UVicSPACE: Research & Learning Repository

Faculty of Science

Faculty Publications

This is a pre-copyedited author-produced PDF of an article published in *Annals of Botany* following peer review. The article is:

Insights from the pollination drop proteome and the ovule transcriptome of *Cephalotaxus* at the time of pollination drop production

Cary Pirone-Davies, Natalie Prior, Patrick von Aderkas, Derek Smith, Darryl Hardie, William E. Friedman, and Sarah Matthews

May 2016

The version of record is available online at:

http://dx.doi.org/10.1093/aob/mcw026

Citation for this paper:

Pirone-Davies, C., Prior, N., von Aderkas, P., Smith, D., Hardie, D., Friedman, W.E. & Matthews, S. (2016). Insights from the pollination drop proteome and the ovule transcriptome of *Cephalotaxus* at the time of pollination drop production. *Annals of Botany*, 117(5), 973-984.

ORIGINAL ARTICLE

Insights from the pollination drop proteome and the ovule transcriptome of *Cephalotaxus* at the time of pollination drop production

Cary Pirone-Davies[1*], Natalie Prior[2], Patrick von Aderkas[2], Derek Smith[3], Darryl Hardie[3], William E. Friedman[1,5], and Sarah Mathews[4*]

1. The Arnold Arboretum of Harvard University, Boston, MA, USA
2. University of Victoria, Victoria, BC, Canada
3. UVic Genome BC Proteomics Centre, Victoria, BC, Canada
4. CSIRO, Centre for Australian National Biodiversity Research, Canberra, Australia
5. Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, MA, USA

Pollination drop proteins and ovule transcripts of *Cephalotaxus*

carypirone@fas.harvard.edu; sarah.mathews@csiro.au

ABSTRACT

- Background and Aims

Most gymnosperms produce an ovular secretion during reproduction, the pollination drop. Once thought to be a mere landing site for pollen, the secretion is now known to contain a suite of ions and compounds, including proteins. These components provide nutrition for the developing pollen and play diverse roles during pollination. Proteins have been identified in the drops of species of *Chamaecyparis*, *Juniperus*, *Taxus*, *Pseudotsuga*, *Ephedra*, and *Welwitschia*, and are likely involved in the conversion of sugars, defense against pathogens, and pollen growth and development.

- Methods

Here, we use mass spectrometry to identify proteins in the pollination drops of *Cephalotaxus sinensis and C. koreana*. We also document the transcripts present in the ovules of *C. sinensis* at the time of pollination drop production

- Key Results

Several identified proteins in *Cephalotaxus* have been reported previously in the drops of other gymnosperms and likely function in defense, polysaccharide metabolism, and pollen tube growth and guidance. Other proteins appear to be unique to *Cephalotaxus*, and their putative functions include starch and callose degradation, among others. Identified transcripts spanned a range of functional categories and some may be involved in drop formation, ovule development, and conspecific pollen recognition.

- Conclusions

The pollination drop proteome expands our understanding of the function of the pollination drop during *Cephalotaxus* reproduction, and provides insight into how pollination drop proteins vary among species of *Cephalotaxus*. The transcriptome data provide a framework for understanding multiple metabolic processes that occur within the ovule and the pollination drop just before fertilization. This is the first published transcriptome of any gymnosperm ovule.

keywords: *Cephalotaxus*, proteome, transcriptome, development, gymnosperm, pollen selection

INTRODUCTION

In most extant gymnosperms, pollination relies on the wind-mediated transfer of pollen to ovulate cones. In many taxa, an ovular secretion, the pollination drop, extends beyond the micropyle and forms a liquid surface that serves as a landing site for pollen. Subsequently, the drop withdraws and transports pollen grains into the ovule where they germinate, form pollen tubes, and ultimately release sperm that fertilize eggs (Singh, 1978, Doyle and O' Leary, 1935). The sole function of the pollination drop was once thought to be transport of pollen into the ovule. We now know, however, that these ovular secretions contain a suite of organic and inorganic compounds including sugars, amino acids, organic acids, calcium, and proteins (Ziegler, 1959, Seridi-Benkaddour and Chesnoy, 1988, Carafa et al., 1992). The proteins that have been discovered in the drops suggest that in addition to nourishing the developing pollen, they play diverse roles in the pollination process (Gelbart and von Aderkas, 2002, Nepi et al., 2009, Coulter et al., 2012).

