Interaction Between the Seed-Chalcid Wasp, *Megastigmus spermotrophus* and its Host, Douglas-fir (*Pseudotsuga menziesii*)

by

Kathleen Louise Donaleshen
B.Sc., University of Victoria, 2012

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

MASTER OF SCIENCE

in the Department of Biology

Supervisory Committee

Interaction Between the Seed-Chalcid Wasp, *Megastigmus spermotrophus* and its Host, Douglas-fir (*Pseudotsuga menziesii)*

by

Kathleen Louise Donaleshen
B.Sc., University of Victoria, 2012

Supervisory Committee

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

Dr. Jürgen Ehlting (Department of Biology)
Co-Supervisor

Dr. Steve Perlman (Department of Biology)
Departmental Member

Abstract

Supervisory Committee
Dr. Patrick von Aderkas (Department of Biology)
Co-Supervisor

Dr. Jürgen Ehlting (Department of Biology)
Co-Supervisor

Dr. Steve Perlman (Department of Biology)
Departmental Member

Megastigmus spermotrophus is a parasitic chalcid wasp that spends most of its life in the seed of its host, Douglas-fir (*Pseudotsuga menziesii*). The adult female wasp lays its eggs into the megagametophyte deep within the ovule; the larva prevents an unpollinated ovule from aborting, redirecting resources to feed itself. Host-site selection pressures that influence female oviposition depend on a number of factors. Morphological characteristics of Douglas-fir cones including seed size, seed location, and scale thickness were measured for every ovuliferous scale. Seeds infested by *M. spermotrophus* as well as seeds fused to galls intiated by a competing conophyte, *Contarinia oregonensis* were noted. Using a generalized linear mixed effects model, I found that seed position, and the presence of *C. oregonensis*, were strong predictors of *Megastigmus* infestation. The percent of *M. spermotrophus* infested seed was higher in the apical and basal regions of the cone where seeds were smaller, scales were thinner and *C. oregonensis* were less frequently found. *M. spermotrophus* was also found to exploit seeds in regions of the cone, where seeds rarely complete development. These data suggest that competitors may not be the only factor influencing infestation; factors of cone morphology are also important.

Douglas-fir seed does not show any anatomically detectable defense response to *Megastigmus* attack. To study mechanisms of host manipulation and defense response of the seed I took a genomics approach. Four types of ovules/seeds were studied: 1. pollinated & uninfested, 2. pollinated & infested, 3. unpollinated & uninfested, and 4. unpollinated and infested. A *de novo* reference transcriptome was assembled. Transcripts were annotated based on sequence similarity to genes of *Pinus taeda, Arabidopsis thaliana, Nasonia vitripennis*, and the UniProt database. Expression values were estimated based on the alignment of the original reads back onto the reference transcriptome. Differentially expressed transcripts were identified. Oviposition of *M. spermotrophus* caused changes in expression of Douglas-fir transcripts. Functional classification of differentially expressed transcripts between infested and uninfested seed revealed genes with possible roles in wounding, but none specific to herbivory. Infested treatments had more transcripts similarly expressed to pollinated than unpollinated seeds suggesting that *M. spermotrophus* is capable of manipulating gene expression. These transcripts had functional roles related to seed storage, cell division and growth, solute transport, hormone signalling, and programmed cell death among others. Overall, this study reveals a select set of genes that may be involved in stress response to wounding and also genes important for seed development and maturation.

Table of Contents

List of Tables

List of Figures

Acknowledgments

First and foremost, I thank my supervisors, Dr. Patrick von Aderkas and Dr. Jürgen Ehlting. You have both been excellent mentors, being incredibly supportive and encouraging throughout my time at UVic. Patrick and Jürgen, thank you for inspiring my scientific thought and for gradually pushing me in the right direction. This has been a wonderful opportunity. I have truly enjoyed every ounce of this project. Also, I would like to thank Dr. Steve Perlman for sharing his expertise and knowledge. I would like to thank all of my lab mates in the von Aderkas and Ehlting labs. You have all been delightful to work with. A special thank you to Ian Boyes. You were an excellent teacher. Thank you for your time and patience while I learned the world of Unix. Stefan Little, thank you for all the conversations, tips, and coffee dates. I appreciate your constant enthusiasm and feedback. David Minkley, I am incredibly grateful for your availability and willingness when I was in need of bioinformatics support. I would also like to thank Tim Crowder from Timber West for generously allowing me to conduct my fieldwork at Mt. Newton Seed Orchard. Thank you David Kolotelo from the Surrey Seed Centre for providing *Megastigmus* infested seed for this project. To my mom, Mary Ann Donaleshen, thank you for your unwavering love, support, and encouragement. In times of need you are always at my side. Cameron Freshwater, James Robinson, and Travis Tai, your positivity and eagerness for science were instrumental in helping me discover my passion for research. Thank you for making this a pleasant and positive experience.

CHAPTER 1. INTERACTION BETWEEN DOUGLAS-FIR AND *M. SPERMOTROPHUS*

1.1 Forest sustainability and seed trade

As the world's population continues to rise, so does the demand for quality lumber. This increased global demand for wood products has created an added pressure on the regeneration of natural stands. Until the mid 1900's, forest sustainability was left to natural regeneration (Herman and Lavender, 1999). Although naturally regenerated forests are highly productive, they are unable to adequately sustain themselves at the rate we are processing lumber. In the 1950's tree improvement plans were initiated (Herman and Lavender, 1999). In order to supply enough seed annually for forest regeneration, seed orchards have been established that have genetically improved stock (Plomion et al., 2011). Today, planting programs rely heavily on seed orchards for mass production.

1.2 Introduction of invasive species

During the late 19th century there was a mass global movement and introduction of conifer species. This intercontinental movement of both trees and of wood products has resulted in the establishment of invasive, parasitic seed insects (Roques et al., 2003), in particular, *Megastigmus*. *Megastigmus spermotrophus* (Wachtl) was accidently introduced to Europe when Douglas-fir was brought over from North America (Milliron, 1949). The introduction and spread of these species has resulted in a tremendous ecological impact, e.g. on tree growth, mortality and reproduction. This has altered the ecological services provided by these conifer species (Liebhold et al., 1995). Such

impacts underscore the need for strict international trade guidelines as well as the need for preventative measures and early detection of insect attack (Roques et al., 2003). The potential that an insect may pose a threat ought to be taken into consideration before seed is moved between biomes, let alone continents.

1.3 *Pseudotsuga menziesii*

Douglas-fir is a member of Pinaceae, the largest family of conifers (Allen and Owens, 1972). *Pseudotsuga Menziesii* (Mirb.) Franco has the largest longitudinal range of all members of this genus. Its natural range extends from central British Columbia to California and Mexico. In optimal conditions, Douglas-fir is one of the world's tallest trees; there is historical evidence of individual trees easily exceeding 100 m (Eckenwalder, 2009). Of the conifers in the temperate zone, Douglas-fir is also the most productive. Its rapid growth, along with superior wood quality has undoubtedly made Douglas-fir an economically important lumber species (Bormann, 1984). This species is often planted for its high-density wood, which has superior stress properties. Douglas-fir grows on 4.5 million ha in Canada, and 17 million ha in the United States (Herman and Lavender, 1999).

1.3.1 Exotic forestry

Of the numerous conifers introduced to Europe, *P. menziesii* has become one of the most valued and productive timber species. It was originally introduced to Europe as an ornamental plant by the Scottish botanist David Douglas in 1827 (Milliron, 1949). By

the end of the 19th century, it was widely planted in forest plantations (Plomion, et al., 2011). Today, Douglas-fir grows on over 700, 000 ha of forest in Europe (Herman and Lavender, 1999). Many countries, including France, Germany, the United Kingdom, and the Netherlands have successfully introduced Douglas-fir as a lumber species. France, alone has over 400, 000 ha of planted Douglas-fir, accounting for more than half of the total coverage in Western Europe.

Both coastal and interior varieties of Douglas-fir were introduced into Europe from North America. The interior variety was however, unsuitable for cultivation. Successful introduction of *P. menziesii* can be attributed to the combination of favourable site conditions, availability of a suitable seed source, lack of natural pests, and its ability to adapt to a variety of soils and climates (Bormann, 1984).

A key reason for choosing Douglas-fir as a study organism is that its reproduction is the best understood of all the western conifers. Embryogeny, a field that includes prezygotic development of males and females, as well as embryology – the development of the embryo – has been dissected in a detailed and systematic manner. This has allowed manipulation of phenologies, i.e., developing overhead irrigation to delay pollen release, application schedules of pollen for supplemental mass pollination collection, and accurate breeding programs. The details are important contributing factors affecting field experimentation.

1.3.2 Embryogeny

The reproductive cycle of Douglas-fir extends over a seventeen-month period (Figure 1). Bud primordia are initiated from cells above the nodes at the beginning of April (Owens and Smith, 1964). These will eventually differentiate into vegetative or reproductive buds (Allen and Owens, 1972). The type of bud can be distinguished approximately 15 weeks after initiation. By November, buds enter a period of dormancy that lasts for three months (Allen and Owen, 1972). Bud growth resumes the following March. In the northern range of Douglas-fir, bud break occurs in April.

1.3.2.1 Microsporogenesis

Microsporangia are initiated in early summer, and by the end of autumn are fairly well developed. They remain dormant until the end of February, at which time the microspore mother cells undergo meiosis to produce a tetrad of haploid microspores. Cell division of the microspores produces mature, five-celled pollen grains consisting of two prothallial cells, a stalk cell, a body cell and a tube cell (Allen and Owens, 1972. At this five-cell stage, which typically occurs at the beginning of April, anthesis begins, i.e. pollen is shed (See A, Figure 1). It is not until after pollen germination that the two male gametes will form as a result of the mitotic division of the body cell (Allen and Owens, 1972).

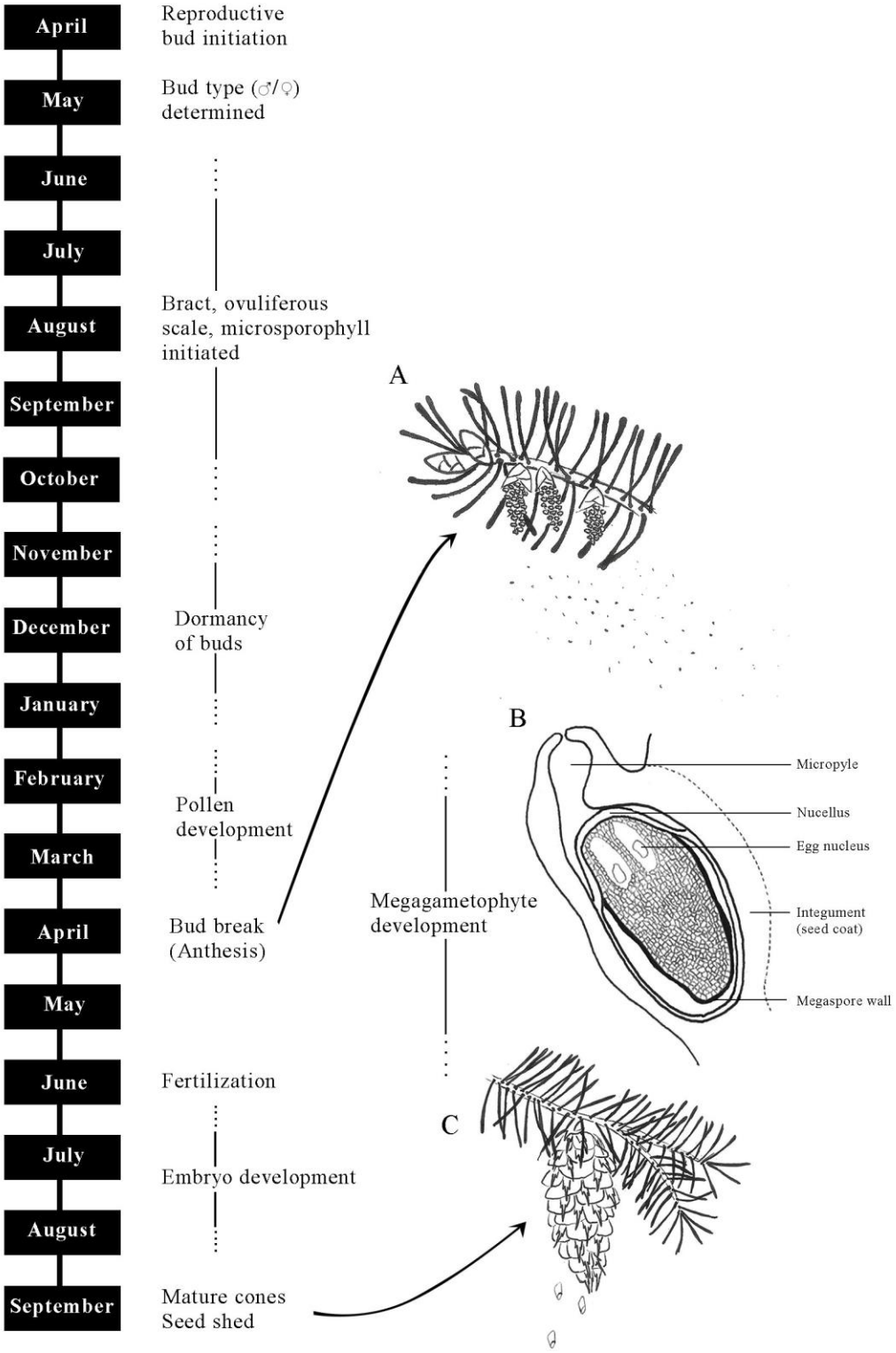


Figure 1. Timeline of Douglas-fir reproduction over a 17-month period, from bud initiation to seed shed. Illustrations of A. anthesis, B. a megagametophyte prior to fertilization, and C. a mature cone dropping seed, are included.

1.3.2.1 Megasporogenesis

Cone buds are initiated in late summer and then go dormant over winter. Megaspore mother cells that have differentiated on distal portions of the ovuliferous scale stay dormant until mid-February, when dormancy ends and ovule development continues (Allen, 1943).

Megaspore mother cells divide by meiosis to produce a tetrad of haploid megaspores; three megaspores degenerate, but the fourth enlarges to become the functional megaspore. Each megaspore undergoes several synchronous, free nuclear divisions, resulting in large, single cells known as coenocytes that have 250-1000 nuclei (Allen and Owens, 1972). This period of free nuclear division lasts up to four weeks (Lawson, 1909). Alveolation occurs in the chalazal end of the megagametophyte, resulting in the formation of numerous cells (Singh, 1978). After continued division these form the mass of prothallial cells. Of these, four to six cells become archegonial initials (Allen, 1943; Allen and Owens, 1972).

Unequal division of each archegonial initial produces a primary neck cell – that undergoes further division to produce a layer of neck cells to the outside, and a large inner central cell. The surrounding prothallial cells subsequently form a jacket layer around the maturing central cell (Owens and Morris, 1990). By June, the central cell divides to form a ventral canal cell and an egg cell (Allen, 1943; Allen and Owens, 1972) (See B, Figure 1).

1.3.2.2 Pollination and fertilization

At maturation, microsporangia dehisce along longitudinal slits, allowing the pollen grains to be shed and then dispersed by wind (Singh, 1978). Mature pollen grains are carried only short distances (Allen and Owens, 1972). Pollen release occurs over a period of two to three weeks in early spring. During pollination, receptive female cones are upright with their ovuliferous scale-bract complexes separated to allow pollen to enter and contact the ovules. The tips of the bracts are bent back and the margins curve upwards, creating a funnel to augment pollen capture (Allen and Owens, 1972). Individual ovules are only receptive for a few days.

The stigmatic hairs on either side of the micropyle catch pollen (Allen and Owens, 1972). These hairs then collapse, engulfing the pollen and forcing the grains inside the micropyle, or ovule opening. After the micropyle is sealed, pollen can no longer enter (Owens et al., 1991).

After several days the engulfed pollen swell and elongate within the micropyle (Allen, 1963). A long period of inactivity follows. After about four to six weeks, a drop is secreted into the micropyle that triggers germination (von Aderkas and Leary, 1999; Poulis et al., 2005). Pollen germination tubes develop and grow into the nucellus (Owens and Morris, 1990). Pollen tubes penetrate the archegonium and release their gametes (Lawson, 1909). Fertilization occurs between one of the male gametes and the egg. The other male gamete serves no function. In *Pseudotsuga*, fertilization occurs from early to mid-June.

1.3.2.3 Embryo development

Embryo development in Douglas-fir has four stages – proembryo, early embryo, late embryo, and dormant embryo development. These occur over several months. During proembryo development a zygote is formed. This stage is very brief. i.e., a few days. The nucleus of the zygote immediately undergoes three free nuclear divisions, followed by cellularization (Allen, 1943). The most apical cells repeatedly divide to form an embryonal mass (Singh, 1978). Below these cells are the suspensor cells whose function is to push the developing embryo out of the archegonium, through the surrounding jacket cells and into the corrosion cavity of the megagametophyte where the embryo will be nourished (Allen, 1943). There is competition among embryos. There are up to six eggs in an ovule but only one embryo will reach maturity. This constitutes a form of mate selection (Wilson and Burley, 1983).

During early embryo development, root and shoot meristems are initiated and the embryo elongates. Meristems and cotyledons are formed during late embryo development (Owens et al., 1992). Before the embryo has completely matured, a seed coat and seed wing develop from the integument and ovuliferous scale, respectively (Allen and Owens, 1972). By September the cones open, seeds are released, and flutter to the ground with the aid of their one wing (See C, Figure 1).

1.3.3 The megagametophyte

The megagametophyte supplies a developing embryo with the nutrients required for growth (Singh, 1978). There is a spatial and temporal zonation of the storage reserves

within a megagametophyte. Following fertilization, the prothallial cells in the central region of the megagametophyte, closest to the embryo, accumulate large storage reserves in the form of starch (Owens et al., 1992). These cells break down to form a nutrient rich liquid that is confined to a cone-like space, called the corrosion cavity. Over a few weeks, the corrosion cavity expands and the prothallial cells on the corrosion cavities margin fill with lipid bodies and large protein bodies (von Aderkas et al., 2005b) and in turn, break down. Many of these cells are multi-nucleate. In white spruce, proteins make up 70% of the storage reserves in the megagametophyte (Misra and Green, 1990). Following germination, the storage proteins are broken down to amino acids that are mobilized to feed the embryo (Misra and Green, 1991). Over the course of Douglas-fir embryogenesis and gametophyte development, levels of hormones, such as auxin, cytokinin and abscisic acid (ABA) are regulated (Chiwocha and von Aderkas, 2002). ABA levels are especially high within the fluid of the corrosion cavity, suggesting a possible role in solute mobilization into the embryo (Carman et al., 2005).

Not all ovules are fertilized. In the absence of pollen, fertilization fails and the gametophyte dies. Degeneration of the megagametophyte begins once the egg has died, which occurs over a period of one to two weeks. Contents of the egg and the megagametophyte are reclaimed by the tree. The presence of an embryo appears to be the factor necessary for continual development as well as for storage deposition of lipids, starch and proteins in the megagametophyte (Orr-Ewing, 1956; Owens et al., 1992).

1.3.4 Factors affecting seed set

Conifer reproduction is complex: it requires development of male and female cones to be synchronized. Douglas-fir seed set is influenced by prezygotic and postzygotic events (Owens et al., 1991). In conifers, low seed set may be due to a number of factors including, low male or female cone production, self-pollination, and poor synchronization of male and female cone phenology. Although pollination can be a significant factor affecting seed set, genetic load will also influence embryo abortion (Owens et al., 1991). It is assumed that conifers have a stringent mechanism for post-zygotic selection compared to pre-zygotic selection (Owens and Morris, 1991). Ovules have been shown to abort if the embryos are absent (Orr-Ewing, 1956). Embryo collapse often occurs due to increased homozygosity of recessive deleterious alleles that may be a result of self-pollination. Thus, cross-pollination is essential for ensuring a good seed crop.

1.4 *Megastigmus*

Parasitic seed insects substantially impact seed yield. There are over 400 phytophagous insect species belonging to seven different orders, responsible for much of the seed loss in conifers following fertilization (Turgeon et al., 1994). Many of these species found in North America and Europe are highly specialized for feeding or developing within conifer seed. Among these insects, species within the genus, *Megastigmus, Dioryctria, Contarinia, Leptoglossus,* and *Barbara* and have become serious pests of commercially important conifers. On an international scale, *Megastigmus* Dalman (Hymneoptera: Chalcidoidea: Torymidae: Megastiminae) has received the most attention due to mass

destruction of seed crops. For this reason and for its interaction with Douglas-fir, *Megastigmus* will be the focus of this thesis.

Megastigmus is a genus of parasitic chalcid wasps distributed throughout many regions of the world (Roques and Skrzypczynska, 2003). The genus divides into parasites of animals and plants. The plant parasites can be divided into three distinct groups, according to whether they attack pinaceaeous, or cupressaceaeous conifers or angiosperms (Roux and Roques, 1996). Four feeding patterns have been observed – facultative and obligate parasitoids, gall formers and seed feeders (Roques and Skrzypczynska, 2003). Over 125 species have been described, of which 41 feed on conifer seed (Grissell, 1999; Roques and Skrzypczynska, 2003). *Megastigmus* species infesting conifers are highly specialized and generally fixed to interactions at the genus level, indicating a loss of genetic variation in the parasite (Auger-Rozenberg et al., 2005). Evolutionary trends indicate there are relatively low levels of genetic variation between *Megastigmus* species attacking species in the family Pinaceae (Auger-Rozenberg et al., 2005). By contrast, at the host family level, there is strong support for the separation of chalcids specialized on Pinaceae from those infesting Cupressaceae (Auger-Rozenberg et al., 2005). The genetic divergence that has occurred within *Megastigmus* suggests speciation within populations (Roux and Roques, 1996).

Megastigmus species are restricted to hosts that are phylogenetically similar. However, the degree of specificity remains unclear. There have been reports of five *Megastigmus* species that are able to switch genera following introduction to a new ecosystem (Roques

and Skrzypczynska, 2003). With the exception of these five species (*M. pinsapinis, M. schimitscheki, M. suspectus, M. amicorum,* and *M. atedius*) the establishment and spread of seed pests into a new environment largely depends on the presence of host species that are congeneric to that of the native hosts (Roques et al., 2003).

1.4.1 *Megastigmus spermotrophus*

Megastigmus spermotrophus is restricted entirely to the range of its host, *P. menziesii*. This is the most extensively studied *Megastigmus*-conifer interaction due to the economic importance of Douglas-fir. It is now widely distributed throughout Europe after its accidental introduction (Milliron, 1949). Previous reports document *M. spermotrophus* damaging over 90 % of Douglas-fir seed crop within European orchards (Rappaport et al., 1993), with an increasing loss at higher altitudes (Mailleux et al., 2008). In Europe, seed damage due to *M. spermotrophus* is much higher than its native range in North America. The reason for such elevated seed predation is the complete absence of parasites of *M. spermotrophus* in introduced areas (Rappaport and Roques, 1991).

The development of *M. spermotrophus* larvae begins after oviposition of an egg (Hussey, 1953). The larva undergoes five instar stages. By the 4th instar stage the larval body becomes arched in form. The length of the pupal phase depends largely on temperature and light conditions (Hussey, 1953). Warmer conditions promote rapid pupation, however, the longevity of emerged adults is significantly lower than larva pupating in a cooler climate. Once emerged, females can reach up to 4.1 mm in body length. Adult females have an amber-coloured body and can be distinguished from males by the

presence of a long, black ovipositor (Milliron, 1949). The adult males are significantly smaller than the females, reaching up to 3.1 mm; their body colour is also more variable, and often a dark yellow.

1.4.2 Behaviour

Adult *Megastigmus* can be found mating among the current years' needles (Hussey, 1953). It has been suggested that volatiles emitted from the host are responsible for the evolution of species-specific parasitoids within the genus (Marion-Poll and Thiéry, 1992). Prior to oviposition, *M. spermotrophus* females are attracted to a number of organic compounds released from Douglas-fir such as heptanol, terpineol, hexanol, and heptanone (Thiéry and Marion-Poll, 1998). Females search for suitable cones by moving their antennae over the surface of the scales. Once an appropriate cone is selected the female penetrates the apical third of an exposed scale with her ovipositor. She releases an egg into the central or chalazal, i.e. basal, prothallial cells of the megagametophyte (von Aderkas et al., 2005a). Although multiple eggs can be deposited into a single megagametophyte, only one reaches maturity (Milliron, 1949; von Aderkas et al., 2005a). This implies that cannibalism occurs among larvae. Female adults have been observed to be more active when the temperature is warm. Once the scales of the cone lignify, they become much harder to penetrate; females can die if they get their ovipositor stuck in a scale or if the ovipositor becomes bent and twisted out of position (Milliron, 1949). Seed coat hardening marks the end of the period of Douglas-fir cone development and is the point where ovules can no longer be attacked.

From May to June, adult females lay their eggs within the ovules of developing Douglas-fir megagametophytes. The timing of insect oviposition depends on the meteorological conditions of a given year. Oviposition extends approximately one week prior to and one week after Douglas-fir fertilization (von Aderkas et al., 2005a). Larval development begins before the megagametophyte has sexually matured, i.e., has developed eggs that are receptive. Within a day of hatching, first instar larvae migrate toward the archegonial region where they begin feeding on the highly nutritive cells surrounding the corrosion cavity. Larval development continues until the end of summer. Continual feeding occurs until the megagametophyte has been consumed in its entirety.

From late summer onwards, 5th instar larvae enter diapause, during which normal development and metabolic activity slows. Activity does not resume for many months (Denlinger, 2002). The factors controlling initiation of diapause in *Megastigmus* are unknown. However, other insects may use photoperiod to measure day length as a cue to arrest development. Several hormones may be involved in this process (Denlinger, 2002; Hahn and Denlinger, 2011). Nutrient stores of lipids, carbohydrates and amino acids are necessary to maintain catabolic and anabolic processes during diapause (Hahn and Denlinger, 2011). Many seed-infesting insects, particularly those in the order Lepidoptera, synthesize proteins prior to the onset of diapause. During diapause, many genes appear to be silenced, while a select few, particularly heat shock proteins are expressed (Denlinger, 2002).

Typically, diapause persists for 9-10 months. In *M. spermotrophus*, diapause can extend up to five years if there is prolonged cold exposure (Hussey, 1953; Turgeon et al., 1994). A chilling period is necessary to both initiate and terminate simple and prolonged diapause in *M. spermotrophus* (Roux et al., 1997). This is common for many insects that undergo diapause during their life cycle. In the silkworm, *Bombyx mori* a two-month chilling period at 5 °C is required (Moribe et al., 2001). Insect emergence experiments indicate that temperature is the most important factor, while light plays a lesser role (Roux et al., 1997).

Most commonly, *M. spermotrophus* pupates and emerges in June of the year following oviposition. The adult chews its way through the testa of the seed. The timing of emergence depends on temperature (Miller, 1916). Males tend to emerge up to one week before females (Hussey, 1953). Once emerged, adults may live up to 4 weeks.

1.4.3 Invasive species

The life cycle of *M. spermotrophus* facilitates their introduction and establishment in exotic countries due to the length of its life spent within seed. These invaders exhibit large invasive potential and have a tremendous impact on seed productivity. *Megastigmus spermotrophus* is highly invasive because of a number of adaptations (Turgeon et al., 1994):

i) Larval development accounts for the majority of the life cycle, which takes place solely inside the seed. This means that *M. spermotrophus* can easily go undetected through seed trade. In the western Palearctic region 8 of the 21

species present are exotic; their introduction is attributed to the movement of seed (Roques and Skrzypczynska, 2003). International seed trade is extending the range of these species, allowing them to overcome the natural barriers that would prevent their expansion.

ii) *Megastigmus spermotrophus* is able to undergo parthenogenesis, – a form of asexual reproduction, where eggs can develop in the absence of fertilization but they will only develop into males – which allows further spread when sex ratios are unequal or skewed (Boivin et al., 2007).

iii) *Megastigmus spermotrophus* infestation rates exceed that of expected filled seed (Niwa and Overhulser, 1992). *Megastigmus spermotrophus* can develop in unpollinated, unfertilized ovules (von Aderkas et al., 2005a). Thus, in periods of exceptionally low pollen production or poor pollination, insect reproduction is not affected.

iv) *Megastigmus* is capable of prolonged diapause for up to five years, allowing it to survive years with poor cone crops or unfavorable conditions (Turgeon et al., 1994; Roux et al., 1997).

v) There are relatively few competitors in areas that have been invaded (Roques *et al.*, 2006). *Mesopolobus spermotrophus*, a chalcid wasp from the family Pteromalidae is a known parasitoid of *M. spermotrophus*. It is widely distributed throughout the natural range of Douglas-fir, in North America, but has never been observed in France where Douglas-fir is planted on thousands of hectares and where *M. spermotrophus* has become a highly successful invader (Mailleax et al., 2008). Initial egg load of females and age-specific fecundity are important traits

influencing the displacement of competitors (Boivin et al., 2007). However, individuals that can access and exploit hosts before competitors have a significant advantage. The infestation rates of *M. spermotrophus* that have been documented are much higher outside their native range. Previous estimations were inflated because of over-estimation of empty seed (Rappaport et al., 1993). In North American seed orchards, damage from *M. spermotrophus* does not exceed 10%, while the infestation rates in exotic countries such as France and Belgium have reached 95% (Rappaport and Roques, 1991). Regardless of the presence of competition, parasitoids only eliminate about 15% of *M. spermotrophus* larvae. Therefore, parasitoids may not be an effective solution for controlling these invasive pests (Mailleux et al., 2008).

1.5 Interaction between *M. spermotrophus* and *P. mensiezii*

A striking adaption that is strongly linked to the life strategy of *M. spermotrophus* has allowed the wasp to manipulate the physiology of Douglas-fir. Since the accumulation of storage products within the megagametophyte rapidly occurs after fertilization, one would expect female wasps to select only fertilized seed. Originally, *Megastigmus* was assumed to selectively oviposit in fertilized seed (Hussey, 1955). Unfertilized ovules or even those with degenerate embryos appear to be unfavourable for larval development due to break down and death of the megagametophyte (Owens et al., 1991; Orr-Ewing, 1956). Therefore, the discovery that female wasps successfully infest not only fertilized but also unfertilized ovules (von Aderkas et al., 2005a) raised interesting questions.

Once an ovule has been fertilized, it is recognized as a seed. However, this can be confusing when referring to *M. spermotrophus* oviposition because the insect can lay eggs into both fertilized and unfertilized ovules. For simplicity sake, from this point onwards, I will refer to ovules as seeds and unfertilized ovules as unfertilized seed.

Recent studies have shown that oviposition in unfertilized seed not only prevents the abortion process that would normally occur in the megagametophyte, but also induces accumulation of storage products in a similar manner to what occurs in unifested, fertilized seed (von Aderkas et al., 2005b). The presence of larva does not appear to influence megagametophyte development; accumulation of storage products continues. Hormone analysis of parasitized seed suggests that the insects may be inducing profiles similar to that of normal megagametophyte development. Abscisic acid (ABA) was higher in infested, unpollinated megagametophytes compared to non-infested megagametophytes and closer to levels found in fertilized seed (Chiwocha et al., 2006). Thus, ABA levels may be an important factor inhibiting normal abortion. Hormone regulation during seed development is well established. For example, programmed cell death (PCD) in aleurone cells of barley is hormonally regulated (Bethke et al., 1999). Gibberellic acid induces aleurone cell death, while ABA acts as a negative regulator of PCD. During seed development of *Brassica napus*, ABA is an important promoter of nutrient deposition (Finkelstein et al., 1985).

1.6 Defence response

The ability of *M. spermotrophus* to alter hormone profiles within the megagametophyte also suggests the possibility that hormones or their analogs are able to over-ride defense. The long-term associations that have existed between conifers and invading insects may have allowed for the evolution of constitutive and induced defense responses of the host. Methyl jasmonate has been found to play a significant role in the signaling cascade of defense mechanisms in conifers and angiosperms (Hudgins et al., 2004). However, most work on conifers has focused on terpene-induced responses, such as those elicited by stem-boring insects (Hudgins et al., 2004; Miller et al., 2005).

To date, little is known about how megagametophyte tissue defends itself. When parasitized by *M. spermotrophus*, the megagametophyte of Douglas-fir does not show any physical sign of defense; instead, it behaves as though it were feeding an embryo (von Aderkas et al., 2005b). However, it seems highly unlikely that there would be no regulation of a defense response. A defense response may be coupled between sporophyte and gametophyte. In *Arabidopsis*, communication between the gametophyte and sporophyte is necessary to facilitate normal development of the embryo sac (Bencivenga et al., 2011). This is achieved through hormone regulation of haploid and diploid tissues. Plants have also been shown to develop a defense response to microbial pathogens (Ponce de Leon and Montesano, 2013). Initially, microbial molecular signals are recognized by plant receptors, which will activate an immune response in the presence of a pathogen. However, some pathogens have evolved mechanisms towards their host by releasing virulence factors that target the immune response and inhibit plant

defense. It is possible that *Megastigmus* introduces microbial pathogens into its host. Paulson et al. (2014) showed that *M. spermotrophus* has many species of bacteria that are found in larvae and adults. The introduction of pathogens via *M. spermotrophus* could potentially be involved in overriding Douglas-fir defense.

1.7 Similarities between gall formers and seed parasites

A variant of plant defense to insect attack is the situation found in galls. Through a complex interaction, gall-inducing wasps are capable of changing the phenology of plant tissue in order to obtain nutrients and protection for developing larva (Harper et al., 2004). It was previously suggested that wasps forming galls on angiosperms may have evolved from seed parasites (Ronquist and Liljeblad, 2001). Recent molecular phylogenetic reconstruction of Chalcidoidea indicates that there have been evolutionary shifts between gall-inducers and seed-feeders (Munro et al., 2011). Although seed feeders and gall-forming insects have different strategies of host invasion, the similarities in their life history traits may help in understanding the mechanisms involved in parasitic invasion. Gall development is thought to be heavily controlled by the insect as well as by hormones (Mapes and Davies, 2001). The functions of galls are thought to achieve various outcomes. It has been hypothesized that galls are either an enhanced nutritive source, or are a microenvironment to protect the galler from unfavourable conditions such as desiccation, or are a barrier against parasite attack (Stone and Schönrogge, 2003). Similar to the megagametophytes of conifers, the content within galls is highly nutritive for a developing larva. Galls are thought to have more nutrients than the surrounding tissue. Inner gall tissue contains high protein content (Schönrogge et al., 2000), in

particular, biotin carboxyl carrier proteins, which are associated with lipid production during development (Harper et al., 2004). Although the megagametophyte and gall act as a physical barrier from the external environment, galls contain high levels of phenolics, which rapidly decline in the absence of a larva (Hartley, 1998). The presence of phenolics may act as a defense against predators or may be necessary for gall-larva development. The question arises whether *Megastigmus'* interaction with Douglas-fir seed is a form of gall, in which the insect manipulates the host into providing food and shelter.

1.9 Thesis objectives

In this thesis I explore both ecological and genetic factors contributing to how *M. spermotrophus* has successfully invaded Douglas-fir seed. In the second chapter I study host site selection in Douglas-fir cones by *M. spermotrophus*. I measure morphological characteristics of Douglas-fir cones and observe the presence or absence of a competing conophyte. Competition is an important factor influencing site selection by invading insects. However, to understand resource use by invading insects it is also necessary to consider host morphology.

In the third chapter of this thesis I study internal changes at the genetic level that occur throughout megagametophyte development and which may be attributable either to plant defense or manipulation by *M. spermotrophus*. Gall-inducing insects are able to manipulate host physiology to create active sinks in host tissue, where sinks would not normally occur. Similar to gall-inducing insects, *M. spermotrophus* can redirect nutrient

reservoirs in aborting seeds. The mechanism by which this manipulation takes place is unknown. Additionally, the defense response of Douglas-fir seed to *M. spermotrophus* infestation is unclear. One of the ways to study this plant-insect interaction is to use massive-parallel sequencing. In the third chapter, I use transcriptomics to search for differentially expressed transcripts with the aim of identifying genes related to defense in infested seed compared to uninfested seed. I also identify Douglas-fir transcripts that have been manipulated by *M. spermotrophus*. Identifying differentially expressed transcripts between infested and uninfested samples will allow me to determine parasite influence at the genetic level. This will provide useful information into understanding defense investments of haploid tissue as well as providing insights into host manipulation by parasites.

CHAPTER 2. SITE SELECTION FOR OVIPOSITION INTO DOUGLAS FIR CONES BY *MEGASTIGMUS SPERMOTROPHUS*

2.1 Introduction

Intercontinental movement of Douglas-fir (*Pseudotsuga menziesii*) seed from its native range in North America has resulted in the establishment of an invasive, parasitic seed insect, *M. spermotrophus*, in Europe. The range of this chalcid wasp follows that of its host, expanding throughout the regions where Douglas-fir has been introduced (Roques and Skrzypczynska, 2003). Previous reports have estimated that *M. spermotrophus* destroys up to 95% of seed crop in France, compared to less than 10% of seed crop loss in its native range in North America (Rappaport and Roques, 1991). The high infestation rates in Europe cause major economic losses for seed orchard production. These differences in infestation rates in the native and introduced regions are not well understood, but may be closely associated to host-site selection pressures of females to maximize fitness.

Plant-insect relationships have been extensively studied. However, many questions about their associations remain unanswered, particularly those regarding host site selection by ovipositing females. Survivorship of offspring is directly affected by the quality of a site that an egg is laid in (Thompson, 1987). For conophyte insects whose larval offspring are capable of moving freely within a cone and are not bound within specific structures, i.e. a seed, site selection for oviposition is less critical for larval success (Price, 1977; Quiring and McNeil, 1987). In the case of the chalcid wasp, *M. spermotrophus*, which

spends most of its life entirely within the seed, there is selection pressure for females to oviposit into the right location, a nutrient rich seed (Jaenike, 1978). This strategy where females choose oviposition sites based on expected larval performance is also recognized as the "mother knows best" strategy. Female site preference should favour high quality seeds, that will provide the most nutrition (Jaenike, 1978), i.e. large, fertilized seeds, that contains more storage reserves than unfertilized seeds (Fidgen et al., 1998; von Aderkas et al., 2005b). *Megastigmus spermotrophus* females are either poor mothers or they know something "better than best," because they oviposit in fertilized seed and unexpectedly in unfertilized seed. This begs the question of what else is involved in site selection. Are host condition, host defense, competition, host chemistry, or genetic covariance involved in female assessment of oviposition sites (Courtney and Kibota, 1990; Jaenike, 1990, Thompson and Pellmyr, 1991)?

Thompson (1987) discusses a number of general selection pressure hypotheses to explain the relationship between preference and performance in phytophagous insects. These hypotheses have been explored individually in numerous studies of plant feeding insects; however, very few apply to the relationships that conophytes have with their hosts. For the purpose of this introduction, I will consider two of the hypotheses that may apply to *Megastigmus.* The first is the enemy-free space hypothesis and the other is the patch dynamics hypothesis.

i) *Enemy-free space hypothesis*: The performance of larvae may be affected by the presence of predators, or other competitors (Thompson, 1987).

Natural selection may favour enemy/competitor-free space to avoid overcrowding and over-exploitation of resources (Prokopy et al., 1984). In this scenario, insects preferentially choose sites where competition is minimal to maximize offspring success (Thompson, 1987). Since sedentary insects, such as *M. spermotrophus,* can only utilize a single seed, inter- and intraspecific competition are likely to be significant predictors of female selection (Birch, 1957). Interspecific competition is common amongst specialist cone feeders. For instance, Douglas-fir has numerous species of cone pests all competing for seed (Ruth, 1980). An example of a Douglas-fir cone pest is *C. oregonensis* Foote, a cone-gall midge whose eggs are laid near the base of Douglas-fir cone scales. Larvae of *C. oregonensis* feed on scale tissue and subsequently a gall is formed around the larvae, fusing the scale to its adjacent developing seed (Ruth, 1980). Oviposition by *C. oregonensis* occurs approximately one month prior to *M. spermotrophus* infestation. Assuming that all seed within a cone is equally available to all Douglas-fir seed predators, resource use by *M. spermotrophus* would be displaced by its numerous competitors. This was demonstrated by Rappaport and Roques (1991), who showed that *M. spermotrophus* resource utilization within a cone changes in areas where natural predators and competitors are absent. In introduced areas, with no competitors, *M. spermotrophus* occupies the seed located in the central portion of Douglas-fir cones, whereas in its native range, in North America, *M. spermotrophus* occupies more seed in the apical and basal ends of a cone.

Intraspecific competition has been observed more frequently in the literature and is specifically reported for *M. spermotrophus*, where multiple larvae have been oviposited

into a single ovule (von Aderkas et al., 2005a). Intraspecific competition may occur either by interference or by exploitation (Birch, 1957). Interference competition occurs when conspecifics share a single resource and as a result competitor mortality increases. A large cost occurs with this form of competition; surviving larvae have reduced fitness because the amount of resource available during late development is depleted by conspecifics (Birch, 1957; Quiring and McNeil, 1983). Competition by interference is likely to occur when quality resources are limited (Birch, 1957). Exploitative competition occurs when conspecifics compete for the same resource but avoid each other to increase their own fitness (Birch, 1957). This is the most common form of competition observed in *Megastigmus*.

The mechanism as to how insects recognize unoccupied space is not completely understood. A number of cues related to vision and olfactory response have been proposed, however, these cues are thought to be variable among species and between sexes (Turgeon et al., 1994). Several studies have indicated that pheromone detection is the predominant cue for host, competitor and predator recognition (Turgeon et al., 1994; Miller and Border, 1984). However, most research on pheromone marking has been demonstrated in phytophagous insects (Nufio and Papaj, 2001). To deter other gravid females, the leaf miner, *Agromyza frontella* leaves pheromone markers at oviposition sites (McNeil and Quiring, 1983). Using the sensillae in the antennae, foraging *A. frontella* females recognize exploited hosts by the presence of marking pheromones from conspecifics (Quiring and McNeil, 1987). Although the use of pheromones for marking oviposition sites has not been demonstrated in *M. spermotrophus*, these insects can detect

volatiles emitted from their host, Douglas-fir (Marion-Poll and Thiéry, 1992). It is possible that *M. spermotrophus* detect pheromones left from a previous ovipositing female to identify exploited and unexploited seed by conspecifics. The ability of *M. spermotrophus* to detect pheromones has however, not yet been tested.

ii) *Patch dynamics hypothesis*: Within different geographical sites, host preference may vary due to the genetic differences and abundance of an individual host (Thompson, 1987).

