The Microbial Associates and Putative Venoms of Seed Chalcid Wasps (Hymenoptera: Torymidae: *Megastigmus*)

by

Amber Rose Paulson
BSc, Vancouver Island University, 2007

A Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of

MASTER OF SCIENCE

in the Department of Biology

Supervisory Committee

The Microbial Associates and Putative Venoms of Seed Chalcid Wasps
(Hymenoptera: Torymidae: *Megastigmus*)

by

Amber Rose Paulson
BSc, Vancouver Island University, 2007

Supervisory Committee

Dr. Steve Perlman, Co-supervisor
(Department of Biology)

Dr. Patrick von Aderkas, Co-supervisor
(Department of Biology)

Dr. Juergen Ehlting, Departmental Member
(Department of Biology)

Supervisory Committee

Dr. Steve Perlman, Co-supervisor
(Department of Biology)

Dr. Patrick von Aderkas, Co-supervisor
(Department of Biology)

Dr. Juergen Ehlting, Departmental Member
(Department of Biology)

Abstract

Conifer seed-infesting chalcids of the genus *Megastigmus* (Hymenoptera: Torymidae) are important forest pests. At least one species, *M. spermotrophus* Wachtl, has been shown to be able to manipulate the seed development of its host, Douglas-fir (*Pseudotsuga menziesii*) in remarkable ways, such as redirecting unfertilized ovules that would normally abort. The mechanism of host manipulation is currently unknown. Microbial associates and venoms are two potential mechanisms of host manipulation. Microbial associates are emerging as an important player in insect-plant interactions. There is also evidence that venoms may be important in gall-induction by phytophagous wasps. PCR and 16S rRNA pyrosequencing was used to characterize the microbial associates of *Megastigmus* and transcriptomic sequencing was used to identify putative venoms that were highly expressed in female *M. spermotrophus*. The common inherited bacterial symbionts *Wolbachia* and *Rickettsia* were found to be prevalent among several populations of *Megastigmus* spp. screened using a targeted PCR approach. A member of the Betaproteobacteria, *Ralstonia*, was identified as the dominant microbial associate of *M. spermotrophus* using 16S rRNA pyrosequencing. The transcriptome of *M. spermotrophus* was assembled *de novo* and three putative venoms transcripts were identified as highly expressed in females. One of these putative venoms transcripts, Aspartylglucosaminidase, (AGA) appears to have originated through gene duplication within the Hymenoptera and has been identified as a major venom component of two divergent parasitoid wasps. AGA was identified as a promising candidate for further investigation as a potential mechanism of early host manipulation by *M. spermotrophus*.

Table of Contents

List of Tables

Supplemental Tables

List of Figures

Supplemental Figures:

Acknowledgments

I would first like to express my deepest gratitude to my supervisors, Dr. Steve Perlman and Dr. Patrick von Aderkas for their guidance and support throughout this entire process. Thank you Steve and Patrick for allowing me this wonderful experience, I learned so much and had a marvelous time doing it! The kind encouragement I received from both of you is very much appreciated and all of your thoughtful feedback has helped me to become a better biologist. I would also like to acknowledge Dr. Juergen Ehlting for providing mentorship in next generation sequencing technology. Next, I would like to thank all of the wonderful French colleagues involved in the MAC-BI project, including Dr. Thomas Boivin, Dr. Marie-Anne Auger-Rozenberg, Dr. Jean-Noël Candau and Dr. Alain Roques. It was an honour to work with each of you and I learned so much about *Megastigmus* from this collaborative experience. I would also like to thank all of my amazing lab mates from the Perlman, von Aderkas and Ehlting labs for being awesome. I would like to give a special thank you to Dr. Belaid Moa for all of your help running bioinformatics programs on WestGrid. Thank you to the seed orchard managers, Tim Crowder, Don Piggott and the late Tim Lee, as well as, Dave Kolotelo from the Surrey Seed Centre for providing infested seed for this project. Thank you to my parents Phillip and Debbie Paulson for always encouraging my scientific curiosity. Last but definitely not least, I would like to thank my wonderful partner and rock, Mike Ashbee, for being a source of unwavering support and encouragement each and every day.

THE EVOLUTION OF PHYTOPHAGOUS HYMENOPTERA

1.1 General evolutionary themes within the Hymenoptera

The ants, bees, wasps and sawflies form the Hymenoptera, one of the most successful animal radiations of all time, with 115,000 described species and many more yet to be described (LaSalle and Gauld 1993, Grissell 1999a). The Hymenoptera is an immensely diverse group with phytophagous, parasitic, predatory and eusocial members, comprising a vital component of terrestrial ecosystems. Bees likely represent the most important of all angiosperm pollinators and ants have been recognized as a primary component of arthropod biomass in terrestrial ecosystems around the world. Hymenoptera is traditionally divided into the Apocrita (the thin-waisted wasps) nested within the paraphyletic Symphyta grade (the broad-waisted wasps or sawflies) (Ronquist et al. 1999, Vilhelmsen 2001) (Figure 1). The Apocrita can be further divided into the Aculeata (stinging wasps), which are nested within the paraphyletic Parasitica grade (parasitoid wasps) (Ronquist et al. 1999, Sharkey 2007). The Aculeata have a modified stinging ovipositor that is used defensively or in prey capture, while the ovipositor of parasitic wasps is used for laying eggs in or on their hosts.

The evolution of parasitism within the Hymenoptera was the single most important shift giving rise to an explosive radiation (Wiegmann et al. 1993, Whitfield 2003, Davis et al. 2010, Heraty et al. 2011). As a result of this successful shift, the majority of hymenopterans are specialized parasites known as parasitoids. Parasitoids are characterized by having a free-living adult stage and a larval stage that develops on or within an animal host (usually another insect), ultimately killing it (Eggleton and Gaston 1990, Eggleton and Belshaw 1992). This important transition likely occurred within the Vespina (Orussidae + Apocrita) (Heraty et al. 2011). The Orussidae are parasitoids of wood-boring beetles and are considered the sister group to Apocrita (Ronquist et al. 1999, Vilhelmsen 2001, Schulmeister 2003, Davis et al. 2010) . Parasitoid Hymenoptera are key regulators of phytophagous insect populations, important indicators of ecosystem health and essential components of several biological control programs (LaSalle and Gauld 1993). Parasitoid wasps are also a large component of Earth's biodiversity, for example, the superfamily Chalcidoidea is likely the most diverse group of insects, containing up to an estimated 0.5 million species (Noyes 2013).

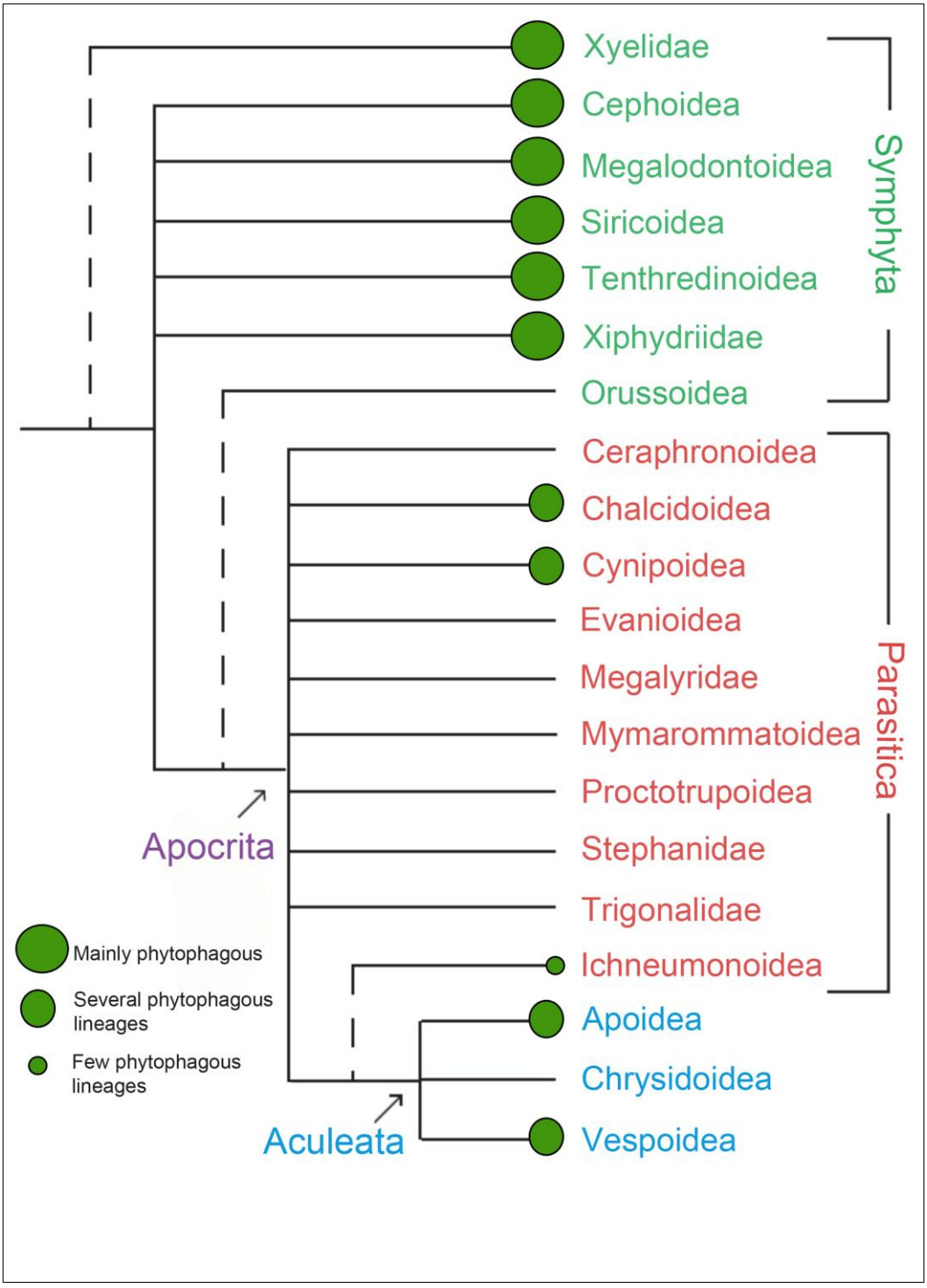


Figure 1: A summary of hymenopteran relationships adapted from (Davis et al. 2010). Terminal taxa are superfamilies or those families not assigned to a superfamily. Dashed lines represent hypothetical sister group relationships. Green circles denote taxa containing phytophagous lineages.

While the factors driving the origin of the parasitoid lifestyle in Hymenoptera are not well understood, the most widely accepted hypothesis is that parasitic Hymenoptera evolved from mycophagous origins (Eggleton and Belshaw 1992, Whitfield 2003). It has been proposed that the earliest parasitoids looked very much like the Siricoidea (Sharkey 2007, Vilhelmsen et al. 2010). Most siricids harbour symbiotic fungi that they inject into dead wood during oviposition; the developing larva subsequently feeds on the fungus for nutrition. Some siricids, however, lack symbiotic fungi and steal the burrows and fungi of related species. In the mycophagous origin theory, this strategy of stealing evolved directly into parasitism (Eggleton and Belshaw 1992). Indeed, parasitoids of wood-boring insects are found to be basal in many parasitic lineages, such as the Evanoidea, the Ichneumonoidea and the Cynipoidea (Eggleton and Belshaw 1992).

Subsequent to the parasitoid lifestyle, there have been other notable successful radiations in the Hymenoptera, including repeated secondary reversals to phytophagy, such as nectar- and pollen-feeding in bees, as well as gall-making and seed-feeding in some Parasitica (Whitfield 2003, Heraty et al. 2011). Finally, two important transitions in the Hymenoptera are the evolution of provisioning within the Aculeata and the subsequent independent evolution of eusocial behaviour in the some bees, wasps and ants (Andersson 1984, Pilgrim et al. 2008).

In comparison to other hymenopteran innovations and life history strategies, such as complex social behaviour and the parasitoid lifestyle, there have been few syntheses on the evolution of phytophagy in Hymenoptera. This introductory chapter aims to provide a synopsis on the evolution of phytophagous Hymenoptera. I will survey the wide range of plant-feeding guilds within the Hymenoptera, including foliage-eaters, wood-borers, stem-borers, leaf-miners, and pollen- and nectar-feeders. I will focus in particular on internal parasitism of plants (endophytophagy), such as gall-inducers and seed parasites, as they represent a specialized group of plant feeders that exhibits very intimate associations with their host plants.

1.2 Phytophagous Hymenoptera

1.2.1 Basal phytophagous Hymenoptera: Symphyta

The vast majority of Symphyta, the most basal Hymenoptera, are phytophagous (Vilhelmsen 2001). Many types of plant-feeding guilds are found within this paraphyletic suborder, also

known as the sawflies. The larvae of most sawflies forage on exposed vegetation (Sharkey 2007), but in many symphytan groups the larvae have adapted to feed internally on plant tissues through leaf-mining, gall-induction, stem-boring or wood-boring (Roininen et al. 2005). It is thought that phytophagy in sawflies likely arose from saprophagous ancestors (Malyshev 1968, Roskam 1992). The larvae of Cephoidea (stem sawflies), Siricoidea (horntails) and Xiphyridoidea (wood wasps) are primarily xylophagous, feeding internally on wood. The stem sawflies mainly develop within herbaceous plants, while the horntails and wood wasps introduce a symbiotic fungus during oviposition and then the larvae feed on the fungal infected wood (Sharkey 2007).

The ability to induce galls within the sawflies has evolved independently six to ten times, and the majority of gall-inducing species belong to the tribe Nematini within the Tenthredinidae (Roininen et al. 2005). Sawflies induce relatively simple galls on leaves, buds, shoots or berries of mainly *Salix* and *Populus* (especially *Euura* and *Pontania*) (Roskam 1992). The oviposition behaviour is very important to the sawfly gall-induction process, with the injection of colleterial fluid stimulating at least the initial growth of the gall (McCalla et al. 1962, Smith 1970). The larvae of gall-inducing sawflies are caterpillar-like and they live and behave similarly to leaf and stem miners as they do not require pre-digested food, feeding on the inside of the gall, which resembles wound callus tissue (Rohfritsch 1992).

1.2.2 The Secondary evolution of phytophagy: Apocrita

Several lineages within the Apocrita, especially within the Parasitica, independently reverted back to phytophagy from a parasitoid life style. The habit of gall-forming has evolved independently in several groups of ancestrally parasitic Hymenoptera, including the family Braconidae (Austin and Dangerfield 1998), several families of Chalcidoidea (Munro et al. 2011) and the family Cynipidae (Ronquist and Liljeblad 2001). Also, the habit of seed-feeding is found within the Chalcidoidea (Munro et al. 2011). Inquilines are also common among many of the previously mentioned phytophagous lineages and these are phytophagous species that have lost the ability to initiate galls *de novo*, but still reside within galls induced by other insects, retaining some ability to influence gall tissue to produce nutritive cells (Ronquist 1994, Brooks and Shorthouse 1998).

1.2.3 Gall-induction in the Parasitica

The induction of galls is a specialized feeding strategy that is found not only in the primitive Symphyta, but also in select lineages within the Parasitica. In a broad sense, a gall is any pathological excrescence produced by a specific reaction to the presence and activity of a foreign organism, in which the modified tissues of the plant serve as the shelter and a nutrition source for the causative agent (Dreger-Jauffret and Shorthouse 1992). Insects that are capable of inducing plant galls are highly specialized herbivores, being able to over-ride normal plant development by instigating unusual gene expression in adjacent plant cells. Unlike the galls induced by sawflies, the galls induced by parasitic Hymenoptera are comparatively more complex, with the inner gall tissue being comprised of nutritive tissues on which the larvae feeds, concentrically surrounded by several discrete layers (Rohfritsch 1992).

1.2.4 Cynipidae

After the gall midges (Diptera: Cecidomyiidae), gall wasps in the family Cynipidae form the second largest radiation of gall-inducing insects, with currently over 1,300 described species of gall-inducers and inquilines (Liljeblad and Ronquist 1998, Ronquist et al. 1999). The cynipid gall wasps belong to the superfamily Cynipoidea, which contains mainly parasitoid species. Cynipid gall wasps produce arguably the most complex and well-organized insect-induced galls (Cornell 1983), with easily recognizable galls found on oaks and roses (Ronquist and Liljeblad 2001). The inner-most layer is made of nutritive tissues, which surround the developing larva, followed by concentric layers of starch, sclerenchyma, cortex, peripheral vascular tissue and epidermis (Rohfritsch 1992). Most cynipid gall wasp species are very specific with respect to the location of the gall, with most species targeting leaves and buds, but some species target stems, catkins and roots (Dreger-Jauffret and Shorthouse 1992). The surfaces of cynipid galls are commonly covered with hairs, fleshy or spiny outgrowths, scales and/or sticky resins (Stone and Cook 1998). The initial induction of gall growth is likely a result of wounding by the ovipositor, the lytic action of the eggs on surrounding plant tissues and/or ovipositional secretions (Rohfritsch 1992, Stone et al. 2002, Leggo and Shorthouse 2006); however, the completion of gall growth requires the activity of larval feeding (Leggo and Shorthouse 2006). Many oak and sycamore gall wasps alternate between a sexual generation in the spring and an asexual generation in the fall with each generation targeting different host tissues on the same or closely

related host species (Roskam 1992). The galls of Cynipidae support diverse and ecologically closed communities of inquilines and parasitoids, providing an important model system for community-level ecological research (Csóka et al. 2005).

The earliest theory on the origin of phytophagy within the Cynipoidea was proposed by Kinsey (1920). He proposed that gall-induction evolved from primitive phytophagous cynipids that were non-gall making inhabitants of herbaceous plants and that gall-induction on woody plants, such as rose and oak, was a derived trait evolving later among gall-inducing cynipids. Examples of extant primitive cynipids belong to the genera *Aulacidea* and *Phanacis*, in which species either cause conspicuous multi-chambered stem swellings or no outward deformity in herbaceous Asteraceae (Ronquist and Liljeblad 2001). In other early work, Wells (1921) also speculated that the first cynipid galls were multi-chambered stem swellings and that more complex single-chambered and detachable galls evolved later within the Cynipidae. A later, conflicting theory emerged from the work of Malyshev (1968), who argued that the first cynipids were likely associated with higher woody plants rather than the more recently evolved asters. Malyshev also suggested that the first galls were induced in reproductive buds or developing seeds, and that gallers evolved from seed rather than stem-feeders. The latter theory has received less support from later phylogenetic investigations (Roskam 1992, Ronquist and Liljeblad 2001). A recent analysis of cynipid phylogenetics, based on an extensive morphological dataset, found that the earliest cynipids likely induced single-chambered galls within the reproductive organs of herbaceous members of the poppy family or possibly the mint family. This analysis also suggested that the colonization of woody hosts has only occurred three times within the Cynipidae (Ronquist and Liljeblad 2001). Roskam (1992) suggested that inquilines form a monophyletic group that arose from one gall-inducing host and later radiated to attack other hosts. However, a more recent phylogenetic analysis by Nylander et al. (2004), combining morphological and molecular data, found that inquilines may not form a deeply nested monophyletic group among gall-inducing lineages. These findings support an alternative hypothesis that inquilinism evolved several times among the Cynipidae.

1.2.5 Chalcidoidea

The superfamily Chalcidoidea is an extremely diverse group, constituting one-third of all described parasitic Hymenoptera (Lasalle and Gauld 1991). Most phytophagous chalcids are

either seed-feeders or gall-inducers (Roskam 1992). However, there are potentially a few representative stem boring chalcids (*Tetramesa* (Eurytomidae) and maybe *Aiolomorphus* (Eurytomidae)) (Lasalle 2005). The galls of chalcids contain an inner layer of differentiated nutritive tissues that surround the developing larva (van Staden et al. 1977). Adjacent to the vascular bundles a wall of sclerenchyma forms behind the nutritive layer (Rohfritsch 1992). Gall initiation is thought to be caused by the eggs and the ovipositor wound. Often a swelling of the attacked organ forms opposite to the oviposition scar (Rohfritsch 1992).

Gall-induction evolved independently at least fifteen different times within Chalcidoidea. The majority of gall-inducers belong to Agaonidae, Eurytomidae or Torymidae. A few cases are found among the Eulophidae, Pteromalidae and Tanaostigmatidae (Lasalle 2005). Little is known about the general biology of gall-inducing species. There is also a lack of knowledge about the phylogenetic relationships among the Chalcidoidea which limits our ability to develop evolutionary theories to predict the transitions leading to phytophagy within this group (Lasalle 2005, Munro et al. 2011). It is likely that many gall-forming chalcids arose from progenitors that are parasitoids of gall-inducers, probably via an inquiline intermediate step (Wharton and Hanson 2005). There are few examples of extant transitionary stages, such as facultative parasitoids of gall-inducers in the genus *Eurytoma* (Eurytomidae). Here, the wasp larva first feed on its insect host and then on host-derived gall tissues (Zerova and Fursov 1991). *Paragaleopsomyia cecidobroter* Gordh & Hawkins (Eulophidae) is another example of an extant transitionary stage, in which the larva develops within independent endogalls with the host gall (Hawkins and Goeden 1982).

Megastigmus (Torymidae), *Eurytoma* (Eurytomidae), Melanosomellini (Pteromalidae), *Tetrastichinae* (Eulophidae) and *Tanaostigmodes* (Tanaostigmatidae) are lineages that include gall-associated species (gall-inducers, inquilines, or parasitoids of gall-inducers), as well as seed-feeders. The presence of gall-associated species and seed-feeders within several lineages suggests that there have been potential shifts between seed-feeding and gall-induction over evolutionary time (Lasalle 2005). It has been proposed that some lineages of gall-inducing chalcids likely evolved from phytophagous ancestors, such as seed-feeders (Malyshev 1968, Wharton and Hanson 2005). A recent molecular phylogenetic re-construction of the Chalcidoidea found that phytophagous groups were scattered across the tree and rarely formed

basal clusters within their respective lineages, providing evidence against the phytophagous ancestor hypothesis (Munro et al. 2011).

1.2.6 Agaonidae

The monophyletic family Agaonidae includes all of the fig-pollinating wasps, which form an intimate mutualism with figs (Kjellberg et al. 2005). The fig-pollinating wasps are parasites of plant reproductive tissue that have evolved a mutualistic relationship with their host. The *Ficus*-agaonid wasp association is one of the classical examples of insect-plant mutualism and co-evolution (Weiblen and Bush 2002). The fig-pollinating wasps demonstrate extreme host specificity, specialized morphology and life cycles that are completely synchronized with fig reproductive phenology (Wiebes 1979, Weiblen 2002). Pollen-carrying females enter receptive figs, pollinating internal flowers as well as laying eggs in a few of them; the larvae subsequently develop within galled ovules and feed on endosperm (Kjellberg et al. 2005). Many chalcids and a few braconids have secondarily evolved to exploit the *Ficus*-agaonid association as gall-makers, inquilines or parasitoids (Cook and Rasplus 2003).

1.2.7 Seed-feeding chalcids

Seed-feeding represents an alternative endophytophagous life-style that is found primarily in two chalcid lineages, *Megastigmus* (Torymidae) and *Eurytoma* (Eurytomidae), as well as, Melanosomellini (Pteromalidae), *Tetrastichinae* (Eulophidae) and Tanaostigmatidae (Lasalle 2005). The larvae of seed-infesting chalcids develop within plant ovules, gaining access to a highly nutritious food source. Seed-feeding represents a very intimate interaction between insect and host plant, involving synchrony between plant reproductive phenology and oviposition (Rouault et al. 2004) and in some cases the manipulation of normal seed development (von Aderkas et al. 2005a). In contrast to gall-induction, there is no development of abnormal plant tissues during seed-feeding.

Eurytoma is a very wide-spread genus that contains diverse larval feeding guilds including mainly parasitoids, as well as gall-inducers, inquilines, seed-feeders and facultative parasitoids of gall-inducers (Lasalle 2005). The seed-feeding larvae of some *Eurytoma* species, such as the almond seed wasp *E. amygdali* Enderlein, are pests of stone fruits (Zerova and Fursov 1991).

The genus *Megastigmus* (Torymidae) also includes seed-feeders, facultative parasitoids of gall-inducers, parasitoids and gallers (Grissell 1999b). More than half of the 134 currently described species of *Megastigmus* are tree and shrub seed-feeders (Grissell 1999b, Auger-Rozenberg and Roques 2012), several of which are invasive pests of conifers (Roques and Skrzypczyńska 2003). In general, very little is known about host manipulation by seed-feeders, with *M. spermotrophus* Wachtl being the most widely studied. *M. spermotrophus* is a pest of Douglas-fir, *Pseudotsuga menziesii*. This species is known to influence normal seed development for its own reproductive success. Not only does *M. spermotrophus* re-direct unfertilized ovules that would normally abort, the developing larva acts like a 'surrogate' embryo, obtaining nourishment from the continued accumulation of storage reserves in the megagametophyte (von Aderkas et al. 2005a, b). The ability of *M. spermotrophus* to re-direct unfertilized ovules to continue development can be partially explained by changes in seed hormone levels, especially cytokinins (Chiwocha et al. 2007).

1.2.8 Braconidae

The Braconidae is one of the largest families among the Hymenoptera and until recently it was believed that members of the family Braconidae were exclusively parasitoids. It is now known that at least three groups of braconids, within the genera *Allorhogas, Mesostoa* and *Monitoriella*, are able to induce plant galls (de Macêdo and Monteiro 1989, Infante et al. 1995, Austin and Dangerfield 1998, Centrella and Shaw 2010); gall-induction is thought to have evolved independently in these three groups. Due to their recent discovery, very little is known about gall-inducing braconids, which tend to have very inconspicuous galls (Wharton and Hanson 2005). In all species of gall-inducing *Allorhogas*, larval feeding alone induces gall formation within young fruits of legumes. The role of accessory gland secretions of the ovipositing female does not seem to play a role in gall induction in this genus (de Macedo and Monteiro 1989). It is unknown whether ovipositional secretions are involved in gall initiation in other gall-inducing braconid species (Wharton and Hanson 2005). The galls produced by *Allorhogas dypistus* Marsh are relatively simple (de Macêdo et al. 1998) compared to the woody galls formed by *Mesostoa kerri* Austin & Wharton (Austin and Dangerfield 1998) and the leaf galls caused by *Monitoriella elongata* Hedqvist (Infante et al. 1995).

1.2.9 Phytophagy within the Aculeata

Several successful shifts back to phytophagy are also evident within the Aculeata, such as nectar- and pollen-feeding species of bees and wasps (Whitfield 2003, Danforth et al. 2006). The bees are arguably the most diverse group of aculeate Hymenoptera and they are of great ecological importance as pollinators of natural and agricultural plant communities. Bees are dependent on flower resources during both larval and adult stages (Neff and Simpson 1993). Phytophagy is also present in another very ecologically important and diverse aculeatan group: the ants (Formicidae). Many ant species are omnivores that forage opportunistically (Hölldobler and Wilson 1990). Also, the leaf-cutters ants of the tribe Attini tend fungal gardens that are fertilized with leaf fragments supplied by the ants (Mueller et al. 1998). Finally, many tropical ant lineages are almost exclusively herbivorous, and feed on plant exudates and honeydew excreted by phloem-feeding insects, such as treehoppers and scale insects (Buckley 1987, Heil and McKey 2003).

1.3 Nutritional considerations of phytophagous hymenopterans

Phytophagous Hymenoptera have evolved a variety of morphological, physiological and developmental adaptations required to consume plant materials. Also, symbiosis with microbes is wide-spread among several phytophagous aculeate and wood-feeding Hymenoptera.

Phytophagous insects consume suboptimal food, with dilute nutrients trapped within an indigestible matrix of cellulose, lignin and secondary metabolites designed to deter feeding (Schoonhoven et al. 2005). Though many herbivorous insects are known to possess intrinsic cellulases (Davison and Blaxter 2005), symbiotic microbes are also thought to contribute to the digestion of wood and other cellulose-rich diets (Douglas 2009). For example, two groups of xylophagous Hymenoptera, the woodwasps and the horntails, rely on a symbiotic fungus for cellulose-digestion and/or nutrition during larval stages (Kukor and Martin 1983, Šrůtka et al. 2007). Woodwasps have also been found to be associated with cellulose degrading bacteria (Adams et al., 2011). Leaf-cutter ants in the genus *Atta* have also formed a symbiotic relationship with fungi, in which the ants cultivate and consume a mutualistic fungus on a substrate of foraged leaf fragments (Weber 1966).

Accessing the rich nutrients stored within pollen grains is also difficult for animals due to an extremely recalcitrant outer coat (Roulston and Cane 2000). The honeybee, *Apis mellifera* Linnaeus, is known to be associated with a distinct microbiota (Jeyaprakash et al. 2003, Mohr and Tebbe 2006, Olofsson and Vásquez 2008, Martinson et al. 2011, 2012, Moran et al. 2012). This association suggests that symbiotic relationships are important for both bee health (Olofsson and Vásquez 2008, Martinson et al. 2011) and pollen coat digestion.

Some herbivorous diets, such as sap, are particularly nutrient-poor, lacking essential amino acids required by insects. Insects that feed exclusively on plant sap, such as aphids, whiteflies and other hemipterans, harbour obligate bacterial endosymbionts that supply them with essential amino acids and vitamins (Douglas 2009). Arboreal herbivorous ants that subsist mainly on sugary plant exudates and hemipteran honeydew secretions are similarly nutrient-limited. These harbour gut symbionts, which aid in nutrition. These symbiotic gut microbes include bacteria that are related to nitrogen-fixing root-nodule bacteria (van Borm et al. 2002, Russell et al. 2009, Anderson et al. 2012). Carpenter ants in the genus *Camponotus* have an obligate endosymbiont, the gammaproteobacteria *Blochmannia*, which is found in host-derived bacteriocytes (Degnan et al. 2004). The sequenced genomes of *B. floridanus* and *B. pennsylvanicus* suggest that the obligate symbiont provides nutritional upgrading by providing essential amino acids (Gil et al. 2003, Degnan et al. 2005). There is also evidence that *Blochmannia* plays a role in nitrogen recycling by encoding urease (Feldhaar et al. 2007).

Compared to nitrogen-deficient honeydew, plant exudates or indigestible pollen and wood, gall-tissue provides a richer source of nutrients. The nutritional hypothesis of the adaptive significance of galls suggests that gall-tissue is notably more nutritious, but contains less defensive compounds than unmodified plant tissue (Price et al. 1986, 1987). This hypothesis is widely accepted, yet few experiments have demonstrated that gall-inducers are able to manipulate nutrient levels within the gall (Hartley and Lawton 1992, Gange and Nice 1997, Koyama et al. 2004, Diamond et al. 2008) and that the nutritional manipulation actually benefits the gall-inducer (Koyama et al. 2004). Similarly, the endosperm within seeds likely provides seed-feeding larva with a rich source of nutrients with few defensive compounds. Analogous to the nutritional hypothesis of galls, is the selective feeding hypothesis, which suggests that leaf-miners are selective feeders that target plant tissues with a higher nutrient content and reduced

structural and/or chemical defense content (Kimmerer and Potter 1987, Connor and Taverner 1997). The nutritional role of microbes in leaf-mining, gall-forming or seed-feeding insects has not been widely studied.

1.4 Challenges associated with endophytophagy

The following discussion will primarily focus on galling and seed-feeding wasps, which are arguably the two most predominant endophytophagous habits found among Apocrita. Leaf-mining hymenoptera will largely be ignored, as they constitute a somewhat rare and understudied group.

Although the nutritive cells of galls and the storage cells of seed endosperm provide the developing larva with a nutrient rich diet, endophytophagous Hymenoptera have had to evolve a variety of adaptations in order to exploit plants in such an intimate manner. Females must oviposit eggs inside the host plant tissues. All gall-inducing symphytans possess a laterally compressed ovipositor with a serrated ventral surface that is used to saw into plant tissues (Vilhelmsen 2000). The adaptive potential of the ovipositor to exploit different hosts in different habitats was likely ‘a key’ factor in the evolution of parasitic Hymenoptera (Quicke 1997).

Once inside the plant, the initial act of host manipulation and gall induction is a critical period. It is at this stage that the insect gains control of plant tissues and redirects physiological processes and morphogenesis for its own advantage. The induction of the gall and the differentiation of plant tissues is typically a result of several factors, including ovipositional secretions, ovipositional wounding and specific activities of both egg and early instar larva (Rohfritsch 1992). In many sawflies, the injection of colleterial fluid is all that is required for gall formation (McCalla et al. 1962, Smith 1970, Price 1992). The role of ovipositional secretions in chalcid and cynipid galls is less clear, because continuous larval feeding is usually critical for galls to reach their maximum size (Rohfritsch 1992, Leggo and Shorthouse 2006).

The build-up of waste is another key challenge associated with larval development within a confined space made of plant tissues. Endoparasitoid wasps that develop within their host also face this same challenge. Consequently, all higher Hymenoptera (Apocrita) have evolved a blind larval midgut, in which the hind gut and the midgut do not join together until the end of the last

larval instar and the excretion of the fecal pellet is delayed until either pupation or adult eclosion (Wharton et al. 2004). The blind gut was likely pre-adaptive for endophytophagy, since not only does it enhance nutrient assimilation, but also prevents chamber fouling. Alternatively, gall-inducing sawflies chew a hole in the gall and then excavate their waste outside of the gall to avoid chamber fouling.

Plants have evolved direct and indirect chemical defense in response to herbivory and these defense responses present another challenge for phytophagous Hymenoptera that develop within plant tissues. It is thought that hymenopteran gall-inducers are able to avoid direct chemical plant defenses, since gall nutritive tissue usually contains few secondary metabolites (Nyman and Julkunen-Tiitto 2000). However, a study by Hartley (1998) found that the phenolic content of galled tissues was actually higher compared to normal plant tissues, and that perhaps phenols played a role in gall development by influencing plant growth pathways. It has also been suggested that some gall-inducing wasps can redirect these defensive compounds to outer gall tissues, which could deter other organisms from consuming or entering the gall (Allison and Schultz 2005). The underlying mechanisms by which gall-inducing wasps avoid or manipulate plant chemical defenses remains unknown. Plants are able to generate a diversity of signals and attacked plants may produce volatiles in response to phytophagous insects, which act indirectly to attract natural enemies (Thaler 1999, Wei et al. 2007). The indirect response of plants to galling insects has not been studied extensively. One gall wasp, *Antistrophus rufus* Gillette (Cynipidae), has been shown to alter the ratio of volatiles that its host plant emits during larval development, which provides olfactory cues for mate location (Tooker et al. 2002, Tooker and Hanks 2004).

Endophytophagous insects likely interfere with normal plant hormones, but the exact mechanism is unknown. Cytokinins and auxins are important plant hormones that are involved in cell division and the regulation of various processes associated with nutrient translocation, active growth, metabolism and plant development (Sakakibara 2006); these hormones have long been suspected to be important in the formation of insect-induced galls (Elzen 1983, Mapes and Davies 2001a). Changes in cytokinin levels are at least partially responsible for developmental re-direction of ovules by the seed-feeder *M. spermotrophus* (Chiwocha et al. 2007).

Endophytophagous insects manipulate plant hormone pathways by either producing endogenous cytokinin, modifying exogenous storage cytokinin or by another unknown mechanism. A study of tetrastichine (Chalcidoidea: Eulophidae) galls on *Erythrina latissima* found that most of the cytokinin activity occurred in the larvae as opposed to gall tissue. This same study also found that the leaf laminae contained mainly inactive cytokinin, while the larvae contained mostly active cytokinin, implying that the insect was able to sequester and modify exogenous storage cytokinin (van Staden and Davey 1978). Recently, the glands of adult willow-sawfly were found to have extremely high levels of an active cytokinin, t-zeatin riboside (Yamaguchi et al. 2012). Alternatively, microbial symbionts may alter cytokinin levels. To my knowledge the role of microbial associates in host manipulation by gall-inducing and seed-feeding Hymenoptera has not yet been investigated. However, there have been several interesting studies in a lepidopteran leaf-miner, *Phyllonorycter blancardella* Fabricius. This species causes a characteristic green island phenotype, where leaf senescence is delayed. Green islands associated with *P. blancardella* contain high levels of cytokinins (Giron et al. 2007). The green islands were eliminated following antibiotic treatment, implying that a microbial associate was involved in manipulation of the plant (Kaiser et al. 2010, Body et al. 2013).

1.5 Study-system: *Megastigmus*

The goal of this thesis is to develop *Megastigmus* as a model for understanding the mechanisms involved in the manipulation of seed development in conifers. This intimate plant-insect association presents an excellent opportunity for further exploration of the role of microbial associates and venomous secretions in host manipulation by endophytophagous Hymenoptera. Seed-feeding chalcids represent a fascinating and understudied plant parasite, having evolved specialized adaptations to manipulate seed development via unknown mechanisms. The widely distributed genus *Megastigmus* currently contains 134 species, with more than 72 seed feeders (Grissell 1999b, Auger-Rozenberg and Roques 2012). Of the remaining species, the majority are gall-inducers, inquilines or facultative parasitoids of other gall-inducing wasps (Grissell 1999b).

Recent surveys of seed-feeding *Megastigmus* from Europe found that 11 of 21 species were of exotic origin (Roques and Skrzypczyńska 2003, Auger-Rozenberg et al. 2006). The majority of introduced *Megastigmus* found in Europe are associated with conifers native to the Nearctic

(Roques and Skrzypczyńska 2003). Global and inter-regional movement of seeds with reduced phytosanitary regulations (i.e., minimal x-ray screening of seed products) is emerging as a common route for the accidental introduction of seed pests (Roques and Skrzypczyńska 2003). In addition to being a pest of seed orchards, the introduction of *Megastigmus* could potentially have negative impacts on the regeneration potential of native host plants due to increased seed losses in natural stands. The introduction of *Megastigmus* could also negatively affect insect biodiversity by increasing competition for seed resources in introduced areas (Fabre et al. 2004).

Many species of conifer-associated *Megastigmus* possess adaptations that contribute to their invasiveness (Roques et al. 2003). Many forest tree populations produce heavy seed crops at irregular intervals; this phenomenon is known as masting (Silvertown 1980). Accordingly most species of conifer-associated *Megastigmus* have evolved the ability to remain viable for up to five years in a state of extended larval diapause (Turgeon et al. 1994, Roques et al. 2003). The larvae are highly robust during this final larval instar increasing the chances of survival during seed harvest, processing, shipment and storage. Parthenogenetic reproduction also increases the chances of successful establishment in an introduced area. Some species of *Megastigmus* reproduce asexually due to infection by a vertically transmitted bacterium, *Wolbachia* (Boivin and Candau 2007, Boivin et al. 2008, 2013). Some species of *Megastigmus* are species-specific, while others exhibit generic-level host specificity. Conifer-associated *Megastigmus* have demonstrated host-preference plasticity in introduced areas; for example, both *M. pinus* Parfitt and *M. rafni* Hoffmeyer can now develop on *Abies alba* in southern France (Roques and Skrzypczyńska 2003, Auger-Rozenberg et al. 2006) and *M. schimitscheki* Novitzky switched from the Cypriot endemic *Cedrus brevifolia* to *C. atlantica* when introduced to France (Auger-Rozenberg et al. 2012).

At least one species of *Megastigmus* is known to oviposit in both unfertilized and fertilized ovules (Rouault et al. 2004). Originally it was thought that female *M. spermotrophus* selected fertilized ovules for oviposition since unfertilized ovules do not normally accumulate storage reserves (Hussey 1955); however, seed infestation levels of this species exceeded the expected amount of filled seed, suggesting that females must have also been ovipositing in unfertilized ovules (Rappaport and Roques 1991, Niwa and Overhulser 1992). Although Douglas-fir usually aborts unfertilized ovules, wasps were able to prevent this abortion and direct continued

accumulation of storage reserves (von Aderkas et al. 2005a, b). Chiwocha et al. (2007) implicated cytokinins as providing a partial explanation for failure of the megagametophyte to abort in the absence of a viable embryo, suggesting that the presence of the larvae could induce similar hormone profiles to those observed during normal seed development. The cytokinins were not characterized as being endogenous to the insect. The mechanism(s) resulting in continued redirection of unfertilized ovules and maintenance of storage reserves in the megagametophyte after the embryo has been consumed have yet to be discovered.

1.6 Thesis objectives

In this thesis, I explore two possible mechanisms contributing to seed-feeding adaptations in *Megastigmus* – microbes and venoms. The first data chapter of this thesis (Chapter 2) will focus on the identification and further characterization of the microbial associates of *Megastigmus* with the long term goal of understanding their role in host manipulation and host nutrition. Gall-inducing and seed-feeding insects form intimate associations with their host plants, using unknown mechanisms to manipulate normal physiological processes of the host. Endophytophagy has evolved independently within several parasitic hymenopteran lineages. Their parasitoid ancestors evolved numerous strategies to exploit their animal hosts, including producing a diverse cocktail of venoms that are injected into hosts along with eggs. The relatively well-studied Douglas-fir seed chalcid *M. spermotrophus* provides an interesting opportunity to investigate possible the role of venoms in plant manipulation. In the second data chapter of this thesis (Chapter 3), I use transcriptomic approaches to identify putative venoms in female wasps, with the long term goal of identifying mechanisms of early host manipulation.

Using molecular and next-generation sequencing approaches to characterize the microbial associates and identify putative venoms of *M. spermotrophus* will provide important information that can be used to develop control and management strategies for this invasive species. Furthermore, elucidation of nutritional aspects and potential mechanisms of host manipulation of seed-feeding chalcids will also contribute to our understanding of the unique biology and ecology associated with endophytophagy, a successful feeding strategy employed by diverse insect plant pests.

Chapter 2. CULTURE-INDEPENDENT SURVEY OF THE MICROBIAL ASSOCIATES OF THE SEED-CHALCID WASPS (GENUS: *MEGASTIGMUS*) USING DIRECTED PCR SCREENING AND 454 PYROSEQUENCING

2.1 Introduction

One of the major reasons that insects are the most diverse and abundant animals on Earth is due to their coevolution with plants (Schoonhoven et al. 2005). Indeed, insects have evolved myriad strategies to successfully feed on plants. Only recently have we come to appreciate the role of microbial symbionts of phytophagous insects in contributing to the evolutionary success and diversification of their hosts (Janson et al. 2008, Feldhaar 2011), for example by providing essential metabolites and vitamins (Dillon and Dillon 2004, Douglas 2009, Engel and Moran 2013, Nakabachi and Ishikawa 1999, McCutcheon and Moran 2007), breaking down cell wall components, such as lignocellulose (Warnecke et al. 2007), recycling nitrogenous waste (Whitehead et al. 1992) and detoxifying plant secondary metabolites (Genta et al. 2006, Adams et al. 2013).

Associations between insects and heritable (i.e. maternally transmitted) microbial symbionts are ubiquitous and extremely diverse in nature (Dale and Moran 2006, Moran et al. 2008). Perhaps insects that feed exclusively on plant sap provide the most profound example of the importance of inherited microbes shaping plant-insect interactions. All sap feeding insects possess obligate symbionts that provide their hosts with essential nutrients that are otherwise missing from this extremely limited diet (Feldhaar 2011). These obligate nutritional symbionts are usually found within specialized host-derived organs called bacteriomes (Baumann 2005). Obligate symbionts typically have extremely reduced genomes compared to their free-living and pathogenic relatives (Moran et al. 2008) and they often exhibit strict co-speciation with their host lineages, indicative of an ancient association stabilized by strict vertical transmission from mother to offspring (Wernegreen 2002). The best studied primary endosymbiont is the obligate nutritional symbiont of aphids, the gammaproteobacterium *Buchnera aphidicola*, which has been stably transmitted in aphids for 150-280 million years (Baumann et al. 1997, Douglas 1998).

Insects also commonly form associations with facultative heritable endosymbionts that are not necessary to the development and reproduction of the host. As a result of their maternal transmission, these symbionts have evolved diverse strategies to persist in their hosts, including manipulating reproduction, for example by inducing parthenogenesis (Stouthamer et al. 1999). Other facultative symbionts increase host fitness under certain conditions, and it is in this regard that they are potentially important in mediating plant-insect interactions. For example, facultative inherited symbionts of pea aphids have been implicated in facilitating the colonization of novel host plants (Tsuchida et al. 2004, Henry et al. 2013).