Proteomic studies of pollination drops have been performed in *Pseudotsuga menziesii* (Poulis et al., 2005), *Taxus* × *media* (O'Leary et al., 2007), *Juniperus communis*, *Juniperus oxycedrus*, and *Chamaecyparis lawsoniana*, one gnetophyte, *Welwitschia mirabilis* (Carafa et al., 1992, Wagner et al., 2007), and several species of *Ephedra* (von Aderkas et al., 2015). Thaumatin-like proteins, chitinases, invertase, galactosidase, peroxide, and subtilisin-like protease, among others, have been detected (Table 1). Arabinogalactan proteins were documented in *Taxus* x *media* using immunological methods (O'Leary et al., 2004), and acid phosphatase was identified in *Welwitschia mirabilis* (Carafa et al., 1992). Putative functions of the proteins include the conversion of sugars, cleavage of polysaccharides, defense against pathogens, and a variety of roles associated with the expansion and growth of pollen tubes (Table 1) (Gelbart and von Aderkas, 2002, Poulis et al., 2005, Wagner et al., 2007). Chitinases in *Pseudotsuga menziesii* have antifungal activity (Coulter et al., 2012). Some of these proteins are conserved among taxa, such as chitinase and glucosidase, while others, such as galactosidase, have been observed in only a single species (to date).

Developmental studies suggest that pollen selection conceivably could occur within pollination drops. In the ovules of several species of Pinaceae, differences in pollen germination and pollen tube growth and development were observed, depending on whether the pollen was conspecific, heterospecific (McWilliam, 1959, Fernando et al., 2005), or heterogeneric (von Aderkas et al., 2012) with respect to the ovule. The underlying mechanisms controlling these differential responses of pollen and pollen tubes are unknown, however. It is possible that nutritional requirements differ for each type of male gametophyte, leading some to thrive in certain pollination drops while others languish or die (Gelbart and von Aderkas, 2002). Alternatively, a protein-protein interaction may occur between the pollen and the ovule, beginning with pollen recognition and culminating in the destruction or inhibition of the growth of certain pollen types. Protein-protein interactions between pollen and/or pollen tubes and ovules, have been well documented in angiosperms, including those of the incompatibility reactions of S-locus proteins. Self-incompatibility (SI) systems are found in at least 100 angiosperm families (Igic et al., 2008), and comprise diverse molecular mechanisms, the most well characterized of which are the sporophytic self-incompatibilty system (SSI) found in members of the Brassicaceae, and the gametopythic Si systems (GSI) found in the Papaveraceae, Solanaceae, Plantaginaceae, and Rosaceae [(for reviews, see (Iwano and Takayama, 2012, Takayama and Isogai, 2005, Franklin-Tong, 2008)].

Proteomic studies are needed in additional gymnosperms to better understand pollination drop composition and function throughout conifers, cycads, *Ginkgo*, and gnetophytes. Here, we documented the proteins present in the pollination drops of *Cephalotaxus*, for which no data were previously available. *Cephalotaxus*, or "plum yew" (Figure 1), is native to southern and eastern Asia and includes 8-11 species (Bassett et al., 2005). It is the sole genus in Cephalotaxaceae, although some taxonomic treatments do not recognize Cephalotaxaceae, but consider *Cephalotaxus* to be a divergent genus within the Taxaceae (Rai et al., 2008). Regardless, the divergence of *Cephalotaxus* from Taxaceae sensu stricto is ancient, having occurred about 150 Ma (Leslie et al., 2012). We focused on *Cephalotaxus sinensis* Rehder and Wilson and *C. koreana* Nakai. We also generated RNA-Seq data to document the transcripts present in ovules of *C. sinensis* at the time of pollination drop production and to provide additional insight on ovule metabolic processes and pollination drop functions. This is the first transcriptome assembled from a *Cephalotaxus* ovule at the time of pollination drop production, and the first transcriptome of any gymnosperm ovule.