Seed orchards have many genetically distinct superior clones of a given species both to provide genetic gain and to maintain genetic diversity (Krugman, 1986; Libby, 1986). These have become important research sites because of the ability to control for genetic differences and to test susceptibility as a heritable trait. Female site selection preference and larval performance is greatly influenced by the timing of seed development, which is one of the differences observed among genotypes. Some genotypes are more susceptible than others to seed chalcids (Roques, 1981; Blatt and Borden, 1998). Susceptibility to insect invasion may be heritable (Schowalter et al., 1986; Schowalter and Haverty, 1989).

Noticeable patch differences of insect attack can occur between genetically identical hosts at small scales, such as in a single orchard (Schowalter and Haverty, 1989). Patch differences also occur on a larger scale between geographically different sites. For example, cone damage of *Pinus cembra* is greater at lower elevations and among trees that are isolated, compared to trees at higher elevation and found within stands (Dormont

and Roques, 1999). Both *M. spermotrophus* and *C. oregonensis* show changes in abundance across their longitudinal range (Schowalter et al., 1985). In addition, significant differences between yearly crop sizes will influence the distribution of attack rate (Caron and Powell, 1988; Roques, 1986; McClure et al., 1998). Attack rate is more evenly distributed in years of a light cone crop relative to the number of emerging chalcids, compared to mast years, when heavy cone crops are attacked patchwise (Roques, 1986).

2.1.1 Objectives

The aim of this chapter is to describe patterns of *M. spermotrophus* resource use related to cone morphology of Douglas-fir. There has been extensive research on the interaction between *M. spermotrophus* and Douglas-fir, including host preference; however, relatively little is known about the host-related factors involved in oviposition site selection by the chalcid. Rappaport and Roques (1991) proposed that chalcid site selection is largely influenced by the presence or absence of competing conophyte insects. This is the most widely accepted hypothesis to explain how *M. spermotrophus* occupies the entire seed niche of Douglas-fir in Europe, where competitors are absent. The hypothesis of "enemy-free space" may not be the only factor influencing female choice of *Megastigmus* oviposition. Other factors, such as host development and physiology, crop size, accessibility, and range may be contributing to host suitability (Courtney and Kiobta, 1990).

For plant-insect relationships that involve a specific obligate plant host, it is important to understand the spatial characteristics related to the morphology of the host that make it an optimal habitat for insect success. In this study, I examine the presence of a competing insect and multiple morphological characteristics of Douglas-fir cones including seed size, and scale thickness in relation to the position of occupied seed to improve our understanding of host site selection.

2.2 Materials and methods

2.2.1 Seed samples

Five Douglas-fir trees known to be highly infested with *Megastigmus* in previous years were selected from the University of Victoria campus (48°27'42.90" N, 123°18'37.50" W). Six cones were randomly selected from each tree at the end of August, just prior to cone maturity. Cones were dissected scale-by-scale beginning at the base and working towards the tip. The following aspects were measured or noted: (*i*) total scale number, (*ii*) total potential seed, (*iii*) seed location, (*iv*) *Megastigmus* infested seed, (*v*) seed fused to galls formed by *Contarinia*, (*vi*) filled seed, and (*vii*) seed length. *C. oregonensis* infestation was scored by the presence of distinct galls on the dissected fused scale/seed complexes. *M. spermotrophus* infestation was determined by X-raying seed using a Faxitron N 4355A at the Tree Seed Centre, Surrey, British Columbia. Seeds were exposed to 20 kV for 12 seconds. Using the X-ray images, the seeds were categorized as filled, empty, or infested by *M. spermotrophus* (Figure 2). Seed length was measured using ImageJ software (Schneider *et al.*, 2012). Seed location within a cone was determined according to scale order.

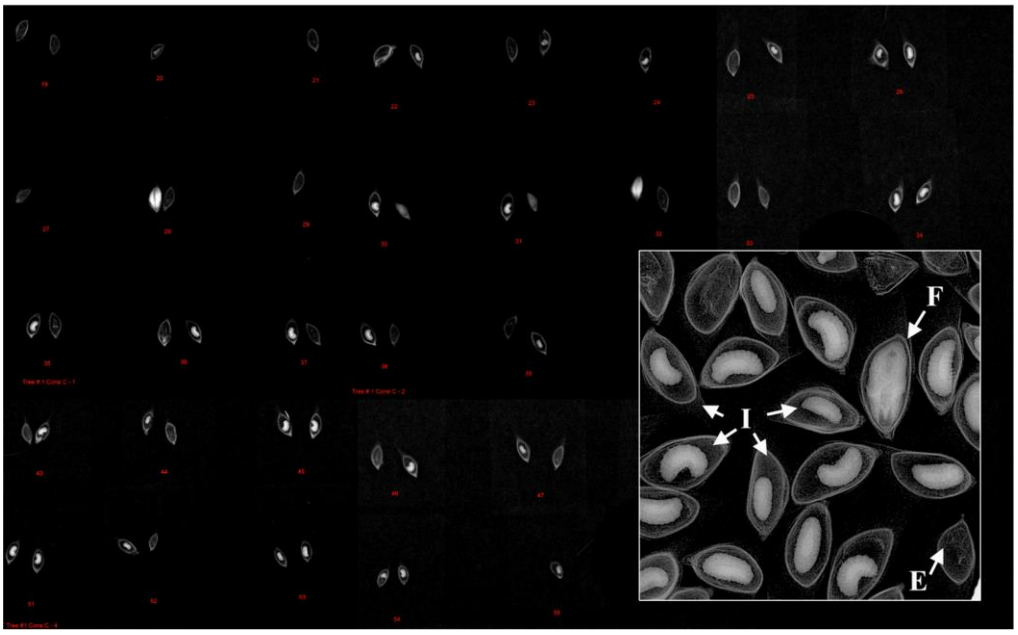


Figure 2. Radiograph of *Pseudotsuga menziesii* seed dissected from a single cone from the basal to proximal end. Radiographs were imaged using a Faxitron N 4355A. Seeds were identified as either infested with *Megastigmus spermotrophus* (I), filled seed with a Douglas-fir embryo (F), or empty, aborted seed (E).

The most basal scale represented position one. These positions were standardized across all cones by turning the scale number into a proportion of the total number of scales per cone.

2.2.2 Seed data analysis

Megastigmus infestation relative to seed position within a cone was investigated by fitting a generalized linear mixed-effects model (GLMM) with quadratic regression and a binomial distribution using lme4 (Bates et al., 2014) package in R version 3.0.2 (R Core Team, 2013). The outcome variable of the model was whether an individual seed was infested or not infested by *M. spermotrophus*. The model included the following variables: seed position, seed length, the interaction between seed position and seed length, presence/absence of *C. oregonensis* infested seed and tree and cone as random variables. Both random and fixed effects had random intercepts.

Collinearity was observed between seed length and seed position (r^2 = 0.976, df= 11, P = $5.37e^{-10}$). To avoid effects of collinearity, seed length and the interaction between seed length and seed position were removed from the model. The validity of the model was verified using a likelihood ratio test with Akaike Information Criterion (AIC), in which the null model with only random effects was compared to alternative models with fixed and random effects. The model with the lowest AIC value was kept as the final model. Results were considered significant at the $\alpha = 0.05$ level. The relationship between percent of infested seed and seed position was determined using the coefficient of determination.

2.2.3 Scale Samples

Six Douglas-fir trees located around the Forest Biology Compound at the University of Victoria Campus were selected for sampling. The trees selected were not all used in the seed study due to low cone production during the scale-study sampling season. However, these trees were known to be highly infested with *M. spermotrophus* in previous years. Five cones from each tree were randomly selected early in June, which was when *Megastigmus* were seen ovipositing in Douglas-fir cones. The length and diameter of each cone was determined and the cones were dissected scale by scale to measure scale thickness in relation to scale position along the cone. The total number of scales was determined. Scale thickness was only measured on scales with seeds. Measurements were taken directly above and between the two seeds with a calliper.

2.2.4 Scale data analysis

To investigate the relationship between scale thickness and scale position along a cone, a linear mixed effects model (LMM) with quadratic regression was fitted using lme4 (Bates et al., 2014) and lmerTest packages in R, version 3.0.2 (R Core Team, 2013). The outcome variable of the model was scale thickness. The model included scale position, cone diameter, cone length, and the interaction between cone diameter and cone length, with tree and cones as random variables. Both random and fixed effects had random intercepts.

Collinearity was observed between cone length and cone diameter (r = -0.690) (r^2 = 0.424, df = 28, $P = 9.836e^{-5}$). To avoid effects of collinearity, cone length and the

interaction between cone length and cone diameter were removed from the model. The validity of the model was verified using a likelihood ratio test with Akaike Information Criterion (AIC), in which the null model with only random effects was compared to models with fixed and random effects. The model with the lowest AIC value was kept as the final model. Results were considered significant at the $\alpha = 0.05$ level. The relationship between cone length and diameter was determined using the coefficient of determination.

2.3 Results

2.3.1 Seed length and oviposition site selection

Douglas-fir seed size and seed position were highly correlated ($r^2 = 0.976$, df= 11, $P = 5.37e^{-10}$). Seed size changed along the length of a cone; larger seeds were found in the central portion and gradually became smaller towards the distal and proximal ends (Figure 3). The GLMM revealed that both seed position and *C. oregonensis* presence were significant predictors of *M. spermotrophus* infestation (Table 1). The estimated probability of a seed being infested by *Megastigmus* was lowest in seeds located at approximately the 0.4^{th} position along a cone (Figure 4). As seed position moved away from the 0.4^{th} position, towards 0 or 1, the probability of infestation increased. There was a high correlation between the amount of infested seed and seed position ($r^2 = 0.678$, df =11, $P = 0.002$). The number of chalcid-infested seed was greatest in the apical and basal ends of a cone where seeds were smaller (Figure 5A).

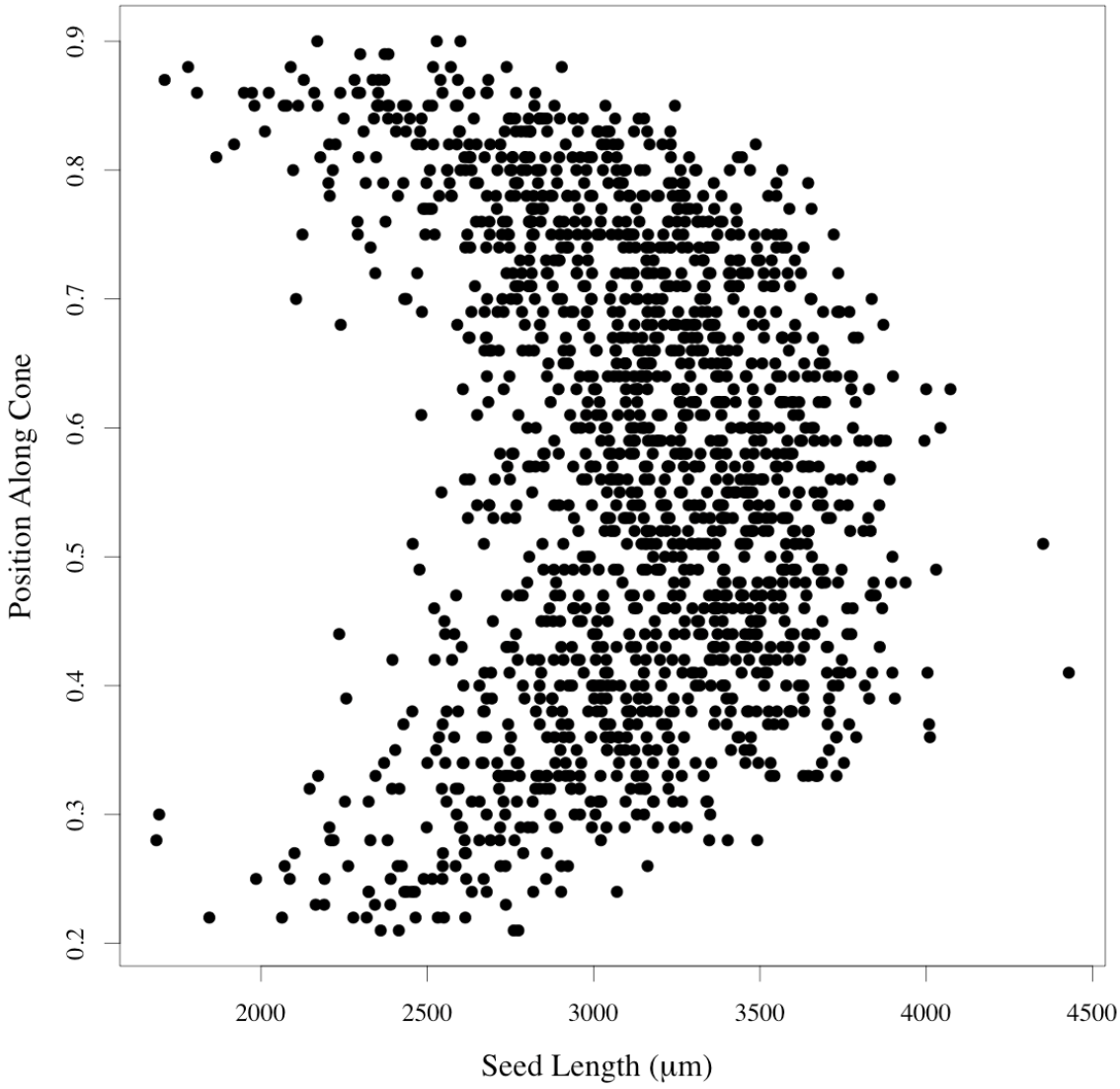


Figure 3. Scatterplot of seed length (μm) at every position along a cone from the basal to apical end, for all seed collected.

Table 1. Fixed-effects of the Generalized Linear Mixed-Effects Model fit with quadratic regression and a binomial distribution, explaining *M. spermotrophus* infestation in Douglas-fir seed based on seed position and *C. oregonensis* presence per cone.

	Model: *Megastigmus* infested ~ Position + Position^2 + *Contarinia* + Tree/Cone		
Fixed Effects	**Estimate**	**Std. Error**	**P**
Intercept	-0.59	0.84	0.48
Position (linear)	-4.97	2.49	0.046*
Position (quadratic)	5.67	2.18	9.26e-3*
Contarinia	-1.18	0.26	7.86e-6*
Deviance		1525.1	
AIC		1537	

* indicates significant predictor variables

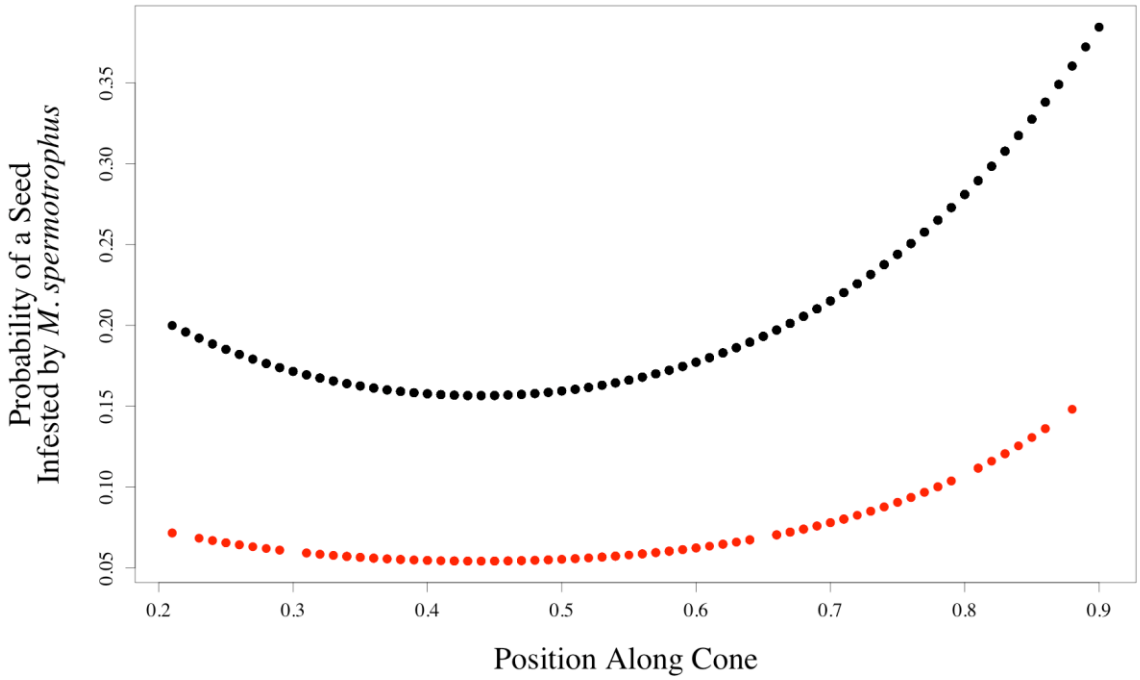


Figure 4. Scatterplot of the Generalized Linear Mixed-Effects Model fit with quadratic regression and a binomial distribution at a population level, illustrating the probability of a seed infested by *M. spermotrophus* at every position along a cone in seed that is fused to a *C. oregonensis* gall (red) and in seed where galls from *C. oregonensis* are absent (black).

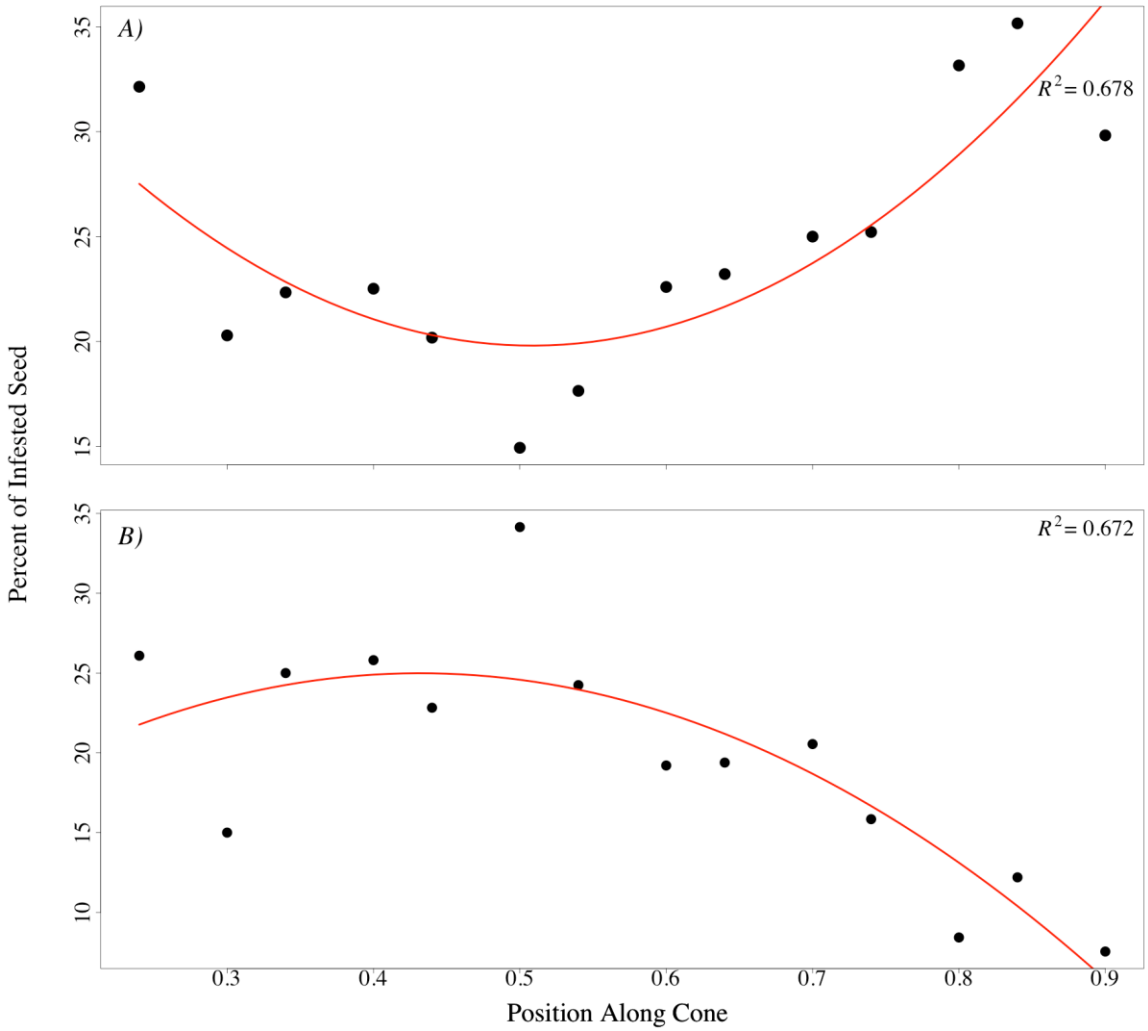


Figure 5. Scatterplot of the number of A) *M. spermotrophus* infested seed and B) *C. oregonensis* galled seed over the total number of seed for each position along a cone. Infestation numbers were summed within 14 bins of proportional distance across all cones.

Megastigmus females more frequently oviposited into smaller seed. However, for any given position along a cone, there was no significant difference between the size of seed that were infested by *M. spermotrophus* and seed that was not infested (Figure 6).

Contarinia gall presence on a seed influenced the probability of infestation by *Megastigmus*. Seeds fused to galls had a lower probability of being infested by *Megastigmus* compared to seeds without galls (Figure 4). The abundance of *Contarinia* within a cone was also correlated to seed position (r^2 = 0.672, df =11, P = 0.002). *Contarinia* distribution within a cone was inversely related to *Megastigmus* infestation; there was a greater number of *Contarinia* in the central portion of the cone, with very few in the apex (Figure 5B).

Megastigmus were infesting the extreme ends of cones where there were no full, healthy seed produced (Figure 7), whereas *Contarinia* were not found in these areas. There were very few full seeds where *Contarinia* had formed a gall between the seed and the scale (Figure 7). The majority of *Contarinia* infected seeds were empty.

Both tree and cone effects accounted for variance in the model (Table 2). The distribution and number of *Megastigmus* filled seed differed between cones, within and between trees (Figure 8). Infestation was not consistent among all cones from a single tree. Some cones sampled from trees 2 and 5 had either only a single infested seed or none at all. Overall, cones collected from trees 2 and 5 had fewer chalcid-infested seeds compared to cones from trees 1, 3, and 4 (Table 3).

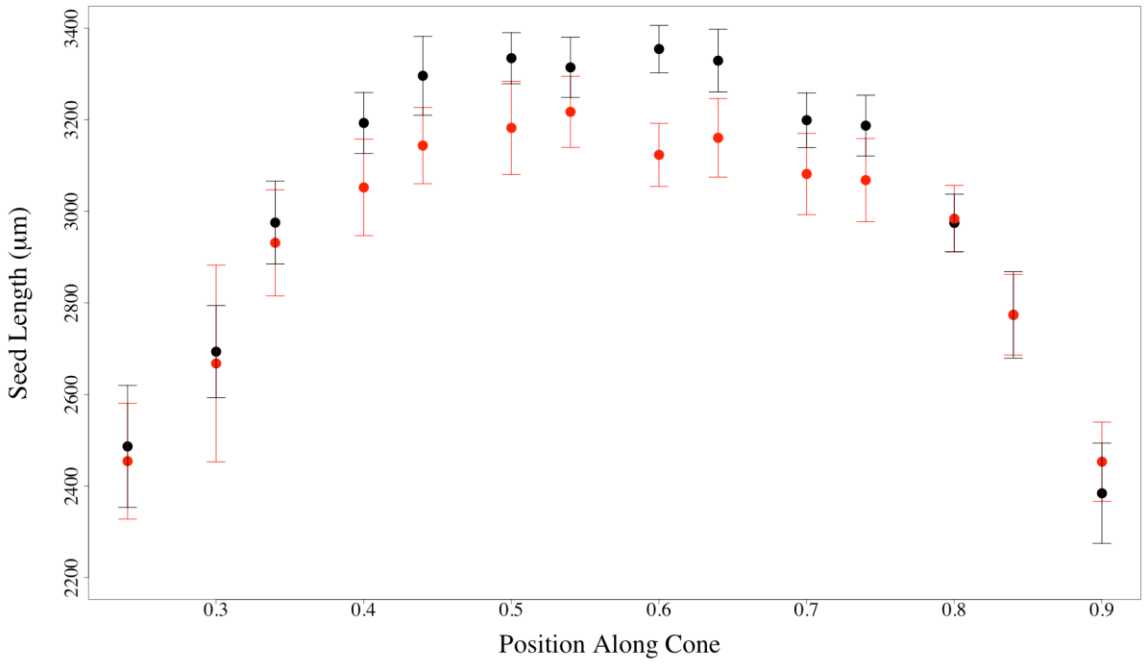


Figure 6. Scatterplot of the average seed size (µm) along a cone, comparing *M. spermotrophus* infested seed (red) and uninfested seed (black). Mean seed lengths and standard error were determined by seed position, which was divided into 14 bins of proportional distance across all cones.

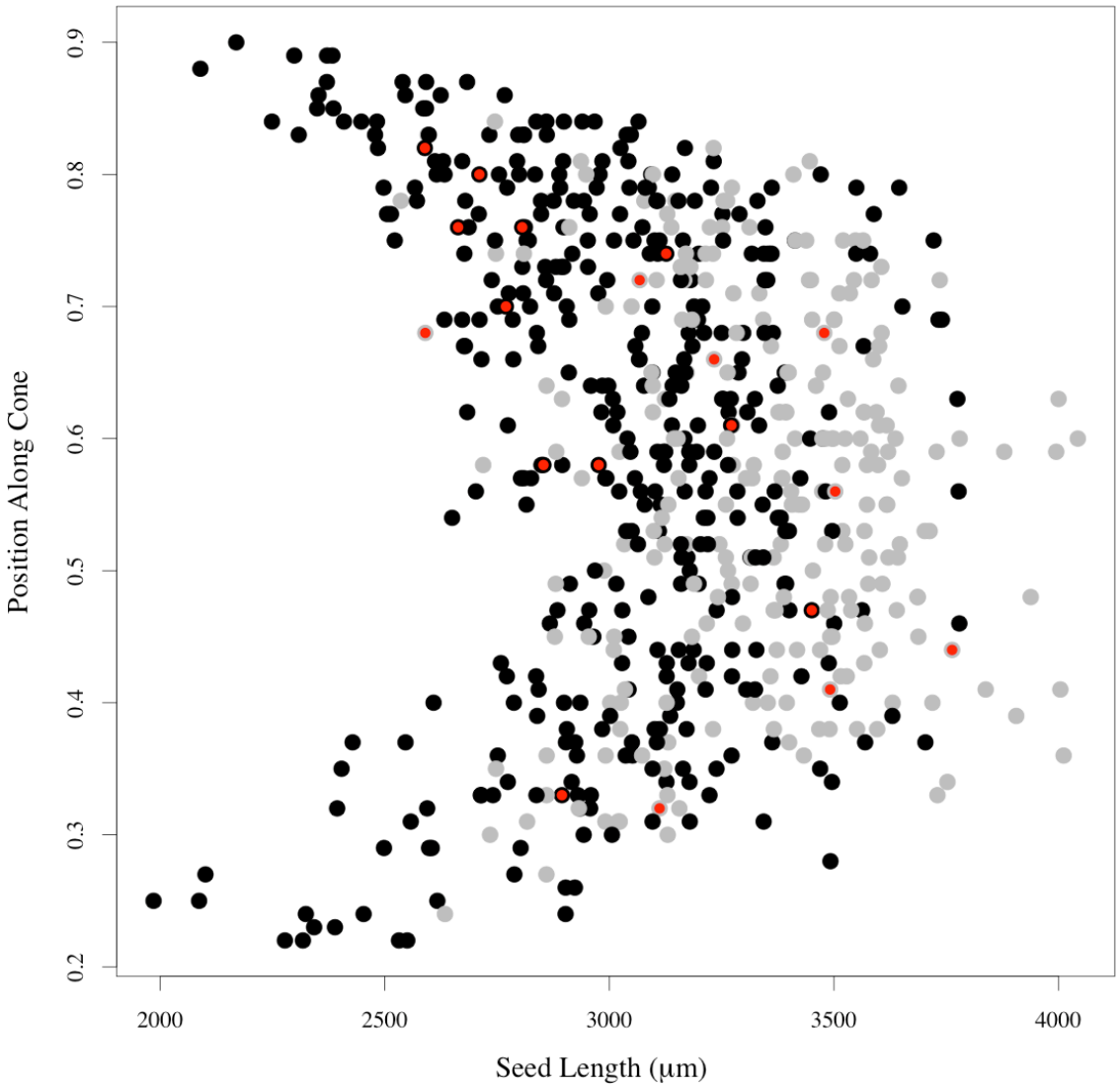


Figure 7. Scatterplot of seed length (μm) and seed position along a cone from the basal to apical end for all seeds collected, including: full seed, (grey); *M. spermotrophus* infested seed (black); and seed fused to *C. oregonensis* galls (red). All seeds lacking an embryo or uninfested by *M. spermotrophus* were removed.

Table 2. Random effects of the Generalized Linear Mixed-Effects Model fit with a binomial distribution, showing the variance accounted for by cone and tree.

	Model: *Megastigmus* infested ~ Position + Position^2 + *Contarinia* + Tree/Cone	
Random Effects	**Variance**	**Std. Dev**
Cone:Tree	0.48	0.69
Tree	1.20	1.10
Number of observations: 1634 Groups: Cone: 30; Tree: 5		

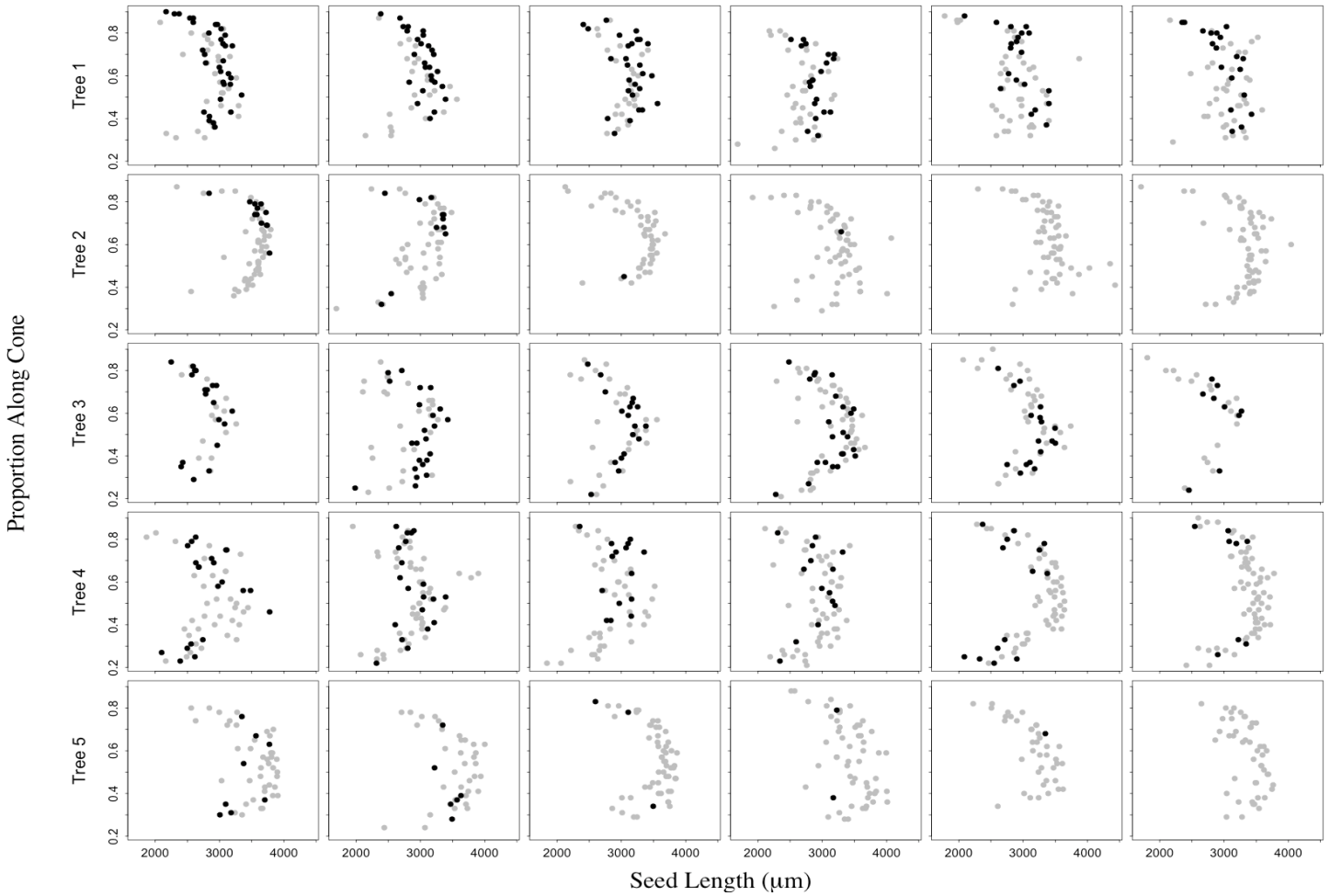


Figure 8. Scatterplot of seed length (μm) along the length of a cone from the basal to apical end, showing the distribution of *Megastigmus* infested seed (black points) compared to empty and filled seed (grey). Each panel represents an individual cone.

Table 3. Percentage of seed with absolute values and prevalence of *Megastigmus* and *Contarinia* larvae among sampled trees

	Percent of Seed Per Tree				
	1	**2**	**3**	**4**	**5**
Unfilled seed *	95% (319)	78% (267)	90% (246)	76% (297)	90% (260)
Undamaged/filled seed	4% (15)	22% (75)	10% (27)	22% (87)	9% (26)
Damage by					
Megastigmus	44% (147)	7% (25)	40% (109)	23% (92)	7% (20)
Contarinia	13% (45)	15% (50)	1% (3)	19% (75)	18% (52)
Megastigmus + *Contarinia*	3% (10)	0% (1)	0% (0)	2% (6)	1% (3)
Total Number of Seed	335	344	273	392	290

***Seed without an embryo (may be empty or may include a larvae)**

The distribution of infested seed within cones from trees 2 and 5 was localized to the apical and basal ends, whereas the distribution was more uniform among cones from trees 1, 3, and 4 (Figure 8). All trees sampled contained over 75 % of unfilled seed of which approximately 25 – 60 % was a result of infestation from either *M. spermotrophus* or *C. oregonensis* (Table 3). Relatively few seeds per tree (<3%) were infested by both *M. spermotrophus* and *C. oregonensis*.

2.3.2 Scale thickness

The LMM indicated that proportional distance along a cone and cone diameter were significant predicators of scale thickness (Table 4). Scale thickness was strongly correlated with scale position (r^2 = 0.963, df = 11, P = 0.002). Within a cone, scales were the thickest at the 0.5^{th} position along a cone and became thinner towards the basal and apical ends (Figure 9). The random intercept LMM accounted for baseline-differences in scale thickness among all cones sampled. Since the proportional distance did not differ between cones, random slopes were not necessary.

Table 4. Fixed effects of the Linear Mixed-Effects Model fit with quadratic regression and a binomial distribution, explaining scale thickness in Douglas-fir trees based on scale position within a cone and cone diameter

	Model: Scale thickness ~ Position + Position^2 + Cone diameter + Tree/Cone		
Fixed Effects	**Estimate**	**Std. Error**	**P**
Intercept	-0.64	0.21	0.01
Position	4.94	0.11	<2e-16*
Position^2	-4.99	0.11	<2e-16*
Cone diameter	0.07	0.01	1.09e-08 *
Deviance		-809.5	
AIC		-795.5	

* indicates significant predictor variables

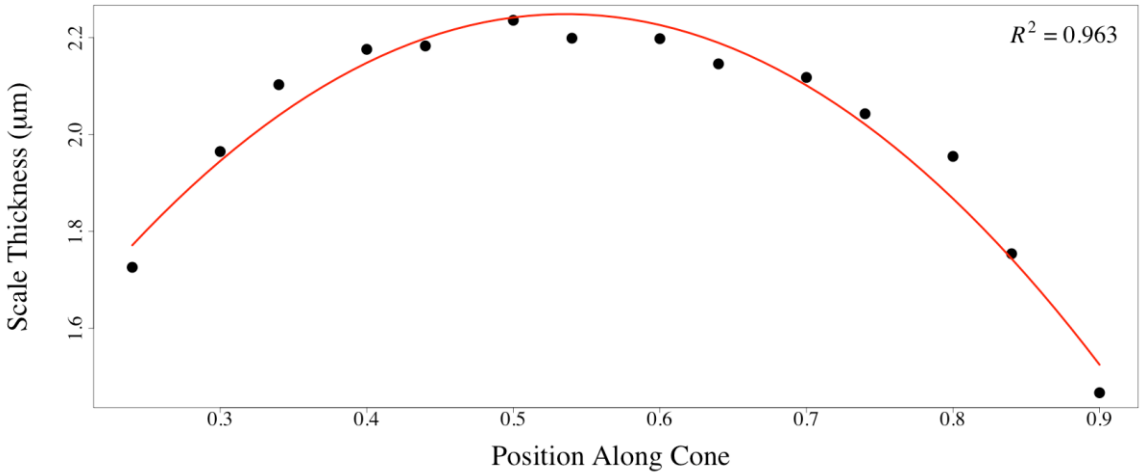


Figure 9. Mean scale thickness across all sampled scales along the proportional length of a Douglas-fir cone from the basal to apical end. Average values of scale thickness were determined across 14 bins of cone position.

2.4 Discussion

2.4.1 Patterns of *Megastigmus* resource allocation in Douglas-fir cones

My study shows that *M. spermotrophus* infested seeds in both the proximal and distal ends of a cone more frequently than in the central region, with a slight preference towards the distal tip. This finding is consistent with Lessman (1974) who reports attack rates to be the greatest at the apex of Douglas-fir cones. Rappaport and Roques (1991) also report this bimodal distribution of cone attack in Douglas-fir cones found within their native range.

The distribution of *Megastigmus* infested seed within a cone parallels the quadratic relationship between seed size and seed position. *Megastigmus* more frequently occupy the smallest seed within a cone. This finding is rather surprising because *Megastigmus* larval movement is restricted to an individual seed. It would seem more favourable for adult females to oviposit into high quality hosts with large seed to increase the success of offspring development (Thompson 1987). Other conophyte species have been shown to preferentially select hosts based on host size. For example, females of the spruce cone worms, *Strobilomyia neanthracina* and *S. appalachensis* laid more eggs on longer cones than shorter ones in developing trees of *Picea glauca* and *P. mariana*, respectively (Fidgen et al., 1998). This resulted in developing larvae consuming a higher volume of large seed, ultimately maximizing female fitness.

In contrast to *Strobilomyia* species, *Megastigmus* have been observed to heavily parasitize seed found within narrow, short cones, compared to cones that are longer and

wider (Lessman, 1974). Rapport and Roques (1991) tested this hypothesis, looking at infestation rates between large and small cones. Cone size was a significant factor in attack rate; smaller cones were more heavily infested. Additionally, *Megastigmus* seed infestation was greater in the apical and basal regions of the cone (Rapport and Roques, 1991). In short, *Megastigmus* attack both smaller seed and smaller cones. Benefits associated with infesting smaller seed must be greater than the benefits of providing a large food source for developing larvae.

Selection of small seed by *Megastigmus* raises the question of whether or not *Megastigmus* is manipulating seed size after oviposition. For *M. spermotrophus* there is a short, two-week window for females to lay eggs that fluctuates around the time of Douglas-fir fertilization. At this time Douglas-fir seeds have reached their maximum volume. Since there is no significant difference between size of infested and non-infested seed at approximately the same location in the cone (Figure 6), *Megastigmus* are likely not influencing seed size. Additionally, the relationship between seed size and seed position remains quadratic, with the smallest seed located in the apical and basal ends of a cone, regardless of infestation. If *Megastigmus* were manipulating seed size, this quadratic relationship between seed size and seed position would not exist in infested cones. Seed length is highly correlated with seed position. *Megastigmus* may be choosing position, rather than size. In other words, seed length may just be a by-product of selection.

In addition to the finding that *Megastigmus* infest small seed at the tips of the cone, this is the first study to report *Megastigmus* exploiting Douglas-fir seeds at the extreme ends, where seeds rarely completely develop. *M. spermotrophus* oviposits into regions where there are no full seed at cone maturity. In all *Pinaceae*, the most apical and basal ovuliferous scales of female cones are sterile. These ovuliferous scales either i) do not produce ovules ii) produce ovules that never fully develop or iii) produce ovules that will never be pollinated (Owens and Blake, 1986). On average, three-quarters of *Pinus* cones are made up of ovuliferous scales without seed (Sarvas ,1962; Palowski, 2000). This phenomenon is not nearly as extreme in Douglas-fir, but it also has approximately 10-15 underdeveloped ovules at either of the cones extremities (Owens and Blake 1986).

Why do conifers produce seeds that will never develop? One hypothesis is that the abundance of sterile seed acts as a defense against insect herbivory by reducing the loss of viable seed (Traveset,1993). In *Dipterocarpus*, it is hypothesized that selfed seeds are retained on the tree during predation to mitigate predation losses of crossed seed (Ghazoul and Satake, 2009). Thus, the fact that *M. spermotrophus* does not discriminate between fertilized and unfertilized seed may be advantageous for Douglas-fir (von Aderkas et al., 2005a). An alternative hypothesis is based on cone construction. The small ovuliferous scales at the most apical and basal ends of the cone may be insufficient for ovules to completely develop (von Aderkas, pers. comm.).