Gut microbes also play important roles in plant-insect interactions. Some herbivorous insects are associated with essential communities of microbes found within the chambers (e.g. termite, cockroach) (Bracke et al. 1979, Breznak 1982) or crypts (e.g. true bugs) (Glasgow 1914) of the gut. Unlike intracellular symbionts, which are transmitted from mother to offspring via transovarial methods (Dale and Moran 2006), several posthatch transmission mechanisms have evolved to ensure transmission of obligate gut associates from generation to generation, such as egg-smearing (Jones et al. 1999), coprophagy (Nalepa et al. 2001) and capsule-mediated transmission (Hosokawa et al. 2005). In addition, some true bugs have evolved the ability to acquire their essential gut microbes *de novo* every generation from the environment (Kikuchi et al. 2007, Olivier-espejel et al. 2011, Shibata et al. 2013).

The importance of obligate gut associates in plant-insect interactions is emerging as an active area of research. For example, when the symbiont capsule from a stinkbug pest of soybean, *Megacopta punctatissima*, is exchanged with a non-pest species, *M. cribraria*, there is an increase in fitness of the non-pest species on soybean and a decrease in fitness of the pest species on soybean. This implies that the obligate symbiont dictates the pest status of the host (Hosokawa et al. 2007). Since some of the major lineages of gut symbionts have only recently been discovered and characterized, we are still in early days in our understanding of how associated microbial communities are able to shape plant-insect interactions (Frago et al. 2012).

Some endophytophagous insects, such as seed-feeders, gallers and leaf-miners, have evolved the ability to manipulate plants in complex ways, permitting the larval stage access to internal plant tissues with relatively high nutrient content and low defense response. This very interesting

feeding lifestyle has evolved independently in several insect orders. However little is known about the complex adaptations involved in such an intimate association between insect and plant; even less is known about the role symbiotic microbes might have in these interesting systems.

Gallers, miners and to some extent seed-feeders, cause physiological and morphological modifications of host plant tissue, including differentiation of additional tissues (gall formation), *in situ* up-regulation and synthesis of proteins and sugars, translocation of nutrients to the insect feeding site and the formation of green islands (photosynthetically active areas surrounding leaf-mining insects during leaf senescence) (Stone and Schönrogge 2003, Giron et al. 2007, 2013, Schwachtje and Baldwin 2008). It is widely believed that cytokinins (CKs) and auxins are important in the formation and maintenance of insect galls and green islands. The exact mechanisms, including the role of microbes, are unknown. The synthesis of CKs and auxins are also important for gall formation by phytopathogenic bacteria, viruses and fungi (Jameson 2000) and the induction of nodule organogenesis by symbiotic nitrogen-fixation (Frugier et al. 2008). Increased levels of several CKs are found within the green island tissues in the *Malus domestica/Phyllonorycter blancardella* Fabricius leaf-mining system. The types of CKs involved in this plant-insect interaction are similar to those used by bacteria to manipulate plant physiology (Jameson 2000, Sakakibara 2006, Giron et al. 2007, Kaiser et al. 2010, in Giron et al. 2013), suggesting that microbes may be an important factor. When leaf-miners were treated with antibiotics, the green-island phenotype failed to appear. This suggests that bacterial symbionts, and perhaps *Wolbachia* (a known symbiont of *P. blancardella*), might be involved in manipulation of the plant (Kaiser et al. 2010).

Seed chalcid wasps of the genus *Megastigmus* (Hymenoptera: Torymidae) provide an interesting system to explore the role of microbes in nutrition and host manipulation of endophytophagous insects. The genus *Megastigmus* contains 134 described species, of which more than 72 are tree and shrub seed feeders; the remaining species are thought be mainly parasitoids of gall insects (Grissell 1999b, Auger-Rozenberg and Roques 2012). Seed infesting species of *Megastigmus* undergo their development within the seeds of plants, obtaining nourishment from the developing embryo and storage reserves within the megagametophyte (Roques and Skrzypczyńska 2003). *M. spermotrophus* Wachtl is the best studied species. It is a major pest of Douglas-fir (*Pseudotsuga menziesii*). This insect has the ability to manipulate the seed

development of Douglas-fir for its own reproductive success. First, *M. spermotrophus* can re-direct unfertilized ovules that normally abort to continue developing. Ovules do not redirect resources back to the mother plant, but feed the insect (von Aderkas et al. 2005a). Second, the developing larva acts like a 'surrogate' embryo, causing the continued accumulation of storage reserves in the megagametophyte, which provides nourishment for the larvae (von Aderkas et al. 2005b). The re-direction of unfertilized ovule development by the presence of the parasite can be partially explained by changes in seed hormone levels, especially CKs (Chiwocha et al. 2007). It is generally suspected that all *Megastigmus* species infesting Pinaceae hosts can manipulate seed development (Rouault et al. 2004).

2.1.1 Objectives

The aim of this chapter is to characterize the microbial symbionts of *Megastigmus*, with the long-term goal of understanding their role in host nutrition and manipulation. Little is known about the microbial symbionts of endophytophagous insects, let alone *Megastigmus*. Bansal et al. (2011) conducted a systematic survey of the associated bacteria of the Hessian fly, *Mayetiola destructor* and some cynipid oak gallwasps have been surveyed for *Wolbachia* (Rokas et al. 2002). The facultative inherited symbiont *Wolbachia* has been recently implicated in causing parthenogenetic reproduction in *Megastigmus* (Boivin et al. 2008). In this study, I used two approaches. First, I screened a large sample of *Megastigmus* species for common heritable endosymbionts, using symbiont-specific primers. Next, I used 16S rRNA Roche 454 pyrosequencing to perform an unbiased and in-depth survey of the microbes associated with different developmental stages of *M. spermotrophus*.

2.2 Materials and methods

2.2.1 Insect samples

Several species of *Megastigmus* and their parasitoids were screened for common heritable symbionts using PCR. Adult insects were reared from seeds that were collected from forest stands in France, Greece, Denmark and Turkey from 1997 to 2011; detailed information on sample species is listed in Table 1. Also, larvae of *M. spermotrophus* were dissected from infested seed collected in 2011 from seed orchards located throughout British Columbia. Adult *M. spermotrophus* were reared from this same seed. Any *Eurytoma* sp. parasitoids that emerged

Table 1: *Megastigmus* spp. and parasitoids screened for common heritable symbionts using PCR

Species	Host plant	Year	Location	Number	Sample Type
	Family: Pinaceae				
M. schimitscheki Novitzky	*Cedrus atlantica*	2010	Petit Luberon, FR	15	Female
M. schimitscheki	*Cedrus atlantica*	2009	Mont Ventoux, FR	14	Female
M. schimitscheki	*Cedrus atlantica*	2010	Saou, FR	14	Female
M. schimitscheki	*Cedrus atlantica*	2010	Gap, FR	15	Female
M. schimitscheki	*Cedrus atlantica*	2008	Barjac, FR	15	Female
M. schimitscheki	*Cedrus libani*	2005	Turkey	9	Female
M. rafni Hoffmeyer	*Abies alba*	2009	Lespinassière, FR	15	Female
M. rafni	*Abies alba*	2009	Pardailhan, FR	15	Female
M. rafni	*Abies alba*	2010	Ventouret, FR	15	Female
M. rafni	*Abies alba*	2004	Doubs, FR	9	Female
M. rafni	*Abies nordmanniana*	2000	Rold Skov, DK	9	Female
M. rafni	*Abies grandis*	2012	Vancouver Island, CAN	16	Female
M. rafni	*Abies grandis*	2012	Vancouver Island, CAN	10	Male
M. milleri Milliron	*Abies grandis*	2012	Vancouver Island, CAN	16	Female
M. milleri	*Abies grandis*	2012	Vancouver Island, CAN	10	Male
M. spermotrophus Wachtl	*Pseudotsuga menziesii*	2011	British Columbia, CAN	26	Female
M. spermotrophus	*Pseudotsuga menziesii*	2011	British Columbia, CAN	10	Larvae
	Family: Cupressaceae				
M. watchli Seitner	*Cupressus sempervirens*	2011	Sallèles du Bosc, FR	15	Female
M. watchli	*Cupressus sempervirens*	2011	Monfavet, FR	15	Female
M. watchli	*Cupressus sempervirens*	2011	Ruscas, FR	16	Female
M. watchli	*Cupressus sempervirens*	1997	Aghois Ioannis, GR	10	Female
M. bipuncatatus Swederus	*Juniperus sabina*	2011	Briançon, FR	10	Female
M. bipuncatatus	*Juniperus sabina*	2011	Pallon, FR	13	Female
M. bipuncatatus	*Juniperus sabina*	2011	Pallon, FR	10	Male
M. amicorum Bouček	*Juniperus phoenicea*	2011	Petit Luberon, FR	8	Female
M. amicorum	*Juniperus phoenicea*	2011	Luberon, FR	15	Female
M. amicorum	*Juniperus phoenicea*	2011	Luberon, FR	10	Male
M. amicorum	*Juniperus oxycedrus*	2009	Corsica, FR	10	Female
M. amicorum	*Juniperus oxycedrus*	2011	Corsica, FR	10	Female
M. amicorum	*Juniperus oxycedrus*	2011	Corsica, FR	9	Male
Parasitoids of *M. spermotrophus*					
Eurytoma sp.	-	2011	British Columbia, CAN	7	-
Mesopolobus sp.	-	2011	British Columbia, CAN	16	-

were also collected. Wild adult female *M. spermotrophus* were collected from trees located on the University of Victoria campus in Victoria, BC (48°27'42.90"N, 123°18'37.50"W). Whole insect samples were stored in 95 % ethanol at -20 °C until DNA extraction.

For 16S rRNA pyrosequencing, *M. spermotrophus* and their parasitoids were obtained in 2011 from heavily infested seed from the Mt. Newton Seed Orchard, located in Saanichton, BC (48°35'54.00"N, 123°25'56.87"W). The seeds were placed at room temperature to hasten the development of larvae and adult emergence. Larvae as well as approximately one-week-old pupae were extracted from surface sterilized seeds. Adult female *M. spermotrophus* and adult *Eurytoma* sp. were collected upon emergence about two and three weeks later, respectively. Samples of uninfested ovules were also collected from surface sterilized seeds.

2.2.2 DNA extraction

Whole insects were rinsed several times with sterile water and allowed to air dry. The samples were then placed individually into 2 mL Micro tubes (Sarstedt) with 100 µL of PrepMan Ultra Reagent (Applied Biosystems, USA) and approximately twenty 1.0mm dia. zirconia or silica beads (BioSpec Products). Samples were homogenized using the Mini-Beadbeater-16 (BioSpec Products) on maximum (3450 oscillations/min) for two 20-30 second cycles separated by 30 seconds of centrifugation at 13,000 x g. The samples were then incubated at 100 °C for ten minutes, then cooled to room temperature for one minute, then centrifuged for three minutes at 13,000 x g and transferred into new Eppendorf tubes. DNA samples used for pyrosequencing were purified by precipitation in cold isopropanol and then washed with 70 % ethanol and re-suspended in TE buffer (pH = 7.5). The NanoDrop 2000 Spectrophotometer (Thermo Scientific) was used to determine the DNA concentration and quality. The quality of the DNA extract was also checked by successful PCR amplification of the mitochondrial cytochrome oxidase subunit I (COI) gene with primers LCO1490 (5’-GGTCAACAAATCATAAAGATATTGG-3’) and HCO2198 (5’-TAAACTTCAGGGTGACCAAAAAATCA-3’) (Folmer et al. 1994), which is a commonly used DNA check for invertebrates. All DNA extracts were stored at -20 °C.

2.2.3 Directed PCR

Directed PCRs were conducted using either Invitrogen or ABM PCR Taq and reagents. A selection of targeted primer pairs were used to screen the samples for the presence of common

heritable symbionts (Table 2). The following positive controls for each common heritable symbiont primer set were used: *Drosophila neotestacea* (*Wolbachia* and *Spiroplasma* positive), *Macrosteles quadrilineatus* (*Arsenophonus* and *Cardinium* positive) and *Ctenocephalides felis* (*Rickettsia* positive). Sterile water was used as a negative control. Positive PCR products were separated on 1 % agarose gel, stained with eithidium bromide and visualized under UV light.

Five microlitres of DNA from each individual extraction within a sample subset were pooled (total of 32 pooled samples) and then screened using each primer set. If a positive PCR product was amplified from a pooled sample then each individual sample was screened for presence or absence of the corresponding symbiont using the same primer set. Positive PCR products were validated by sequencing representative amplicons in both directions. Purification and sequencing of PCR products were completed at Macrogen USA (Maryland). Forward and reverse sequences were aligned using MUSCLE and manually edited using the software Geneious (v6.1.3) (Biomatters) to create high-quality consensus sequences.

The quality of the template DNA from each pooled sample was checked by amplifying a 900bp amplicon using *Megastigmus* specific COI primers Ana (5'-TCCAAAAATTGCAAATACAGC-3') and Will (5'-TTCCTGATATAGCATTTCCTCG-3') (Auger-Rozenberg, M-A., pers. comm.). Individual samples that failed to yield amplicons for specific primers were also validated using Ana/Will. To confirm species identity, COI sequences were generated using Ana/Will for one representative female from each of the *M. amicorum* Bouček and *M. bipunctatus* Swederus samples.

In order to test for non-random associations between host sex and infection frequency, Fisher's exact tests for independence and correlation analysis were performed in R (v2.15.1) (R Development Core Team 2013).

Nucleotide sequence from the citrate synthase gene (gltA) was obtained from *Rickettsia* positive samples using the following primers: Rp877p (5'-GGGGACCTGCTCACGGCGG-3') and Rp1258n (5'-ATTGCAAAAAGTACAGTGAACA-3'), which generates a small 381bp amplicon (Roux et al. 1997). The gltA gene was used for phylogenetic analysis of *Rickettsia*.

Table 2: Targeted primers for PCR screening of *Megastigmus* spp. and parasitoids for common heritable symbionts

Target Symbiont	Target gene	Primer name	Primer sequence (5'-3')	Product size (bp)	Temperature Profile	Refs.*
Wolbachia	ftsZ	ftsZ_F1	ATYATGGARCATATAAARGATAG	1200	94°C 3:00; 35x(94°C 0:30, 54°C 0:45, 70°C 1:30); 70°C 10:00	1
		ftsZ_R1	TCRAGYAATGGATTRGATAT			
Arsenophonus	16S	ArSF	GGGTTGTAAAGTACTTTCAGTCGT	580-800	95°C 2:00; 35x(94°C 0:30, 52°C 0:30, 72°C 1:30); 72°C 5:00	2
		ArSR3	CCTYTATCTCTAAAGGMTTCGCTGGATG			
Cardinium	16S	CLOf	GCGGTGTAAAATGAGCGTG	450	94°C 4:00; 35x(94°C 0:40, 57°C 0:40, 72°C 0:45); 72°C 5:00	3
		CLOr1	ACCTMTTCTTAACTCAAGCCT			
Spiroplasma *poulsonii, kunkelii, citri*	p58	p58-f	GTTGGTTGAATAATATCTGTTG	793	95°C 3:00; 35x(94°C 1:00, 54°C 1:00, 72°C 1:30); 72°C 10:00	4
		p58-r	GATGGTGCTAAATTATATTGAC			
Spiroplasma *ixodetis*	16S	SpixoF	TTAGGGGCTCAACCCCTAACC	810	95°C 2:00; 35X (94°C 0:30, 52°C 0:30, 72°C 1:30); 72°C 5:00	2
		SpixoR	TCTGGCATTGCCAACTCTC			
Rickettsia	16S	RSSUf	GCTTTCAAAACTACTAATCTA	380	95°C 3:00; 35x(95°C 0:60, 51°C 0:45, 72°C 1:00); 72°C 5:00	5
		RSSUr	AAAAGCATCTCTGCGATCCG			
Microsporidia	18S	18S-MicroF	CACCAGGTTGATTCTGCC	~500	94°C 3:00; 30x(94°C 1:00, 53.7°C 1:00, 72°C 1:30); 72°C 5:00	6
		NemopopR	CGGTACAACGGTCTCTA			7

*References: 1: Baldo et al. 2006; 2: Duron et al. 2008; 3: Weeks et al. 2003; 4: Montenegro et al. 2005; 5: von der Schulenburg et al. 2001; 6: Baker et al. 1995; 7: Alex Ardila-Garcia and Naomi Fast, unpublished.

In order to obtain longer 16S rRNA fragments for phylogenetic analysis from the *Spiroplasma* strain infecting *Eurytoma*, general 16S rRNA amplicons were generated using the primers 63F (5'-CAGGCCTAACACATGCAAGTC-3') (Marchesi et al. 1998) and 907R (5'-CCGTCAATTCCTTTRAGTTT-3') (Schabereiter-Gurtner et al. 2003). Amplicons were then cloned using the Strataclone kit with Solopack Competent cells (Stratagene). Transformation was validated with PCR using M13F (5'- CACGACGTTGTAAAACGAC-3') and M13R (5'-GGATAACAATTTCACACAGG-3'). Eight clones were sent for sequencing and one representative *Spiroplasma* 16S rRNA sequence was used for further analysis. Attempts to clone longer *Ralstonia* 16S rRNA fragments were not successful.

2.2.4 Bacterial tag-encoded FLX amplicon pyrosequencing

Three replicates of five sample types were submitted for bacterial tag-encoded FLX 454-pyrosequencing (bTEFAP): *M. spermotrophus* larvae, pupae and adult females, *Eurytoma* sp. adults and *P. menziesii* ovules. Inhibitor removal and bTEFAP were completed by MR. DNA Laboratories (Shallowater, TX). Inhibitor removal involved the use of the PowerClean DNA Clean-up kit (MO BIO Laboratories, Inc., Carlsbad, CA) according to the manufacturer's protocol. The methods used for bTEFAP are previously described in Palavesam et al. (2012) and Shange et al. (2012) and were originally described by Dowd et al. (2008). Briefly, a single-step PCR was done using the following temperature profile: 94 °C for 3 minutes, followed by 28 cycles of 94 °C for 30 seconds, 53 °C for 40 seconds and 72 °C for 1 minute, with a final elongation step at 72 °C for 5 minutes using HotStarTaq Plus Master Mix Kit (Qiagen, Valencia, CA). The 16S universal bacterial primers 27Fmod (5'-AGRGTTTGATCMTGGCTCAG-3') and 519Rmodbio (5'-GTNTTACNGCGGCKGCTG-3') were used to amplify a 500 bp region of the 16S rRNA gene spanning the V1-V3 regions The PCR products from each of the different samples were mixed in equal concentrations and then purified using Agencourt Ampure beads (Agencourt Bioscience Corporation, MA, USA). Following the manufacturer's guidelines, sequencing was conducted using the Roche 454 FLX titanium platform (Roche, Indianapolis, IN).

Typically the 27F/519R primer set is not ideal for characterizing bacterial 16S rRNA sequence from plant tissue due to contamination by chloroplast DNA contamination (Wang et al. 2008,

Aires et al. 2012). In this case however, ovule samples were sequenced with this primer set in order to see if any trace endophytic bacteria could be found after post-sequencing removal of plastid sequences. In order to fully characterize the endophytic microbial community of the ovules, an alternative primer set or an alternative enrichment process would need to be used instead.

2.2.5 Qiime pipeline

The 454 generated Standard Format Flowgram (SFF) file was converted into a SFF text file using Mothur (v1.23.0) (Schloss et al. 2009). The open source software package Quantitative Insights Into Microbial Ecology (QIIME v1.6.0) was used to process the sequence data (Caporaso et al. 2010). The raw sequencing data was filtered using the following parameters: minimum sequence length of 100 bp, maximum sequence length of 2,000 bp and maximum homopolymer region of eight. Also, any sequences with an average quality score below 25 or any ambiguous bases were discarded. This filtering step reduced the number of total sequences from 81,207 to 60,543. The 454 data were then denoised to reduce the number of erroneous operational taxonomic units (OTUs) (Reeder and Knight 2010). Chimera detection was done independently of QIIME by implementing UCHIME through the USEARCH (v6.0.307) program (Edgar et al. 2011). The sequences were compared against the Gold database (http://www.drive5.com/usearch/manual/otupipe.html, downloaded February 13, 2013). Chimeric sequences (1,190 or 1.97 %) were gleaned from the data set.

OTUs were picked with the UCLUST method with the optimal option indicated. Similar sequences were clustered at the default level of 0.97 (Edgar 2010). Taxonomy was assigned to representative sequences using the RDP Classifier 2.2 method at the 0.9 confidence level (Wang et al. 2007). Taxonomies were based on the Greengenes database (ftp://greengenes.microbio.me/greengenes_release/gg_12_10/gg_12_10_otus.tar.gz, downloaded February 1, 2013) (Werner et al. 2012, McDonald et al. 2012).

Originally, the PyNast method was used to align the representative sequences to a pre-aligned database; however this method resulted in poor overall alignment. Alternatively, representative sequences were aligned to a Stockholm format reference of pre-aligned sequences and secondary structures using Infernal (Nawrocki et al. 2009). The aligned sequences were filtered to remove

common gap positions, with the gap filter threshold set to 0.8 and the entropy threshold set to 0.10. An approximately-maximum-likelihood phylogenetic tree was created using FastTree 2.1.3 (Price et al. 2010). An OTU table in Biom format was created and then split at the highest taxonomic ranking to remove unclassified OTUs (likely remnant chimeric sequences). Singletons were removed from the Biom table. Alpha diversity results were generated using with rarefaction depth set to 5,000.

Data exploration, visualization and analyses were performed in R (v3.0.1) (R Development Core Team 2013) on RStudio (v0.97.336) (www.rstudio.com, downloaded August 5, 2013), mainly using the *Phyloseq* R-package (v1.5.19) (Mcmurdie and Holmes 2012). Data were rarefied to an equal sampling depth of 1,962 prior to community analysis. Initial correspondence analysis and biplots were generated using the *Ade4* R-package (v1.5-2) (Dray and Dufour 2007). Principle component analysis was completed using unweighted and weighted UniFrac distances (Lozupone and Knight 2005, Hamady et al. 2010).

2.2.6 Phylogenetic analysis

A *Wolbachia* phylogeny was re-constructed using *Wolbachia* ftsZ sequence generated in this study and a sample of ftsZ sequences obtained from GenBank. A sample of ftsZ sequences obtained from an independent study of *Wolbachia* in parthenogenetic *Megastigmus* was also included (Boivin et al. 2013). Sequences were aligned using ClustalW, visually inspected and trimmed when necessary. A maximum-likelihood tree was generated using the Tamura 3-parameter model plus gamma distributed rates among sites (best substitution model identified by MEGA), with MEGA 5.1 (Tamura et al. 2011), bootstrapped 500 times. *Wolbachia* supergroups were previously defined by Casiraghi et al. (2005). Phylogenetic analysis of *Rickettsia* gltA was performed, using the same methods as the *Wolbachia* tree, with grouping conventions from Weinert et al. (2009).

A *Ralstonia* 16S rRNA phylogeny was generated using sequence from the most abundant OTU from the pyrosequencing data and from representative *Ralstonia* species and outgroup sequences, obtained from GenBank. Maximum likelihood analysis was performed as in *Wolbachia* and *Rickettsia*, except using the Tamura-Nei model with invariant sites and gamma rate distribution. A *Spiroplasma* 16S rRNA phylogeny was generated using the same methods as in *Wolbachia*,

except the maximum likelihood tree was generated using the general time reversible model and gamma distributed rates among sites.

2.3 Results

2.3.1 Inherited symbiont screening

Wolbachia was detected in *M. amicorum* and *M. bipunctatus; Rickettsia* was also detected in these two species, as well as in *M. milleri* (Table 3). DNA extractions from uninfected individuals were validated by amplifying an insect gene, COI. The prevalence of *Rickettsia* and *Wolbachia* varied among females and males of the three wasp species, with infection frequencies of 54 - 100 % (females) and 56 - 90 % (males) for *Rickettsia*, and 38 - 100 % (females) and 33 - 80 % (males) for *Wolbachia* (Table 3). There was no significant difference in the frequency of infection in males and females in any of the *Megastigmus* species infected with either *Wolbachia* or *Rickettsia*, based on Fisher's exact test of independence with respect to host sex (Table 4). For the species with mixed infections of *Wolbachia* and *Rickettsia* (i.e. *M. amicorum* and *M. bipunctatus*) there was no association between symbionts, i.e., there were not more or less single or double infections than expected by chance, based on Fisher's exact test of independence with respect to symbiont infection status (Table 5).

Arsenophonus, *Cardinium*, *Spiroplasma* and Microsporidia were not detected in *Megastigmus* samples screened using PCR with symbiont-specific primers.

2.3.2 Phylogenetic analysis

Wolbachia from *M. amicorum* and *M. bipunctatus* both fall in supergroup A (Figure 2). The ftsZ sequence from *M. bipunctatus Wolbachia* was very closely related to other group A *Wolbachia* infecting parthenogenetic *Megastigmus* (0.6% sequence divergence). *M. amicorum Wolbachia* sequences from *J. oxycedrus* and *J. phoenicea* were identical, and were very closely related to other group A *Wolbachia* infecting parthenogenetic *Megastigmus* (0.6% sequence divergence). *Rickettsia* gltA sequences from *M. amicorum* and *M. bipunctatus* were identical and fall into a group of bacteria allied with *R. felis*. *Rickettsia* from *M. milleri* Milliron is distantly related, falling into the *Bellii* group (Figure 3).

Table 3: Prevalence of *Wolbachia* and *Rickettsia* in *Megastigmus* spp. screened in this study.

Megastigmus Species	**Host plant**	**Sampling Year**	**Sampling Location**	**n**	**Sex**	**Wolbachia Positive**	**Rickettsia Positive**
M. bipuncatatus	*Juniperus sabina*	2011	Briançon, FR	10	Female	90 % (9)	100 % (10)
M. bipuncatatus	*Juniperus sabina*	2011	Pallon, FR	13	Female	38 % (5)	54 % (7)
M. bipuncatatus	*Juniperus sabina*	2011	Pallon, FR	10	Male	50 % (5)	60 % (6)
M. amicorum	*Juniperus phoenicea*	2011	Petit Luberon, FR	8	Female	100 % (8)	100 % (8)
M. amicorum	*Juniperus phoenicea*	2011	Luberon, FR	15	Female	100 % (15)	93 % (14)
M. amicorum	*Juniperus phoenicea*	2011	Luberon, FR	10	Male	80 % (8)	70 % (7)
M. amicorum	*Juniperus oxycedrus*	2009	Corsica, FR	10	Female	70 % (7)	80 % (8)
M. amicorum	*Juniperus oxycedrus*	2011	Corsica, FR	10	Female	80 % (8)	100 % (10)
M. amicorum	*Juniperus oxycedrus*	2011	Corsica, FR	9	Male	33 % (3)	56 % (5)
M. milleri	*Abies grandis*	2012	Buckley Bay, CAN	16	Female	-	75 % (12)
M. milleri	*Abies grandis*	2012	Buckley Bay, CAN	10	Male	-	90 % (9)

Table 4: Fisher's Exact Test for Independence for endosymbiont prevalence with respect to host sex

Wolbachia	**Sample**	**P-value***
	M. bipunctatus on *J. sabina*	0.7066
	M. amicorum on *J. phoenicea*	0.0852
	M. amicorum on *J. oxycedrus*	0.0478
Rickettsia	**Sample**	**P-value**
	M. bipunctatus on *J. sabina*	0.6904
	M. amicorum on *J. phoenicea*	0.0613
	M. amicorum on *J. oxycedrus*	0.0559
	M. milleri on *A. grandis*	0.3463

*Bonferroni adjusted p-value cutoff = 0.0071429

Table 5: Fisher's Exact Test for Independence for *Wolbachia* and *Rickettsia* within a host

Sample	**P-value**
M. bipunctatus on *J. sabina*	0.4421
M. amicorum on *J. phoenicea*	1.0000
M. amicorum on *J. oxycedrus*	0.7884

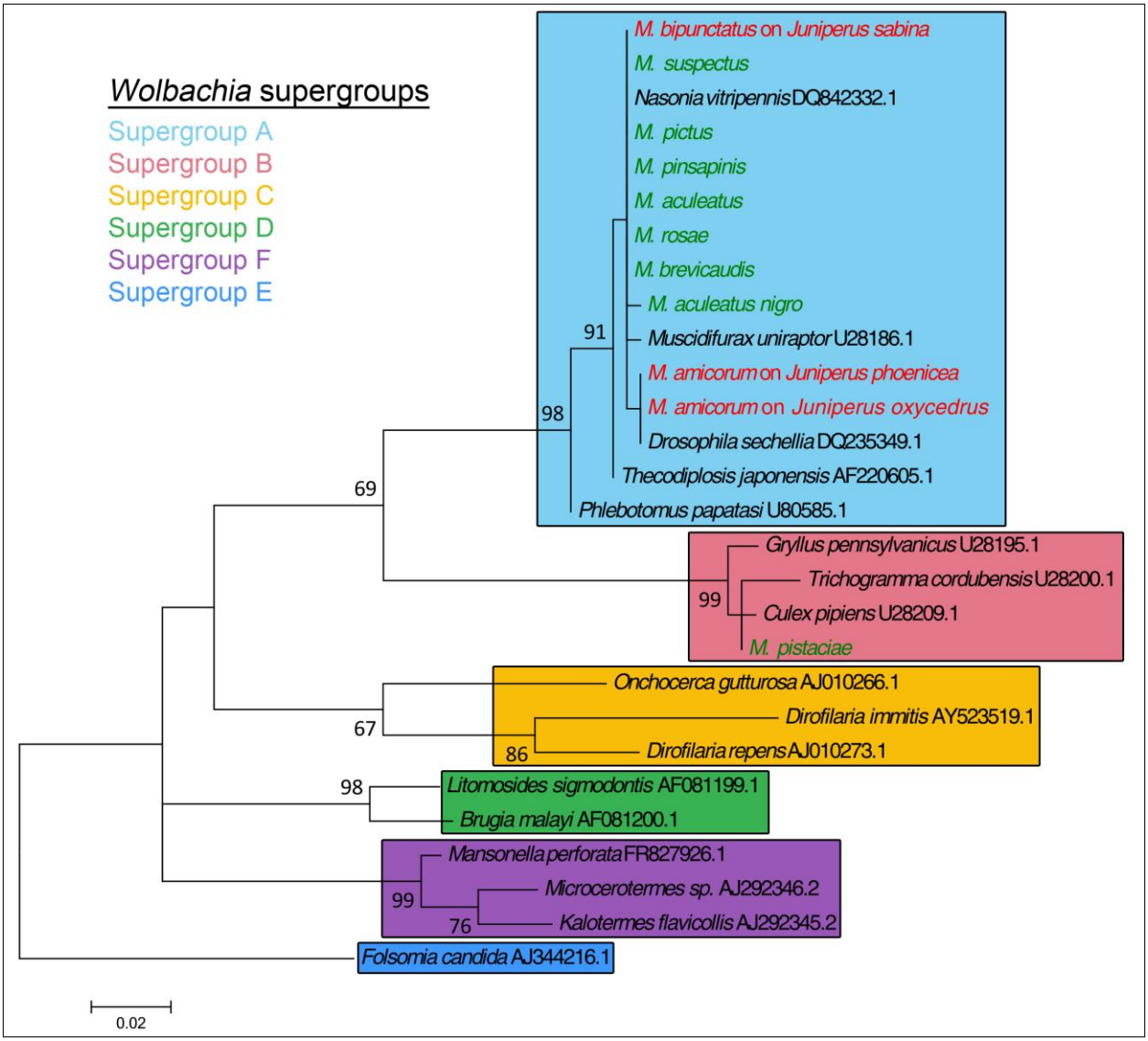


Figure 2: Maximum likelihood phylogeny for *Wolbachia* ftsZ sequence using the Tamura 3-parameter model plus gamma distributed rates among sites model. Sequences generated by this study are highlighted in red and sequences previously obtained from parthenogenetic species of *Megastigmus* (Boivin et al. 2013) are highlighted in green. Numbers next to the nodes indicate percentage of bootstrap support from 500 bootstrap replicates. Nodes without numbers received less than 65 % bootstrap support.

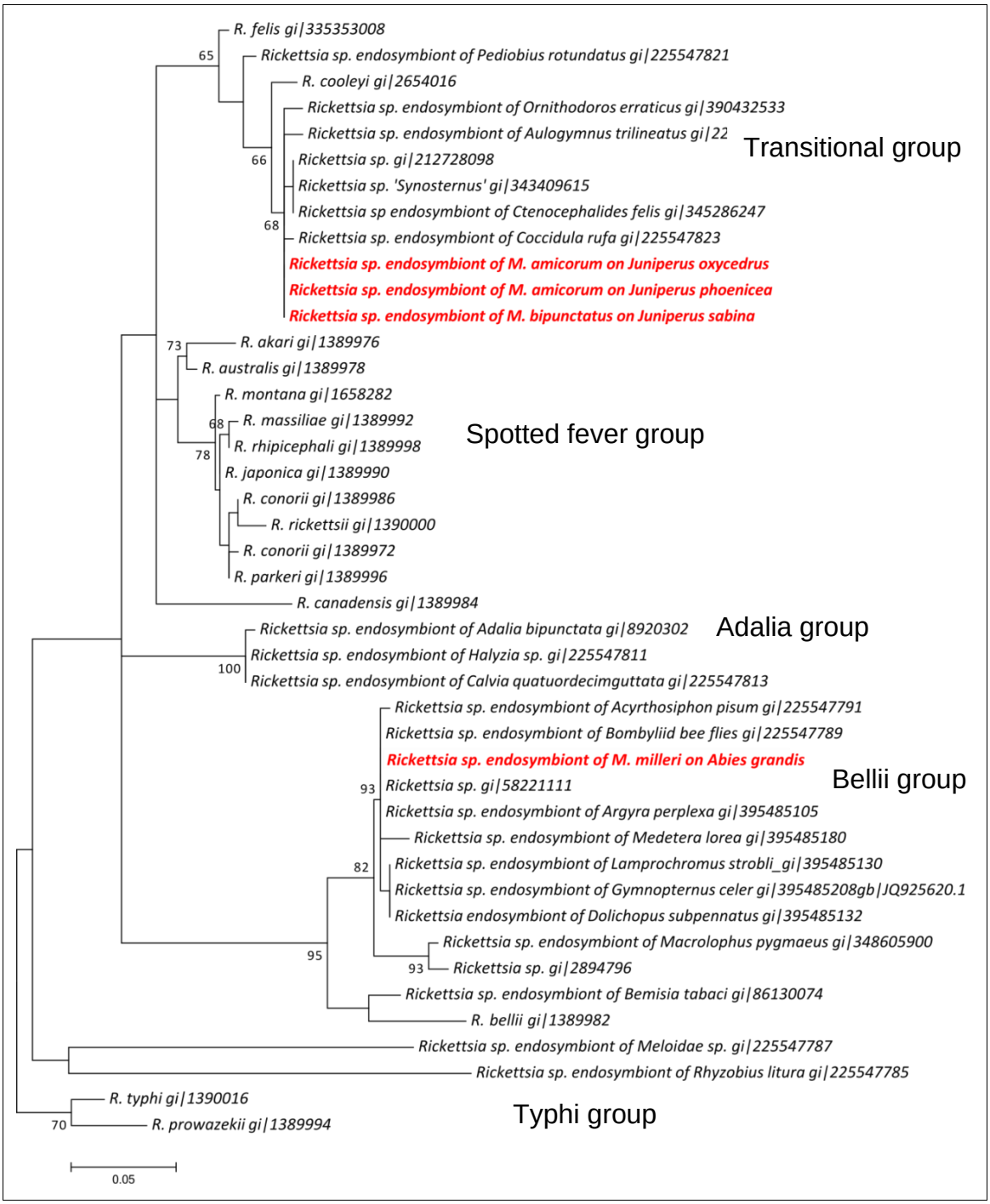


Figure 3: Maximum likelihood phylogeny for *Rickettsia* citrate synthase sequence using the Tamura 3-parameter model plus gamma distributed rates among sites model. The sequences generated by this study are highlighted in red. Numbers next to the nodes indicate percentage of bootstrap support from 500 bootstrap replicates. Nodes without numbers received less than 65 % bootstrap support.

2.3.3 Microbial associates of *M. spermotrophus*

Bacterial 16S rRNA pyrosequencing of *M. spermotrophus* (Adult female, larvae and pupae), *Eurytoma* sp. (Adult) and *P. menziesii* ovules generated 81,207 raw reads with an average length of 422bp. Quality and chimera filtering removed approximately 27 % of the reads. The assignment of OTUs resulted in 352 unique bacterial clusters after the removal of singletons. A total of 160 OTUs were assigned to the genus level. The maximum and minimum number of sequences per sample was 6,130 and 1,962, respectively (Table 6).

In order to identify possible outliers (i.e., samples that contain unusual or unexpected OTUs), the *M. spermotrophus* microbiome data were visualized using a correspondence analysis biplot (Figure 4). Correspondence analysis enables visualization of categorical data (Sourial et al. 2010), with both microbial OTUs and sample types represented in the biplot. One pupal sample (P1) and one female sample (F4) were found to be associated with distinct OTUs that did not cluster with the remaining samples. Sample P1 had a relatively elevated species richness compared to the other samples, likely originating from environmental contamination (data not shown). Sample F4 contained bacteria typical of human contamination. Subsequently these two samples were removed from further analysis.

Rarefaction analysis showed that for most of the *M. spermotrophus* samples the number of observed OTUs no longer exponentially increased after an approximate sampling depth of 3,000 sequences (Figure 5). The average sequencing depth was 3,616 sequences per sample (Table 6). The average number of observed species was 60 ± 13 and the average Chao1 species diversity estimate was 71 ± 25 (Figure 5).

The core microbiome of *M. spermotrophus* is comprised of bacterial OTUs that have a total relative abundance of 0.5 % or greater (Table 7). Associated with the *M. spermotrophus* microbiome are fifteen major OTUS assigned to five diverse bacterial classes: Betaproteobacteria, Gammaproteobacteria, Actinobacteria, Firmicutes and Alphaproteobacteria. Over 60 % of the sequences from the *M. spermotrophus* samples were assigned to the genus *Ralstonia* sp. (61.57 %). Other major OTUs were assigned to the genera *Acinetobacter* and *Corynebacterium* representing 17.20 % and 4.44 % of total relative abundance, respectively (Figure 6).

Table 6: Summary of sequence data from tag-encoded FLX 454-pyrosequencing of 16S rRNA from *M. spermotrophus*, *Eurytoma* sp. and *P. menziesii* ovule samples.

454 Sequence summary	
Number of raw input reads:	81,207
Raw length min/max/avg:	100/633/422.3
Number of filtered reads:	60,543
Filtered length min/max/avg:	72/605/394.3
Number of chimeric sequences:	1,190
Number of singletons:	183
Number of unclassified sequences:	4,932
OTU Assignments	
Number of OTUs assigned to kingdom	352
Number of OTUS assigned to phylum	255
Number of OTUS assigned to class	248
Number of OTUS assigned to order	233
Number of OTUs assigned to family	217
Number of OTUs assigned to genus	160
Seqs/sample summary	
Number of Samples:	15
Min:	1,962
Max:	6,130
Median:	2,974
Mean:	3,616
Standard Deviation:	1,415

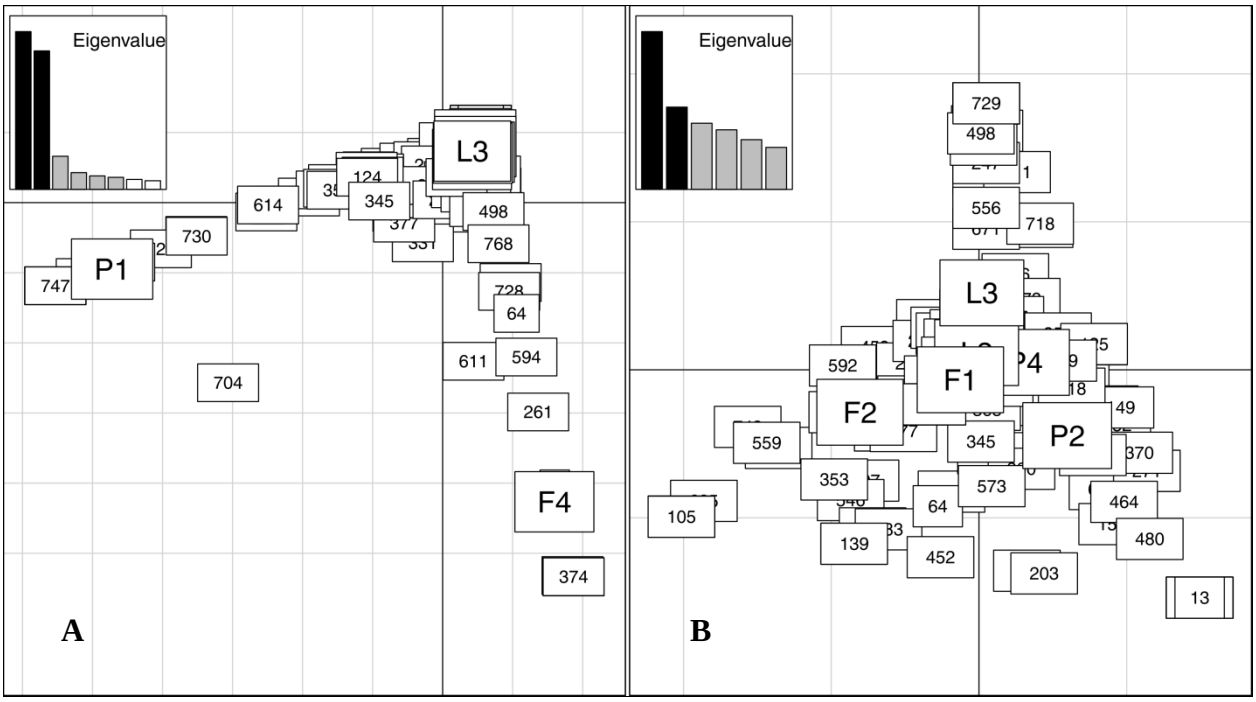


Figure 4: Correspondence analysis of the bacterial diversity associated with *M. spermotrophus* based on 16S rRNA sequences from pyrosequencing represented in a biplot, showing rows (OTUs) and columns (samples) simultaneously. Eigenvalues plotted against factor numbers are given in the top left hand corners. A) Analysis including all samples, B) analysis with outliers, F4 and P1, removed.

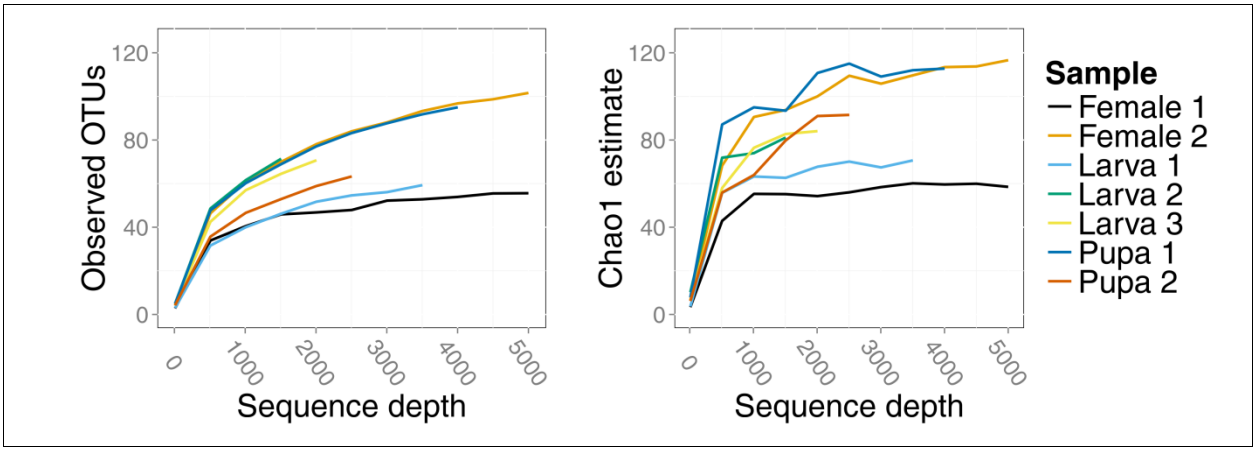


Figure 5: Observed species and Chao1 species diversity estimator rarefaction curves for bacteria associated with different life stages of *M. spermotrophus*, based on 16S rRNA pyrosequencing.

Table 7: Major bacterial OTUs associated with *M. spermotrophus* (greater than 0.5 % average relative abundance) based on 16S rRNA amplicons from pyrosequencing.