MATERIALS AND METHODS

Pollination Drop and Ovule Collection

Pollination drops and ovules were collected from species of *Cephalotaxus* at the Arnold Arboretum of Harvard University (Boston, MA). Pollination drops were collected from each of *C. koreana* and *C. sinensis* using a flame-drawn capillary tube in April 2011. Drops were pooled to obtain a minimum of 100 µl per species and were stored at -20°C until analysis. These are referred to as "outdoor" samples in the remainder of the paper. Additionally, branches of approximately 20 cm in length were cut from *C. koreana* plants, transferred to cups of water, and maintained in greenhouses of the Arnold Arboretum. Temperatures were kept at 12-13°C during the day and 4°C at night to mimic outdoor temperatures. At the time of pollination drop production, drops were collected and pooled as above. These are referred to as “indoor” samples. In April 2012, cones were collected from *C. sinensis* at the time of pollination drop production and were frozen immediately in liquid nitrogen.

Protein Preparation, Electrophoresis, Mass spectrometry, and Data Analysis

Sample preparation, electrophoresis, and protein sequencing followed previous methods (Prior et al., 2013). In brief, 1-D SDS-PAGE was performed, and gels were stained with Coomassie blue. Bands were excised from gels, digested with trypsin, and analyzed at the University of Victoria Genome BC Proteomics Center via HPLC-MS/MS. HPLC was performed on a RP nano-analytical column Magic C-18AQ (Michrom BioResources Inc, Auburn, CA). The chromatography system was coupled on-line with an LTQ Orbitrap Velos mass spectrometer equipped with a Nanospray II source (Thermo Fisher Scientific).

The raw data files were searched using Thermo Scientific Proteome Discoverer software version 1.2 (Thermo Fisher Scientific) with Mascot version 2.2.1 (Matrix Science, Boston, Massachusetts, USA) and PEAKS Studio v. 6 (Bioinformatics Solutions Inc., Ontario, Canada) against the UniProt-SwissProt and Uniprot-Trembl databases. Several post-translational modifications were tested for during searches. These included the fixed carbamidomethylation of cysteines when iodoacetamide was used in sample processing, oxidation of methionine, and deamidation (N, Q). Due to the paucity of gymnosperm data available in these databases, de novo discovery of peptides was also performed using PEAKS Studio. SPIDER homology searches implemented in PEAKS Studio were also performed to compensate for sequencing errors and mutations which may otherwise inhibit the correct identification of peptides (Han et al., 2005). PEAKS “In Chorus” was used to summarize the various search results. In both

MASCOT and PEAKS, peptides were accepted as correctly identified if their scores had values of at least $p < 0.01$. The false discovery rate, as determined by a decoy database search, was kept below 1%. Proteins were considered as correctly identified if they contained at least one unique peptide that fulfilled the above criteria. We did not include “uncharacterized protein” hits in this report.

RNA-Seq

Ovules, including the nucellus and integuments, were dissected from cones on ice and pooled into a single sample. RNA was extracted using a modified version of a previous protocol (Chang et al., 1993). Library preparation and sequencing was performed by the FAS Center for Systems Biology (Harvard University). In brief, RNA was amplified using the PrepX SPIA RNA-Seq Library Kit (Nugen), sheared, and made into 200 bp insert libraries using the PrepX ILM DNA Library Preparation kit (IntengenX, Pleasanton, CA). Samples were sequenced on an Illumina HiSeq 2000.

Quality of sequences was assessed using FastQC (Andrews). Adapters were removed using CutAdapt (Martin, 2011), and low quality sequences were trimmed using Sickle (Joshi and Fass, 2011) and Trimmomatic (Lohse et al., 2012). Trimmed sequences were assembled using Trinity (Grabherr et al., 2011) with default parameters. Similarity searches were performed using BLASTx against the NCBI non-redundant database (nr) and Uniprot (SwissProt and TReMBL) databases. Annotation was performed using blast2go (Conesa et al., 2005) against the Gene Ontology (GO) database.

RESULTS

Proteome

Proteins ranged in MW from approximately 7 to 50 kDa. Six bands from *C. koreana* indoor and outdoor samples, and eight bands from *C. sinensis* were sequenced (Figure 2). A total of 30 proteins from *C. sinensis* (outdoor), 32 from *C. koreana* (outdoor), and 18 from *C. koreana* (indoor) were identified (Table 2). Peptides can be found in Supplementary Table 1 **[Supplementary Information]**. Six of these proteins were detected in all three samples: chitinase IV, peroxidase, thaumatin-like protein (TLP), pollen allergen CJP-38, alpha-galactosidase, and luminal binding protein (BiP). Acidic endochitinase and neutral ceramidase were shared by both *C. koreana* indoor and outdoor samples, while an additional seven proteins were shared between *C. sinensis* outdoor and *C. koreana* outdoor (Table 2).

