPQGRYRNLYFVYSFLV
LNIIIRQFLFAQFGTIIMIWGDIEKMVSCGALLMTNIAHAFKVVVMLRHQKSIQILL
DSLQDELFTRNNERFKFIAYEYTWKGIFHHVAYQSFGTVAVLCWVFTSISDLIMK
HMKRLPIPLWYPFNVTNTPAFELITLQQTVGVISGCFHNVAMDTLITGLITVACCQ
FELLKTNIITIGSNVNHEFCKISHNEINVTNVEDNENKVWQEMRKCIAHNNKIIEFS
KEIQSLFGTAIFLQFLANCVIICLTTFNMSQTTVFVPSEILGTVAYTCCMVYQIFIY
CWHGNELWLQSESVLRMSFTSNWWEHNRRYKKALQMVMLRSSRPVILTAGKL
LKLSLETFVAFYECPIHFLPF

>CcOr51
MEILPLCFKFWTICGLWRPVRYYSSVSKFFYDSYSCVTVLVLFSFTLLEFVDIALN
FGNVDDFTSNAFLFLTMIGVCYKVAHTFMKRLQIICIRDMLYDEICRPENMEEQR
ILDNCANACRSNTVNLGIIIMSSIVFFLVGPLLHDVNDRVLPFRAWLPYSLSSLPVF
YLTYLHQSLAIISAGLVSVASDTFISGLMIQACGQLDILKCRLANLHESHNFEMAI
NDTFHREEAILKMCIRHHNHIFKFAALVEKTFNHVVLFQCFISAFVVCTTTLTLLN
HQLLSIEFITIIVYFLSIVFQLLVYCWYGNEVILKSIEVTDAVFEMDWIRLPASMKK
DLIMIMMRAERAIKFTSGYVLTLSLDSYVAILKLSYTTFNVLQRSS

>CcOr52
MEILPLCFKLWTIGGFWRPLDYYSPMSKFLYNCYSFVSFLGLYSLMLFEFLDIVFN
FGNLDDFTSIAFMFMTVLSDCLKAANTLGKRSRIIEIQDMLTDETCIAHDTEEQMI
LERSAKVCRSNTQYLGILIMSFVVFLFVGSLLFNVNERILPVKVWLPYSLSSLTSFS
LSYLYQIWVSMLAAFMATANDTYITGLMIQICAQLDILKYRLLSLPRSGDSERNA
NSSVSQKEIILKMCIRHHNHIIKVAATVEETFSSIIFFQCFVSTTVICVATFILLKSDF
MSVEFIITLMYFPPLLFELLIYCWYGNEVTLKSTEIRDAVFEMDWIPLPVSMKKDL
ILVILRAGRTIKFTSAYVLTLSLDSFMAVSIHTHYF

>CcOr53
MDYEYNPEEAFAINFGFLRVDGAWPTGITNPILRFIINEVSFIVVKSVICQNNYTVS
VMNCCRFLYRIYGFCVLSIFTVLYFSTEIILMTVIWGDVQAMVDCSFLFLTHLSHS
VKLTNFLLRQKRLDRMIFQLREPMFVPKRPEHIAIMKAAVLNAERETKMFLTITF
GTIICWSILPLLSQKEVKELPLMGWFPFDTTQSPAYGFTYTYQVISVLLNACVNA

MMDTMTSGLLILMGAQLDMLKENLQSLTDDASTMSQAEMKYIYNVRDVFQIGI
MAQFLASSIIICLTCFKYHQHL

>CcOr54
MDISGATKVLQWNNRLLRFLGLWPFDLNNVKFMFFFTYVTVQCFLQYAGLLEYI
SDLNYVVANLTETIIINMLIFKISIYRINTRQLRELIQNIEKDFSTELYNTADEMTIFL
KYNSLSKTIVQCISIVCGATAVLFYIRNSSSAFLMPYHIRLFFNLTEARTYYIVYAC
EIFMIPLVTCGYAGPSCLLITLVLHICGQVSILTCQVENLIKNPETIHHQLKQIVLKH
RRLINLCMNLNSTYATFLLQELIGITLLLCLGGYNIIATPILDETGHFLAFLLYTVT
VIFQLLGFCYIGECLRNESKSLCDAFYN

>CcOr56
MIIPGATKVLQWNHWLLGFIGLWPFDLNNSKFIFFFAYAMVHTFLQYGDLIEHIS
DLNHVVANLTETIIINMLIFKMSIYRINTRQLRELIQNIEKDFSTELYNTADEMTIFL
KYNSLSRTIVQCFSIMCLISPILFYIHPLLSHLLAYNDSMGNSSIAFVFPIHFRLFFNL
TEERTYYIIYACEILLVPTCTCGYNGPICLMITLVLHTCGQISILASQVKSMIHDPK
AVHQQLKQIVIKHRRVISLVANLQSAYSAILLPEVSGMTFVICLGSYNVITTSAVT
DSSKFLKFLFYILTLTFQLFSLCYIGECLITESTNLYNAFCNYEWYNVSPDHAKLL
VMCLLRSQRPLTLTTGKFFTFSLESFRIKTCGAKLIVKVNSNLIST