Phylum	**Class**	**Order**	**Family**	**Genus**	**Percent total relative abundance**
Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	*Ralstonia*	55.86
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	*Acinetobacter*	16.28
Actinobacteria	Actinobacteria	Actinomycetales	Corynebacteriaceae	*Corynebacterium*	3.41
Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	*Ralstonia*	3.12
Proteobacteria	Betaproteobacteria	Burkholderiales	Oxalobacteraceae	*Ralstonia*	2.59
Proteobacteria					1.29
Actinobacteria	Actinobacteria	Actinomycetales	Corynebacteriaceae	*Corynebacterium*	1.03
Actinobacteria	Actinobacteria	Actinomycetales			0.95
Proteobacteria	Gammaproteobacteria	Pseudomonadales	Moraxellaceae	*Acinetobacter*	0.92
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	*Anaerococcus*	0.79
Firmicutes					0.74
Proteobacteria					0.73
Proteobacteria	Betaproteobacteria				0.72
Firmicutes	Clostridia	Clostridiales	Clostridiaceae	*Anaerococcus*	0.52
Proteobacteria	Alphaproteobacteria	Rhizobiales	Bradyrhizobiaceae		0.50

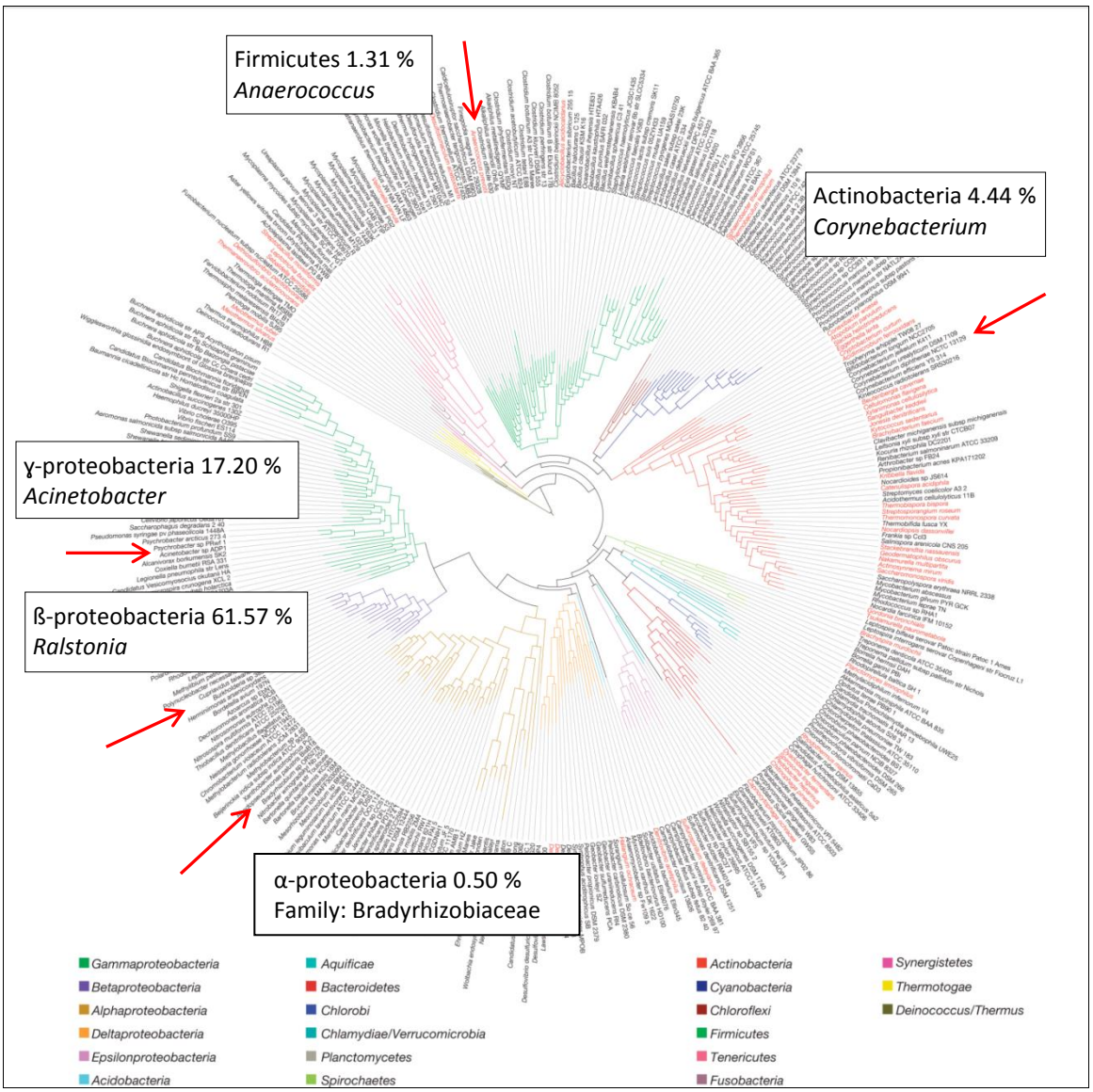


Figure 6: Concatenated bacterial maximum-likelihood phylogenetic tree with major OTUs associated with *M. spermotrophus* indicated by the red arrows. Percentage of total relative abundance is provided for each of the major phylotypes. Phylogeny was created by Wu et al. (2009).

Further investigation using BLAST searches against the Ribosomal Database Project (http://rdp.cme.msu.edu/) and GenBank's 16S ribosomal RNA sequence database revealed that all but one of the major OTUs not assigned to the genus level were actually *Acinetobacter*, *Corynebacterium*, or *Ralstonia*. The unknown Firmicutes is most closely related to *Turicibacter*, a strictly anaerobic gram-positive bacteria in the family Erysipelotrichaceae (Ludwig et al. 2008); this OTU represents 0.74 % of the total relative abundance of the 16S rRNA sequences in the *M. spermotrophus* samples.

The relative abundance of the major OTUs from the different developmental stages of *M. spermotrophus* was mostly conserved (Figure 7). The total relative abundance of OTUs from the class Betaproteobacteria (all in the genus *Ralstonia*) ranged from 46.4 - 72.3 %. One female sample contained only a very small proportion of OTUs assigned to the class Gammaproteobacteria (0.36 % relative abundance) while the total relative abundance of Gammaproteobacteria ranged from 12.7 - 33.1 % in the remaining samples. The total relative abundance of all OTUs within the class Actinobacteria (all in the genus *Corynebacterium*) ranged from 1.9 - 7.1 %.

Principle coordinate analysis (PCoA) on UniFrac phylogenetic distances was performed to compare the microbial communities associated with the different developmental stages of *M. spermotrophus* (Figure 8). Unweighted UniFrac analysis only considers the fraction of phylogenetic branch lengths shared between two samples; weighted UniFrac analysis assigns weights to the branch lengths based on abundance. The core microbiome from the different developmental stages did not form distinct clusters based on the unweighted UniFrac analysis. The core microbiomes of the different developmental stages were all tightly clustered together in the weighted UniFrac analysis.

A maximum likelihood phylogeny for *Ralstonia* was created using 16S rRNA sequence from the most abundant *Ralstonia* OTU in the pyrosequencing data set (Figure 9). Strong bootstrap support clusters the *Ralstonia* isolated from *M. spermotrophus* with the human pathogen *R. pickettii* (sequence divergence = 3.3 %).

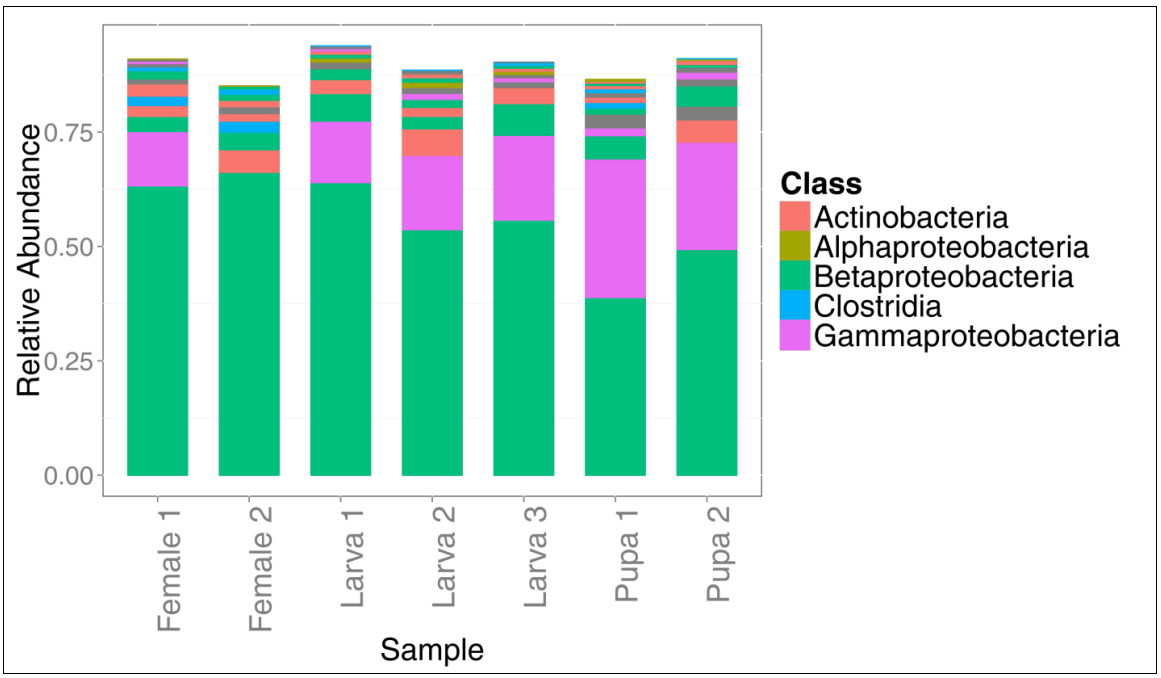


Figure 7: Relative abundance of major bacterial OTUs associated with larvae, pupae and adult *M. spermotrophus* (total relative abundance greater than or equal to 0.5 %) based on 16 rRNA sequence from pyrosequence. Unknown classes are coloured grey.

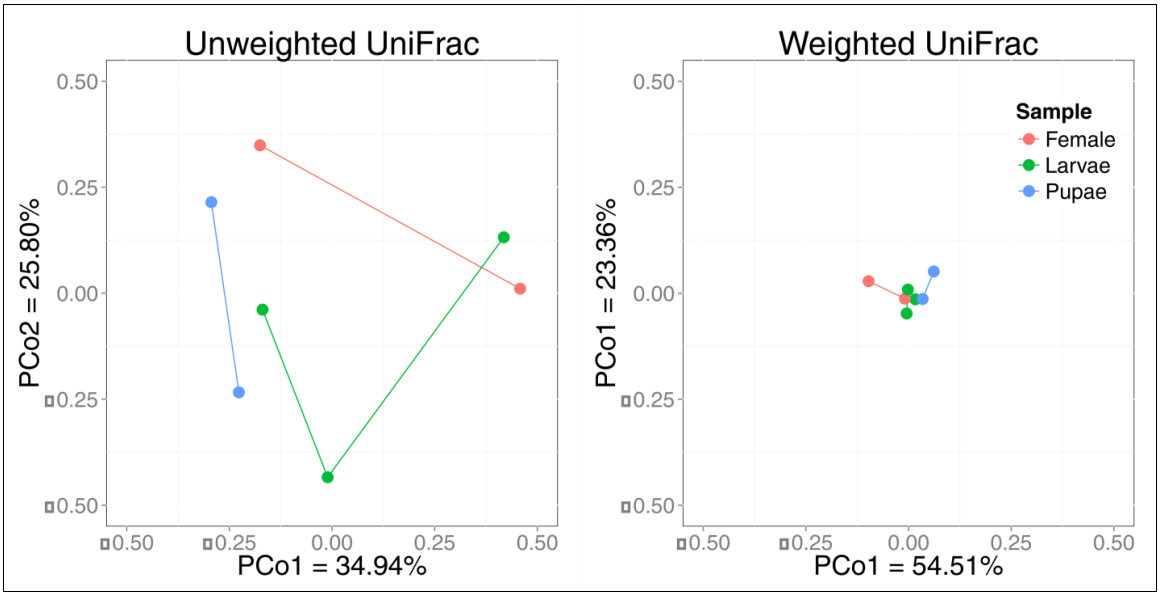


Figure 8: Analysis of phylogenetic distances (UniFrac) for all OTUs associated with different developmental stages of *M. spermotrophus* based on 16S rRNA amplicon pyrosequence.

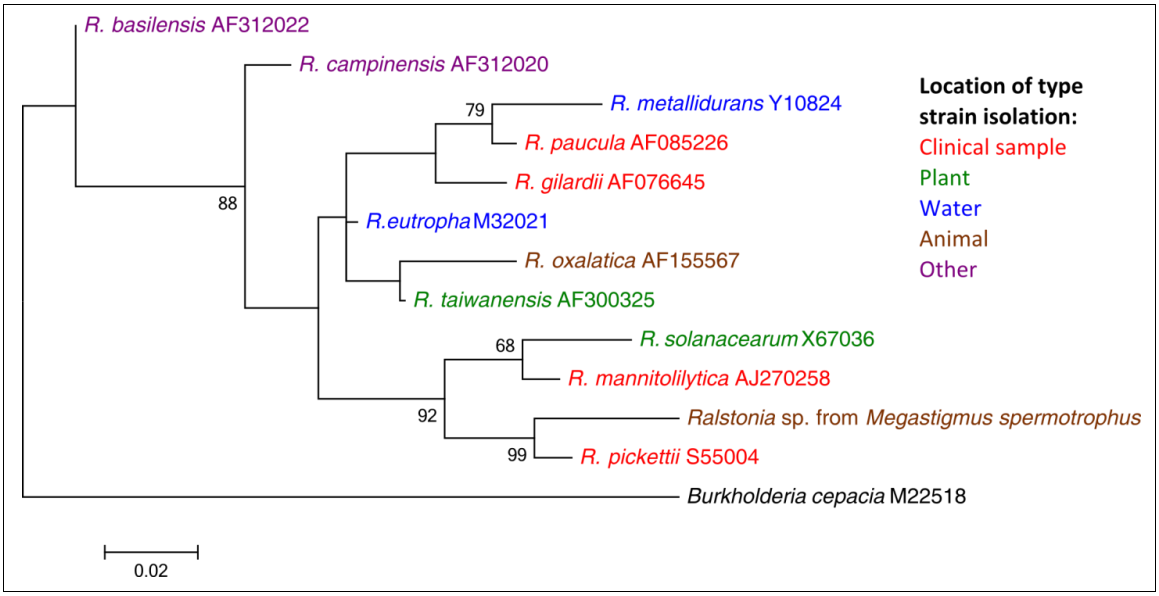


Figure 9: Maximum likelihood phylogeny for *Ralstonia* 16S rRNA sequence found using the Tamura-Nei model of nucleotide substitution with invariant sites and gamma rate distribution. Numbers next to the nodes indicate percentage of bootstrap support from 500 bootstrap replicates. Nodes without numbers received less than 65 % bootstrap support.

Ovule samples were dominated by plastid rRNA (99.0 %); the remaining OTUs included *Ralstonia* (0.8 %) and *Acinetobacter* (0.2 %). The *Eurytoma* parasitoid samples were dominated by one OTU, which is allied with inherited *Spiroplasma* in the *Ixodetis* group (Figure 10). The majority of the 16S samples were *Spiroplasma* (99.6%); the remaining OTUs were *Ralstonia*.

2.4 Discussion

2.4.1 Common heritable symbionts

My study provides the first report of *Wolbachia* from a sexual species of *Megastigmus* and the first report of *Rickettsia* from any *Megastigmus* species. Surveying a large number of *Megastigmus* revealed a patchy distribution of these inherited symbionts, with three species infected with *Rickettsia*, and two of these same species infected with *Wolbachia*. The finding of *Wolbachia* is not surprising, since it is a very common heritable symbiont that displays a variety of reproductive and mutualistic phenotypes in a diverse array of hosts. More surprising was the isolation of two divergent strains of likely non-sex ratio distorting *Rickettsia*.

Wolbachia, in the Alphaproteobacteria, is the most common intracellular bacterial symbiont of insects, and is estimated to infect 40 % of terrestrial arthropods (Zug and Hammerstein 2012). *Wolbachia* are transmitted maternally, in the egg cytoplasm, and many strains have evolved strategies to increase the frequency of infected female hosts in the population by manipulating host reproduction, for example by causing cytoplasmic incompatibility or by distorting sex ratios by killing males or inducing parthenogenetic reproduction (Werren et al. 2008).

A recent study found a perfect association between *Wolbachia* and parthenogenesis (i.e. clonal production of females) in *Megastigmus*, with 10/10 asexual and 0/15 sexual species infected (Boivin et al. 2013). Treating the asexual *Wolbachia*-positive *M. pinsapinis* with the antibiotic tetracycline restored the production of males, strongly suggesting that *Wolbachia* is the causative agent of thelytoky in asexual species of *Megastigmus* (Boivin et al. 2008, 2013). Parthenogenesis-inducing *Wolbachia* are common in Hymenoptera and have been characterized in several parasitoid (Stouthamer 1997) and several cynipid gall wasps (Plantard et al. 1998, Rokas et al. 2002).

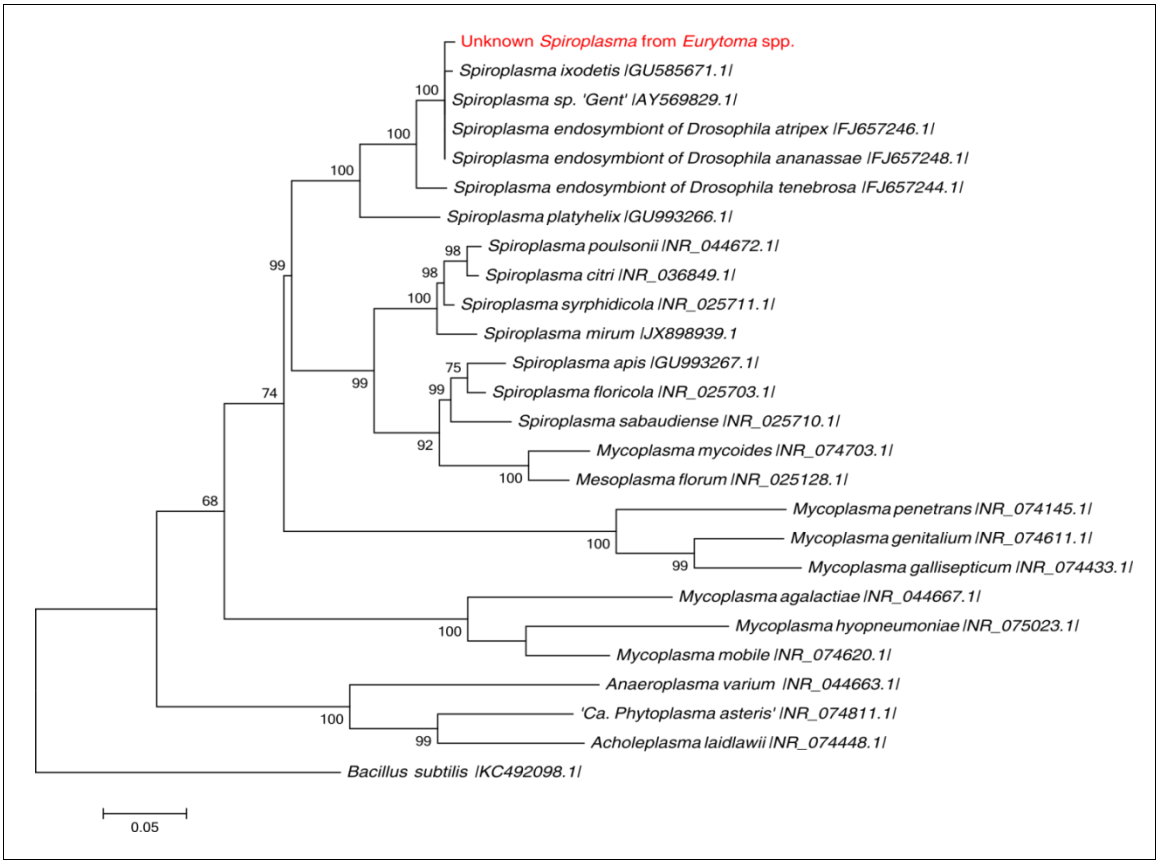


Figure 10: Maximum likelihood phylogeny for *Spiroplasma* 16S rRNA sequence found using the general time reversible model of nucleotide substitution with gamma distributed rates. The sequence generated in this study is highlighted in red. Numbers next to the nodes indicate percentage of bootstrap support from 500 bootstrap replicates. Nodes without numbers received less than 65 % bootstrap support.

Both *M. amicorum* and *M. bipunctatus* were found to be infected with *Wolbachia,* with infection frequencies ranging from 38 – 100 % and 33 – 80 % in females and males, respectively. Other samples of these species were previously collected from different localities, but PCR screens for *Wolbachia* were negative (Auger-Rozenberg, M-A., pers. comm.). Both of these *Megastigmus* species infest Cupressaceae hosts in the genus *Juniperus*; *M. amicorum* has been collected from *J. phoenicea* and *J. oxycedrus* in the West Palearctic and *J. thurifera* in Corsica; *M. bipunctatus* has been collected from *J. communis* and *J. sabina* in the West Palearctic and *J. thurifera* in Morocco, Spain and continental France (Roques and Skrzypczyńska 2003).

Based on the ftsZ gene, the *Wolbachia* strains from *M. amicorum* and *M. bipunctatus,* as well as all the strains infecting thelytokous *Megastigmus* except one, are found in supergroup A. However, ftsZ evolves very slowly and it would be useful to sequence a number of more rapidly evolving genes to determine whether strains infecting sexual and asexual *Megastigmus* are closely related. Other studies have shown that closely related *Wolbachia* strains can cause different phenotypes. In tropical butterflies of the genus *Acraea*, *Wolbachia* strains were identical in sequence at two loci, yet neither exhibited sex ratio-distortion or cytoplasmic incompatibility (Jiggins et al. 2001, 2002). The *Wolbachia* strain harboured by *M. pistaciae* is divergent from the others, as it found in supergroup B; *Megastigmus* species found on host plants in the family Anacardiaceae are believed to be quite divergent from other *Megastigmus* species. The majority of *Wolbachia* strains found in arthropods belong to supergroups A and B (Casiraghi et al. 2005).

A situation that is comparable to the presence of *Wolbachia* in both sexual and asexual species of *Megastigmus* occurs in *Trichogramma* parasitoid wasps (Chalcidoidea). In this example several species of *Trichogramma* are infected with *Wolbachia*, which causes thelytokous parthenogenesis (Stouthamer et al. 1990). However, *T. bourarachae* is a sexual species in which *Wolbachia* is known to cause increased fecundity (Girin and Boulétreau 1995, Vavre et al. 1999b). In contrast to the *Megastigmus* situation where both phenotypes of *Wolbachia* align with supergroup A, the *Wolbachia* associated with the sexual *T. bourarachae* aligns with supergroup A, while the thelytokous causing strains align with supergroup B based on partial ftsZ sequence (Vavre et al. 1999b). Grenier et al. (1998) were able to horizontally transfer parthenogenetic inducing *Wolbachia* into a novel host species of *Trichogramma*. However, in the novel host only

partial induction of thelytoky was observed, which suggests that the expression of parthenogenesis is dependent on symbiont density and/or symbiont-host interactions.

What is the phenotype caused by *Wolbachia* infecting *M. amicorum* and *M. bipunctatus*? *Wolbachia* is not likely causing sex ratio distortion in these populations because males were found to be frequently infected. The variable rate of infection in both males and females does not support the idea that *Wolbachia* may be exhibiting weak cytoplasmic incompatibility, yet further experimentation would be required to confidently rule out this mode of sexual manipulation. Alternatively, the *Wolbachia* strains associated with *M. amicorum* and *M. bipunctatus* may be conferring some selective advantage to their hosts under specific environmental conditions. For example, some *Wolbachia* infections have been shown to increase host fecundity (e.g. Weeks et al. 2007), while others have been shown to protect their hosts against RNA viruses (Teixeira et al. 2008).

In addition to the discovery of *Wolbachia* in *M. amicorum* and *M. bipunctatus*, the directed PCR screen also revealed a co-infection with *Rickettsia*, another common heritable alphaproteobacterial endosymbiont, as well as a strain of *Rickettsia* infecting *M. milleri*. Bacteria in the genus *Rickettsia* are well known for being insect-vectored vertebrate pathogens, such as the causal agents of Rocky Mountain spotted fever (*R. rickettsiae*) and typhus (*R. typhi*). However, recent surveys have uncovered many *Rickettsia* that are vertically transmitted symbionts of diverse arthropods, most of which do not feed on vertebrates (Perlman et al. 2006). Two modes of reproductive manipulation have been documented in these inherited *Rickettsia*: male-killing (von der Schulenburg et al. 2001, Lawson et al. 2001) and parthenogenesis-induction (Hagimori et al. 2006, Giorgini et al. 2010). *Rickettsia* has been isolated from gall wasps (Weinert et al. 2009) and it is known to induce parthenogenesis in some species of parasitoid wasps (Hagimori et al. 2006, Giorgini et al. 2010). Phylogenetic analysis based on the citrate synthase gene places the *Rickettsia* infecting *M. amicorum* and *M. bipunctatus* in the 'transitional' group, which also includes strains that infect the gall wasps, *Aulogymnus trilineatus* and *Pediobius rotundatus* (Weinert et al. 2009). The strain infecting *M. milleri* is found in the unrelated Bellii group.

As in the case of *Wolbachia*, it is difficult to be certain of the effect of *Rickettsia* on *Megastigmus*. The presence of infected males rules out the possibility that these strains of *Rickettsia* are sex-ratio distorters. Cytoplasmic incompatibility is not a likely option since there have been no previous reports of *Rickettsia* causing this type of reproductive manipulation. Potentially the interaction with the host could be more complex, suggestive of a balancing act between positive and negative effects on host fitness. For example, in pea aphids, *Rickettsia* infections have been shown to reduce body mass and fecundity (Chen et al. 2000, Sakurai et al. 2005), yet some isolates can confer protection against pathogenic fungi (Łukasik et al. 2013).

Rickettsia and *Wolbachia* co-infections were common in *M. amicorum* and *M. bipunctatus*. Similar co-infection has been reported in the chestnut weevil (Toju and Fukatsu 2011), the date stone beetle (Zchori-Fein et al. 2006), and the predatory bug, *Macrolophus pygmaeus* (Machtelinckx et al. 2012). In theory, the evolution of a cooperative relationship between co-existing vertically transmitted symbionts can develop if selection favors multiple infections in the daughters of multiply infected mothers (Vautrin et al. 2008, Vautrin and Vavre 2009). Recent studies have demonstrated that co-infection by vertically transmitted symbionts is frequent (Duron et al. 2008, Gottlieb et al. 2008, Toju and Fukatsu 2011). Little is known about the function of multiple associations on host ecology and evolution. Although a co-infection between *Wolbachia* and *Rickettsia* was identified in two juniper-infesting *Megastigmus* species, both symbionts are likely independent of each other based on statistical analysis.

Phylogenetic analyses of *Wolbachia* and other facultative symbionts in diverse arthropod hosts provides strong evidence of horizontal transmission over evolutionary timescales (Moran et al. 2008), with closely related symbionts infecting distantly related hosts. There are however, some examples of horizontal transmission over more recent, ecological timescales. In a study of pea aphid facultative symbiont phylogenies, Henry et al. (2013) found evidence that recent horizontal transmission events were common and that facultative symbionts had an impact on the ecological niche of the host. Since, both *M. amicorum* and *M. bipunctatus* harbour nearly identical strains of *Wolbachia* and *Rickettsia*, it seems feasible that these secondary symbionts were recently acquired by horizontal transmission. Ecologically mediated pathways, such as shared feeding and breeding sites or parasitoids are possible horizontal transmission routes for *Wolbachia* in some arthropod communities (Vavre et al. 1999a, Huigens et al. 2004, Haine et al.

2005, Sintupachee et al. 2006). Juniper seeds are shared breeding and feeding sites for *Megastigmus*, moths and mites (Roques et al. 1984). This presents an opportunity for horizontal transmission of secondary symbionts through competitive interactions within the juniper seed. There is also evidence that horizontal transmission of *Rickettsia* can be plant-mediated, with transmission via the phloem (Caspi-Fluger et al. 2012), suggesting that secondary symbionts could also be horizontally transmitted from phloem feeders attacking junipers. The patchy distribution of secondary symbionts found on juniper-infesting species of *Megastigmus* provides at least some indication that the particular habitat of these species may provide increased opportunities for horizontal transmission.

2.4.2 Microbial associates of *M. spermotrophus*

Bacterial 16S rRNA amplicon pyrosequencing was used to characterize the bacterial associates of three different developmental stages of *M. spermotrophus* (non-feeding fifth-instar larvae, pupae and recently emerged adult females). Rarefaction analysis indicated that commonly occurring bacterial associates where captured, but rare OTUs may have been missed. The sequence depth was likely sufficient for characterizing the major components of the bacterial community associated with *M. spermotrophus*. The *M. spermotrophus* species richness estimate (60 ± 13 OTUs) fell within the range of other studies of insect microbiomes. Based on a meta-analysis by Colman et al. (2012) pollenivorous and predacious Hymenoptera (bees and wasps) harbour distinct bacterial communities with the lowest level of species richness (11.0 ± 5.4 OTUs/sample) and termites harbour the highest species diversity (89.5 ± 61.2 OTUs/sample). A recent study estimated the diversity of bacteria associated with parasitoid wasps from the genus *Nasonia* ranged from 14 to 38 bacterial OTUs (Brucker and Bordenstein 2012). Many previous insect microbiome surveys, including the *Nasonia* study, were done using 16S rRNA clone sequencing rather than 454 pyrosequencing. A recent study found that even after stringent quality filtering, 454 pyrosequencing identified substantially more OTUs than traditional techniques (Kautz et al. 2013). The higher estimated microbial diversity of *M. spermotrophus* compared to other Hymenoptera may be partially explained by the increased OTU identification power of 454 pyrosequencing compared to 16S clone sequencing.

Despite a relatively high overall richness, only fifteen major OTUs are present with a total relative abundance of 0.5 % or greater (Table 7). The core bacterial community of *M.*

spermotrophus can thus be considered to have a somewhat low diversity, characterized by bacterial OTUs that are commonly found associated with insect guts. The major OTUs associated with *M. spermotrophus* can be grouped into five distinct phylotypes: Betaproteobacteria (mostly *Ralstonia*), Gammaproteobacteria (mostly *Acinetobacter*), Actinobacteria (*Corynebacterium*), Firmicutes (mostly *Anaerococcus*) and Alphaproteobacteria (family Bradyrhizobiales) (Figure 6). Most of these OTUs are related to bacteria that have been previously reported in insect guts, with *Acinetobacter* and *Corynebacterium* especially common (e.g. Brucker and Bordenstein 2012, Ishak et al. 2011). All of the major OTUs identified below the order level are bacteria that commonly occur in the environment, such as in soil (Janssen 2006) and in the rhizospere (da Rocha et al. 2009). Similar results are commonly found with microbial associates of insects. For example, the microbial symbionts of *Tetraponera* ants are closely related to nitrogen-fixing root nodule bacteria (van Borm et al. 2002).The giant mesquite bug, *Thasus neocalifornicus* Brailovsky and Barrera acquires an important mutualistic gut symbiont *de novo* every generation from the soil (Olivier-espejel et al. 2011). It is likely that many of these bacterial associates reside within the digestive tract of *M. spermotrophus* where they could contribute to the decomposition of seed storage molecules and nitrogen recycling and upgrading of nonessential amino acids. Further study would be required to identify other internal microhabitats where associates may be localized.

The *M. spermotrophus* microbiome appears to be highly conserved across development, as demonstrated by the UniFrac analysis, with all of the samples tightly grouped. The conservation of the bacterial community throughout *M. spermotrophus* development is interesting because insect guts do not provide a stable environment for microbes (assuming that the majority of these bacteria are found within the digestive tract). Like most higher Hymenoptera, the larvae of *M. spermotrophus* have a blind digestive system with the midgut and hind gut only uniting during the last larval instar. Prior to pupation all of the built-up wastes are voided in a fecal pellet, termed the meconium (Sharkey, 2007). During metamorphosis the larval midgut epithelium is discarded and replaced by a new pupal epithelium (Hakim et al. 2010). How might *M. spermotrophus* then maintain its major bacterial associates throughout development? Some insects, like true bugs, termites and cockroaches, have crypts or paunches associated with the gut that are thought to enhance persistence of the microbiota (Engel and Moran 2013). This

physiological feature is not well characterized in the Hymenoptera, with the exception of some ants (Bution and Caetano 2010). Some insects that lack specialized gut structures, such as grasshoppers and caterpillars, have been shown to harbour an indigenous microflora despite a rapid rate of food throughput (Dillon and Dillon 2004). In contrast to *Megastigmus*, a recent survey of microbial associates of three *Nasonia* species found that bacterial species richness increased with development (Brucker and Bordenstein 2012).

A single OTU assigned to the genus *Ralstonia* comprised over 50 % of all sequences from the *M. spermotrophus* samples. *Ralstonia* was also found to be associated with Douglas-fir ovules and the parasitoid, *Eurytoma*. The high abundance and persistence of *Ralstonia* throughout host development is a strong indicator that this bacterium is an important associate of *M. spermotrophus*. The genus *Ralstonia* contains species from ecological diverse niches, such as the plant pathogen *R. solanacearum*, the opportunistic human pathogen *R. pickettii* and the environmental isolate *R. eurytropha* (Brenner et al. 2005). A maximum likelihood phylogeny placed *M. spermotrophus* associated *Ralstonia* in a cluster with the human pathogen *R. pickettii*. To my knowledge, my study is the first report of *Ralstonia* being a very abundant and likely important component of an insect microbiome.

Ralstonia spp. have been previously reported from microbial surveys of insects, including the cotton bollworm (not published; accession # EU124821), *Bartonella*-positive fleas (Jones et al. 2008) and an omnivorous carabid beetle (Lundgren and Lehman 2010) . Recently, Husnik et al. (2013) also report the horizontal transfer of one *Ralstonia* gene into the genome of the mealybug *Planococcus citri*. Also, *R. oxalatica* was isolated from the alimentary canal of an Indian earthworm (Vaneechoutte et al. 2004). A recent meta-analysis of 16S clone-library studies of insect associated microbes found that Betaproteobacteria contributed over 50 % to all sequences from Hymenoptera (Colman et al. 2012). The most common bacterial phylotype identified from solitary bee species, was a Betaproteobacteria from the genus *Burkholderia* (Martinson et al. 2011), which is closely related to *Ralstonia*. *Burkholderia* spp. have also been identified as important mutualists of some phytophagous true bugs (suborder Heteroptera), where they reside in gut crypts (Kikuchi et al. 2005, 2007, 2011, Olivier-espejel et al. 2011, Shibata et al. 2013).

The *Ralstonia* associated with *M. spermotrophus* could potentially play a nutritional role. Many insects subsist on nitrogen-limited diets and mutualistic microbes have been suggested to promote increased availability of nitrogen through a variety of ways (Douglas 2009). The developing *M. spermotrophus* larva consumes the megagametophyte, which contains all of the seed storage reserves, primarily in the form of triacylglycerols and nitrogen rich proteins (King and Gifford 1997, Stone and Gifford 1999). Parasitism by *M. spermotrophus* results in the formation of a nutrient sink, in which the larva and associated microbes are nourished by storage reserves of the megagametophyte that were originally intended to provide nourishment for a developing seedling or be re-absorbed by the mother plant in the event of megagametophyte abortion.

In Loblolly pine, more than half of the nitrogen in megagametophytes comes from the amino acid arginine (Todd and Gifford 2002). Insects use the enzyme arginase to hydrolyze arginine into ornithine and urea (Pant 1988). Excretion of urea would result in the substantial loss of nitrogen, especially since larvae must undergo extended periods of diapause. Very few insects are known to produce urease, the enzyme required to convert urea into ammonium for subsequent amino acid biosynthesis (Rosenthal et al. 1977). *Ralstonia* or other microbial associates of *M. spermotrophus* might play an important role in nitrogen recycling by producing urease or other key enzymes missing from the host genome. Nitrogen recycling by symbionts has been shown to be important during diapause in the shield bug, *Parastrachia japonensi* (Kashima et al. 2006). Also, *Blattabacterium* sp., a maternally transmitted symbiont found in the fat bodies of cockroaches, can recycle nitrogen from urea and ammonia, into glutamate, using urease and glutamate dehydrogenase (Sabree et al. 2009).

Ralstonia's high prevalence throughout *M. spermotrophus* development also suggests that the association is likely very important to the host. It is also tempting to speculate that this symbiont could potentially play a role in plant manipulation. Another *Ralstonia* species, *R. taiwanensi*, has been shown to be capable of nodulating and fixing nitrogen in *Mimosa* spp. The ability to nodulate roots implies the ability to manipulate plant physiology with phytohormones, such as CK and IAA. It is unknown if the *M. spermotrophus*-associated *Ralstonia* contains any genes necessary to manipulate Douglas-fir seed development.

Now that *Ralstonia* has been identified as a likely symbiont of *M. spermotrophus*, further surveys would be helpful in determining its prevalence in other populations of *M. spermotrophus*, in other *Megastigmus* species, and in associated plants. The development of a PCR screening tool would be required to conduct such widespread surveys, since pyrosequencing can be cost prohibitive on a large scale. Furthermore, a specifically designed primer set could be linked to fluorescent probes to conduct fluorescent *in situ* hybridization to localize *Ralstonia* within *M. spermotrophus.* Full-length 16S rRNA sequences and other genes would also be useful in understanding the phylogenetic affinities of this strain.

Despite being respectively dominated by plastid and *Spiroplasma* symbiont sequences, my ovule and *Eurytoma* parasitoid samples also contained the major OTUs (i.e. *Ralstonia* and *Acinetobacter*) found in *M. spermotrophus*. The presence of major OTUs from all three sample types provides clues to the distribution and transmission of the *Megastigmus* microbiome. This suggests that the major components of the *M. spermotrophus* microbial community could be derived from the environment, which, for developing wasps, is the ovule. The association of *Ralstonia* with the parasitoid *Eurytoma* is not surprising, since *Eurytoma* ultimately consumes *M. spermotrophus* and lives within the ovule, which are both associated with *Ralstonia*. Also, *Acinetobacter* and *Corynebacterium* have been previously observed and cultured from surface sterilized seeds and ovules (Mundt and Hinkle 1976, Mukhopadhyay et al. 1996, Hallmann et al. 1997). This suggests that *Ralstonia* and *Acinetobacter* may be present on the surface of and maybe even inside Douglas-fir ovules. Some insects have evolved mechanisms to transmit non-heritable microbial associates from one generation to the next, such as egg smearing (Hosokawa et al. 2007) or the selective uptake of insect gut bacteria from the environment every generation (Kikuchi et al. 2007, Olivier-espejel et al. 2011). Further analysis would be required to determine if the microbiome of *M. spermotrophus* is actively transmitted from mother to offspring, or acquired from endophytic bacteria.

2.5 Conclusions

In this study two different approaches were used to survey *Megastigmus* for microbial symbionts. Both approaches provided novel findings. The directed PCR screens identified the presence of two common heritable symbionts, *Wolbachia* and *Rickettsia*; these are not distorting

sex ratios in the species in my study. Pyrosequencing was used to characterize the core microbiome of *M. spermotrophus*, which is dominated by *Ralstonia*, a microbe that has never been characterized as an important microbial associated of an insect. Interestingly, *Ralstonia* was also present in ovule and *Eurytoma* samples, indicating its prevalence within the niche of the ovule and potential horizontal transmission route from host to parasitoid. Many new questions are inspired by these findings, such as, how is the microbiome of *M. spermotrophus* maintained and transmitted? How widespread is the association with *Ralstonia*? What is the effect of heritable symbionts in sexual *Megastigmus*?

This initial characterization of microbial associates of *Megastigmus* did not provide any insight into the potential involvement in host manipulation, although the maintenance of a consistent microbiome from larvae to adult suggests that microbes may be vital to the development and reproduction of *M. spermotrophus*. The diet of *M. spermotrophus* consists of nitrogen, protein and lipid-filled megagametophyte tissue, making it balanced compared to other herbivorous insects. Therefore, the microbial associates of *M. spermotrophus* do not likely play a role in supplementing this already rich diet with missing essential nutrients. Alternatively, microbial associates may be important for recycling the nitrogenous waste product urea, especially since *M. spermotrophus* must survive periods of prolonged diapause. *Ralstonia* may not be a key associate of *Megastigmus* species in general, but rather a microbe that is found in the seed environment that encodes enzymes required for the catabolism of seed storage molecules or other essential pathways required for the seed feeding life style of *M. spermotrophus*.

Chapter 3. *DE NOVO* TRANSCRIPTOME ASSEMBLY AND PUTATIVE VENOM DISCOVERY IN THE DOUGLAS-FIR SEED CHACLID, *MEGASTIGMUS SPERMOTROPHUS* (HYMENOPTERA: TORYMIDAE).

3.1 Introduction

Hymenoptera are among the most diverse and successful terrestrial animals on the planet. This 'mega-diverse' insect order includes important and well-known groups, such as ants, wasps, and bees. Parasitoid wasps represent the most diverse lineage within the Hymenoptera. A parasitoid characteristically has a free-living adult stage and a juvenile stage that develops on or within an animal host (usually another insect), ultimately killing it (Eggleton and Belshaw 1992). Parasitoids have evolved to attack a wide range of hosts, from spider egg cases to aquatic insects and butterfly eggs (Lasalle and Gauld 1991). Endophytophagy, the habit of feeding on plant tissues from within, has evolved independently within several lineages of Hymenoptera. Leaf-mining, leaf-rolling, gall-induction, stem-boring and wood-boring are feeding habits found among the most basal Hymenoptera, the Symphyta or sawflies (e.g. Roskam 1992). Plant parasitism has evolved independently numerous times from parasitoid wasp lineages. For example, we find the secondary evolution of galling in the Braconidae, Cynipidae and Chalcidoidea, as well as seed-feeding in the Chalcidoidea (Austin and Dangerfield 1998, Ronquist and Liljeblad 2001, Munro et al. 2011). Work on the phylogenetic relationships among and within the major parasitoid lineages is still ongoing, making it difficult to understand the key transitions that lead to phytophagy (Lasalle 2005, Munro et al. 2011).

Megastigmus spermotrophus Wachtl is a seed-feeding chalcid (Torymidae) that is a pest of Douglas-fir, *Pseudotsuga menziesii*. *M. spermotrophus* lays its egg within developing ovules and is able to redirect the development of unfertilized ovules that would normally abort (von Aderkas et al. 2005a). After the egg hatches the larva consumes the developing plant embryo, yet the megagametophyte continues to accumulate storage products on which the larva feeds (von Aderkas et al. 2005b). *M. spermotrophus* is able to co-opt the conifer female reproductive tissue for its own reproductive success at the expense of the host, demonstrating a unique method of manipulating seed development.

How *M. spermotrophus* alters Douglas-fir seed development is unknown. There are a number of possible mechanisms, such as, larval salivary secretions or modification of the host genome as demonstrated by the gall-inducing bacterium, *Agrobacterium tumefaciens*. Data from hormone profiling suggested that the failure of the megametophyte to abort in unpollinated infested treatments may be partially explained by changes in cytokinins (Chiwocha et al. 2007). Cytokinins and other phytohormones have been shown to be involved the development of insect galls and green islands caused by leaf-mining insects (Mapes and Davies 2001a, 2001b, Giron et al. 2007, Yamaguchi et al. 2012). Microbial symbionts have also been shown to be involved in cytokinin mediated host manipulation by a leaf-mining caterpillar (Kaiser et al. 2010, Body et al. 2013).

Another possibility is that *Megastigmus* uses venom to manipulate its hosts. Many animals are venomous including snakes, lizards, jellyfish, sea anemones, sea urchins, sea snails, mammals, spiders, centipedes, scorpions, fish and insects (Fry et al. 2009).There has been some controversy as to how to best define venom, and it has been argued that the traditional definition (a toxic fluid that causes sudden paralysis or death of prey or host) does not adequately describe many venoms, such as those produced by endophytophagous and parasitoid wasps. Fry et al. (2009, 2012) have proposed a modernized definition of venom that is broad enough to include all hymenopteran venoms. This new definition defines venoms as secretions that are produced in specialized glands and delivered into a target by a puncture and release, which facilitate feeding or defense by interfering with normal physiological or biochemical processes. Venoms serve a diverse range of functions in the Hymenoptera. Bees, some vespid wasps and ants use venoms in defense (van Marle and Piek, 1986) but other vespid wasps use venom to kill prey (Robertson 1968). Parasitoid wasp venoms are known to disrupt host cells or tissues, enhance other virulence factors, induce paralysis, modify host metabolism and physiology, interfere with host development and/or suppress the host immune response (Vinson and Iwantsch 1980, Moreau et al. 2002, Moreau and Guillot 2005, Rivers et al. 1999). Little is known about the composition of parasitoid venom, although recently, several large-scale transcriptomic and/or proteomic surveys have been performed (Parkinson et al. 2003, Vincent et al. 2010, de Graaf et al. 2010, Zhu et al. 2010, Dorémus et al. 2013, Colinet et al. 2013). These studies have shown that parasitoid wasp venoms are complex and diverse. They have many components, including small peptides,

neurotoxins, amines and larger enzymes (Asgari and Rivers 2011). In the last two decades there has been a surge in venom-based drug discovery programs (King 2011). With rapid advances in next generation sequencing platforms we will likely see continued drug-bioprospecting of unstudied venomous lineages for novel drug compounds (Casewell et al. 2013).