Transcriptome

Paired end sequencing yielded 314,781,368 paired-end reads from a 200 bp insert library, encompassing 21.8 GB of data. After stringent quality filtering, only reads of QV higher than 20 were accepted (mean QV=37). Sequences have been deposited in the NCBI SRA database under the accession numbers SRP058054.

Using the Trinity de novo assembly program, 402,215 transcripts were assembled with N50 = 390, with 56,370 transcripts above 500 bp and 17,201 transcripts above 1000 bp (Figure 3). To identify putative homologues in other species, transcripts were searched against the NCBI non-redundant database (nr) and the combined Swissprot (SP) and TREMBL (TR) Uniprot databases using BLASTx and an e-value cutoff of 10^{-5}. 49,277 transcripts had significant matches against the nr database, 27,837 against SP, and 47,558 against TR [**Supplementary Information]**. In total, 49,769 transcripts were matched to putative homologues using this approach.

BLAST searches depend on several factors that impose limits on interpretations. Short reads are rarely matched to known genes. For example, only 7.8% of all transcripts less than 500 bp in length had blast hits against the nr database, but 36.0% of transcripts sized 500 bp or greater had hits. Also, the paucity of gymnosperm data available in the databases limits the overall number of BLAST hits.

Transcripts with BLAST matches against nr were annotated using Gene ontology (GO) terms to predict transcript function. 9,644 transcripts were annotated [**Supplementary Information]**. Gene ontology (GO) annotations classify the function of transcripts into three categories: biological processes (bp), cellular components (cc) and molecular functions (mf). Within these categories, the greatest numbers of transcripts were assigned to the sub-categories “binding” and “catalytic” (molecular processes), “cellular process” and “metabolic process” (biological processes) and “cell” and “cell part” (cellular component) (Figure 4).

DISCUSSION

Proteome of the pollination drops of C. koreana and C. sinensis

We found a diversity of proteins in the pollination drops of *C. koreana and C. sinensis*, with potential roles in defending nutrient-rich pollination drops from pathogens, promoting and supporting pollen tube growth, metabolism during drop production, and response to stress. Some of these proteins likely are actively secreted into the pollination drop (O'Leary et al., 2007), while others may enter the drop as a result of nucellar breakdown (von Aderkas et al., 2015). As pollination drops form in *Cephalotaxus* and in some other gymnosperms, the most micropylar cells of the nucellus break down (C.Pirone-Davies, unpubl.res.)(O'Leary et al., 2004). Thus, we expect that some of the proteins detected here are by-products of this degradation.

Several of the proteins we detected in *Cephalotaxus* are found in the pollination drops of other gymnosperms, while others are reported here for the first time (Table 2). Defensive proteins are conserved across all species examined to date. Chitinases have been detected previously in two species of *Juniperus*, *Ephedra foeminea*, *E. trifurca*, and *Welwitschia mirabilis* (Wagner et al., 2007), TLPs have been detected in species of *Juniperus, E. minuta*, and *Taxus x* *media* (O'Leary et al., 2007), peroxidase in *Pseudotsuga menziesii, E. compacta, E. likiangensis, and E. trifurca,* and galactosidase in *E. minuta, E. compacta, E. trifurca,* and *P. menziesii* (Poulis et al., 2005). Three of the six proteins shared among all three samples are pathogenesis-related (PR) proteins, i.e., they are induced in the presence of a pathogen. PR-proteins are classified into diverse families, including chitinases (PR-3), thaumatin-like proteins (TLPs) (PR-5), and peroxidases (PR-9) (van Loon et al., 2006). In addition to defensive functions, these proteins are involved in various physiological processes.