>CcOr58
MGNTDMKFVTNIHRLVLAPIGLWPDVENIKKDKVLKFKIISSILLVLVFVYLPQTI
LLVTRIDDFDLVIQILATGEIVCSCALIKIIILYFNRQALGKLLKYMEEDWMNSVN
ANEETKGILLKKGKFVRRFCISNILFSYLTFLLRVLVKIQFHITKTSECETHRGLEFF
FPSYFPKYMLKSPNFELIFIGQIVATVIVINIYIGCDNFLVLLIMHVSARSTILRALL
RDLPLKVHPKDSIKFMKELGIIVRRHEHLNRFASMIDDNFNIILLAQMTLLAALLC
FLCFQLIMKVQVQNAEVSTMEIVFLIFFILATLIPLYIYCYFGEILQHENEKIRNAA
YECLWYIREPIDARSVILIMDRTVRPMKLTAGKFRSFTMSTFTDILKTSMSYLSVI
LATMQEPVDN

>CcOr59
MDHEIEQFKIKTTVQNKNFDFDMEYAIKLNRWLLKPFGIWPLNSSSTKFDRGISVI
SPSICCLLLFFVMIPSFVAMFVSEKDFKGKLEIVGPTSFIIMVVLKYFFLITRGDTLK
MCIDTLISDWRNVHAKHTFLPVLTAKSLARGNKLTIRLVLFFSCNFILFDPYVRPIF
EIVYIAHCFCSVIVYSITTEISSLAAKFVIHACGQCEIVIPLQNIIDSDKRCLNVVEK
NLLLLFCDIYMRSGKIIILQVKGLLERPPFYGFLTILRMNIQYIFVTNVEDLLNQICF
VEFLVGTTVYMGKRKCTEAHYSLFLWTLLNSILHLYMKKCHQIEEPSYMIH

>CcOr60
MDISGAADILTWNKWVFSFLGIWPLDDSNKLFLFHYIYFMIHTCLEYVELFKYLN
NLEYVVANLTENLVITLTFLKISAYRINKNELRKMSNDINDDFKEESYKDSEEKA
LFLEYNAIAQKFMKVGIPITLIAAVLYYLRPLTGHVIVDTSRNTSHSYIMPYRITLF
FEITNFQMYVFMYAFEILNVPIIAFGFIGTDCLLITLVLHLCGQLAVLSHQVRNLT
MNPQNFQNEFRVIVMKHLRLIRISKSLDTAFNQILIQQLVGMSLLLCLAGYNLIA

MSDSGQNVHLITFIFYACSVSMLLFAYCLIGEYLINEIQDAFYNSTWYNRPPYQLE
LFLICMVFVQRPLVLTAGKIYSLSLSTFARVSK

>CcOrco
MMKFKQQGLVADLMPNIRHMQFSGHFMFNYYNDTGGSTKLFHTIYCSIHLFLIL
LQFGLCCVNLTLERADVDDLTANTITVLFFAHSIIKLAYFAVRSKLFYRTLGIWNN
PNSHPLFAESNARYHAIALTKMRRLLAAVGAATILTVCAWTGITFVGDSVKKVT
DPVTNETMTVEIPRLMLRSWYPYDASHGMAHVLTLIYQFYFLLITTMDANSLDV
LFCSWLLFACEQLQHLKQIMKPLMELSATLDTVVPHTNELFKAGSTDHLRDTQG
TQPMAPPPNENMLDMDLRGIYSNRQDFTATFRTAAGMNFNGGVGPNGLTKKQE
MLVRSAIKYWVERHKHIVRLVTAIGDAYGVALLFHMLITTVSLTLLAYQATKVN
TVDVYAATVIGYVLYTLGQVFLFCIFGNRLIEESSSVMEAAYSCHWYDGSEEAKT
FVQIVCQQCQKAMSISGAKFFTVSLDLFASVLGAVVTYFMVLVQLK

>CcP450_1
MFTDILLAIVVLLFVACILSGGKSKNYPPGPYSWPLVGNIYQLRKLACTHGGQHL
ALLELSRRYSSDVIALRLGRQNMIAVSGYEGIQTILNGQEYEGRPWNEFIKIRNMG
MRKGITMSDGPEWKEIRAWVVRSLRSVGFGRREMSVMIKDEVVHVVENLKVG
GVRSMKPVIAPAVINVLWTMATGKRLCEGPRLQYFMELMERRAKAFDMSGGIL
STFPWMRYLAPEFSGYNVLMVFNEEIKNFLMDIINEHKENYTPGNEDDLIDMFLG
EMYSGKGAKAGFTEDQLVMILMDLFIAGITTTTVTLDFLFLNMVMHVDKQKRL
QQEVDAIVGRERLPELSDRPNLPYTEAVLTESQRICMVTPVIGPRRVLSDTELLGY
TIPENTTVLINVYSVHMDPLHYPEPEKFKPERFIKDGVYVPDENLMLFGRGRRRC
PGEALARSALFLLFVGILQKFVLLPVPGKQTPTLDIVPGLTISPKPYEVLVVPRDN
CTNGTSKLID

>CcP450_2
MLVERAAQWAWQAMGDTRTDVFYTFLVFLVAFALVRGAQWLRYIRSLPPGPW
GLPVLGYLPFLKSDVHLQFGKLAKKYGSIFSARLGSQFVVVISDYKTIRDTFRREE
FTGRPHTEFSNILGGYGIINTEGALWKHQRRFIQEKLRSFGLTYVGAGKKIMQSKI
MREVDTFLQGLSLRRGAPTDMSGSLGMSISNVICSIIMGVRFHHGETRFKRFMDL
lEEGFKLFGSIVAVNFIPVMRYLPFHKTICNKISANHKEMANFFQQAVNQHRATF
DAGNVRDLVDTYLLEIEKAKVEGREATLFEGMNHDRQIQQILGDLFSAGMETVK
TTLEWAVILMLHHPEAARAVQEELDQVVGRSRIPGLEDQPFLPVTEATILEVLRR
SSIVPMGTTHATTRDVTLNGFTIPAGAQVVPLLHAVHMDPELWTEPEAFRPSRFL
SSEGQVTKPEYFMPFGVGRRMCLGDVLARMELFLFFSSLLHTFDLRLPEGASLPS
LRGNVGVTVTPDPFKVCLVQRTLEGDLADVECTDGNGPLRNVGSH