Why does *M. spermotrophus* occupy the smallest seeds at the tips of a cone and additionally sterile seed that will never fully develop? Oviposition by *M. spermotrophus*

appears to contradict the theory that "Mother knows best". There are other examples of insects that do not choose optimal sites for their offspring. For example, the gall midge, *Dasineura marginemtorquens* does not differentiate between resistant and non-resistant genotypes of the willow, *Salix viminalis*; on resistant genotypes, larvae cannot initiate gall formation. Consequently, they die shortly after hatching (Larsson et al., 1995). Although small seed appears marginally suitable for *Megastigmus* larval development because it offers a reduced food source, small seed may confer an adaptive advantage (Björkmann et al., 1997; Gripenberg et al., 2010).

Results from this study suggest that *Megastigmus* adult females do not choose oviposition sites based on food availability for the offspring, but rather by accessibility. Infestation of Douglas-fir cones is greatest in the distal and proximal ends, where scales are the thinnest. The length of the ovipositor is a limiting factor to female success and will thus influence female selection. For example, ovipositor length in species of fig wasps found in the genus *Apocrypta* differ depending on the thickness of the fig wall in which each species deposits its eggs; as a result, oviposition behaviour and selection strategy differs among species (Zhen et al., 2005). Parasitism of a shoot-galling sawfly, *Euura lasiolepis* by *Lathrostizus euurae* wasps was lower in larger galls where larvae were out of reach of the ovipositor (Craig et al., 1990). The ovipositor of *M. spermotrophus* may be limited to areas of Douglas-fir cones where seeds are more exposed as a result of the ovuliferous scales being thinner. The seeds in the middle region may be more difficult to access. Not only are the scales the thickest in the central portion of the cone, the ovipositor may have to penetrate through multiple scales to access the seed in this region.

The angle in which the female's ovipositor penetrates through the scales is not known. Therefore, it is difficult to know with any precision, which seed, judging from the entry site, has been targeted by the ovipositor. Lessman (1974) showed that *Megastigmus* prefer to access the seed by penetrating along the margins of the ovuliferous scales. Structural limitations of the ovipositor have likely influenced localized site preferences of where the female inserts into the ovuliferous scale.

2.4.2 Response to competition: Does *Megastigmus* search for enemy-free space?

Given the complexities revealed by my study of ovipositional site selectivity, considering conophyte behaviour becomes a bigger challenge than I had previously thought. The concept of enemy-free space applies, in part, to ovipositional preferences. The presence of *C. oregonensis* is a significant factor contributing to seed selection within a cone in which *Megastigmus* will oviposit. *Megastigmus* rarely oviposit into seed on galled scales. This finding is consistent with Schowalter et al. (1985). Although, *Contarinia* does not directly compete for seed, gall formation fuses the seed to its adjacent scale, thereby allowing the larvae to obtain nutrients from the seed (Miller, 1982). *Megastigmus* may likely search for enemy-free space, to avoid competition. Many species use strategies to avoid over-exploitation of resources by assessing both conspecifics and competitors (Prokopy et al., 1984). Interference intraspecific competition occurs in seed-feeding insects. For example, if multiple *Megastigmus* larvae exist within a single seed, only one will survive (Milliron, 1949; von Aderkas et al., 2005a). In some insects, intraspecific 'enemy' avoidance may even be an adaptive trait. Larvae of the spruce cone worms, *Strobilomyia laricis* and *S. viaria* had 40 % greater

chance of survival in cones that were not occupied by a conspecific (McClure et al., 1998). Phytophagous insects have even demonstrated adaption to avoiding predators. *Oleria onega* adult female butterflies preferentially lay their eggs on plant surfaces other than their usual host to avoid predation (De-silva et al., 2011).

Interspecific competition from other cone feeders influences female choice for oviposition sites. Support for the idea comes from *Megastigmus* avoidance of *Contarinia* galled seeds. This is consistent with the hypothesis proposed by Rappaport and Roques (1991), stating that *Megastigmus* oviposition may largely be determined by the presence of other competing seed feeders. It appears that *Megastigmus* and *Contarinia* partition Douglas-fir resources differently within a cone. *Contarinia* gall formation is heaviest in the central portion of the cone and lightest in the apex, whereas *Megastigmus* infestation is the opposite (Figure 4). However, *Contarinia* is not the only insect competing for resources within Douglas-fir cones. Lepidopteran and other dipteran species are important competitors that may be influencing *M. spermotrophus* preference for seeds at the apical and basal ends of a cone (Bates et al., 2000; Schowalter et al., 1985). Species that feed in the centre of the cone may have the largest impact. For example, *Strobilomyia* cone maggots feed mainly in the central region of *Picea* cones, where the seed is heaviest (Fidgen et al., 1998). This would push other competitors to the extremities.

Of the two kinds of competition, interspecific and intraspecific, there is little support for the latter. Intraspecific competition is not a big driver of host site selection in

Megastigmus. For other insects it can be important. A recent study showed that females of the green-veined white butterfly, *Pieris napi*, are attracted to oviposition sites already holding conspecifics (Raitanen et al., 2014). This would increase interference competition, and decrease fitness of surviving larvae (Birch, 1957; Quiring and McNeil, 1983). In cases such as *P. napi*, the cost of interference competition must have benefits that outweigh the cost of interspecific competition. Since not all insects use enemy avoidance strategies when selecting oviposition sites, other selection pressures must be considered.

2.4.3 Response to patch dynamics: Does host distribution contribute to female preference?

In this study, tree and cone differences are important variables explaining variation in *Megastigmus* infestation. Distribution of *Megastigmus* within a cone is not consistent between trees and between cones. This means that there is variability among individual trees. *Megastigmus* infest some trees more readily than others. Blatt and Borden (1998) showed that *Megastigmus* adult females have clonal preferences. Tree genotype was not controlled for in our study, even though this may be account for differences in attack rates. Even within a single tree the distribution of infestation differs between highly infested and lightly infested cones; highly infested cones were uniformly attacked, whereas lightly infested cones were attacked more at the proximal and distal ends. Patch dynamics may be important factors contributing to variation in attack rate. Douglas-fir seed loss by *C. oregonensis* and *M. spermotrophus* changes significantly between tree positions in an orchard (Schowalter and Haverty, 1989). Patch differences are thought to

also occur on individual trees. Cone maturity increases towards the tops of trees and cones found on the southern side of trees may be more mature than those found on the northern side (von Aderkas, pers. comm). There is a small window of opportunity for *Megastigmus* to oviposit; therefore, variability in cone development within and between trees is also an important factor in female site selection.

2.5 Conclusions

In this study, competitive displacement along with multiple morphological characteristics of Douglas-fir cones were used to assess preference in oviposition sites by *Megastigmus*. Competition and seed position were significant predictors of infestation. Overall, the proximal and distal ends of cones had a higher percent of infested seed than the central region. Additionally, *Megastigmus* adult females oviposit into seed that would not normally complete development. As Rappaport and Roques (1981) suggested, preferentially ovipositing into seed at the extreme ends of cones may just be an evolved response to competition. Even though seed resources are smaller and potentially have low nutritive value at the ends of cones, the ability of *Megastigmus* to manipulate its host allows it to utilize areas that are undesirable to other insects. Small seeds may only be marginally suitable for larval development, however, oviposition into small seed is likely an adaptive response to find enemy-free space. To determine if *Megastigmus* use preference – performance strategies when selecting oviposition sites, an interesting study would be to evaluate larval survivorship as well as sex ratios between larvae occupying seeds from the tips of Douglas-fir cones compared to larvae found in seed within the central region of a cone. Perhaps adult *M. spermotrophus* females invest more into their

female offspring compared to their male offspring and consequently lay male eggs in small seed that may be less desirable for offspring success.

Strong correlations between seed position and seed length as well as position and scale thickness suggest that competition may not be the only predictor of infestation. Accessibility of seed to *Megastigmus* is equally important in locating suitable oviposition sites. Contrary to these findings, Rapport and Roques (1991) report *Megastigmus* to preferentially occupy seed in the central portion of cones in Douglas-fir trees outside their native range. If ovipositor length is truly a limitation to *Megastigmus* oviposition, it would be necessary to compare scale thickness of cones between trees located in native and exotic habitat. Although studying differences between trees is important in explaining variability of infestation, it would also be interesting to study variance in *Megastigmus* ovipositor length. Perhaps ovipositor lengths differ among individuals penetrating scales in the central region of cones compared to those penetrating the extremities.

This study shows that many factors contribute to host selection by *Megastigmus*. Conceivably there is a hierarchy of preference in which sites are ranked according to resources availability, and presence of competitors. Many phytophagous insects have been shown to rank resources (Jaenike, 1978; Craig et al., 1989). They can have a fixed range of suitability or be more flexible when resources are limited. *Megastigmus* may prefer large seed if it is attainable but when competitors are present, females may choose to oviposit into smaller seed at the extreme ends of cones. Selective pressures discussed

in this study attributing to *Megastigmus* host selection will likely vary from site to site. Insight into the relationship between preference and performance is critical to understanding distribution of insect populations and understanding host shifts.

CHAPTER 3. PARASITIZED DOUGLAS-FIR OVULES: INSIGHT INTO PLANT DEFENSE AND ITS MANIPULATION BY *MEGASTIGMUS SPERMOTROPHUS*

3.1 Introduction

Douglas-fir seed is parasitized by *M. spermotrophus* (Hussey, 1955). To understand the possible interactions between plant and insect, it is first necessary to consider how plants defend themselves and how insects overcome these defenses. Insects are not aseptic organisms, but often carry microorganisms and viruses. Consequently, the plant must respond to a community of invaders, which poses an interesting set of challenges. This introduction will highlight some defense responses of plants before addressing the objectives of this study.

Plants' ability to detect and respond to a multitude of invaders, including herbivores and pathogens has been of pivotal importance for their survival and evolution. Angiosperms are without a doubt the most-well studied systems. Responses such as wounding and "non-self" recognition described in flowering plants appear to be fairly conserved throughout the evolution of land plants (Winter et al., 2014; Oliver et al., 2009). Recently, groups ancestral to angiosperms, such as bryophytes, have been shown to share similarities with angiosperms in perception, signalling, and activation of defense (Ponce de Leon et al., 2007; Ponce de Leon and Montesano, 2013; Wang et al., 2010).

Plants have developed a two-tiered defense strategy that includes constitutive and inducible responses (Freeman and Beattie, 2008). Induced defense systems include

hypersensitive responses (HR) and systemic acquired resistance (SAR). The latter develops in plants that have been previously exposed to specific pathogens or herbivores. Once primed, the plant exhibits higher resistance to subsequent attacks (Conrath, 2006). This system is analogous to aquired immune system responses in animals. Acquired resistance can be both local and systemic (Conrath, 2006). Constitutive defenses in plants include physical defenses, such as cuticle layers, casparian strips, hairs and other morphological structures that decrease an invasive organism's ability to penetrate into the nutrient-rich plant interior (Hanley et al., 2007; Serrano et al., 2014). Chemical defenses in plants include feeding deterrents, such as minerals (silica), crystals (calcium oxalate), terpenes, phenolics, and alkaloids, amongst others (Hanley et al., 2007; Hudgins et al., 2003; Zhao et al., 2011; War et al., 2012). Defense compounds can be distributed among cells and concentrated in hairs, resin ducts or specialized cells.

In Douglas-fir seed there are three kinds of constitutive defenses; a lignified seed coat, that is relatively impermeable, a nucellus that contains defense compounds (antimicrobial proteins) (Poulis et al., 2005) and a megaspore wall, which is a thick physical barrier of sporopollenin coupled to glycoproteins (Pettitt, 1977). Inducible defenses are more diverse, as they include upregulation of gene products by a number of molecular systems.

Plants produce hormones that act upon signalling networks for not only developmental processes but also stress responses. Jasmonic acid (JA), ethylene and salicylic acid (SA), have been implicated in defense signalling and wound responses (Broekaert et al., 2006; Creelman and Mullet, 1995; Ralph et al., 2006). As is typical for many plant hormones,

they also play regulatory roles in plant development. The overlap or even redundancy in these regulatory roles poses challenges in the interpretation of gene expression studies, which is best illustrated with an example. Abscisic acid (ABA) initiation and maintenance are important for both seed and bud dormancy as well as for regulation of drought responses (Kermode, 2005). ABA also has important roles in seeds, inducing embryogenesis and seed maturation, dormancy and preventing germination (Kermode, 2005; Chiwocha and von Aderkas, 2002). ABA is a hormone found in all land plants but over the course of evolution many functions have been acquired.

An important defense response that involves many signalling networks is wounding. It is a defense response conserved in land plants. Wounding is a response to tissue damage from biotic or abiotic stress and is thought be mediated through jasmonate signalling (Titarenko et al., 1997). Initially, subcellular organization and structural changes occur in chromatin, which triggers rRNA gene expression (Pecinka and Scheid, 2012). Consequently there are elevated levels of transcription, translation and secretion in affected cells. Active cell division results in a seal over a wound site and repairs tissues including vascular transport networks, mechanical tissue and storage tissue. Wounding responses cause mass changes in gene expression; Ralph et al., (2006) showed that 11% of the transcriptome is regulated by wounding. High throughput methods have been used to identify wound-responsive genes (Durrant et al., 2000; Reymond et al., 2000). Defense responses occur directly at wound sites or can be triggered in surrounding, undamaged tissue through SAR.

Although herbivores and parasites such as *Megastigmus* cause wounding, it is not known whether they can suppress defense responses of their host. Plants generally perceive herbivores by effector molecules that trigger biosynthesis of secondary metabolites (Howe and Jander, 2008). Larvae of *Pieris brassicae*, a lepidopteran herbivore, release proteins in their saliva that suppress induction of herbivore defense genes (Consales et al., 2012). Some herbivore associated molecular patterns (HAMPS) that are perceived by plants include the enzymes glucose oxidase, β-glucosidase, as well as fatty acid amino acid conjugates, sulphur containing fatty acids, cell wall fragments, and peptides digested from plant proteins (Consales et al., 2012; Mattiacci et al, 1995; Alborn et al., 2007; Mithöfer and Boland, 2008). These are not general elicitors, but are particular to specific plant-insect interactions.

When herbivores carry microorganisms within, these fellow-travellers may be employed to fool plant defences. Herbivores can exploit bacteria within their saliva to deceive plant defense responses, which switch from herbivore to pathogen defense. For example, Colorado potato beetle larvae secrete bacteria in their saliva to avoid antiherbivore defenses in tomato plants (Chung et al., 2013). The beetle is, in effect, diverting resources away from the appropriate plant immune response. In this example, it is the bacterial-rich oral secretions of the larvae that elicit plant response. The plant perceives a pathogen attack, instead of an insect attack. Insects may benefit from vectoring plant pathogens because the plant's response to the insect is conflicted by the plant's simultaneous response to pathogens; the conflict is manifested in antagonisitic cross-talk between signal transduction pathways (Koornneef and Pieterse, 2008). Insect-associated

microbes can cause changes in host plant physiology. For example, the lepidopteran leaf-miner, *Phyllonocrycter blancardella* causes green island phenotypes in apple leaves in order to delay leaf senescence (Kaiser et al., 2010). When the leaf-miners are treated with antibiotics, the green island phenotypes do not form, suggesting that microbial associates may be involved in manipulation of leaf senescence. Microbes associated with the larvae of a beetle, *Diabrotica vigifera* are thought to be involved in supressing defense genes in maize roots (Barr et al., 2010). *Megastigmus* have been shown to harbour microbial associates including, *Wolbachia, Rickettsia, and Ralstonia* (Paulson et al., 2014).

Defense responses of conifers have been studied with respect to wood-boring and phloem-feeding insects (Miller et al., 2005; Byun-McKay et al., 2006) but no attention has been given to seed feeders. Nothing is known about molecular responses in seeds of gymnosperms. An understanding of how gymnosperms respond to conophytes may provide many clues to the evolution of seed defenses of angiosperms. Seed isolation within carpels may have provided an evolutionary advantage to flowering plants, which are the dominant seed plants today. The fact that gymnosperm ovules are exposed but angiosperm ovules are hidden visibly in a protective structure represents a different opportunity for seed insects.

In a sense, this opportunity defines the plant-attacking members of *Megastigmus*, of which, some species are phytophagous, gall makers of angiosperms and other species are seed parasites of gymnosperms. Gall formation represents interplay between host and

insect; the insect manipulates the host to make a hollow shelter and provide and supply an inner wall of nutritional cells to feed the developing larva (Stone and Schönrogge, 2003). The insect does not kill its hosts, but consumes a controlled amount. The host can reproduce, although at reduced rates. The host species pays a price in loss of fitness. Seed parasites exact a greater cost in fitness as *Megastigmus* adult females lay their eggs directly into developing seed and cause that seed to fail. However, just as in the gall, the seed accommodates the larvae, feeding it until it is able to metamorphose, mature and reproduce.

In the case of the gall, it is certain that the insect must manipulate its host's gene expression to force the host to make structural and nutritional alterations (Nabity et al., 2013). The case of the seed parasite is less clear. Not only does the structure remain unaltered, but so does the nutritional status of the developing seed. Insect larvae have merely supplanted the plant embryo. Maternal care is provided by the plant to the insect.

The interaction between Douglas-fir seed and the chalcid wasp, *Megastigmus spermotrophus* is an interesting system to study. Developing larvae of *M. spermotrophus* feed on megagametophyte tissue, rich in starch, triacylglercols, and nitrogen rich proteins. Similarly to galls, *M. spermotrophus* can induce changes in plant tissue that redirects the movement of storage products (von Aderkas et al., 2005a).

Douglas-fir megagametophytes that have been parasitized by *M. spermotrophus* do not appear to defend themselves. There is no obvious visual sign of a hypersensitive

response leading to cell death in the megagametophyte due to insect attack and herbivory. Instead, the megagametophyte behaves as though it is carrying an embryo – the developing larvae act as surrogates – and the megagametophyte continues to accumulate storage reserves (von Aderkas et al., 2005b). The only visible sign of *M. spermotrophus* activity is the formation of small resin drops on scale tissue at the site of oviposition. This is likely a response to wounding; however, this response may only appear in sporophytic tissue. It is unclear if the megagametophyte also responds to wounding.

3.1.2 Objectives

The focus of the work presented in this chapter is to identify defense responses in haploid Douglas-fir tissue parasitized by *M. spermotrophus*. An additional aim of this work is to identify Douglas-fir genes that are influenced by *M. spermotrophus* infestation, in order to understand the process of manipulation within parasitized plant tissue.

To identify Douglas-fir genes responding to infestation and genes that may be manipulated by *M. spermotrophus,* I used an RNA-seq approach. This method combines deep sequencing technologies with transcriptomics. RNA-seq is a revolutionizing tool for studying gene expression, especially for studying non-model organisms, whose genetic sequences have yet to be determined (Wang et al., 2009). Advancements in computational methods have made *de novo* assembly possible in the absence of a reference genome. This study takes advantage of large plant databases such as the *Arabidopsis* information resource (TAIR) and Uniprot, along with a recently sequenced parasitoid wasp (*Nasonia vitripennis*) genome that is a close relative to *M.*

spermotrophus, and bioinformatics to identify potential candidates involved in plant defense and manipulation.

The objectives of this study were three-fold; 1) to use histochemical techniques to characterize accumulation of storage products during stages of megagametophyte development, 2) to construct a *de* novo assembly of parasitized Douglas-fir ovules from samples of infested and uninfested megagametophytes that have been pollinated and unpollinated and 3) to obtain expression data for each treatment to identify highly differentially expressed transcripts that are involved in Douglas-fir megagametophyte defense response to infestation and transcripts that are manipulated by *M. spermotrophus.*

3.2 Materials and methods

3.2.1 Study system

3.2.1.1 The host

Douglas-fir trees were sampled from the Timber West seed orchard in Saanichton, BC, Canada (48°35'54.00" N, 123°25'56.87" W). Six ramets with abundant female cones were selected from a first generation clone, no. 969. Immature male buds were manually removed from selected branches. The emasculated branches were covered with paper bags to isolate their female buds from pollination. When female strobili became receptive, half of the bagged branches were randomly selected for manual pollination using a polymix of six pollen genotypes. To ensure successful pollination, the polymix was injected twice within a week. Two weeks after pollination, pollen exclusion bags were replaced with insect mesh exclusion bags.

3.2.1.2 The parasite

Megastigmus spermotrophus were raised from infested seed provided by the Surrey Tree Seed Centre. Insects from lot 63289 were reared in darkness in a 19 °C chamber. After initial emergence of the males, the seed was placed into a 15 °C chamber with no light until the females emerged. Adult females (n=490) were collected in vials, 10 insects per vial. They were fed 10 % sugar water and placed in an 8 °C chamber overnight, before using them for field trials. Mating time was not taken into consideration because all adult females are able to lay eggs parthenogenetically. *Megastigmus spermotrophus* eggs will develop regardless of whether they have been fertilized; however, unfertilized eggs can only develop into male offspring.

Of the pollinated and unpollinated bags in the orchard, approximately half were randomly selected for *Megastigmus* introduction. Insects were added to these bags one day following feeding. Bags with less than 20 cones were each given 10 female wasps, while bags with 20 or more cones were each given 20 wasps. To ensure a high success of infestation, a second round of *Megastigmus* with five female wasps per bag were added a week later.

3.2.2 Tissue collection

Five days following *Megastigmus* introduction, a minimum of 40 megagametophytes were dissected every 2-3 days from each of the four treatments – pollinated (P), unpollinated (U), pollinated-infested (PI), and unpollinated-infested (UI). Typically, dissections were performed on two freshly harvested cones at a time. Once their

megagametophytes were removed and frozen, anther two cones were picked and dissected. Identification of infested megagametophytes in PI and UI treatments was determined by looking for eggs or larvae under a dark-field microscope. Each dissected sample was immediately flash- frozen in liquid nitrogen. Tissue was stored at -80 °C until RNA extractions were performed. Cones selected each day for dissection were randomly chosen using a random number generator. Seven samples were collected for each treatment over a three-week period, along with a single sample of the integument and ovuliferous-scale-bract complex (OSBC) (Table 5). Sampling dates correspond with stages of Douglas-fir development, which include corrosion cavity formation, fertilization, embryogenesis and abortion of unpollinated megagametophytes.

3.2.3 RNA-sequencing

3.2.3.1 RNA extractions

RNA was extracted using a modified procedure from Kolosova et al. (2004). Samples of megagametophytes were ground into a fine powder in liquid nitrogen using a mortar and pestle and stored at -80 °C. At least 65 mg of ground tissue from each sample were used to obtain a minimum of 20 μg of RNA for sequencing**.**

Ground tissue was added to 1 mL of an extraction buffer (200 mM Tris-HCl pH 8.0, 1.5 % lithium dodecylsulphate, 300 mM lithium chloride, 10 mM EDTA disodium salt, 1 % sodium deoxycholate, and 1 % tergitol NP-40) along with 1 mM aurintricarboxylic acid, 10 mM dithiothreitol, 5 mM thiourea, and 2 % w/v polyvinylpolypyrrolidone and immediately flash-frozen in liquid nitrogen.

Table 5 - Summary of megagametophyte (pollinated [P], unpollinated [U], pollinate and infested [P+I] and unpollinated and infested [U+I]) and ovule tissue (Integument, and ovuliferous scale bract complex [OSBC]) collected over seven sampling dates during the 2012 field season.

18-Jun	**21-Jun**	**25-Jun**	**28-Jun**	**03-Jul**	**06-Jul**	**10-Jul**
P	P	P	P	P	P	P
U	U	U	U	U	U	U
P+I	P+I	P+I	P+I	P+I	P+I	P+I
U+I	U+I	U+I	U+I	U+I	U+I	U+I
					Integument	
					OSBC	

Samples were thawed and immediately centrifuged at 18,000 X *g* for 15 min at 4 °C. The supernatant was added to 75 µL of 3.3 M sodium acetate (pH 6.1) and 750 µL of chilled isopropanol. These samples were incubated at -80 °C for 1.5 hr to precipitate RNA. Samples were thawed at room temperature and immediately centrifuged at 8,000 X *g* for 30 min at 4 °C. Following centrifugation, the supernatant was discarded and the pellet was resuspended over ice in 400 µL of TE Buffer (10 mM Tris (pH 8.0) and 1 mM EDTA (pH 8.0)) and 400 µL of 5M NaCl. The suspension was vortexed with 200 µL of cetyltrimethylammonium bromide (CTAB) and incubated at 65 °C for 5 min. Following incubation, the solutions were extracted two times with 500 µL of 25:1 chloroform:isopentanol and centrifuged at 5,000 X *g* for 6 min at 4 °C. The aqueous phase was mixed with 100 µL 8M lithium chloride, and incubated overnight at 4 °C.

Solutions were thawed and immediately centrifuged at 18,000 X *g* for 30 min at 4 °C. The supernatant was discarded and the RNA pellet was washed with 500 µL of 70 % ethanol, followed by another centrifugation at 18,000 X *g* for 10 min at 4 °C. The pellets were dried and resuspended in 30 µL of nuclease-free water.

3.2.3.2 Removal of genomic DNA and quality analysis

A DNase digest, adapted from Invitrogen (USA) was performed on extracted RNA to remove contaminating genomic DNA. DNase reaction volumes were 100 µL, containing 20 µg of RNA, 10 µL of DNase buffer (Invitogen, USA), 2 µL of RNase (Invitrogen, USA) inhibitor and 2 µL of DNase I (Invitrogen, USA). Solutions were incubated for 30 min at 37 °C. Reactions were stopped with the addition of 20 µL of 0.1 M EDTA (pH

8.0). Solutions were extracted with 110 µL of 25:24:1 phenol:chloroform:isopentanol, vortex-mixed and centrifuged at 14,000 X *g* for 10 min. The aqueous phase was extracted twice with one volume of 24:1 chloroform:isopentanol and centrifuged at 14000 X *g* for 10 min. RNA was precipitated by adding 10 µL of 3.3 M sodium acetate (pH 6.1) and 120 µL of chilled isopropanol to the supernatant. The solution was incubated for 2 hr at -20 °C and centrifuged at 14,000 X *g* for 30 min at 4 °C. The supernatant was discarded and the pellet was washed with 70 % ethanol. A final centrifugation at 14,000 X *g* for 10 min at 4 °C was done. The pellet was dried and resuspended in 30 µL of nuclease-free water and stored at -80 °C.

The total RNA and the purity of mRNA extracted were tested using the BioRad Experion (Hercules, California, USA) RNA StdSens Analysis kit according to the manufacturer's instructions.

3.2.3.4 Library construction and sequencing

Complementary DNA library construction and paired-end, strand-specific RNA sequencing on the Illumina Sequencing platform (HISeq 2000) was performed at the Michael Smith Genome Science Centre in Vancouver, Canada. A total of 33 samples were sequenced across eight lanes, each holding four multiplexed samples. In addition, one lane had five multiplexed samples. Sequenced samples included megagametotphyes from all sampling dates along with three technical replicates for samples with extra tissue, (June 28_PI, July 3_P, and July 6_U) and a sample of integument and OSBC. Sequence length size exclusion was targeted for between 200 and 300 base pairs (bp) to

ensure a paired-end read length of 100 bp. Sequence data were packaged into BAM files containing basecalls, i.e., both reads and quality scores.

3.2.3.5 Read quality analysis and trimming

BAM files were converted to FASTQ files using Picard-tools, SamtoFastq package (v1.98) (Li et al., 2009). FASTQ files were assessed for the following: per base sequence quality, per base sequence content, and sequence duplication using FASTQC (v) (Andrews, 2012). Reads were quality filtered using Trimgalore! (v.0.2.5) (Krueger, 2012) with a Phred cutoff value of 20. Illumina barcodes and adapters were removed in this process. Reads were considered of low quality and discarded if either their phred quality score was lower than 20, if they were shorter than 36 bp, or if one of the pairs was removed as a results of these cutoff values. To ensure quality improvement, trimmed reads were visually evaluated from figures produced with FASTQC.

3.2.3.6 *De novo* assembly

Trimmed strand specific, paired-end reads from 33 libraries were pooled to assemble a Douglas-fir/*M.spermotrophus* (larvae) *de novo* transcriptome assembly using Trinity (Grabherr et al., 2011). Prior to assembly, trimmed reads were normalized for sequence coverage to reduce run times using Trinity's *in silico* normalization. The minimum and maximum k-mer coverage for normalization was 1 and 30, respectively. A Trinity assembly was created with the normalized reads.

3.2.3.7 Annotation

BLAST databases were made for the *Nasonia vitripennis* (Munoz-Torres et al., 2011) protein sequence collection, *Pinus taeda* genome (Zimin et al., 2014), *Arabidopsis* protein sequence collection (Lamesch et al., 2011), and the UniProt/SwissProt sequences. Transcripts from the Douglas-fir/*M. spermotrophus* assembly were assigned annotations by querying the databases using BLASTX 2.2.29+. For each query, the cutoff e-value was set to 10^{-5}. Only the best hit was retained for each query. TAIR and UniProt accessions were linked to each transcript with a BLAST hit. Transcripts with hits to *P. taeda* were identified.

3.2.3.8 Read mapping, counting and normalization

Transcript expression profiles were quantified using RSEM software package (v1.2.0) (Li and Dewey, 2011). Both forward and reverse reads were aligned to quantify the estimated coverage counts. Sequence length was normalized within libraries using RSEM. Transcripts with estimated coverage counts of 0 across all samples were removed. FPKM values were lifted by a value of 1 and converted to $\log_2 FPKM_1$ resulting in a minimal $\log_2 FPKM_1$ of 0. Since only three samples had technical replicates, the expression data for these samples were averaged.

3.2.3.9 Post-assembly filtering

The dataset was reduced by discarding transcripts in the assembly that had low expression across the majority of samples. Fold change for each transcript across all

samples was calculated by taking the maximum $\log_2 FPKM_1$ expression value and dividing it by the minimum $\log_2 FPKM_1$ value. The averaged expression value across all samples was determined. A fold change cutoff value was calculated for each transcript using the following equation:

$$\text{fold change cut-off} = \left(\frac{3}{\text{average } \log_2 FPKM_1} \right) + 2$$

Transcripts were kept if the fold change was greater than the fold change cut-off (FC > FCcut-off). Transcripts were further filtered to include only those that had an open reading frame of 300 bp or more. Open reading frame counts were determined using EMBOSS getorf (v 6.6.0.0) (Rice et al., 2000).

Taxonomic IDs from NCBI were linked to transcripts of the reduced dataset that had an annotation. Using NCBI taxonomic lineage, each transcript was further classified according to its sequence similarity to one of the following groups: plant (Viridiplantae), insect (Insecta), other (including bacteria, fungi, mammals, crustaceans, etc) or unknown. If a transcript was categorized as "other" or "unknown" but had sequence similarity to *Pinus* or *Nasonia,* the transcripts were put into plant and insect groups, respectively. Transcripts classified as insect but also had sequence similarity to *Pinus* were evaluated based on bitscore; the transcript was placed into the category with the higher bitscore.

3.2.3.10 Differential expression profiling and candidate gene search

For the purpose of this thesis, only transcripts with sequence similarity to plants that had an annotation were selected for further analysis. To identify transcripts differentially

changing over megagametophyte development between P, U, PI and UI treatments, quadratic regression analysis using time course expression data for each treatment was performed. Transcripts were categorized into a regression pattern if their expression data significantly fit the model ($p < 0.05$, $r^2 > 0.9$). Transcripts fitting a regression model were used to search for candidate genes involved in plant defense or genes that were manipulated by *M. spermotrophus*.

Using the results from the quadratic regression, candidate genes were identified to address the objectives of the study as follows:

1. Transcripts involved in megagametophyte defense response to *M. spermotrophus* were defined as transcripts where PI and UI were at least 8-fold ($\Delta\log_2 > 3$) greater than P or U, respectively at a minimum of 2 sampling dates.
2. Transcripts involved in seed development that were manipulated by *M. spermotrophus* were defined as transcripts where P, PI, and UI were at least 8-fold ($\Delta\log_2 > 3$) greater than U at a minimum of 2 sampling dates. Additionally, the fold change expression values for the transcripts manipulated by the seed chalcid had to significantly fit into a quadratic pattern. Transcripts were also considered to be manipulated if U samples were at least 8-fold ($\Delta\log_2 > 3$) larger than P, PI and UI at a minimum of 2 sampling dates. The fold change value for these transcripts also had to significantly fit a quadratic pattern. Thus, manipulated transcripts by *M.*

spermotrophus were ones where the UI treatment was not similarly expressed to U treatment.

Filtered transcripts from each of the objectives were hierarchically clustered based on mean centred $\log_2[FPKM]_1$ values, using Genesis version 1.7.6 (Sturn et al., 2003). Transcripts were clustered based on their expression profiles within each treatment.

3.2.4 Histology

Over the duration of the 2012 field season, ovules from Douglas-fir genotype 969 were collected once a week from May until July. In 2013 an additional genotype, 970 was used for collections. In both years, select genotypes were chosen for their abundance of female cones. Megagametophytes from ovules infested with *M. spermotrophus* were also collected from Douglas-fir trees found on the University of Victoria campus. These samples were chosen to illustrate the stage of megagametophyte development that corresponded to RNA-seq collections. Samples were fixed in 2.5 % glutaraldehyde in 0.05 M sodium phosphate buffer and stored at 4 °C until required. Samples were rinsed in 0.05 M sodium phosphate buffer, dehydrated through a series of graded ethanol solutions and embedded in glycol methacrylate (Technovit 7100, Marivac, Canada). Using a Leica SM2400 sledge microtome equipped with a tungsten-carbide knife, samples were sectioned at 4 µm thickness. Sections were either stained with a combination of Ponceau Red 2S and Azure Blue for proteins and cell wall, respectively, or with a combination of Toluidine Blue O and Lugol's for a general metachromatic stain as well as starch. In addition, some sections were stained with a combination of Safranin

O, Azure Blue for cell walls and with Lugol's for starch. Staining procedures were according to Gutmann (1995). Images were compiled and adjusted for brightness and contrast in Adobe Photoshop CS6, version 13.0x64.

3.3 Results

3.3.1 Histological samples

Young megagametophytes undergoing free nuclear division did not accumulate starch or proteins within coenocyte cells (Figure 10). At the stage of free nuclear division, starch only accumulated in integument cells surrounding the micropyle of the ovule. Following alveolation and cellularization of the megagametophyte starch begins to build up. Just prior to fertilization at the central cell stage, megagametophytes had accumulated starch granules at the base of the central cells and expanded into the location of the eventual corrosion cavity (Figure 11). Starch was continually deposited after infestation by *M. spermotrophus* (Figure 11). Once the central cell further divided into the egg and ventral cell, starch deposition accumulated further into the megagametophyte where the corrosion cavity would form (Figure 12). Protein production was still not observed in cells at the egg stage.

During early embryo development, starch granules become more numerous and prominent within each prothallial cell that surrounded the formed corrosion cavity (Figure 13). Protein bodies also began to accumulate within prothallial cells around the corrosion cavity (Figure 13). Additionally, multinucleate cells were observed. Ovules that were not fertilized only accumulated starch (Figure 14).

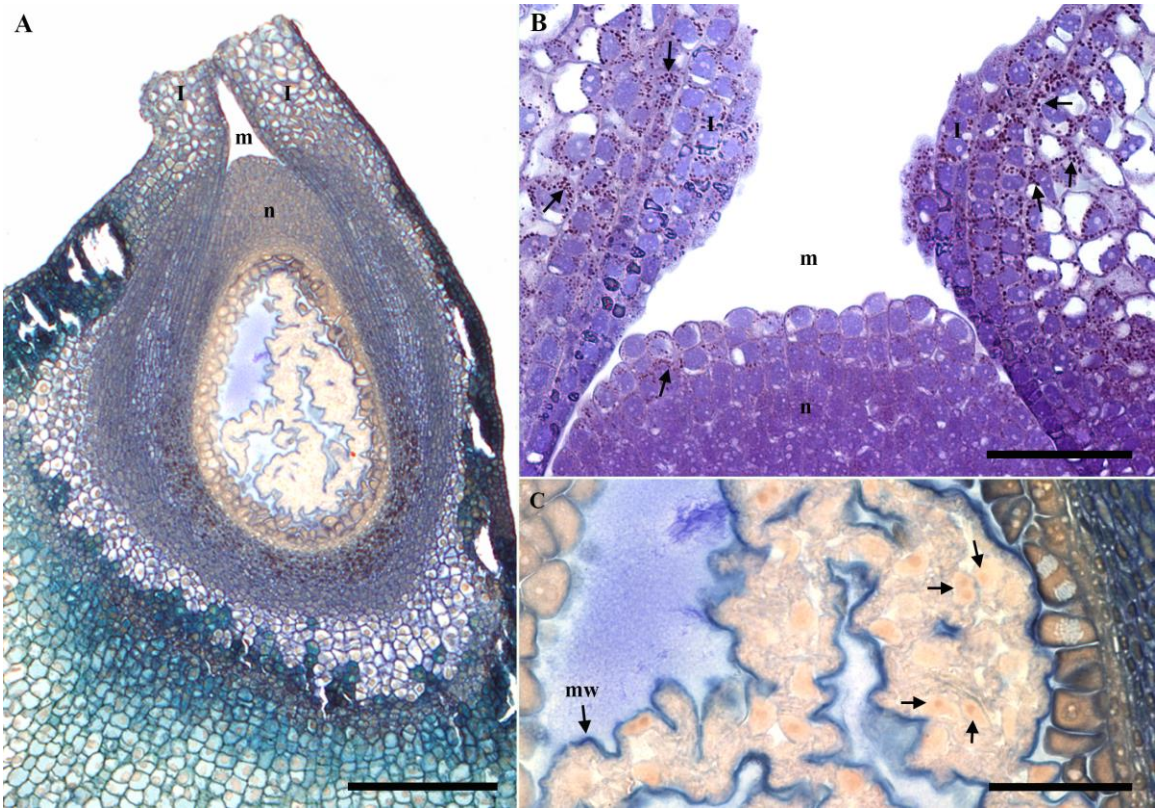


Figure 10. Female gametophyte development in mid-May, during free nuclear division, several weeks after pollination. A. Longitudinal section of an ovuliferous scale, showing the integument (I), nucellus (n), micropyle (m) and development of the megagametophyte, stained with Ponceau Red 2S-Azure Blue. B. Integument and nucellus, showing starch accumulation (arrows) around the micropyle, stained with Toludine Blue O-Lugol's IKI. C. Section of the developing ovule, showing the megaspore wall (mw) and multiple nuclei (arrows) within a single cell, stained with Ponceau Red 2S-Azure Blue. The scale bar represents 500 μm for A, and 100 μm for B and C.

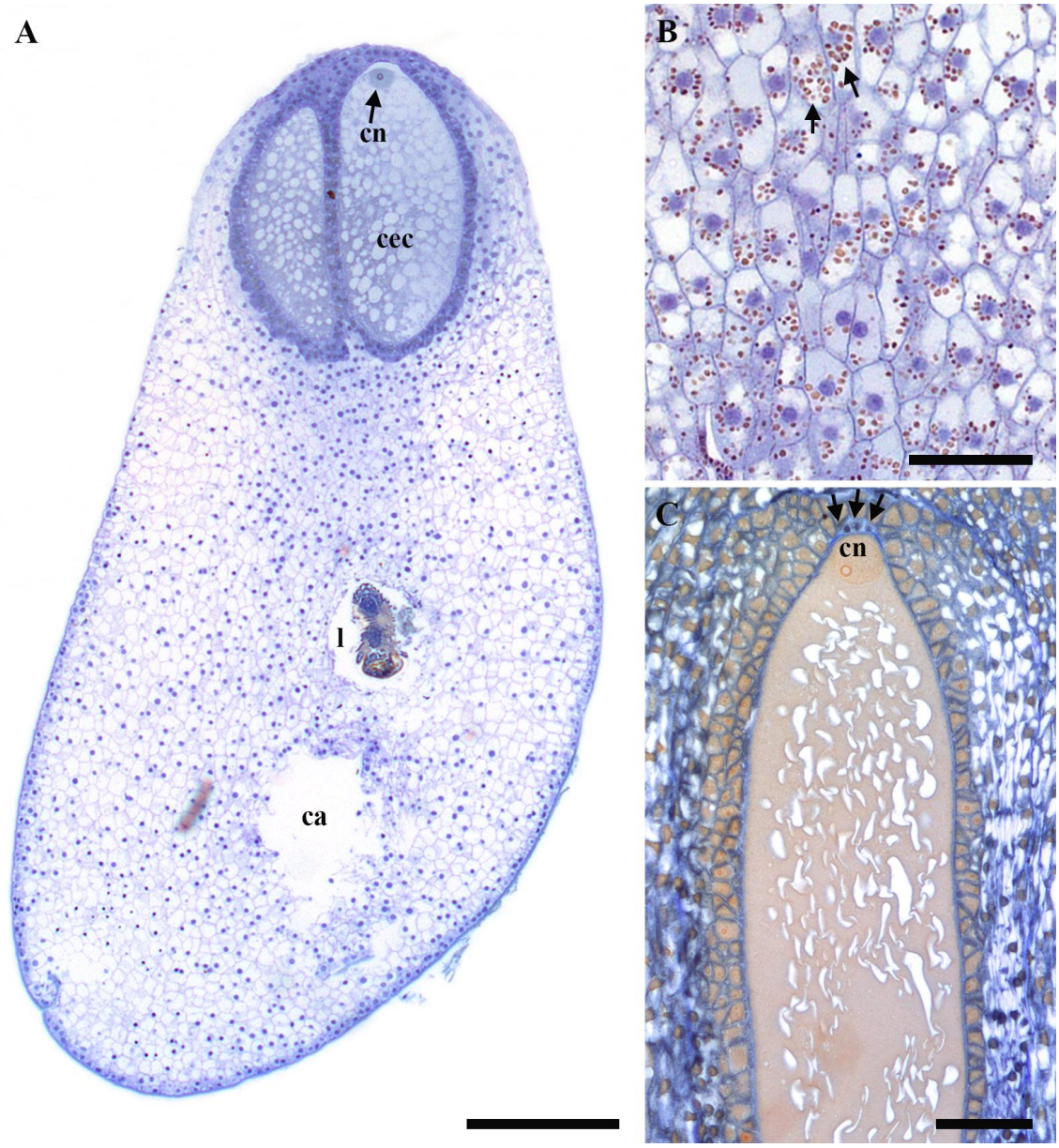


Figure 11. Female gametophyte infested with *M. spermotrophus* larva in mid-June, during central cell development. A. Longitudinal section of a megagametophyte, showing a central cell (cec), central cell nucleus (cn), larva (I) and consumed gametophyte tissue (ca), stained with Ponceau Red 2S-Azure Blue. B. Section of the megagametophyte, showing the accumulation of starch (arrows) in prothallial cells, stained with Ponceau Red 2S-Azure Blue. C. Section of a pollinated gametophyte, showing the central cell with the central cell nucleus, and neck cells (arrows), stained with Toludine Blue O-Lugol's IKI. The scale bars represent 500 μm for A, and 100 μm for B and C.