Peroxidases are involved in auxin metabolism, lignin and suberin formation, and the linking of cell wall components (Passardi et al., 2005), while TLPs accumulate in tissue in response to some environmental stresses and developmental cues (Liu et al., 2010). Some TLPs can inhibit the formation of ice crystals (Hon et al., 1995). Chitinases break down various polymers, particularly chitin, and are involved in growth and developmental processes (Grover, 2012). Chitin is the primary component of fungal cell walls, and thus chitinases often function as antifungals, as has been shown in the pollination drops of *Pseudotsuga menziesii* (Coulter et al., 2012). In *Cephalotaxus* and other conifers, and *Welwitschia*, it is likely that one function of some or all of these proteins is to defend the nutrient rich drop from pathogens (Gelbart and von Aderkas, 2002).

Galactosidase enzymes are present in various plant tissues where they metabolize a variety of polysaccharides and are involved in fruit ripening, growth, and the hydrolysis of lactose. In flowering plants, galactosidases are observed in both the stigma exudate and in the pollen, and are hypothesized to loosen the cell wall components of the intine and assist in pollen germination and elongation (Hruba et al., 2005, Rejon et al., 2013). It has been suggested that galactosidase and xylosidase present in the pollination drops of *Pseudotsuga menziesii* may also help loosen the cell walls of pollen via degradation of xyloglucan support chains of the pollen intine, thus promoting pollen tube growth (Poulis et al., 2005).

Pollen allergen CJP-38 is homologous to β-1,3-glucanase, which degrades β-1,3-glucan, or callose. Functions of β-1,3-glucanase are diverse, and include the regulation of plasmodesmata (Levy et al., 2007), hydrolysis of β-1,3-glucan in fungal cell walls and pollen tube growth in angiosperms (Kotake et al., 2000, Sela-Buurlage et al., 1993). In the pollination drops of *Cephalotaxus*, CJP-38 may have anti-fungal activity, or, if it originates from nucellar cells, it may regulate plasmodesmata, and thus cell-cell communication within the ovule. It is unclear whether CJP-38 may be involved in pollen tube growth in *Cephalotaxus* as it is in angiosperms. Callose is present in the pollen tubes of some gymnosperms (Yatomi et al., 2002), but it is not ubiquitous, and its distribution within the tube varies depending on species and developmental stage [for review, see (Fernando et al., 2010)]. In contrast, the angiosperm tube wall is composed predominantly of callose, and callose septae, or plugs, are formed throughout the tube as it grows (Abercrombie et al., 2011).

The luminal binding protein (BiP) is located in the endoplasmic reticulum (ER) lumen, where it assists in the proper folding of proteins (Boston et al., 1996, Galili et al., 1998). Based on its subcellular location, BiP likely is present in the drop as a result of nucellar breakdown. It is conserved across all three samples, suggesting that it may have an important metabolic role at the time of pollination drop production. BiP activity increases during biotic and abiotic stress responses to pathogens, nutrient deficiency, temperature changes, and water stress (Alvim et al., 2001). It has been proposed that increased BiP production is needed to support an increase in the synthesis of PR proteins (Jelitto-Van Dooren et al., 1999). Numerous PR proteins are present in the pollination drop, thus it is possible that BiP could support their synthesis.

Other proteins were present in both the outdoor samples of *C. koreana* and *C. sinensis*, but not the indoor *C. sinensis* sample (Table 2). Cupa3, Cup s, and Jun r are pollen allergens known from *Cupressus arizonica*, *Cupressus sempervirens*, and *Juniperus rigida*, respectively (Cortegano et al., 2004, Breiteneder and Mills, 2005), and are members of the PR-5 protein family. These proteins may play a role in defense. Alpha amylase hydrolyzes starch, thus mobilizing energy for growth and development (Huang et al., 1992). Starch has been observed to accumulate in the nucellar cells of some gymnosperms just before drop formation (Carafa et al., 1992, Takaso and Owens, 1995), and in some species, it decreases at the time of pollination drop production (Owens and Simmons, 1987).

In both the indoor and outdoor samples of *C. koreana*, we detected acidic endochitinase, another defensive protein. Neutral ceramidase also was detected in these samples. Little is known about the function of ceramidase in plants, but it may be involved in cell signaling and development (Pata et al., 2010). Other proteins were detected in only one of the samples examined (Table 2), although technical replicates are needed to verify that they were not present in other samples (Elias et al., 2005, Ham et al., 2008). Many of these other proteins are involved in cellular metabolic processes, and may have been released from nucellar cells. Several of these proteins were recently detected in pollination drops of seven species of *Ephedra* (von Aderkas et al., 2015).