>CcP450_3
MFLPAILFLIILMLFFYLGSRKPKGLPPGPRWWPILGSTLEVIRLRKQTGYLYKTLS
VLSKKYGPVFAVKIGIDTIVFLNDYNSIRSMLTNEDCDGRPTGHFYKMRTQGKSQ
GKICVLVTDGRLWVEQRRFVLRHLREFGFGRTTMASLIEEEAQHLVNHFKRLLLS
ANNPVSESRQQIKSCSNNNGQIYQLISDTNINNTMMTHEKDNIESDTKKRKNEGM
TIEDAYVKADDYDEVRKISQSSGMIISMHDAFGIPVLNTLWRMMTGKRYNPDDA

ELKYLQKIATQLLKDVDMVGSTFSYFPILRYLAPEMSGYKSFVATHERIWRFLKK
ELENHKTNINPEKPKDLMDAYINILQTESYSKTFSGMLISKINIPIKDAITSIIFISEQ
QLLAICLDLFMAGSETTSKALSFGFLYIVLNPDVQRKAQEEIDLIIGHKQFPRLSDR
VRMPYNDAIVLESVRMFMGRTMGIPHRALRDTYITGYRIPKVNENTMIVPNFNG
VLMNNFWKDPEVFRPERFINSDGKISIPDQYLPFSFGKHRCMGEGLAKSNIFVITT
TLLQTFRFSVVPGDKKPSTEIVDGVTAGPKPFRVLVTPRI

>CcP450_5
KLSVHSGRQGVDPFETLHHCLGNLMNRIVFGKDYREDDVVWKWLRHLQEEGV
KHIGVAGPLNFLPILRLLPRFSKVFDYLIDGKRKTHQVYKQIIAEHRNQVDNGEET
MDSVLAAFDDEMRRRMNSDDPGHFTEPQFLHFLADLFGASTDTTLATIRWFLLF
MVAYPEEQLRIQQELDAVLRGKNVSLDDRSALPYLEAAIAETQRIRSVVPVGIPH
GTTEDTLIAGYEVPKGSMVVPLQWAIHMNPKYWQDPTKFKPERFIAQDGSLAKP
EAFLPFQTGKRMCLGEDLAKMMLFLFTGRILQRLNLTSPEDQQLDLTGECGITMT
PKPYRLIFTSRS

>CcP450_6
MDVGAFWILFTCATILAMIARNRLRSGPTEEEGDTVAWEKGKLRKEKDQADDIK
KDKRAGRMKKLLNGPRALPLFGSLHLLASPGGPFEAFTRLARKYGDIYEIRLGVA
TCVVVSSYSLVKEVLITKGSHFGGRPDFLRFHQLFGGDRNNSLALCDWSDLQMK
RRSLARSFCSPRRGSPQQEELSRVATLEASQFMSALTNEESQDVLRGERSLKPLLL
GAVANMFTRYMCSTRFSYADQEFCKIVRTFDEIFWDINQGYAVDFLPWLRPFYA
GHMKRLNNWATGIRTFILKRIIDRHRASLDSVNGVPRDFTDALLLHLESPGSGLS
WEHIIFELEDFLGGHSAIGNLVMLILANSVVHSEVQSKIQAECDTVLAQPGRSRAL
VTLEDRADMPYTESVIWETLRISSSPIVPHVATQDTEIGGYNVAKDTVVFVNNYE
LNLGESYWGPGSREFRPERFLKYSVPEAAGDSGEKPKVRRPEHFVPFSTGKRTCI
GQRLVQGFTFVLVTAIFSRFDISQASEDLASRLLPGCVAVPPDPFHLVLVPRAGKS
DD

>CcP450_7
MSCGLDTGYGVSMVLLKWDRTPAEVWIYLGILVFICLLYAFKDYIRILRFIFSLDG
PVALPVLGNANYVMRKDLLHIVSHNVYNDYGPIFRFWLTFLPYVTLLEPADIQC
VLSSAKHTEKIFFYRLLDNFLGKGLITRDVVTWKGHRKILQPAFHLHLLERFTKSF
SECATNLMNKIINEGGRDVNITSFVNDSVYDILNETVLGVTLASGKNKQLDSDDD
MPFRKGQVVAPYRITRPWLLIEWIYQLTSIGRAEQKQRDELFNACHKMMERTRQ
NIMKNGSTITDSSQNAVKKISLLEFMVEISGKHPEFTDEDIINECCTFMLAGQDSV
GTATAMTLFLLANHQDWQRNCFEELEEIFGKDERSPTMKDFRDMKCLEMCIKET
LRLFPPVPFFARTLGEDVVLGKQVIPAGCGVFIMPYATHRLAHHFPDPHTFNPQR
FDPEVSNKMHPYAFLPFSAGPRNCIGNKFAMLEMKSMISAIIRKCHLGPVTGKED
IRPKFRLTVRAEGGLWVKITPRSEFRYQPYMNL