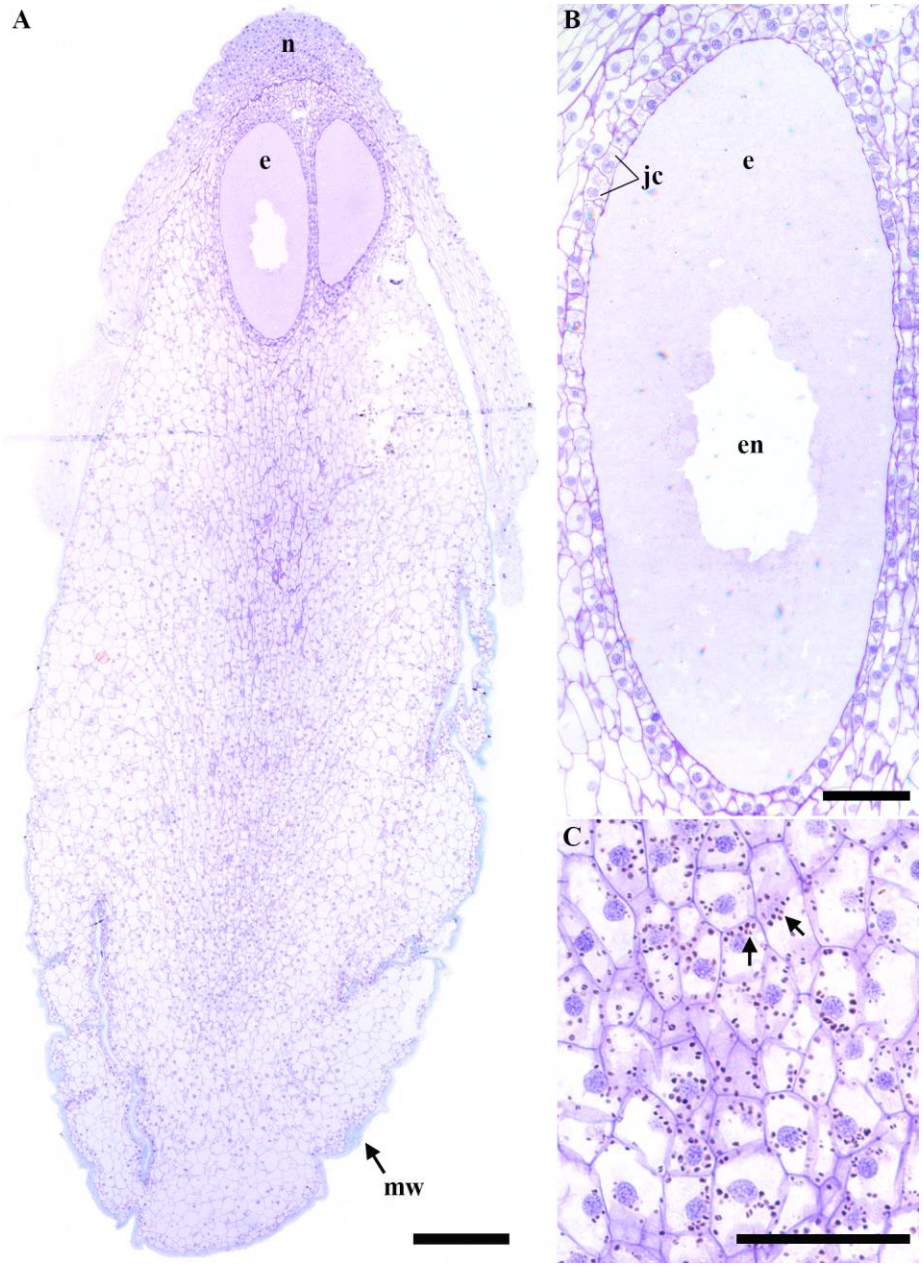


Figure 12. Pollinated female gametophyte in mid-June, during egg development, around the time of fertilization. A. Longitudinal section of a megagametophyte, showing the megaspore wall (mw), nucellus (n) and the egg cell (e). B. Egg cell, showing the egg nucleus (en), and jacket cells (jc). C. Section of the megagametophyte showing starch accumulation (arrows). All sections were stained with Toludine Blue O-Lugol's IKI. The scale bars represent 500 μm for A, and 100 μm for B and C.

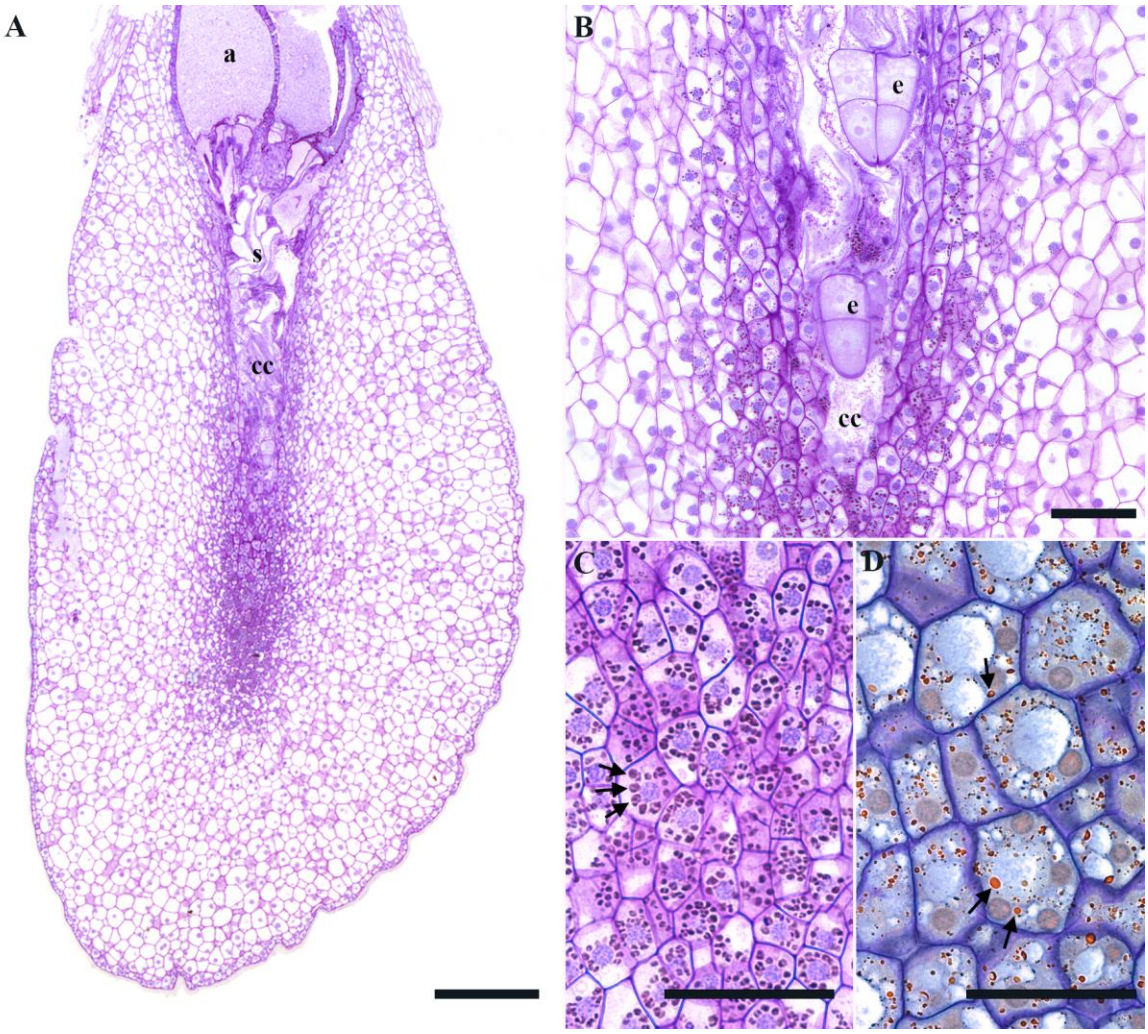


Figure 13. Fertilized female gametophyte in early July during early embryo development. A. Longitudinal section of a megagametophyte, showing the archegonia (a), suspensor cells (s), and corrosion cavity (cc), stained with Toludine Blue O-Lugol's IKI. B. Section of the megagametophyte, showing the two embryos (e) being pushed to the tip of the corrosion cavity by the suspensor cells, stained with Toludine Blue O-Lugol's IKI. C. Tip of the corrosion cavity, showing accumulation of starch (arrows) in the prothallial cells, stained with Toludine Blue O-Lugol's IKI. D. Section of the megagametophyte, showing protein bodies (arrows) within the prothallial cells, stained with Ponceau Red 2S-Azure Blue. Cells with arrows also have multiple nuclei. Scale bars represent 500 μm for A, and 100 μm for B-D.

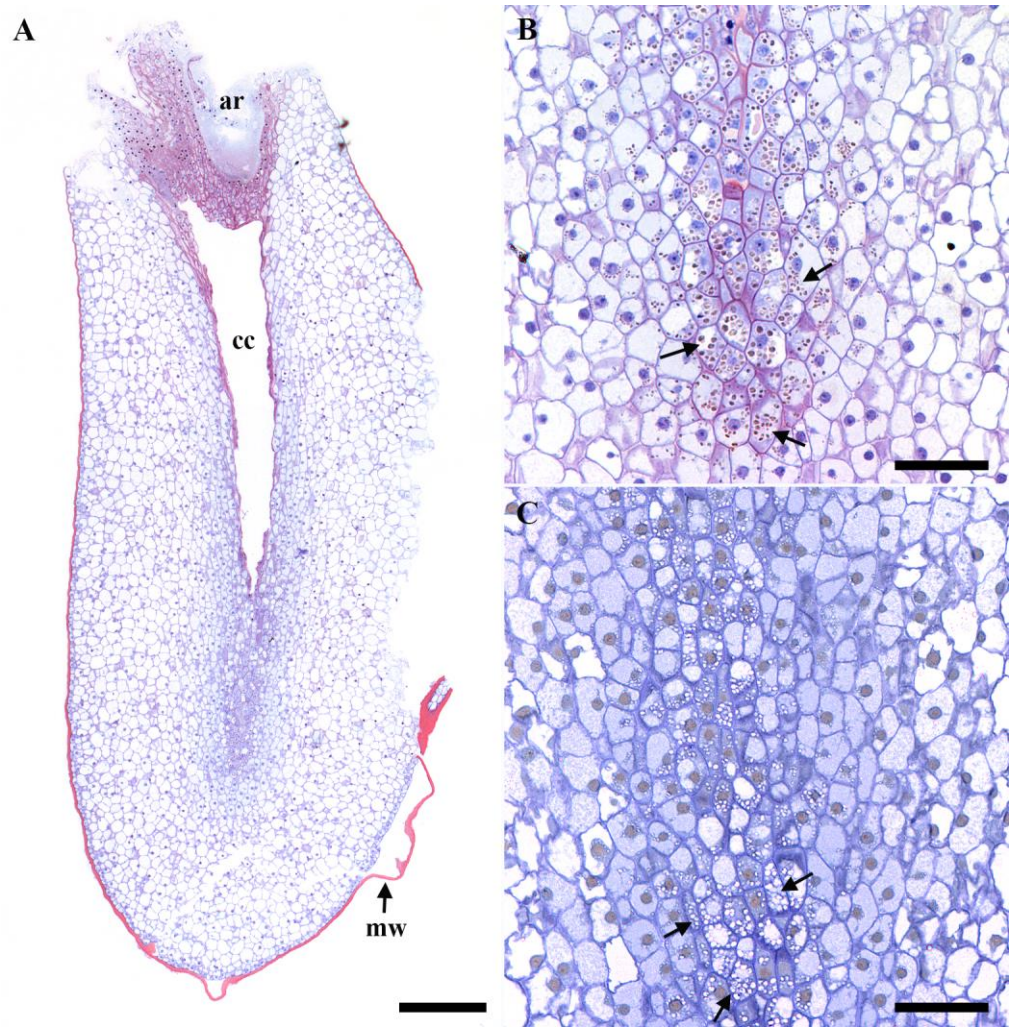


Figure 14. Unpollinated female gametophyte in late June, shortly after the time of expected fertilization. A. Longitudinal section of a megagametophyte, showing the archegonia with no egg cells (ar), corrosion cavity (cc), and megaspore wall (mw), stained with Safranin O-Azure Blue-Lugol's IKI. B. Section of the megagametophyte, just above the corrosion cavity, showing accumulation of starch (arrows), stained with Safronin O-Azure Blue-Lugol's IKI. C. Section of the megagametophyte, showing starch (arrows), but no protein accumulation, stained with Ponceau Red 2S-Azure Blue. The scale bars represent 500 μm for A, and 100 μm for B and C.

Unfertilized ovules had deep corrosion cavities with starch granules only present in cells below the cavity. There were no observed unfertilized sections that had protein bodies within the prothallial cells. At no point during megagametophyte and embryo development were phenolic deposits observed in cells.

3.3.2 RNA-seq analysis of Douglas-fir megagametophyte development

3.3.2.1 *de Novo* assembly

Raw read pair counts ranged from 29 billion (June 25 UI) to 173 billion (June 21 P). Filtering low quality reads using FASTQC resulted in an overall 2.8% loss of raw reads (Table 6). Trimmed read pair counts ranged from 28 billion to 168 billion. Approximately 320 billion reads were combined from all libraries to create a *de novo* assembly.

The Douglas-fir/*M. spermotrophus* assembly consisted of 382,101 putative genes, derived from contigs corresponding to 738,221 isoforms. Of the total number of putative genes, 351,150 genes were composed of single contigs, whereas 30,951 were composed of multiple isoforms.

3.3.2.2 Post-assembly filtering

There was a greater number of transcripts with low expression ($\log_2$ $FPKM_1$), compared to transcripts with high expression. The majority of low expression transcripts did not pass post-assembly filtering (Figure 15).

Table 6. Read-pair counts per library, prior to and following removal of low quality read-pairs using Trim Galore!. Read-pairs were assessed using FASTQC.

Library	Raw Read Count (x 10^6)	Trimmed Read Count (x 10^6)	% Reads lost
June 18 P	107	104	2.9
June 18 U	115	111	3.2
June 18 P+I	113	110	2.4
June 18 U+I	103	101	2
June 21 P	173	168	3
June 21 U	147	143	3
June 21 P+I	80	79	2
June 21 U+I	76	74	2.4
June 25 P	83	82	1.9
June 25 U	109	106	2.8
June 25 P+I	89	87	2
June 25 U+I	29	28	2.5
June 28 P	71	69	2.8
June 28 U	91	89	2
June 28 P+I	76	74	2.9
June 28 U+I	150	147	2.1
July 3 P	118	114	3.1
July 3 P tr	98	96	2.7
July 3 U	111	109	1.9
July 3 P+I	82	80	2.5
July 3 U+I	69	67	2.6
July 6 P	103	100	2.8
July 6 U	64	61	3.6
July 6 U tr	113	110	2.4
July 6 P+I	72	70	2.8
July 6 U+I	113	110	2
July 6 Integ	102	100	1.7
July 6 OSBC	97	95	2.3
July 10 P	123	117	4.8
July 10 U	122	118	3
July 10 P+I	106	103	3.3
July 10 U+I	90	86	5.3
July 10 U+I tr	123	119	3.6
Total	3,317	3,226	2.8

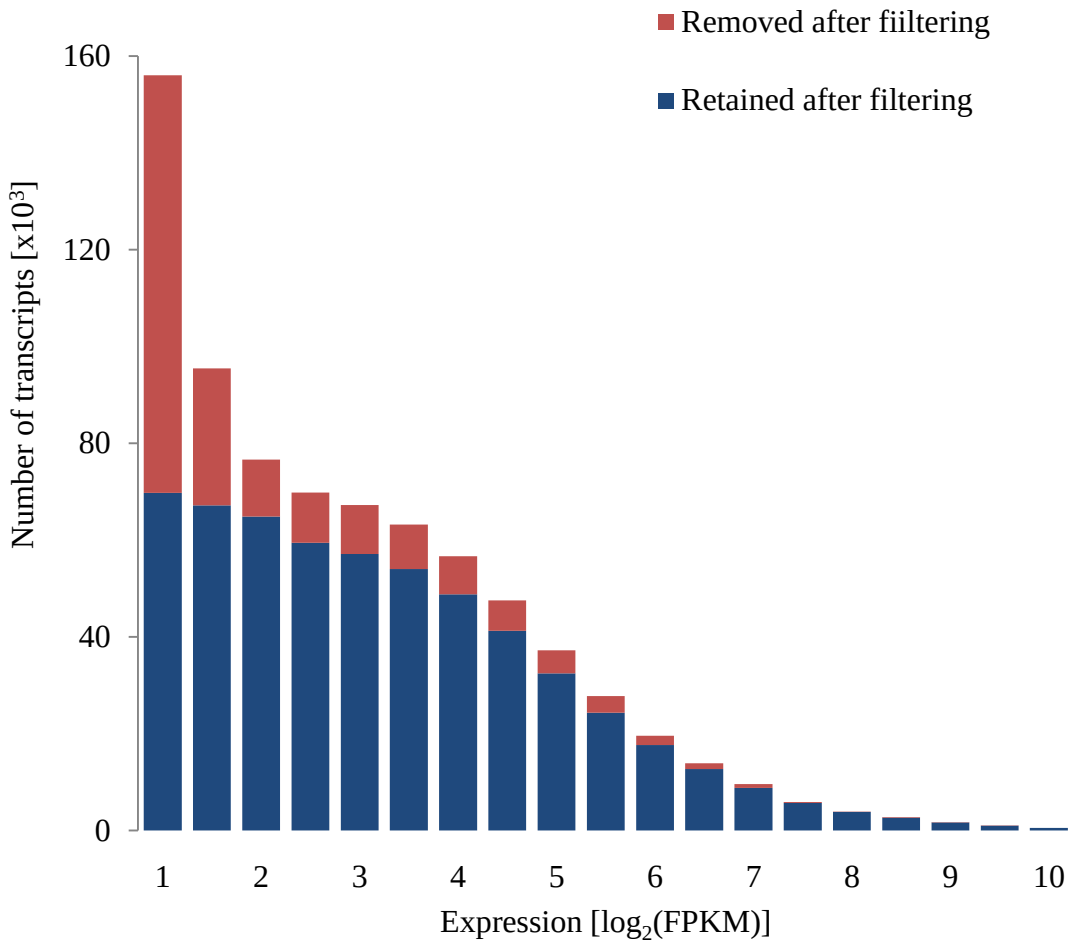


Figure 15. Histogram showing the total number of transcripts in the Trinity assembly across $\log_2$ $FPKM_1$ expression levels. Red bars represent the fraction of transcripts removed by filtering, using a floating fold change algorithm, where FC > FCcut for all samples across a transcript. Transcripts retained after filtering are represented by blue bars, where $n_{retained}$ = 28,182.

Filtering transcripts from the Trinity assembly using the floating fold change algorithm resulted in a total of 28,182 transcripts, of which over half (15,535) had sequence similarities to plants and 10% (2,715) had sequence similarities to insects (Figure 16). Very few of the filtered genes had BLAST hits to contaminating organisms; less than 1% had sequences similar to bacteria, fungi, or viruses (Figure 16). Approximately a third of the filtered transcripts did not have sequence similarity to any organism in the databases/genomes blasted against. None of the unknown transcripts had blast hits to *P. taeda*. These transcripts may be unique to both *M. spermotrophus* and Douglas-fir, which lack sequenced genomes. However, the majority of these transcripts may be artefacts of assembly compilation.

Of the 15,535, Douglas-fir transcripts, 14,010 had sequence similarity to genes from the *P. taeda* geome. Plant transcripts were further reduced to include ones that had a high likelihood of representing genes by including only those with a minimal open reading of 300 bp or more. A total of 11,730 plant transcripts remained, of which 10,826 had blast hits to a *P. taeda* sequence.

3.3.2.3 Insect annotated transcripts

Approximately 2,700 transcripts with sequence similarity to insects passed post-assembly filtering. Of these transcripts, 2,458 had blast hits to *N. vitripennis*. Overall, transcripts identified as insect genes were uniquely clustered from transcripts identified as plant genes (Figure 17).

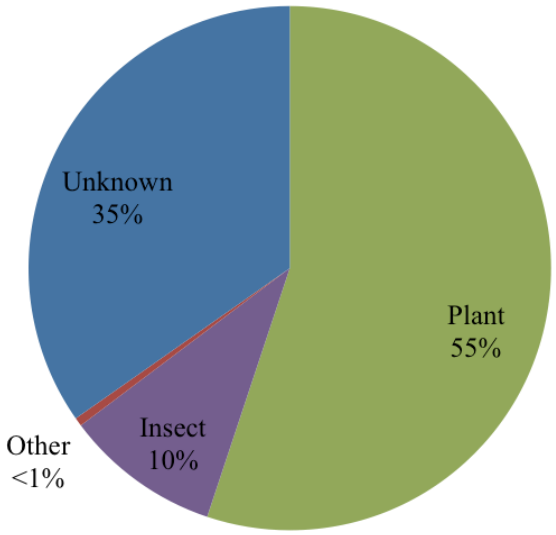


Figure 16. Pie chart showing the taxonomic distribution of transcripts after filtering out lowly expressed transcripts from the original assembly (n=28,182).

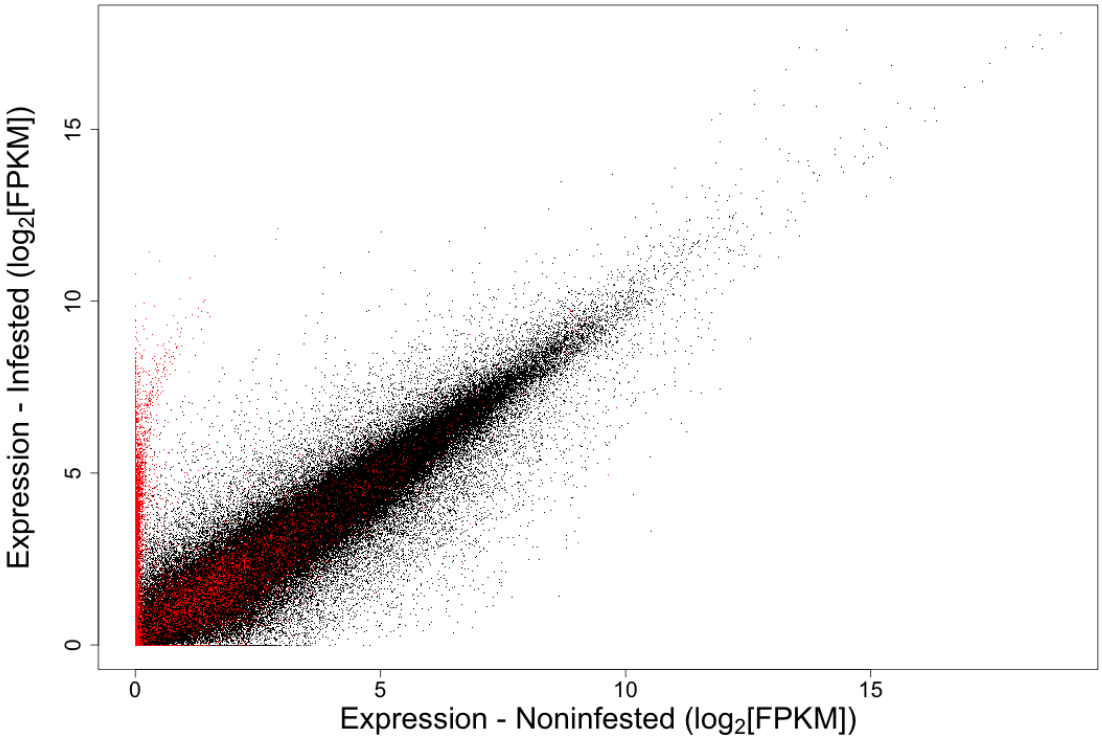


Figure 17. Scatterplot showing transcript expression ($\log_2$ $FPKM_1$) of infested and uninfested megagametophytes by *M. spermotrophus.* Transcripts with plant annotations are illustrated in black. Transcripts with insect annotations are illustrated in red. The total number of expression points for plant and insect transcripts are 434,908 and 76,020, respectively.

However, some transcripts identified as insect genes were expressed in non-infested samples (Figure 17). Of the insect genes there were considerably more log_2FPKM_1 expression points (28,499) unique to infested samples, compared to log_2FPKM_1 expression points found in uninfested samples that overlapped with plant gene expression (9,511). About half (4,783) of the expression points of insect genes overlapping with plant gene expression had log_2FPKM_1 values less than 0.1. These points may be a result of error in RSEM expression, or error due to non-normalization of GC content between sequenced lanes. The remaining insect expression points that overlapped with plant expression may represent missannotation or highly conserved genes among plants and insects.

Regardless of the marginal overlap between plant and insect gene expression, the sum of all insect transcripts was larger and increased over the sampling period for infested samples, compared to uninfested samples. The latter had consistently low gene expression across all sampling dates (Figure 18). Similar to insect transcripts, a proportion of plant transcripts were likely missannotations or expression levels were falsely enriched. These false positives likely represent the expression values of uninfested samples seen in Figure 18.

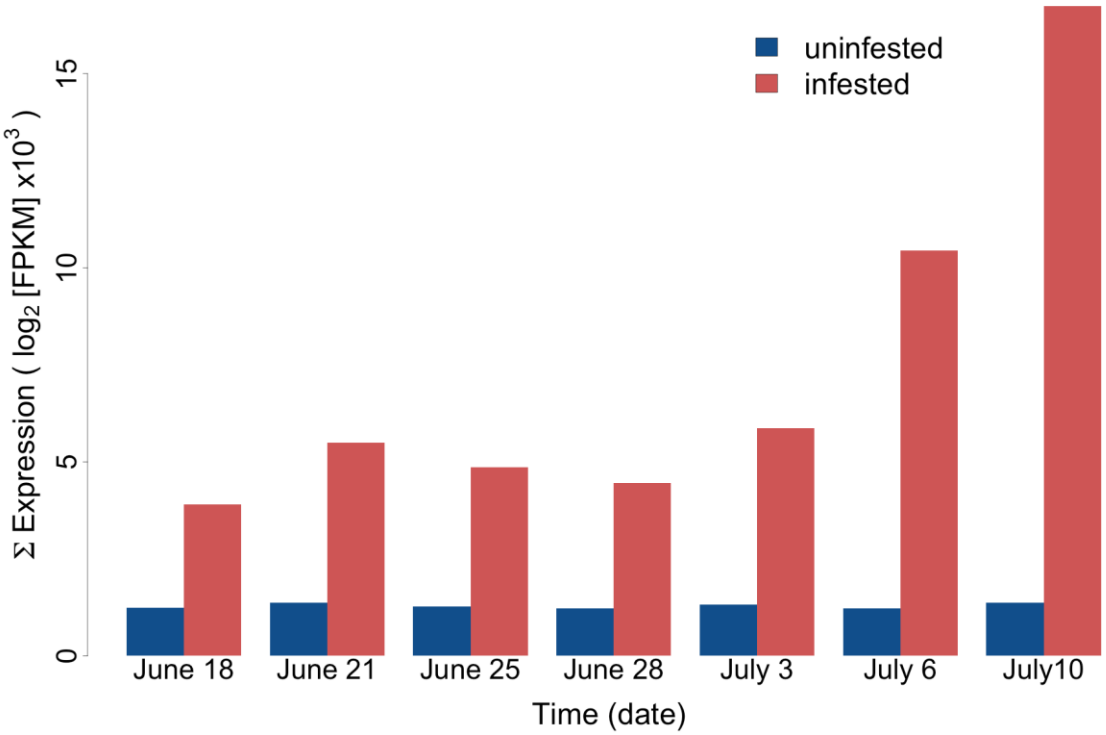


Figure 18. Histogram of transcripts identified as insect origin, showing the sum of transcript expressions (log_2 $FPKM_1$) at each sampling date for megagametophytes that have been infested (red bars) and uninfested (blue bars) by *M. spermotrophus*. Uninfested expression values are likely false positives or represent missannotations.

3.3.2.4 Quadratic regression of plant annotated transcripts

Quadratic regression analysis identified 10,027 plant transcripts out of 11,730 that significantly fit a quadratic pattern. The majority of transcripts were either up- or down-regulated with the greatest number following a linear trend (Figures 19, 20, 21). Relatively fewer transcripts were transiently expressed compared to transcripts that were up- or down-regulated between megagametophyte sample types (Figure 21). However, pollinated-infested and unpollinated-infested accounted for the majority of the transiently expressed transcripts, specifically transiently-upregulated (Figure 21).

Among transcripts that fit a quadratic pattern, 1,610 had similar expression patterns between all megagametophyte sample types (Figure 22). Unpollinated megagametophytes had the greatest number of uniquely expressed transcripts (2926). Infested treatments (PI and UI) shared 722 transcripts with similar expression profiles. These represent all the plant transcripts within the megagametophyte that are responding to *M. spermotrophus* infestation. If there is a plant defense response to *M. spermotrophus*, they are likely to be represented among these transcripts.
P and PI treatments had 694 transcripts with similar expression profiles, whereas U and UI had only 158 transcripts with similar expression patterns (Figure 22).
There are more transcripts similarly expressed between PI, UI, and P (905) than between PI, UI and U (283).

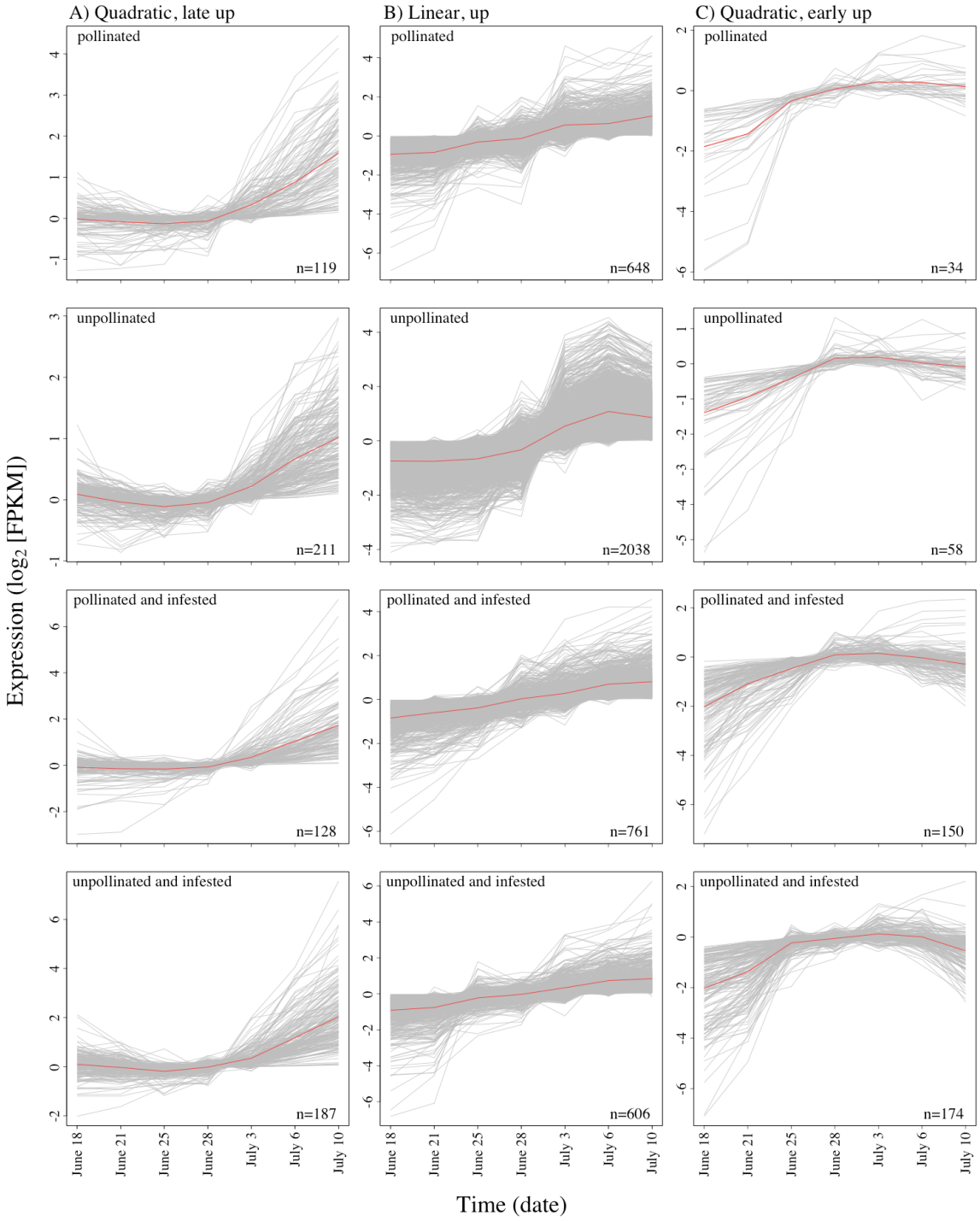


Figure 19. Quadratic regression of differentially expressed (mean centered $\log_2[\text{FPKM}]_1$) transcripts up-regulated (categories A-C) during megagametophyte development in P, U, PI and UI samples across 7 sampling dates. Average expression level of transcripts within each category is shown in red, whereas individual transcripts are shown in grey.

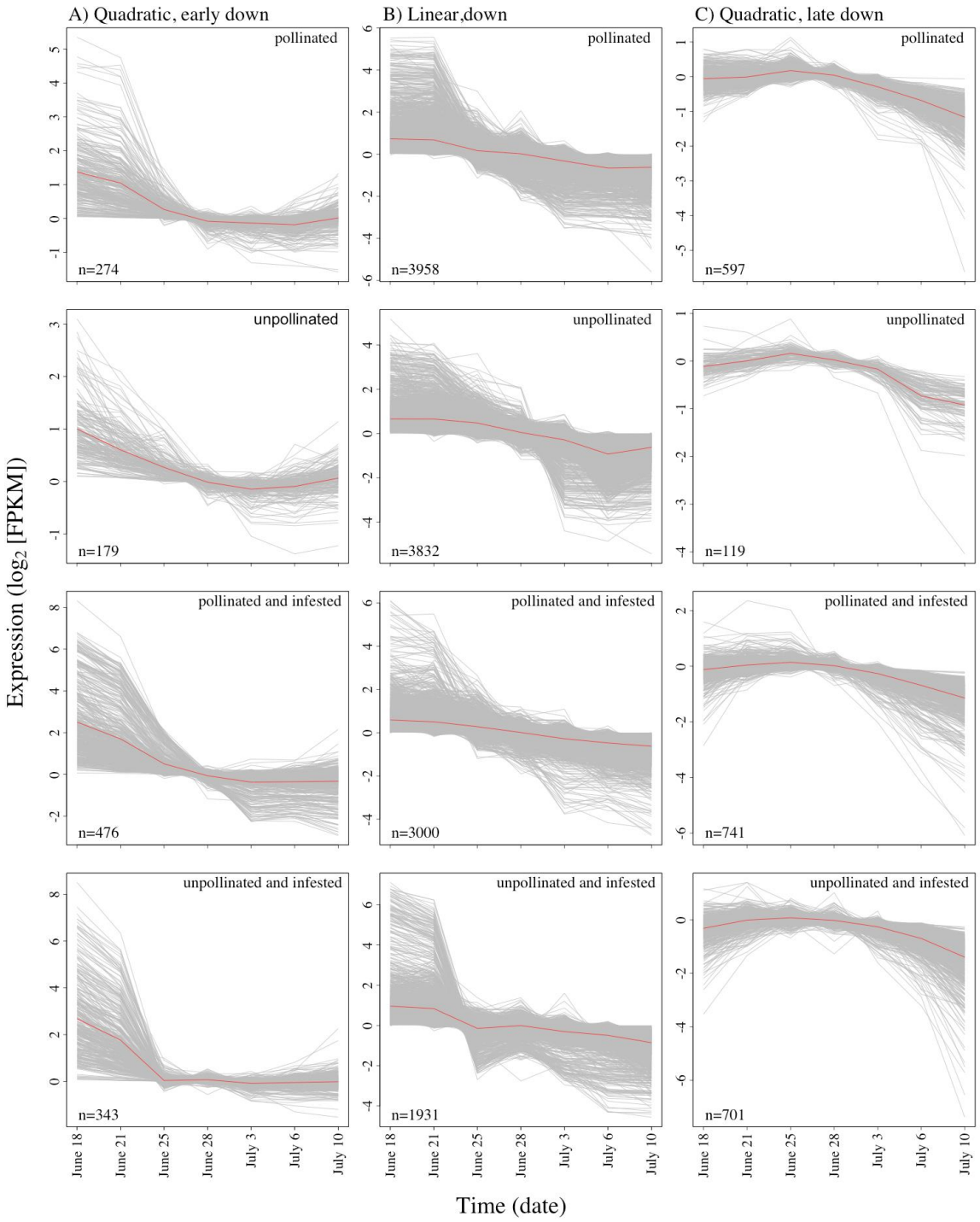


Figure 20. Quadratic regression of differentially expressed (mean centered $\log_2[\text{FPKM}]_1$) transcripts down-regulated (categories A-C) during megagametophyte development in P, U, PI and UI samples across 7 sampling dates. Average expression level of transcripts within each category is shown in red, whereas individual transcripts are shown in grey.

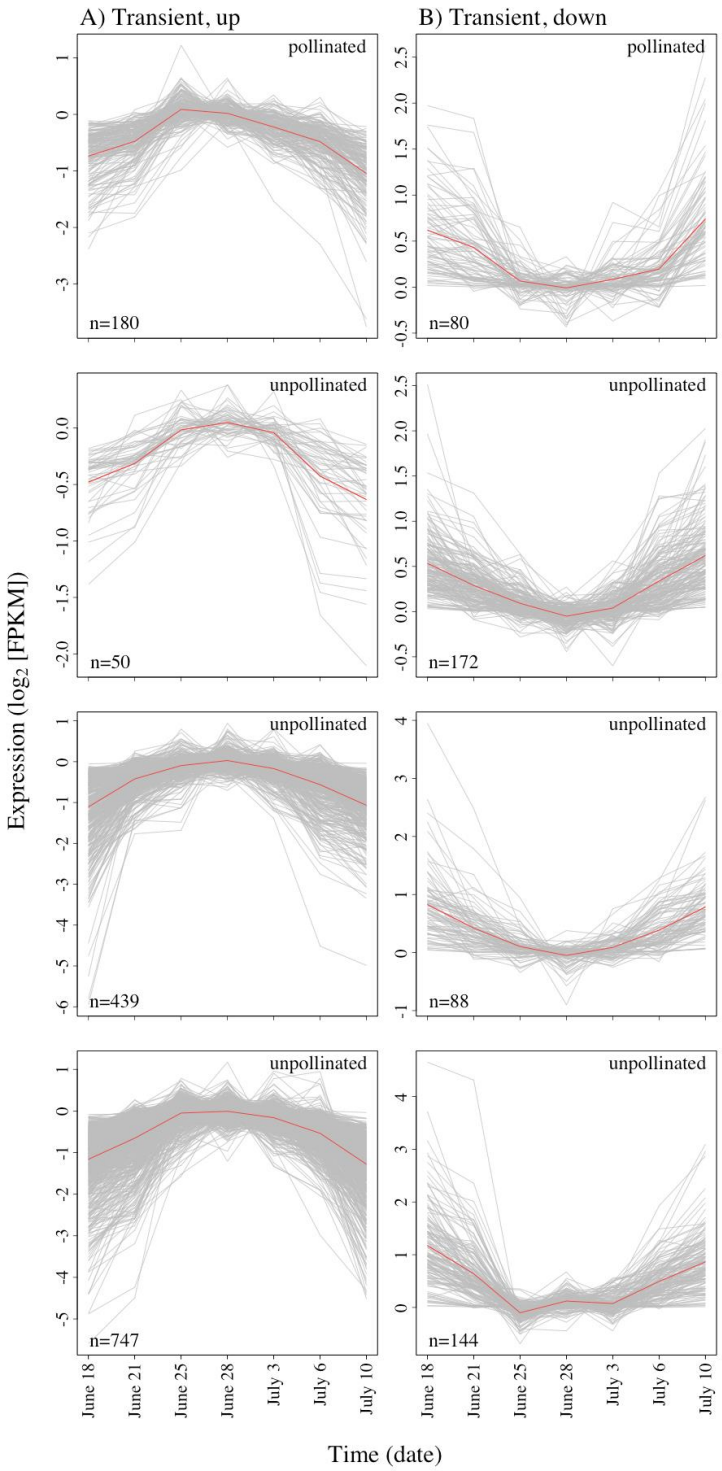


Figure 21. Quadratic regression of differentially expressed (mean centered $\log_2[\text{FPKM}]_1$) transcripts transiently expressed (categories A and B) during megagametophyte development in P, U, PI and UI samples across 7 sampling dates. Average expression level of transcripts within each category is shown in red, whereas individual transcripts are shown in grey.