Given the importance and diverse functions of venoms within the Hymenoptera, it would be surprising if venoms were not involved in the manipulation of host plant tissues by phytophagous wasps. In the case of *M. spermotrophus*, we hypothesize that venomous secretions may play a role in early host manipulation (i.e., the redirection of unfertilized ovules), potentially through interference of normal phytohormone pathways. At least some evidence exists to support the notion that gall-inducing wasps produce ovipositional secretions and that these secretions are associated with the induction of galls in sawflies (Tenthredinidae), fig-wasps (Agaonidae) and cynipid wasps (Cynipidae) (McCalla et al. 1962, Price 1992, Kjellberg et al. 2005, Leggo and Shorthouse 2006). Such ovipositional secretions may be considered venoms based on the above definition, since they are introduced through a wound caused by the ovipositor and interfere with normal physiological and biochemical processes to facilitate feeding.

Colleterial fluid secreted by the accessory glands of the female genitalia of willow-gall sawflies has been found to be important in the formation of the procecidium, or pre-formed gall (McCalla et al. 1962, Smith 1970, Price 1992). Early work by Hovanitz (1959) focused on a leaf galler in the genus *Pontania*, in which he demonstrated that the presence of the egg was not necessary for the first phase of gall development. In *P. proxima* the colleterial fluid is sufficient for complete gall growth, where the larva hatches only after the gall is fully developed (Rohfritsch 1992). However, galls caused by other *Pontania* species require additional stimulus from larval feeding in order for the gall to reach maximum size (Smith 1970). An extremely high concentration of t-zeatin riboside, an active cytokinin, was recently found in the glands of adult *Pontania* sawflies, suggesting that plant hormone analogs may be an important component of colleterial secretions (Yamaguchi et al. 2012).

The initiation of cynipid galls is believed to involve several factors including wounding by the ovipositor, application of ovipositional fluids, activity of the egg and activity of freshly hatched larvae (Rohfritsch 1992, Leggo and Shorthouse 2006). However, the role of ovipositional

secretions in cynipid gall development is unclear, since other authors report that gall induction results from egg and larval secretions and not from maternal secretions (Stone et al. 2002). Leggo and Shorthouse (2006), observed in *Diplolepis triforma* Shorthouse & Ritchie that the part of the egg which was embedded into the plant tissue was covered in an "amorphous [and] dark-staining substance", and that a cluster of dense cells rapidly appeared in this area, which continued to proliferate for 2-3 days. In early induction of cynipid galls the cells directly adjacent to the egg lyse to form a chamber prior to the larva hatching in most species (Csóka et al. 2005). It has also been suggested that the ovipositional secretions in Cynipidae form a mucopolysaccharide layer around the egg that may have a protective role against noxious effects of the surrounding plant tissue (Kraińska 1966). According to a recent study on the morphological evolution of the venom apparatus from cynipoid wasps, most phytophagous species have a larger venom apparatus than inquilines and parasitoids (Vårdal 2004, 2006).

Ovipositional secretions have also been found to be associated with gall induction in fig-wasps. Many species of fig-pollinating wasps induce galls on fig ovules (Weiblen 2002). The female wasp lays an egg in between the integument (ovule covering) and the nucellus (where the embryo sac develops) along with several drops of secreted fluid. The gall rapidly forms from the nucellus tissue, probably due to the drops of fluids, before the egg hatches (Kjellberg et al. 2005). The association of ovipositional secretions and gall induction by other chalcid families has not been very well studied. A very early investigation of the internal anatomy of a phytophagous chalcid from the genus *Harmolita* revealed the presence of a well-developed poison apparatus, leading James (1926) to speculate that secretions from the poison apparatus were injected during oviposition and that the fluid initiated and/or caused the gall to form.

3.1.1 Objectives

The focus of this chapter is to identify putative venom components that are highly expressed in female *M. spermotrophus*, which may be injected into the developing megagametophyte during oviposition, with the long-term objective of characterizing mechanisms of early host manipulation in the *M. spermotrophus*/Douglas-fir association.

In order to identify putative maternal venom components, I took advantage of two resources. Recently, *Nasonia vitripennis*, became the first parasitoid wasp and chalcid to have its genome

sequenced (Werren et al. 2010). *N. vitripennis* is an ectoparasitoid of flesh fly pupae. Supplemental to the genome, a recent study identified 79 constituents of *Nasonia* venom, obtained by a combination of bioinformatics and proteomics (de Graaf et al. 2010). Both *N. vitripennis* and *M. spermotrophus* belong to the superfamily Chalcidoidea. The availability of a sequenced genome combined with a diverse set of *N. vitripennis* venom protein sequences provided an excellent tool to investigate the possibility that *M. spermotrophus* may share homologous venom components.

In order to identify *Megastigmus* venom components, I used RNA-seq. This method combines next generation sequencing with transcriptomics. RNA-seq is emerging as a revolutionary tool in the study gene expression (Wang et al. 2009). Additionally, advanced computational methods can be used to assemble transcriptomes *de novo,* in the absence of a reference genome. *De novo* transcriptome assembly generates valuable genomic information and permits gene expression analysis in non-model organisms (Bräutigam et al. 2011). In this study the transcriptome of *M. spermotrophus* was assembled *de novo*. Bioinformatics were used to identify potential candidate venom constituents from the transcriptome, based on homology with the *N. vitripennis* venom database. Differential expression analysis was used to identify putative venom transcripts that were highly expressed in adult female wasps, compared to adult males and larvae.

3.2 Materials and Methods:

3.2.1 RNA extraction

Adult *M. spermotrophus* males and females were collected upon emergence and larvae were extracted from heavily infested seed from the Mt. Newton Seed Orchard, located in Saanichton, BC (48°35'54.00"N, 123°25'56.87"W). Wild females were collected from trees located on the University of Victoria campus in Victoria, BC (48°27'42.90"N, 123°18'37.50"W). All insects were collected on liquid nitrogen and then stored at -80 °C. Approximately 10-20 individuals were placed into 2 ml Micro tubes (Starstedt) with one volume buffer RLT (Qiagen), 1/100 volume beta-mercaptoethanol and three 3.5mm dia. glass beads (BioSpec Products). Samples were homogenized using the Mini-Beadbeater (BioSpec Products) at half-speed for 90 seconds. The homogenate was centrifuged at 1,300 x g for 3 minutes. Total RNA was extracted using RNeasy (Qiagen), followed by on-column DNase digestion and RNA cleanup, using the

manufacture's guidelines. Next, the RNA extract was purified using an isopropanol precipitation followed by a 100 % ethanol wash and then re-suspended in RNase-free water. The RNA extract was separated on a 1 % agarose gel stained with SYBR Safe (Invitrogen) and visualized under UV light. The RNA quality and quantity was determined using a Nanodrop 2000 instrument (Thermo Scientific) and RNA quality was further analyzed using an Experion Electrophoresis Station (Bio-Rad).

Complimentary DNA library construction with oligo(dT) primers, library fragmentation, size exclusion purifications (target average sequence length of 300 bp) and sequencing on the Illumina sequencing platform (HISeq 2000) were conducted by the BC Cancer Agency Genome Sciences Centre, Vancouver, Canada.

3.2.2 Short read filtering and *de novo* assembly

Short reads were first quality filtered with Trimmomatic (v0.22) (Lohse et al. 2012) with the following parameters: minimum leading quality of three, minimum trailing quality of 20, minimum read length of 36 and sliding window of four bases with a minimum quality of 20. Filtered reads were then assessed using FASTQC (v0.10.1) to verify quality improvements (Andrews 2010). The short reads were assembled *de novo* using the trans-ABySS (v1.3.2) pipeline using k-mer values of 30, 35, and even values from 52-96 (Robertson et al. 2010). The assembly was further clustered using CD-HIT-EST with the default sequence identity threshold of 0.95 (v4.6) (Li and Godzik 2006). Additional clustering was performed by using TIGR-TGICL (Pertea et al. 2003) with the Cap3 specific overlap percent identity cut-off set to 98. Only contigs larger than 100 bases were used in subsequent analysis.

3.2.3 Annotation

The *M. spermotrophus* transcriptome contig set was annotated with BLASTX (v2.2.27+) against the NCBI non-redundant (nr) database with an E-value cut-off of 10^{-5}. Any contigs without BLASTx hits were then annotated with BLASTn using the NCBI nucleotide database. Corresponding taxonomic information was linked to each annotation by mapping the taxon identifier to the NCBI taxonomy database.

A query of sixty-four *N. vitripennis* venom protein sequences including proteases/peptidases, protease inhibitors, carbohydrate metabolism, DNA metabolism, glutathione metabolism, esterases, recognition/binding, others and unknowns was obtained from GenBank (Table S1). The venom proteins included in the query were originally generated by de Graaf et al. (2010) using both bioinformatic and proteomic approaches. In the bioinformatic approach a query of 383 protein sequences from previously known adult hymenopteran venom proteins was used to identify venom protein homologs from the *N. vitripennis* genome using BLASTp. In the proteomic approach crude *N. vitripennis* venom was analyzed using two methods of two-dimensional liquid chromatography-mass spectrometry. In order to identify putative venom transcript homologs, the *N. vitripennis* venom protein query was compared to the *M. spermotrophus* transcriptome using tBLASTn, with an E-value cut-off of 10^{-15}.

Exploratory plots were constructed with R (v3.0.1) (R Development Core Team 2013) in RStudio (v0.97.551) (www.rstudio.com, downloaded August 5, 2013).

3.2.4 Microsporidia investigation

Wild adult female *M. spermotrophus* were collected from trees located on the University of Victoria campus in Victoria, BC (48°27'42.90"N, 123°18'37.50"W). Whole body insects were flash frozen in liquid nitrogen and stored at -80°C. DNA was extracted from whole bodies using the Prepman/beadbeater method outlined in Chapter 2. Samples were screened for the presence of microsporidia using 18S ribosomal RNA primers: 18-MicroF (5'-CACCAGGTTGATTCTGCC-3') (Baker et al. 1995) and NemopopR (5'-CGGTACAACGGTCTCTA-3') (Alex Ardila-Garcia and Naomi Fast, unpublished) using ABM PCR Taq and reagents. PCR reactions were conducted using the following temperature profile: 94 °C for 3 minutes, followed by 30 cycles of 94 °C for 1 minute, 53.7 °C for 1 minute and 72 °C for 1.5 minutes, with a final elongation step at 72 °C for 5 minutes. The quality of the template DNA was verified by PCR amplification of the COI gene with primers LCO1490 (5'-GGTCAACAAATCATAAAGATATTGG-3') and HCO2198 (5'-TAAACTTCAGGGTGACCAAAAAATCA-3') (Folmer et al. 1994).

To obtain a longer Microsporidia 18S rRNA sequence for phylogenetic analysis, amplicon products from 18-MicroF and 1537r (5'-TTATGATCCTGCTAATGGTTC-3') (Baker et al.

1995) were cloned using the Strataclone kit with Solopack Competent cells (Stratagene).

A phylogeny was re-constructed using 18S rRNA sequence from Microsporidia infecting *Megastigmus* and a broad sample of microsporidia 18S rRNA sequences collected from GenBank. Sequences were aligned using ClustalW, visually inspected and trimmed when necessary. A maximum-likelihood tree was generated using a general time reversible and gamma distributed rates among sites model of nucleotide substitution, with MEGA 5.1 (Tamura et al. 2011). Branch support was assessed by bootstrapping the data 500 times.

3.2.5 Differential expression

Transcript expression was quantified using the RSEM software package (v1.2.0) (Li and Dewey 2011), aligning forward reads only and providing a mean fragment length of 300 bp. Only transcripts with NCBI annotations were considered for expression analysis. Expected count values were normalized using the conditional quantile normalization (CQN) R package (v1.7.0) (Hansen et al. 2012) . Differential expression analysis was implemented using the non-parametric statistical analysis package Noiseq (v2.0.0) (Tarazona et al. 2011). The Noiseq-sim feature was utilized to simulate technical replicates with the following parameters: size of simulated samples equal to twenty percent of sequencing depth, five simulation replicates and allowance of two percent variability. Differential expression probability was increased from 0.8 to 0.9 to account for the lack of technical replicates. Bioinformatics packages were implemented using R (v3.0.1) in RStudio (v0.97.551).

3.2.6 Aspartylglucosaminidase protein phylogeny

A phylogeny was re-constructed using representative translated aspartylglucosaminidase (AGA) sequences from the transcriptome and a sample of insect AGA protein and translated nucleotide sequences from GenBank. The *Nasonia* genome and sequenced genomes of bees and ants (Munoz-Torres et al. 2011) were scanned to ensure all AGA copies were included in the phylogeny. The AGA phylogeny was generated using similar methods to the Microsporidia tree, except the WAG model plus gamma distributed rates among sites was used and the phylogeny was tested with 100 replicate bootstraps.

3.3 Results

3.3.1 Short read filtering and *de novo* assembly

Illumina sequencing of four *M. spermotrophus* whole insect cDNA libraries (larva, adult male, naïve adult female and wild adult female) generated 236,985,595 paired-end reads of 100 bp in length, equating to 47.40 giga-bases of total sequence. Fewer reads were sequenced from the wild adult female. Investigation of base-call quality using FastQC identified reduced phred scores at the ends of the reads. Subsequently, quality filtering removed approximately 15.4 % of the reads prior to assembly, resulting in a mixed population of paired- and single-end reads (Table 8).

The transcriptome of *M. spermotrophus* was assembled *de novo* using multiple *k*mer values (Figure S1). Each of the individual *k*mer assemblies was combined using trans-ABySS, resulting in 1,361,656 assembled contigs. These assembled contigs were first clustered using the program CD-HIT-EST, which generated 296,711 clusters. A second clustering program, TIGR-TGICL, was used, resulting in 44,176 clusters and 149,236 singletons. Removal of all contigs less than or equal to 100 bp resulted in a final contig set of 143,306 transcripts (Table 9). The transcripts ranged in length from 101 (minimum contig length) up to 32,049 bp, with a N50 of 2,420 bp. The entire length of the transcriptome totalled 118,105,899 bp with an average contig length of 824 bp.

3.3.2 Annotation

From the transcriptome 1,639 contigs had significant similarity (E-value cut-off = 10^{-15}) to 41 venom protein sequences from the *N. vitripennis* venom query. All of the carbohydrate metabolism, DNA metabolism and glutathione metabolism venom proteins from the *N. vitripennis* query were represented in the *M. spermotrophus* transcriptome. The unknown venom category had the smallest proportion of significantly similar sequences from the *M. spermotrophus* transcriptome. In some cases, annotations representing proteins with normal physiological functions from the NCBI non-redundant protein database had a smaller E-value than previously assigned venom protein annotations. Subsequently only 21 venom protein annotations were included in the final *M. spermotrophus* annotation set after annotations with the most significant E-value were assigned to each transcript (Figure 11).

Table 8: Illumina sequencing output for the *Megastigmus spermotrophus* whole insect cDNA

Library:	Larva	Male	Naïve Female	Wild Female	**Total**
Total number raw paired-end reads:	62,396,989	64,991,169	63,560,118	46,037,319	**236,985,595**
Raw paired-end read total length (Gbp):	12.48	13.00	12.71	9.21	**47.40**
Total number filtered paired-end reads:	49,883,162	52,200,155	50,592,674	36,311,907	**188,987,898**
Total number filtered single-end reads:	6,023,891	6,299,194	6,135,548	4,478,658	**22,937,291**

Table 9: *Megastigmus spermotrophus* transcriptome clustering results

Clustering Method	Number of Contigs	N50 (min:200bp)
Trans-ABySS	1,361,656	1,690
CD-HIT-EST	296,711	1,570
TIGR-TGICL	193,412	2,420
Singleton Removal	143,306	2,420

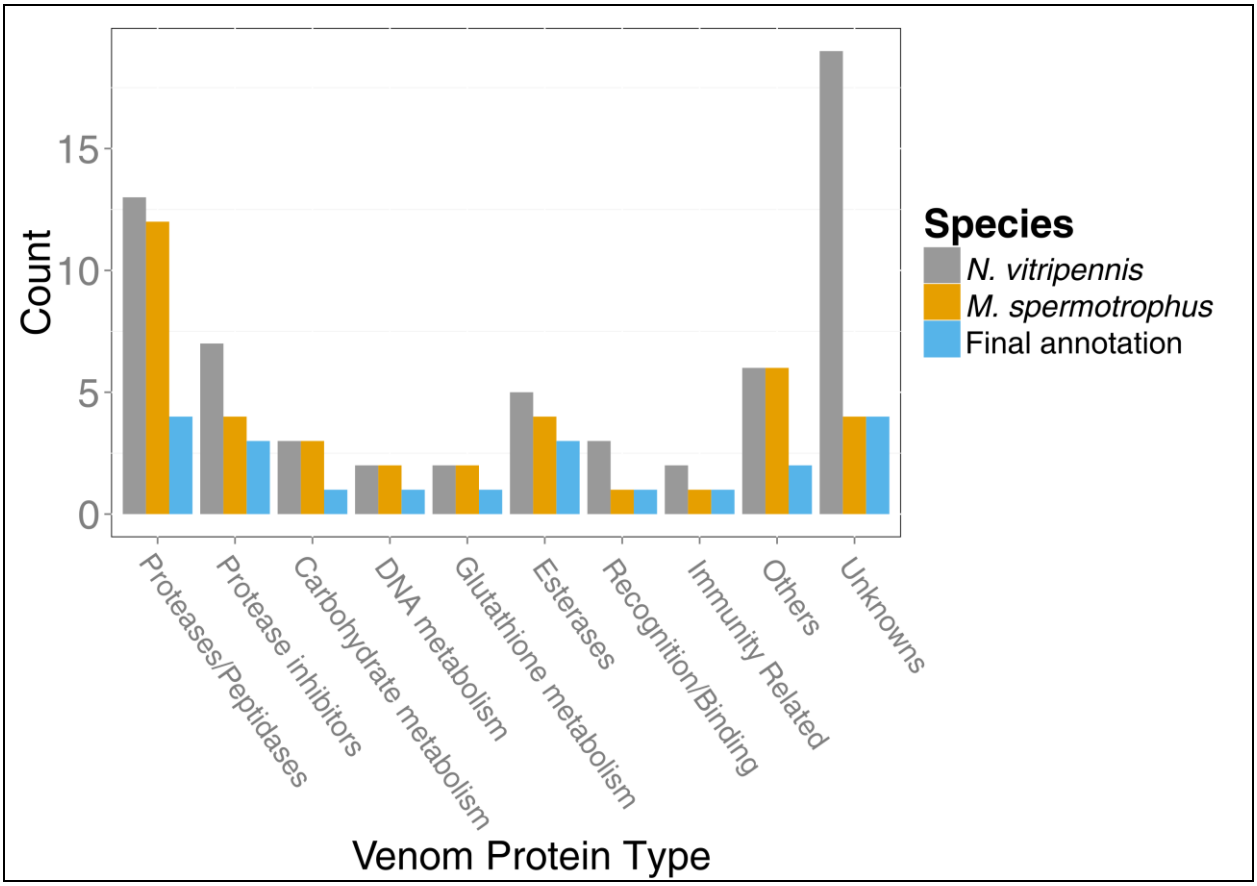


Figure 11: Number of *N. vitripennis* venom proteins, significantly similar transcripts and final venom transcript annotation from the *M. spermotrophus* transcriptome, organized by venom type (E-value cut-off 1 x 10^{-15}).

Transcripts were primarily annotated using NCBI's non-redundant protein database (BLASTx) and secondarily from the NCBI's nucleotide collection (BLASTn). From the entire transcriptome, 39,601 (27.6 %) showed a significant match (E-value less than or equal to 10^{-5}) to sequences in the protein database, with an average amino acid identity of 71.8 % (SD = 20.4) for the top-hits. Additionally, 3,077 (2.1 %) of the remaining contigs were assigned annotations from the nucleotide database. Many of the annotations were redundant, being assigned to multiple contigs (average annotation assignment = 2.1, standard deviation = 3.5), resulting in 20,284 (14.2 %) total non-redundant annotations (Table 10). A large proportion of the annotations were 'hypothetical' or 'predicted' proteins with no known function. The majority of smaller transcripts (less than 600 bp) were not assigned an annotation. Almost all of the transcripts larger than the N50 value were assigned an annotation (Figure 12).

The majority of non-redundant annotations assigned from the nr protein database were from insects (77.3 %). However, prokaryotes were also well represented in the annotation set (14.8 %) (Figure 13); A single genus of Betaproteobacteria, *Ralstonia*, represented 10.6% of all non-redundant protein annotations. Also, the intracellular parasite, microsporidia represented 49.5 % of all fungal annotations (1.5 %).

The jewel wasp, *N. vitripennis*, represented the greatest proportion of insect annotations (61.2 %); however other Hymenoptera were also represented in the transcriptome (i.e. *Apis* spp., *Bombus* spp., *Camponotus floridanus*, *Megachile rotundata*, *Harpegnathos saltator*, *Solenopsis invicta* and *Acromyrmex echinatior*) (Figure 14).

3.3.3 Transcript expression

The software package RSEM was used to generate expression values by mapping the forward reads from all of the libraries onto the assembled contigs and calculating expected counts. In total 202,977,319 forward reads were processed, with 5,980,555 (3.0 %) read mapping failures. As the mean fragment length was set to 300 bp, subsequently expression values were only calculated for contigs of length 300 bp or longer.

Table 10: *Megastigmus spermotrophus* transcriptome annotation results (E-value cut-off = 10^{-5})

Database:	NCBI nr Protein	NCBI nucleotide*	**Total**
Total Annotations	39,600	3,077	**42,677**
Total Non-redundant	18,610	1,674	**20,284**

*NCBI nucleotide annotations only assigned to contigs which were not previously assigned a NCBI protein annotation.

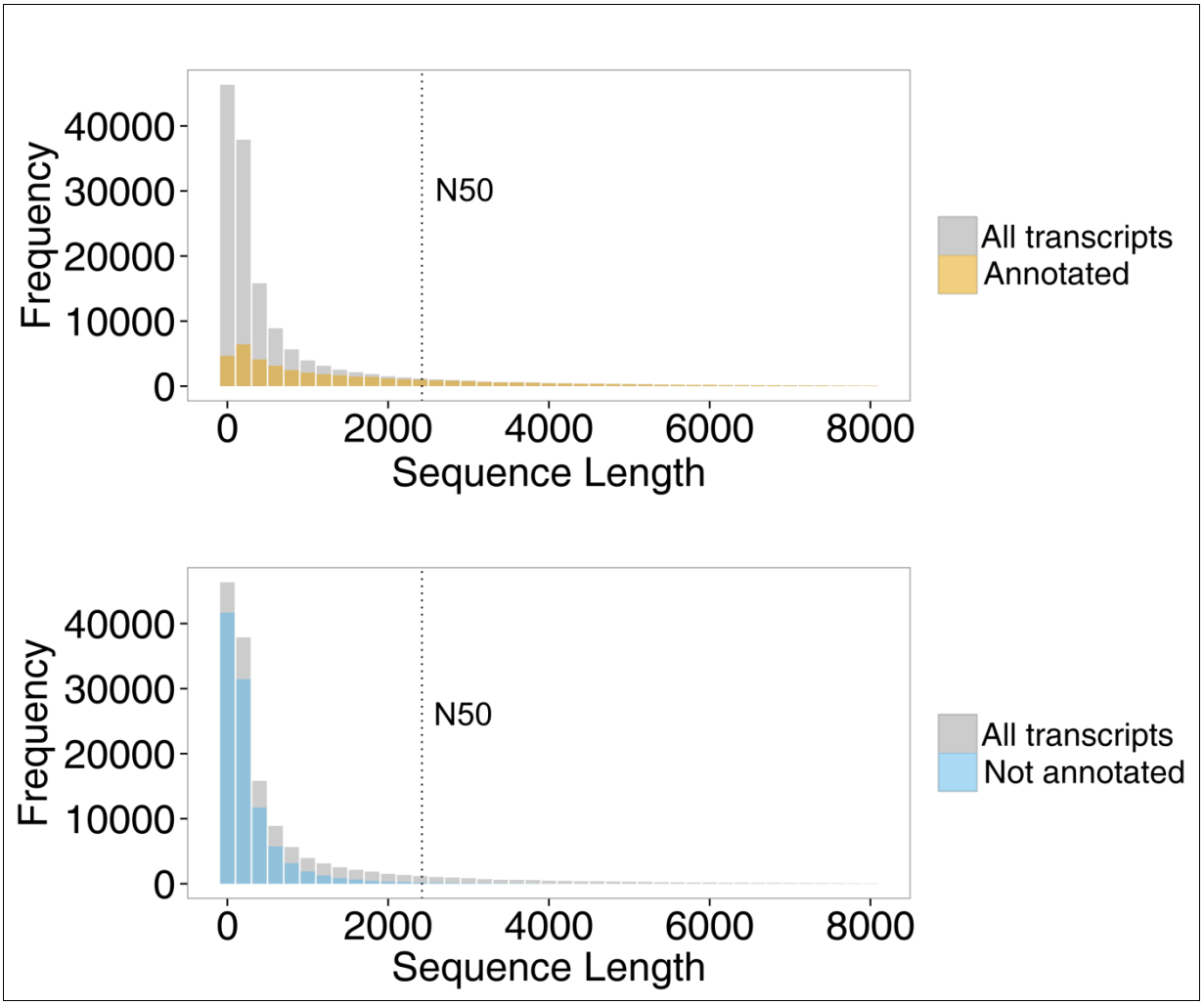


Figure 12: Assembled transcript length frequency histogram of the *Megastigmus spermotrophus* transcriptome (200 bp bin-size). All assembled transcripts up to 8,000 bp in size are included. Top: annotated transcript length frequency histogram is over-laid in yellow. Bottom: Non-annotated transcript length frequency histogram is over-laid in blue.

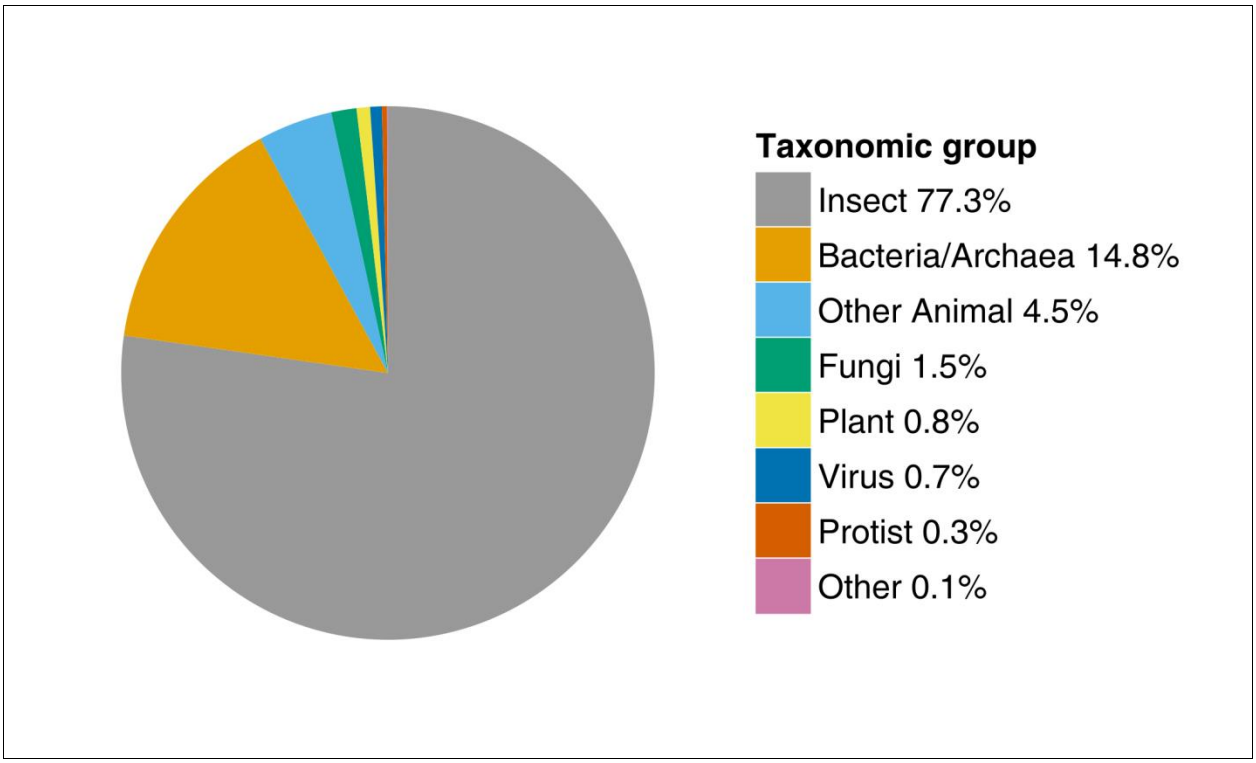


Figure 13: Taxa distribution of the BLASTx matches of the *M. spermotrophus* transcriptome. The taxon distribution and proportions of non-redundant contig BLASTX matches against the protein database (E-value cutoff $<10^{-5}$).

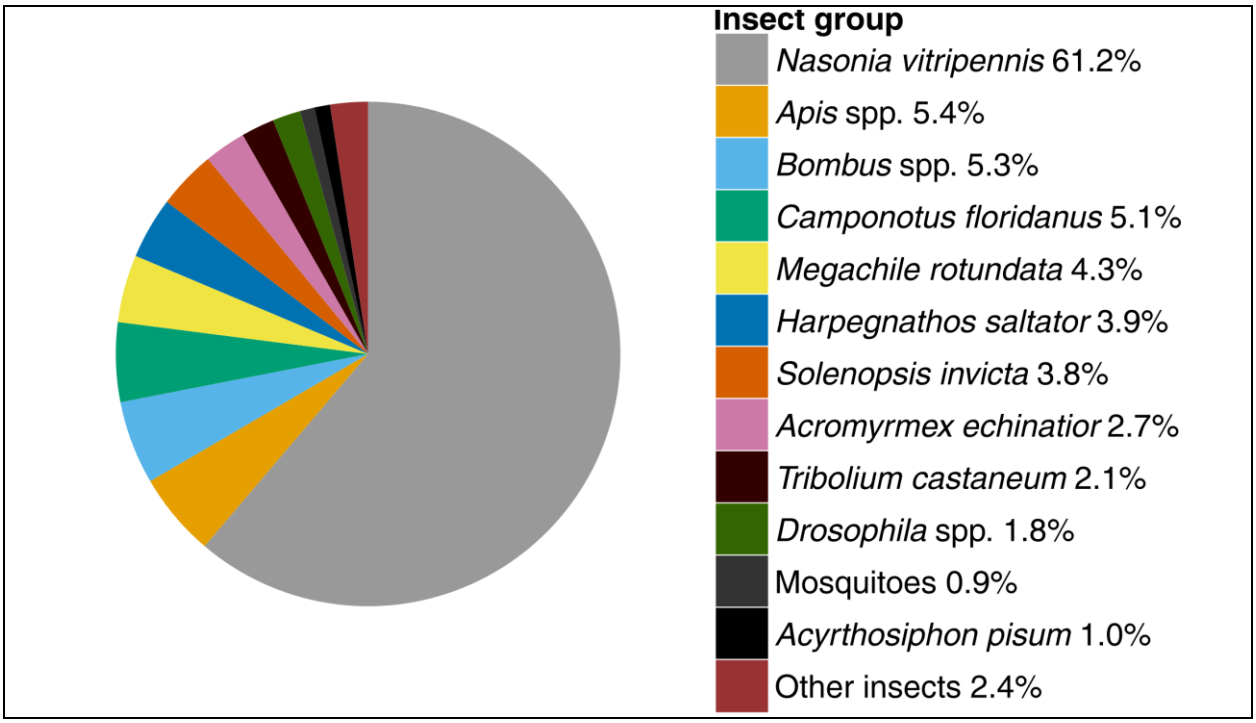


Figure 14: Insect distribution of the BLASTx matches of the *M. spermotrophus* transcriptome. The insect distribution and proportions of non-redundant contig BLASTX matches against the protein database (E-value cutoff $<10^{-5}$).

The expression data from the female and non-female libraries were found to be strongly biased by contig length since almost 100 % of the variability within the data is explained by fitting a cubic spline regression model, which follows a fairly linear pattern (Figure S2). Based on this analysis longer contigs were more likely to have a higher expected count. The expression data from the female and non-female libraries were also found to be strongly biased by GC content since almost 78 % and 71 % of the variability within the data can be explained by fitting a cubic spline regression model, respectively (Figure S3). The GC content bias was similar for both female and non-female libraries, except the lowest GC content point in the non-female library has a remarkably high mean expression value. The fitted model predicted that contigs with GC content close to 0.4 had a higher expected count.

The number of annotated transcripts from each of the libraries, and how the number of annotated transcripts changes with increasing sequence depth, based on expected counts, was investigated using a saturation plot. At this sequencing depth 72.6 - 76.2 % of annotated transcripts were obtained from the libraries. The larvae generated the greatest number of annotated transcripts compared to the other three library types. The wild adult female library had the lowest number of annotated transcripts and the lowest sequencing depth of all the libraries (Figure S4).

3.3.4 Normalization

Expected count data were transformed using CQN. Compared to other common normalization techniques, such as trimmed mean of M (TMM) and upper quartile (UQ), CQN resulted in the standardization of the distribution of counts for all four libraries. With the data normalized using CQN, 20 % of the transcripts account for approximately 90 % of the reads in all of the libraries. In contrast, the count distribution varied among the libraries with non-normalized data and data normalized using TMM and UQ (Figure S5). The CQN data still contained some expression level to transcript length bias; however the fitted model was more complicated, less linear and explained less of the variability in the data compared to the non-normalized data (Figure S2). The effect of GC content on expression was reduced with the CQN data, but the fitted model was still fairly similar to the non-normalized data (Figure S3).

3.3.5 Microsporidia investigation

The mean expression of transcripts annotated as insect, prokaryote, other animal, virus, plant, and protist remained conserved between the four *M. spermotrophus* transcriptome libraries (data not shown). Notably, the mean normalized expression of fungal transcripts was higher in wild adult females compared to larvae, males and naïve females (Figure 15). All microsporidian transcripts were expressed exclusively in wild females, partially accounting for the increased expression of fungal sequences in this particular library.

Microsporidia was detected in one out of six wild *M. spermotrophus* by targeted PCR screening. To place the microsporidian identified in the single wild female *M. spermotrophus* and in the transcriptome in a phylogenetic context, a maximum likelihood tree was generated based on the 18S rRNA sequence (Figure 16). The unknown microsporidian was allied with the *Nosema/Vairimorpha* clade.

3.3.6 Differential expression analysis

The NOISeq-sim algorithm identified 404 transcripts that are likely differentially expressed in *M. spermotrophus* females compared to larvae and males (Figure 17). This analysis will focus on the 243 transcripts that are likely highly expressed in females, since putative venom transcripts are more likely present in this specific set of transcripts (Table S2).

Three putative venom transcripts were identified as likely being highly expressed in female *M. spermotrophus*: Aspartylglucosaminidase precursor (AGA), gram-negative bacteria binding 1-2 precursor (GNB) and venom protein R precursor (venom R). Three contigs were annotated as AGA, of which, two were highly expressed in females. Six contigs were annotated as GNB, of which, four were highly expressed in females. Two contigs were identified as venom R, of which, one was highly expressed in females (Figure 18). Two transcripts were annotated as a putative paralog of AGA, referred to as L-asparaginase or L-asparaginase-like; these paralog transcripts were not found to be highly expressed in females. Three transcripts were annotated as the putative paralog of GNB, beta-1,3-glucan-binding protein; these three transcripts were all highly differentially expressed in females. There were no putative paralogs of venom R.

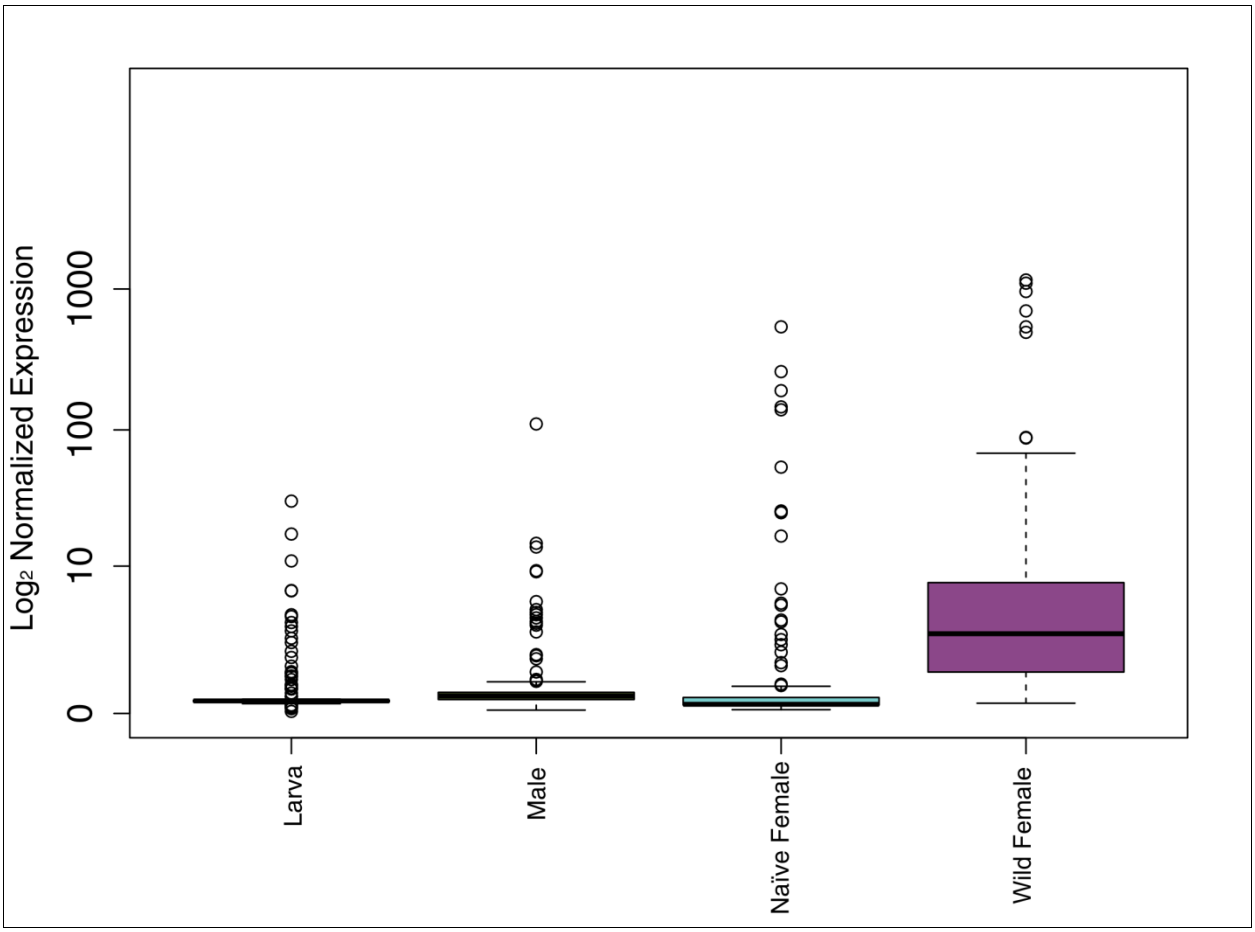
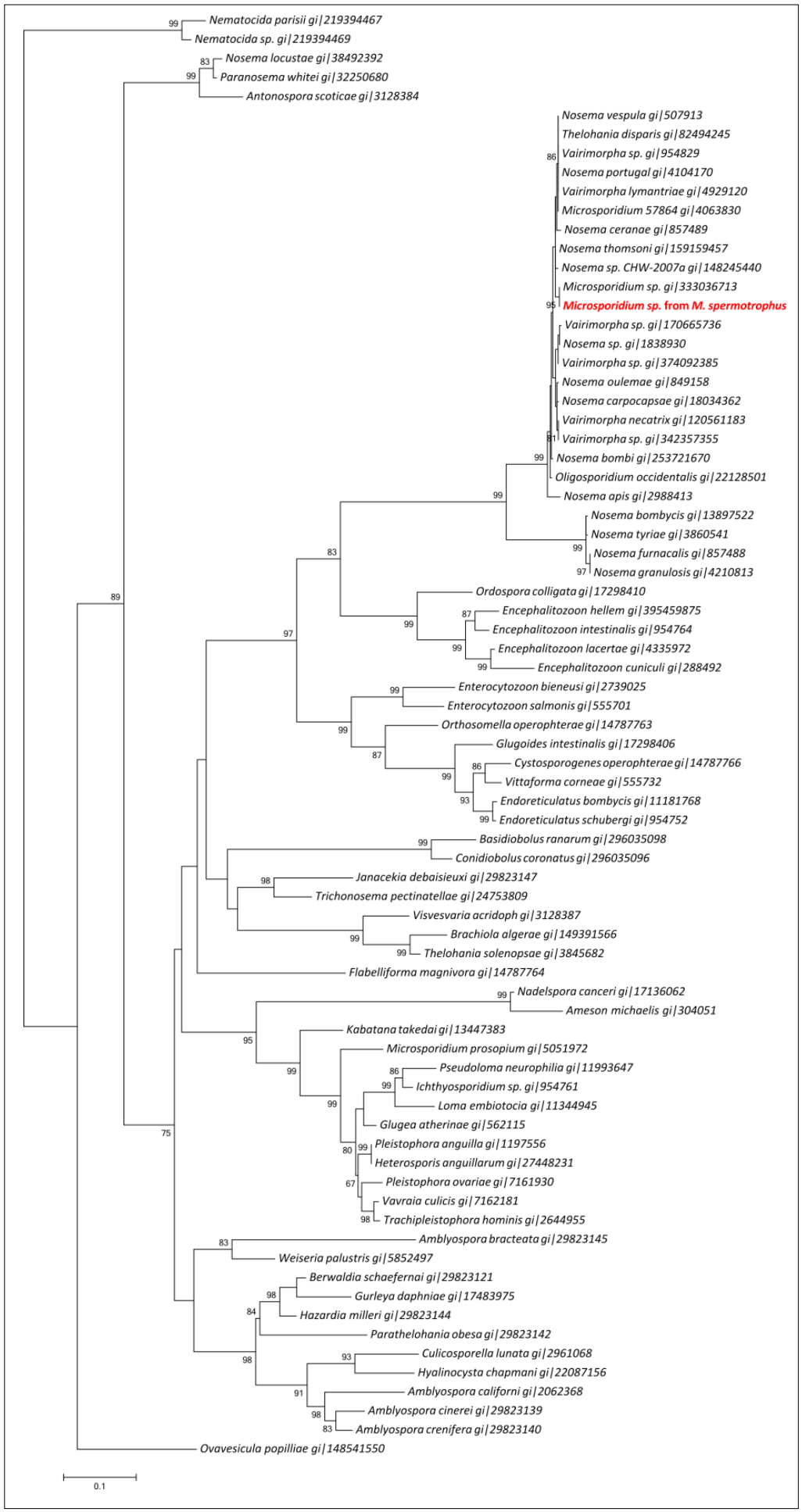


Figure 15: Box plot of $\log_2$ expression values from fungal transcripts from the *M. spermotrophus* transcriptome.

Figure 16: Maximum likelihood phylogeny for microsporidian 18S rRNA sequence using the general time reversible and gamma distributed rates among sites model of nucleotide substitution. The sequence generated by this study is highlighted in red. Numbers next to the nodes indicate percentage of bootstrap support from 500 bootstrap replicates. Nodes without numbers received less than 65 % bootstrap support.