>CcP450_8
LFQRLMAMRFLWGHQEGINKAWLGNHPYVFLSKATTVEPILGSNRHIDKSSDYR
FLQPWLGTGLLTGSGRKWHNRRKILTPTFHFKILEDFVEIFAEQTELLVNKVSCEI

GQDSFNIFPYVTLCALDIICETAMGRQVFAQYDSDSDYVRAVYEIGSIIQTRQATL
WYHPDCLFRLTSLYKKHQECIRILHGFSNKVIAERRREIKRTRNETVSEDEGRKRL
AFLDLLIEASQNGTSLSDEDIREEVDTFMFEGHDTTSSAVSWTLYLLGCHPNIQNE
VVAELADIFGENDNGRRATLRDLQSMKCLERCIKETLRLYPSVPLLARRISEDVRI
GKFLIPEGTTAMVILPMLHRDPAVFPNPDKFDPNRFLPENCTGRHPYAYIPFSAGP
RNCIGQKFALLEEKVIISGVLRKFKIEAAERREEISLTGELVIRAKNGLHVRVTRRE
R

>CcP450_9
MLLYVVLSLLILIIILIVPYHLIVHKSGNGKLINKIPGPKAYPILGNLLEFVVPFEEL
WHVVRGMNEKYYPIYRIWRFSHSVINIRHPDDIEAIIANSKHIEKSLIYRFLQPWF
GTGLLTSTGRKWQQRRRILTPAFHFNILTEYSKTFIEHAEGLVNSLKAEGEETVKD
IVPLITTVTLSSICETAMGTPVQEDEDFQTKFQFAVHKMGAIVVSRVFHPWLFND
WLYKFTPTFRLQTKVLDILHGFSLKIINIRQQYHDATGGKYLNQFSTDKKNIDGN
EVKDSQRKRLSMLDLLIGVSRNTGQIDLAGIREEVDTFIFEGHDTTSMAICFIIQLL
AEHKDIQEKARKEIDEILKQNDGEVTITEIQQFAYLDRCIKESLRLYPSVPFMSRKL
SEDLQLKNYLLPAGTSVHIHTFDLHRDPNFWPDPLVFDPDRFLPEKIQGRHPFCY
LPFSGGPRNCIGQKFAMLELKAIIGLLLHNFYLEAIDRPSETRFLPDLILRPAHPIRV
RFVTRRK

>CcP450_10
MSSAAAPEVVPGSVAAVAAAGLSASTVFFALLIPALVLYYVYFRISRRHLLELAE
KIPGPQGYPVIGNALDFIGSSDTIFKNIYQKSFEFDQVIKIWIGPKLVIFLMDPRDVE
IILSSHVYIDKSDEYRFFKPWLGNGLLISTGPKWRAHRKLIAPTFHLNVLKSFIDLF
NANSRAVVEKMRREQGKEFDCHDYMSELTVEILLETAMGVSKTTQERSGFEYA
MAVMKMCDILHLRHTKVWLRPDWLFNLTKYGKDQIQLLEIIHGLTKRVIKRKKE
EYKTGKRSVINPAVATIQSEMKAAAKGKESKNSSTTVVEGLSFGQSAGLKDDLD
VDDNDVGEKKRLAFLDLLIEAGQNGVVLTDEEVKEQVDTIMFEGHDTTAAGSSF
FLSMMGTHPDIQERVIQELDEIFGDSDRPATFQDTLEMKYLERCLMETLRMYPPV
PIIARELKTDLKLASGDYTVPAGATVVITTFKMHRQPNIYPNPDNFDPDNFLPEKT
ANRHYYAFVPFSAGPRSCVGRKYAMLKLKIVLSTILRNFRIRSDIKESDFRLQADII
LKRAEGFKIRLEPRKPGIKA

>CcP450_11
MKDNLSNKILSMLYAGKKWYTRRKLITPSFHSGLLEQYLKNSIREADILVTKLRA
EVNKNGFNIVPYAKLAALDVICVSAMGYHINAQSNCTNEYVLAVDKAARITQER
FVNIWISLDVIFKRTTAGKEFKKALEIMDRFTEKVITGRKAEWRSKRDGNFNEFP
KKRKALLDMLLEMSDDGKILSDEDIREEVNTFMFAGHDTVATSVSWILYALGRF
PKYQNLILEEYDRIIGSNDITLEDLNKLDWLEACIKESWRLYPVTPLIARQISDPLR
LRDTDIPSGSTILINSYILNRDPRHYAQPEMYYPERFLPDKPKPPPFAFIPFSAGSRN
CIGAKFALIEVKVMILRILNAYNIHSLDSEDKLRLTAELVLANTGGLRLAITPRHSP
A

>CcP450_12

MAVSVTGLILCSAILCGLIWIILYHVKRLRMYKAAAKYPGPVAWPFLGNAYYFIG
KNEDILNKIIKLMNTYPSPLRFWLGQRLFFAVYDAEEMKTVFLSPKTIEKEDLYKF
ARPWLGTGLFTAPASKWRVHRKLILPTFNQKILESFVEVFSAQSQIMVKNMEVE
VDGEEFDIFNYISLCTLDIICETAMGVAVKAQTERNSRYVESAKRAFEVIFTRMFK
IWLHPDIIFNHTQLGADQRTCVEYLHSLTNDVIQKKKQTFHPVNGKIANSEVVEN
QPNQRKAFLDLLMELSNNGTKFTDEELREEVDTMMIAGNDTTATVNCFALLMM
ANFPEIQDRVYQELYDIYGMDDPDDSPIKHEDLHRMEYLERVIKETLRLFPVGPIL
VRRVIDDLNIGTHTLPKGSSVVLGIFHVHRSEKYWPDPLKFDPDRFLPEEVAKRH
PFAFTPFSAGPRNCLGLKYAMMAMKALLGTVLRRYIIKKDHMTPISEMKLKADV
MLKPVEPIKFRIERRIPKIGCID