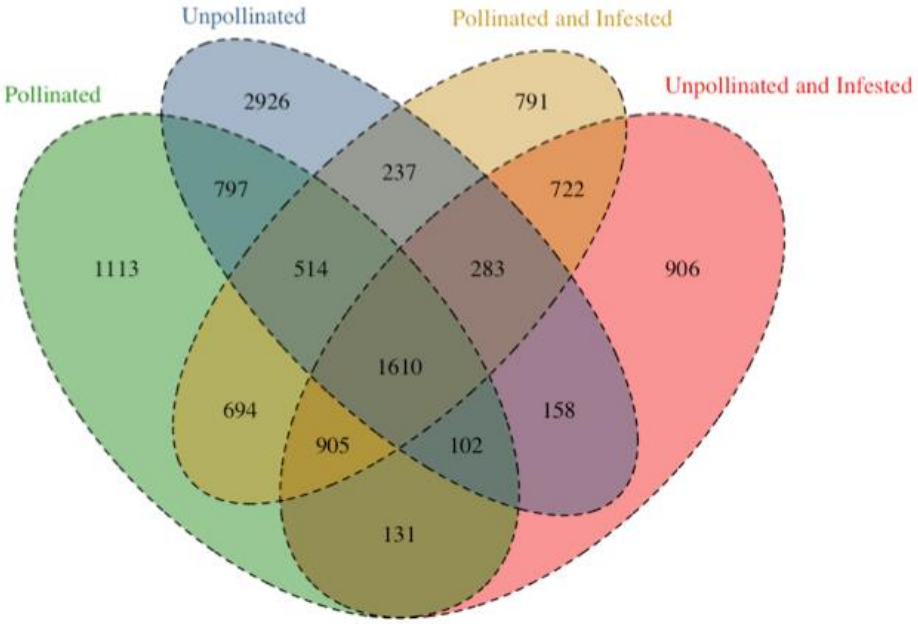


Figure 22. Venn diagram of differentially expressed plant transcripts between P, U, PI and UI megagametophytes. Transcripts were included if they significantly fit into a quadratic regression pattern at the $\alpha = 0.05$ level. Transcripts were considered to be similar between treatments if they shared expression patterns of being up- down- or transiently expressed.

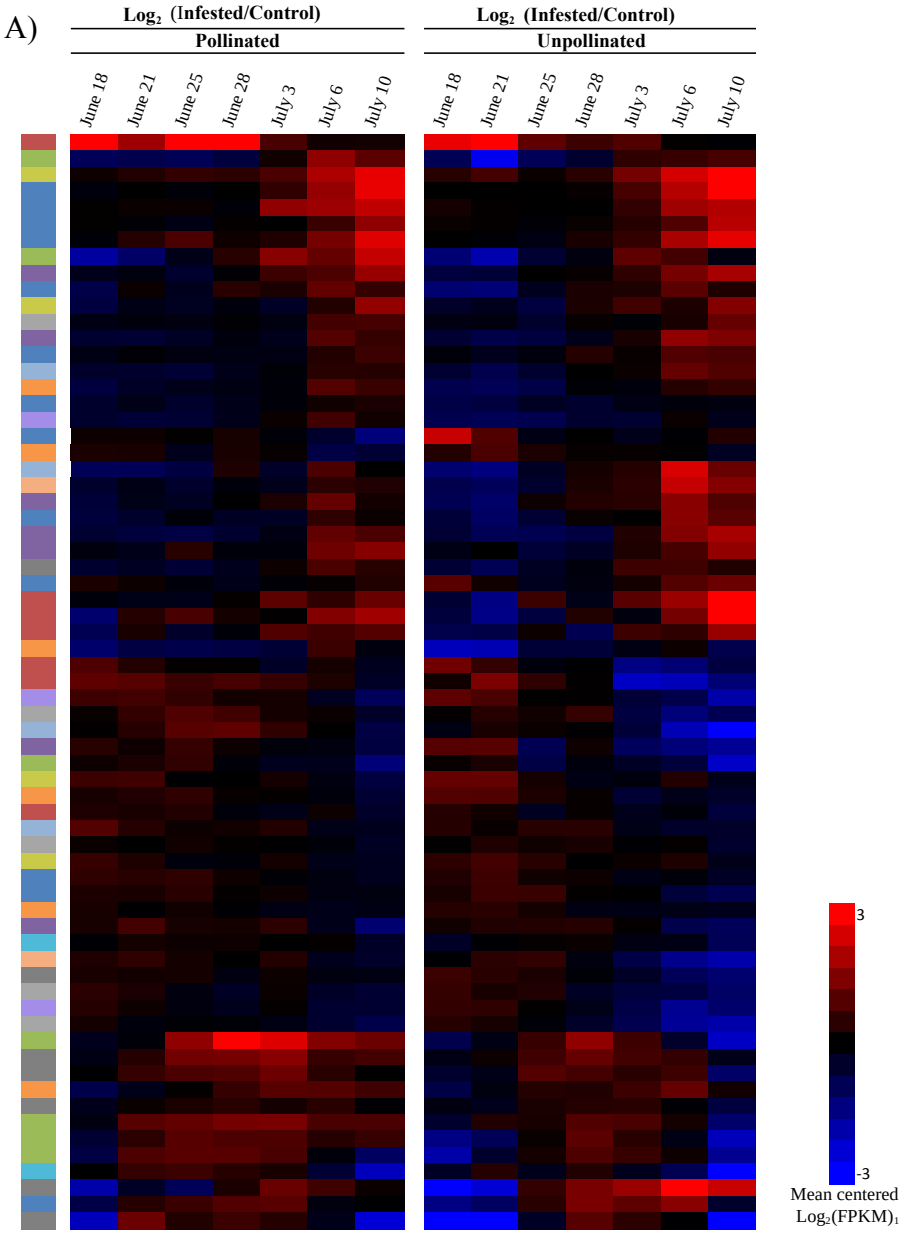
Since there are few genes similarly expressed between UI and U, a number of genes may be manipulated by *M. spermotrophus*. If there are genes manipulated by the insect, they will likely have similar expression patterns to the P treatment and will be represented in the 905 transcripts with unique expression in PI, UI and P and the 2926 transcripts unique to the U treatment.

3.3.2.5 Candidate gene search

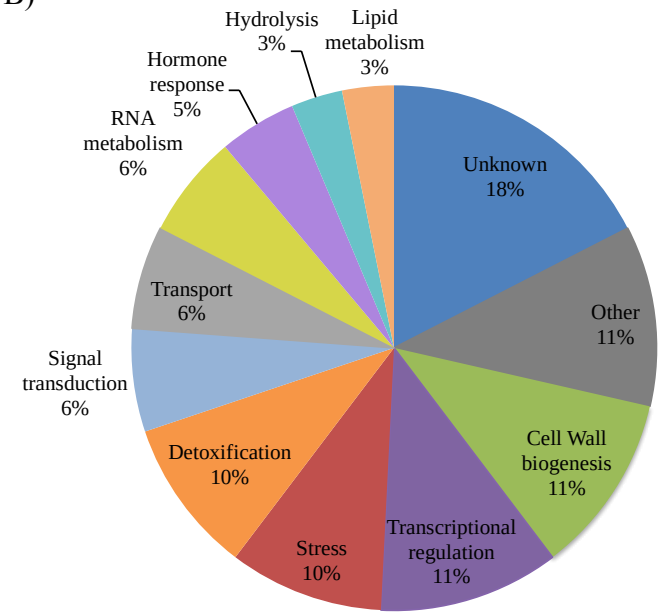
There were a total of 89 transcripts responding to *M. spermotrophus* infestation that were highly differentially expressed, of which 68 were identified in the *P. taeda* genome (Figure S1). Transcripts differentially expressed in PI and UI compared to P and U, respectively were down-regulated, up-regulated and some transiently expressed (Figure 23A). Many of these transcripts had functions related to a diverse set of biological roles. Prominent functional groupings included cell biosynthesis and remodeling, transcriptional regulation, stress response, and detoxification (Figure 23B). However, the majority of transcripts responding to *M. spermotrophus* had an unknown function. There were a few notable transcripts involved in stress and hormone production with interesting expression patterns between infested and non-infested megagametophytes. Putative roles of these transcripts will be discussed.

Figure 23. A. Putative Douglas-fir transcripts responding to *M. spermotrophus* infestation. Heat map illustrating differentially expressed transcripts between P, U, PI and UI treatments. Only transcripts with a minimal of 8-fold ($\Delta\log_2 > 3$) higher expression in PI and UI megagametophyte samples compared to P and U samples, respectively, were included. Transcripts represent $\log_2[\text{FPKM}]_1$ fold change of infested treatments over their respective control. Transcripts with annotation hits to both plants and insects were removed. B. Functional groups of transcripts responding to *M. spermotrophus* infestation.

A)


Log2 (Infested/Control)
Pollinated
Log2 (Infested/Control)
Unpollinated
June 18
June 21
June 25
June 28
July 3
July 6
July 10
3
-3
Mean centered
Log2(FPKM)i

B)

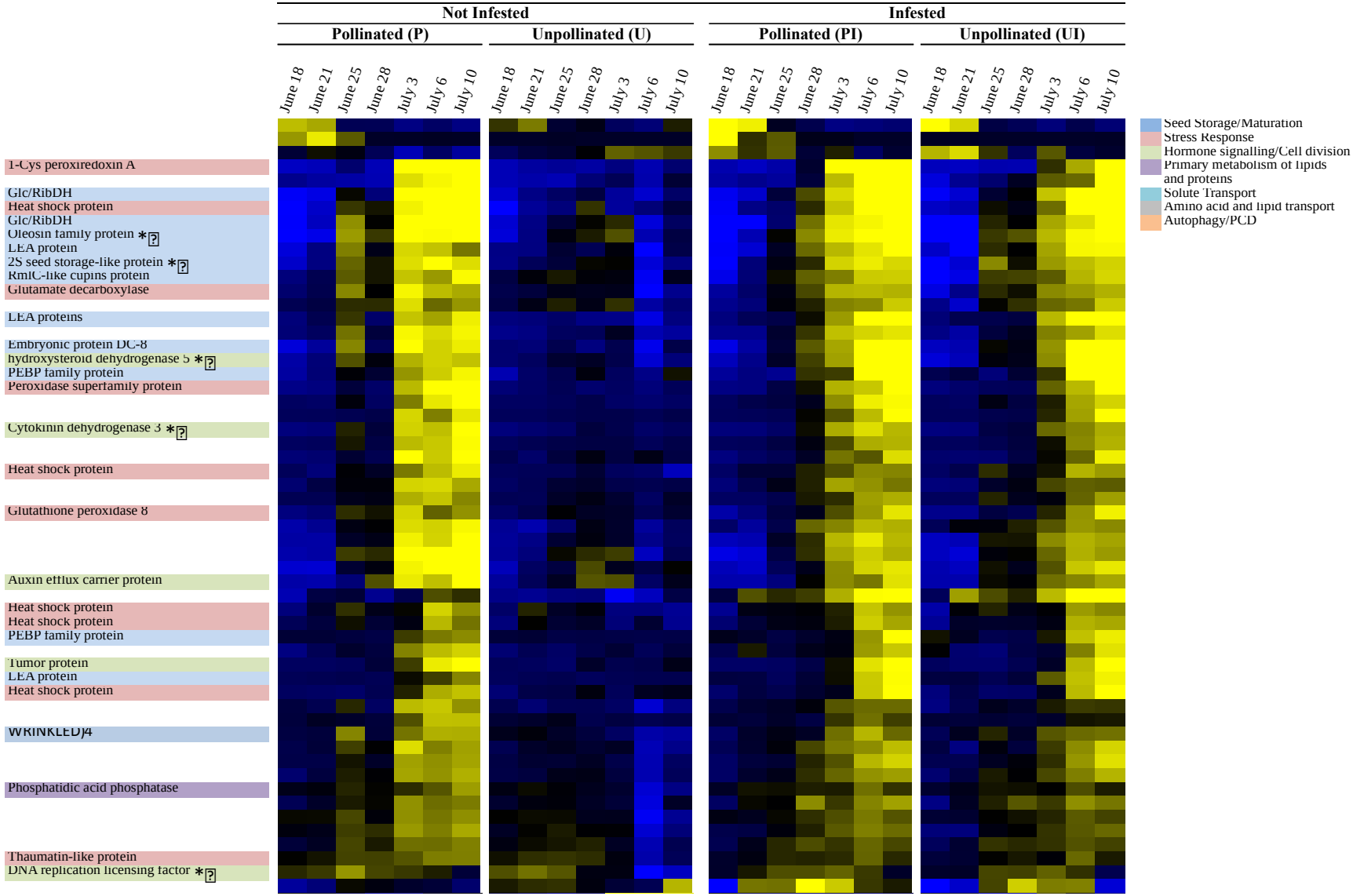

Unknown 18%
Other 11%
Cell Wall biogenesis 11%
Transcriptional regulation 11%
Stress 10%
Detoxification 10%
Signal transduction 6%
Transport 6%
RNA metabolism 6%
Hormone response 5%
Hydrolysis 3%
Lipid metabolism 3%

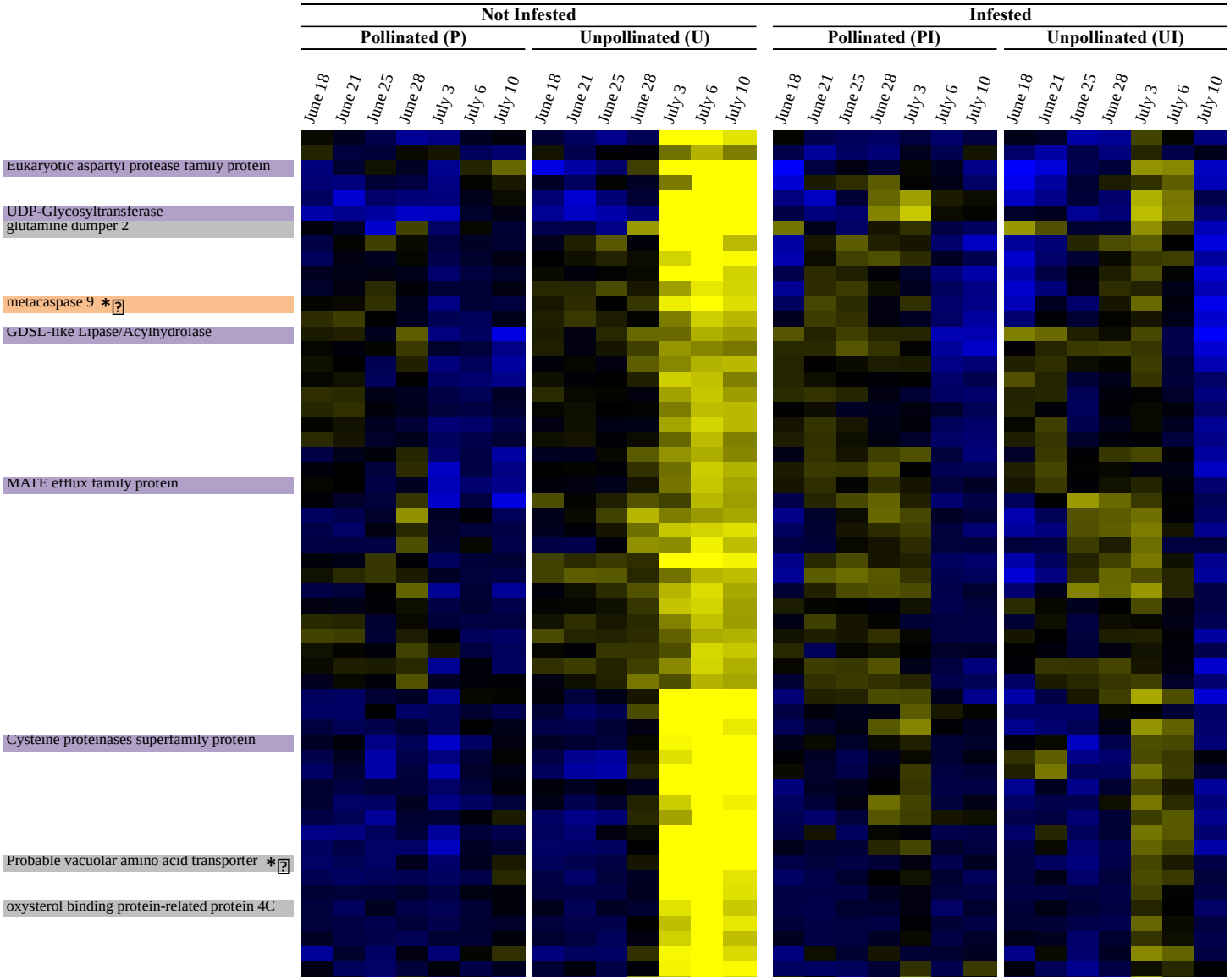
Three hundred highly differentially expressed transcripts involved in megagametophyte development were influenced by *M. spermotrophus* infestation (Figure 24A). Two hundred and fifty-four of the three hundred transcripts were identified in the *P. taeda* genome (Figure S1). The majority of these transcripts (234) were upregulated late in the U treatment with no obvious expression pattern in P, PI and UI treatments (Figure 24A). Transcripts with distinct expression in the U treatment were diverse in function, predominantly related to primary and secondary metabolism, solute transport, stress response and cell wall biosynthesis and remodeling (Figure24B). A smaller fraction of the transcripts expressed in the U treatment were involved in detoxification, amino acid transport, programmed cell death / autophagy, lipid transport and seed maturation (Figure 24B).

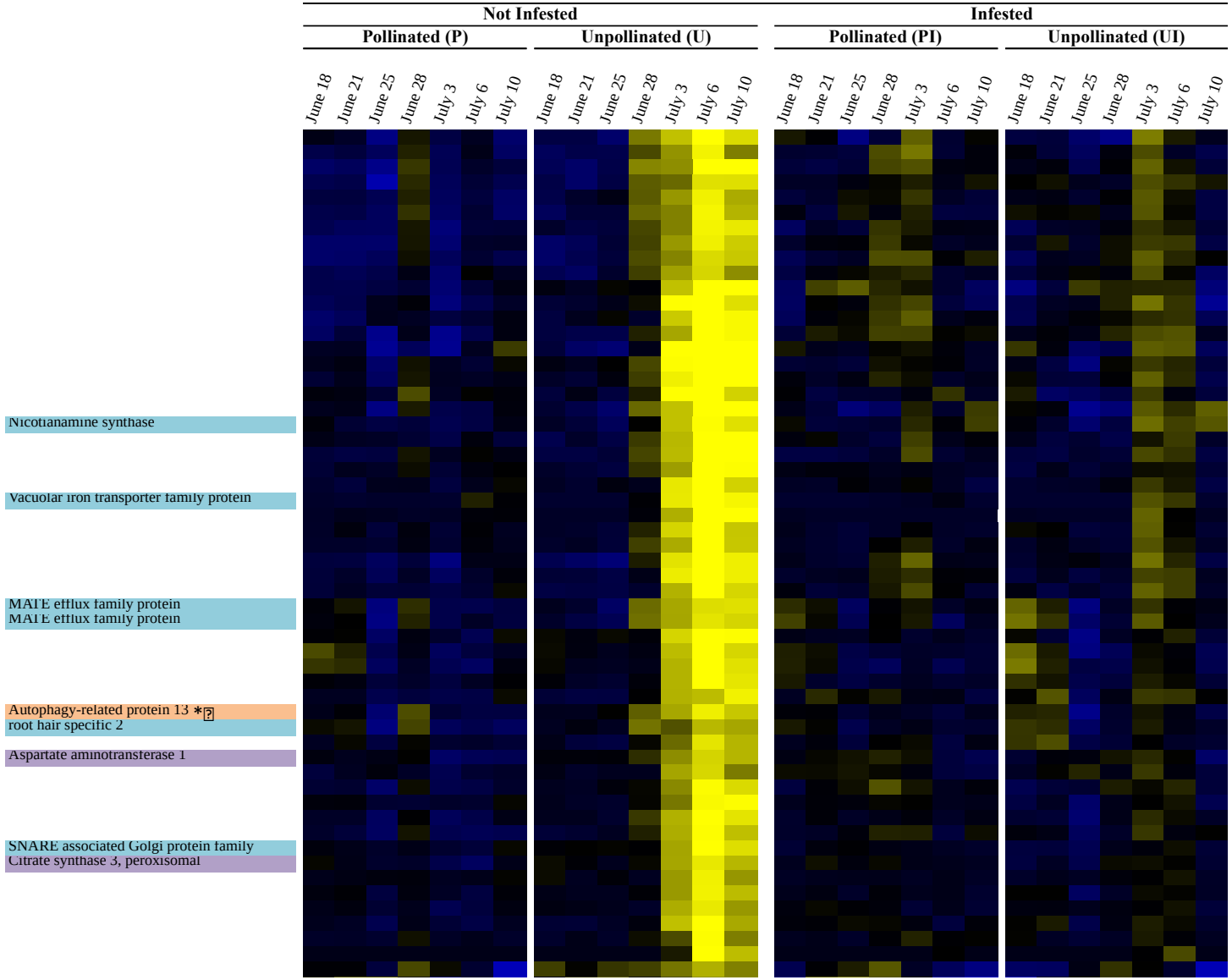
Of the three hundred transcripts manipulated by *M. spermotrophus*, 56 transcripts were identified to be involved in normal seed development of pollinated ovules. The majority of these transcripts had low expression in unpollinated samples but were similarly upregulated late in P, PI and UI samples (Figure 24A). Most of these transcripts are induced by June 28th. The largest portion of these transcripts were functionally characterized as putative seed storage / maturation genes, while the remainder were putatively involved in stress response, cell division, cell wall biogenesis and transcriptional regulation (Figure 24C).

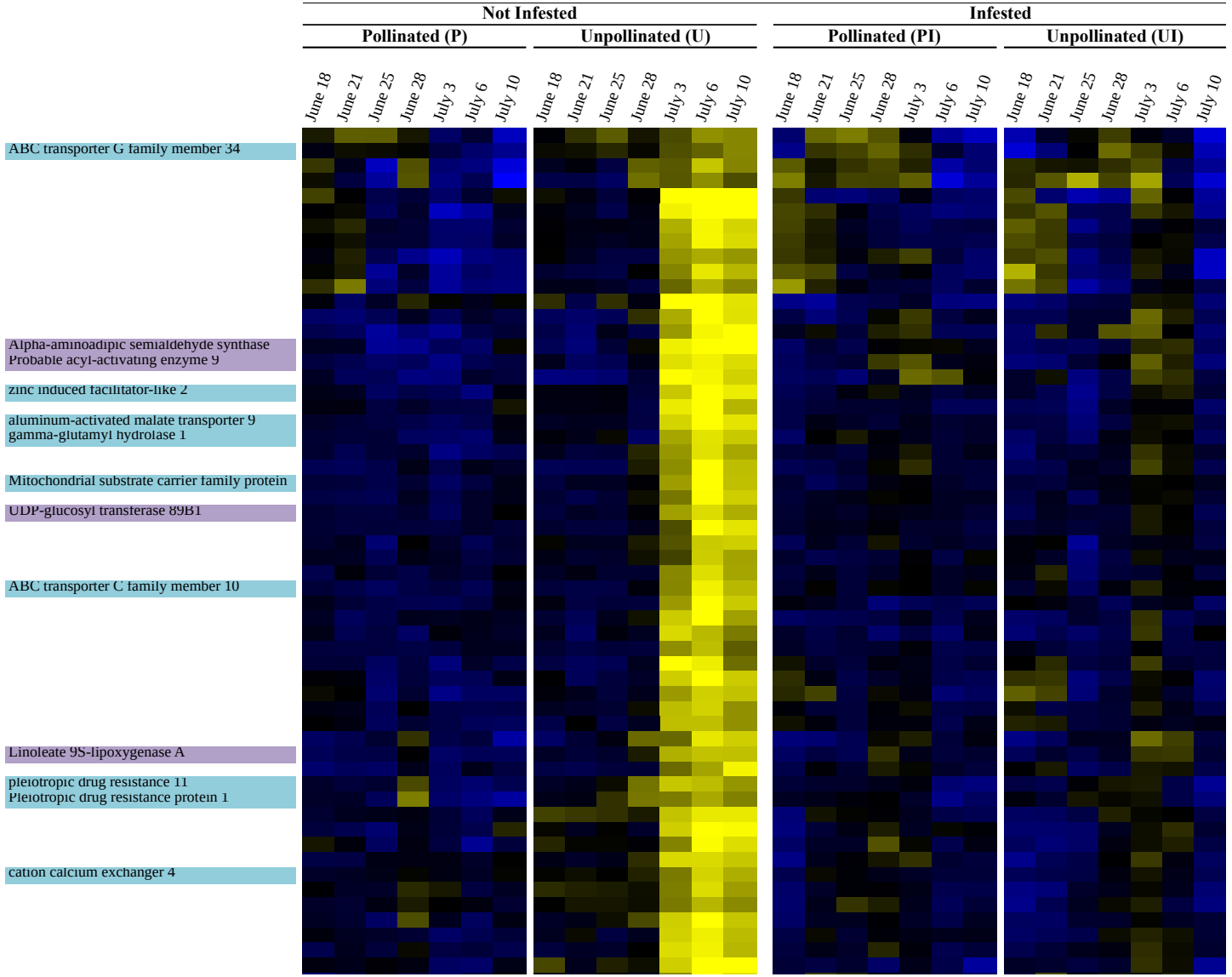
Figure 24. A. Putative Douglas-fir transcripts responding to *M. spermotrophus* infestation. Heat map illustrating differentially expressed transcripts between P, U, PI and UI treatments. Only transcripts with a minimal of 8-fold ($\Delta\log_2 > 3$) higher expression in PI and UI megagametophyte samples compared to P and U samples, respectively, were included. Additionally selected transcripts had to significantly fit a quadratic pattern at $\alpha = 0.05$. Transcripts represent $\log_2[\text{FPKM}]_1$ fold change of infested treatments over their respective control. Transcripts were manually assigned to functional groups for those with similarly recognized patterns in P, PI and UI treatments (B) and for transcripts with unique expression patterns to the U treatment (C). Transcripts with an asterisk will be discussed in further detail.

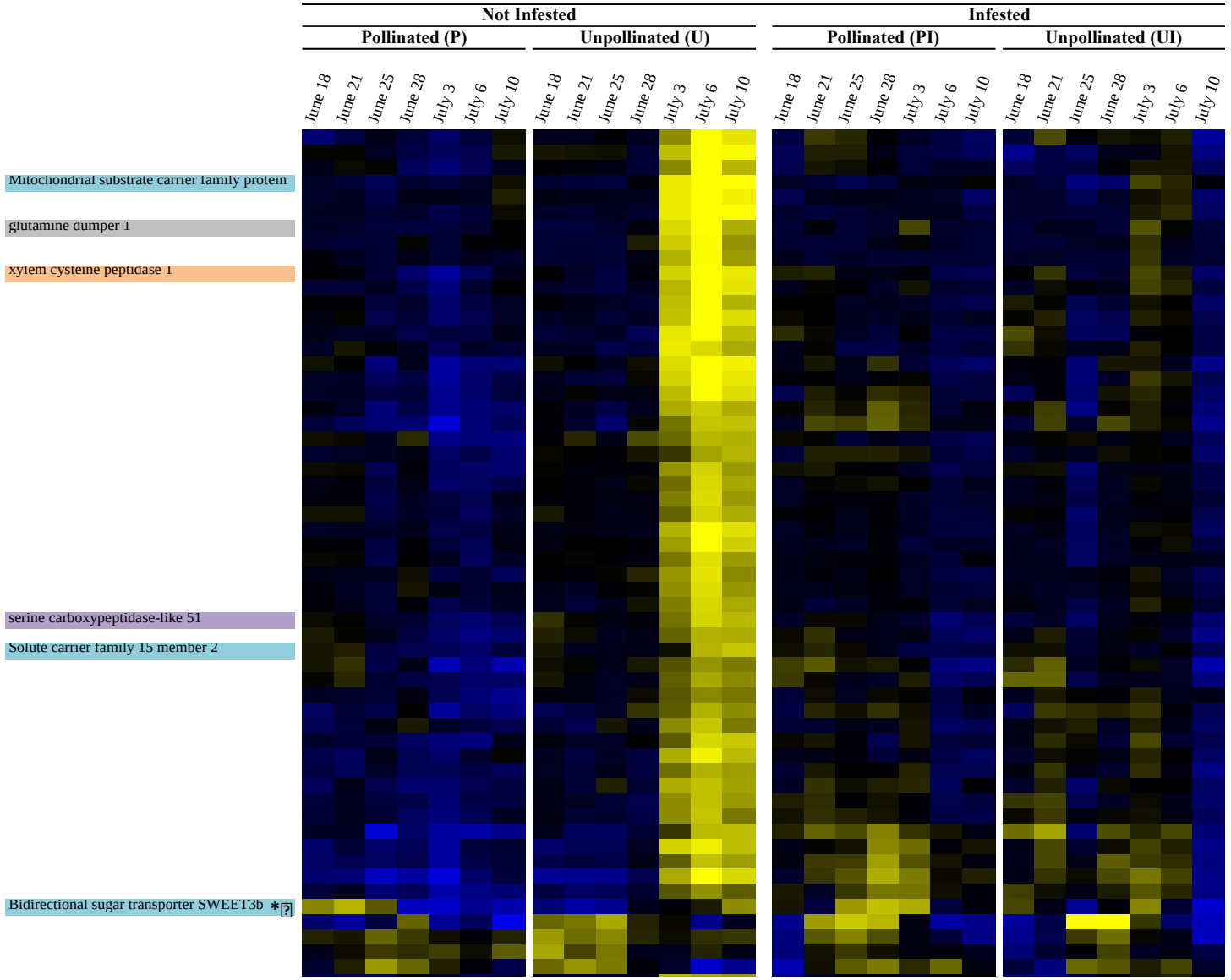
A.

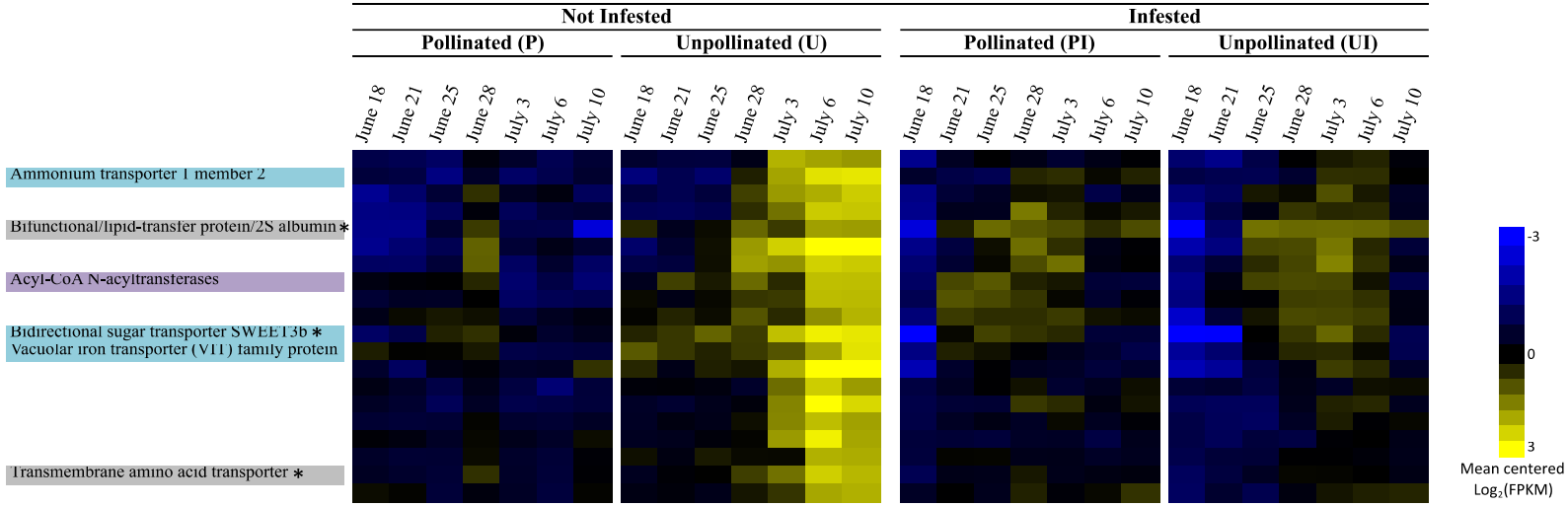

Not Infested
Infested
Pollinated (P)
Unpollinated (U)
Pollinated (PI)
Unpollinated (UI)
June 18
June 21
June 25
June 28
July 3
July 6
July 10
1-Cys peroxiredoxin A
Glc/RibDH
Heat shock protein
Glc/RibDH
Oleosin family protein *?
LEA protein
2S seed storage-like protein *?
RmlC-like cupins protein
Glutamate decarboxylase
LEA proteins
Embryonic protein DC-8
hydroxysteroid dehydrogenase 5 *?
PEBP family protein
Peroxidase superfamily protein
Cytokinin dehydrogenase 3 *?
Heat shock protein
Glutathione peroxidase 8
Auxin efflux carrier protein
Heat shock protein
Heat shock protein
PEBP family protein
Tumor protein
LEA protein
Heat shock protein
WRINKLED4
Phosphatidic acid phosphatase
Thaumatin-like protein
DNA replication licensing factor *?
Seed Storage/Maturation
Stress Response
Hormone signalling/Cell division
Primary metabolism of lipids and proteins
Solute Transport
Amino acid and lipid transport
Autophagy/PCD


Not Infested
Infested
Pollinated (P)
Unpollinated (U)
Pollinated (PI)
Unpollinated (UI)
June 18
June 21
June 25
June 28
July 3
July 6
July 10
Eukaryotic aspartyl protease family protein
UDP-Glycosyltransferase
glutamine dumper 2
metacaspase 9 *?
GDSL-like Lipase/Acylhydrolase
MATE efflux family protein
Cysteine proteinases superfamily protein
Probable vacuolar amino acid transporter *?
oxysterol binding protein-related protein 4C

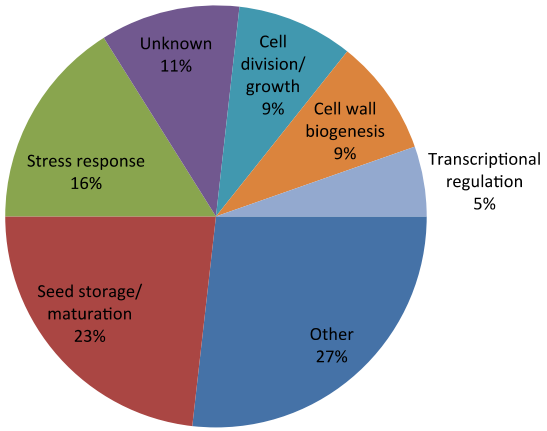

Not Infested
Infested
Pollinated (P)
Unpollinated (U)
Pollinated (PI)
Unpollinated (UI)
June 18
June 21
June 25
June 28
July 3
July 6
July 10
Nicotianamine synthase
Vacuolar iron transporter family protein
MATE efflux family protein
MATE efflux family protein
Autophagy-related protein 13 *
root hair specific 2
Aspartate aminotransferase 1
SNARE associated Golgi protein family
Citrate synthase 3, peroxisomal


Not Infested
Infested
Pollinated (P)
Unpollinated (U)
Pollinated (PI)
Unpollinated (UI)
June 18
June 21
June 25
June 28
July 3
July 6
July 10
ABC transporter G family member 34
Alpha-aminoadipic semialdehyde synthase
Probable acyl-activating enzyme 9
zinc induced facilitator-like 2
aluminum-activated malate transporter 9
gamma-glutamyl hydrolase 1
Mitochondrial substrate carrier family protein
UDP-glucosyl transferase 89B1
ABC transporter C family member 10
Linoleate 9S-lipoxygenase A
pleiotropic drug resistance 11
Pleiotropic drug resistance protein 1
cation calcium exchanger 4

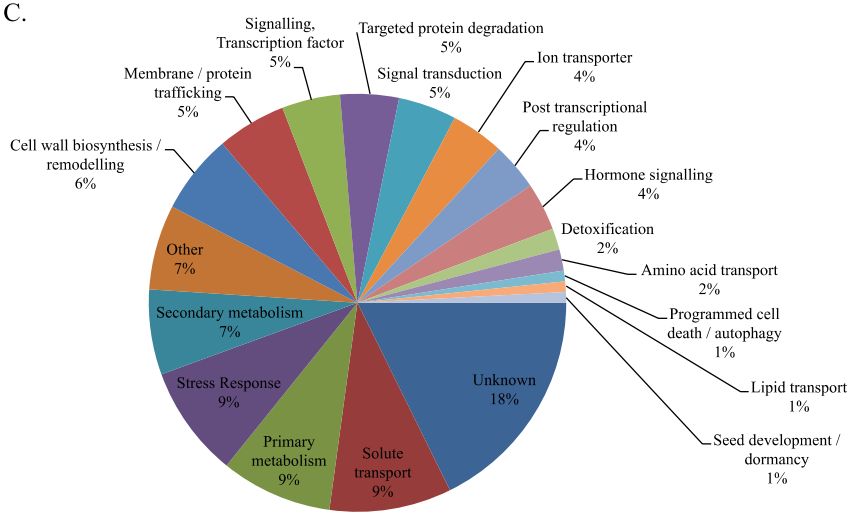

Not Infested
Infested
Pollinated (P)
Unpollinated (U)
Pollinated (PI)
Unpollinated (UI)
June 18
June 21
June 25
June 28
July 3
July 6
July 10
Mitochondrial substrate carrier family protein
glutamine dumper 1
xylem cysteine peptidase 1
serine carboxypeptidase-like 51
Solute carrier family 15 member 2
Bidirectional sugar transporter SWEET3b *?


Not Infested
Infested
Pollinated (P)
Unpollinated (U)
Pollinated (PI)
Unpollinated (UI)
June 18
June 21
June 25
June 28
July 3
July 6
July 10
Ammonium transporter 1 member 2
Bifunctional/lipid-transfer protein/2S albumin *
Acyl-CoA N-acyltransferases
Bidirectional sugar transporter SWEET13b *
Vacuolar iron transporter (VIT) family protein
Transmembrane amino acid transporter *
-3
0
3
Mean centered
Log2(FPKM)

B.


Unknown
11%
Cell
division/
growth
9%
Cell wall
biogenesis
9%
Transcriptional
regulation
5%
Stress response
16%
Seed storage/
maturation
23%
Other
27%

C.


Signalling,
Transcription factor
5%
Targeted protein degradation
5%
Membrane / protein
trafficking
5%
Signal transduction
5%
Ion transporter
4%
Post transcriptional
regulation
4%
Cell wall biosynthesis /
remodelling
6%
Hormone signalling
4%
Detoxification
2%
Other
7%
Amino acid transport
2%
Programmed cell
death / autophagy
1%
Secondary metabolism
7%
Lipid transport
1%
Stress Response
9%
Unknown
18%
Seed development /
dormancy
1%
Primary
metabolism
9%
Solute
transport
9%

3.4 Discussion

3.4.1 Douglas-fir response to *M. spermotrophus*

Sixty-seven Douglas-fir transcripts were identified as highly differentially expressed between infested and non-infested megagametophytes. This result suggests that Douglas-fir responds to *M. spermotrophus* infestation. In many plant-insect interactions, both egg deposition and herbivory can rapidly induce host-plant defenses (Gouhier-Darimont et al., 2013; Bonaventure, 2012). Specific response to wounding and herbivory include, but are not limited to emission of volatile chemicals, oxidative bursts, callus deposition, increased production of secondary metabolites and programmed cell death of damaged cells (Hilker et al., 2005; Liu et al., 2010; Bruessow and Reymond 2007; Ralph et al., 2006). Intense induction of genes involved in ethylene and jasmonate pathways in the presence of insects are strong indicators of wound and herbivory signalling in plants (Howe and Jander, 2008). Transcriptional profiling of *Arabidopsis* after wounding revealed a number of transcription factors including AP2, WRKY and MYB that are induced upon JA and ethylene signalling (Cheong et al., 2002).

My dataset of highly differentially expressed transcripts between infested and non-infested samples did not identify many genes that are clear indicators of plant defense to herbivory. If Douglas-fir megagametophytes were responding to herbivory by *M. spermotrophus* larvae, I would have expected to see a strong induction of enzymes and transcription factors involved in ethylene and jasmonic acid pathways, for example, lipoxygenase, allene oxide synthase and allene oxide cylcase AP2, and WRKY in the infested treatments. This was not the case. The absence of wound tissue responses in

any sections supports the gene expression work. The dataset, however, revealed a select few transcripts predicted to be involved in hormone signalling. These include ETHYLENE INSENSITIVE 3-like protein (EI3), XERICO and ABI3-interacting protein 2 (AIP2) (comp186517_c1, comp202130_c0, and comp206474_c0, respectively).

AIP2 has been described as a negative regulator of ABA, whereas XERICO is a precursor to ABA synthesis (Zhang et al., 2005; Ko et al., 2006). In my dataset, the transcript encoding XERICO was initially highly expressed at the time of insect oviposition and gradually declined to levels similar to expression found in the pollinated samples. This result is consistent with the ABA trends reported by Chiwocha et al. (2007) for *M. spermotrophus*-infested Douglas-fir seed. Increased levels of ABA have been reported in Douglas-fir megagametophytes during early embryogenesis and in mature seeds entering dormancy (Chiwocha and von Aderkas 2002; Bianco et al., 1997). ABA induction has additionally been shown to be an important hormone mediating both, biotic and abiotic stresses in conifers (De Diego, et al., 2013; Liu et al., 2013). However, stress induction through ABA synthesis has only been studied in non-reproductive tissues of conifers. It is widely accepted that ABA synthesis stimulates protein deposition during megagametophyte development (Misra et al., 1993; Leal et al., 1995; Chatthai and Misra, 1998). Thus, increased levels of ABA in megagametophytes after insect oviposition suggests a possible role in manipulation by the insect to increase the production of proteins.