Nematocida parisii gi|219394467
Nematocida sp. gi|219394469
Nosema locustae gi|38492392
Paranosema whitei gi|32250680
Antonospora scoticae gi|3128384
Nosema vespula gi|507913
Thelohania disparis gi|82494245
Vairimorpha sp. gi|954829
Nosema portugal gi|4104170
Vairimorpha lymantriae gi|4929120
Microsporidium 57864 gi|4063830
Nosema ceranae gi|857489
Nosema thomsoni gi|159159457
Nosema sp. CHW-2007a gi|148245440
Microsporidium sp. gi|333036713
Microsporidium sp. from M. spermotrophus
Vairimorpha sp. gi|170665736
Nosema sp. gi|1838930
Vairimorpha sp. gi|374092385
Nosema oulemae gi|849158
Nosema carpocapsae gi|18034362
Vairimorpha necatrix gi|120561183
Vairimorpha sp. gi|342357355
Nosema bombi gi|253721670
Oligosporidium occidentalis gi|22128501
Nosema apis gi|2988413
Nosema bombycis gi|13897522
Nosema tyriae gi|3860541
Nosema furnacalis gi|857488
Nosema granulosis gi|4210813
Ordospora colligata gi|17298410
Encephalitozoon hellem gi|395459875
Encephalitozoon intestinalis gi|954764
Encephalitozoon lacertae gi|4335972
Encephalitozoon cuniculi gi|288492
Enterocytozoon bieneusi gi|2739025
Enterocytozoon salmonis gi|555701
Orthosomella operophterae gi|14787763
Glugoides intestinalis gi|17298406
Cystosporogenes operophterae gi|14787766
Vittaforma corneae gi|555732
Endoreticulatus bombycis gi|11181768
Endoreticulatus schubergi gi|954752
Basidiobolus ranarum gi|296035098
Conidiobolus coronatus gi|296035096
Janacekia debaisieuxi gi|29823147
Trichonosema pectinatellae gi|24753809
Visvesvaria acridoph gi|3128387
Brachiola algerae gi|149391566
Thelohania solenopsae gi|3845682
Flabelliforma magnivora gi|14787764
Nadelspora canceri gi|17136062
Ameson michaelis gi|304051
Kabatana takedai gi|13447383
Microsporidium prosopium gi|5051972
Pseudoloma neurophilia gi|11993647
Ichthyosporidium sp. gi|954761
Loma embiotocia gi|11344945
Glugea atherinae gi|562115
Pleistophora anguilla gi|1197556
Heterosporis anguillarum gi|27448231
Pleistophora ovariae gi|7161930
Vavraia culicis gi|7162181
Trachipleistophora hominis gi|2644955
Amblyospora bracteata gi|29823145
Weiseria palustris gi|5852497
Berwaldia schaefernai gi|29823121
Gurleya daphniae gi|17483975
Hazardia milleri gi|29823144
Parathelohania obesa gi|29823142
Culicosporella lunata gi|2961068
Hyalinocysta chapmani gi|22087156
Amblyospora californi gi|2062368
Amblyospora cinerei gi|29823139
Amblyospora crenifera gi|29823140
Ovavesicula popilliae gi|148541550
0.1

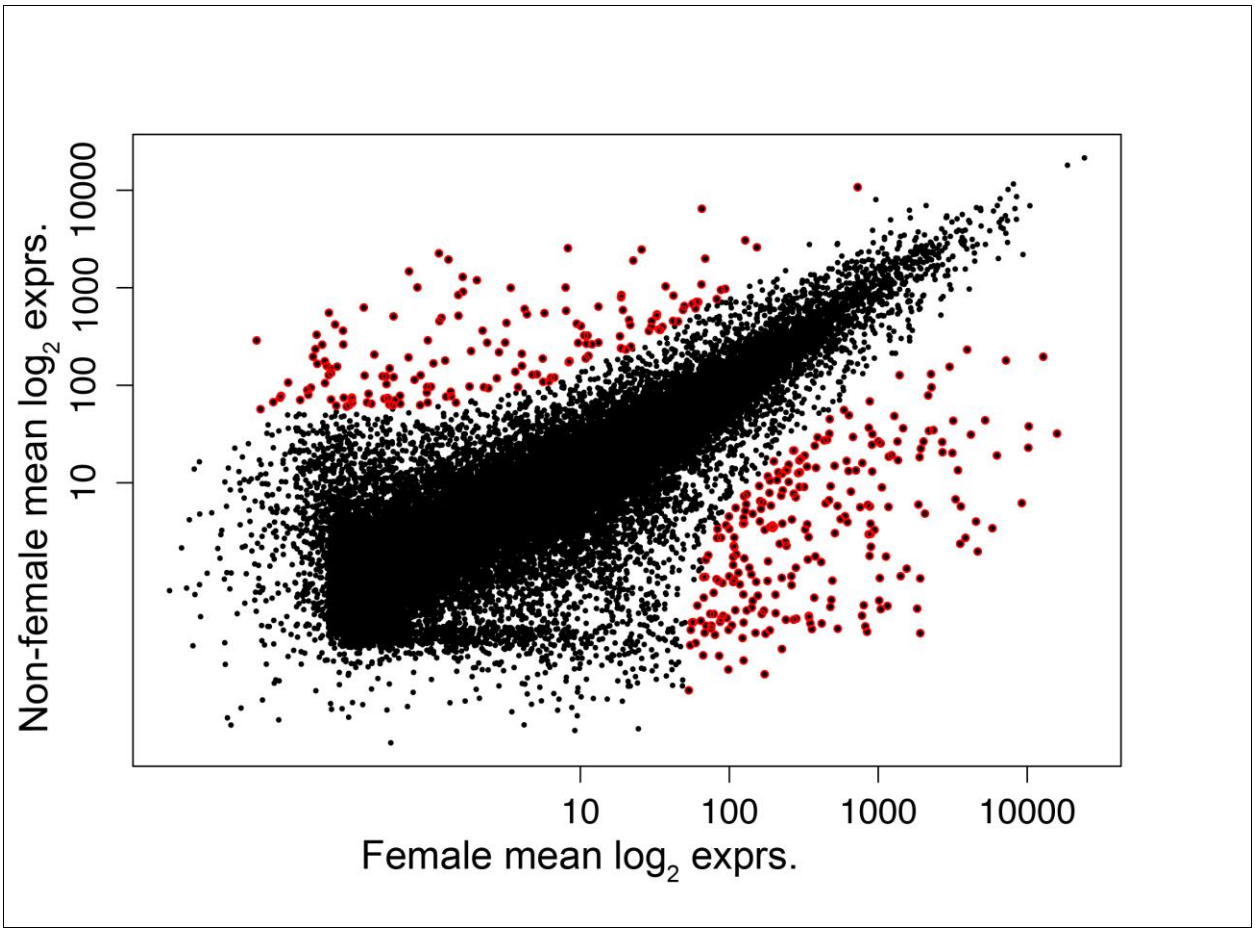


Figure 17: Log_2 mean normalized expression values from female (wild and naïve) and non-female (larva and adult male) transcriptome libraries of *M. spermotrophus*. Likely differentially expressed features are highlighted in red, as determined by the NOISeq-sim algorithm.

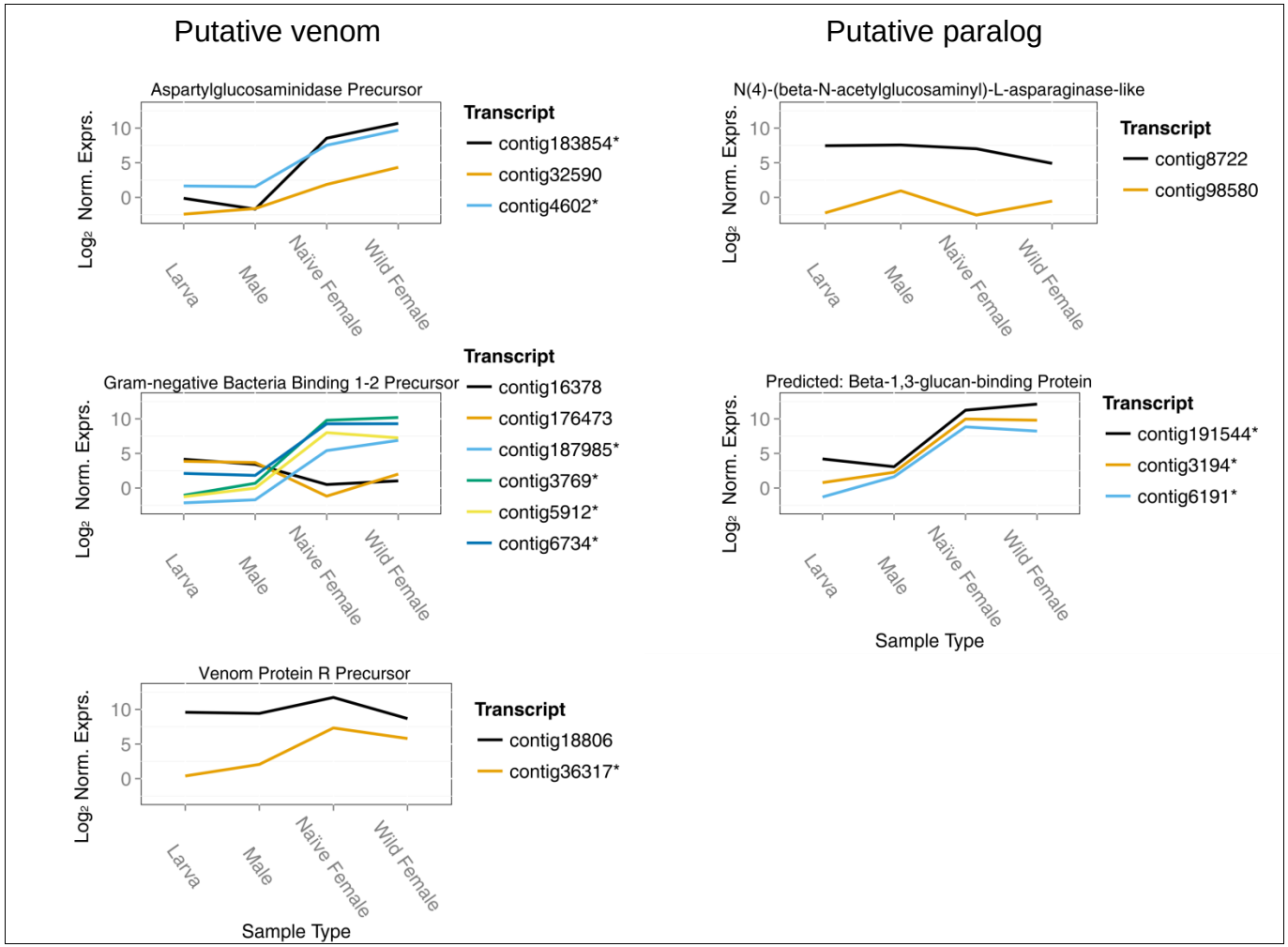


Figure 18: Expression profiles for three putative venom transcripts and their respective putative paralogs from the *M. spermotrophus* transcriptome. *Asterisk denotes contigs that are likely highly differentially expressed in females based on non-parametric statistical analysis from NOISeq-sim.

The evolutionary relationships of AGA and its asparaginase paralog were reconstructed using maximum likelihood methods (Figure 19). This phylogeny shows that AGA has been duplicated in Hymenoptera, but not in other insects. The copy that is differentially expressed in *Megastigmus* females is homologous to the copy that has been characterized as a venom precursor in *N. vitripennis.*

3.4 Discussion:

3.4.1 Potential venom transcripts in the *M. spermotrophus* transcriptome

We hypothesize that, as in their parasitoid relatives, venom proteins are produced and stored in glands associated with the reproductive tract of female wasps and then injected into the developing megagametophyte during oviposition and that these venom proteins may play an important role in early host manipulation (i.e.; ovule redirection), possibly by influencing the regulation and/or production of phytohormones. A diverse set of 41 candidate venom transcripts were identified in the *M. spermotrophus* transcriptome using a query of venom protein sequences from *N. vitripennis.* This result suggests that parasitoid and phytophagous wasps may share a large number of homologous venom proteins. It is interesting to note that dissections of *M. spermotrophus* revealed the presence of a glandular structure that appears to drain into the base of the ovipositor (Paulson A., personal observation). Although, these two species have quite different life histories (one species develops on fly pupae and the other inside Doulas-fir ovules) they share venom sequences homology. Twenty-one *N.vitripennis* venom protein annotations were included in the final *M. spermotrophus* transcriptome annotation set, including venom proteins from all categories. Many are thought to have evolved from proteins with normal physiological functions (Casewell et al. 2013), but there is also evidence that venom toxins can be co-expressed in venom glands and other tissues (Casewell et al. 2012). There was a high degree of sequence similarity between venom protein sequences and their corresponding 'physiological' protein sequence, resulting in some of the initial venom transcript annotations being replaced by non-venom transcript annotations that had lower E-value scores. Since the objective of this study is to identify putative venom transcripts involved in early host manipulation, the next step is to identify venom transcripts that are likely highly expressed in females using differential expression analysis.

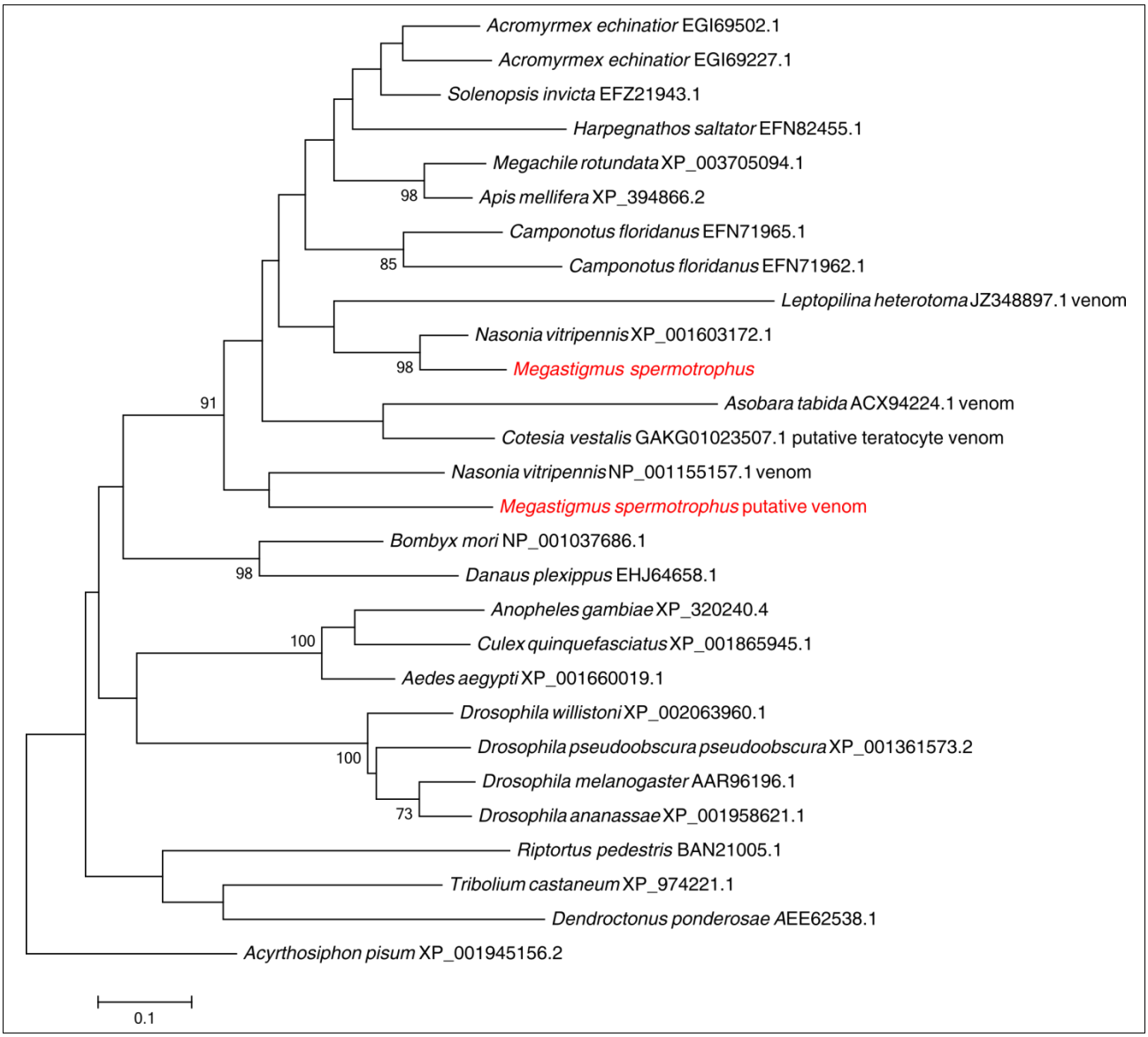


Figure 19: Maximum likelihood phylogeny for Aspartylglucosaminidase protein sequence from insects using a WAG amino acid replacement model with gamma distributed rates among sites. Numbers next to the nodes indicate percentage of bootstrap support from 100 bootstrap replicates. Nodes without numbers received less than 65 % bootstrap support.

3.4.2 Narrowing down candidate venom transcripts: Differential expression in adult females

Candidate venom transcripts that may play a role in early host manipulation would be expected to be highly expressed in adult female *Megastigmus*. Differential expression analysis of the *M. spermotrophus* transcriptome revealed three candidate venom transcripts that are highly expressed in females, aspartylglucosaminidase precursor (AGA), gram-negative bacteria binding 1-2 precursor (GNB), and venom protein R precursor. Analysis of expression patterns and phylogenetic analysis of AGA suggest that this is a particularly promising candidate.

Aspartylglucosaminidase Precursor – Two contigs annotated as AGA are highly differentially expressed in females and a third contig follows a similar pattern. On the other hand, transcripts of AGA's paralog, N-(4)-(beta-N-acetylglucosaminyl)-L-asparaginase (a synonym for AGA), which is not a venom protein in *Nasonia*, were not differentially expressed in females. The distinctive expression patterns of AGA and its corresponding 'physiological' paralog provide a good indication that AGA may be a venom protein of *M. spermotrophus*, warranting further investigation.

AGA has been identified as a venom protein of several species of parasitoids, including *N. vitripennis*. Initially it was identified as a major venom constituent of the *Drosophila* parasitoid *Asobara tabida* (Braconidae) (Moreau et al. 2004, Vinchon et al. 2010). The function of AGA in the venom of *A. tabida* remains unknown; however induction of transient paralysis of *D. melanogaster* host larvae, or diversion of host metabolism have been suggested as a potential mechanism of venomous AGA (Moreau et al. 2002, 2004). More recently, transcriptomic analysis of the venom apparatus identified AGA as a major venom constituent of *Leptopilina heterotoma* (Figitidae) (Colinet et al. 2013), another *Drosophila* parasitoid (Schlenke et al. 2007). In contrast to *L. heterotoma*, AGA was not found to be a major component in the venom of closely related *L. boulardi* (Colinet et al. 2013), demonstrating the highly dynamic nature of wasp venom. Finally, the NCBI transcriptome shotgun assembly database was searched using *M. spermotrophus* AGA venom transcript sequence as a query. A significant match to a transcribed RNA sequence from an unpublished transcriptome study of teratocytes from *Cotesia plutellae* (Braconidae) was found (accession #GAKG01023507.1). Teratocytes are specialized cells that develop from extra-embryonic tissues during embryogenesis and interfere with host defense

(Dahlman 1990). Potentially AGA could be produced and secreted by teratocytes, since digestive enzymes, such as acid phosphatase, esterase and leucine amino peptidase have been previously reported from teratocyte preparations (Strand et al. 1986). Young teratocytes are known to possess the cell architecture necessary for production of secretory products (Dahlman 1990).

3.4.3 Evolution of venom AGA in Hymenoptera

From bacteria to humans, AGA is typically a lysosomal enzyme, involved in the digestion of glycoproteins (Tarentino et al. 1995, Tenhunen et al. 1995, Liu et al. 1996). AGA acts on glycosylated asparagines by hydrolyzing the ßN glycosidic linkage between the asparagine residue and the N-acetylglucosamine moiety. Human AGA production and cellular transportation has been studied extensively because deleterious mutations to the AGA gene results in a lysosomal storage disease, called aspartylglucosaminuria (Saarela et al. 2001). The recruitment of lysosomal enzymes into venomous secretions is not a new concept and the diversity of hydrolases found in hymenopteran venoms provides motivation for the idea that at least some ancient lysosomal genes may have evolved into components of wasp venom (Vinchon et al. 2010). Further support for this idea is provided by the recent finding of another major lysosomal enzyme, acid phosphatase, from the venoms of two parasitoid wasps, *Pimpla hypochondriaca* Retzius (Dani et al. 2005) and *Pteromalus puparum* (Zhu et al. 2010).

The evolution of many venom toxins is thought to occur by a process of ‘birth and death’ (Nei et al. 1997). In this process a gene that encodes a normal ‘physiological’ protein that has an important bioactivity or regulatory function is duplicated. The duplicate copy becomes selectively expressed in the venom gland. (Kordiš and Gubenšek 2000, Fry et al. 2003). The birth and death process can result in multigene families that contain functional genes and non-functional pseudogenes (Nei et al. 1997). Further investigation of AGA and its paralog within the *N. vitripennis* genome (Munoz-Torres et al. 2011) revealed evidence of a gene duplication. The *N. vitripennis* AGA venom protein (NP_001155157.1) and the asparaginase putative paralog (XP_001603172.1) are located directly adjacent to each other on chromosome 1, scaffold 53. Duplication of the AGA gene only occurs within the Hymenoptera based on a survey of all available hymenopteran genomes and insect AGA sequences from GenBank databases.

Gram-negative Bacteria Binding 1-2 Precursor – Four of six transcripts related to *Nasonia* GNB were highly expressed in adult females compared to larvae and adult males; the other two contigs did not show expression differences. Both GNB and its paralog β-1,3-glucan binding protein (βGBP) belong to a family of recognition proteins known as the gram-negative bacteria-binding proteins (GNBPs). Some of these proteins have a strong affinity for lipopolysaccharides from gram-negative bacteria while others have an affinity for β-1,3-glucan from fungi (Kim et al. 2000, Ochiai 2000). GNB was first characterized as a venom component using proteomics of the *N. vitripennis* venom gland. GNBPs were not previously known to be associated with insect venom (de Graaf et al. 2010). In *Megastigmus*, all of the transcripts annotated as βGBP were also highly expressed in females, possibly as a response to an immunological challenge (see section on interesting microbes from the transcriptome below). It is therefore possible that GNB also has an immunological and not necessarily venomous function in *M. spermotrophus*, but additional investigation may prove otherwise.

Venom R – One of two transcripts related to *Nasonia* venom R protein were highly expressed in *M. spermotrophus* females, while the other transcript was not. This protein was previously identified from venom extract of *N. vitripennis* using a proteomics approach (de Graaf et al. 2010), and has no similarity to any known protein, so it is difficult to make predictions with respect to its function. Further investigation would be required to determine if venom R could be a putative venom protein of *M. spermotrophus*.

3.4.4 Limits of the transcriptome in venom transcript detection

While our approach was useful for targeting potential adult venom constituents, the early larval instars are also likely very important in host manipulation. Larval secretions during feeding are known to be critical in gall formation in cynipid gall wasps (Leggo and Shorthouse 2006). Perhaps, the feeding larva of *M. spermotrophus* continues to manipulate seed development via salivary secretions. The *M. spermotrophus* transcriptome is missing this key development stage, which limited our ability to identify any candidate venom proteins that may be secreted by the larvae. Furthermore, the identification of putative venom transcripts from the *M. spermotrophus* transcriptome was done using a collection of venom protein sequences identified from adult female *N. vitripennis*, further limiting our ability to detect venom-like sequences that may be uniquely expressed during parasitoid larval development.

The bioinformatics methods employed in this study were likely not very sensitive to detect rare and lowly expressed venom transcripts in female wasps for three reasons: 1) the expression data was considerably noisy as a result of moderate redundancy and mappability issues, 2) one of the female libraries was sequenced to a lesser depth than the other libraries, and 3) the lack of technical replicates required the use of stringent non-parametric statistical analysis to identify differential expression.

In many instances, not all of the redundantly annotated contigs were found to be differentially expressed in females. In some cases, this may have been caused by different expression patterns of legitimate isoforms or splice variants of the same protein, or more likely a consequence of noisy data.

Although a number of annotated transcripts were identified as differentially expressed in females many of them are hypothetical proteins with no known function. However, included among the differentially expressed transcripts were genes that would be expected to be highly expressed in females, including transcripts associated with oogenesis, such as vitellogenin-like and vitellogenin (Guidugli et al. 2005), vitellogenin receptor (Schonbaum et al. 2000), nanos (Forbes and Lehmann 1998) and maternal effect protein oskar (Lehmann and Nüsslein-Volhard 1986). Also, several transcripts associated with odor perception, such as chemosensory protein CSP-1 (Pelosi et al. 2006) and putative odorant binding protein 70 (Vogt et al. 1999) were highly expressed in females. The expression of odor perception related genes may be essential for locating susceptible Douglas-fir ovules by females.

3.4.5 General features of the *M. spermotrophus* transcriptome

Both the stochastic process of next generation sequencing and the heuristic approach of transcriptomic and RNA-seq methodologies are probable sources of noise and bias associated with the *M. spermotrophus* transcriptome. The following sections provide a brief description of the general features of the transcriptome, including inherent sources of bias and noise.

De novo assembly - Comparison of *de novo* transcriptome features demands caution, since the sequencing method, sequencing depth, assembly methodologies and post-assembly processing strategies can be highly variable among studies. *De novo* assembly and post-assembly clustering

of the *M. spermotrophus* transcriptome generated 143,306 contigs. The number of contigs is comparable to other insect transcriptomes constructed from Illumina generated sequences, such as that of soybean aphid (253,603 contigs) (Liu et al. 2012), *Anopheles funestus* (46,987 contigs) (Crawford et al. 2010), brown planthopper (1,921,675 contigs) (Xue et al. 2010), oriental fruit fly (484,628 contigs) (Shen et al. 2011) and salt marsh beetle (65,766 contigs) (van Belleghem et al. 2012). The post-sequencing clustering and removal of singletons resulted in the *M. spermotrophus* having a relatively large average contig length of over 800 bp.

Annotation - As expected, the majority of annotations assigned to the *M. spermotrophus* transcriptome were from the model parasitoid wasp, *N. vitripennis*, which provided some validation that the transcriptome contain a large representation of insect transcripts. Non mRNAs, such as prokaryotic (mostly bacteria) and ribosomal RNA were also represented in the *M. spermotrophus* transcriptome. Transcriptomes often include a small portion of non mRNAs, which is likely a result of inefficiencies with the mRNA enrichment process; some library preparation protocols include two rounds of mRNA enrichment to reduce the presence of non mRNA transcripts (Tariq et al. 2011).

A large proportion of transcripts from the *M. spermotrophus* transcriptome were not assigned an annotation (70 %). They were generally short length transcripts. They likely represent small, non-coding RNA that are not included in the NCBI protein or nucleotide databases. Other short transcripts may represent partially assembled rare transcripts. Low coverage transcripts are less likely to be annotated for two reasons. First, their rarity results in under representation in current databases. Second, their short length translates into short query coverage, which makes it harder to generate an e-value that is considered significant enough to assign an annotation. In addition, novel transcripts that are essential to the unique ecology of seeding feeding may not have been annotated because they are likely not present in the NCBI databases.

Over half of the annotations assigned to the *M. spermotrophus* transcriptome were redundant, being assigned to at least more than one contig. Redundancy adds noise to transcriptome data (Yang and Smith 2013). It was unclear how much the assembly method, assembly errors, allelic variation, variation in post-translational modification and genomic DNA contamination contributed to the redundancy. Redundancy is an issue that has been found in other *de novo*

insect transcriptomes; for example, less than half of the annotated transcripts represented unigenes in salt marsh beetle transcriptome (van Belleghem et al. 2012). An even higher level of redundancy was observed in the deeply sequenced soybean aphid transcriptome, in which, only 8 % of total annotations were non-redundant (Liu et al. 2012).

Expression and normalization - Sequencing depth was less for the wild female sample. Subsequently, the number of detected features (annotated contigs) was reduced for this developmental stage. This may have negatively affected the ability to detect rarely expressed venom transcripts from this particular library since rare transcripts may have only been partially sequenced, if they were sequenced at all.

The presence of redundant transcripts may have further exacerbated the number of misalignments. Only forward reads were mapped back onto the *M. spermotrophus* transcriptome in order to generate expression levels. The alignment of single-reads, instead of paired-end reads, likely created additional noise in the expression level data; alignment of single-end reads improves the total number of reads that align, but also increases the chance of misalignments.

The expected count data was found to have strong length and GC content bias. The length bias may be partially attributed to the reduced chance of rare transcripts being fully assembled, resulting in low coverage of many short reads. The GC content bias was not likely indicative of a biological function, but rather a reflection of the high proportion of *M. spermotrophus* transcripts with GC content close to 0.4. Normalization provided some correction for length and GC content bias. More importantly, normalization potentially improved the ability to detect differentially expressed features between the different developmental stages by reducing variation in the count distributions between libraries.

3.4.6 Interesting microbes from the transcriptome

While most transcripts were clearly of insect origin (77.3 %), there were a number of microbial transcripts (14.8 %). Of these, two sources stand out in particular. Approximately 76 % of the bacterial transcripts from the transcriptome were assigned to a member of the Betaproteobacteria, *Ralstonia*. Using 16S rRNA pyrosequencing, this microbe was found to be a major component of the *M. spermotrophus* microbiome (Chapter 2). The high prevalence of

Ralstonia associated transcripts provides further evidence that this is likely an important microbial associate of *M. spermotrophus*. Approximately 23 % of the bacterial transcripts from the transcriptome were assigned to Gammaproteobacteria; bacteria from this group were also prevalent in the *M. spermotrophus* microbiome study. While the majority of bacterial transcripts in the transcriptome appear to represent normal microflora, it is also possible that accidental bacterial contamination from the environmental or laboratory sources may have accounted for some additional bacterial associated transcripts.

This is the first report of a microsporidian infecting *Megastigmus* as Microsporidia transcripts were detected in the wild female library. The overall expression of fungal transcripts was higher in the wild female library. Microsporidia infection in wild females was confirmed by targeted PCR screening. Microsporidia are common intracellular parasites and pathogens of animals, including insects. Their use in classical biological control has been studied for two well-known forest pests: the gypsy moth (Weiser & Novotnỳ, 1987) and the spruce budworm (Wilson and Kaupp 1976). Recent molecular work has placed the phylum Microsporidia within the kingdom Fungi (Hirt et al. 1999, Capella-Gutiérrez et al. 2012). Phylogenetic analysis of the 18S small subunit rRNA gene showed that the *Megastigmus* parasite lies in the *Nosema/Varimorpha* clade, which includes important parasites of terrestrial insects (Solter et al. 2012). Much future work is required to assess the prevalence, pathogenicity, host range and transmission of this microsporidian in *Megastigmus*.

The immunity-related transcripts that were highly expressed in females suggest a host reaction to a pathogen, such as Microsprodia. These proteins included akirin-like protein (Goto et al. 2008), gram-negative bacteria binding 1-2 precursor and beta-1,3-glucan binding proteins (Kim et al. 2000), antimicrobial peptide defensin 1-2 precursor (Hoffman and Hetru 1992), and hymenoptaecin-1 isoform 2 precursor (Casteels et al. 1993). The differential expression of immunity related transcripts, especially fungal specific immunity binding proteins, suggests that the microsporidian identified in *Megastigmus* may be pathogenic.

3.5 Conclusions

In this study the *de novo* transcriptome of *M. spermotrophus* was mined to identify putative venom transcripts that were highly expressed in female wasps, with the long-term goal of

identifying mechanisms of host manipulation. Although the majority of transcripts did not have significant matches to NCBI databases, many of those that were annotated were assigned matches from *N. vitripennis* and other insects. Additionally, 21 candidate venom transcripts were annotated in the *M. spermotrophus* transcriptome from a query of *N. vitripennis* venom protein sequences.

Of three putative venom transcripts that were identified from the transcriptome and that appeared to be highly expressed in females, AGA and its lysosomal paralog demonstrated the clearest expression pattern. Furthermore, the recent finding of AGA as a major venom constituent in two other parasitoid wasps provided additional support that AGA is a putative venom component of *M. spermotrophus*. AGA exists as a single copy across all insects except Hymenoptera, where it appears to have duplicated, with one copy apparently evolving a specialized role as a venom protein.

There are many avenues to explore as a result of AGA's identification. Is AGA produced in the acid gland? Is AGA injected into the megagametophyte with the egg? What is the role of AGA in the venom of *M. spermotrophus*? Does it act on developing megagametophyte tissue or activate other unknown venomous proteins via catabolism of glycoproteins? Is venomous AGA present in other species of *Megastigmus* or gall wasps?

There are tools available that could be used to answer some of these new questions. AGA has been detected from the venom gland and reservoir extracts of *A. tabida* and *L. heterotoma*, respectively using an immunological probe (Moreau et al. 2004, Colinet et al. 2013). This may prove useful for localization of AGA in either the *M. spermotrophus* acid gland or freshly parasitized ovules. Quantitative real-time PCR (qPCR) could be applied to validate the expression levels of AGA in female wasps. It would also be useful to obtain venom AGA sequence from other species of *Megastigmus* or other phytophagous wasps for comparative phylogenetic analysis. Proteomics approaches could be used to characterize the biochemical properties of the AGA precursor by analyzing crude venom extract with two-dimensional liquid chromatography-mass spectrometry.

The application of RNA-seq in the detection of venom transcripts from a phytophagous wasp provided novel and intriguing results. The idea that oviposition secretions are important to gall induction by sawflies, cynipid wasps and fig-pollinating wasps has been previously suggested (McCalla et al. 1962, Price 1992, Kjellberg et al. 2005, Leggo and Shorthouse 2006). Also, gall-inducing cynipids and chalcids are known to have well developed venom glands (James 1926, Vårdal 2004, 2006). The finding of several homologous venom sequences between the parasitoid *N. vitripennis* and the seed-feeder *M. spermotrophus* supports the idea that phytophagous hymenopteran lineages have likely adapted the parasitoid venom machinery for manipulating plants. The high expression of AGA and Venom R in female *M. spermotrophus* warrants the need for further investigation of the potential role of venomous secretions in early host manipulation by seed chalcid wasps.

Bibliography

Adams, A. S., F. O. Aylward, S. M. Adams, N. Erbilgin, B. H. Aukema, C. R. Currie, G. Suen, and K. F. Raffa. 2013. Mountain Pine Beetles colonizing historical and naïve host trees are associated with a bacterial community highly enriched in genes contributing to terpene metabolism. Applied and Environmental Microbiology 79:3468–3475.

Aires, T., N. Marbà, E. A. Serrao, C. M. Duarte, and S. Arnaud-Haond. 2012. Selective elimination of chloroplastidial DNA for metagenomics of bacteria associated with the green algae *Caulerpa taxifolia* (Bryopsidophyceae). Journal of Phycology 48:483–490.

Allison, S. D., and J. C. Schultz. 2005. Biochemical responses of chestnut oak to a galling cynipid. Journal of Chemical Ecology 31:151–166.

Anderson, K. E., J. A. Russell, C. S. Moreau, S. Kautz, K. E. Sullam, Y. Hu, U. Basinger, B. M. Mott, N. Buck, and D. E. Wheeler. 2012. Highly similar microbial communities are shared among related and trophically similar ant species. Molecular Ecology 21:2282–2296.

Andersson, M. 1984. The evolution of eusociality. Annual Review of Ecology and Systematics 15:165–189.

Andrews, S. 2010. FastQC: A quality control tool for high throughput sequence data. http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.

Asgari, S., and D. B. Rivers. 2011. Venom proteins from endoparasitoid wasps and their role in host-parasite interactions. Annual Review of Entomology 56:313–335.

Auger-Rozenberg, M.-A., T. Boivin, E. Magnoux, C. Courtin, A. Roques, and C. Kerdelhué. 2012. Inferences on population history of a seed chalcid wasp: invasion success despite a severe founder effect from an unexpected source population. Molecular Ecology 21:6086–6103.

Auger-Rozenberg, M.-A., C. Kerdelhué, E. Magnoux, J. Turgeon, J.-Y. Rasplus, and A. Roques. 2006. Molecular phylogeny and evolution of host-plant use in conifer seed chalcids in the genus *Megastigmus* (Hymenoptera: Torymidae). Systematic Entomology 31:47–64.

Auger-Rozenberg, M.-A., and A. Roques. 2012. Seed wasp invasions promoted by unregulated seed trade affect vegetal and animal biodiversity. Integrative Zoology 7:228–246.

Austin, A. D., and P. C. Dangerfield. 1998. Biology of *Mesostoa kerri* (Insecta: Hymenoptera: Braconidae: Mesostoinae), an endemic Australian wasp that causes stem galls on *Banksia marginata*. Australian Journal of Botany 46:559–569.

Baker, M. D., C. R. Vossbrinck, E. S. Didier, J. V Maddox, and J. A. Shadduck. 1995. Small subunit ribosomal DNA phylogeny of various Microsporidia with emphasis on AIDS related forms. Journal of Eukaryotic Microbiology 42:564–570.

Baldo, L., J. C. Dunning Hotopp, K. A. Jolley, S. R. Bordenstein, S. A. Biber, R. R. Choudhury, C. Hayashi, M. C. J. Maiden, H. Tettelin, and J. H. Werren. 2006. Multilocus sequence typing system for the endosymbiont *Wolbachia pipientis*. Applied and Environmental Microbiology 72:7098–7110.

Bansal, R., S. Hulbert, B. Schemerhorn, J. C. Reese, R. J. Whitworth, J. J. Stuart, and M.-S. Chen. 2011. Hessian fly-associated bacteria: transmission, essentiality, and composition. PloS One 6:e23170.

Baumann, P. 2005. Biology bacteriocyte-associated endosymbionts of plant sap-sucking insects. Annual Review of Microbiology 59:155–189.

Baumann, P., N. A. Moran, and L. Baumann. 1997. The evolution and genetics of aphid endosymbionts. BioScience 47:12–20.

Van Belleghem, S. M., D. Roelofs, J. Van Houdt, and F. Hendrickx. 2012. *De novo* transcriptome assembly and SNP discovery in the wing polymorphic salt marsh beetle *Pogonus chalceus* (Coleoptera, Carabidae). PloS One 7:e42605.

Body, M., W. Kaiser, G. Dubreuil, J. Casas, and D. Giron. 2013. Leaf-miners co-opt microorganisms to enhance their nutritional environment. Journal of Chemical Ecology 39:969–977.

Boivin, T., J. Candau, A. Roques, and M.-A. Auger-Rozenberg. 2008. First finding of *Wolbachia* in thelytokous seed chalcids (Hymenoptera: Torymidae, *Megastigmus* sp). XXIII International Congress of Entomology. Durban, Republic of South Africa.

Boivin, T., and J.-N. Candau. 2007. Flow cytometric analysis of ploidy levels in two seed-infesting *Megastigmus* species: applications to sex ratio and species determination at the larval stage. Entomologia Experimentalis et Applicata 124:125–131.

Boivin, T., H. Henri, F. Vavre, C. Gidoin, J.-N. Candau, E. Magnoux, A. Roques, and M. A. Auger-Rozenber. 2013. Epidemiology of thelytoky induced by the endosymbiotic Wolbachia across seed-specialized wasp species: host plant specialization matters. Manuscript submitted for publication.

Van Borm, S., A. Buschinger, J. J. Boomsma, and J. Billen. 2002. *Tetraponera* ants have gut symbionts related to nitrogen-fixing root-nodule bacteria. Proceedings of the Royal Society of London B: Biological Sciences 269:2023–2027.

Bracke, J. W., D. L. Cruden, and A. J. Markovetz. 1979. Intestinal microbial flora of the of the American cockroach, *Periplaneta americana* L. Applied and Environmental Microbiology 38:945–955.

Bräutigam, A., T. Mullick, S. Schliesky, and A. P. M. Weber. 2011. Critical assessment of assembly strategies for non-model species mRNA-seq data and application of next-

generation sequencing to the comparison of C_3 and C_4 species. Journal of Experimental Botany 62:3093–3102.

Brenner, D. J., N. R. Krieg, and J. T. Staley (Eds.). 2005. Volume 2, Part C, The Proteobacteria; The Alpha-, Beta-, Delta-, and Epsilonproteobacteria. Bergey's Manual of Systematic Bacteriology, 2nd edition. Springer, New York.

Breznak, J. A. 1982. Intestinal microbiota of termites and other xylophagous insects. Annual Review of Microbiology 36:323–343.

Brooks, S. E., and J. D. Shorthouse. 1998. Developmental morphology of stem galls of *Diplolepis nodulosa* (Hymenoptera: Cynipidae) and those modified by the inquiline *Periclistus pirata* (Hymenoptera: Cynipidae) on *Rosa blanda* (Rosaceae). Canadian Journal of Botany 76:365–381.

Brucker, R. M., and S. R. Bordenstein. 2012. The roles of host evolutionary relationships (genus: *Nasonia*) and development in structuring microbial communities. Evolution 66:349–362.

Buckley, R. 1987. Interactions involving plants, Homoptera, and ants. Annual Review of Ecology and Systematics 18:111–135.

Bution, M. L., and F. H. Caetano. 2010. Symbiotic bacteria and the structural specializations in the ileum of *Cephalotes* ants. Micron 41:373–381.

Capella-Gutiérrez, S., M. Marcet-Houben, and T. Gabaldón. 2012. Phylogenomics supports microsporidia as the earliest diverging clade of sequenced fungi. BMC Biology 10:47.

Caporaso, J. G., J. Kuczynski, J. Stombaugh, K. Bittinger, F. D. Bushman, E. K. Costello, N. Fierer, A. G. Peña, J. K. Goodrich, J. I. Gordon, G. A. Huttley, S. T. Kelley, D. Knights, J. E. Koenig, R. E. Ley, C. A. Lozupone, D. Mcdonald, B. D. Muegge, M. Pirrung, J. Reeder, J. R. Sevinsky, P. J. Turnbaugh, W. A. Walters, J. Widmann, T. Yatsunenko, J. Zaneveld, and R. Knight. 2010. QIIME allows analysis of high-throughput community sequencing data. Nature Methods 7:335–336.

Casewell, N. R., G. A. Huttley, and W. Wüster. 2012. Dynamic evolution of venom proteins in squamate reptiles. Nature Communications 3:1–10.

Casewell, N. R., W. Wüster, F. J. Vonk, R. A. Harrison, and B. G. Fry. 2013. Complex cocktails: the evolutionary novelty of venoms. Trends in Ecology and Evolution 28:1–11.

Casiraghi, M., S. R. Bordenstein, L. Baldo, N. Lo, T. Beninati, J. J. Wernegreen, J. H. Werren, C. Bandi, and C. Bandi. 2005. Phylogeny of *Wolbachia pipientis* based on gltA, groEL and ftsZ gene sequences: clustering of arthropod and nematode symbionts in the F supergroup, and evidence for further diversity in the *Wolbachia* tree. Microbiology 151:4015–4022.

Caspi-Fluger, A., M. Inbar, N. Mozes-Daube, N. Katzir, V. Portnoy, E. Belausov, M. S. Hunter, and E. Zchori-Fein. 2012. Horizontal transmission of the insect symbiont *Rickettsia* is plant-mediated. Proceedings of the Royal Society B: Biological Sciences 279:1791–1796.

Casteels, P., C. Ampe, F. Jacobs, and P. Tempst. 1993. Functional and chemical characterization of hymenoptaecin, antibacterial polypeptide that is infection-inducible in Honeybee (*Apis mellifera*). The Journal of Biological Chemistry 268:7044–7054.

Centrella, M. L., and S. R. Shaw. 2010. A new species of phytophagous braconid *Allorhogas minimus* (Hymenoptera: Braconidae: Doryctinae) reared from fruit galls on *Miconia longifolia* (Melastomataceae) in Costa Rica. International Journal of Tropical Insect Science 30:101–107.

Chen, D., C. B. Montllor, and A. H. Purcell. 2000. Fitness effects of two facultative endosymbiotic bacteria on the pea aphid, *Acyrthosiphon pisum*, and the blue alfalfa aphid, *A. kondoi*. Entomologia Experimentalis et Applicata 95:315–323.

Chiwocha, S., G. Rouault, S. Abrams, and P. von Aderkas. 2007. Parasitism of seed of Douglas fir (*Pseudotsuga menziesii*) by the seed chalcid, *Megastigmus spermotrophus*, and its influence on seed hormone physiology. Sexual Plant Reproduction 20:19–25.

Colinet, D., E. Deleury, C. Anselme, D. Cazes, J. Poulain, C. Azema-Dossat, M. Belghazi, J.-L. Gatti, and M. Poirié. 2013. Extensive inter- and intraspecific venom variation in closely related parasites targeting the same host: the case of *Leptopilina* parasitoids of *Drosophila*. Insect Biochemistry and Molecular Biology 43:601–611.