>CcP450_13
MSWSYTRRLIYGSLITPWKYFDFHMISVTLIILIGVWLAKVILEWRNRRIRLIEFSK
KLPGPPTLPLIGNALAFACSSTETLNVVTSLCNYGPIFRVWMGPKLFVAITDPRDY
EVILTSPKASHKGAFYRFMKPFIGEGLVSGSGPTHRTHRKIIMPMLNVRVLEVFIQ
YFNEHSKYCANQLEEMVDTGEFDILPFTAHCAADITLETVMGIPGTVQRGGYKQ
LMHWAERMYEVIHTRIMKVWLHPEWIYSWTEYRAQEKIGQNVIQGFTESAIERK
KKEHYALERGALISNRPRTMILEQLITYVENTHVMNDEELRDEIYTVFTAAQDTT
GLVSAFALLLLGMHSDTQDKVRAELKDVVGTKDITMEMIPELKYLEMVIKETLR
LFPIAPMFVRELTGDVELESCTLPKGCSVAMVPFVTHRSPKYWMEPEKFIPERFLP
ENSQNRHHYSFVPFSAGTRSCIGQKYAIMSMKIIIGTIVRKYQITAKGSMETIRLKT
DISVRSVDGYKVSITRTT

>CcP450_14
MAALEILCSVAVVFILFYYYVTRNYNFWKQRHVAGPEPMFPFGNFKEITFGQLPI
GILIKQYYDQFKGEPMVGIFVRSLPTLIVRDPELVKDVLIKDFTSFPNRGIHYNTK
VDPLSENLFMLKETKWRPLRRKLTPVFTSGKLKQMFYLLLECGDVFEKYLDHLS
EKNEPIECRDITAKFTTEVIGTCAFGLKTDAIADDNSEFRKMGIRVFSTSFMKTLQ
FILRQSMPWLYNILGKCMADWVVIDFFTKSIKDTIEHRKQNNIVRHDFIDLLIDMK
EHSDNMSDVDVTDNLIAAQAFIFFVAGFETSSTTMSNALYELALNQDIQDKLRQE
IKEEIMKNNGELTYEAVKDMKYLHMVFQETLRKYPPATVMFRVVNQPYTFRGT
NVTIPKDHRVIIPIYGLHHDPEVYPSPETFDPERFTAENVKARHPMYFIPFSDGPRN
CIGERFAVYQTKVGLIKVLINHKVDVCEKTKIPYEFNTRAFILTPKDGIYLKFSKLS
A

>CcP450_15
MQVERYRKNFAIERRHFCTRMTVIFEQWVNYWTIGLLTVIIVIVYVYVKRHHSYF
QKMGIPYIESHILVGNRGPVVFKQRSLFDHSQVLYNSIPDAKYVGIYDFAIPAILIR
DVELIKDVGVKHSENFLNHKILSDPQLDPLFANNLFNLQNDKWRKMRALLTPAF
TSSKMKMMFDMITKCATDFADFIEKQSPIVVDVKDGFARYTNDVIAGAAFGITV
NSMKERNNEFYLMGKEATTSGFVELMKMILYQASPTLAKKLKANVVSHRAVNY
FCNIIESTIATRKTQGIQRSDMIQLLMEAKDSGDKLDLTIEDITAQAFIFFFAGFDTS
SSTMSFAAYSLATNPEIQKKVQAEVDDILKKCNGKVTYEHLKEMKYLEAVVSET
LRLYPAAAEIDRVCSKTFELPPARPAARSVTVKPGTVILFPVYAIHRDPQYHPNPD

KFDPDRFLNENEKINPLAYFPFGLGPRYCIGTRFALMEVKLVLFYLLMQCNFEPC
EKTTIPMKFSTANFLPRPLNGHWLKIVPRKGIKRL

>CcP450_16
MVLLIQSATSRLGHRLSRSCRSLCSEALNATGHVTSSISGDLDRAQSYDLVPGPKP
IPILGNTWRLLPVIGQYQISDVAKISQMFYDEYGKIVRLTGLIGRPDLLFVYDADEI
EKIYRQEGPTPFRPSMPCLVHYKSEVRKDFFGKLPGVVGVHGEPWREFRTRVQK
PVLQPQTVRKYIGPIEVVTDDFIQRIEKIKRDDCELPADFDNEIHKWALECIGRVA
LDVRLGCLGDKLPLDSEPQKIIDAAKYALRNVAVLELKAPFWRYIPTPLWTRYV
KNMNYFIEICMKHIDVAMERLKTKKALSELDLSLIERILAKESDPKMAYILALDLI
LVGIDTISMAVCSILYQLAIRPEEQEKIYQELKMILPDPTVSLTTKHLDQAVYTKA
FVREVFRVYSTVIGNGRTLQHDTVICGYKIPKGVQVVFPTLVTGNMEEYVTDSKT
FKPSRWIKNDNNEKLHPFASLPYGYGARMCLGRRFADLEIQVLLAKLIRSYKLEY
HHQPLKYKVTFMYAPDGELKFKLTKR