The transcriptome of C. sinensis at the time of pollination drop production

In our RNA-Seq data, we also detected transcripts with potential roles in pollination drop and ovule processes including pollen recognition, pollination drop formation, and ovule development. There was a degree of correlation between the identity of transcripts and the proteins detected by MS-MS [**Supplementary Information]**, but we do not expect a one to one correlation between transcripts and proteins for several reasons. The pollination drop proteome represents portions of both the pollen and ovule proteomes, and includes just a fraction of the total ovule proteome. Also, the synthesis and degradation of mRNA and protein are differently affected by various factors including numerous post-transcription regulatory processes that affect mRNA stability, the timing of protein synthesis, and rates of protein turnover [for review, see (Vogel and Marcotte, 2012)].

Pollen Recognition

We detected several transcripts that may be involved in pollen recognition, based on blast results and GO annotations (Supplementary Table 5). Surprisingly, these included two S-locus lectin protein kinases, one g-type lectin S-receptor-like serine threonine-protein kinase, and two predicted proteins with a GO annotation of “”. S-locus proteins determine the specificity of pollen rejection in angiosperm self-incompatibility systems (SI) (Takayama and Isogai, 2005), and the S-receptor kinase gene (SRK) is the female determinant of the SI system in *Brassica* (Stein et al., 1991). SRK belongs to the diverse receptor-like kinase protein family (RLK) (Shiu and Bleeker, 2001), and is the best characterized member of the subfamily S-domain RLK (SRLK) (Shiu and Bleecker, 2003, Xing et al., 2013). This subfamily is differentiated based on the presence of an extracellular S-domain composed of 3 sub-domains, B_Lectin, SLG, and PAN_APPLE, one of which, SLG, is responsible for binding the male determinant during SI reactions (Kemp and Doughty, 2007). Apart from SRK, the functions of SRLK members are unknown, but based on their presence in non-reproductive tissues and self-fertilizing species, are predicted to be involved in roles other than pollen recognition, including development and defense (Bassett et al., 2005, Dwyer et al., 1994). Further studies are needed to determine the function of these four transcripts in *C. sinensis*. However, the possibility that they may be involved in self-incompatibility reactions or other forms of pollen recognition is intriguing.

Self-incompatibility is widespread in angiosperms, and some authors have suggested that SI also occurs in gymnosperms (Kormatuk, 1999, Runions and Owens, 1998, Owens et al., 1990). It is well established that self-pollinated conifers have lower rates of seed set than those that are cross-pollinated. This trend is generally accepted to be the result of embryo inviability due to lethal recessive genes or competition between outcrossed and selfed embryos in seeds with multiple fertilization products (simple polyembryony). However, in species of *Larix* (Kosinski, 1986), *Abies* (Kormatuk, 1999), *Picea* (Runions and Owens, 1998), and *Thuja* (Owens et al., 1990), low seed set in selfed individuals was sometimes correlated not with embryo abortion, but with the failure of pollen tube growth in the nucellus or archegonium, suggesting the occurrence of SI. However, in all studies, pollen tube growth varied greatly both between ovules of the same individual and also between individuals, and some selfed ovules progressed as normal to result in viable seed. Furthermore, observations were made on a small number of ovules and in some cases it was unclear how frequently aborted growth occurred. Strong evidence does support the occurrence of interspecific pollen selection in the pollination drop. In several cross-pollination studies, conspecific pollen germinated and grew better in the pollination drop than heterospecific or heterogeneric pollen (McWilliam, 1959, von Aderkas et al., 2012, Fernando et al., 2005). However, these studies do not distinguish whether this was because nutritional needs were better met by conspecific pollination drop, or whether signaling among pollen and ovule might be occurring, or both.

Interestingly, observations during various other stages of reproduction suggest that signaling can occur between the pollen and ovule. These include the dependence of ovule development on the germination and growth of the pollen tube in species of *Pinus* and *Tsuga* (Owens et al., 2005, Dogra, 1967, Owens and Blake, 1983), and the retraction of the pollination drop in the presence of conspecific, but not heterospecific pollen in *Juniperus communis*

(Mugnaini et al., 2007). Although *Cephalotaxus* is dioecious, the presence of S-locus transcripts suggests a possible molecular basis for the recognition of pollen. SI systems help promote outcrossing and are often cited as a unique feature of angiosperms that contributes to reproductive isolation and speciation (Stebbins, 1957, Jain, 1976). The presence of similar systems in gymnosperms would require a revision of this hypothesis.