Colman, D. R., E. C. Toolson, and C. D. Takacs-Vesbach. 2012. Do diet and taxonomy influence insect gut bacterial communities? Molecular Ecology 20:5124–5137.

Connor, E. F., and M. P. Taverner. 1997. The evolution and adaptive significance of the leaf-mining habit. Oikos 79:6–25.

Cook, J. M., and J.-Y. Rasplus. 2003. Mutualists with attitude: coevolving fig wasps and figs. Trends in Ecology and Evolution 18:241–248.

Cornell, H. V. 1983. The secondary chemistry and complex morphology of galls formed by the Cynipinae (Hymenoptera): Why and how? American Midland Naturalist 110:225–234.

Crawford, J. E., W. M. Guelbeogo, A. Sanou, A. Traoré, K. D. Vernick, N. Sagnon, and B. P. Lazzaro. 2010. *De novo* transcriptome sequencing in *Anopheles funestus* using Illumina RNA-seq technology. PloS One 5:e14202.

Csóka, G., G. N. Stone, and G. Melika. 2005. Biology, Ecology, and Evolution of Gall-inducing Cynipidae. Pages 573–642 *in* A. Raman, C. W. Schaefer, and T. M. Withers, editors. Biology, Ecology, and Evolution of Gall-inducing Arthropods, Volume 2. Science Publishers, Inc., Enfield, USA.

Dahlman, D. L. 1990. Evaluation of teratocyte functions: An overview. Archives of Insect Biochemistry and Physiology 13:159–166.

Dale, C., and N. A. Moran. 2006. Molecular interactions between bacterial symbionts and their hosts. Cell 126:453–465.

Danforth, B. N., S. Sipes, and J. Fang. 2006. The history of early bee diversification based on five genes plus morphology. Proceedings of the National Academy of Sciences 103:15118–15123.

Dani, M. P., J. P. Edwards, and E. H. Richards. 2005. Hydrolase activity in the venom of the pupal endoparasitic wasp, *Pimpla hypochondriaca*. Comparative Biochemistry and Physiology. 141B:373–381.

Davis, R. B., S. L. Baldauf, and P. J. Mayhew. 2010. The origins of species richness in the Hymenoptera: Insights from a family-level supertree. BMC Evolutionary Biology 10:109.

Davison, A., and M. Blaxter. 2005. Ancient origin of glycosyl hydrolase family 9 cellulase genes. Molecular Biology and Evolution 22:1273–1284.

Degnan, P. H., A. B. Lazarus, and J. J. Wernegreen. 2005. Genome sequence of *Blochmannia pennsylvanicus* indicates parallel evolutionary trends among bacterial mutualists of insects. Genome Research 15:1023–1033.

Degnan, P., A. Lazarus, C. Brock, and J. Wernegreen. 2004. Host-symbiont stability and fast evolutionary rates in an ant-bacterium association: Cospeciation of *Camponotus* species and their endosymbionts, *Candidatus Blochmannia*. Systematic Biology 53:95–110.

Diamond, S. E., C. P. Blair, and W. G. Abrahamson. 2008. Testing the nutrition hypothesis for the adaptive nature of insect galls: Does a non-adapted herbivore perform better in galls? Ecological Entomology 33:385–393.

Dillon, R. J., and V. M. Dillon. 2004. The gut bacteria of insects: nonpathogenic interactions. Annual Review of Entomology 49:71–92.

Dorémus, T., S. Urbach, V. Jouan, F. Cousserans, M. Ravallec, E. Demettre, E. Wajnberg, J. Poulain, C. Azéma-Dossat, I. Darboux, J.-M. Escoubas, D. Colinet, J.-L. Gatti, M. Poirié, and A.-N. Volkoff. 2013. Venom gland extract is not required for successful parasitism in the polydnavirus-associated endoparasitoid *Hyposoter didymator* (Hym. Ichneumonidae) despite the presence of numerous novel and conserved venom proteins. Insect Biochemistry and Molecular Biology 43:292–307.

Douglas, A. E. 1998. Nutritional interactions in insect-microbial symbiosis: Aphids and their symbiotic bacteria *Buchnera*. Annual Review of Entomology 43:17–37.

Douglas, A. E. 2009. The microbial dimension in insect nutritional ecology. Functional Ecology 23:38–47.

Dowd, S. E., T. R. Callaway, R. D. Wolcott, Y. Sun, T. McKeehan, R. G. Hagevoort, and T. S. Edrington. 2008. Evaluation of the bacterial diversity in the feces of cattle using 16S rDNA bacterial tag-encoded FLX amplicon pyrosequencing (bTEFAP). BMC Microbiology 8:125.

Dray, S., and A.-B. Dufour. 2007. The ade4 package: Implementing the duality diagram for ecologists. Journal of Statistical Software 22:1–20.

Dreger-Jauffret, F., and J. D. Shorthouse. 1992. Diversity of Gall-inducing Insects and Their Galls. Pages 8–33 *in* J. D. Shorthouse and O. Rohfritsch, editors. Biology of Insect-induced Galls. Oxford University Press, New York, USA.

Duron, O., D. Bouchon, S. Boutin, L. Bellamy, L. Zhou, J. Engelstädter, and G. D. Hurst. 2008. The diversity of reproductive parasites among arthropods: *Wolbachia* do not walk alone. BMC Biology 6:27.

Edgar, R. C. 2010. Search and clustering orders of magnitude faster than BLAST. Bioinformatics 26:2460–2461.

Edgar, R. C., B. J. Haas, J. C. Clemente, C. Quince, and R. Knight. 2011. UCHIME improves sensitivity and speed of chimera detection. Bioinformatics 27:2194–2200.

Eggleton, P., and R. Belshaw. 1992. Insect parasitoids: An evolutionary overview. Philosophical Transactions: Biological Sciences 337:1–20.

Eggleton, P., and K. J. Gaston. 1990. "Parasitoid" species and assemblages: Convenient definitions or misleading compromises? Oikos 59:417–421.

Elzen, G. W. 1983. Minireview: Cytokinins and insect galls. Comparative Biochemistry and Physiology 76:17–19.

Engel, P., and N. A. Moran. 2013. The gut microbiota of insects - diversity in structure and function. FEMS Microbiology Reviews 37:699–735.

Fabre, J.-P., M.-A. Auger-Rozenberg, A. Chalon, S. Boivin, and A. Roques. 2004. Competition between exotic and native insects for seed resources in trees of a Mediterranean forest ecosystem. Biological Invasions 6:11–22.

Feldhaar, H. 2011. Bacterial symbionts as mediators of ecologically important traits of insect hosts. Ecological Entomology 36:533–543.

Feldhaar, H., J. Straka, M. Krischke, K. Berthold, S. Stoll, M. J. Mueller, and R. Gross. 2007. Nutritional upgrading for omnivorous carpenter ants by the endosymbiont *Blochmannia*. BMC Biology 5:48.

Folmer, O., M. Black, W. Hoeh, R. Lutz, and R. Vrijenhoek. 1994. DNA primers for amplification of mitochondrial cytochrome c oxidase subunit I from diverse metazoan invertebrates. Molecular Marine Biology and Biotechnology 3:294–299.

Forbes, A., and R. Lehmann. 1998. Nanos and Pumilio have critical roles in the development and function of *Drosophila* germline stem cells. Development 125:679–690.

Frago, E., M. Dicke, and H. C. J. Godfray. 2012. Insect symbionts as hidden players in insect-plant interactions. Trends in Ecology and Evolution 27:705–711.

Frugier, F., S. Kosuta, J. D. Murray, M. Crespi, and K. Szczyglowski. 2008. Cytokinin: Secret agent of symbiosis. Trends in Plant Science 13:115–120.

Fry, B. G., N. R. Casewell, W. Wüster, N. Vidal, B. Young, and T. N. W. Jackson. 2012. The structural and functional diversification of the Toxicofera reptile venom system. Toxicon 60:434–448.

Fry, B. G., K. Roelants, D. E. Champagne, H. Scheib, J. D. A. Tyndall, G. F. King, T. J. Nevalainen, J. A. Norman, R. J. Lewis, R. S. Norton, C. Renjifo, and R. C. R. de la Vega. 2009. The toxicogenomic multiverse: convergent recruitment of proteins into animal venoms. Annual Review of Genomics and Human Genetics 10:483–511.

Fry, B. G., W. Wüster, R. M. Kini, V. Brusic, A. Khan, D. Venkataraman, and A. P. Rooney. 2003. Molecular evolution and phylogeny of elapid snake venom three-finger toxins. Journal of Molecular Evolution 57:110–129.

Gange, A. C., and H. E. Nice. 1997. Performance of the thistle gall fly, *Urophora cardui*, in relations to host plant nitrogen and mycorrhizal colonization. New Phytologist 137:335–343.

Genta, F. A., R. J. Dillon, W. R. Terra, and C. Ferreira. 2006. Potential role for gut microbiota in cell wall digestion and glucoside detoxification in *Tenebrio molitor* larvae. Journal of Insect Physiology 52:593–601.

Gil, R., F. J. Silva, E. Zientz, F. Delmotte, F. González-Candelas, A. Latorre, C. Rausell, J. Kamerbeek, J. Gadau, B. Hölldobler, R. C. H. J. van Ham, R. Gross, and A. Moya. 2003. The genome sequence of *Blochmannia floridanus*: Comparative analysis of reduced genomes. Proceedings of the National Academy of Sciences 100:9388–9393.

Giorgini, M., U. Bernardo, M. M. Monti, A. G. Nappo, and M. Gebiola. 2010. *Rickettsia* symbionts cause parthenogenetic reproduction in the parasitoid wasp *Pnigalio soemius* (Hymenoptera: Eulophidae). Applied and Environmental Microbiology 76:2589–2599.

Girin, C., and M. Boulétreau. 1995. Microorganism-associated variation in host infestation efficiency in a parasitoid wasp, *Trichogramma bourarachae* (Hymenoptera: Trichogrammatidae). Experientia 51:398–401.

Giron, D., E. Frago, G. Glevarec, C. M. J. Pieterse, and M. Dicke. 2013. Cytokinins as key regulators in plant-microbe-insect interactions: Connecting plant growth and defence. Functional Ecology 27:599–609.

Giron, D., W. Kaiser, N. Imbault, and J. Casas. 2007. Cytokinin-mediated leaf manipulation by a leafminer caterpillar. Biology Letters 3:340–343.

Glasgow, H. 1914. The gastric caeca and the caecal bacteria of the Heteroptera. Biological Bulletin 26:101–170.

Goto, A., K. Matsushita, V. Gesellchen, L. E. Chamy, D. Kattenkeuler, O. Takeuchi, J. A. Hoffmann, S. Akira, M. Boutros, and J.-M. Reichart. 2008. Akirins, highly conserved nuclear proteins, required for NF-kB dependent gene expression in *Drosophila* and mice. Nature Immunology 9:97–104.

Gottlieb, Y., M. Ghanim, G. Gueguen, S. Kontsedalov, F. Vavre, F. Fleury, and E. Zchori-Fein. 2008. Inherited intracellular ecosystem: symbiotic bacteria share bacteriocytes in whiteflies. The FASEB Journal 22:2591–2599.

De Graaf, D. C., M. Aerts, M. Brunain, C. A. Desjardins, F. J. Jacobs, J. H. Werren, and B. Devreese. 2010. Insights into the venom composition of the ectoparasitoid wasp *Nasonia vitripennis* from bioinformatic and proteomic studies. Insect Molecular Biology 19:11–26.

Grenier, S., B. Pintureau, A. Heddi, C. Jager, and C. Louis. 1998. Successful horizontal transfer of *Wolbachia* symbionts between *Trichogramma* wasps. Proceedings of the Royal Society of London B: Biological Sciences 265:1441–1445.

Grissell, E. E. 1999a. Hymenopteran biodiversity: Some alien notions. American Entomoligist 45:235–244.

Grissell, E. E. 1999b. An annotated catalog of world Megastigminae (Hymenoptera: Chalcidoidea: Torymidae). Contributions of the American Entomological Institute 31:1–92.

Guidugli, K. R., M.-D. Piulachs, X. Bellés, A. P. Lourenço, and Z. L. P. Simões. 2005. Vitellogenin expression in queen ovaries and in larvae of both sexes of *Apis mellifera*. Archives of Insect Biochemistry and Physiology 59:211–218.

Hagimori, T., Y. Abe, S. Date, and K. Miura. 2006. The first finding of a *Rickettsia* bacterium associated with parthenogenesis induction among insects. Current Microbiology 52:97–101.

Haine, E. R., N. J. Pickup, and J. M. Cook. 2005. Horizontal transmission of *Wolbachia* in a *Drosophila* community. Ecological Entomology 20:464–472.

Hakim, R. S., K. Baldwin, and G. Smagghe. 2010. Regulation of midgut growth, development, and metamorphosis. Annual Review of Entomology 55:593–608.

Hallmann, J., A. Quadt-Hallmann, W. F. Mahaffee, and J. W. Kloepper. 1997. Bacterial endophytes in agricultural crops. Canadian Journal of Microbiology 43:895–914.

Hamady, M., C. Lozupone, and R. Knight. 2010. Fast UniFrac: facilitating high-throughput phylogenetic analyses of microbial communities including analysis of pyrosequencing and PhyloChip data. The ISME Journal 4:17–27.

Hansen, K. D., R. A. Irizarry, and Z. Wu. 2012. Removing technical variability in RNA-seq data using conditional quantile normalization. Biostatistics 13:204–216.

Hartley, S. E. 1998. The chemical composition of plant galls: are levels of nutrients and secondary compounds controlled by the gall-former? Oecologia 113:492–501.

Hartley, S. E., and J. H. Lawton. 1992. Host-plant manipulation by gall-insects: A test of the nutrition hypothesis. Journal of Animal Ecology 61:113–119.

Hawkins, B. A., and R. D. Goeden. 1982. Biology of a gall-forming *Terastichus* (Hymenoptera: Eulophidae) associated with gall midges on saltbush in southern California. Annals of the Entomological Society of America 75:444–447.

Heil, M., and D. McKey. 2003. Protective ant-plant interactions as model systems in ecological and evolutionary research. Annual Review of Ecology, Evolution and Systematics 34:425–453.

Henry, L. M., J. Peccoud, J.-C. Simon, J. D. Hadfield, M. J. C. Maiden, J. Ferrari, and H. C. J. Godfray. 2013. Horizontally transmitted symbionts and host colonization of ecological niches. Current Biology 23:1–5.

Heraty, J., F. Ronquist, J. M. Carpenter, D. Hawks, S. Schulmeister, A. P. Dowling, D. Murray, J. Munro, W. C. Wheeler, N. Schiff, and M. Sharkey. 2011. Evolution of the hymenopteran megaradiation. Molecular Phylogenetics and Evolution 60:73–88.

Hirt, R. P., J. M. Logsdon, B. Healy, M. W. Dorey, W. F. Doolittle, and T. M. Embley. 1999. Microsporidia are related to Fungi: Evidence from the largest subunit of RNA polymerase II and other proteins. Proceedings of the National Academy of Sciences 96:580–585.

Hoffman, J. A., and C. Hetru. 1992. Insect defensins: Inducible anitbacterial peptides. Immunology Today 13:411–415.

Hölldobler, B., and E. O. Wilson. 1990. The Ants. Belknap, Cambridge, Mass.

Hosokawa, T., Y. Kikuchi, X. Y. Meng, and T. Fukatsu. 2005. The making of symbiont capsule in the plataspid stinkbug *Megacopta punctatissima*. FEMS microbiology ecology 54:471–477.

Hosokawa, T., Y. Kikuchi, M. Shimada, and T. Fukatsu. 2007. Obligate symbiont involved in pest status of host insect. Proceedings of the Royal Society B: Biological Sciences 274:1979–1984.

Hovanitz, W. 1959. Insect and plant galls. Scientific American 201:151–162.

Huigens, M. E., R. P. de Almeida, P. A. H. Boons, R. F. Luck, and R. Stouthamer. 2004. Natural interspecific and intraspecific horizontal transfer of parthenogenesis-inducing *Wolbachia* in *Trichogramma* wasps. Proceedings of the Royal Society of London B: Biological Sciences 271:509–515.

Husnik, F., N. Nikoh, R. Koga, L. Ross, R. P. Duncan, M. Fujie, M. Tanaka, N. Satoh, D. Bachtrog, A. C. C. Wilson, C. D. von Dohlen, T. Fukatsu, and J. P. McCutcheon. 2013. Horizontal gene transfer from diverse bacteria to an insect genome enables a tripartite nested mealybug symbiosis. Cell 153:1567–1578.

Hussey, N. W. 1955. The life histories of *Megastigmus spermotrophus* Wachtl. (Hym. Chalcidoidea) and its principal parasite, with descriptions of the developmental stages. Transactions of the Royal Entomological Society of London 106:133–151.

Infante, F., P. Hanson, and R. Wharton. 1995. Phytophagy in the Genus *Monitoriella* (Hymenoptera: Braconidae) with description of new species. Annals of the Entomological Society of America 88:406–415.

James, H. C. 1926. The anatomy of a British phytophagous Chalcidoid of the genus *Harmolita* (Isosoma). Proceedings of the Zoological Society of London 96:75–182.

Jameson, P. 2000. Cytokinins and auxins in plant-pathogen interactions – An overview. Plant Growth Regulation 32:369–380.

Janson, E. M., J. O. Stireman, M. S. Singer, and P. Abbot. 2008. Phytophagous insect-microbe mutualisms and adaptive evolutionary diversification. Evolution 62:997–1012.

Janssen, P. H. 2006. Identifying the dominant soil bacterial taxa in libraries of 16S rRNA and 16S rRNA genes. Applied and Environmental Microbiology 72:1719–1728.

Jeyaprakash, A., M. A. Hoy, and M. H. Allsopp. 2003. Bacterial diversity in worker adults of *Apis mellifera capensis* and *Apis mellifera scutellata* (Insecta: Hymenoptera) assessed using 16S rRNA sequences. Journal of Invertebrate Pathology 84:96–103.

Jiggins, F. M., J. K. Bentley, M. E. Majerus, and G. D. Hurst. 2001. How many species are infected with *Wolbachia*? Cryptic sex ratio distorters revealed to be common by intensive sampling. Proceedings of the Royal Society of London B: Biological Sciences 268:1123–1126.

Jiggins, F. M., J. K. Bentley, M. E. N. Majerus, and G. D. D. Hurst. 2002. Recent changes in phenotype and patterns of host specialization in *Wolbachia* bacteria. Molecular Ecology 11:1275–1283.

Jones, K. G., P. F. Dowd, and M. Blackwell. 1999. Polyphyletic origins of yeast-like endocytobionts from anobiid and cerambycid beetles. Mycological Research 103:542–546.

Jones, R. T., K. F. McCormick, and A. P. Martin. 2008. Bacterial communities of *Bartonella*-positive fleas: diversity and community assembly patterns. Applied and Environmental Microbiology 74:1667–1670.

Kaiser, W., E. Huguet, J. Casas, C. Commin, and D. Giron. 2010. Plant green-island phenotype induced by leaf-miners is mediated by bacterial symbionts. Proceedings of the Royal Society B: Biological Sciences 277:2311–2319.

Kashima, T., T. Nakamura, and S. Tojo. 2006. Uric acid recycling in the shield bug, *Parastrachia japonensis* (Hemiptera: Parastrachiidae), during diapause. Journal of Insect Physiology 52:816–25.

Kautz, S., B. E. R. Rubin, J. A. Russell, and C. S. Moreau. 2013. Surveying the microbiome of ants: comparing 454 pyrosequencing with traditional methods to uncover bacterial diversity. Applied and Environmental Microbiology 79:525–534.

Kikuchi, Y., T. Hosokawa, and T. Fukatsu. 2007. Insect-microbe mutualism without vertical transmission: A stinkbug acquires a beneficial gut symbiont from the environment every generation. Applied and Environmental Microbiology 73:4308–4316.

Kikuchi, Y., T. Hosokawa, and T. Fukatsu. 2011. An ancient but promiscuous host-symbiont association between *Burkholderia* gut symbionts and their heteropteran hosts. The ISME Journal 5:446–460.

Kikuchi, Y., X. Meng, and T. Fukatsu. 2005. Gut symbiotic bacteria of the genus *Burkholderia* in the broad-headed bugs *Riptortus clavatus* and *Leptocorisa chinensis* (Heteroptera: Alydidae). Applied and Environmental Microbiology 71:4035–4043.

Kim, Y. S., J. H. Ryu, S. J. Han, K. H. Choi, K. B. Nam, I. H. Jang, B. Lemaitre, P. T. Brey, and W. J. Lee. 2000. Gram-negative bacteria-binding protein, a pattern recognition receptor for lipopolysaccharide and beta-1,3-glucan that mediates the signaling for the induction of innate immune genes in *Drosophila melanogaster* cells. The Journal of Biological Chemistry 275:32721–32727.

Kimmerer, T. W., and D. A. Potter. 1987. Quality of specific leaf tissues and selective feeding by a specialist leafminer. Oecologia 71:548–551.

King, G. F. 2011. Venoms as a platform for human drugs: translating toxins into therapeutics. Expert Opinion on Biological Therapy 11:1469–1484.

King, J. E., and D. J. Gifford. 1997. Amino acid utilization in seeds of Loblolly pine during germination and early seedling growth. 1. Arginine and arginase activity. Plant Physiology 113:1125–1135.

Kinsey, A. C. 1920. Phylogeny of cynipid genera and biological characteristics. Bulletin American Museum of Natural History 42:357–402.

Kjellberg, F., E. Jousselin, M. Hossaert-McKey, and J.-Y. Rasplus. 2005. Biology, Ecology, and Evolution of Fig-pollinating Wasps (Chalcidoidea, Agaonidae). Pages 539–572 *in* A. Raman, C. W. Schaefer, and T. M. Withers, editors. Biology, Ecology, and Evolution of Gall-inducing Arthropods, Volume 2. Science Publishers, Inc., Enfield, USA.

Kordiš, D., and F. Gubenšek. 2000. Adaptive evolution of animal toxin multigene families. Gene 261:43–52.

Koyama, Y., I. Yao, and S.-I. Akimoto. 2004. Aphid galls accumulate high concentrations of amino acids: A support for the nutrition hypothesis for gall formation. Entomologia Experimentalis et Applicata 113:35–44.

Kraińska, M. 1966. Histochemical study of acid mucopolysaccharides in the oviduct gland of *Cynips folii*. Folia Histochemica et Cytochemica 4:103–110.

Kukor, J. J., and M. M. Martin. 1983. Acquisition of digestive enzymes by siricid woodwasps from their fungal symbiont. Science 220:1161–1163.

Lasalle, J. 2005. Biology of Gall Inducers and Evolution of Gall Induction in Chalcidoidea (Hymenoptera: Eulophidae, Eurytomidae, Ptermalidae, Tanaostigmatidae, Torymidae). Pages 507–537 *in* A. Raman, C. W. Schaefer, and T. M. Withers, editors. Biology, Ecology, and Evolution of Gall-inducing Arthropods, Volume 1 and 2. Science Publishers, Inc., Enfield, USA.

Lasalle, J., and I. D. Gauld. 1991. Parasitic Hymenoptera and the biodiversity crisis. Redia 74:315–334.

LaSalle, J., and I. D. Gauld. 1993. Hymenoptera: Their diversity, and their impact on the diversity of other organisms. Pages 1–26 *in* J. LaSalle and I. D. Gauld, editors. Hymenoptera and Biodiversity. C.A.B. International, Wallingford, UK.

Lawson, E. T., T. A. Mousseau, R. Klaper, M. D. Hunter, and J. H. Werren. 2001. *Rickettsia* associated with male-killing in a buprestid beetle. Heredity 86:497–505.

Leggo, J. J., and J. D. Shorthouse. 2006. Development of stem galls induced by *Diplolepis triforma* (Hymenoptera: Cynipidae) on *Rosa acicularis* (Rosaceae). Canadian Entomologist 138:661–680.

Lehmann, R., and C. Nüsslein-Volhard. 1986. Abdominal segmentation, pole cell formation, and embryonic polarity require the localized activity of oskar, a maternal gene in *Drosophila*. Cell 47:141–152.

Li, B., and C. N. Dewey. 2011. RSEM: accurate transcript quantification from RNA-seq data with or without a reference genome. BMC Bioinformatics 12:323.

Li, W., and A. Godzik. 2006. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics 22:1658–1659.

Liljeblad, J., and F. Ronquist. 1998. A phylogenetic analysis of higher-level gall wasp relationships (Hymenoptera: Cynipidae). Systematic Entomology 23:229–252.

Liu, S., N. P. Chougule, D. Vijayendran, and B. C. Bonning. 2012. Deep Sequencing of the transcriptomes of soybean aphid and associated endosymbionts. PLoS One 7:e45161.

Liu, Y., G. S. Dunn, and N. N. Aronson. 1996. Purification, biochemistry and molecular cloning of an insect glycosylasparaginase from *Spodoptera frugiperda*. Glycobiology 6:527–536.

Lohse, M., A. M. Bolger, A. R. Fernie, J. E. Lunn, M. Stitt, and B. Usadel. 2012. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. Nucleic Acids Research 40:W622–627.

Lozupone, C., and R. Knight. 2005. UniFrac: A new phylogenetic method for comparing microbial communities. Applied and Environmental Microbiology 71:8228–8235.

Ludwig, W., K.-H. Schleifer, and W. B. Whitman. 2008. Revised Road Map to the Phylum Firmicutes. Bergey's Manual of Systematic Bacteriology. Vol. 3. Springer, New York.

Łukasik, P., M. van Asch, H. Guo, J. Ferrari, and H. C. J. Godfray. 2013. Unrelated facultative endosymbionts protect aphids against a fungal pathogen. Ecology Letters 16:214–218.

Lundgren, J. G., and R. M. Lehman. 2010. Bacterial gut symbionts contribute to seed digestion in an omnivorous beetle. PloS One 5:e10831.

De Macêdo, M. V, M. C. P. Pimentel, and R. C. Vieira. 1998. Response of *Pithecellobium tortum* Martius (Leguminosae) seeds to the attack of the phytophagous Braconid *Allorhogas dyspistus* Marsh (Hymenoptera: Braconidae). Journal of Hymenoptera Research 7:274–279.

De Macêdo, M. V., and R. F. Monteiro. 1989. Seed predation by a braconid wasp, *Allorhogas* sp. (Hymenoptera). Journal of New York Entomological Society 97:358–362.

Machtelinckx, T., T. Van Leeuwen, T. Van De Wiele, N. Boon, W. H. De Vos, J.-A. Sanchez, M. Nannini, G. Gheysen, and P. De Clercq. 2012. Microbial community of predatory bugs of the genus *Macrolophus* (Hemiptera: Miridae). BMC Microbiology 12:S9.

Malyshev, S. I. 1968. Genesis of the Hymenoptera and the Phases of their Evolution. Methuen & Co. Ltd. (English translation of Russian publication.), London, UK.

Mapes, C. C., and P. J. Davies. 2001a. Indole-3-acetic acid and ball gall development on *Solidago altissima*. New Phytologist 151:195–202.

Mapes, C. C., and P. J. Davies. 2001b. Cytokinins in the ball gall of *Solidago altissima* and in the gall forming larvae of *Eurosta solidaginis*. New Phytologist 151:203–212.

Marchesi, J. R., T. Sato, A. J. Weightman, T. A. Martin, J. C. Fry, S. J. Hiom, D. Dymock, and W. G. Wade. 1998. Design and evaluation of useful bacterium-specific PCR primers that amplify genes coding for bacterial 16S rRNA. Applied and Environmental Microbiology 64:795–799.

Van Marle, J., and T. Piek. 1986. Morphology of the Venom Apparatus. Pages 17–44 *in* T. Piek, editor. Venoms of the Hymenoptera. Academic Press, London.

Martinson, V. G., B. N. Danforth, R. L. Minckley, O. Rueppell, S. Tingek, and N. A. Moran. 2011. A simple and distinctive microbiota associated with honey bees and bumble bees. Molecular Ecology 20:619–628.

Martinson, V. G., J. Moy, and N. A. Moran. 2012. Establishment of characteristic gut bacteria during development of the honeybee worker. Applied and Environmental Microbiology 78:2830–2840.

McCalla, D. R., M. K. Genthe, and W. Hovanitz. 1962. Chemical nature of an insect gall growth-factor. Plant Physiology 37:98–103.

McDonald, D., M. N. Price, J. Goodrich, E. P. Nawrocki, T. Z. DeSantis, A. Probst, G. L. Andersen, R. Knight, and P. Hugenholtz. 2012. An improved Greengenes taxonomy with explicit ranks for ecological and evolutionary analyses of bacteria and archaea. The ISME Journal 6:610–618.

Mcmurdie, P. J., and S. Holmes. 2012. Phyloseq: A bioconductor package for handling and analysis of high-throughput phylogenetic sequence data. Pages 235–246 Pacific Symposium on Biocomputing.

Mohr, K. I., and C. C. Tebbe. 2006. Diversity and phylotype consistency of bacteria in the guts of three bee species (Apoidea) at an oilseed rape field. Environmental Microbiology 8:258–272.

Montenegro, H., V. N. Solferini, L. B. Klaczko, and G. D. D. Hurst. 2005. Male-killing *Spiroplasma* naturally infecting *Drosophila melanogaster*. Insect Molecular Biology 14:281–287.

Moran, N. A., A. K. Hansen, J. E. Powell, and Z. L. Sabree. 2012. Distinctive gut microbiota of honey bees assessed using deep sampling from individual worker bees. PloS one 7:e36393.

Moran, N. A., J. P. McCutcheon, and A. Nakabachi. 2008. Genomics and evolution of heritable bacterial symbionts. Annual Review of Genetics 42:165–90.

Moreau, S. J. M., A. Cherqui, G. Doury, F. Dubois, Y. Fourdrain, L. Sabatier, P. Bulet, J. Saarela, G. Prévost, and P. Giordanengo. 2004. Identification of an aspartylglucosaminidase-like protein in the venom of the parasitic wasp *Asobara tabida* (Hymenoptera: Braconidae). Insect Biochemistry and Molecular Biology 34:485–492.

Moreau, S. J. M., A. Dingremont, G. Doury, and P. Giordanengo. 2002. Effects of parasitism by *Asobara tabida* (Hymenoptera: Braconidae) on the development, survival and activity of *Drosophila melanogaster* larvae. Journal of Insect Physiology 48:337–347.

Moreau, S. J. M., and S. Guillot. 2005. Advances and prospects on biosynthesis, structures and functions of venom proteins from parasitic wasps. Insect Biochemistry and Molecular Biology 35:1209–1223.

Mueller, U. G., S. A. Rehner, and T. R. Schultz. 1998. The evolution of agriculture in ants. Science 281:2034–2038.

Mukhopadhyay, K., N. K. Garrison, D. M. Hinton, C. W. Bacon, G. S. Khush, H. D. Peck, and N. Datta. 1996. Identification and characterization of bacterial endophytes of rice. Mycopathologia 134:151–159.

Mundt, J. O., and N. F. Hinkle. 1976. Bacteria within ovules and seeds. Applied and Environmental Microbiology 32:694–698.

Munoz-Torres, M. C., J. T. Reese, C. P. Childers, A. K. Bennett, J. P. Sundaram, K. L. Childs, J. M. Anzola, N. Milshina, and C. G. Elsik. 2011. Hymenoptera Genome Database: Integrated community resources for insect species of the order Hymenoptera. Nucleic Acids Research 39:D658–D662.

Munro, J. B., J. M. Heraty, R. A. Burks, D. Hawks, J. Mottern, A. Cruaud, J.-Y. Rasplus, and P. Jansta. 2011. A molecular phylogeny of the Chalcidoidea (Hymenoptera). PLoS One 6:e27023.

Nalepa, C. A., D. E. Bignell, and C. Bandi. 2001. Detritivory, coprophagy, and the evolution of digestive mutualisms in Dictyoptera. Insectes Sociaux 48:194–201.

Nawrocki, E. P., D. L. Kolbe, and S. R. Eddy. 2009. Infernal 1.0: inference of RNA alignments. Bioinformatics 25:1335–1337.

Neff, J. L., and B. B. Simpson. 1993. Bees, pollination systems and plant diversity. Pages 143–168 *in* J. Lasalle and I. D. Gauld, editors. Hymenoptera and Biodiversity. C.A.B. International, Wallingford, UK.

Nei, M., X. Gu, and T. Sitnikov. 1997. Evolution by the birth-and-death process in multigene families of the vertebrate immune system. Proceedings of the National Academy of Sciences 94:7799–7806.

Niwa, C. G., and D. L. Overhulser. 1992. Oviposition and development of *Megastigmus spermotrophus* (Hymenoptera: Torymidae) in unfertilized Douglas-fir seeds. Journal of Economic Entomology 85:2323–2328.

Noyes, J. S. 2013. Universal Chalcidoidea Database website. Available: www.nhm.ac.uk/entomology/chalcidoids/index.html. Accessed September 2, 2013.

Nylander, J., F. Ronquist, J. Huelsenbeck, and J. L. Nieves-aldrey. 2004. Bayesian phylogenetic analysis of combined data. Systematic Biology 53:47–67.

Nyman, T., and R. Julkunen-Tiitto. 2000. Manipulation of the phenolic chemistry of willows by gall-inducing sawflies. Proceedings of the National Academy of Sciences 97:13184–13187.

Ochiai, M. 2000. A pattern-recognition protein for ß-1,3-glucan: Rhe binding domain and the cDNA cloning of ß-1,3-glucan recognition protein from the silkworm, *Bombyx mori*. Journal of Biological Chemistry 275:4995–5002.

Olivier-espejel, A. S., Z. L. Sabree, K. Noge, and J. X. Becerra. 2011. Gut microbiota in nymph and adults of the giant mesquite bug (*Thasus neocalifornicus*) (Heteroptera: Coreidae) is dominated by *Burkholderia* acquired *de novo* every generation. Environmental Entomology 40:1102–1110.

Olofsson, T. C., and A. Vásquez. 2008. Detection and identification of a novel lactic acid bacterial flora within the honey stomach of the honeybee *Apis mellifera*. Current Microbiology 57:356–363.

Palavesam, A., F. D. Guerrero, A. M. Heekin, J. Wang, S. E. Dowd, Y. Sun, L. D. Foil, and A. A. Pérez de León. 2012. Pyrosequencing-based analysis of the microbiome associated with the Horn Fly, *Haematobia irritans*. PloS One 7:e44390.

Pant, R. 1988. Nitrogen excretion in insects. Proceedings: Animal Sciences 97:379–415.

Parkinson, N. M., C. M. Conyers, J. N. Keen, A. D. MacNicoll, I. Smith, and R. J. Weaver. 2003. cDNAs encoding large venom proteins from the parasitoid wasp *Pimpla hypochondriaca* identified by random sequence analysis. Comparative Biochemistry and Physiology 134C:513–520.

Pelosi, P., J.-J. Zhou, L. P. Ban, and M. Calvello. 2006. Soluble proteins in insect chemical communication. Cellular and Molecular Life Sciences 63:1658–1676.

Perlman, S. J., M. S. Hunter, and E. Zchori-Fein. 2006. The emerging diversity of *Rickettsia*. Proceedings of the Royal Society B: Biological Sciences 273:2097–2106.

Pertea, G., X. Huang, F. Liang, V. Antonescu, R. Sultana, S. Karamycheva, Y. Lee, J. White, F. Cheung, B. Parvizi, J. Tsai, and J. Quackenbush. 2003. TIGR Gene Indices clustering tools (TGICL): A software system for fast clustering of large EST datasets. Bioinformatics 19:651–652.

Pilgrim, E. M., C. D. von Dohlen, and J. P. Pitts. 2008. Molecular phylogenetics of Vespoidea indicate paraphyly of the superfamily and novel relationships of its component families and subfamilies. Zoologica Scripta 37:539–560.

Plantard, O., J. Rasplus, I. Le Clainche, and M. Solignac. 1998. *Wolbachia*-induced thelytoky in the rose gallwasp *Diplolepis spinosissimae* (Giraud) (Hymenoptera: Cynipidae), and its consequences on the genetic structure of its host. Proceedings of the Royal Society of London B: Biological Sciences 265:1075–1080.

Price, M. N., P. S. Dehal, and A. P. Arkin. 2010. FastTree 2 - Approximately maximum-likelihood trees for large alignments. PloS One 5:e9490.

Price, P. W. 1992. Evolution and Ecology of Gall-inducing Sawflies. Pages 208–224 *in* J. D. Shorthouse and O. Rohfritsch, editors. Biology of Insect-induced Galls. Oxford University Press, New York, USA.

Price, P. W., G. W. Fernandes, and G. L. Waring. 1987. The adaptive nature of insect galls. Environmental Entomology 16:15–24.

Price, P. W., G. L. Waring, and G. W. Fernandes. 1986. Hypotheses on the adaptive nature of galls. Proceedings of The Entomological Society of Washington 88:361–363.

Quicke, D. L. J. 1997. Parasitic Wasps. Chapman & Hall, London, UK.

R Development Core Team. 2013. R: A language and environmnet for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Rappaport, N., and A. Roques. 1991. Resource use and clonal differences in attack rate by the douglas-fir seed chalid *Megastigmus spermotrophus* Wachtl (Hymenoptera: Torymidae) in France. The Canadian Entomologist 123:1219–1228.

Reeder, J., and R. Knight. 2010. Rapidly denoising pyrosequencing amplicon reads by exploiting rank-abundance distributions. Nature Methods 7:668–669.

Robertson, G., J. Schein, R. Chiu, R. Corbett, M. Field, S. D. Jackman, K. Mungall, S. Lee, H. M. Okada, J. Q. Qian, M. Griffith, A. Raymond, N. Thiessen, T. Cezard, Y. S. Butterfield, R. Newsome, S. K. Chan, R. She, R. Varhol, B. Kamoh, A.-L. Prabhu, A. Tam, Y. Zhao, R. A. Moore, M. Hirst, M. A. Marra, S. J. M. Jones, P. A. Hoodless, and I. Birol. 2010. *De novo* assembly and analysis of RNA-seq data. Nature Methods 7:909–912.

Robertson, P. 1968. A morphological and functional study of the venom apparatus in representatives of some major groups of Hymenoptera. Australian Journal of Zoology 16:133–166.

Da Rocha, U. N., L. van Overbeek, and J. D. van Elsas. 2009. Exploration of hitherto-uncultured bacteria from the rhizosphere. FEMS Microbiology Ecology 69:313–328.

Rohfritsch, O. 1992. Patterns in Gall Development. Pages 60–86 *in* J. D. Shorthouse and O. Rohfritsch, editors. Biology of Insect-Induced Galls. Oxford University Press, New York.

Roininen, H., T. Nyman, and A. Zinovjev. 2005. Biology, Ecology, and Evolution of Gall-inducing Sawflies (Hymenoptera: Tenthredinidae and Xyelidae). Pages 467–494 *in* A. Raman, C. W. Schaefer, and T. M. Withers, editors. Biology, Ecology, and Evolution of Gall-inducing Arthropods, Volume 2. Science Publishers, Inc., Enfield, USA.

Rokas, A., R. J. Atkinson, J.-L. Nieves-Aldrey, S. A. West, and G. N. Stone. 2002. The incidence and diversity of *Wolbachia* in gallwasps (Hymenoptera; Cynipidae) on oak. Molecular Ecology 11:1815–1829.

Ronquist, F. 1994. Evolution of parasitism among closely related species: Phylogenetic relationships and the origin of inquilinism in gall wasps (Hymenoptera, Cynipidae). Evolution 48:241–266.

Ronquist, F., and J. Liljeblad. 2001. Evolution of the gall wasp-host plant association. Evolution 55:2503–2522.

Ronquist, F., A. P. Rasnitsyn, A. Roy, K. Eriksson, and M. Lindgren. 1999. Phylogeny of the Hymenoptera: A cladisitic reanalysis of Rasnitsyn's (1988) data. Zoologica Scripta 28:13–50.

Roques, A., J. P. Raimbault, and F. Goussard. 1984. La colonisation des cônes et galbules des genévriers méditerranéens par les insectes et acariens et son influence sur les possibilités de régénération naturelle des essences. Ecologia Mediterranea 10:147–169.

Roques, A., and M. Skrzypczyńska. 2003. Seed-infesting chalcids of the genus *Megastigmus* Dalman, 1820 (Hymenoptera: Torymidae) native and introduced to the West Palearctic region: Taxonomy, host specificity and distribution. Journal of Natural History 37:127–238.

Roques, A., J. Sun, M.-A. Auger-Rozenberg, and O. Hua. 2003. Potential invasion of China by exotic insect pests associated with tree seeds. Biodiversity and Conservation 12:2195–2210.

Rosenthal, G. A., D. H. Janzen, D. L. Dahlman, S. Url, N. Carolina, and C. Hill. 1977. Degradation and detoxification of canavanine by a specialized seed predator. Science 196:658–660.

Roskam, H. C. 1992. Evolution of the Gall-inducing Guild. Pages 34–49 *in* J. D. Shorthouse and O. Rohfritsch, editors. Biology of Insect-induced Galls. Oxford University Press, New York.

Rouault, G., J. Turgeon, J.-N. Candau, A. Roques, and P. von Aderkas. 2004. Oviposition strategies of conifer seed chalcids in relation to host phenology. Naturwissenschaften 91:472–480.

Roulston, T. H. I., and J. H. Cane. 2000. Pollen nutritional content and digestibility for animals. Plant Systematics and Evolution 222:187–209.

Roux, V., E. Rydkina, M. Eremeeva, and D. Raoult. 1997. Citrate synthase gene comparison, a new tool for phylogenetic analysis, and its application for the rickettsiae. International Journal of Systematic Bacteriology 47:252–261.

Russell, J. A., C. S. Moreau, B. Goldman-huertas, M. Fujiwara, D. J. Lohman, and N. E. Pierce. 2009. Bacterial gut symbionts are tightly linked with the evolution of herbivory in ants. Proceedings of the National Academy of Sciences 106:21236–21241.

Saarela, J., M. Laine, C. Oinonen, C. von Schantz, A. Jalanko, J. Rouvinen, and L. Peltonen. 2001. Molecular pathogenesis of a disease: structural consequences of aspartylglucosaminuria mutations. Human Molecular Genetics 10:983–995.

Sabree, Z. L., S. Kambhampati, and N. A. Moran. 2009. Nitrogen recycling and nutritional provisioning by *Blattabacterium*, the cockroach endosymbiont. Proceedings of the National Academy of Sciences 106:19521–19526.

Sakakibara, H. 2006. Cytokinins: activity, biosynthesis, and translocation. Annual Review of Plant Biology 57:431–449.

Sakurai, M., R. Koga, T. Tsuchida, X. Meng, and T. Fukatsu. 2005. *Rickettsia* symbiont in the Pea Aphid *Acyrthosiphon pisum*: Novel cellular tropism, effect on host fitness , and interaction with the essential symbiont *Buchnera*. Applied and Environmental Microbiology 71:4069–4075.

Schabereiter-Gurtner, C., W. Lubitz, and S. Rölleke. 2003. Application of broad-range 16S rRNA PCR amplification and DGGE fingerprinting for detection of tick-infecting bacteria. Journal of Microbiological Methods 52:251–260.

Schlenke, T. A., J. Morales, S. Govind, and A. G. Clark. 2007. Contrasting infection strategies in generalist and specialist wasp parasitoids of *Drosophila melanogaster*. PLoS Pathogens 3:1486–1501.

Schloss, P. D., S. L. Westcott, T. Ryabin, J. R. Hall, M. Hartmann, E. B. Hollister, R. A. Lesniewski, B. B. Oakley, D. H. Parks, C. J. Robinson, J. W. Sahl, B. Stres, G. G. Thallinger, D. J. Van Horn, and C. F. Weber. 2009. Introducing mothur: Open-source, platform-independent, community-supported software for describing and comparing microbial communities. Applied and Environmental Microbiology 75:7537–7541.

Schonbaum, C. P., J. J. Perrino, and A. P. Mahowald. 2000. Regulation of the vitellogenin receptor during *Drosophila melanogaster* oogenesis. Molecular Biology of the Cell 11:511–521.

Schoonhoven, L. M., J. J. A. van Loon, and M. Dicke (Eds.). 2005. Insect-plant Biology 2nd Edition. Oxford University Press, New York, USA.

Von der Schulenburg, J. H. G., M. Habig, J. J. Sloggett, K. M. Webberley, D. Bertrand, G. D. D. Hurst, and M. E. N. Majerus. 2001. Incidence of male-killing *Rickettsia* spp. (alpha-Proteobacteria) in the Ten-Spot Ladybird Beetle *Adalia decempunctata* L . (Coleoptera: Coccinellidae). Applied and Environmental Microbiology 67:270–277.