>CcP450_17
MSYQRAMLVRSALLAGAKGNVLNIGLIERYASTVLQADVKSPEETEWNEAKSFD
EIPGPRALPIIGNMFRFLPHIGEFGNQPVMEQMRNLKKQYGPIVKLDGIPHRRTVV
FLFDPELSEKMYRTEGRWPTRIAMETMALYREERPELYKGNYGLASSQGKDWH
NFRSKVNQHMMQPRTIKPHVGQINEVAQDFIEKMRKLRDPETLELPKSFNNEMN
KWALESICAIALDQRLGCLEFDLPTDSEPQIMINAVHDMFDLFYKLEILPSLWKV
YRTRNLKKLFHCFDIMSGISKRYISLAREKLRESTSMGNEELSMLQRLVAIDEQT
ANVMIMDMMLAGIDTTSNVAASALYHIATNPQAQEKLHQEAAKVLPEKDSPVT
YEALNKIPYTKACIKEALRMSPIAIGNLRTMQKDIVLEGYKIPKPHWKSLLKNKS
VISRPFCYITRIRVSAIWRQTSYGGQFEIDSPCHLVVMYRRMLSVIVMTIICLQF

>CcP450_18
MNYILELSFGAAIILVLLIRAKIKKYNYFKERGVPTLSIYAGLTNTWKLFTKKISFP
DMVMDLYNAHPNAKYIGIFDQGSPVFIIRDPEVVKLVTVKNFTVFLDHKGFVKE
EVDPLLGGGLFNMKGGKWKEMRAVISPTFTTSKMKCMFELMSKCGQDFVEFLE
KEKQPLVLEMKDVFTRYANDVIATTAFGVTVNSIHDKENEFYMMGVEATSFTPL
RSFKTILFQWYPKIMKFLKFKIVSDKVSNFFRDIVITTINTRKEQNIHRPDLLNILM
HARMKDPDRELTHDEIICQAFLFFFAGFETVASVMSFVAYHLAIYPEVQEKLRKE
VDGVCGRENGQLIYEELGEMKYMEMVLNETMRLIPPSVLTDRVAVEDFHLPAN
DEKSPPIHIEKGTAIWIPIAGFQKDPKYFAEPEKFDPERFDSDHKNNILPFTFIPFGV
GPRSCPANRFAFMEIKLILINILKRFVIKPCEKTSIPMKFAKGTITVVAEKGIWLSFK
PRKNIE

>CcP450_19
MRERVHPAQYESYVASETCRSVLAMEYLWHALQCVLVIVGLTLAALIYRQYTY
WKKRGIPYMEVIPLIGNQAPIFFKLKTFPELIMDMYNRYPNARFIGMMDFNTPLV
FIRDPQLIRDIGVKNFDNFPDHMSFVDEKVDKLFGKNIFSLRGERWKEMRATLSP
SFTASKMKFMFELISKTSKDFVTYLLEHPEETKSIETKDAFTRYTNDVIASAAFGIS
VNSMREKDNDFYIHGKEATNFEGFWKIMKFVMFRVCPSILRLMDQGFVTRDTAT

FFRTLIQDAVKARDEKGIVRPDMIHLLMQARDNERVSLSTDDIVAQAFIFFVAGF
DTSSTLMCYMLHELVLNPDVQQKLREEVDQVTSGAEDEVTYEMLQKMKYMDM
VISETLRKHPPVPAIDRVCVKSFQLPQSMPGYSDVTVEKGSVAWFPVHALHYDP
QYFPDPYKFDPERFNDENKDKIVPYTYLPFGLGPRMCIGNRFALMETKIIVTHLV
QNFMFERTEKTVDPDKFTLQSFSITPKEGFWLRLKPRKI

>CcP450_20
MSVLSDIWDLKFYLLAGAIGAIYLYVKLVIFKYFDRRGIPYERPTFPFGNVSGVFR
GKMSIGEACAELYKKHRQSPYFGMFISINPALVINDPEIIRHVLTKDFSHFHDRGIY
VDEKVDPLSGHLFSLPGEKWKNLRTKLTPTFTSGKIKQMFPIVNQCGVELGDCVE
KVLKIDETVEIKDIVARYTTDVISSVAFGLETNSLKEPNNEFRRCGQLVFSSRPLTT
LLSLFMPSLLRILKLPNTDTRVTNFFTKVFSDTVNYRKQNNIKRHDMLNMLMQL
MDKGYVETDDANEIPKGKAVNDSKKLEMTEATAQAFVFFVAGFETSSSTVTYCI
FEMACHPEIQEKVYKDIDQALKKHGGLTYDSLAEMEYLSQAVDETLRLYPTLPIL
NRECTREYDVPGTNLHISKGDAIVIPVLGIHKDPNIYPDPEKFDPERFSKENKGKIH
SHAYLPFGEGPRICIGKRFGLIQAKVALVSLLSRFRFSTNSKTEIPVKHSPGTIVMC
PDHGVNLRVEMR