Ethylene is induced in conifers in response to abiotic stress. Phenolic and terpenoid defenses are mediated by the production of ethylene (Hudgins and Franceschi, 2004). Ethylene precursors have been shown to be upregulated in gymnosperms following wounding (Hudgins et al., 2006). In my dataset EI3 is upregulated in infested seeds compared to unifested seeds. EI3 has been demonstrated to be an essential transcription factor for ethylene signalling in *Arabidopsis* (Chao et al., 1997). Differences in transcription levels between infested and non-infested seeds indicate a potential response to wounding. However, ethylene is an important hormone necessary for many developmental processes including cell differentiation and growth (Dolan, 1997). Perhaps EI3 induction is influencing proliferation of cell development in pollinated megagametophytes. Ethylene may be another important hormone that manipulates megagametophyte fate.

Among genes associated with stress, I identified a gibberellin stimulated-like (GSL2) gene, previously called Snakin-2 (comp122638_c0) (Figure S1). GSLs are cysteine-rich peptides that were first isolated from potato tubers where they were shown to have antimicrobial activity (Berrocal-Lobo et al., 2002). Unlike GSL1, which is constitutively expressed in multiple tissues over the course of development, GSL2 is strongly induced after wounding (Segura et al., 1999; Berrocal-Lobo et al., 2002). Its expression is localized to wound sites. A recent study confirmed that over-expression of GSL2 in potato plants increased resistance to disease (Mohan et al., 2014). In my study, GSL2 is expressed only in insect-infested seeds. GSL2 is highly expressed at early sampling dates when the megagmetophyte has been initially infested by *M. spermotrophus*. This was

followed by an abrupt decline by June 25 and July 3 in UI and PI samples, respectively. Induced expression around the time of insect oviposition suggests a possible response to wounding. However, it is unclear whether the rapid decline in expression of GSL2 is attributed to host manipulation and thus down regulation of gene activity or whether the gene only responds to mechanical wounding and the change in expression marks the time where *M. spermotrophus* may have stopped ovipositing. The latter argument is not consistent for the PI treatment; gene expression remains high until July 3. It is not likely that *M. spermotrophus* are continuing to oviposit past June 28th due to the lignification of the ovuliferous scales, making pentration difficult. This suggests that GSL2 may be responding to more than just wounding, e.g., herbivory or pathogen infection. To determine if GSL2 expression responds in a similar manner to only wounding, it would be necessary to repeat this experiment and include a treatment where megagametophytes are artificially wounded, in the absence of the insect.

An additional transcript with predicted involvment in wounding is annotated as a LATERAL ORGAN BOUNDARIES (LOB) protein (comp216793_c1). In Arabidpsis, LOB transcription factors act downstream of auxin-response factors to control callus formation (Fan et al., 2012). Callus formation is a common repair mechanism at wound sites and has been observed in numerous plant tissues (Ikuechi et al., 2013). Megagametophytes growing *in vitro* on media that have been sliced often produce callus at cut sites (von Adkeras and Bonga, 1988). LOB is not expressed at high levels in pollinated or unpollinated megagametophytes, however it is transiently expressed in infested seeds, with peak expression at June 28th. This suggests a possible role in callus

production, which may be a result of physical wounding i.e., damage of cells from ovisposition and larval feeding.

Wounding and herbivory pathways significantly overlap (Cheong et al., 2002), which makes it difficult to determine whether a plant is responding to mechanical damage or damage induced by insect feeding (Howe and Jander, 2008). There is an increasing body of literature supporting the view that plants can detect insect herbivores in a manner similar to pathogens, i.e. via a mechanism known as herbivore-associated molecular patterns (HAMP) (Mithöfer, and Boland, 2008; Arimura et al., 2011). Several plant species recognize elicitors present in insect oral secretions (Bonaventure, 2012). Since host-specific herbivores often carry unique elicitors, identification of HAMP mediated responses can be difficult to study.

Although there are few transcripts with possible roles in wounding, it appears that Doulgas-fir megagametophytes do not mount a typical herbivory-defense response against *M. spermotrophus* invasion. There is a limited number of transcripts highly differentially expressed between infested and uninfested megagametophytes, but those typically seen to be induced by herbivory are not among them.

3.4.2 Douglas-fir transcripts manipulated upon infection by *M. spermotrophus*

A diverse set of two-hundred and ninety-nine Douglas-fir transcripts from the Douglas-fir/*M. spermotrophus* transcriptome were found to be highly differentially expressed in P, PI and UI treatments compared to the U treatment. These data clearly illustrate that a

large number of putative Douglas-fir transcripts have been manipulated by *M. spermotrophus*. Interestingly, I found more transcripts with similar expression profiles between P, PI and UI megagametophytes than between U, PI and UI megagametophytes. This suggests that at the gene level infested treatments, regardless of pollination, behave more like pollinated than unpollinated megagametophytes. Previously, manipulation of megagametophyte development by *M. spermotrophus* was demonstrated physiologically. Accumulation of storage products and hormone concentrations were maintained in infested megagametophytes to levels found in normal developing ovules (von Aderkas et al., 2005b; Chiwocha et al., 2007). Therefore I expected to find a number of transcripts related to seed storage and hormone signalling. The following section will highlight transcripts of particular interest whose expression was manipulated in insect infested seed to levels found in pollinated seed.

Transcripts with similar expression patterns between P, PI and UI treatments -
Most of the identified transcripts with similar expression in P, PI and UI treatments were not expressed in megagametophytes until either June 25 or June 28th. This abrupt change correlates with the timing of Douglas-fir fertilization as observed from the histological developmental study we carried out in parallel.

Since the UI treatment is unlike the aborting treatment, both in genetic expression and physiology, it is assumed that *M. spermotrophus* is capable of manipulating overall plant gene expression within the seed. Currently we do not understand the process of manipulation of Douglas-fir seed by *M. spermotrophus* but studies on gall formation may

provide insights. Regulation of hormones such as cytokinins within plant tissue is suggested to be involved in the formation of insect-induced galls. However, increased cytokinin levels in galls may be restricted to the larvae itself, rather than the plant tissue (Elzen 1983). Gene expression analysis showed that the glands of adult willow sawflies had high concentrations of cytokinins (Yamaguchi et al., 2012). These results suggest that insects are capable of producing endogenous cytokinins. Cytokinins directly influence cell division (Werner et al., 2001). These are not the only genes involved in hormone expression that we detected. In my dataset, I identified transcripts annotated as proteins involved in the synthesis of brassinosteroids as well as cytokinins. These included a hydroxysteroid dehydrogenase (HSD) and a cytokinin dehydrogenase (comp 212371_c0 and comp206029_c0, respectively) (Figure S1). Recently an HSD was identified in plants and shown to behave like a brassinosteroid, a plant hormone involved in vascular differentiation and seed germination (Li et al., 2007). The HSD was up-regulated late in P, PI and UI treatments, with no expression in the U treatment. *Arabidopsis* plants overexpressing HSD showed increased growth. They were also more tolerant to saline and osmotic stress during seed dormancy. Manipulation of key hormone pathways involved in cell division and development such as cytokinins and HSDs may be essential for *M. spermotrophus'* success.

As expected, a large proportion (23%) of transcripts expressed similarly in the pollinated treatment and infested treatments were categorized to have a functional role related to seed storage and seed maturation. These transcripts include multiple LEA proteins, an oleosin protein (comp186025_c0), and a 2S seed storage protein (comp217681_c0)

among others. Accumulation of 2S albumin initially occurs in megagametophytes after fertilization, followed by the zygotic embryo (Chatthai and Misra, 1998). My histological sections of Douglas-fir megagametophytes showed that proteins did not visibly accumulate until early embryogenesis. However, storage transcripts appear to be upregulated a weak earlier, i.e. at the time of fertilization. Boyes et al. (in press) showed that transcription of storage proteins also increased in unpollinated megagametophytes of Douglas-fir even though proteins were not being produced in visible protein bodies. They hypothesized that storage proteins are post-transcriptionally regulated. This may explain the week-long gap between transcript expression and protein body presence in developing megagametophytes. In addition it may explain why there are few 2S seed storage proteins similarly expressed in the pollinated and both infested treatments of my study. Transcript expression of seed proteins may in fact be similar between all four treatments. Not only did storage transcripts spike in pollinated samples, their expression levels were similar in insect treatments, regardless of pollination. This result is consistent with the findings of von Aderkas et al. (2005a) who showed that megagametophytes infested with *M. spermotrophus* synthesized protein bodies to the same levels found in normal developing ovules. *Megastigmus spermotrophus* may be influencing post-transcriptional regulation of storage proteins so that protein bodies become actively synthesized.

Another storage product that increases during seed development is lipids. Accumulation of lipids is an essential energy source for embryos in germinating seeds (Huang, 1992). In Douglas-fir, lipid bodies only accumulate following fertilization (von Aderkas et al.,

2005ab, Owens et al., 1993). The synthesis of lipids is thought to be regulated by oleosin proteins (Siloto et al., 2006). Suppression of oleosin in *Arabidopsis* seeds caused embryos to develop abnormally large oil bodies, which were shown to disrupt storage organelles and postpone germination (Siloto et al., 2006). The timing of oleosin transcript expression shown in my study corresponds with increased transcription of storage production. This adds support to Siloto et al.'s (2006) hypothesis that oleosins function by preventing oil bodies from coalescing.

In addition to the accumulation of storage products within the megagametophyte, I observed endoreduplication in prothallial cells. Most cells were coenocytes, which, by definition have active DNA synthesis and karyokinesis in the absence of cytokinesis. Multinucleate cells were observed in fertilized ovules with early embryos and in ovules infested by *M. spermotrophus* (von Aderkas et al., 2005). Interestingly, a transcript encoding a DNA replication licencing factor, MCM3 (comp207431_c2) with transient expression in P, PI, and UI treatments was identified. Initially, this transcript was also expressed in unpollinated megagametophytes, however, it quickly declines by the point where fertilization would occur (Fig24A). DNA-licensing factors are necessary for regulating progression through the cell cycle (Sabelli and Larkins, 2007). To prevent multiple rounds of DNA replication per cycle, DNA-licensing factors are down-regulated (Castellano et al., 2004). The expression of this transcript in my dataset peaks around fertilization, when parenchyma storage cells become multinucleated.

Endoreduplication is prevalent in terminally differentiated plant tissue with high metabolic activity, such as megagametophytes and endosperms of gymnosperms and angiosperms, respectively (von Aderkas et al., 2005; Pichot and El Maataoui, 1997; D'Amato, 1984). There are a number of reasons as to the purpose of endoreduplication in seeds (Grime and Mowforth, 1982; Larkins et al., 2001; Bourdon et al., 2012; Sabelli et al., 2013). A commonly accepted theory is that it up-regulates gene expression by increasing the number of DNA templates, thereby increasing transcription and translation (Bourdon et al., 2012). It may also be a mechanism for storing nucleotides as a form of nitrogen cycling for developing embryos (Larkins et al., 2001). Both hypotheses would suggest a role in increased storage accumulation and protein synthesis in developing megagametophytes. Manipulation of the cell cycle by *M. spermotrophus* may be important for sequestering amino acids for protein synthesis.

Transcripts uniquely expressed in unpollinated megagametophytes -

The largest number of transcripts categorized as being manipulated are induced in unpollinated megagametophytes on July 3rd. This point in megagametophyte sampling corresponds to early embryogenesis in pollinated ovules. These transcripts likely represent genes involved in the abortion process of unfertilized ovules. However, the process of seed abortion is not well understood and may be impossible to tease apart from regular organized destruction of cells by programmed cell death (PCD). Not only is PCD involved in abortion, it is integral to normal megagametophyte and embryonic development. PCD is involved in corrosion cavity formation (Vuosku et al., 2009), elimination of multiple embryos in polyembryonic ovules (Filonova et al., 2002),

separation of embryos from suspensor cells (Suarez et al., 2004) and breakdown of cells to feed a developing embryo following germination (He and Kermonde, 2003a,b).

The developmental study revealed the progressive increase in the corrosion cavity prior to fertilization in both pollinated and unpollinated megagametophytes. This poses challenges for detecting transcripts specifically involved in the breakdown of cells during abortion. Despite the possibility of a reduced transcriptional signal between megagametophyte treatments, three transcripts involved in PCD were identified in unpollinated megagametophytes. A transcript encoding a protein described as a metacaspase was strongly induced following fertilization (comp201934_c1). In somatic embryos of *Picea* abies, metacaspases activated PCD of suspensor cells as embryos entered late embryogeny (Suarez et al., 2004). An additional transcript described as an autophagy related-protein (comp209706_c0), was also upregulated late in unpollinated megagametophytes. The lack of expression of these transcripts in the pollinated treatment and both infested treatments indicates a role in abortion.

PCD is essential for the breakdown of cells in the megagametophyte during seed abortion but the transport of stored products back into the mother tree is equally important. Of the transcripts unique to unpollinated ovules, the largest portion functioned in solute transport of sugars, lipids and amino acids. During ovule development, megagametophytes are active sink tissues before and after fertilization (von Aderkas et al., 2005b). Starch granules accumulate prior to fertilization in pollinated and unpollinated megagametophytes. However, in unfertilized ovules, storage reserves that

have accumulated in megagametophytes are recycled back into maternal tissue (von Aderkas et al., 2005b). Therefore, it is not surprising that post-fertilization solute transport increases in unpollinated ovules. Among these solute transporters, I identified a bidirectional sugar transporter (SWEET3b , comp205371_c0), amino acid transporters (comp202242_c0, comp192033_c0, comp211541_c0) and a lipid transporter (comp103395_c0). The presence of these specific transporters suggests bidirectional movement of solutes into and out of the megagametophyte.

Unlike diploid tissue, haploid megagametophytes do not have vascular tissue. In fact, with the exception of cycads and gingko, vascular tissue does not reach ovules of gymnosperms (Gifford and Foster, 1989). Studies of carbohydrate breakdown and transport back into the diploid sporophyte are lacking for seed plants. Unloading of solutes into seeds has however, received attention in systems of angiosperms. In angiosperms, vascular tissue ends at the seed coat, where there is symplastic unloading of solutes such as sucrose and amino acids from the phloem into the integument (Werner et al., 2011). A simple uptake experiment using pea seeds revealed that amino acid and sucrose release from the seed coat was a passive process requiring facilitated membrane transport (de Jong et al., 1997). Distribution of solutes within the integument occurs via plasmodesmata (Patrick and Offler, 2001; Werner et al., 2011). However, transport into endosperms and developing embyros requires solute specific transporters (Khan et al., 2014). High expression of transcripts encoding genes involved in sugar, amino acid and lipid transport within the chalazal region of angiosperm seed coats demonstrates a finer level of control over nutrient filling, which is greater than mere diffusion (Khan et al.,

2014). By analyzing microarrays of laser dissections, Chen et al. (2015) show that the sucrose efflux from the integument to the endosperm and subsequently into the embryo occurs through a series of SWEET transporters, specific to each tissue. They propose a model pathway of sucrose transport that enters the embryo via the micropylar end of the seed coat. In gymnosperms, solute transport is likely to occur at the nucellar end of the megagametophyte. This is the connection point between sporophyte and gametophyte, and coincidentally where the relatively impermeable megaspore wall is the thinnest. The use of laser dissection, alongside massive parallel sequencing would be useful techniques in understanding mechanisms of transport, and programmed cell death within specific target areas of the megagametophyte.

3.4.3 Limitations of the transcriptome and Differential expression analysis in detecting transcripts related to defense and transcripts manipulated by *M spermotrophus.*

Gene expression results provide clear evidence that *Megastigmus* alters Douglas-fir transcript expression in genes related to storage, and less so in genes related to defense but there are certain caveats to consider. There are no anatomical indications of megagametophytes detecting and mounting a response to *M. spermotrophus*. For example, structures typically found in galls, such as nutritive layers, and wound layers, are not seen. Upregulation of phenolics is not seen either. This does not mean that megagametophytes are incapable of a more concerted defense response. While my approach was useful in targeting highly differentially expressed transcripts between infested and uninfested megagametophytes, it is possible that the timing of sampling

itself contributed to the somewhat inconclusive result. Adult female chalcids were allowed two days to oviposit before megagametophyte collections began. In other host-parasite systems, transcriptional studies have shown that changes may occur within minutes of wounding (Leon et al., 2000). Studies observing changes in gene expression due to mechanical or herbivore wounding are often completed within hours of initiation (Koo et al., 2009). Nevertheless, in our system *M. spermotrophus* repeatedly oviposits, and I would expect initial changes in gene expression to have been detected. However, pooling itself may not solve this problem, as mixing multiple megagametophytes per sample may have made transcriptional signals weaker.

It may be necessary to identify elicitors unique to *M.* spermotrophus in order to understand Douglas-fir defense responses to this chalcid. This would require analyzing the suite of insect-specific genes. Furthermore, identification of specific elicitors may be challenging because all protein sequences were identified based on their resemblance to adult female *N. vitripennis* and insect protein sequences available in the Uniprot/Swissprot protein database. The lack of reference genomes of Douglas-fir and *M. spermotrophus* also limits the ability to detect and identify sequences that are unique to either species.

Another explanation for the low defense response rests on the possibility that plant defense was suppressed during initial egg deposition by *M. spermotrophus* or during insect feeding. The hypothesis that host defense can be suppressed by insect saliva, or microbial associates harboured by feeding insects has received much recent attention

(Thompson and Goggin, 2006; Consales et al., 2012; Chung et al., 2013). It is possible that changes in expression between infested and non-infested samples may reflect transcriptional reprogramming of Douglas-fir by *M. spermotrophus*. There may only have been a brief window where Douglas-fir transcripts reflected responses to HAMPS. However, the RNA-seq data does not indicate a strong presence of microbial products, and HAMPS may not have even been detected.

3.4.3 Bias generated from the transcriptome assembly and analysis

The following sections provide a brief description of bias and noise generated in the transcriptome assembly process.

De novo assembly – The *de novo* Trinity assembly was comprised of 382,101 contigs with 351,150 unique transcripts. Compared to other conifer transcriptomes, this Douglas-fir/*Megastigmus* assembly is significantly larger. Illumina sequencing of *P. monticola* needles generated an assembly of 39,439 unique transcripts (Lui, et al., 2013). The transcriptome of loblolly pine, which was compiled from multiple tissue types, including both haploid and diploid material, produced an assembly of 149,992 transcripts (Lorenz et al., 2012), approximately two thirds smaller than mine. Differences in assembly sizes can arise due to the combination of tissue types sequenced, quantity of sequenced tissue, sequence depth, sequencing technology, and assembly methods. The transcriptome generated in this study was compiled from 33 individual libraries, which included tissue from *M. spermotrophus* larvae, Douglas-fir haploid micro and megagametophytes, diploid Douglas-fir embryos, and diploid Dougals-fir maternal tissue, including

integument and scales. Using multiple tissue types from two generations, two diverse organisms, and a large number of libraries has added complexity and noise, and undoubtedly influenced assembly size.

Annotations – As expected, the majority of post-assembly filtered transcripts were of plant origin. Of 11,730 plant transcripts, 10,826 had sequence similarity to the *P. taeda* genome, which provided some validation of the plant annotations. The presence of two distinct species within the assembly did, however, make annotations difficult. Multiple transcripts had sequence similarity to both plant and insect. When deciding species origin of these transcripts, I compared E-values and bit scores. Using E-values is potentially problematic since they are assigned based on quality of the alignment, length of the sequence queried, and the size of the database. Comparing E-values between databases of different sizes, i.e. the *Nasonia* genome compared to Swissprot/Uniprot protein database could result in missannotations, especially when E-values between the databases are similar.

While most highly expressed transcripts were of plant origin, there were a considerable number of transcripts that did not have sequence similarity to any organism (35%). It is unusual to see an abundance of highly differentially expressed transcripts that cannot be identified to a species. Some of these transcripts may represent unique sequences from Douglas-fir or *M. spermotrophus,* while others may represent non-coding RNA that are not included in the TAIR and Swissprot/Uniprot protein databases. Considering that

unannotated transcripts, by definition, lacked sequence similarity to *P. taeda*, they may have been artefacts of assembly compilation.

Transcript expression– Sequencing depth was not uniform for all 33 libraries. Five samples of megagametophytes infested with *M. spermotrophus* were pooled into one lane. The remaining 7 lanes consisted of only 4 libraries. Thus, the total number of contigs of the 5 infested samples would have been reduced and underrepresented in the assembly.

An additional problem arises when there are differences in the total number of reads for each organism. In this study the number of reads for plant tissue was significantly greater than the number of reads for insect tissue. Since expression is based on the number of reads present, FPKM expression values of insect transcripts would be low. To obtain accurate expression values, it is necessary to realign transcripts of insect and plant origin back to the assembly, separately. Failure in realigning the plant transcripts has likely increased the chance of false positives in my data.

DE analysis - Although quadratic regression was a useful technique for targeting highly differentially expressed transcripts across a time-series of megagametophyte treatments, it may have been too stringent to identify transcripts involved in plant defense and/or manipulation. Transcripts with expression deviating from a regression pattern were categorized as having no distinct pattern and were not used for further analysis. In some cases, transcripts deviated from the mean at only one time point, yet these were still

excluded. Linking GO annotations would have been useful to additionally search for transcripts that did not significantly fit a regression pattern. The use of GO terms such as "seed storage", "solute transport", "stress", etc to search for candidate genes may have identified transcripts involved in plant defense against herbivory or other transcripts that were manipulated by *M. spermotrophus* that did not meet the requirements of quadratic regression.

3.4.6 Prospects and conclusions.

In this study, a *de novo* transcriptome assembly of Douglas-fir megagametophytes infested by *M. spermotrophus* larvae was generated and expression data obtained to identify i) putative genes related to plant defense, and ii) putative genes manipulated by the insect. The majority of transcripts were annotated as plant origin and many had sequence similarities to transcripts of the *P. taeda* genome. I identified 56 transcripts highly differentially expressed between infested and uninfested treatments. These represented putative transcripts related to defense. Insect infested megagametophytes were found to have transcript expression profiles more similar to pollinated than unpollinated samples. Further, three hundred of these transcripts were hypothesized to be manipulated as a result of insect parasitism.

Of the transcripts differentially expressed between infested and uninfested megagmetophytes, there were a few candidates suggesting a response of Douglas-fir megagametophytes to wounding. These include SNAKIN2, LOB, and the enzymes, XERICO, AIP2, and EI3, which are involved in ABA and ethylene regulation. The

identification of SNAKIN2, which is a protein implied to be involved in wounding and pathogen infection is an intriguing result and is a putative gene worth pursuing in order to understand plant defense. SNAKIN2 is an excellent candidate to test whether Douglas-fir megagametophytes are responding to mechanical wounding or if in fact megagametophytes are responding to herbivory or possibly pathogens that could be introduced by *M. spermotrophus*.

There were no transcripts that indicated recognition of *M. spermotrophus* presence by Douglas-fir. Therefore, it is difficult to know whether wound response was due to oviposition by *M. spermotrophus* eggs or if it was due to continued feeding by the larvae. The idea that plants can separate mechanical wound response from herbivory has been previously suggested (Howe and Jander, 2008). Larval secretions from herbivores are well known to elicit responses by host plants and may be involved in suppression of host plant defense (Consales et al., 2012; Chung et al., 2013). I would be interested in using proteomic approaches to identify proteins present in larval secretions. This would be useful for identifying insect-specific elicitors, or enzymes that may be involved in insect feeding.

These data provide insight into megagametophyte development of Douglas-fir, particularly highlighting processes that may be important for abortion of unpollinated ovules. The act of fertilization and presence of genes specific to embryos is thought to be compulsory for the progression of megagametophyte development (Pagnussat et al., 2005). This plant-parasite interaction demonstrates that neither fertilization nor an

embryo are necessary. Unpollinated megagametophytes infested with *M. spermotrophus* continue development. Multiple studies suggest that this insect manipulates gene expression of a Douglas-fir megagametophyte so that unpollinated ovules behave as though they are fertilized, thus enabling the continuation of storage accumulation.

At its simplest, manipulation may involve little more than this insects presence. The insect substitutes for an embryo. For example, in developing ovules that have been fertilized, the embryo becomes an additional sink. Similarly, insect larvae that consume nutrients stored in the megagametophyte, maintain the gradient that defines this sink. Sucrose transport is highly dependent on osmolality (Patrick and Offler, 2001). Since most nutrients are transported towards sink tissue through phloem, it is suggested that solute transport occurs by bulk flow (Zhang et al., 2007). Concentration differences of solutes in sink tissue compared to source tissue may be the most important factor controlling the rate of transport from source to sink (Patrick and Offler, 2001). Thus, the ability of a seed to sequester solutes based on its demand will influence post-phloem transport.

The presence of an embryo within the megagametophyte is drives the passive transport of solutes within the phloem tissue towards the seed. A hypothesis that may explain why unpollinated, infested megagametophytes do not abort is that the consumption activity of the larvae replaces the role of an embryo by becoming a new sink. This consequently increases the demand of solute assimilation in the megagametophyte.

There are tools available that could be used to get a better understanding of solute transport within megagametophytes of conifers. The use of laser dissections and microarrays may be useful in targeting genes directly involved transport mechanisms of solutes such as sucrose, lipids and amino acids. It would be particularly interesting to compare expression of genes related to transport in infested and uninfested megagametophytes of Douglas-fir. This may provide insight into the abortion process of seed plants.

The application of RNA-seq has been useful in identifying differentially expressed transcripts between pollinated and unpollinated megagametophytes, both infested and uninfested by *M. spermotrophus*. Although this study did not reveal any transcripts involved in the mechanism of manipulation, there were a number of transcripts that were manipulated in the presence of *M. spermotrophus*. The idea that *M. spermotrophus* manipulates its host has previously been shown through histochemical and histological studies (von Aderkas et al., 2005a,b; Chiwocha and von Aderkas, 2002; Chiwocha et al., 2006). The finding that infested megagametophytes had more transcripts similarly expressed to pollinated than unpollinated treatments provides further evidence of host manipulation at a genetic level. Identification of solute transporters along with genes involved in autophagy in unpollinated samples may be useful in understanding the abortion process and also development of megagametophytes. The defense investment of Douglas-fir megagametophytes remains inconclusive. Although there are transcripts differentially expressed between infested and uninfested megagametophytes, very few have roles that may be related to wounding; most of these transcripts cannot be attributed

to a specific biological activity. There is a possibility that Douglas-fir defense responses are suppressed. Analysing the insect and microbial transcripts from this transcriptome may provide further insight into this complex plant-parasite interaction.

Bibliography

Alborn, H.T., Hansen, T.V., Jones, T.H., Bennet, D.C., Tumlinson, J.H., Schmelz, E.A., and Teal, P.E.A. 2007. Disulfooxy fatty acids from the American bird grasshopper *Schistocerca americana*, elicitors of plant volatiles. Proceedings of the National Academy of Science of the United States of America 104(32): 12976-12981.

Allen, G.S. 1943. The embryogeny of *Psuedotsuga taxifolia* (Lamb) Britt. American Journal of Botany 30(8): 655-661.

Allen, G.S. 1963. Origin and development of the ovule in Douglas-fir. Forest Science 9(4): 386-393.

Allen, G.S., and Owens, J.N. 1972. The life history of Douglas-fir. Environmental Canada Forest Services.

Andrews, S. 2012. Version 0.10.1 2013-05-03. The Babraham Institue, Cambridge UK. Website http://http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ [accessed 2013]

Arimura, G-I., Ozawa, R., and Maffei, M.E. 2011. Recent advances in plant early signaling in response to herbivory. International Journal of Molecular Science 12(6): 3723-3739.

Auger-Rozenberg, M.A., Kerdelhué, C., Magnoux, E., Turgeon, J., Rasplus, J-Y., and Roques, A. 2005. Molecular phylogeny and evolution of host-plant use in conifer seed chalcids in the genus *Megastigmus* (Hymenoptera: Torymidae). Systematic Entomology 31(1): 47–64.

Barr, K.L., Hearne, L.B., Briesacher, S., Clark, T.L., and Davis, G.E. 2010. Microbial symbionts in insects influence down-regulation of defense genes in maize. PLoS One 5(6): 11339.

Bates, D.M., M. Maechler, B. Bolker, and S. Walker. 2014. Lme4: Liner mixed-effects models using Eigen and S4. R package version 1.0-6.

Bates, S.L., Borden, J.H., Savoie, A., Blatt, S.E., Lait, C.G., Kermode, A.R., and Bennett, R.G. 2000. Impact of feeding by *Leptoglossus occidentalis* (Hemiptera: Coreidae) on the major storage reserves of mature Douglas-fir (Pinaceae) seeds. Canadian Entomologist 132(1): 91-102.

Bencivenga, S., Colombo, L., and Masiero, S. 2011. Cross talk between the sporophyte and the megagametophyte during ovule development. Sexual Plant Reproduction 24(2): 113–121.

Bethke, P.C., Lonsdale, J.E., Fath, A., and Jones, J.L. 1999. Hormonally regulated programmed cell death in barley aleurone cells. The Plant Cell 11(6): 1033-1045.

Berrocal-Lobo, M., Seguara, A., Moreno, M., López, G., García-Olmedo, F., and Molina, A Snakin-2, an antimicrobial peptide from potato whose gene is locally induced by wounding and responds to pathogen infection. Plant Physiology 128: 951-961.

Bianco, J., Garello, G., and Le Page-Degivry, M.T. 1997. *De novo* ABA synthesis and expression of seed dormancy in gymnosperm: *Pseudotsuga menziesii*. Plant Growth Regulation 21: 115-119.

Birch, L.C. 1957. The meanings of competition. The American Naturalist 91: 5-18.

Björkman, C., Larsson, S., Bommarco, R. 1997. Oviposition preferences in pine sawflies: A trade-off between larval growth and defense against natural enemies. Oikos 79: 45-52.

Blatt, S.E., and Borden, J.H. 1998. Interactions between the Douglas-fir seed chalcid, *Megastigmus spermotrophus* (Hymenoptera: Torymidae), and the western conifer seed bug, *Leptoglossus occidentalis* (Hemiptera: Coreidae). The Canadian Entomologist 130(6): 775-782.

Boivin, T., Rouault, G., Chalon, A., and Candau, J. N. 2007. Differences in life history strategies between an invasive and a competing resident seed predator. Biological Invasions 10(7): 1013–1025.

Bonaventure, G. 2012. Perception of insect feeding by plants. Plant Biology 14: 872-880.

Bormann, B. T. 1984. Douglas-fir: An American wood. USDA Forest Service Report FS-235, Washington, D.C.

Bourdon, M., Pirrello, J., Cheniclet, C., Coriton, O., Bourge, M., Brown, S., Moïse, A., Peypelut, M., Rouyère, V., Renaudin, J-P, Chevalier, C., and Frangne, N. 2012. Evidence for karyoplasmic homeostasis during endoreduplication and a ploidy-dependent increase in gene transcription during tomato fruit growth. Development 139: 3817–3826.

Boyes, I.G., Little, S.A., von Aderkas, P., and Ehlting, J. A transcriptomic resource for Douglas-fir seed development and analysis of transcription during late megagametophyte development. [In Press]

Broekaert, W.F., Delaure, S.L., De Bolle, M.F.C., and Cammuel, B.P.A. 2006. The role of ethylene in host-pathogen interactions. Annual Reviews of Phytopathology 44: 393-416.

Bruessow, F., and Reymond, P. 2007. Oviposition-induced changes in *Arabidopsis* genome expression: anticipating your enemy? Plant Signal Behaviour 2: 165-167.

Byun-McKay, A., Godard, K-A., Toudefallah, M., Martin, D.M., Alfaro, R., King, J., Bohlmann, J., and Plant, A.L. 2006. Wound-induced terpene synthase gene expression in sitka spruce that exhibit resistance or susceptibility to attack by the white pine weevil. Plant Physiology 140: 1009-1021.

Carman, J. G., Reese, G., Fuller, R. J., Ghermay, T. and Timmis, R. 2005. Nutrient and hormone levels in Douglas-fir corrosion cavities, megagametophytes, and embryos during embryony. Canadian Journal of Forest Resources 35: 2447-2456.

Caron, G.E, and G.R. Powell. 1989. Cone size and seed yield in young *Picea mariana* trees. Canadian Journal of Forest Research 19:351-358.

Castellano, del Mar., Boniotti, M.B., Caro, E., Schnittger, A., and Gutierrez, C. DNA replication licensing affects cell proliferation or endoreplication in a cell type-specific manner. The Plant Cell 16(9): 2380-2393.

Chatthai, M., and Misra, S. 1998. Sequence and expression of embryogenesis-specific cDNAs encoding 2S seed storage proteins in *Pseudotsuga menziesii* [Mirb.] Franco. Planta 206(1): 138-145.

Chao, Q., Rothenberg, M., Solano, R., Roman, G., Terzaghi, W., and Ecker, J.R. 1997. Activation of the ethylene gas response pathway in Arabidopsis by the nuclear protein ETHYLENE-INSENSITIVE3 and related proteins. Cell 89: 1133–1144.

Chen, L.Q., Lin, I.W., Qu, X-Q., Sosso, D., McFarlane, H.E., Londoño, A., Samuels, A.L., and Frommer, W.B. 2015. A cascade of sequentially expressed sucrose transporters in the seed coat and endosperm provides nutrition for the *Arabidopsis* embryo. The Plant Cell 27: 607-619.

Cheong, Y.H., Chang, H-S., Gupta, R., Wang, X., Zhu, T., and Luan, S. 2002. Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in *Arabidopsis*. Plant Physiology 129(2): 661-677.

Chiwocha, S., Rouault, G., Abrams, S., and von Aderkas, P. 2006. 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(1): 19–25.

Chiwocha, S., and von Aderkas, P. 2002. Endogenous levels of free and conjugate forms of auxin, cytokinins and abscisic acid during seed development in Douglas fir. Plant Growth Regulation 36: 191-200.

Chung, S.H., Rosa, C., Scully, E.D., Peiffer, M., Tooker, J.F., Hoover, K., Luthe, D.S., and Felton, G.W. 2013. Herbivore exploits orally secreted bacteria to suppress plant defenses. Proceedings of the National Academy of Science of the United States of America 110(39): 15728-15733.

Conrath, U. 2006. Systemic acquired resistance. Plant Signaling and Behavior 1(4): 179-184.

Consales, F., Schweizer, F., Erb, M. Gouhier-Darimont, C., Bodenhausen, N., Bruessow, F., Sobhy, I., Reymond, P. 2012. Insect oral secretions suppress wound-induced responses in *Arabidopsis*. Journal of Experimental Botany 63(2): 727-737.

Courtney, S.P., and Kibota, T.T. 1990. Mother doesn't know best: selection of hosts by ovipositing insects. In Bernays, E.A. [ed.], Insect-Plant Interactions, vol. II. CRC, Boca Raton. FL, 161-187.

Craig, T.P., Itami, J.K., and Price, P.W. 1989. A strong relationship between oviposition preference and larval performance in a shoot-galling sawfly. Ecology 70(6): 1691-1699.

Craig, T. P., Itami, J.K., and Price, P.W. 1990. The window of vulnerability of a shoot-galling sawfly to attack by a paraistoid. Ecology 71(4): 1471-1482.

Creelman, R.A., and Mullet, J.E. 1995. Jasmonic acid distribution and action in plants: regulation during development and response to biotic and abiotic stress. Proceedings of the National Academy of Sciences of the United States of America 92(10): 4114-4119.

D'Amato, F. 1984. Role of polyploidy in reproductive organs and tissues. In: Johri B.M., ed. Embryology of Angiosperms. New York: Springer Verlag, 523–566.

Denlinger, D. L. 2002. Regulation of diapause. Annual Review of Entomology 47: 93-122.

De Diego, N., Rodríguez, J.L., Dodd, I.C., Pérez-Alfocea, F., Moncaleán, P., and Lacuesta, M. 2013. Immunolocalization of IAA and ABA in roots and needles of radiata pine (*Pinus radiata*) during drought and rewatering. Tree Physiology 33: 537-549.

De Jong, A., Koerselman-Kooij, J.W., Schuurmans, J.A.M.J., and Borstlap, A.C. 1997. The mechanisms of amino acid efflux from seed coats of developing pea seeds as revealed by uptake experiments. Plant Physiology 114: 731–736.

De-silva, D.L., Vásquez, S.A., and Mallet, J. 2011. Selection for enemy-free space: eggs placed away from host plant increase survival of a neotropical ithomiine butterfly. Ecological Entomology 36: 667-672.

Dolan, L. 1997. The role of ethylene in the development of plant form. Journal of Experimental Botany 48(307): 201-210.

Dormont, L., and Roques, A. 1999. A survey of insects attacking seed cones of *Pinus cembra* in the Alps, the Pyrenees and Massif Central. Journal of Applied Entomology 123:65-72

Durrant, W.E., Rowland, O., Piedras, P., Hammond-Kosack, K.E., and Jones, J.D.G. 2000. cDNA-AFLP reveals a striking overlap in race-specific resistance and wound response gene expression profiles. Plant Cell 12(6): 963-977.

Eckenwalder, J. E. 2009. Conifers of the world – the complete reference. Timber Press Inc., USA.

Elzen, G.W. 1983. Cytokinins and insect galls. Comparative Biochemistry and Physiology 76(1): 17-19.

Fan, M., Xu, C., Xu, K., Hu, Y. 2012. LATERAL ORGAN BOUNDARIES DOMAIN transcription factors direct callus formation in *Arabidopsis* regeneration. Cell Research 22(7): 1169-1180.

Fidgen, L.L., Quiring, D.T., and Sweeney, J.D. 1998. Effect of cones size on adult and larval foraging behavior of *strobilomyia neanthracina* and *Strobilomyia appalachensis* (Diptera: Anthomyiidae). Enviornmental Entomology 27(4): 877-884.

Filonova, L.H., von Arnold, S., Daniel, G., and Bozhkov, P.V. 2002. Programmed cell death eliminates all but one embryo in a polyembryonic plant seed. Cell Death and Differentiation 9: 1057-1062.

Finkelstein, R.R., Tenbarge, K.M., Shumway, J.E., and Crouch, M.L. 1985. Role of ABA in the maturation of rapeseed embryos. Plant Physiology 78: 630-636.

Freeman, B.C., and Beattie, G.A. 2008. An overview of plant defenses against pathogens and herbivores. The Plant Health Instructor. DOI: 10.1094/PHI-I-2008-0226-01

Gouhier-Darimont, C., Schmiesing, A., Bonnet, C., Lassueur, S., and Reymond, P. 2013. Signalling of *Arabidopsis thaliana* response to *Pieris brassicae* eggs shares similarities with PAMP-triggered immunity. Journal of Experimental Botany 64(2): 665-674.

Ghazoul, J., and Satake, A. 2009. Nonviable seed set enhances plant fitness: the sacrificial sibling hypothesis. Ecology 90(2): 369-377.

Grabherr, M.G., Haas, B.J. Yassour, M., Levin, J.Z. Thompson, D.A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q., Chen Z., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B.W., Nusbaum, C., Lindblad-Toh, K., Friedman, N., and Regev, A. 2011. Full length transcriptome assembly form RNA-Seq data without a reference genome. Nature Biotechnology 29: 644-652.

Grime, J.P., and Mowforth, M.A. 1982. Variation in genome size—an ecological interpretation. Nature 299: 151–153.

Gripenberg, S., Mayhew, P.J., Parnell, M., and Roslin, T. 2010. A meta-analysis of preference-performance relationships in phytophagous insects. Ecology Letters 13: 383-393.

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

Gutmann, M. 1995. Improved staining procedures for photographic documentation of phenolic deposits in semi-thin sections of plant tissue. Journal of Microscopy. 179: 277-281.

Hahn, D. A., and Denlinger, D. L. 2011. Energetics of insect diapause. Annual Review of Entomology 56: 103–121.

Hanley, M.E., Byron, B.L., Fairbanks, M.M., Rafferty, and C.M. 2007. Plant structural traits and their role in anti-herbivore defence. Perspectives in Plant Ecology, Evolution and Systematics 8: 157-178.

Harper, L., Schönrogge, K., Lim, K., Francis, P., and Lichtenstein, C. 2004. Cynipid galls: insect-induced modifications of plant development create novel plant organs. Plant. Cell and Environment 27(3): 327–335.

Hartley, S. 1998. The chemical composition of plant galls: are levels of nutrients and secondary compounds controlled by the gall-former? Oecologia 113(4): 492–501.

He, X., Kermode, A.R. 2003a. Nuclease activities and DNA fragmentation during programmed cell death of megagametophyte cells of white spruce (*Picea glauca*) seeds. Plant Molecular Biology 51: 509–521.

He, X., Kermode, A.R. 2003b. Proteases associated with programmed cell death of megagametophyte cells after germination of white spruce (*Picea glauca*) seeds. Plant Molecular Biology 51: 729-744.

Hermann, R. K. and Lavender, D. P. 1999. Douglas-fir planted forests. New Forests 17: 53-70.

Hilker, M., Stein, C., Schröder, R., Varama, M., and Mumm, R. 2005. Insect egg deposition induces defense responses in *Pinus sylvestris*: characterization of the elicitor. Journal of Experimental Biology 208: 1849-1854.

Howe, G.A., and Jander, G. 2008. Plant immunity to insect herbivores. Annual Review of Plant Biology 59: 41-66.

Huang, A.H.C. 1992. Oilbodies and oleosins in seeds. Annual Review of Plant Physiology in Plant Molecular Biology 43: 177–200.

Hudgins, J.W., Christiansen, E., and Franceschi, V.R. 2004. Induction of anatomically based defense responses in stems of diverse conifers by methyl jasmonate: a phylogenetic perspective. Tree Physiology 24: 251-264.

Hudgins, J.W., and Franceschi, V.R. 2004. Methyl jasmonate-induced ethylene production is responsible for phloem defense responses and reprogramming of conifer stem cambial zone for traumatic resin duct formation. Plant Physiology 135: 2134-2149.

Hudgins, J.W., Krekling, T., and Francheschi, V.R. 2003. Distribution of calcium oxaloate crystals in the secondary phloem of conifers: a constitutive defence mechanism? New Phytologist 159: 677-690.

Hudgins, J.W., Ralph, S.G., Franceschi, V.R., and Bohlmann, J. 2006. Ethylene in induced conifer defense: cDNA cloning, protein expression, and cellular and subcellular localization of 1-aminocyclopropane-1-carboxylate oxidase in resin duct and phenolic parenchyma cells. Planta 224(4): 865-877.