Schulmeister, S. 2003. Review of morphological evidence on the phylogeny of basal Hymenoptera (Insecta), with a discussion of the ordering of characters. Biological Journal of the Linnean Society:209–243.

Schwachtje, J., and I. T. Baldwin. 2008. Why does herbivore attack reconfigure primary metabolism? Plant Physiology 146:845–851.

Shange, R. S., R. O. Ankumah, A. M. Ibekwe, R. Zabawa, and S. E. Dowd. 2012. Distinct soil bacterial communities revealed under a diversely managed agroecosystem. PloS One 7:e40338.

Sharkey, M. J. 2007. Phylogeny and classification of Hymenoptera. Zootaxa 1668:521–548.

Shen, G.-M., W. Dou, J.-Z. Niu, H.-B. Jiang, W.-J. Yang, F.-X. Jia, F. Hu, L. Cong, and J.-J. Wang. 2011. Transcriptome analysis of the oriental fruit fly (*Bactrocera dorsalis*). PloS One 6:e29127.

Shibata, T. F., T. Maeda, N. Nikoh, K. Yamaguchi, K. Oshima, M. Hattori, T. Nishiyama, M. Hasebe, T. Fukatsu, Y. Kikuchi, and S. Shigenobu. 2013. Bacterial symbiont of the Bean bug *Riptortus pedestris*. Genome Announcements 1:1–2.

Silvertown, J. W. 1980. The evolutionary ecology of mast seeding in trees. Biological Journal of the Linnean Society 14:235–250.

Sintupachee, S., J. R. Milne, S. Poonchaisri, V. Baimai, and P. Kittayapong. 2006. Closely related *Wolbachia* strains within the pumpkin arthropod community and the potential for horizontal transmission via the plant. Microbial Ecology 51:294–301.

Smith, E. L. 1970. Biosystematics and morphology of Symphyta. II. Biology of gall-making nematine sawflies in the California region. Annals of the Entomological Society of America 63:36–51.

Solter, L. F., J. J. Becnel, and D. H. Oi. 2012. Microsporidian Entomophathogens. Pages 221–263 *in* F. E. Vega and H. K. Kaya, editors. Insect Pathology, 2nd edition. Academic Press, Waltham, MA.

Sourial, N., C. Wolfson, B. Zhu, J. Quail, J. Fletcher, S. Karunananthan, K. Bandeen-Roche, F. Béland, and H. Bergman. 2010. Correspondence analysis is a useful tool to uncover the relationships among categorical variables. Journal of clinical epidemiology 63:638–646.

Šrůtka, P., S. Pažoutová, and M. Kolařík. 2007. *Daldinia decipiens* and *Entonaema cinnabarina* as fungal symbionts of *Xiphydria* wood wasps. Mycological Research 111:224–231.

Van Staden, J., and J. E. Davey. 1978. Endogenous cytokinins in the laminae and galls of *Erythrina latissima* leaves. Botanical Gazette 139:36–41.

Van Staden, J., J. E. Davey, and A. R. Noel. 1977. Gall formation in *Erythrina latissima*. Zeitschrift für Pflanzenphysiologie 84:283–294.

Stone, G. N., and J. M. Cook. 1998. The structure of cynipid oak galls: patterns in the evolution of an extended phenotype. Proceedings of the Royal Society B: Biological Sciences 265:979–988.

Stone, G. N., and K. Schönrogge. 2003. The adaptive significance of insect gall morphology. Trends in Ecology and Evolution 18:512–522.

Stone, G. N., K. Schönrogge, R. J. Atkinson, D. Bellido, and J. Pujade-Villar. 2002. The population biology of oak gall wasps (Hymenoptera: Cynipidae). Annual Review of Entomology 47:633–668.

Stone, S. L., and D. J. Gifford. 1999. Structural and biochemical changes in Loblolly Pine (*Pinus taeda* L .) seeds during germination and early seedling growth . II . Storage triacylglycerols and carbohydrates. International Journal of Plant Sciences 160:663–671.

Stouthamer, R. 1997. *Wolbachia*-induced parthenogenesis. Pages 102–124 *in* S. L. O'Neill and J. H. Werren, editors. Influential Passengers. Oxford University Press, New York, USA.

Stouthamer, R., J. A. J. Breeuwer, and G. D. D. Hurst. 1999. *Wolbachia pipientis*: Microbial manipulator of arthropod reproduction. Annual Review of Microbiology 53:71–102.

Stouthamer, R., R. F. Luck, and W. D. Hamilton. 1990. Antibiotics cause parthenogenetic *Trichogramma* (Hymenoptera/Trichogrammatidae) to revert to sex. Proceedings of the National Academy of Sciences 87:2424–2427.

Strand, M. R., S. M. Meola, and S. B. Vinson. 1986. Correlating pathological symptoms in *Heliothis virescens* eggs with development of the parasitoid *Telenomus heliothidis*. Journal of Insect Physiology 32:389–402.

Tamura, K., D. Peterson, N. Peterson, G. Stecher, M. Nei, and S. Kumar. 2011. MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. Molecular Biology and Evolution 28:2731–2739.

Tarazona, S., F. García-Alcalde, J. Dopazo, A. Ferrer, and A. Conesa. 2011. Differential expression in RNA-seq: A matter of depth. Genome Research 21:2213–2223.

Tarentino, A. L., G. Quinones, C. R. Hauer, L. M. Changchien, and T. H. Plummer. 1995. Molecular cloning and sequence analysis of *Flavobacterium meningosepticum* glycosylasparaginase: a single gene encodes the alpha and beta subunits. Archives of Biochemistry and Biophysics 316:399–406.

Tariq, M. A., H. J. Kim, O. Jejelowo, and N. Pourmand. 2011. Whole-transcriptome RNAseq analysis from minute amount of total RNA. Nucleic Acids Research 39:e120.

Teixeira, L., A. Ferreira, and M. Ashburner. 2008. The bacterial symbiont *Wolbachia* induces resistance to RNA viral infections in *Drosophila melanogaster*. PLoS Biology 6:e1000002.

Tenhunen, K., M. Laan, T. Manninen, A. Palotie, L. Peltonen, and A. Jalanko. 1995. Molecular cloning, chromosomal assignment, and expression of the mouse aspartylglucosaminidase gene. Genomics 30:244–250.

Thaler, J. S. 1999. Jasmonate-inducible plant defences cause increased parasitism of herbivores. Nature 399:686–688.

Todd, C. D., and D. J. Gifford. 2002. The role of the megagametophyte in maintaining loblolly pine (*Pinus taeda* L.) seedling arginase gene expression in vitro. Planta 215:110–118.

Toju, H., and T. Fukatsu. 2011. Diversity and infection prevalence of endosymbionts in natural populations of the chestnut weevil: Relevance of local climate and host plants. Molecular Ecology 20:853–868.

Tooker, J. F., and L. M. Hanks. 2004. Stereochemistry of host plant monoterpenes as mate location cues for the gall wasp *Antistrophus rufus*. Journal of Chemical Ecology 30:473–477.

Tooker, J. F., W. A. Koenig, and L. M. Hanks. 2002. Altered host plant volatiles are proxies for sex pheromones in the gall wasp *Antistrophus rufus*. Proceedings of the National Academy of Sciences 99:15486–15491.

Tsuchida, T., R. Koga, and T. Fukatsu. 2004. Host plant specialization governed by facultative symbiont. Science 303:1989.

Turgeon, J. J., A. Roques, and P. De Groot. 1994. Insect fauna of coniferous seed cones: diversity, host plant interactions and management. Annual Review of Entomolgoy 39:179–212.

Vaneechoutte, M., P. Kämpfer, D. B. Thierry, E. Falsen, and G. Verschraegen. 2004. *Wautersia* gen. nov., a novel genus accommodating the phylogenetic lineage including *Ralstonia eutropha* and related species, and proposal of *Ralstonia* [*Pseudomonas*] *syzygii* (Roberts *et al.* 1990) comb. nov. International Journal of Systematic and Evolutionary Microbiology 54:317–327.

Vårdal, H. 2004. From Parasitoids to Gall Inducers and Inquilines - Morphological Evolution in Cynipoid Wasps. Ph.D. Thesis. Uppsala University, Sweden.

Vårdal, H. 2006. Venom gland and reservoir morphology in cynipoid wasps. Arthropod Structure and Development 35:127–136.

Vautrin, E., S. Genieys, S. Charles, and F. Vavre. 2008. Do vertically transmitted symbionts co-existing in a single host compete or cooperate? A modelling approach. Journal of Evolutionary Biology 21:145–161.

Vautrin, E., and F. Vavre. 2009. Interactions between vertically transmitted symbionts: Cooperation or conflict? Trends in Microbiology 17:95–99.

Vavre, F., F. Fleury, D. Lepetit, P. Fouillet, and M. Boulétreau. 1999a. Phylogenetic evidence for horizontal transmission of *Wolbachia* in host-parasitoid associations. Molecular Biology and Evolution 16:1711–1723.

Vavre, F., C. Girin, and M. Boulétreau. 1999b. Phylogenetic status of a fecundity-enhancing Wolbachia that does not induce thelytoky in *Trichogramma*. Insect Molecular Biology 8:67–72.

Vilhelmsen, L. 2000. The ovipositor apparatus of basal Hymenoptera (Insecta): Phylogenetic implications and functional morphology. Zoologica Scripta 29:319–345.

Vilhelmsen, L. 2001. Phylogeny and classification of the extant basal lineages of the Hymenoptera (Insecta). Zoological Journal of the Linnean Society 131:393–442.

Vilhelmsen, L., I. Mikó, and L. Krogmann. 2010. Beyond the wasp-waist: structural diversity and phylogenetic significance of the mesosoma in apocritan wasps (Insecta: Hymenoptera). Zoological Journal of the Linnean Society 159:22–194.

Vincent, B., M. Kaeslin, T. Roth, M. Heller, J. Poulain, F. Cousserans, J. Schaller, M. Poirié, B. Lanzrein, J.-M. Drezen, and S. J. M. Moreau. 2010. The venom composition of the parasitic wasp *Chelonus inanitus* resolved by combined expressed sequence tags analysis and proteomic approach. BMC Genomics 11:693.

Vinchon, S., S. J. M. Moreau, J. M. Drezen, G. Prévost, and A. Cherqui. 2010. Molecular and biochemical analysis of an aspartylglucosaminidase from the venom of the parasitoid wasp *Asobara tabida* (Hymenoptera: Braconidae). Insect Biochemistry and Molecular Biology 40:38–48.

Vinson, B. S., and G. F. Iwantsch. 1980. Host regulation by insect parasitoids. The Quarterly Review of Biology 55:143–165.

Vogt, R. G., F. E. Callahan, M. E. Rogers, and J. C. Dickens. 1999. Odorant binding protein diversity and distribution among the insect orders, as indicated by LAP, an OBP-related protein of the true bug *Lygus lineolaris* (Hemiptera, Heteroptera). Chemical Senses 24:481–495.

von Aderkas, P., G. Rouault, R. Wagner, R. Rohr, and A. Roques. 2005a. Seed parasitism redirects ovule development in Douglas fir. Proceedings of the Royal Society B: Biological Sciences 272:1491–1496.

von Aderkas, P., G. Rouault, R. Wagner, S. Chiwocha, and A. Roques. 2005b. Multinucleate storage cells in Douglas-fir (*Pseudotsuga menziesii* (Mirbel) Franco) and the effect of seed parasitism by the chalcid *Megastigmus spermotrophus* Wachtl. Heredity 94:616–622.

Wang, H.-X., Z.-L. Geng, Y. Zeng, and Y.-M. Shen. 2008. Enriching plant microbiota for a metagenomic library construction. Environmental Microbiology 10:2684–2691.

Wang, Q., G. M. Garrity, J. M. Tiedje, and J. R. Cole. 2007. Naïve bayesian classifier for rapid assignment of rRNA sequences into the new bacterial taxonomy. Applied and Environmental Microbiology 73:5261–5267.

Wang, Z., M. Gerstein, and M. Snyder. 2009. RNA-Seq: a revolutionary tool for transcriptomics. Nature Reviews Genetics 10:57–63.

Warnecke, F., P. Luginbühl, N. Ivanova, M. Ghassemian, T. H. Richardson, J. T. Stege, M. Cayouette, A. C. McHardy, G. Djordjevic, N. Aboushadi, R. Sorek, S. G. Tringe, M. Podar, H. G. Martin, V. Kunin, D. Dalevi, J. Madejska, E. Kirton, D. Platt, E. Szeto, A. Salamov, K. Barry, N. Mikhailova, N. C. Kyrpides, E. G. Matson, E. a Ottesen, X. Zhang, M. Hernández, C. Murillo, L. G. Acosta, I. Rigoutsos, G. Tamayo, B. D. Green, C. Chang, E. M. Rubin, E. J. Mathur, D. E. Robertson, P. Hugenholtz, and J. R. Leadbetter. 2007. Metagenomic and functional analysis of hindgut microbiota of a wood-feeding higher termite. Nature 450:560–565.

Weber, N. A. 1966. Fungus-growing ants. Science 153:587–604.

Weeks, A. R., M. Turelli, W. R. Harcombe, K. T. Reynolds, and A. A. Hoffmann. 2007. From parasite to mutualist: Rapid evolution of *Wolbachia* in natural populations of *Drosophila*. PLoS Biology 5:e114.

Weeks, A. R., R. Velten, and R. Stouthamer. 2003. Incidence of a new sex-ratio-distorting endosymbiotic bacterium among arthropods. Proceedings of the Royal Society of London B: Biological Sciences 270:1857–1865.

Wei, J., L. Wang, J. Zhu, S. Zhang, O. I. Nandi, and L. Kang. 2007. Plants attract parasitic wasps to defend themselves against insect pests by releasing hexenol. PloS One 2:e852.

Weiblen, G. D. 2002. How to be a fig wasp. Annual Review of Entomology 47:299–330.

Weiblen, G. D., and G. L. Bush. 2002. Speciation in fig pollinators and parasites. Molecular Ecology 11:1573–1578.

Weinert, L. A., J. H. Werren, A. Aebi, G. N. Stone, and F. M. Jiggins. 2009. Evolution and diversity of *Rickettsia* bacteria. BMC Biology 7:6.

Weiser, J., and J. Novotnỳ. 1987. Field application of *Nosema lymantriae* against the gypsy moth, *Lymantria dispar* L. Journal of Applied Entomology 104:58–62.

Wells, B. W. 1921. Evolution of zoocecidia. Botanical Gazette 71:358–377.

Wernegreen, J. J. 2002. Genome evolution in bacterial endosymbionts of insects. Nature Reviews: Genetics 3:850–861.

Werner, J. J., O. Koren, P. Hugenholtz, T. Z. DeSantis, W. A. Walters, J. G. Caporaso, L. T. Angenent, R. Knight, and R. E. Ley. 2012. Impact of training sets on classification of high-throughput bacterial 16s rRNA gene surveys. The ISME Journal 6:94–103.

Werren, J. H., L. Baldo, and M. E. Clark. 2008. *Wolbachia*: master manipulators of invertebrate biology. Nature reviews. Microbiology 6:741–751.

Werren, J. H., S. Richards, C. A. Desjardins, O. Niehuis, J. Gadau, and J. K. Colbourne. 2010. Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. Science 327:343–348.

Wharton, R. A., and P. E. Hanson. 2005. Biology and Evolution of Braconid Gall Wasps (Hymenoptera). Pages 495–505 *in* A. Raman, C. W. Schaefer, and T. M. Withers, editors. Biology, Ecology, and Evolution of Gall-inducing Arthropods, Volume 2. Science Publishers, Inc., Enfield, USA.

Wharton, R., L. Vilhelmsen, and G. A. P. Gibson. 2004. Characterizing basal apocritans (Hymenoptera: Apocrita). Proceedings of the Russian Entomological Society 75:17–23.

Whitehead, L. F., T. L. Wilkinson, and A. E. Douglas. 1992. Nitrogen recycling in the pea aphid (*Acyrtosiphon pisum*) symbiosis. Proceedings of the Royal Society of London B: Biological Sciences 250:115–117.

Whitfield, J. B. 2003. Phylogenetic insights into the evolution of parasitism in Hymenoptera. Advances in Parasitology 54:69–100.

Wiebes, J. T. 1979. Co-evolution of figs and their insect pollinators. Annual Review of Ecology and Systematics 10:1–12.

Wiegmann, B. M., C. Mitter, and B. Farrell. 1993. Diversification of carnivorous parasitic insects: Extraordinary radiation or specialized dead end? The American Naturalist 142:737–754.

Wilson, G., and J. Kaupp. 1976. Application of *Nosema fumiferanae* and *Pleistophora schubergi* (Microsporidia) against the spruce budworm in Ontario, 1976. Canadian Forest Service, Sault Ste. Marie, Ontario. Information Report IP-X-15:1–14.

Wu, D., P. Hugenholtz, K. Mavromatis, R. Pukall, E. Dalin, N. N. Ivanova, V. Kunin, L. Goodwin, M. Wu, B. J. Tindall, S. D. Hooper, A. Pati, A. Lykidis, S. Spring, I. J. Anderson, P. D'haeseleer, A. Zemla, M. Singer, A. Lapidus, M. Nolan, A. Copeland, C. Han, F. Chen, J.-F. Cheng, S. Lucas, C. Kerfeld, E. Lang, S. Gronow, P. Chain, D. Bruce, E. M. Rubin, N. C. Kyrpides, H.-P. Klenk, and J. A. Eisen. 2009. A phylogeny-driven genomic encyclopaedia of Bacteria and Archaea. Nature 462:1056–1060.

Xue, J., Y.-Y. Bao, B.-L. Li, Y.-B. Cheng, Z.-Y. Peng, H. Liu, H.-J. Xu, Z.-R. Zhu, Y.-G. Lou, J.-A. Cheng, and C.-X. Zhang. 2010. Transcriptome analysis of the brown planthopper *Nilaparvata lugens*. PloS One 5:e14233.

Yamaguchi, H., H. Tanaka, M. Hasegawa, M. Tokuda, T. Asami, and Y. Suzuki. 2012. Phytohormones and willow gall induction by a gall-inducing sawfly. New Phytologist 196:586–595.

Yang, Y., and S. A. Smith. 2013. Optimizing *de novo* assembly of short-read RNA-seq data for phylogenomics. BMC Genomics 14:328.

Zchori-Fein, E., C. Borad, and A. R. Harari. 2006. Oogenesis in the date stone beetle, *Coccotrypes dactyliperda*, depends on symbiotic bacteria. Physiological Entomology 31:164–169.

Zerova, M. D., and V. N. Fursov. 1991. The Palaearctic species of *Eurytoma* (Hymenoptera: Eurytomidae) developing in stone fruits (Rosaceae: Prunoideae). Bulletin of Entomological Research 81:209–219.

Zhu, J.-Y., Q. Fang, L. Wang, C. Hu, and G.-Y. Ye. 2010. Proteomic analysis of the venom from the endoparasitoid wasp *Pteromalus puparum* (Hymenoptera: Pteromalidae). Archives of Insect Biochemistry and Physiology 75:28–44.

Zug, R., and P. Hammerstein. 2012. Still a host of hosts for *Wolbachia*: analysis of recent data suggests that 40% of terrestrial arthropod species are infected. PloS One 7:e38544.

Appendix 1: Supplementary data

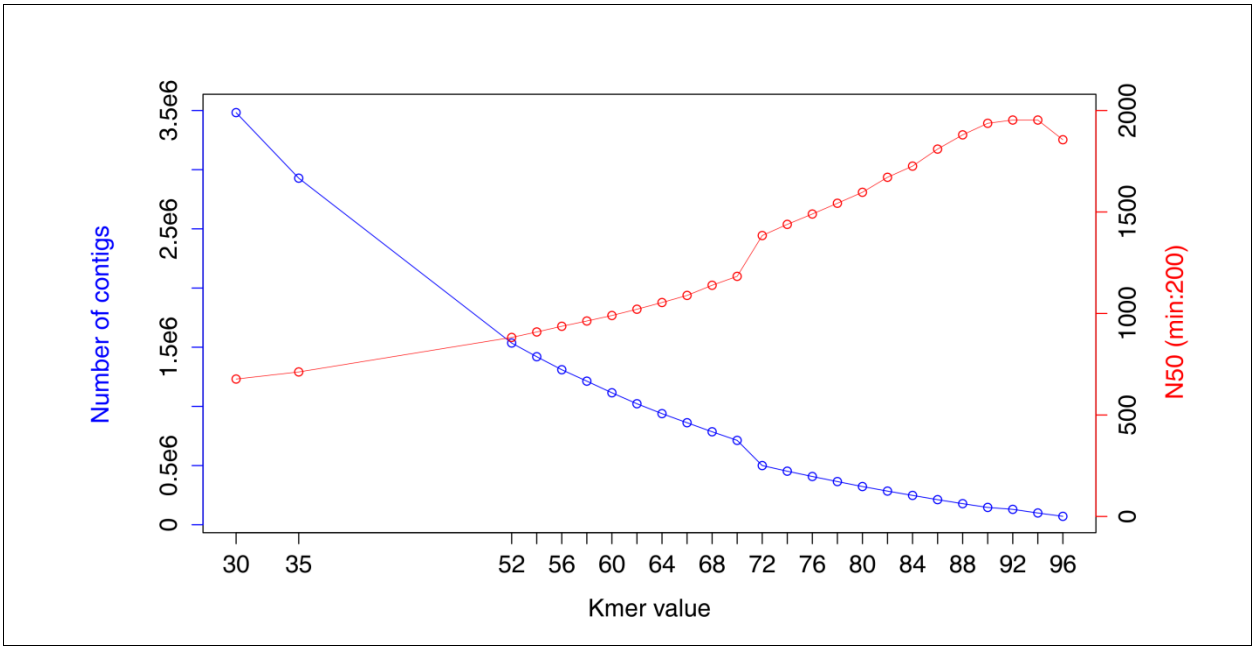


Figure S1: *M. spermotrophus* ABySS multi-k *de novo* transcriptome assembly results: Number of contigs (blue) and N50 value (red).

Table S1: *Nasonia vitripennis* venom query

Protein Name	Accession number
Proteases and peptidases	
metalloprotease-like precursor	gi|238859601|ref|NP_001155006.1|
serine protease precursor	gi|238859621|ref|NP_001155015.1|
serine protease precursor	gi|283046805|ref|NP_001164348.1|
serine protease 16 precursor	gi|239050219|ref|NP_001155077.1|
serine protease homolog 21 precursor	gi|239049675|ref|NP_001155060.1|
serine protease 22 precursor	gi|239048216|ref|NP_001155043.1|
serine protease homolog 29 precursor	gi|238859623|ref|NP_001155016.1|
serine protease 33 precursor	gi|238859625|ref|NP_001155017.1|
serine protease homolog 42 isoform 2 precursor	gi|239050294|ref|NP_001155079.1|
serine protease homolog 42 isoform 1 precursor	gi|239050264|ref|NP_001155078.1|
serine protease 50 precursor	gi|239050201|ref|NP_001155076.1|
serine protease 96 precursor	gi|238859618|ref|NP_001155014.1|
serine protease 97 precursor	gi|239048185|ref|NP_001155042.1|
Protease inhibitors	
cysteine-rich/KU venom protein precursor	gi|238859591|ref|NP_001154998.1|
cysteine-rich/pacifastin venom protein 1 precursor	gi|238859589|ref|NP_001154997.1|
cysteine-rich/pacifastin venom protein 2 precursor	gi|238859587|ref|NP_001154996.1|
cysteine-rich/TIL venom protein 2 precursor	gi|238859641|ref|NP_001155022.1|
Kazal type serine protease inhibitor-like venom protein 1 precursor	gi|238859585|ref|NP_001154995.1|
Kazal type serine protease inhibitor-like venom protein 2 precursor	gi|283046811|ref|NP_001164350.1|
small serine proteinase inhibitor-like venom protein precursor	gi|239050434|ref|NP_001155083.1|
Carbohydrate metabolism	
chitinase 5 precursor	gi|239050479|ref|NP_001155084.1|
glucose dehydrogenase-like venom protein	gi|239050555|ref|NP_001155086.1|
glucose dehydrogenase-like venom protein	gi|239050502|ref|NP_001155085.1|
DNA metabolism	
endonuclease-like venom protein precursor	gi|239050560|ref|NP_001155087.1|
inosine-uridine preferring nucleoside hydrolase-like precursor	gi|239787140|ref|NP_001155172.1|
Glutathione metabolism	
gamma-glutamyl cyclotransferase-like venom protein isoform 1 precursor	gi|239735530|ref|NP_001155144.1|
gamma-glutamyl cyclotransferase-like venom protein isoform 2	gi|239735532|ref|NP_001155145.1|
Esterases	
venom acid phosphatase-like precursor	gi|239735537|ref|NP_001155147.1|
venom acid phosphatase-like precursor	gi|239735535|ref|NP_001155146.1|
multiple inositol polyphosphate phosphatase-like venom protein precursor	gi|239787860|ref|NP_001155183.1|
carboxylesterase clade B, member 2 precursor	gi|239735550|re|NP_001155148.14|
lipase-like venom protein precursor	gi|238908542|ref|NP_001155039.1|
lipase A-like precursor	gi|238859582|ref|NP_001154991.1|

Table S1 (Continued)

Protein Name	Accession number
Recognition/binding proteins	
gram-negative bacteria binding protein 1-2 precursor	gi|239735554|ref|NP_001155149.1|
chitin binding protein-like venom protein precursor	gi|283046793|ref|NP_001164343.1|
GOBP-like venom protein precursor	gi|239735556|ref|NP_001155150.1|
low-density lipoprotein receptor-like venom protein precursor	gi|239047943|ref|NP_001155040.1|
Immunity related proteins	
calreticulin precursor	gi|239735560|ref|NP_001155151.1|
C1q-like venom protein precursor	gi|239735565|ref|NP_001155152.1|
Others	
aminotransferase-like venom protein 1 precursor	gi|239735567|ref|NP_001155153.1|
aminotransferase-like venom protein 2 precursor	gi|239735577|ref|NP_001155156.1|
antigen 5-like protein 1 precursor	gi|239735572|ref|NP_001155154.1|
aspartylglucosaminidase precursor	gi|239787108|ref|NP_001155157.1|
laccase-like precursor	gi|239787112|ref|NP_001155159.1|
venom laccase precursor	gi|239787110|ref|NP_001155158.1|
Unknowns	
venom protein D precursor	gi|239787138|ref|NP_001155171.1|
venom protein F precursor	gi|239787114|ref|NP_001155160.1|
venom protein G precursor	gi|283046795|ref|NP_001164344.1|
venom protein H precursor	gi|238908530|ref|NP_001155027.1|
venom protein I precursor	gi|283046801|ref|NP_001164346.1|
venom protein J precursor	gi|283046803|ref|NP_001164347.1|
venom protein K precursor	gi|238908532|ref|NP_001155028.1|
venom protein L precursor	gi|238908534|ref|NP_001155029.1|
venom protein M precursor	gi|238908536|ref|NP_001155030.1|
venom protein N precursor	gi|283046809|ref|NP_001164349.1|
venom protein O precursor	gi|238908540|ref|NP_001155031.1|
venom protein Q precursor	gi|239787116|ref|NP_001155161.1|
venom protein R precursor	gi|239787122|ref|NP_001155164.1|
venom protein T precursor	gi|239787126|ref|NP_001155166.1|
venom protein U precursor	gi|239787134|ref|NP_001155170.1|
venom protein V precursor	gi|239048037|ref|NP_001155041.1|
venom protein X precursor	gi|239787128|ref|NP_001155167.1|
venom protein Y precursor	gi|239787130|ref|NP_001155168.1|
venom protein Z precursor	gi|239787132|ref|NP_001155169.1|

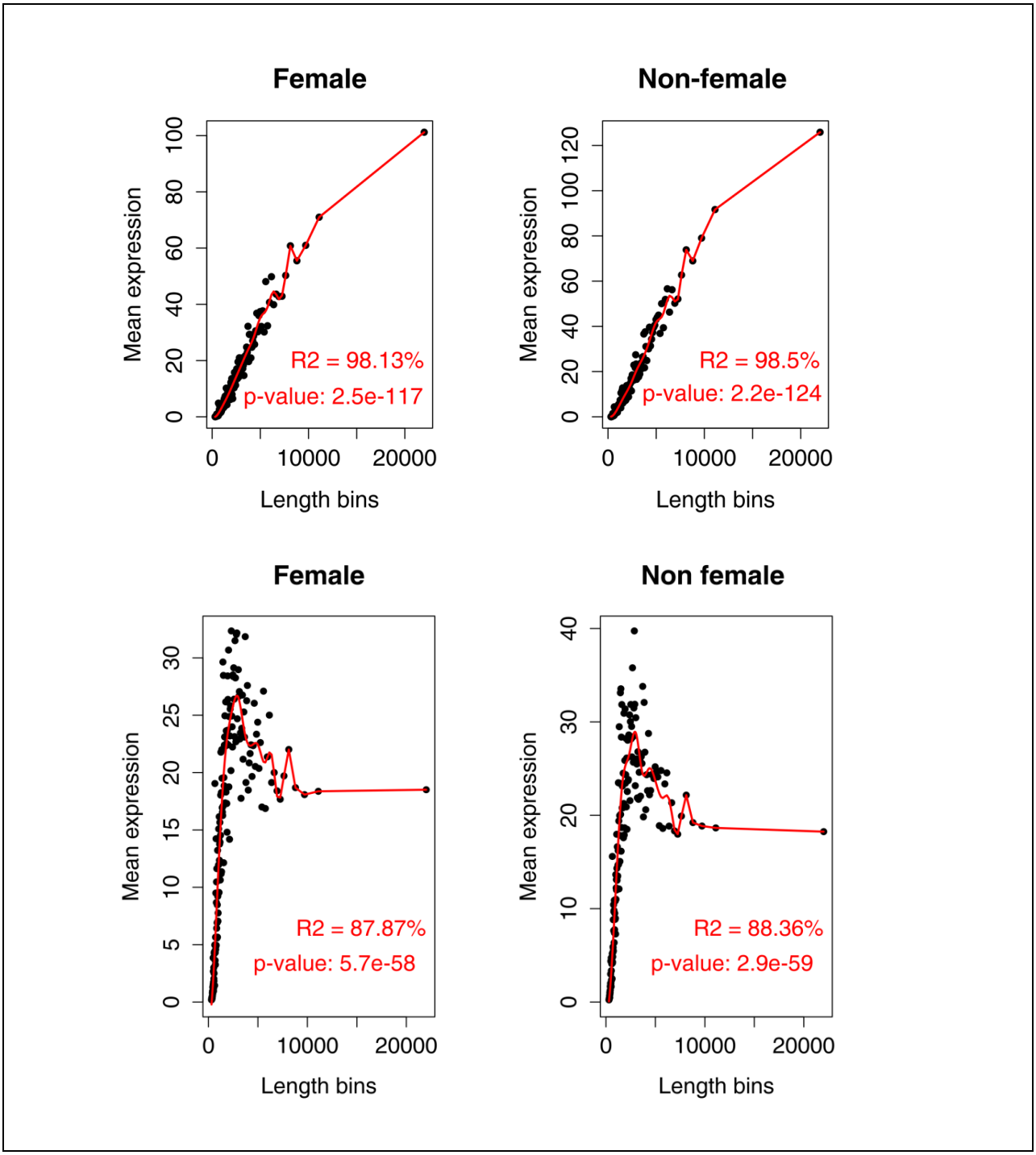


Figure S2: Mean expression/length bias exploratory plots. Top: Mean expression versus contig length, fitted with a cubic spline regression model as implemented in NOISeq for *M. spermotrophus* transcriptome data. Bottom: Mean expression versus contig length, fitted with a cubic spline regression model as implemented in NOISeq for *M. spermotrophus* transcriptome CQN normalized data.

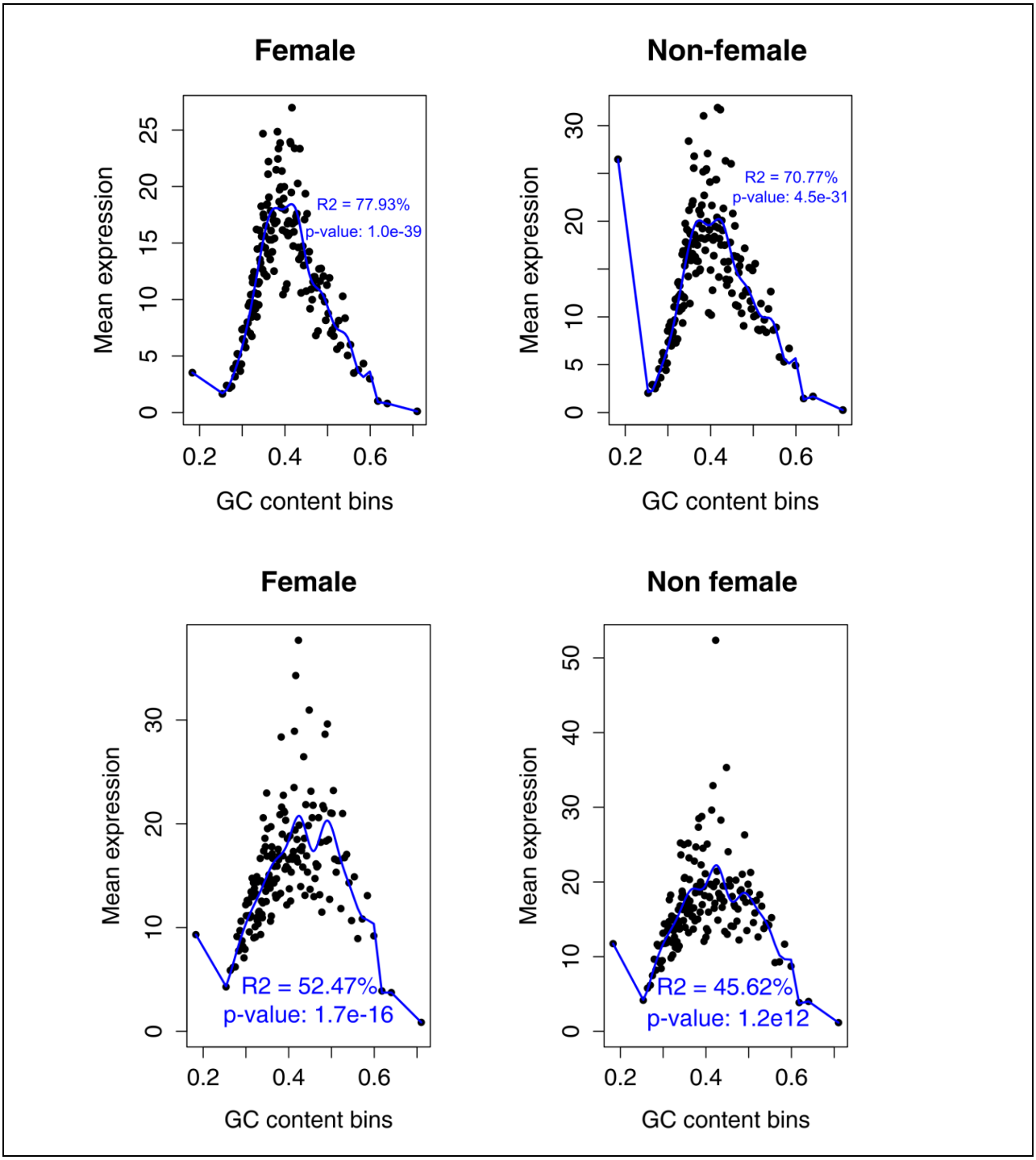


Figure S3: Mean expression/GC content bias exploratory plots. Top: Mean expression versus GC content, fitted with a cubic spline regression model as implemented in NOISeq for *M. spermotrophus* transcriptome data. Bottom: Mean expression versus GC content, fitted with a cubic spline regression model as implemented in NOISeq for *M. spermotrophus* transcriptome CQN normalized data.

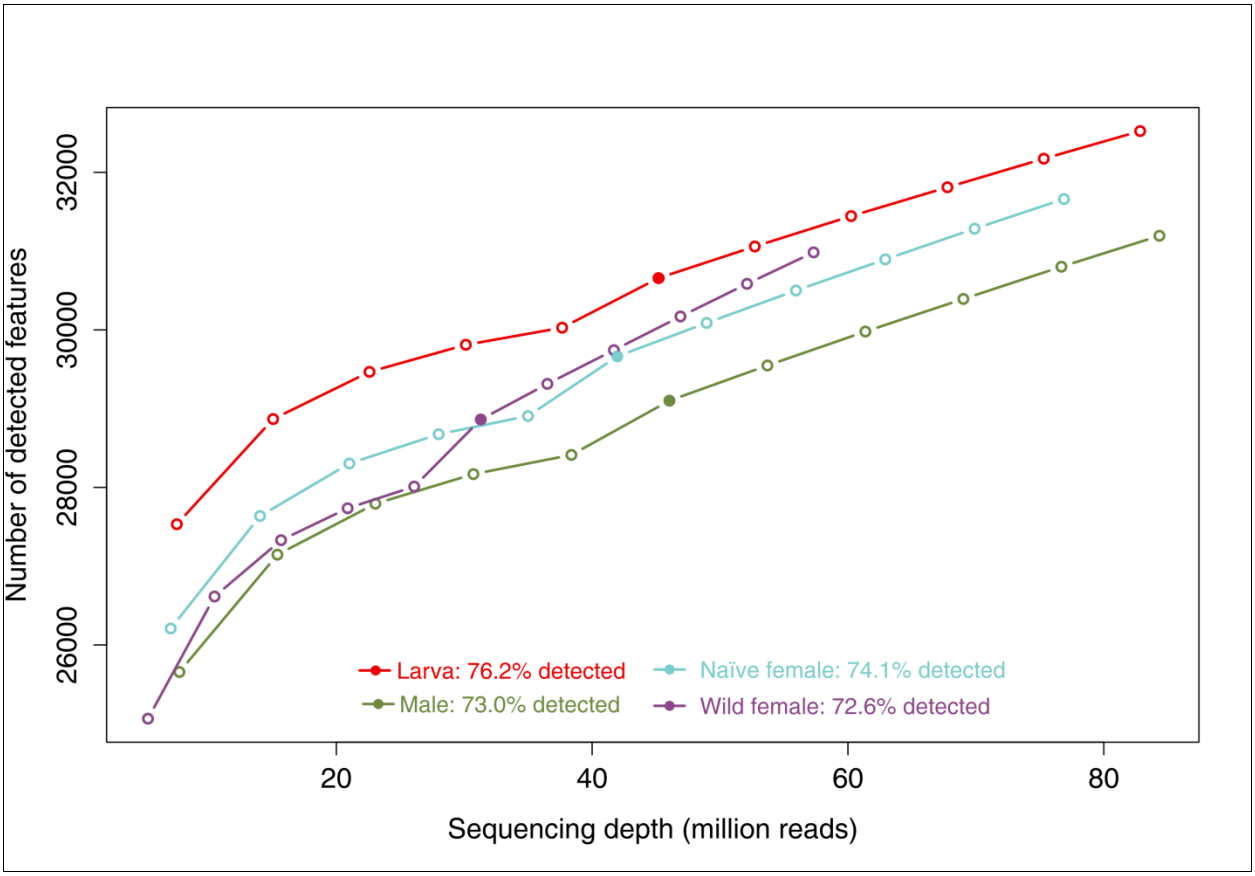


Figure S4: Saturation plot for all annotated contigs from the *M. spermotrophus* transcriptome, including zero counts (greater than or equal to 300 bp).

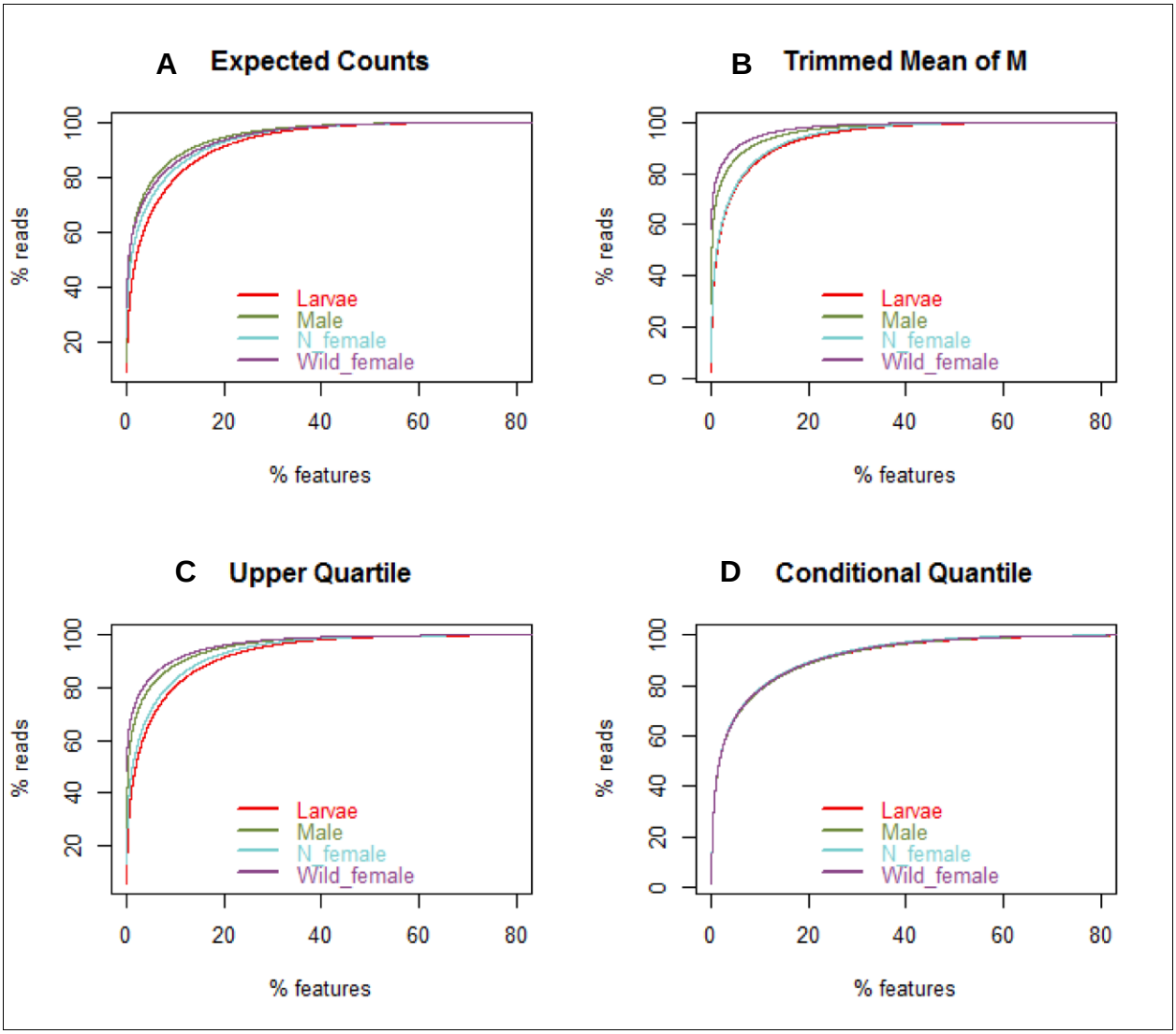


Figure S5: Comparison of count distribution plots of *M. spermotrophus* transcriptome data; A: Expected counts (no normalization), B: Trimmed mean of M normalization, C: Upper quartile normalization and D: Conditional quantile normalization.