>CcP450_21
MEFLSADGIGITLALVATAISLFYIYAKYKLSYWKRRGVESLPTNLIFGNFKDAVL
FRTAPGWHLGNLHKAAKKDVPYLGFYIFHKPCFLLRDPEVIKQILIRDFENFSDR
HFAGSEQKDSIGMKNLFGIKNPNWKYLRAKITPTLTRGKLKQMFPLMLETGEPM
MKYLENQPENKNGVKVIDAQELSYKYTTDLIASIALGTKMDSFHYPNEEFSKAV
MEFFHGFKRMVALVTVFFMPELVEMIGTRMLFNSSFVKKVFWEAMESRERTGN
KRGDFIDSLLQLKNGEQNSDYKFEGENLLYQSGTFFSGFESSSTTMSFTLMELAN
HPEYQERVRQDINKAIDAHGFTFDAFNEMKYLDQAIAEGVRLHPPVSTIDRYTRK
DYKIPGTDIILEKGTPIYISLYGLQEDPRFFKDHSTFNPDRFTDDSKVADAYIPFGS
GPRMCVGMKVGQLHAKVVIAMLLREYEIWQKPEDKSILDPRSTFTAAANGIVLQ
FRKLVK

>CcP450_22
METSTSHLLIYVVTFILSLVGLLYIYIRYKLSYWSRRGVKTPPVHLLFGNFKDTITF
KKAPGNTLRDIHENADADDPYIGFYIFQKRYLMLRNLDIIKQIFIKDFDLFPNRNF
GGKREIDSVGLINLLGIRQPGWKYLRQKITPFITGQKLRRMIPLMVDSGYPMTNFI
DDKSTDSEGWKEVELKDISSRYSTDVISSLAFGVTTNSFNKNDTAFWKAGLKILS
GTKRNIILMIVFLLPELIDIIYPIMTGPAQFFREVFWDSMNTREKAGYKRGDLIDSL
LQLKNAQQNPIFKFEGDNLLGQAVAFYVAGYEASSTAIAFTLYELARHPEYQER
VFQEIQTEISKNGIDYNSINEMSFLDKVVIETLRMHPPLPLVDRFAIEDYKIPGTDLI
IEKGTPVIVSTNGINRDPKYYSNPDEFDPERTQEGKEELIGASLTFGIGPRSCVGQR
IGTLMTKVGVITILSNYELSYKCKENEDFNAITIFTAAADGVVVHLKKREKPIKIS
VNEWATD

>CcP450_23

MGALTGHWIGDGLVLMFSIAAVFYLYMTRKFNYWKKRGVLEFKPVPFLGNFGE
CLFFRKSALDFIKDYYNNSKGLKYMGFYILDQPFLMLRDPELIKHILVKDFNYFH
DRYTSTGKKDILGNANLFTIKNPAWRFLRGKLTPFFTSGKMKKMFELMLAVNKD
LDTYLESLDIDSKGREMEMKELCAMYTTDLIGVTAYGLRVNSLNNPDAEFRACG
RQIFNFNWRRGMELACTFFAPKIVDLFGFKFFSKESTEFLRRAFWDTIDERVKSG
VKRNDLIDLLIDLREQTKNGENKEFRFDGDNLVAQAAIFFTGGFETSSSAMSFSL
YELALQPEIQRKLRVEILEALEDKNGELTYDMVANLPYLNMVVSETLRKYPPLPF
LDRETNMDYKLPGTDLIVEKGTPVYIPLQGLHYDPQYFPDPEKFDPERFSEENKK
NIPPCVYMPFGEGPHICIGLRFGLLQTKLGLIKILSKYEISACDKTNVPIVFNKKAL
VIAAEGGITLKIKKLMKVPG

>CcP450_24
MSSIILELMNSWVFIIAAIAALYVYMNRQGNYFKKLGIPYLKSYMLVGNRGPVIF
RQRSSIEHEQALYNSMPDAKYFGMYDYSSPAIMVRDIELIREIGSKHFDNFVDHK
SMFDPELDSLLSSTVFFLREDKWRRMRSLLSPAFTSSKIKTMFGLISNCATNLADF
VVKQQEPIILDMKDLFSRCTNDVIASTAFGINLNSMVDRQNEFYKMGREVTDFGF
LRSSKIFCMQMWPKLARKLKASLLTDRMVDFFHQIIATTVATRRAQGIQRPDLIH
LLMQASESETSSKLTVDDMTAQAYIFFFAGFDTSSTTMSFAAYSLAANPDVQEKL
QTEIDEALEKCNGEVTYDYLKEMKYLDAVISETLRMYPPVTELDRLCTKPFELPP
AKEGGKPVFVKRGTCFLIPVNGIHRDPKFFPNPDKFDPERFLNEKEKINSSAYIPFG
VGPRTCIGARFAMIELKLVLFHLLKMCNLKPCAKTCIPMKFCTKTFMPAPATGH
WLQIERRK

>CcP450_25
LWNWPVISIFSRRDIETVLRRASRYPLRPPSEVISYYRQSRPDRYTNLGLVNEQGIT
WHKLRSALTPELMGANTVLGFLPALNDVTDDFVDLIRSQRSGSKVVGFEELAYR
MGLESTCTLILGRRLDFLKPDSGSLTATLAEAVRVNFTASRDAFYGLPFWKLLPT
KAYKQLVESEDVIYDTISELMETTIWEKQDDAKDEIVEAVFQSILREKNLDIKDK
KAAIIDFIAAGIHTLGNTLVFLFYLIGRDAKVQRRLYEEAASLAPPGCNLVAEDLR
GAMYLRACITEAFRMIPTTPCIARILDEPIELSGYNLNAGTVVLLQTWIAGLDEEN
FKNASQWMPERWLEPTSPHSPLLVSPFGAGRRICPGKRFTEQALQLILAKVVREF
EIIADEELGLQFEFILAPKGPVSLSFRDRLECT