Hussey N.W. 1953. The life histories of *Megastigmus spermotrophus* Wachtl (Hymenoptera: Chalcidoidea) and its principle parasite with description of the developmental stages. Transactions of the Royal Entomological Society of London 106: 133-151.

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

Ikeuchi, M., Sugimoto, K., and Iwase, A. 2013. Plant Callus: Mechanisms of induction and repression. The Plant Cell 25(9): 3159-3173.

Jaenike, J. 1978. An optimal oviposition behaviour in phytophagous insects. Theoretical Population Biology 14: 350-356.

Jaenike, J. 1990. Host specialization in phytophagous insects. Annual Review of Ecologoy Systems 21: 243-273.

Kaiser, W., Huguet, E., Casas, J. Commin, C., and Giron, D. 2010. Plant green-island phenotype induced by leaf-miners is mediated by bacterial symbionts. Proceedings of the Royal Society of Biological Sciences 277(1692): 2311-2319.

Kermode, A.R. 2005. Role of abscisic acid in seed dormancy. Journal of Plant Growth and Regulation 24: 319-344.

Khan, M.A., Castro-Guerrero, N., and Mendoza-Cozatl, D.G. 2014. Moving toward a precise nutrition: preferential loading of seeds with essential nutrients over non-essential toxic elements. Frontiers in Plant Science 5: 51.

Ko, J-H., Yang, S.H., and Han, K-H. 2006. Upregulation of an Arabidopsis RING-H2 gene, XERICO, confers drought tolerance through increased abscisic acid biosynthesis. The Plant Journal 47: 343-355.

Kolosova, N., Miller, B., Ralph, S., Ellis, B.E., Douglas, C., Ritland, K., and Bohlmann, J. 2004. Isolation of high quality RNA from gymnosperm and angiosperm trees. BioTechniques 36: 821-824.

Koo, A., Gao, X., Jones, A.D., How, G.A. 2009. A rapid wound signal activates the systemic synthesis of bioactive jasmonates in Arabidopsis. The Plant Journal 59: 974-986.

Koornneef, A., and Pieterse, C.M.J. 2008. Cross talk in defense signaling. Plant Physiology 146(3): 839-844.

Krueger, F. 2012. Version 0.2.5 2012-10-19. The Babraham Institute, Cambridge UK. Website http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/ [Accessed 2014].

Krugman, S.L. 1986. The ethical question. Journal of Forestry 84: 40-41.

Lamesch, P., Berardini, T., Swarbreck, D.L.D., Wilks, C., Sasidharan, R., Muller R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., Karthikeyan, A.S., Lee, C.H., Nelson, W.D., Ploetz, L., Singh, S., Wensel, A., and Huala, E. 2011. The Arabidopsis information resource (TAIR): improved gene annotation and new tools. Nucleic Acids Research 2011: 1-9.

Larkins, B.A., Dilkes, B.P., Dante, R.A., Coelho, C.M., Woo, Y-M., and Liu, Y. 2001. Investigating the hows and whys of DNA endoreduplication. Journal of Experimental Botany 52(355): 183-192.

Lawson, A.A. 1909. The gametophytes and embryo of *Pseudotsuga Douglasii*. Annals of Botany 23: 163-178.

Larsson, S., C. Glynn, and S. Hoglund. 1995. High oviposition rate of *Dasineura marginemtorquens* on *Salix viminalis* genotypes unsuitable for offspring survival. Entomologia Experimentalis et Applicata 77: 263.270.

Leal, I., Misra, S., Attree, S.M., and Fowke, L.C. 1995. Effect of abscisic acid. osmoticum, and desiccation on 11S storage protein gene expression in somatic embryos of white spruce. Plant Science 106: 121-128.

León, J., Rojo, E., and Sánchez-Serrano, J.J. 2000. Wound signalling in plants. Journal of Experimental Botany 52(354): 1-9.

Lessman, D. 1974. Ein beitrag zur verbreitung and lebenenweise von *Megastigmus spermotrophus* Wachtl. (Hym. Torymidae). Acta Univ. Agric., Rocz. XXXVI, Sper c. 272: 275-286.

Li, F., Asami, T., Wu, X., Tsang, E.W., and Culter, A.J. 2007. A putative hydroxysteroid dehydrogenase involved in regulating plant growth and development. Plant Physiology 145: 87-97.

Li, B., and Dewey, C.N. 2011. RSEM: Accurate transcript quantification from RNA-Seq data with or without a reference genome. BioMed Central Bioinformatics 12: 323.

Li H., Handsaker B., Wysoker A., Fennell T., Ruan J., Homer N., Marth G., Abecasis G., Durbin R., and 1000 Genome Project Data Processing Subgroup. 2009. The sequence alignment/map (SAM) format and SAMtools. Bionformatics 25: 2078-9.

Libby, W. J. 1986. Clonal propagation. Journal of Forestry 84: 37-42.

Liebhold, A.M., MacDonald, W.L., Bergdahl, D., and Mastro, V.C. 1995. Invasion of exotic forest pests: A threat to forest ecosystems. Forest Science 41(2): 1-50.

Liu, J-J. Sturrock, R.N., and Benton, R. 2013. Transcriptome analysis of *Pinus monticola* primary needles by RNA-seq provides novel insight into host resistance to *Cronartium ribicola*. BMC Genomics 14: 884-899.

Liu, X., Williams, C.E. Nemacheck, J.A., Wang, H., Subramanyam, S., Zheng, C., and Chen, M-S. 2010. Reactive oxygen species are involved in plant defense against a gall midge. Plant Physiology 152(2): 985-999.

Lorenz, W.W., Ayyampalayam, S., Bordeaux, J.M., Howe, G.T., Jermstad, K.D., Neale, D.B., Rogers, D.L., and Dean, J.F.D. 2012. Conifer DBMagic: A database housing multiple de novo transcriptome assemblies for twelve diverse conifer species. Tree Genetics and Genomes 8: 1477-1485.

Mailleux, A.C., Roques, A., Molenberg, J-M., and Gregorie, J-C. (2008). A North American invasive seed pest, *Megastigmus spermotrophus* (Wachtl) (Hymenoptera: Torymidae): Its populations and parasitoids in a European introduction zone. Biological Control 44: 137-141.

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

Marion-Poll, F. and Thiery, D. 1992. Olfactory sensitivity of Dougas-fir seed chalcids to plant odors. IIUFRO Cone and Seed Insects Conference, Beijing (China): 1-8.

Mattiacci, L., Dicke, M., and Posthumus, M.A. 1995. β-Glucosidase: An elicitor of herbivore-induced plant odor that attracts host-searching parasitic wasps. Proceedings of the National Academy of Science of the United States of America 92(6): 2036-2040.

McClure, M., Quiring, D. T., and Turgeon, J. J. 1998. Proximate and ultimate factors influencing oviposition site selection by endoparasites on conifer seed cones: two sympatric dipteran species on larch. Entomolgia Experimentalis et Applicata 87: 1-13.

McNeil, J.N., and Quiring, D.T. 1983. Evidence of an oviposition-deterring pheromone in the alfalfa blotch leaf-miner, *Agromyza frontella* (Rond.) (Diptera: Agromyzidae). Environmental Entomolgy 12: 990-992.

Miller, J.M. 1916. Oviposition of *Megastigmus spermotrophus* in the seed of Douglas-fir. Journal of Agricultural Research 6(3): 65-73.

Miller, G.E., and Borden, J.H. 1984. Reproductive behaviour of the Dougals-fir cone gall midge, *Contarinia oregonensis* (Dipetera: Cecidomyiidae). The Canadian Entomologist 116: 607-618.

Miller, B., Madilao, L.L., Ralph, S.R., and Bohlman, J. 2005. Insect-induced conifer defense. White pine weevil and methyl jasmonate induce traumatic resin, de novo formed volatile emissions, and accumulation of terpenoid sythnase and putative octadecanoid pathway transcripts in Sitka spruce. American Society of Plant Biologists 137(1): 369-382.

Miller, G.E. 1982. Biology, sampling and control of the Douglas-fir cone gall midge, *Contarinia oregonensis* Foote (Diptera: Cecidomyiidae), in Douglas-fir seed orchards in British Columbia. Dissertation, Simon Fraser University.

Milliron, H. E. 1949. Taxonomic and biological investigations in the genus *Megastigmus* with particular reference to the taxonomy of the Nearctic species (hymenopter: chalcidoidea; Callimomidae). American Midland Naturalist 41(2): 257-420.

Misra, S. and Green, M.J. 1990. Developmental gene expression in conifer embryogenesis and germination. I seed proteins and protein body composition of mature embryo and the megagametophyte of White Spruce (*Picea glacua* [Moench] Voss). Plant Science 68: 163-173.

Misra, S. and Green, M.J. 1991. Developmental gene expression in conifer embryogenesis and germination. II Crystaloid protein synthesis in the developing embryo and megagametophyte of White Spruce (*Picea glacua* [Moench] Voss). Plant Science 78: 61-71.

Misra, S., Attree, S.M., Leal, I., and Fowke, L.C. 1993. Effect of abscisic acid, osmoticum, and desiccation on the synthesis of storage proteins during the development of white spruce. Annals of Botany 71: 11-22.

Mithöfer, A., and Boland, W. 2008. Recognition of herbivory-associated molecular patterns. Plant Physiology 146(3): 825-831.

Mohan, S., Meiyalaghan, S., Latimer, J.M., Gatehouse, M.L., Monaghan, K.S. Vanga, B.R., Pitman, A.R. Jones, E.E., Conner, A.J., and Jacobs, J.M.E. 2014. GSL2 over-expression confers resistance to *Pectobacterium atrosepticum* in potato. Theoretical Applied Genetics 127: 667-689.

Moribe, T., Yamashita, O., and Yaginuma, T. 2001. A novel cold inducible gene, encoding a protein with a BAG domain similar to silencer of death domains (SODD/BAG-4), isolated from Bombyx diapause eggs. European Journal of Biochemistry 268(12): 3432-42.

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

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

Nabity, P.D., Haus, M.J., Berenbaum, M.R., and DeLucia, E.H. 2013. Leaf-galling phylloxera on grapes reprograms host metabolism and morphology. Proceedings of the National Academy of Science of the United States of America 110(41): 16663-16668.

Niwa, C. and Overhulser, D.L. 1992. Oviposition and development of *Megastigmus spermotrophus* (Hymenoptera: Torymidae) in unfertilized Douglas-fir seed. Jounral of Economic Entomology 85(6): 2323 – 2328.

Nufio, C.R. and Papaj, D.R. 2001. Host marking behaviour in phytophagous insects and parasitoids. Entomologia Experimentalis et Applicata 99: 273-293.

Oliver, J.P., Castro, A., Gaggero, C., Cascón, T., Schmelz, E.A., Castresana, C., Ponce de León, I. 2009. *Pythium* infection activates conserved plant defense responses in mosses. Planta 230: 569-579.

Orr-Ewing, A.L. 1956. An investigation into the effects of self-pollination of *Pseudotsuga menziesii* (Mirb.) Franco. Ph.D. dissertation. Vancouver, B.C. University of British Columbia.

Owens, J.N., and Blake, M.D. 1986. Production de semences forestiéres: revue bibliographique et suggestions de recherché. Rapport d'information PI-X-53. Institut Forestier National de Petawawa. Service Canadien des Forets, Chalk River, Ont.

Owens, J.N., Colangeli, A.M., and Morris, S.J. 1991. Factors affecting seed set in Douglas-fir (*Pseudotsuga menziesii*). Canadian Journal of Botany 69: 229-238.

Owens, J.N., and Morris, S.J. 1990. Cytological basis for cytoplasmic inheritance in *Pseudotsuga menziesii*. 1. Pollen tube and archegonial development. American Journal of Botany 77(4): 433-445.

Owens, J.N., and Morris, S.J. 1991. Cytological basis for cytoplasmic inheritance in *Pseudotsuga menziesii*. 2. Fertilization and proembryos development. American Journal of Botany 78(11): 1515-1527.

Owens, J.N., Morris, S.J., and Misra, S. 1993. The ultrastructural, histochemical, and biochemical development of the post-fertilization megagametophyte and the zygotic embryo of *Pseudotsuga menziesii*. Canadian Journal of Forest Research 23(5): 816-827.

Owens, J.N., Simpson, S.J., and Misra, S. 1992. The ultrastrucural, histochemical and biochemical development of the post-fertilization megagametophyte and the zygotic embryo of *Pseudotsuga menziesii*. Canadian Journal of Forest Resources 23: 816-827.

Owens, J.N., and Smith, F.H. 1964. The initiation and early development of the seed cone of Douglas-fir. Canadian Journal of Botany 42: 1031-1047.

Pagnussat, G.C., Yu, H.J., Ngo, Q.A., Rajani, S., Mayalagu, S., Johnson, C.S., Capron, A., Xie, L.F., Ye, D., and Sundaresan, V. 2005. Genetic and molecular identification of genes required for female gametophyte development and function in *Arabidopsis*. Development 132: 603–614.

Palowski, B. 2000. Seed yield from polluted stands of *Pinus sylvestris* L. New Forests 20(1): 15-22.

Patrick, J.W., and Offler, C.E. 2001. Compartmentation of transport and transfer events in developing seeds. Journal of Experimental Botany 52: 551-564.

Paulson, A.R., von Aderkas, P., and Perlman, S.J. 2014. Bacterial associates of seed parasitic-wasps (Hymenoptera: *Megastigmus*). BMC Microbiology 14: 224

Pecinka, A., and Scheid, O.M. 2012. Stress-induced chromatin changes: a critical view on their heritability. Plant Cell Physiology 53(5): 801-808.

Pettitt, J.M. 1977. The megaspore wall in gymnosperms: Ultrastructure in some zooidogamous forms. Proceedings of the Royal Society of London. Series B, Biological Sciences 195(1121): 497-515.

Pichot, C., and El Maataoui, M. 1997. Flow cytometric evidence for multiple ploidy levels in the endosperm of some gymnosperm species. Theoretical and Applied Genetics 94: 865–870.

Plomion, C., Bousquet, J., and Kole, C. 2011. Genetics, genomics and breeding of conifers. Science Publishers, Enfield, NH.

Ponce de Leon, I., and Montesano, M. 2013. Activation of defense mechanisms against pathogens in mosses and flowering plants. International Journal of Molecular Science 14(2): 3178-3200.

Ponce de León, I., Oliver, J.P., Castro, A., Gaggero, C., Bentancor, M., and Vidal, S. 2007. *Erwinia carotovora* elicitors and *Botrytis cinerea* activate defense responses in *Physcomitrella patens*. BMC Plant Biology 7: 52-62.

Poulis, B.A.D., O'Leary, S.J.B, Haddow, J.D., and von Aderkas P. 2005. Identification of proteins present in the Douglas fir ovular secretion: an insight into conifer pollen selection and development. International Journal of Plant Sciences 166: 733–739.

Price, P.W. 1977. Concepts on the evolutionary biology of parasites. Evolution 31(2): 405-420.

Prokopy, R.J., Roitber, B.D., and Averill, A.L. 1984. Resource partitioning, pp. 301-330. In Bell W.J., and Carde R.T. [eds.], The chemical ecology of insects. Chapman and Hall, London.

Quiring, D.T., and McNeil, J.N. 1983. Exploitation and interference intraspecific larval competition in the dipteran leaf miner, *Agromyza frontell* (Rondani). Canadian Journal of Zoology 62: 421-427.

Quiring, D.T., and McNeil, J.N. 1987. Foraging behaviour of a Dipteran leaf miner on exploited and unexploited hosts. Oecologia 73: 7-15.

Raitanen, J., Forsman, J.T., Kivelä, S.M. Mäenpää, M.I., and Välimäki, P. 2014. Attraction to conspecific eggs may guide oviposition site selection in a solitary insect. Behavioral Ecology 25(1): 110-116

Ralph, S.G., Yueh, H., Friedmann, M., Aeschliman, D., Zeznik, J.A., Nelson, C.C., Butterfield, Y.S., Kirkpatrick, R., Liu, J., Jones, S.J., Marra, M.A., Doulgas, C.J., Ritland, K., and Bohlmann, J. 2006. Conifer defence against insects: microarray gene expression profiling of Sitka spruce (*Picea sitchensis*) induced by mechanical wounding or feeding by spruce budworms (*Choristoneura occidentalis*) or white pine weevils (*Pissodes strobe*) reveals large-scale changes of the host transcriptome. Plant Cell Environment 29(8): 1545-1570.

Rappaport, N., Mori, S., and Roques, A.1993. Estimating effect of *Megastigmus spermotrophus* (Hymenoptera: torymidae) on Douglas-fir seed production: The new paradigm. Journal of Economic Entomology 86(3): 845-849.

Rappaport, N., and Roques. 1991. Resource use and clonal differences in attack rate by the Douglas-fir seed chalcid, *Megastimgus spermotrophus* Wachtl (Hymenoptera: Torymidae), in France. Canadian Entomologist 123(6): 1219-1228.

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

Reymond, P., Weber H., Damond, M., and Farmer E.E. 2000. Differential gene expression in response to mechanical wounding and insect feeding in *Arabidopsis*. Plant Cell 12: 707-720.

Rice, P., Longden, I., and Bleasby, A. 2000. The European molecular biology open software suite. 2000. Trends in Genetics 16: 276-277.

Ronquist, F. and Liljeblad, J. 2001. Evolution of the gall wasp-host plant association. Evolution 55(12): 2503-2522.

Roques, A. 1981. Distribution and biology of *Megastigmus spermotrophus* Wachtl (Hym chalcid Torymidae) and other insects related to cones in forest stands and seed orchards of Douglas-fir *Pseudotsuga menziesii* (Mirb) Franco in France. Acta Oecologica-oecologia Applicata 2(2): 161-180.

Roques, A. 1986. Dynamique d'infestation des nouveaux vergers à graines de Douglas du sud de la France par le chalcidien *Megastigmus spermotrophus* Wachtl (Hymenoptera, Torymidae). In Proceedings of the 18th International Union of forestry Research Organiziation (IUFRO) World Congress, Ljubljana, Yugoslavia. 685-694.

Roques, A., Auger-Rozenber, M-A., and Boivin, S. 2006. A lack of native congeners may limit colonization of introduced conifers by indigenous insects in Europe. Canadian Journal of Forest Research 36(2): 299-313.

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

Roux, G., Roques, A., Colombari, F., and Menu, F. 1997. Effect of temperature and photoperiod on diapause development in a Douglas fir seed chalcid, *Megastigmus spermotrophus*. Oecologia 111(2): 172–177.

Roux, G. and Roques, A. 1996. Biochemical genetic differentiation among seed chalcid species of genus *Megastigmus* (Hymenoptera: Torymidae). Experientia 52: 522-530.

Ruth, D.S. 1980. A guide to insect pests in Douglas-fir seed orchards. Canadian Forestry Service, Pacific Forest Research Centre, Victoria, British Columbia.

Sabelli, P.A., and Larkins, B.A. 2007. The endoreduplication cell cycle: Regulation and function. In Plant Cell Monograph, Volume 9: Cell Division Control in Plants, eds. D.P.S. Verma and Z. Hong. Springer, Berlin, Germany, 75-100.

Sabelli, P.A., Liu, Y., Dante, R.A., Lizarraga, L.E., Nguyen, H.N., Brown, S.W., Klingler, J.P., Yu, J., LaBrant, E., Layton, T.M., Feldman, M., and Larkins, B.A. 2013. Control of cell proliferation, endoreduplication, cell size, and cell death by the retinoblastoma-related pathway in maize endosperm. Proceedings of the National Academy of Sciences of the United States of America 110(19): E1827–E1836.

Sarvas, R. 1962. Investigations on the flowering and seed crop of *Pinus silvestris*. Communicationes Instituti Forestalia Fennica 53(4): 1-198.

Segura, A., Moreno, M., Madueño, Molina, A., García-Olmedo, F. 1999. Snakin-1, a peptide from potato that is active against plant pathogens. Molecular Plant-Microbe Interactions 12(1): 16-23.

Serrano, M., Coluccia, F., Torres, M., L'Haridon, F., and Métraux, J-P. 2014. The cuticle and plant defense to pathogens. Frontiers in Plant Science 5: 274.

Schneider, C.A., Rasband, W.S., Eliceiri, K.W. 2012. NIH Image to ImageJ: 25 years of image analysis. Nature Methods 9: 671-675.

Schönrogge, K., Harper, L.J., and Lichtenstein C.P. 2000. The protein content of tissues in cynipid galls (Hymenoptera: Cynipidae): similarities between cynipid galls and seeds. Plant, Cell and Environment 23: 215-222.

Schowalter, T.D., and Haverty, M.I. 1989. Influence of host genotype on Douglas-fir seed losses to *Contarinia oregonensis* (Diptera: Cecidomyiidae) and *Megastigmus spermotrophus* (Hymenoptera: Torymidae) in Western Oregon. Environmental Entomology 18(1): 94-97.

Schowalter, T.D., Haverty, M.I., Dombrosky, S.A., and Sexton, J. 1986. Response of Douglas-fir cone gall midge and Douglas-fir seed chalcid to host plant genotype. In Roques, A. (Ed.), Proceedings of the Second IUFRO Cone and Seed Insects Working Party Converence, Briancon, France, 217- 223.

Schowalter, T.D., Haverty, M.I., and Koerber, T.W. 1985. Cone and seed insects in Douglas-fir, *Pseudotsuga menziesii* (Mirb.) Franco, seed orchards in the western United States: Distribution and relative impact. Canadian Entomologist 117:1223-2330.

Siloto, R.M.P., Findlay, K., Lopez-Villalobos, A., Yeung, E.C., Nykiforuk, C.L., and Moloney, M.M. 2006. The accumulation of oleosins determines the size of seed oilbodies in *Arabidopsis*. Plant Cell 18: 1961-1974.

Singh, H. 1978. Embryology of gymnosperms. In Gebrüder Borntraeger Stuttgart, Gebrüder Borntraeger, Berlin.

Stone, G.N., and Schönrogge, K. 2003. The adaptive significance of insect gall morphology. Trends in Ecology and Evolution 18(10): 512–522.

Sturn, A., Quackenbush, J., and Trajanoski, Z. 2002. Genesis: cluster analysis of microarray data. Bioinformatics 18: 207-208.

Suarez, M.F., Filonova, L.H., Smertenko, A., Savenkov, E.I., Clapham, D.H., von Arnold, S., Zhivotovsky, B., and Bozhkov, P.V. 2004. Metacaspase-dependent programmed cell death is essential for plant embryogenesis. Current Biology 14: 339–340.

Titarenko, E., Rojo, E., León, J., and Sánchez-Serrano, J.J. 1997. Jasmonic acid-dependent and independent signaling pathways control wound-induced gene activation in *Arabidopsis thaliana*. Plant Physiology 115: 817-826.

Thiéry, D., and Marion-Poll, F. 1998. Electroantennogram responses of Douglas-fir seed chalcids to plant volatiles. Journal of insect physiology 44: 483–490.

Thompson, K. 1987. The resource ratio hypothesis and the meaning of competition. Functional Ecology 1: 297-303.

Thompson, G.A., and Goggin, F.L. 2006. Transcriptomics and functional genomics of plant defence induction by phloem-feeding insects. Journal of Experimental Botany 57: 755–766.

Thompson, J.N., and Pellmyr, O. 1991. Evolution of oviposition behaviour and host preference in Lepidoptera. Annual Review of Entomology 36: 65-89.

Traveset, A. 1993. Deceptive fruits reduce seed predation by insects in *Pistacia terebinthus* L. (Anacardiaceae). Evolutionary Ecology 7: 357-361.

Turgeon, J.J., Roques, A., and De Groot, P. 1994. Insect fauna of coniferous seed cones: Diversity, host plant interactions and management. Annual Review of Entomology 39: 179–2112.

von Aderkas, P. and J.M. Bonga. 1988. Morphological definition of phenocritical period of initiation of haploid embryongenic tissue from explants of *Larix decidua*. Proceedings of the IUFRO Working Party S2. 07-07 Somatic Cell Genetics of Woody Plants, 29-38. Kluwer Academic Publishers. Dordrecht, The Netherlands.

von Aderkas, P., and Leary, C. 1999. Micropylar exudates in Douglas-fir – timing and volume of production. Secular Plant Reproduction 11: 354-356.

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

von Aderkas, P., Rouault, G., Wagner, R., Chiwocha, S. and Roques, A. 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.

Vuosku, J., Sarjala, T., Jokela, A., Sutela, S., Saaskilahti, M., Suorsa, M, Laara, E., Haggman, H. 2009. One tissue, two fates: different roles of megagametophyte cells during scots pine embryogenesis. Journal of Experimental Botany 60: 1375-1386.

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

Wang, X., Kuang, T., He, Y. 2010. Conservation between higher plants and moss *Physcomitrella patens* in response to the phytohormone abscisic acid: a proteomic analysis. BMC Plant Biology 10: 192-202.

War, A.R., Paulraj, M.G., Ahmad, T., Buhroo, A.A., Hussain, B., Ignacimuthu, S., and Sharma, H.C. 2012. Mechanisms of plant defense against insect herbivores. Plant Signaling and Behavior 7(10): 1306-1320.

Werner, D., Gerlitz, N., and Stadler, R. 2011. A dual switch in phloem unloading during ovule development in *Arabidopsis*. Protoplasma 248: 225-235.

Werner, T., Motyka, V., Strnad, M., and Schmülling, T. 2001. Regulation of plant growth by cytokinin. Proceedings of the National Academy of Science of the United States of America 98(18): 10487-0492.

Wilson, M.F., and Burley, N. 1983. Mate choice in plants: tactics, mechanisms and consequences. Princeton University Press, Princeton, New Jersey, USA.

Winter, P.S., Bowman, C.E., Villani, P.J., Dolan, T.E., and Hauck, N.R. 2014. Systemic acquired resistance in moss: Further evidence for conserved defense mechanisms in plants. PLoSOne 9(7): e101880.

Yamaguchi, H., Tanaka, H., Hasegawa, M., Tokuda, M., Asami, T., and Suzuki, Y. 2012. Phytohormones and willow induction by a gall inducing sawfly. New Phytologist 196(2): 586-595.

Zhang, W-H., Zhou, Y., Dibley, K.E., Tyerman, S.D., Furbank R.T., and Patrick, J.W. 2007. Nutrient loading of developing seeds. Functional Plant Biology 34: 314–331.

Zhang, X., Garreton, V., and Chua, N-H. 2005. The AIP2 E3 ligase acts as a novel negative regulator of ABA signalling by promoting ABI3 degradation. Genes and Development 19(3): 1532-1543.

Zhao, T., Krokene, P., Hu, J., Christiansens, E., Björklund, N., Långström, B., Solheim, H, Borg-Karlson, A-K. 2011. Induced terpene accumulation in Norway spruce inhibits bark beetle colonization in a dose-dependent manner. PLoSOne 6(10): e26649.

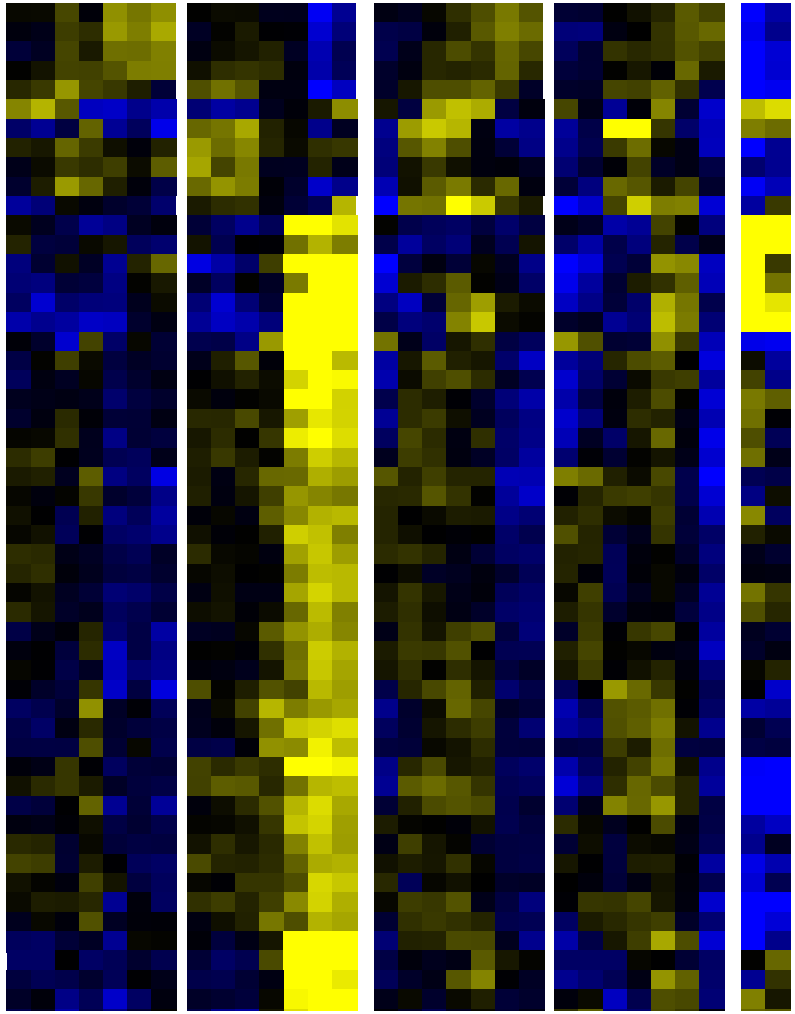
Zhen, W-Q., D-W. Huang, J-H., Xiao, D-R. Yang, C-D. Zhu, and H. Xiao. 2005. Ovipositor length of three Apocrypta species: effect on oviposition behaviour and correlation with syconial thickness. Entomology 33(2):113-120.

Zimin, A., Stevens K., Crepeau, M., Holtz-Morris A., Korabline, M., Marcais, G., Puiu, D., Roberts, M., Wegrzyn, J.L, de Jong, P.J., Neale, D.B., Salzberg, S.L., Yorke, J.A., and Langly, C.H. 2014. Sequencing and assembly of the 22-Gb loblolly pine genome. Genetics 196: 875-890.

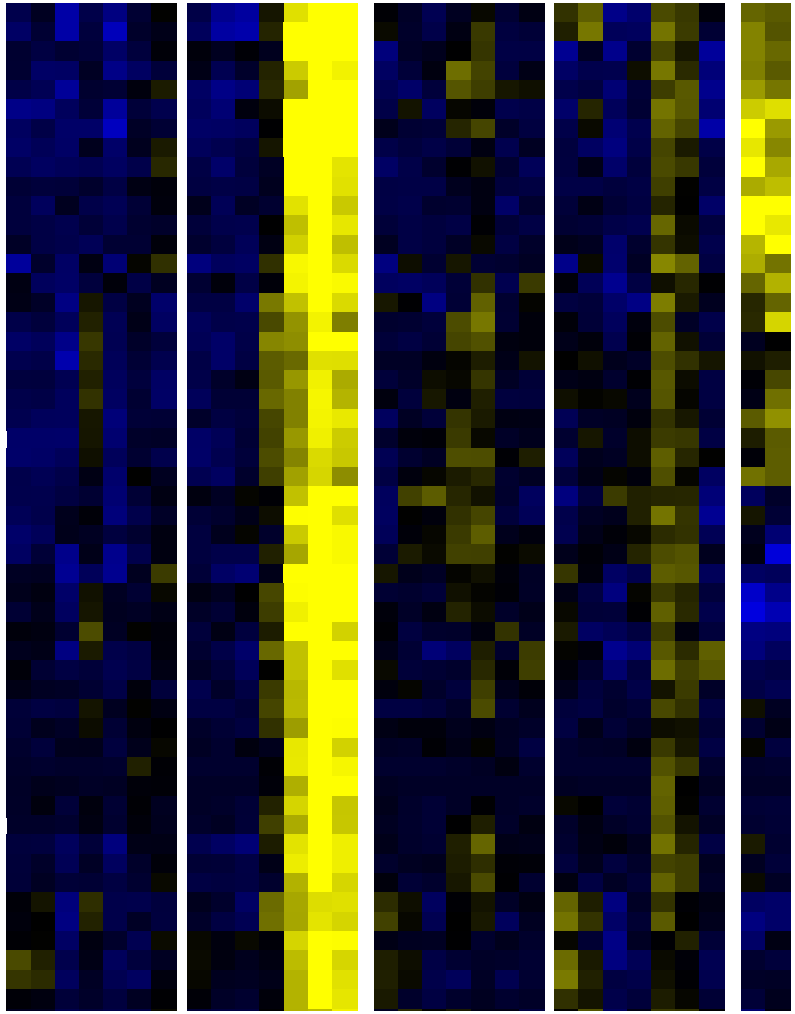
Appendix 1: Supplementary data

Figure S1 Complete list of all Douglas-fir transcripts identified to be either involved in defense or manipulated by *M. spermotrophus.* A heat map with expression data across all samples used to generate the assembly is included along with a gene description, database where the annotation was derived, and the gene identifier from the database. Transcripts with sequence similarity to the *P. taeda* genome were identified.

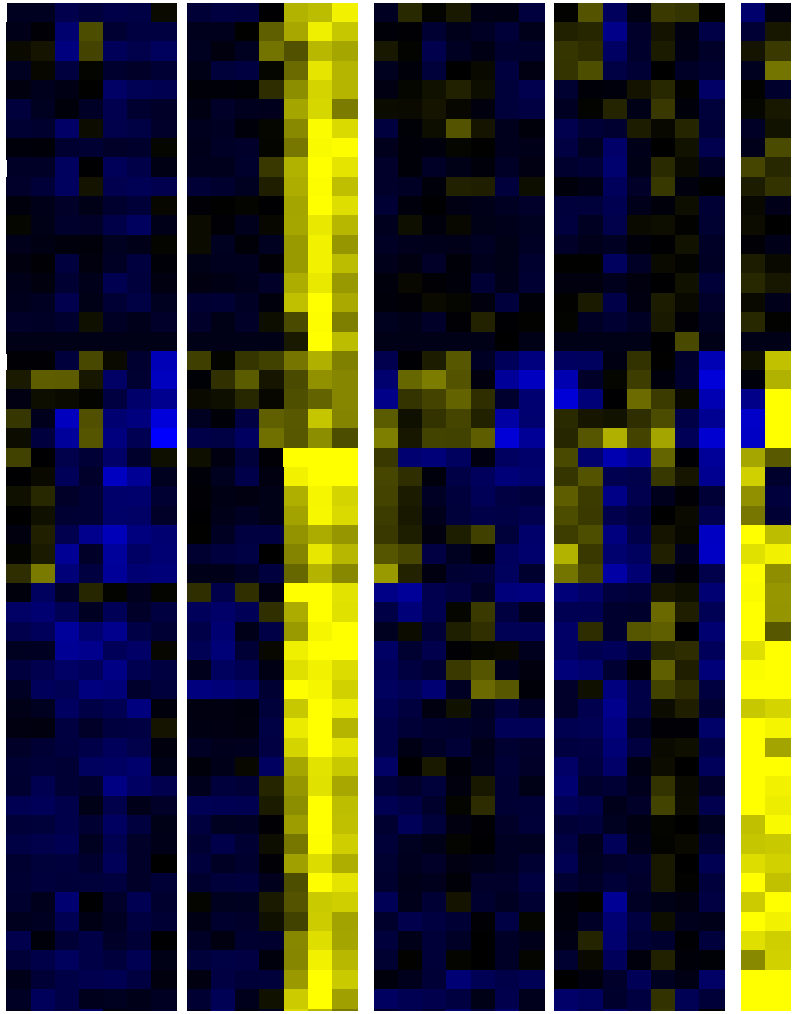
	Not Infested												Infested															
	Pollinated (P)							Unpollinated (U)							Pollinated (PI)							Unpollinated (UI)							Integ OSBC			
Contig ID	June 18	June 21	June 25	June 28	July 3	July 6	July 10	June 18	June 21	June 25	June 28	July 3	July 6	July 10	June 18	June 21	June 25	June 28	July 3	July 6	July 10	June 18	June 21	June 25	June 28	July 3	July 6	July 10	July 6	July 6	Gene description	Identifier	Database	Hit to *taeda*
Putative genes manipulated by M. spermotrophus																																		
comp171794_c1																															alpha carbonic anhydrase 7	AT1G08080.1	TAIR	Yes
comp221345_c0																															ENTH/ANTH/VHS superfamily protein	AT4G25940.1	TAIR	No
comp179431_c1																															RING/U-box superfamily protein	AT5G42200.1	TAIR	No
comp190858_c0																															1-Cys peroxiredoxin A	P0C5C9	Uniprot	Yes
comp208313_c0																															Flavonol synthase/flavanone 3-hydroxylase	Q9ZWQ9	Uniprot	Yes
comp169016_c0																															Glucose and ribitol dehydrogenase	Q5KTS5	Uniprot	Yes
comp202055_c0																															HSP20-like chaperones superfamily protein	AT1G53540.1	TAIR	Yes
comp190076_c0																															Glucose and ribitol dehydrogenase	Q5KTS5	Uniprot	Yes
comp186025_c0																															Oleosin family protein	AT2G25890.1	TAIR	Yes
comp184454_c0																															LEA family protein	AT5G44310.2	TAIR	Yes
comp217681_c0																															2S seed storage-like protein	P26986	Uniprot	No
comp214925_c0																															RmlC-like cupins superfamily protein	AT2G28490.1	TAIR	Yes
comp204497_c0																															Glutamate decarboxylase	Q07346	Uniprot	Yes
comp211027_c1																															Protein of unknown function	AT1G29680.1	TAIR	Yes
comp205381_c0																															LEA protein	AT4G36600.1	TAIR	Yes
comp215723_c0																															AP2/B3-like transcriptional factor family protein	AT3G24650.1	TAIR	Yes
comp220538_c0																															Embryonic protein DC-8	P20075	Uniprot	Yes
comp212371_c0																															hydroxysteroid dehydrogenase 5	AT4G10020.1	TAIR	Yes
comp171788_c0																															PEBP family protein	AT5G62040.1	TAIR	Yes
comp213171_c2																															Peroxidase superfamily protein	AT1G71695.1	TAIR	Yes
comp197459_c0																															O-methyltransferase family protein	AT1G77520.1	TAIR	No
comp210765_c0																															Octicosapeptide/Phox/Bem1p family protein	AT4G05150.1	TAIR	Yes
comp206029_c0																															Cytokinin dehydrogenase 3	Q8LNV6	Uniprot	Yes
comp120136_c0																															enzyme binding;tetrapyrrole binding	AT3G59400.1	TAIR	Yes
comp199402_c2																															Leucine-rich repeat protein kinase family protein	AT2G37050.3	TAIR	Yes
comp202994_c0																															heat shock protein 17.6A	AT5G12030.1	TAIR	Yes
comp119636_c0																															cold, circadian rhythm, and rna binding 2	AT2G21660.1	TAIR	Yes
comp180041_c0																															AMP-dependent synthetase and ligase family protein	AT1G20480.1	TAIR	No
comp213031_c0																															glutathione peroxidase 8	AT1G63460.1	TAIR	Yes
comp222381_c0																															Beta-glucosidase 22	Q60DX8	Uniprot	No
comp185456_c0																															Vacuolar cation/proton exchanger 1a	Q769E5	Uniprot	Yes
comp209654_c0																															methyltransferases superfamily protein	AT5G19530.1	TAIR	No
comp221690_c2																															Subtilase family protein	AT2G05920.1	TAIR	Yes
comp191797_c0																															Auxin efflux carrier family protein	AT5G01990.1	TAIR	Yes
comp122800_c0																															Unknown Protein	AT5G66780.1	TAIR	Yes
comp199145_c0																															17.6 kDa class II heat shock protein	AT5G12020.1	TAIR	Yes
comp215172_c0																															17.5 kDa class I heat shock protein	P04794	Uniprot	Yes
comp184948_c0																															PEBP family protein	AT5G01300.1	TAIR	Yes
comp179393_c0																															Exostosin family protein	AT5G25310.1	TAIR	Yes
comp197077_c1																															translationally controlled tumor protein	AT3G16640.1	TAIR	Yes
comp208331_c0																															LEA domain-containing protein	AT1G72100.1	TAIR	Yes
comp198426_c1																															heat shock protein 21	AT4G27670.1	TAIR	Yes
comp205206_c0																															Ribulose bisphosphate carboxylase	P16031	Uniprot	Yes
comp119243_c0																															NAD(P)-binding Rossmann-fold superfamily protein	AT2G23910.1	TAIR	Yes
comp194171_c8																															Integrase-type DNA-binding superfamily protein	AT1G79700.2	TAIR	Yes
comp192895_c0																															Unknown Protein	AT5G17190.1	TAIR	No
comp218841_c0																															nucleoside triphosphate hydrolase	AT5G60760.1	TAIR	Yes
comp171382_c0																															Putative glutathione formaldehyde enzyme	E1VBT6	Uniprot	Yes
comp118831_c0																															phosphatidic acid phosphatase-related	AT3G15820.1	TAIR	Yes
comp170370_c0																															Unknown Protein	AT1G63060.1	TAIR	Yes