Table S2: Highly expressed transcripts from females determined by NOISeq-sim

Transcript ID	Length	% GC	Mean Female Exprs	Mean Non-female Exprs	M*	D**	Prob	E-value	Annotation	Species	Taxa
contig189140	783	0.44	1915.69	0.28	12.75	1915.41	1.00	3.00E-06	PREDICTED: probable salivary secreted peptide-like	*Megachile rotundata*	Insect
contig191714	651	0.46	1826.29	0.50	11.82	1825.78	1.00	2.00E-12	PREDICTED: hypothetical protein LOC100679258	*Nasonia vitripennis*	Insect
contig20108	3216	0.35	841.13	0.29	11.51	840.84	0.99	5.00E-46	hypothetical protein OXYTRI_14248	*Oxytricha trifallax*	Protista
contig6704	1983	0.52	819.34	0.33	11.28	819.01	0.99	2.00E-23	predicted protein	*Trichoderma reesei QM6a*	Fungi
contig6007	776	0.59	4655.62	1.97	11.21	4653.65	1.00	3.00E-22	hypothetical protein I79_019987	*Cricetulus griseus*	Other Animal
contig34540	1959	0.49	1152.98	0.54	11.07	1152.44	0.99	8.00E-09	JMH-2010 28S ribosomal RNA gene, partial sequence	*Paramblynotus sp.*	Other Animal
contig6002	524	0.58	1040.12	0.50	11.03	1039.62	0.99	2.00E-19	hypothetical protein I79_019987	*Cricetulus griseus*	Other Animal
contig28832	2841	0.40	1913.44	1.04	10.85	1912.40	1.00	6.00E-30	PREDICTED: hypothetical protein LOC100679858	*Nasonia vitripennis*	Insect
contig28833	3891	0.43	778.02	0.43	10.84	777.60	0.99	4.00E-40	PREDICTED: hypothetical protein LOC100679858	*Nasonia vitripennis*	Insect
contig5522	692	0.52	534.52	0.31	10.75	534.21	0.99	6.00E-10	predicted protein	*Nematostella vectensis*	Other Animal
contig192522	2510	0.43	172.63	0.10	10.75	172.53	0.96	0	PREDICTED: sialin-like	*Nasonia vitripennis*	Insect
contig6736	3138	0.50	5835.16	3.41	10.74	5831.75	1.00	4.00E-10	hypothetical protein I79_019987	*Cricetulus griseus*	Other Animal
contig183854	1038	0.43	1014.35	0.61	10.70	1013.73	0.99	8.00E-107	aspartylglucosaminidase precursor	*Nasonia vitripennis*	Insect
contig6309	814	0.59	3552.73	2.36	10.56	3550.37	1.00	3.00E-29	hypothetical protein NEMVEDRAFT_v1g223041	*Nematostella vectensis*	Other Animal
contig45	972	0.40	9204.19	6.18	10.54	9198.00	1.00	3.00E-22	PREDICTED: lipase 3-like	*Nasonia vitripennis*	Insect
contig2744	410	0.60	799.73	0.55	10.52	799.19	0.99	5.00E-15	hypothetical protein I79_019987	*Cricetulus griseus*	Other Animal
contig34788	2149	0.49	3852.68	2.72	10.47	3849.96	1.00	4.00E-50	PREDICTED: hypothetical protein LOC100121611	*Nasonia vitripennis*	Insect
contig192563	1128	0.47	1414.19	1.09	10.34	1413.09	0.99	2.00E-10	predicted protein	*Nematostella vectensis*	Other Animal
contig26255	2108	0.36	225.58	0.19	10.22	225.39	0.97	2.00E-171	PREDICTED: glucose dehydrogenase [acceptor]-like	*Nasonia vitripennis*	Insect
contig6671	608	0.41	1552.84	1.31	10.21	1551.53	0.99	3.00E-34	PREDICTED: hypothetical protein LOC100118223	*Nasonia vitripennis*	Insect
contig41705	1462	0.35	417.15	0.35	10.21	416.79	0.98	2.00E-13	PREDICTED: TD and POZ domain-containing protein 2-like	*Nasonia vitripennis*	Insect

Transcript ID	Length	% GC	Mean Female Exprs	Mean Non-female Exprs	M*	D**	Prob	E-value	Annotation	Species	Taxa
contig6505	339	0.61	358.53	0.31	10.19	358.22	0.98	4.00E-11	predicted protein	*Nematostella vectensis*	Other Animal
contig6742	985	0.54	4533.80	3.99	10.15	4529.81	1.00	2.00E-22	hypothetical protein I79_019987	*Cricetulus griseus*	Other Animal
contig35821	1217	0.36	349.82	0.35	9.95	349.47	0.98	2.00E-08	hypothetical protein SERLA73DRAFT_67532	*Serpula lacrymans var. lacrymans S7.3*	Fungi
contig3769	476	0.39	1024.39	1.05	9.94	1023.34	0.99	5.00E-39	gram-negative bacteria binding protein 1-2 precursor	*Nasonia vitripennis*	Insect
contig28352	1984	0.40	476.76	0.53	9.83	476.23	0.98	2.00E-21	hypothetical protein G5I_12475	*Acromyrmex echinatior*	Insect
contig192089	2164	0.50	124.67	0.14	9.78	124.53	0.95	8.00E-112	PREDICTED: hypothetical protein LOC100119851	*Nasonia vitripennis*	Insect
contig162223	2180	0.41	98.53	0.11	9.77	98.42	0.94	0	PREDICTED: glucose dehydrogenase [acceptor]-like	*Nasonia vitripennis*	Insect
contig5826	1107	0.52	342.73	0.41	9.71	342.32	0.98	2.00E-16	predicted protein	*Trichoderma reesei QM6a*	Fungi
contig192499	3398	0.45	53.44	0.07	9.67	53.37	0.90	0	PREDICTED: hypothetical protein LOC100117378	*Nasonia vitripennis*	Insect
contig2829	1574	0.52	341.75	0.42	9.66	341.33	0.98	4.00E-32	predicted protein	*Trichoderma reesei QM6a*	Fungi
contig6723	844	0.39	484.59	0.62	9.60	483.97	0.98	3.00E-63	PREDICTED: putative fatty acyl-CoA reductase CG5065-like	*Nasonia vitripennis*	Insect
contig6680	354	0.58	281.98	0.40	9.48	281.58	0.98	2.00E-08	hypothetical protein I79_019987	*Cricetulus griseus*	Other Animal
contig40898	1875	0.36	270.73	0.39	9.45	270.34	0.97	1.00E-46	serine protease homolog 90 precursor	*Nasonia vitripennis*	Insect
contig6691	3494	0.35	1127.23	1.74	9.34	1125.48	0.99	5.00E-111	PREDICTED: hypothetical protein LOC100680303	*Nasonia vitripennis*	Insect
contig40609	1699	0.42	175.16	0.27	9.32	174.89	0.96	2.00E-27	PREDICTED: hypothetical protein LOC100117558	*Nasonia vitripennis*	Insect
contig5885	928	0.41	3587.40	5.71	9.30	3581.69	1.00	9.00E-143	PREDICTED: maltase 1-like	*Nasonia vitripennis*	Insect
contig29406	2276	0.31	187.07	0.30	9.29	186.77	0.97	2.00E-62	PREDICTED: carbonic anhydrase 2-like	*Nasonia vitripennis*	Insect
contig10166	1177	0.40	239.68	0.39	9.28	239.29	0.97	4.00E-13	PREDICTED: hypothetical protein LOC100678529	*Nasonia vitripennis*	Insect
contig23851	5443	0.43	371.76	0.65	9.15	371.10	0.98	0	PREDICTED: serine protease nudel-like	*Nasonia vitripennis*	Insect
contig159849	2141	0.37	85.48	0.16	9.07	85.32	0.93	0	PREDICTED: glucose dehydrogenase [acceptor]-like	*Nasonia vitripennis*	Insect
contig178078	2726	0.40	149.82	0.28	9.06	149.54	0.96	0	PREDICTED: P protein-like	*Nasonia vitripennis*	Insect
contig14222	3014	0.38	491.92	0.98	8.97	490.94	0.98	5.00E-36	PREDICTED: hypothetical protein LOC100680141	*Nasonia vitripennis*	Insect

Transcript ID	Length	% GC	Mean Female Exprs	Mean Non-female Exprs	M*	D**	Prob	E-value	Annotation	Species	Taxa
contig180772	1265	0.50	122.71	0.25	8.96	122.46	0.95	1.00E-139	PREDICTED: hypothetical protein LOC100120269	*Nasonia vitripennis*	Insect
contig6134	1016	0.42	15862.38	31.95	8.96	15830.43	1.00	7.00E-17	voucher 1663_03 18S ribosomal RNA gene, partial sequence	*Eupristina verticillata*	Insect
contig2882	1075	0.40	874.58	1.77	8.95	872.81	0.99	4.00E-107	PREDICTED: putative beta-carotene-binding protein-like	*Nasonia vitripennis*	Insect
contig687	584	0.51	3304.37	6.76	8.93	3297.61	1.00	2.00E-11	CHK1 checkpoint-like protein	*Helicoverpa armigera*	Insect
contig24739	1366	0.50	204.63	0.45	8.82	204.18	0.97	4.00E-11	PREDICTED: hypothetical protein LOC100121611	*Nasonia vitripennis*	Insect
contig28242	2556	0.38	10173.69	22.88	8.80	10150.81	1.00	7.00E-37	PREDICTED: B1 protein-like	*Nasonia vitripennis*	Insect
contig6629	3907	0.30	2058.04	4.82	8.74	2053.21	0.99	8.00E-28	PREDICTED: hypothetical protein LOC100121903	*Nasonia vitripennis*	Insect
contig185398	1962	0.49	66.72	0.16	8.69	66.56	0.92	0	PREDICTED: discoidin domain-containing receptor 2-like	*Nasonia vitripennis*	Insect
contig1257	3616	0.35	893.13	2.20	8.67	890.93	0.99	0	PREDICTED: putative fatty acyl-CoA reductase CG5065-like	*Nasonia vitripennis*	Insect
contig189367	1338	0.44	165.06	0.44	8.54	164.62	0.96	2.00E-54	PREDICTED: hypothetical protein LOC100679858	*Nasonia vitripennis*	Insect
contig169807	2688	0.33	170.82	0.48	8.48	170.34	0.96	8.00E-50	hypothetical protein SINV_11875	*Solenopsis invicta*	Insect
contig35295	692	0.33	124.70	0.36	8.44	124.34	0.95	3.00E-16	putative odorant binding protein 70	*Nasonia vitripennis*	Insect
contig5949	2482	0.32	6264.77	19.04	8.36	6245.73	1.00	1.00E-11	PREDICTED: proton myo-inositol cotransporter-like	*Nasonia vitripennis*	Insect
contig193081	2745	0.47	98.84	0.30	8.36	98.54	0.94	0	PREDICTED: protein ECT2-like isoform 3	*Nasonia vitripennis*	Insect
contig1382	1775	0.31	1859.93	5.97	8.28	1853.96	0.99	4.00E-11	PREDICTED: hypothetical protein LOC100679258	*Nasonia vitripennis*	Insect
contig174257	382	0.46	104.18	0.34	8.28	103.84	0.94	5.00E-13	serine-type enodpeptidase, putative	*Aedes aegypti*	Insect
contig187985	415	0.43	80.14	0.27	8.23	79.87	0.93	1.00E-36	gram-negative bacteria binding protein 1-2 precursor	*Nasonia vitripennis*	Insect
contig5346	639	0.38	889.40	2.98	8.22	886.42	0.99	5.00E-32	PREDICTED: hypothetical protein LOC100118223	*Nasonia vitripennis*	Insect
contig192319	2809	0.42	261.97	0.88	8.21	261.09	0.97	0	PREDICTED: sialin-like	*Nasonia vitripennis*	Insect
contig4430	1733	0.39	861.52	2.96	8.19	858.57	0.99	3.00E-08	PREDICTED: probable salivary secreted peptide-like	*Megachile rotundata*	Insect
contig5912	764	0.32	204.12	0.70	8.18	203.41	0.97	4.00E-33	gram-negative bacteria binding protein 1-2 precursor	*Nasonia vitripennis*	Insect
contig3194	889	0.38	943.96	3.28	8.17	940.69	0.99	3.00E-39	PREDICTED: beta-1,3-glucan-binding protein	*Nasonia vitripennis*	Insect
contig174544	1169	0.46	141.63	0.51	8.10	141.11	0.95	1.00E-129	PREDICTED: sialin-like	*Nasonia vitripennis*	Insect

Transcript ID	Length	% GC	Mean Female Exprs	Mean Non-female Exprs	M*	D**	Prob	E-value	Annotation	Species	Taxa
contig25986	2540	0.30	59.49	0.22	8.09	59.27	0.91	4.00E-82	PREDICTED: hypothetical protein LOC100123079	*Nasonia vitripennis*	Insect
contig3492	884	0.31	130.77	0.48	8.09	130.29	0.95	1.00E-07	PREDICTED: uncharacterized protein LOC100900476	*Metaseiulus occidentalis*	Other Animal
contig6226	1054	0.39	10244.39	37.86	8.08	10206.53	0.99	1.00E-29	PREDICTED: hypothetical protein LOC100118223	*Nasonia vitripennis*	Insect
contig8507	1628	0.35	413.94	1.53	8.08	412.41	0.98	9.00E-118	serine protease 48 precursor	*Nasonia vitripennis*	Insect
contig97837	1671	0.38	54.70	0.21	8.04	54.50	0.90	2.00E-22	hypothetical protein DAPPUDRAFT_14900	*Daphnia pulex*	Other Animal
contig191544	664	0.37	3429.20	13.45	7.99	3415.75	0.99	1.00E-37	PREDICTED: beta-1,3-glucan-binding protein	*Nasonia vitripennis*	Insect
contig187976	1732	0.44	68.36	0.28	7.93	68.08	0.92	1.00E-13	hypothetical protein IscW_ISCW005027	*Ixodes scapularis*	Other Animal
contig5996	390	0.57	75.38	0.31	7.92	75.06	0.92	3.00E-13	hypothetical protein I79_019987	*Cricetulus griseus*	Other Animal
contig183961	2614	0.40	96.21	0.40	7.90	95.81	0.94	8.00E-46	PREDICTED: hypothetical protein LOC100678599	*Nasonia vitripennis*	Insect
contig18606	2294	0.39	260.96	1.09	7.90	259.87	0.97	1.00E-166	PREDICTED: transferrin-like	*Nasonia vitripennis*	Insect
contig19033	1505	0.34	143.69	0.61	7.89	143.08	0.95	2.00E-25	PREDICTED: akirin-like	*Megachile rotundata*	Insect
contig30139	917	0.31	86.72	0.37	7.89	86.36	0.93	1.00E-36	serine protease 48 precursor	*Nasonia vitripennis*	Insect
contig29261	2422	0.45	887.49	3.80	7.87	883.69	0.99	7.00E-36	PREDICTED: hypothetical protein LOC100121575	*Nasonia vitripennis*	Insect
contig24367	1994	0.35	75.76	0.33	7.82	75.43	0.92	4.00E-124	cytochrome P450 4AB12 precursor	*Nasonia vitripennis*	Insect
contig7351	1113	0.47	155.52	0.69	7.82	154.83	0.96	7.00E-77	PREDICTED: eukaryotic translation initiation factor 4E transporter-like	*Bombus impatiens*	Insect
contig25633	1149	0.36	109.69	0.48	7.82	109.20	0.94	1.00E-07	PREDICTED: hypothetical protein LOC100678455	*Nasonia vitripennis*	Insect
contig163235	2068	0.37	88.29	0.39	7.81	87.89	0.93	8.00E-49	ankyrin repeat protein	*Trichomonas vaginalis G3*	Protista
contig6191	748	0.35	378.15	1.75	7.75	376.40	0.98	9.00E-39	PREDICTED: beta-1,3-glucan-binding protein	*Nasonia vitripennis*	Insect
contig24891	7031	0.42	225.05	1.05	7.74	224.00	0.97	0	PREDICTED: hypothetical protein LOC100678525	*Nasonia vitripennis*	Insect
contig40608	1981	0.42	90.80	0.42	7.74	90.37	0.93	5.00E-27	PREDICTED: hypothetical protein LOC100117558	*Nasonia vitripennis*	Insect
contig167489	1614	0.45	70.18	0.33	7.72	69.84	0.92	6.00E-124	PREDICTED: zinc finger protein 182-like	*Nasonia vitripennis*	Insect

Transcript ID	Length	% GC	Mean Female Exprs	Mean Non-female Exprs	M*	D**	Prob	E-value	Annotation	Species	Taxa
contig14918	2656	0.39	279.98	1.34	7.71	278.64	0.97	3.00E-65	PREDICTED: putative ankyrin repeat protein RF_0381-like	*Nasonia vitripennis*	Insect
contig40538	1663	0.34	336.74	1.62	7.70	335.12	0.97	6.00E-92	PREDICTED: hypothetical protein LOC100122136	*Nasonia vitripennis*	Insect
contig5514	555	0.41	1170.73	5.65	7.70	1165.08	0.99	5.00E-21	PREDICTED: hypothetical protein LOC100118223	*Nasonia vitripennis*	Insect
contig180752	835	0.46	180.52	0.98	7.53	179.55	0.96	3.00E-28	PREDICTED: histone H1C-like	*Nasonia vitripennis*	Insect
contig30992	1076	0.36	55.15	0.30	7.51	54.85	0.90	2.00E-35	hypothetical protein TcasGA2_TC001485	*Tribolium castaneum*	Insect
contig40137	1470	0.31	78.64	0.44	7.47	78.20	0.92	1.00E-40	PREDICTED: protein roadkill-like	*Nasonia vitripennis*	Insect
contig180754	2783	0.40	124.22	0.71	7.46	123.51	0.95	0	PREDICTED: P protein-like	*Nasonia vitripennis*	Insect
contig184090	936	0.43	64.43	0.38	7.42	64.06	0.91	2.00E-11	PREDICTED: similar to Trypsin alpha	*Tribolium castaneum*	Insect
contig4602	4398	0.39	511.44	3.04	7.40	508.41	0.98	8.00E-129	aspartylglucosaminidase precursor	*Nasonia vitripennis*	Insect
contig39641	1847	0.34	74.49	0.45	7.39	74.04	0.92	5.00E-148	PREDICTED: venom carboxylesterase-6	*Nasonia vitripennis*	Insect
contig191455	1316	0.45	151.41	0.96	7.31	150.46	0.95	6.00E-76	PREDICTED: eukaryotic translation initiation factor 4E transporter-like	*Bombus impatiens*	Insect
contig162524	2304	0.41	57.40	0.36	7.31	57.03	0.90	2.00E-51	ankyrin repeat protein	*Trichomonas vaginalis G3*	Protista
contig6734	816	0.35	621.53	3.94	7.30	617.59	0.98	2.00E-45	gram-negative bacteria binding protein 1-2 precursor	*Nasonia vitripennis*	Insect
contig6098	762	0.48	3146.56	20.19	7.28	3126.37	0.99	6.00E-31	18S ribosomal RNA gene, partial sequence	*Aprostocetus purpureus*	Insect
contig179272	2375	0.29	56.13	0.36	7.28	55.77	0.90	7.00E-34	PREDICTED: hypothetical protein LOC100680289	*Nasonia vitripennis*	Insect
contig1995	1283	0.42	878.98	5.70	7.27	873.28	0.98	2.00E-21	hypothetical protein Bm1_17870	*Brugia malayi*	Other Animal
contig6517	4562	0.32	851.54	5.93	7.17	845.61	0.98	2.00E-39	PREDICTED: uncharacterized protein LOC754317	*Strongylocentrotus purpuratus*	Other Animal
contig13839	3616	0.46	757.05	5.61	7.08	751.44	0.98	0	PREDICTED: peroxidase-like	*Nasonia vitripennis*	Insect
contig4329	776	0.40	4189.70	31.15	7.07	4158.55	0.99	4.00E-18	PREDICTED: hypothetical protein LOC100118223	*Nasonia vitripennis*	Insect
contig6166	4767	0.33	559.07	4.23	7.05	554.84	0.98	2.00E-111	PREDICTED: hypothetical protein	*Hydra magnipapillata*	Other Animal
contig41203	839	0.43	598.57	4.53	7.04	594.03	0.98	1.00E-62	serine protease 52 precursor	*Nasonia vitripennis*	Insect
contig6520	1644	0.35	2700.75	20.55	7.04	2680.19	0.99	5.00E-31	PREDICTED: hypothetical protein LOC100121903	*Nasonia vitripennis*	Insect
contig182051	588	0.40	117.67	0.90	7.03	116.77	0.94	2.00E-13	PREDICTED: hypothetical protein LOC100680141	*Nasonia vitripennis*	Insect

Transcript ID	Length	% GC	Mean Female Exprs	Mean Non-female Exprs	M*	D**	Prob	E-value	Annotation	Species	Taxa
contig8327	3581	0.35	140.71	1.14	6.95	139.57	0.95	0	PREDICTED: peptide transporter family 1-like	*Nasonia vitripennis*	Insect
contig189802	1298	0.49	340.77	2.76	6.95	338.01	0.97	5.00E-80	PREDICTED: probable uridine-cytidine kinase-like isoform 2	*Nasonia vitripennis*	Insect
contig6739	769	0.37	5222.84	43.68	6.90	5179.16	0.99	4.00E-32	PREDICTED: hypothetical protein LOC100118223	*Nasonia vitripennis*	Insect
contig30865	1180	0.42	1057.81	8.96	6.88	1048.85	0.98	2.00E-62	PREDICTED: hypothetical protein LOC100678142	*Nasonia vitripennis*	Insect
contig43048	376	0.59	181.96	1.56	6.87	180.41	0.96	1.00E-64	PREDICTED: histone H3.2-like	*Cricetulus griseus*	Other Animal
contig190414	1099	0.38	61.04	0.54	6.82	60.50	0.90	3.00E-58	PREDICTED: trans-1,2-dihydrobenzene-1,2-diol dehydrogenase-like	*Nasonia vitripennis*	Insect
contig193227	1800	0.31	106.09	0.95	6.80	105.14	0.94	6.00E-82	PREDICTED: protein maelstrom homolog	*Megachile rotundata*	Insect
contig188733	1586	0.34	242.09	2.25	6.75	239.84	0.96	2.00E-71	PREDICTED: hypothetical protein LOC100678085	*Nasonia vitripennis*	Insect
contig35428	1011	0.41	78.45	0.74	6.73	77.72	0.92	7.00E-26	PREDICTED: histone H1C-like	*Nasonia vitripennis*	Insect
contig188031	2419	0.35	107.62	1.01	6.73	106.61	0.94	3.00E-159	PREDICTED: zinc carboxypeptidase A 1-like	*Nasonia vitripennis*	Insect
contig6496	1309	0.32	1895.45	18.28	6.70	1877.17	0.98	3.00E-11	PREDICTED: hypothetical protein LOC100121903	*Nasonia vitripennis*	Insect
contig29859	2120	0.48	67.40	0.66	6.68	66.75	0.91	7.00E-77	PREDICTED: hypothetical protein LOC100644105	*Bombus terrestris*	Insect
contig192773	480	0.44	2688.43	26.18	6.68	2662.24	0.99	6.00E-45	PREDICTED: vitellogenin-like	*Nasonia vitripennis*	Insect
contig10277	747	0.32	136.26	1.34	6.67	134.92	0.95	9.00E-13	PREDICTED: hypothetical protein LOC100678646	*Nasonia vitripennis*	Insect
contig7818	2175	0.39	322.89	3.31	6.61	319.58	0.97	3.00E-152	PREDICTED: zinc carboxypeptidase A 1-like	*Nasonia vitripennis*	Insect
contig44139	1296	0.41	115.08	1.19	6.60	113.90	0.94	5.00E-142	PREDICTED: protein takeout-like	*Nasonia vitripennis*	Insect
contig13256	2680	0.33	238.13	2.46	6.60	235.67	0.96	2.00E-131	maternal effect protein oskar	*Nasonia vitripennis*	Insect
contig189868	1341	0.40	89.91	0.93	6.60	88.98	0.93	4.00E-68	PREDICTED: transmembrane protease serine 9	*Nasonia vitripennis*	Insect
contig5525	3528	0.31	217.80	2.36	6.53	215.45	0.96	5.00E-30	PREDICTED: hypothetical protein LOC100121903	*Nasonia vitripennis*	Insect
contig39785	478	0.64	532.82	5.78	6.53	527.04	0.97	1.00E-87	hypothetical protein SINV_16114	*Solenopsis invicta*	Insect
contig41674	2548	0.28	99.78	1.09	6.52	98.69	0.93	2.00E-101	PREDICTED: hypothetical protein LOC100569507	*Acyrthosiphon pisum*	Insect
contig17906	5514	0.40	333.51	3.74	6.48	329.77	0.97	0	PREDICTED: L-ascorbate oxidase-like	*Nasonia vitripennis*	Insect
contig24148	2542	0.46	84.66	0.98	6.44	83.68	0.92	0	PREDICTED: nose resistant to fluoxetine	*Nasonia vitripennis*	Insect

Transcript ID	Length	% GC	Mean Female Exprs	Mean Non-female Exprs	M*	D**	Prob	E-value	Annotation	Species	Taxa
									protein 6-like		
contig12147	449	0.43	1933.17	22.37	6.43	1910.79	0.98	1.00E-12	PREDICTED: vitellogenin	*Nasonia vitripennis*	Insect
contig1996	576	0.31	655.18	8.11	6.34	647.07	0.98	4.00E-06	PREDICTED: hypothetical protein LOC100121115	*Nasonia vitripennis*	Insect
contig191610	2534	0.40	82.10	1.02	6.33	81.08	0.92	7.00E-18	replicase-associated protein [ringspot virus]	*Solanum violaefolium*	Virus
contig4263	1025	0.33	1356.46	17.05	6.31	1339.41	0.98	5.00E-26	PREDICTED: hypothetical protein LOC100118223	*Nasonia vitripennis*	Insect
contig2788	1245	0.34	2014.22	26.51	6.25	1987.72	0.98	9.00E-29	PREDICTED: hypothetical protein LOC100118223	*Nasonia vitripennis*	Insect
contig30820	3459	0.38	126.07	1.67	6.24	124.40	0.94	3.00E-44	PREDICTED: hypothetical protein LOC100678316	*Nasonia vitripennis*	Insect
contig11713	941	0.36	106.19	1.43	6.22	104.76	0.93	5.00E-112	PREDICTED: chymotrypsin-1	*Nasonia vitripennis*	Insect
contig7410	5541	0.40	3186.43	43.22	6.20	3143.20	0.98	0	PREDICTED: vitellogenin-like	*Nasonia vitripennis*	Insect
contig17463	395	0.27	442.76	6.13	6.17	436.62	0.97	2.00E-33	cytochrome oxidase I	*Megastigmus dorsalis*	Insect
contig30266	2482	0.35	467.71	6.61	6.15	461.11	0.97	2.00E-55	PREDICTED: transmembrane protease serine 9	*Nasonia vitripennis*	Insect
contig25820	1423	0.38	909.99	13.04	6.13	896.96	0.98	7.00E-115	PREDICTED: pyrimidine-specific ribonucleoside hydrolase rihA-like	*Nasonia vitripennis*	Insect
contig1888	2328	0.35	2348.76	34.67	6.08	2314.09	0.98	5.00E-29	PREDICTED: hypothetical protein LOC100118223	*Nasonia vitripennis*	Insect
contig12150	1210	0.41	12821.48	195.98	6.03	12625.51	0.98	8.00E-98	PREDICTED: vitellogenin	*Nasonia vitripennis*	Insect
contig4550	720	0.40	2189.83	34.05	6.01	2155.77	0.98	4.00E-33	PREDICTED: hypothetical protein LOC100118223	*Nasonia vitripennis*	Insect
contig6504	1870	0.37	1230.44	19.19	6.00	1211.25	0.98	3.00E-30	PREDICTED: methionine aminopeptidase 2-like	*Nasonia vitripennis*	Insect
contig5957	1095	0.36	68.90	1.08	6.00	67.82	0.91	3.00E-81	PREDICTED: protein lethal(2)essential for life-like [Nasonia vitripennis]		Insect
contig12149	435	0.44	1176.35	18.52	5.99	1157.83	0.98	4.00E-46	PREDICTED: vitellogenin-like	*Nasonia vitripennis*	Insect
contig8177	1044	0.25	263.41	4.19	5.97	259.22	0.96	8.00E-46	PREDICTED: similar to F59H6.5	*Hydra magnipapillata*	Other Animal
contig29520	1430	0.34	66.95	1.07	5.97	65.88	0.91	3.00E-50	PREDICTED: speckle-type POZ protein B-like	*Nasonia vitripennis*	Insect
contig30845	1463	0.34	106.05	1.72	5.94	104.32	0.93	2.00E-19	antimicrobial peptide defensin 1-2 precursor	*Nasonia vitripennis*	Insect
contig188568	2759	0.54	111.55	1.84	5.92	109.71	0.93	2.00E-173	PREDICTED: hypothetical protein LOC100117267	*Nasonia vitripennis*	Insect
contig193099	2237	0.33	226.84	3.81	5.90	223.03	0.96	5.00E-133	maternal effect protein oskar	*Nasonia vitripennis*	Insect

Transcript ID	Length	% GC	Mean Female Exprs	Mean Non-female Exprs	M*	D**	Prob	E-value	Annotation	Species	Taxa
contig192994	684	0.43	196.25	3.43	5.84	192.82	0.95	2.00E-33	PREDICTED: hypothetical protein LOC100118223	*Nasonia vitripennis*	Insect
contig4530	2119	0.40	199.00	3.53	5.82	195.47	0.95	3.00E-55	PREDICTED: hypothetical protein LOC100123806	*Nasonia vitripennis*	Insect
contig38878	4233	0.39	195.00	3.64	5.74	191.36	0.95	0	PREDICTED: hypothetical protein LOC100120040	*Nasonia vitripennis*	Insect
contig30563	359	0.50	187.50	3.50	5.74	183.99	0.95	5.00E-47	PREDICTED: fatty acid synthase-like	*Nasonia vitripennis*	Insect
contig5948	1551	0.37	708.74	13.38	5.73	695.36	0.97	8.00E-08	b01 predicted protein, mRNA	*Paracoccidioides brasiliensis*	Fungi
contig440	1698	0.40	172.79	3.28	5.72	169.51	0.95	5.00E-35	PREDICTED: hypothetical protein LOC100576236	*Apis mellifera*	Insect
contig578	2390	0.32	479.03	9.22	5.70	469.82	0.97	3.00E-23	PREDICTED: hypothetical protein LOC100121903	*Nasonia vitripennis*	Insect
contig28977	2040	0.36	1346.23	26.48	5.67	1319.75	0.97	4.00E-41	hymenoptaecin-1 isoform 2 precursor	*Nasonia vitripennis*	Insect
contig149749	1168	0.55	107.96	2.18	5.63	105.77	0.93	2.00E-66	PREDICTED: hypothetical protein LOC100743039	*Bombus impatiens*	Insect
contig12189	945	0.41	781.71	15.95	5.62	765.76	0.97	9.00E-15	PREDICTED: hypothetical protein LOC100678434	*Nasonia vitripennis*	Insect
contig39784	554	0.57	631.77	13.20	5.58	618.57	0.97	6.00E-94	PREDICTED: histone H3.1-like	*Saimiri boliviensis boliviensis*	Other Animal
contig30890	849	0.46	107.10	2.25	5.57	104.84	0.93	3.00E-28	PREDICTED: histone H1C-like	*Nasonia vitripennis*	Insect
contig36223	3961	0.34	69.60	1.61	5.43	67.99	0.90	8.00E-99	PREDICTED: hypothetical protein LOC100679988	*Nasonia vitripennis*	Insect
contig2854	2044	0.34	1038.11	25.35	5.36	1012.76	0.97	2.00E-12	PREDICTED: hypothetical protein LOC100678823	*Nasonia vitripennis*	Insect
contig34247	10918	0.45	1468.18	36.25	5.34	1431.93	0.97	0	PREDICTED: hypothetical protein LOC100120269	*Nasonia vitripennis*	Insect
contig184787	1472	0.58	72.20	1.80	5.33	70.41	0.90	4.00E-71	PREDICTED: hypothetical protein LOC100743039	*Bombus impatiens*	Insect
contig12151	4990	0.41	7200.72	179.51	5.33	7021.20	0.97	0	vitellogenin	*Pteromalus puparum*	Insect
contig11056	726	0.57	160.39	4.01	5.32	156.39	0.94	4.00E-26	PREDICTED: histone H1C-like	*Nasonia vitripennis*	Insect
contig188553	1762	0.34	279.35	7.08	5.30	272.26	0.95	6.00E-48	PREDICTED: maltase 1-like	*Nasonia vitripennis*	Insect
contig36317	657	0.39	107.95	2.74	5.30	105.21	0.92	1.00E-09	venom protein R precursor	*Nasonia vitripennis*	Insect
contig34497	745	0.35	1000.23	26.61	5.23	973.62	0.97	2.00E-19	chemosensory protein CSP-1	*Polistes dominulus*	Insect
contig18900	1809	0.41	907.49	24.52	5.21	882.96	0.97	1.00E-180	PREDICTED: stearoyl-CoA desaturase 5-like	*Nasonia vitripennis*	Insect
contig4469	1722	0.40	612.96	16.75	5.19	596.20	0.96	1.00E-145	PREDICTED: chitotriosidase-1-like	*Nasonia vitripennis*	Insect
contig191775	750	0.45	318.08	9.07	5.13	309.01	0.95	8.00E-33	PREDICTED: hypothetical protein	*Nasonia vitripennis*	Insect

Transcript ID	Length	% GC	Mean Female Exprs	Mean Non-female Exprs	M*	D**	Prob	E-value	Annotation	Species	Taxa
									LOC100119484 isoform 2		
contig4769	673	0.38	270.70	7.72	5.13	262.98	0.95	1.00E-09	structural maintenance of chromosomes 2	*Nasonia vitripennis*	Insect
contig23450	612	0.49	508.97	14.92	5.09	494.05	0.96	4.00E-57	GJ23521	*Drosophila virilis*	Insect
contig26508	3046	0.32	190.32	5.82	5.03	184.50	0.94	6.00E-68	protein nanos	*Nasonia vitripennis*	Insect
contig150120	2418	0.50	125.65	3.84	5.03	121.80	0.93	0	PREDICTED: atrial natriuretic peptide-converting enzyme-like	*Nasonia vitripennis*	Insect
contig179260	2577	0.57	123.81	3.79	5.03	120.02	0.93	0	PREDICTED: protein neuralized-like	*Bombus terrestris*	Insect
contig39750	2972	0.37	296.08	9.07	5.03	287.01	0.95	5.00E-71	serine protease 100 precursor	*Nasonia vitripennis*	Insect
contig23660	1319	0.63	88.14	2.72	5.02	85.42	0.91	2.00E-55	PREDICTED: hypothetical protein LOC100114548	*Nasonia vitripennis*	Insect
contig27330	3208	0.45	82.76	2.71	4.93	80.05	0.90	1.00E-110	kruppel	*Nasonia vitripennis*	Insect
contig37971	948	0.50	223.04	7.36	4.92	215.68	0.94	4.00E-74	PREDICTED: eukaryotic translation initiation factor 4E transporter-like	*Bombus impatiens*	Insect
contig167194	1176	0.56	163.24	5.40	4.92	157.84	0.93	4.00E-26	PREDICTED: histone H1C-like	*Nasonia vitripennis*	Insect
contig6483	5655	0.36	143.48	4.80	4.90	138.68	0.93	2.00E-119	PREDICTED: hypothetical protein LOC100124061	*Nasonia vitripennis*	Insect
contig5937	730	0.46	99.95	3.35	4.90	96.60	0.91	9.00E-56	PREDICTED: hypothetical protein LOC100123504	*Nasonia vitripennis*	Insect
contig28113	4617	0.39	914.52	31.57	4.86	882.95	0.96	0	PREDICTED: solute carrier family 22 member 3-like	*Nasonia vitripennis*	Insect
contig7168	3410	0.41	124.78	4.36	4.84	120.43	0.92	0	PREDICTED: hypothetical protein LOC100679702	*Nasonia vitripennis*	Insect
contig8365	2348	0.35	2163.06	78.52	4.78	2084.55	0.96	1.00E-152	PREDICTED: stearoyl-CoA desaturase 5-like isoform 1	*Nasonia vitripennis*	Insect
contig941	4252	0.32	228.28	8.35	4.77	219.93	0.94	2.00E-41	hypothetical protein BRAFLDRAFT_79791	*Branchiostoma floridae*	Other Animal
contig6580	2566	0.34	382.20	14.24	4.75	367.97	0.95	2.00E-08	PREDICTED: uncharacterized protein LOC100888155	*Strongylocentrotus purpuratus*	Other Animal
contig34392	2635	0.36	167.66	6.26	4.74	161.40	0.93	2.00E-165	serine/threonine specific protein phosphatase	*Rattus norvegicus*	Other Animal
contig8268	1141	0.32	1281.40	48.34	4.73	1233.06	0.96	3.00E-15	PREDICTED: hypothetical protein LOC100678728	*Nasonia vitripennis*	Insect
contig189278	1472	0.58	95.12	3.67	4.70	91.45	0.91	4.00E-71	PREDICTED: hypothetical protein LOC100743039	*Bombus impatiens*	Insect
contig188862	2228	0.42	83.37	3.33	4.64	80.04	0.90	0	PREDICTED: tryptophan 5-hydroxylase 1-like	*Nasonia vitripennis*	Insect
contig34349	4584	0.39	127.89	5.15	4.64	122.74	0.92	8.00E-127	PREDICTED: hypothetical protein LOC100678793 isoform 1	*Nasonia vitripennis*	Insect

Transcript ID	Length	% GC	Mean Female Exprs	Mean Non-female Exprs	M*	D**	Prob	E-value	Annotation	Species	Taxa
contig29105	2544	0.49	160.93	6.48	4.63	154.44	0.93	2.00E-26	PREDICTED: hypothetical protein LOC100120441	*Nasonia vitripennis*	Insect
contig7424	1937	0.35	250.43	10.20	4.62	240.23	0.94	3.00E-53	PREDICTED: transmembrane protease serine 9	*Nasonia vitripennis*	Insect
contig2652	655	0.37	2281.02	95.29	4.58	2185.73	0.96	1.00E-26	PREDICTED: hypothetical protein LOC100118223	*Nasonia vitripennis*	Insect
contig190878	1897	0.40	186.07	7.87	4.56	178.20	0.93	0	PREDICTED: vitellogenin	*Nasonia vitripennis*	Insect
contig6194	2285	0.33	864.67	36.63	4.56	828.04	0.95	8.00E-13	PREDICTED: hypothetical protein LOC100124061	*Nasonia vitripennis*	Insect
contig3570	947	0.42	297.54	12.81	4.54	284.73	0.94	3.00E-62	PREDICTED: transmembrane protease serine 9	*Nasonia vitripennis*	Insect
contig38721	1854	0.33	289.55	12.53	4.53	277.01	0.94	5.00E-81	PREDICTED: hypothetical protein LOC100119416	*Nasonia vitripennis*	Insect
contig189231	510	0.46	676.64	29.49	4.52	647.14	0.95	3.00E-45	PREDICTED: fatty acid synthase-like	*Nasonia vitripennis*	Insect
contig172167	617	0.51	332.64	14.65	4.51	317.99	0.94	2.00E-67	GM13182	*Drosophila sechellia*	Insect
contig188500	1149	0.43	99.44	4.50	4.46	94.93	0.91	1.00E-54	PREDICTED: hypothetical protein LOC100678420	*Nasonia vitripennis*	Insect
contig184762	1781	0.44	129.83	6.01	4.43	123.81	0.92	0	PREDICTED: major royal jelly protein 5	*Nasonia vitripennis*	Insect
contig26588	1973	0.35	110.35	5.52	4.32	104.83	0.91	3.00E-158	PREDICTED: elongation of very long chain fatty acids protein 7-like	*Nasonia vitripennis*	Insect
contig7408	913	0.41	209.23	10.63	4.30	198.61	0.93	3.00E-103	vitellogenin	*Pteromalus puparum*	Insect
contig6561	679	0.40	3014.29	154.67	4.28	2859.62	0.95	2.00E-25	PREDICTED: hypothetical protein LOC100117346 isoform 1	*Nasonia vitripennis*	Insect
contig185682	1043	0.54	226.27	12.37	4.19	213.89	0.93	0	PREDICTED: glutamine synthetase 1, mitochondrial-like	*Nasonia vitripennis*	Insect
contig14873	1476	0.46	244.62	13.45	4.18	231.17	0.93	4.00E-139	PREDICTED: facilitated trehalose transporter Tret1-like	*Nasonia vitripennis*	Insect
contig40585	963	0.37	240.49	13.47	4.16	227.02	0.93	2.00E-55	Histone H2B	*Lepeophtheirus salmonis*	Other Animal
contig28173	3801	0.32	187.53	10.52	4.16	177.00	0.92	6.00E-95	hypothetical protein CRE_16968	*Caenorhabditis remanei*	Other Animal
contig23053	2179	0.38	295.04	16.94	4.12	278.11	0.93	2.00E-176	PREDICTED: hypothetical protein LOC100678247	*Nasonia vitripennis*	Insect
contig3242	3056	0.35	2260.99	130.12	4.12	2130.86	0.95	5.00E-131	PREDICTED: hypothetical protein LOC100634292	*Amphimedon queenslandica*	Other Animal
contig6715	10974	0.44	125.61	7.27	4.11	118.34	0.91	8.00E-122	hypothetical protein TcasGA2_TC001491	*Tribolium castaneum*	Insect
contig34168	2883	0.58	131.27	7.62	4.11	123.66	0.91	0	PREDICTED: hypothetical protein LOC100117630 isoform 2	*Nasonia vitripennis*	Insect

Transcript ID	Length	% GC	Mean Female Exprs	Mean Non-female Exprs	M*	D**	Prob	E-value	Annotation	Species	Taxa
contig43813	3187	0.40	302.35	17.60	4.10	284.74	0.93	0	PREDICTED: aminopeptidase N-like	*Nasonia vitripennis*	Insect
contig5653	3362	0.38	158.87	9.27	4.10	149.60	0.91	3.00E-118	PREDICTED: hypothetical protein LOC100123504	*Nasonia vitripennis*	Insect
contig6777	1142	0.38	3954.26	231.59	4.09	3722.67	0.95	4.00E-166	PREDICTED: maltase 1-like	*Nasonia vitripennis*	Insect
contig13203	4836	0.39	320.83	19.02	4.08	301.81	0.93	0	PREDICTED: hypothetical protein LOC100119484 isoform 2	*Nasonia vitripennis*	Insect
contig1793	1990	0.39	213.08	12.82	4.05	200.26	0.92	7.00E-105	PREDICTED: hypothetical protein LOC100678556	*Nasonia vitripennis*	Insect
contig6208	4998	0.38	454.02	27.66	4.04	426.36	0.93	0	PREDICTED: hypothetical protein LOC100124061	*Nasonia vitripennis*	Insect
contig190549	1417	0.38	250.39	15.33	4.03	235.06	0.92	6.00E-64	PREDICTED: facilitated trehalose transporter Tret1-like	*Nasonia vitripennis*	Insect
contig3249	1410	0.44	433.25	27.15	4.00	406.10	0.93	6.00E-90	Histone H3c	*Culex quinquefasciatus*	Insect
contig3551	2325	0.43	374.76	23.94	3.97	350.82	0.93	0	PREDICTED: cytochrome P450 6k1	*Nasonia vitripennis*	Insect
contig5171	1324	0.41	181.42	11.64	3.96	169.78	0.91	2.00E-67	PREDICTED: transmembrane protease serine 9	*Nasonia vitripennis*	Insect
contig18357	7266	0.39	469.71	31.83	3.88	437.87	0.93	0	PREDICTED: nose resistant to fluoxetine protein 6-like	*Nasonia vitripennis*	Insect
contig18820	2835	0.48	391.66	29.24	3.74	362.42	0.92	0	PREDICTED: alkaline phosphatase 4-like	*Nasonia vitripennis*	Insect
contig40259	6192	0.48	635.65	49.25	3.69	586.40	0.92	0	PREDICTED: vitellogenin receptor	*Nasonia vitripennis*	Insect
contig13830	3363	0.33	877.41	68.33	3.68	809.07	0.92	0	PREDICTED: pseudouridylate synthase 7 homolog	*Nasonia vitripennis*	Insect
contig31792	663	0.48	272.44	21.44	3.67	251.01	0.91	3.00E-52	Unknown (protein for MGC:185872)	*Xenopus (Silurana) tropicalis*	Other Animal
contig1103	5303	0.46	268.57	21.37	3.65	247.20	0.91	0	PREDICTED: hypothetical protein LOC100120660	*Nasonia vitripennis*	Insect
contig2411	1749	0.45	199.28	16.58	3.59	182.69	0.90	1.00E-63	PREDICTED: transmembrane protease serine 9	*Nasonia vitripennis*	Insect
contig40453	1623	0.39	1391.30	126.95	3.45	1264.34	0.92	5.00E-121	uncharacterized protein LOC100116503 precursor	*Nasonia vitripennis*	Insect
contig12835	843	0.44	589.00	55.76	3.40	533.24	0.91	2.00E-36	PREDICTED: hypothetical protein LOC100114335 (LOC100114335), mRNA	*Nasonia vitripennis*	Insect
contig7350	6606	0.45	471.97	44.90	3.39	427.07	0.91	0	PREDICTED: eukaryotic translation initiation factor 4E transporter-like	*Bombus impatiens*	Insect

*M is the $\log_2$ ratio the mean female expression and the mean non-female expression.

**D is the difference between the mean female expression and the mean non-female expression