>CcP450_26
IFDILYILQKVLNSQQASYKGRVYRFLHPFIGNGLISGDGNMWRRHRKLLTYTMT
QKVLHGFVNIFDRHSRKLMDKLESQADGGYDFDIFPYIEACTIDIVCEAVMGRLD
VNAQDNADQEIIHYTAKMYKIIFERMTKVWLQSDWIFNQTRYYTEQQRGREVIQ
NFVHTCVSAKRNDQNYTPLLSSNQKSILEHLLESLKSSSNSLTDSELQDEIYTVYI
ASQDTIALISSFATLMLGMHHEAQEKARKEIEEIFQGEDEPLTSESLSQLKYLEAVI
KETIRLFPIAPFLIRHLRGELPLENCILPEDCQVLIAAYVTHRSVDYWREPKKFLPE
RFSAEESSTRHPYAFIPFSGGPMSCIGQKFAMTCLKVILANLLRCYRIDTTWQMSD
LLLNADISVRSINGYRVSLKPHEKAA

>CcP450_27

MSWIWSILFDTKFLLLGTLIISIYLYMKKVVYNYFEARGIPYEKPIVPFGSAHILFV
RRTTFGELFANSYKKFNKYPYHGIFMLNVPNLVINDPEIIRLVLNKDFAKFHDRG
LYFNEKADPLSGHLFVLGGEKWKRLRAKLSPTFTSGKIKQMFPVVSECSQLLAE
HINESMKTDSDFEIKDLLGRYSTDVIASIAFGLDSRCLKDPNSDFRKYGKKIFTGQ
TIRTFLFIFTPRLVDRLGIRLSQSDVNEFFIKTFTEMVQYRKDNNIERSDMLNLLIQ
LMDKGYVESDNVKTDKSNGEVDTRKLGLLEAAAQAFVFFIAGFETSSSTVTYCL
YELALNPEIQEKVRQEIETTIEKYGGLTYESLTDMTYLSQVLDETLRKYPTVPVLN
RQCNEDFMVPTTRQIIPKGLNIVIPVQGLHQNPELYPDPEKFDPDRFSKENAKKIH
PFAYLPFGEGPRVCIGKRFGLIQSKMAIISLLSRFRLSPCAKTTIPVVHCRRNLVMT
PQDGVYLRFDPR

>CcP450_28
LKTNYNNFRIANPSRTLYNYLANRHLILSGDFKIEEVDKVSTCLHAVYGDIVKIEG
LLGRPDMVFIYDASEIERIFRQEEKMPHRPSMPSLNYYKHVLRKEHFGNNAGVIG
VHGESWYNFRSKVQQVMLQPRSARMYIGAIEEASQSFLRRIDRIKDENDETPDDF
LNEIHKWSLESIANVALDVRLGCLDNDAPLETQKLIDAVNVFFKNVGVLELKIPF
WKIFNTPTWQKFVNALDVIVDITRKYTDAALTREKTSGDSEKELSILEKILAVEN
DTKTASILALDLFLVGIDTTSNAVASVLYQLALHPEKQAILFDEVARTLPQRNEKL
ESRHLDNLKYLKACIKETLRMYPVVIGNGRCMTKDTVISGYQIPKGVQVVFQHY
VISNQDKYFPRSKEFLPERWLTGNDVLHAFASLPFGYGRRMCLGKRFADLEMAV
VISKVSNNSQLNPPTKFLLF

>CcP450_30
MENPCETVVDSGALREMPEPKGLPVLGTIVDLIISGGAAKLHLYNDKRHAELGPI
YRERIGSVKAVFINSPDEFRRVFRLEGPKPIHFLPEAWTLYNEIRKCKRGLFFMDG
DEWLYFRRIANKLLLRPKSEEFMSRPCQNAADSLASRWEIYSENGKLIPELEAQL
YQWSIEVMLATLMGRNKWEEYGPNIMRKSENLAKNLHRIFEYSATLSLIPAKLA
MRIKLPIWKKFVGSIDEVMEIVRILVPEIIRLDGDGLTKMILDEGVQGEDLVRIVA
DLILAAGDTTAYSMQWALFLLASHPDVQERLYESIKSLDQDEIPKDPFLKGVIKES
LRLYPTAPFLTRFLPEDNIIGGYRVAKGELLLLSLYSSGRNDANFPEANEFRPERW
IRTSTGEYKGVINPHATLPFALGARSCVGRKLAEYQISLTLAKIVKTFRMDCTNK
DSVEMILHLVPVPSEPIQLLLTKRSWP

>CcP450_31
MDTQIIHSSSANTLDCDNAAATAEILDVPKVRVTPDLTMEKTALPFEEIPGPAILK
LWEKYWKYVPLLGTQLFCSLLINRFTQGRLNWNRNITPLQYLFNEYGPIVRIHGPI
AGNILRCSNRKKIHLYEVVSIFCNITICTIENTDCPDHSLCAEWTDVKVKLEKPFT
QQLSQYFDRLELTSDELVQRIRKIRNRQDEVPGQFIEELTRWSMECFSILMFNRRL
GFLESTGLNVASEPARIIEALTTAHVYLSRCETGFQVWRFFETPFAKKLFGACDVI
DSVIGKYIRQAQNKVRHRTLTPLDDKDTMAEKSSPILEKMLFNERIHPDDISTLL
MDMIVLGVQATSNCQGFLLYYLAKNPRVQRKLFEEIYPILPNRSSTLNEATLKSL
PYLKACLQECLRLRPAFPYITRVLPKTINLHGYTIPKGTYLIMANQIAARREENFE
D