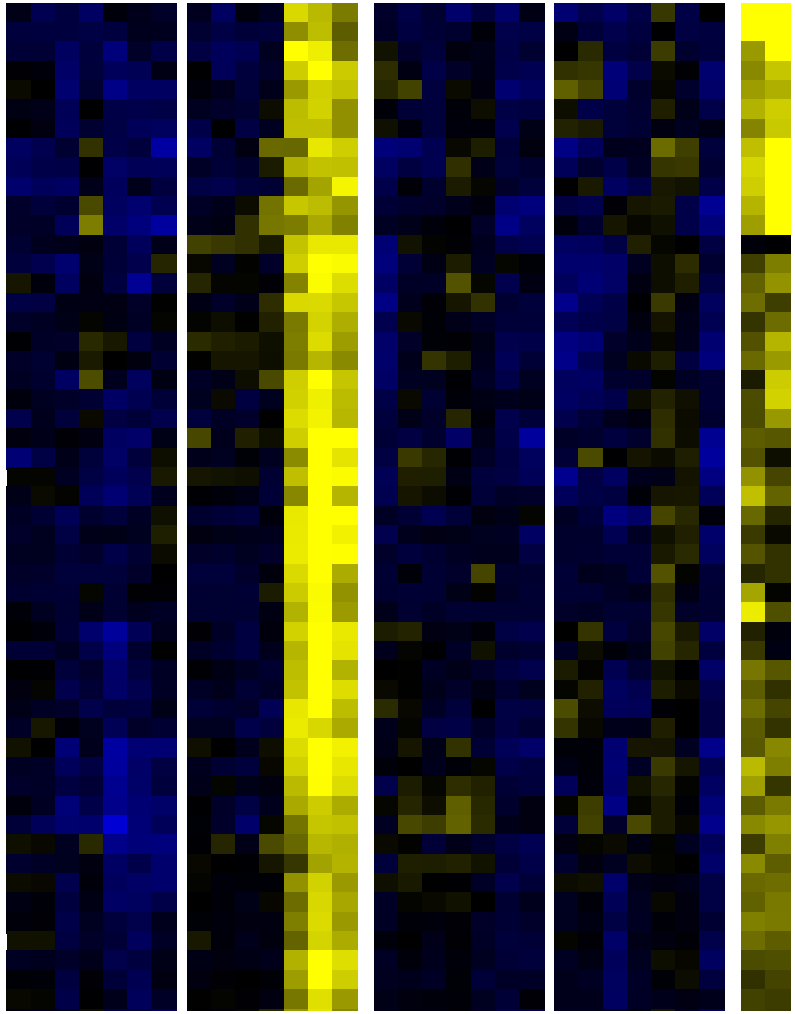
Contig ID	Gene description	Identifier	Database	Hit to *taeda*
comp214500_c0	zinc finger protein 6	AT1G67030.1	TAIR	Yes
comp216697_c0	Erythronate-4-phosphate dehydrog family protein	AT1G01500.1	TAIR	Yes
comp185653_c0	vacuolar iron transporter 1	AT2G01770.1	TAIR	Yes
comp220188_c0	Thaumatin-like protein 1	P83332	Uniprot	No
comp207431_c2	DNA replication licensing factor MCM3 homolog 1	Q43704	Uniprot	Yes
comp219349_c0	Bidirectional sugar transporter SWEET3b	Q5NAZ9	Uniprot	Yes
comp195339_c0	Probable indole-3-acetic acid-amido synthetase	O82333	Uniprot	Yes
comp208744_c0	TRICHOME BIREFRINGENCE-LIKE 36	AT3G54260.1	TAIR	Yes
comp194353_c0	NAD(P)-binding Rossmann-fold superfamily protein	AT4G30470.1	TAIR	Yes
comp192815_c0	Unknown Protein	AT5G36710.1	TAIR	Yes
comp209708_c0	beta-galactosidase 3	AT4G36360.1	TAIR	Yes
comp210486_c1	Dihydroflavonol-4-reductase	P51110	Uniprot	Yes
comp167513_c0	Heavy metal transport/detox superfamily protein	AT1G71050.1	TAIR	Yes
comp205323_c0	Eukaryotic aspartyl protease family protein	AT2G03200.1	TAIR	Yes
comp210067_c0	caleosin-related family protein	AT5G29560.1	TAIR	No
comp210932_c0	Glycosyl hydrolase family protein	AT5G09700.1	TAIR	Yes
comp213336_c1	UDP-Glycosyltransferase superfamily protein	AT2G36780.1	TAIR	Yes
comp120349_c0	glutamine dumper 2	AT4G25760.1	TAIR	No
comp208621_c0	Ras-related protein Rab7	Q43463	Uniprot	Yes
comp220185_c1	C2H2 and C2HC zinc fingers superfamily protein	AT4G17810.1	TAIR	Yes
comp217298_c0	Oryzain alpha chain	P25776	Uniprot	Yes
comp200648_c0	Eukaryotic aspartyl protease family protein	AT2G03200.1	TAIR	Yes
comp201934_c1	metacaspase 9	AT5G04200.1	TAIR	Yes
comp202393_c0	PLATZ transcription factor family protein	AT4G17900.1	TAIR	No
comp212261_c0	GDSL-like Lipase/Acylhydrolase superfamily protein	AT4G01130.1	TAIR	Yes
comp221358_c0	Carbohydrate-binding X8 domain superfamily protein	AT5G35740.1	TAIR	No
comp180883_c0	NAC domain containing protein 100	AT5G61430.1	TAIR	Yes
comp195785_c0	Sec14p-like phosphatidylinositol transfer family protein	AT1G05370.1	TAIR	Yes
comp216159_c0	Major facilitator superfamily protein	AT3G60070.1	TAIR	Yes
comp222828_c0	Retrovirus-related Pol polyprotein	P10978	Uniprot	Yes
comp210967_c1	TRF-like 9	AT3G12560.1	TAIR	Yes
comp218867_c0	Purple acid phosphatase 15	Q9SFU3	Uniprot	Yes
comp220605_c0	RING/U-box superfamily protein	AT5G04460.1	TAIR	Yes
comp213471_c0	Chaperone DnaJ-domain superfamily protein	AT1G80920.1	TAIR	Yes
comp219143_c4	MATE efflux family protein	AT5G65380.1	TAIR	Yes
comp212585_c0	12-oxophytodienoate reductase 7	Q6Z965	Uniprot	Yes
comp207869_c0	Probable xyloglucan endotransglucosylase/hydrolase	Q8LNZ5	Uniprot	Yes
comp205378_c0	Subtilisin-like serine endopeptidase family protein	AT1G20160.1	TAIR	Yes
comp207395_c2	SAUR-like auxin-responsive protein family	AT5G20820.1	TAIR	No
comp208430_c0	Transmembrane amino acid transporter family protein	AT5G16740.1	TAIR	Yes
comp215955_c0	Galactosyltransferase family protein	AT5G57500.1	TAIR	Yes
comp171991_c0	Basic blue protein	P00303	Uniprot	No
comp219920_c1	Cupredoxin superfamily protein	AT1G23010.1	TAIR	Yes
comp206667_c1	anthocyanin 5-aromatic acyltransferase 1	AT5G61160.1	TAIR	No
comp221868_c0	DNAse I-like superfamily protein	AT3G21530.1	TAIR	Yes
comp212952_c2	Walls Are Thin 1	AT1G75500.1	TAIR	Yes
comp217744_c2	UDP-glucosyl transferase 85A2	AT1G22360.1	TAIR	Yes
comp206523_c0	Pectin lyase-like superfamily protein	AT5G19730.1	TAIR	Yes
comp199663_c0	homeobox from Arabidopsis thaliana	AT5G06710.1	TAIR	Yes
comp221598_c2	Toll-Interleukin-Resistance domain family protein	AT1G52900.1	TAIR	No
comp203603_c0	sedoheptulose-bisphosphatase	AT3G55800.1	TAIR	Yes
comp215492_c0	Cysteine proteinases superfamily protein	AT1G02305.1	TAIR	Yes



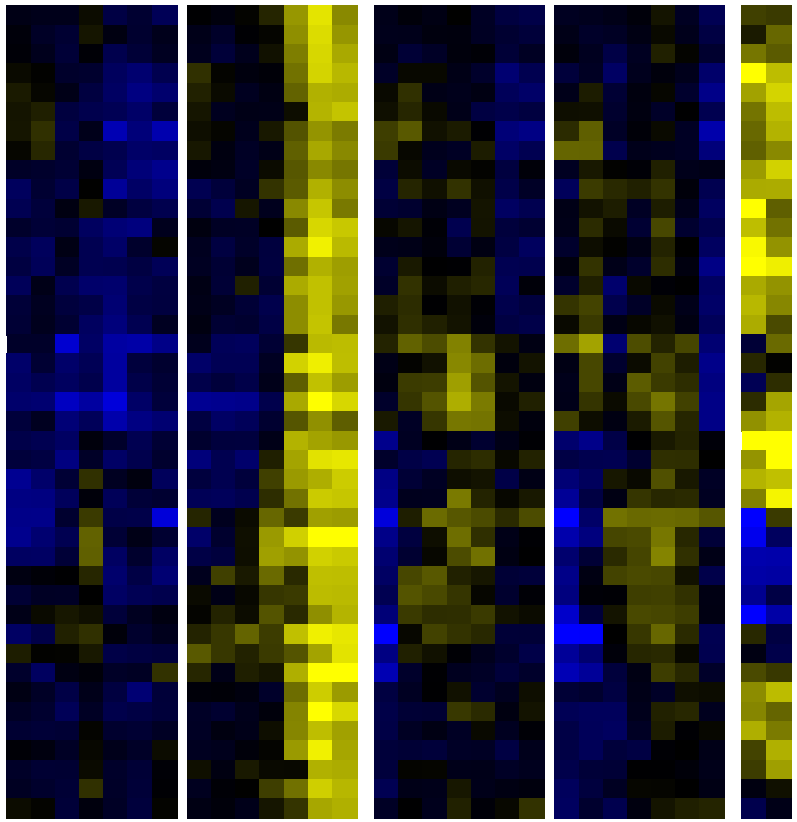
Contig ID	Gene description	Identifier	Database	Hit to *taeda*
comp212815_c1	PLATZ transcription factor family protein	AT1G76590.1	TAIR	Yes
comp212815_c0	PLATZ transcription factor family protein	AT1G32700.1	TAIR	Yes
comp189602_c1	NAC transcription factor ONAC010	A2YMR0	Uniprot	Yes
comp220743_c0	cytochrome P450, family 709	AT2G46950.1	TAIR	Yes
comp202028_c0	GDSL-like Lipase/Acylhydrolase superfamily protein	AT5G08460.1	TAIR	Yes
comp222405_c0	Seed linoleate 9S-lipoxygenase-2	P09439	Uniprot	Yes
comp207788_c0	response to low sulfur 2	AT5G24660.1	TAIR	Yes
comp211541_c0	Probable vacuolar amino acid transporter YPQ2	Q06328	Uniprot	Yes
comp213783_c0	Unknown Protein	AT2G38640.1	TAIR	No
comp207894_c1	LRR and NB-ARC domainsProtein	AT1G61180.2	TAIR	No
comp205418_c0	OSBP-related protein 4C	AT5G57240.1	TAIR	Yes
comp203667_c0	zinc finger protein 8	AT2G41940.1	TAIR	Yes
comp218846_c2	(TIR-NBS-LRR class) protein	AT2G17050.1	TAIR	Yes
comp218799_c2	Leucine-rich receptor-like protein kinase	AT2G24130.1	TAIR	Yes
comp221598_c4	transmembrane receptors;ATP binding	AT1G27170.1	TAIR	Yes
comp220897_c0	Leucine-rich receptor-like protein kinase	AT5G46330.1	TAIR	Yes
comp192161_c0	ATP-dependent zinc metalloprotease FTSH 8	Q8W585	Uniprot	Yes
comp215836_c2	Endosomal targeting protein	AT1G17940.1	TAIR	Yes
comp163036_c0	Heavy metal transport/detox superfamily protein	AT1G01490.1	TAIR	Yes
comp202320_c1	alpha/beta-Hydrolases superfamily protein	AT5G21950.1	TAIR	Yes
comp202320_c0	alpha/beta-Hydrolases superfamily protein	AT2G18360.1	TAIR	Yes
comp210774_c0	Octicosapeptide/Phox/Bem1p family protein	AT3G48240.1	TAIR	Yes
comp214732_c0	unknown protein	AT2G35260.1	TAIR	Yes
comp204942_c0	Catalase	O24339	Uniprot	Yes
comp211118_c0	non-intrinsic ABC protein 3	AT1G67940.1	TAIR	Yes
comp208996_c0	myb domain protein 103	AT1G63910.1	TAIR	Yes
comp209641_c0	HAD superfamily, subfamily IIIB acid phosphatase	AT4G29260.1	TAIR	Yes
comp180129_c1	UPF0496 protein 3	A2XCJ1	Uniprot	Yes
comp218664_c0	Chaperone DnaJ-domain superfamily protein	AT1G56300.1	TAIR	Yes
comp201089_c0	peptidyl-prolyl cis-trans isomerase	AT2G15790.1	TAIR	Yes
comp206682_c1	cellulose synthase A2	AT4G39350.1	TAIR	Yes
comp215130_c0	cellulose synthase 1	AT4G32410.1	TAIR	No
comp199870_c0	Heavy metal transport/detox superfamily protein	AT5G50740.3	TAIR	Yes
comp216295_c0	DNAse I-like superfamily protein	AT3G63240.1	TAIR	Yes
comp193032_c0	Nicotianamine synthase	Q9XGI7	Uniprot	Yes
comp196933_c0	Naringenin,2-oxoglutarate 3-dioxygenase	Q06942	Uniprot	Yes
comp118901_c0	RING-H2 finger A3A	AT2G17450.1	TAIR	No
comp217903_c1	Probable E3 ubiquitin-protein ligase BAH1-like 1	Q7XZZ3	Uniprot	Yes
comp210451_c0	Peptidase M20/M25/M40 family protein	AT4G17830.2	TAIR	Yes
comp196028_c0	Vacuolar iron transporter (VIT) family protein	AT3G43660.1	TAIR	Yes
comp168185_c0	Protein RSI-1	P47926	Uniprot	No
comp182070_c0	Unknown Protein	AT1G54740.1	TAIR	Yes
comp213314_c0	Major facilitator superfamily protein	AT1G18880.1	TAIR	Yes
comp192407_c0	homeobox 3	AT5G15150.1	TAIR	Yes
comp197603_c0	Protein kinase superfamily protein	AT3G07700.3	TAIR	Yes
comp198472_c0	phosphotransferases / phosphatidylinositol kinases	AT4G36080.1	TAIR	No
comp161772_c1	MATE efflux family protein	AT1G71870.1	TAIR	Yes
comp161772_c0	MATE efflux family protein	AT1G71870.1	TAIR	Yes
comp212037_c1	RNase H family protein	AT1G24090.1	TAIR	Yes
comp218052_c1	serine/threonine protein kinase 2	AT3G08720.1	TAIR	Yes
comp218052_c0	serine/threonine protein kinase 2	AT3G08720.1	TAIR	Yes
comp220035_c0	Putative ribonuclease H protein	P0C2F6	Uniprot	No



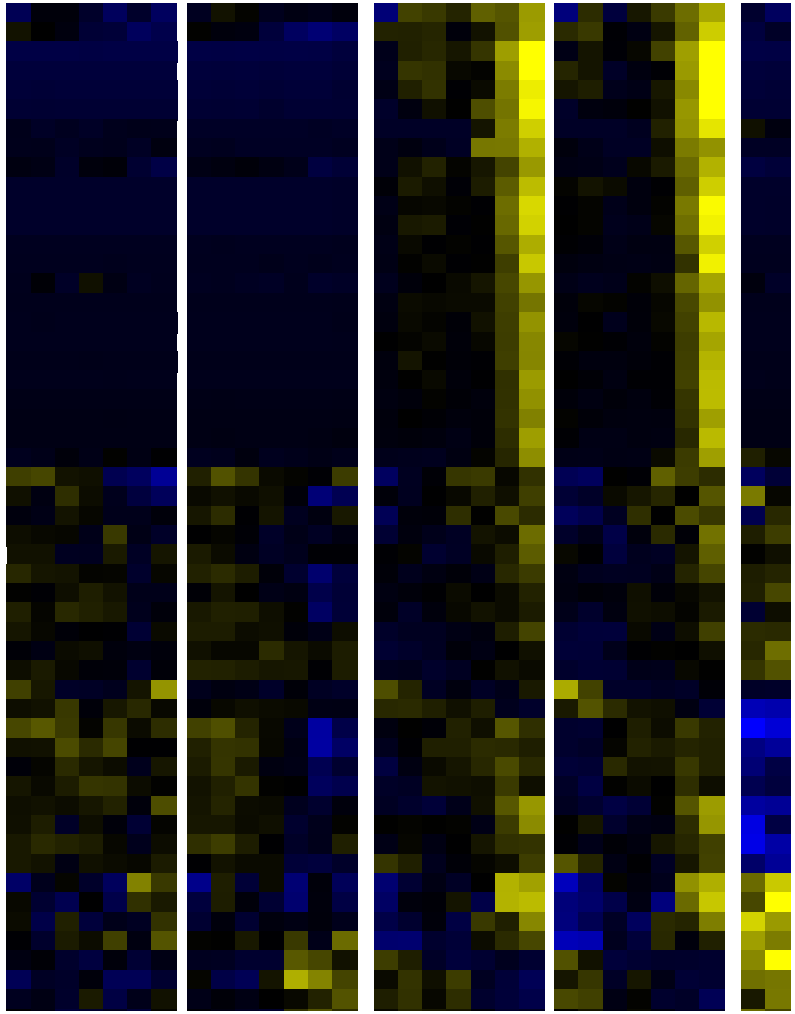
Contig ID	Gene description	Identifier	Database	Hit to *taeda*
comp216793_c1	LOB domain-containing protein 40	AT1G67100.1	TAIR	No
comp209706_c0	Autophagy-related protein 13	AT3G18770.1	TAIR	Yes
comp208182_c0	root hair specific 2	AT1G12950.1	TAIR	Yes
comp212250_c0	Heavy metal transport/detox superfamily protein	AT1G01490.1	TAIR	Yes
comp220339_c1	Aspartate aminotransferase 1	P28011	Uniprot	Yes
comp207550_c1	Unknown Protein	AT1G04000.1	TAIR	Yes
comp186928_c0	Quinone reductase family protein	AT4G27270.1	TAIR	Yes
comp216836_c2	K+ efflux antiporter 3	AT4G04850.2	TAIR	Yes
comp207505_c1	Major facilitator superfamily protein	AT1G62200.1	TAIR	Yes
comp202729_c0	PAS/LOV protein B	AT2G02710.1	TAIR	Yes
comp211317_c1	SNARE associated Golgi protein family	AT4G14950.1	TAIR	Yes
comp207496_c0	Citrate synthase 3, peroxisomal	Q9SJH7	Uniprot	Yes
comp207406_c0	Unknown Protein	AT3G04350.1	TAIR	Yes
comp207893_c2	Unknown Protein	AT5G06440.4	TAIR	Yes
comp204227_c0	DegP protease 1	AT3G27925.1	TAIR	Yes
comp205477_c4	hydrolase superfamily protein	AT2G38740.1	TAIR	Yes
comp188254_c0	BTB/POZ domain-containing protein	AT4G08455.1	TAIR	Yes
comp122874_c1	laccase 5	AT2G40370.1	TAIR	Yes
comp211502_c0	RING/U-box superfamily protein	AT1G72310.1	TAIR	Yes
comp199509_c1	serine-rich protein-related	AT5G11090.1	TAIR	Yes
comp221806_c0	ABC transporter G family member 34	Q7PC87	Uniprot	Yes
comp209459_c0	Beta-phellandrene synthase, chloroplastic	Q9M7D1	Uniprot	Yes
comp195498_c0	Pinene synthase, chloroplastic	Q6XDB5	Uniprot	Yes
comp211467_c0	beta vacuolar processing enzyme	AT1G62710.1	TAIR	Yes
comp124663_c0	inosine-uridine hydrolase family protein	AT5G18860.1	TAIR	Yes
comp209528_c1	inosine-uridine hydrolase family protein	AT5G18870.1	TAIR	Yes
comp223318_c0	inosine-uridine hydrolase family protein	AT5G18860.1	TAIR	Yes
comp213804_c0	Chitinase 5	Q7Y1Z0	Uniprot	Yes
comp216806_c1	myb domain protein 3	AT1G22640.1	TAIR	No
comp219943_c1	Translation initiation factor SUI1 family protein	AT1G54290.1	TAIR	Yes
comp206714_c0	ribonuclease 3	AT1G26820.1	TAIR	Yes
comp120100_c0	Probable deoxyribonuclease RhsA	E0SAK8	Uniprot	Yes
comp218673_c0	glyoxal oxidase-related protein	AT1G19900.1	TAIR	Yes
comp219498_c0	Alpha-aminoadipic semialdehyde synthase	Q9SMZ4	Uniprot	Yes
comp214539_c0	Probable acyl-activating enzyme 9	Q9LPK6	Uniprot	Yes
comp217550_c0	serine/threonine-protein kinase	O64780	Uniprot	Yes
comp208422_c1	zinc induced facilitator-like 2	AT3G43790.3	TAIR	Yes
comp222680_c0	purple acid phosphatase 3	AT1G14700.1	TAIR	Yes
comp201487_c0	aluminum-activated malate transporter 9	AT3G18440.1	TAIR	Yes
comp209400_c0	gamma-glutamyl hydrolase 1	AT1G78660.1	TAIR	Yes
comp169521_c0	arogenate dehydratase 6	AT1G08250.1	TAIR	Yes
comp123457_c0	Protein of unknown function (DUF1677)	AT5G20670.1	TAIR	No
comp212129_c0	Mitochondrial substrate carrier family protein	AT1G07030.1	TAIR	Yes
comp205257_c0	molybdate transporter 1	AT2G25680.1	TAIR	Yes
comp219075_c1	UDP-glucosyl transferase 89B1	AT1G73880.1	TAIR	Yes
comp191225_c0	myb domain protein 46	AT5G12870.1	TAIR	No
comp210183_c0	ARM repeat superfamily protein	AT1G60190.1	TAIR	Yes
comp210001_c1	alpha/beta-Hydrolases superfamily protein	AT5G16120.2	TAIR	Yes
comp214096_c1	Leucine-rich repeat protein kinase family protein	AT5G39390.1	TAIR	No
comp222992_c0	ABC transporter C family member 10	Q9LYS2	Uniprot	Yes
comp194421_c0	Protein of unknown function (DUF 3339)	AT3G48660.1	TAIR	Yes
comp203346_c0	glyoxal oxidase-related protein	AT3G53950.1	TAIR	Yes



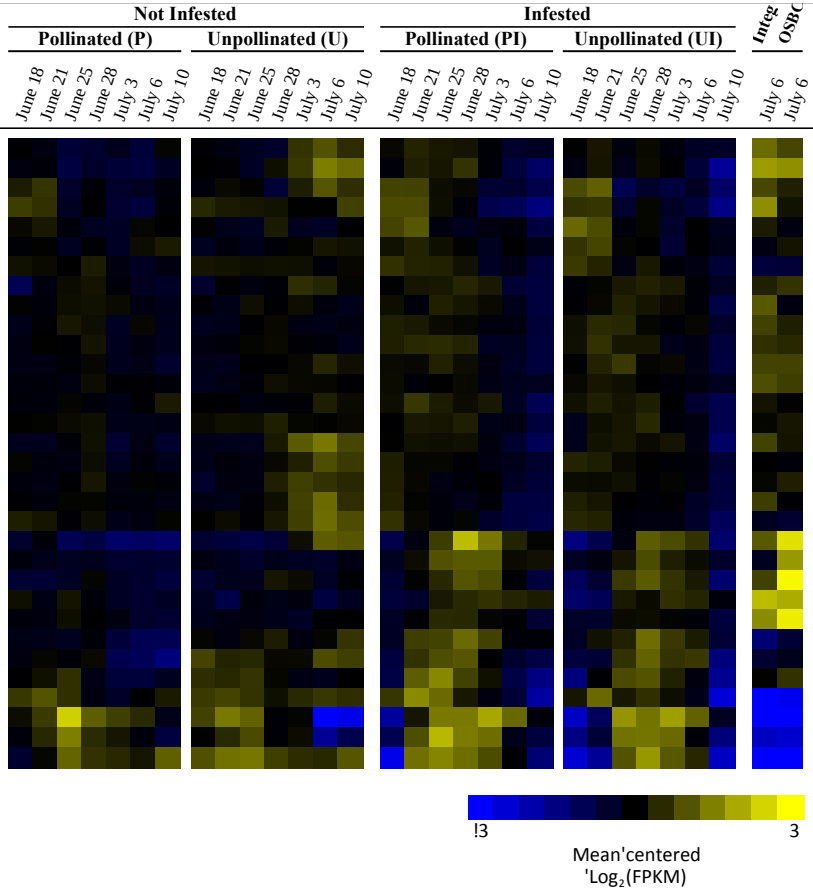
Contig ID	Gene description	Identifier	Database	Hit to *taeda*
comp201557_c1	Nodulin MtN21 /EamA-like transporter family protein	AT5G47470.1	TAIR	Yes
comp199096_c0	EXORDIUM like 3	AT5G51550.1	TAIR	No
comp172409_c0	Pathogenesis-related thaumatin superfamily protein	AT4G38660.1	TAIR	No
comp202384_c0	basic leucine-zipper 42	AT3G30530.1	TAIR	Yes
comp209015_c5	C2H2-like zinc finger protein	AT2G45120.1	TAIR	Yes
comp212720_c1	Protein kinase superfamily protein	AT2G45590.1	TAIR	Yes
comp216000_c2	Major facilitator superfamily protein	AT1G72140.1	TAIR	Yes
comp206756_c0	Homeobox-leucine zipper protein family	AT1G69780.1	TAIR	Yes
comp208380_c0	Linoleate 9S-lipoxygenase A	P38415	Uniprot	Yes
comp209807_c1	Calcium-dependent lipid-binding family protein	AT3G16510.1	TAIR	No
comp222235_c2	pleiotropic drug resistance 11	AT1G66950.1	TAIR	Yes
comp209186_c0	Pleiotropic drug resistance protein 1	Q76CU2	Uniprot	Yes
comp187731_c0	pleiotropic drug resistance 7	AT1G15210.1	TAIR	Yes
comp220431_c1	alpha-L-arabinofuranosidase 1	AT3G10740.1	TAIR	Yes
comp193869_c0	cytochrome P450, family 709	AT2G46950.1	TAIR	Yes
comp210693_c0	Protein of unknown function (DUF607)	AT2G23790.1	TAIR	Yes
comp224795_c0	cation calcium exchanger 4	AT1G54115.1	TAIR	Yes
comp216225_c1	Disease resistance protein (CC-NBS-LRR class) family	AT1G51480.1	TAIR	No
comp216225_c0	NB-ARC domain-containing disease resistance protein	AT4G27190.1	TAIR	No
comp206426_c0	Disease resistance protein (TIR-NBS-LRR class) family	AT3G04220.1	TAIR	Yes
comp211296_c1	transmembrane receptors;ATP binding	AT1G27170.1	TAIR	Yes
comp216565_c0	Disease resistance protein (TIR-NBS-LRR class) family	AT4G14370.1	TAIR	Yes
comp221074_c0	Unknown Protein	AT5G12950.1	TAIR	Yes
comp210769_c0	Family of unknown function (DUF716)	AT1G55230.1	TAIR	Yes
comp205258_c2	NAC domain containing protein 100	AT5G61430.1	TAIR	Yes
comp218250_c1	Leucine-rich receptor-like protein kinase family protein	AT5G46330.1	TAIR	Yes
comp211329_c0	Mitochondrial substrate carrier family protein	AT2G39970.1	TAIR	Yes
comp209957_c2	Beta-glucosidase 25	Q0DA21	Uniprot	Yes
comp201065_c0	Digalactosyldiacylglycerol synthase 1, chloroplastic	Q6DW76	Uniprot	Yes
comp192033_c0	glutamine dumper 1	AT4G31730.1	TAIR	No
comp204446_c0	basic helix-loop-helix DNA-binding superfamily protein	AT2G16910.1	TAIR	Yes
comp123002_c1	glutamine amidotransferase-like superfamily protein	AT4G30550.1	TAIR	Yes
comp212964_c1	xylem cysteine peptidase 1	AT4G35350.1	TAIR	Yes
comp214871_c0	Clp ATPase	AT5G51070.1	TAIR	Yes
comp186915_c1	BTB/POZ domain-containing protein	AT1G01640.1	TAIR	Yes
comp185301_c0	dehydration-induced protein (ERD15)	AT2G41430.1	TAIR	Yes
comp219274_c3	monogalactosyl diacylglycerol synthase 1	AT4G31780.2	TAIR	Yes
comp214594_c0	hAT dimerisation domain-containing protein	AT5G33406.1	TAIR	Yes
comp222060_c0	UDP-Glycosyltransferase superfamily protein	AT2G36970.1	TAIR	Yes
comp216169_c1	Papain family cysteine protease	AT4G39090.1	TAIR	Yes
comp198804_c0	RELA/SPOT homolog 3	AT1G54130.1	TAIR	Yes
comp198230_c0	beta-amylase 3	AT5G18670.1	TAIR	Yes
comp209990_c0	ortholog of sugar beet HS1 PRO-1 2	AT2G40000.1	TAIR	Yes
comp185338_c0	Protein of unknown function, DUF584	AT4G21970.1	TAIR	Yes
comp171495_c0	AIG2-like (avirulence induced gene) family protein	AT3G02910.1	TAIR	Yes
comp211784_c1	Ypt/Rab-GAP domain of gyp1p superfamily protein	AT4G29950.1	TAIR	Yes
comp188107_c0	Unknown Protein	AT1G27290.1	TAIR	Yes
comp192657_c0	Unknown Protein	AT3G26670.1	TAIR	Yes
comp218320_c1	Auxin efflux carrier family protein	AT5G01990.1	TAIR	Yes
comp215774_c0	Carbohydrate-binding protein	AT1G10150.1	TAIR	Yes
comp219854_c1	Unknown Protein	AT2G31130.1	TAIR	Yes
comp203460_c0	Unknown Protein	AT1G06050.1	TAIR	Yes



Contig ID	Gene description	Identifier	Database	Hit to *P. taeda*
comp209121_c0	BTB and TAZ domain protein 4	AT5G67480.2	TAIR	Yes
comp196491_c0	Homeodomain-like superfamily protein	AT1G79430.2	TAIR	Yes
comp208235_c0	AFG1-like ATPase family protein	AT4G30490.1	TAIR	Yes
comp213237_c0	serine carboxypeptidase-like 51	AT2G27920.1	TAIR	Yes
comp189013_c0	plant U-box 17	AT1G29340.1	TAIR	Yes
comp216021_c0	Solute carrier family 15 member 2	P46029	Uniprot	Yes
comp186078_c0	VQ motif-containing protein	AT4G39720.1	TAIR	Yes
comp171269_c0	Glutaredoxin-C1	Q7G8Y5	Uniprot	Yes
comp181411_c0	Zinc finger A20 and AN1 domain-containing protein 6	Q852K5	Uniprot	Yes
comp181899_c0	NDR1/HIN1-like 1	AT3G11660.1	TAIR	Yes
comp198402_c0	RING-H2 finger A1A	AT4G11370.1	TAIR	Yes
comp217600_c1	Pathogenesis-related thaumatin superfamily protein	AT1G73620.1	TAIR	Yes
comp213550_c1	heavy metal atpase 2	AT4G30110.1	TAIR	Yes
comp200066_c0	LEA hydroxyproline-rich glycoprotein family	AT1G65690.1	TAIR	Yes
comp217740_c2	C2H2-like zinc finger protein	AT1G26610.1	TAIR	Yes
comp215255_c0	C2H2-type zinc finger family protein	AT1G02040.1	TAIR	Yes
comp217811_c0	Unknown Protein	AT5G18130.1	TAIR	Yes
comp213183_c0	Homeodomain-like superfamily protein	AT5G37260.1	TAIR	Yes
comp215795_c0	Chaperone DnaJ-domain superfamily protein	AT1G80920.1	TAIR	Yes
comp209159_c0	Galactose oxidase/kelch repeat superfamily protein	AT1G15670.1	TAIR	Yes
comp220761_c0	UDP-Glycosyltransferase superfamily protein	AT1G05675.1	TAIR	Yes
comp197106_c0	Aluminium induced protein with YGL and LRDR motifs	AT5G19140.1	TAIR	Yes
comp219869_c1	Disease resistance protein (CC-NBS-LRR class) family	AT1G12220.1	TAIR	Yes
comp210631_c0	Ammonium transporter 1 member 2	Q9ZPJ8	Uniprot	Yes
comp217389_c1	ethylene responsive element binding factor 1	AT4G17500.1	TAIR	Yes
comp200013_c0	Cytochrome P450 71A1	P24465	Uniprot	Yes
comp103395_c0	Bifunctional inhibitor/lipid-transfer protein/2S albumin	AT1G62500.1	TAIR	No
comp211363_c0	Cytochrome P450 750A1	Q50EK4	Uniprot	Yes
comp220033_c0	Transcription factor HBP-1b(c1)	Q41558	Uniprot	Yes
comp207276_c0	Acyl-CoA N-acyltransferases (NAT) superfamily protein	AT2G23060.1	TAIR	Yes
comp200451_c0	Cytochrome P450 superfamily protein	AT2G26710.1	TAIR	Yes
comp209664_c1	Pectinesterase 2	Q42534	Uniprot	Yes
comp205371_c0	Bidirectional sugar transporter SWEET3b	Q5NAZ9	Uniprot	Yes
comp178663_c0	Vacuolar iron transporter (VIT) family protein	AT3G43660.1	TAIR	Yes
comp215867_c0	phosphate 1	AT3G23430.1	TAIR	Yes
comp202691_c0	Ascorbate-specific transmembrane electron transporter	C4IYS8	Uniprot	Yes
comp208719_c0	Encodes a protein involved in salt tolerance	AT5G02020.1	TAIR	Yes
comp215744_c2	Unknown Protein	AT4G12840.1	TAIR	Yes
comp208429_c1	nodulin MtN21 /EamA-like transporter family protein	AT2G39510.1	TAIR	Yes
comp161964_c0	Maternal effect Embryo arrest 66 MEE66	AT2G02240.1	TAIR	No
comp202242_c0	Transmembrane amino acid transporter family protein	AT5G38820.1	TAIR	Yes
comp204231_c0	Reticulon family protein	AT3G10260.3	TAIR	Yes
Putative genes responding to M. spermotrophus infestation				
comp187226_c1	DEA(D/H)-box RNA helicase family protein	AT1G72730.1	TAIR	No
comp196871_c0	78 kDa glucose-regulated protein	Q90593	Uniprot	Yes
comp122638_c0	Snakin-2	Q93X17	Uniprot	Yes
comp203095_c0	Endoglucanase 15	Q6L4I2	Uniprot	Yes



Contig ID	Gene description	Identifier	Database	Hit to *taeda*
comp222951_c2	Unknown Protein	AT2G01300.1	TAIR	Yes
comp217047_c4	DNA-directed RNA polymerase I subunit rpa1	P15398	Uniprot	Yes
comp120198_c0	TCP-1/cpn60 chaperonin family protein	AT5G20890.1	TAIR	No
comp120398_c0	26S protease regulatory subunit 8	P62196	Uniprot	Yes
comp180046_c0	Mov34/MPN/PAD-1 family protein	AT5G23540.1	TAIR	No
comp180876_c0	aldehyde dehydrogenase 6B2	AT2G14170.1	TAIR	No
comp190764_c0	Em-like protein	Q40864	Uniprot	Yes
comp191416_c0	Em-like protein	Q40864	Uniprot	Yes
comp215422_c0	Unknown	AT2G30480.3	TAIR	Yes
comp203493_c0	nucleoside triphosphate hydrolase	AT3G18600.1	TAIR	No
comp198789_c0	FTSH protease 10	AT1G07510.1	TAIR	No
comp118890_c0	Coatomer subunit delta	Q5XJY5	Uniprot	Yes
comp201333_c0	ATP binding;leucine-tRNA ligases	AT1G09620.1	TAIR	No
comp225315_c0	39S ribosomal protein L17, mitochondrial	Q3T0L3	Uniprot	Yes
comp210813_c4	NADP-malic enzyme 3	AT5G25880.1	TAIR	No
comp215506_c0	ILITYHIA	AT1G64790.2	TAIR	No
comp196046_c0	staurosporin and temperature sensitive 3-like b	AT1G34130.1	TAIR	No
comp119037_c0	Prolyl oligopeptidase family protein	AT1G76140.1	TAIR	No
comp197872_c1	Alanyl-tRNA synthetase	AT1G50200.1	TAIR	No
comp209420_c0	Oxysterol-binding protein 2	Q5QNQ6	Uniprot	Yes
comp189610_c0	methylcrotonyl-CoA carboxylase	AT1G03090.2	TAIR	No
comp206079_c0	cellular apoptosis susceptibility protein (JA biosyn)	AT2G46520.1	TAIR	No
comp171074_c0	Phosphoglucomutase/phosphomannomutase	AT1G70730.3	TAIR	No
comp172126_c0	Em-like protein	Q40864	Uniprot	Yes
comp212559_c0	uclacyanin 1	AT2G32300.1	TAIR	No
comp219529_c1	S-adenosyl-L-homocysteine hydrolase	AT4G13940.1	TAIR	Yes
comp202525_c0	Unknown Protein	AT2G12400.1	TAIR	Yes
comp222114_c0	transposable element	P0C2F6	Uniprot	Yes
comp217821_c1	manganese tracking factor for mitochondrial SOD2	AT4G27940.1	TAIR	Yes
comp216255_c0	NAC domain containing protein 2	AT5G04410.1	TAIR	Yes
comp201087_c0	beta-1,6-N-acetylglucosaminyltransferase	AT5G11730.1	TAIR	Yes
comp195734_c0	Translation elongation factor EFG/EF2	AT1G62750.1	TAIR	Yes
comp216291_c0	Histone-lysine N-methyltransferase SUVR5	O64827	Uniprot	Yes
comp123143_c0	Pentatricopeptide repeat-containing protein	Q9LS25	Uniprot	Yes
comp186517_c1	Ethylene insensitive 3 family protein	AT3G20770.1	TAIR	Yes
comp212698_c0	LEA protein-related	AT5G60530.1	TAIR	No
comp186982_c0	Glutathione S-transferase family protein	AT3G62760.1	TAIR	Yes
comp220199_c0	Probable LRR serine/threonine-protein kinase	C0LGQ9	Uniprot	Yes
comp201711_c0	3-oxoacyl-[acyl-carrier-protein] synthase I	P23902	Uniprot	Yes
comp209348_c0	Protein argonaute 4A	Q9SDG8	Uniprot	Yes
comp198740_c0	O-acetylserine (thiol) lyase isoform C	AT3G59760.1	TAIR	Yes
comp200914_c0	HMG box protein	AT1G04880.1	TAIR	Yes
comp209383_c2	Histone H2A	P19177	Uniprot	Yes
comp203515_c0	x-ray induced transcript 1	AT5G48720.2	TAIR	Yes
comp202372_c0	Unknown	AT2G41350.2	TAIR	Yes
comp212944_c0	Heat shock protein 82	Q08277	Uniprot	Yes
comp199415_c0	17.8 kDa class I heat shock protein	P27396	Uniprot	Yes
comp212411_c0	Galactinol synthase 1	Q9XGN4	Uniprot	Yes
comp218412_c0	Glucose-methanol-choline oxidoreductase	AT1G72970.1	TAIR	Yes
comp215666_c2	Heat shock 70 kDa protein	P26413	Uniprot	Yes
comp207767_c0	transmembrane receptors;ATP binding	AT1G27170.1	TAIR	No
comp202130_c0	RING/U-box superfamily protein	AT2G04240.1	TAIR	No



Contig ID	Gene description	Identifier	Database	Hit to *P. taeda*
comp204988_c0	Amino acid permease family protein	AT3G19553.1	TAIR	Yes
comp218310_c0	Leucine-rich repeat transmembrane protein kinase	AT5G49660.1	TAIR	Yes
comp221114_c0	transcription regulators	AT5G13240.1	TAIR	Yes
comp210411_c1	Plant protein of Unknown	AT3G12060.1	TAIR	Yes
comp217359_c0	RNA-binding (RRM/RBD/RNP motifs) family protein	AT1G13190.1	TAIR	Yes
comp222178_c3	peroxidase 2	AT5G06720.1	TAIR	Yes
comp202508_c0	Transducin/WD40 repeat-like superfamily protein	AT5G03450.1	TAIR	Yes
comp196266_c0	Protein kinase superfamily protein	AT3G57120.1	TAIR	Yes
comp206379_c0	P-loop containing nucleoside triphosphate hydrolase	AT1G03905.1	TAIR	Yes
comp186043_c0	Ribosomal protein L12 protease adaptor protein	AT1G68660.1	TAIR	Yes
comp205412_c0	Protein of unknown function	AT3G50845.1	TAIR	Yes
comp105932_c0	Protein of unknown function	AT4G35980.1	TAIR	Yes
comp208697_c0	Thioredoxin superfamily protein	AT1G07080.1	TAIR	Yes
comp218300_c1	S-domain-2 5	AT4G32300.1	TAIR	Yes
comp219733_c0	Nitrilase/cyanide hydratase N-acyltransferase	AT5G12040.1	TAIR	Yes
comp210136_c2	alpha/beta-Hydrolases superfamily protein	AT5G18630.1	TAIR	Yes
comp203807_c0	tetraspanin10	AT1G63260.1	TAIR	Yes
comp171142_c0	Putative diflavin flavoprotein A 3	Q8YQD8	Uniprot	Yes
comp206474_c0	RING/U-box superfamily protein	AT5G20910.1	TAIR	Yes
comp215116_c1	Metal tolerance protein 5	Q5NA18	Uniprot	Yes
comp199279_c0	NHL domain-containing protein	AT5G14890.1	TAIR	Yes
comp205712_c0	Flavonoid 3',5'-hydroxylase	P37120	Uniprot	Yes
comp208556_c1	Chlorophyll a-b binding protein type 2 member 1A	P15193	Uniprot	Yes
comp194322_c0	Heavy metal transport/detox superfamily protein	AT1G71050.1	TAIR	Yes
comp222570_c2	NFX1-type zinc finger-containing protein 1	Q8R151	Uniprot	Yes
comp204788_c0	Lateral organ boundaries (LOB) domain family protein	AT5G63090.2	TAIR	Yes
comp205917_c0	NAD(P)-binding Rossmann-fold superfamily protein	AT5G58490.1	TAIR	Yes
comp192299_c0	hydroxyproline-rich glycoprotein family protein	AT5G65660.1	TAIR	Yes
comp220181_c0	methyl esterase 13	AT1G26360.1	TAIR	Yes
comp170837_c0	Proliferating cell nuclear antigen	O82134	Uniprot	Yes
comp177783_c0	CASP-like protein RCOM_1504680	B9RA90	Uniprot	No
comp217254_c0	PA-domain containing subtilase family protein	AT1G62340.1	TAIR	Yes