Pollination Drop Formation

We also detected candidate transcripts that may help explain the unknown mechanisms of pollination drop formation and retraction. The pollination drop originates in the nucellus, as shown by immunolocalization studies of pollination drop proteins (Poulis et al., 2005), and then passes through the micropyle to form a droplet. Ultimately, it retracts, bringing pollen into the ovule. In *Cephalotaxus*, drop production and retraction follow a diurnal cycle, with drops produced in the early morning and slowly retracting throughout the day. This has also been observed in other gymnosperms (Tomlinson et al., 1991). Drop retraction is also induced when pollen enters the drop (C Pirone-Davies, The Arnold Arboretum of Harvard University, pers. comm.), as has been noted in other species (Doyle and O' Leary, 1935, Tomlinson et al., 1997, Mugnaini et al., 2007). How the movement of water is regulated is not well understood. Conifer ovules are not vascularized, and thus water associated with pollination drop formation and retraction is not associated with the osmotic potential of the xylem. Pollination drop formation must therefore be controlled by water dynamics within the ovule or cone (O`Leary and von Aderkas, 2005). One explanation is that changes in the osmotic potential of the drop facilitate the movement of water (Coulter et al., 2012, Ziegler, 1959). Sugar is a likely osmotic regulator, as various sugars are present in the pollination drop (Seridi-Benkaddour and Chesnoy, 1988, Nepi et al., 2009), and sugars are proposed to regulate water movement during nectar secretion. We detected several transcripts involved in the transport of sugars [**Supplementary Information]**, including several members of the SWEETS protein family, a superfamily of sugar transporters (Xuan et al., 2013), some of which are involved in nectar secretion (Lin et al., 2014). We also detected transcripts of β-glucosidase and P-loop containing nucleoside triphosphate hydrolases, proteins that have also been shown to be involved in nectar production (Bender et al., 2012). The pollination drop, like nectar, is a liquid secretion from reproductive tissue, and it is possible that the mechanisms controlling secretion in these systems may be similar.

Ovule Development

Finally, we detected several transcripts that are similar to angiosperm genes involved in the development of reproductive organs, and that may be useful for exploring the genetic basis of ovule development. These include *ULTRAPETALA (ULT1*), MADS-box transcription factors, *AP2*-related, *CLAVATA (CLV)*, *WUSCHEL (WUS*), *JOINTLESS*, and *LEAFY (LFY*) (Supplementary Table 5). The roles of several of these loci are understood in angiosperms, but fewer data are available for gymnosperms (for review, see (Mathews and Kramer, 2012). For example, the transcription factor WUS, a member of the WUSCHEL-related homeobox domain (*WOX*) family, interacts with the ligand *CLAVATA3 (CLV3*) and other members of the *CLV* signaling pathway to maintain the floral meristem and the central zone activity of the shoot apical meristem (SAM) (Laux et al., 1996). The expression of WUS in *Arabidopsis* ovules, where it is critical for initiation of the integuments, is consistent with the origin of the nucellus of the ovule from a shoot apical meristem (Gross-Hardt et al., 2002, Mathews and Kramer, 2012), however, consistent with their determinate nature, *CLV3* is not expressed in *Arabidopsis* ovules. It is thus interesting that we detected *CLV* transcripts in the dissected ovules of *C. sinensis*. This could point to the presence of stem cell niche, maintained by the signaling of *WUS* and *CLV*, and the presence of *WUS* in the ovules of *Gnetum* (Nardmann et al., 2009) and *C. sinensis* suggests that the role of WUS during integument formation may be conserved across the seed plants.

MADS-box genes are transcription factors that contain a MADS DNA binding domain, which is conserved across eukaryotes and metazoans (Gramzow et al., 2014). MADS-domain proteins are subdivided into Type 1, or serum response factor (SRF)-like proteins, and Type 2, or myocyte enhancer factor (MEF)-like proteins. Few Type-1 genes have been functionally characterized, but Type-1 genes in *Arabidopsis* are involved in female gametophyte, embryo sac, and seed development [for review, see (Gramzow and Theissen, 2010)]. Type-1 genes were detected in several members of the Pinaceae as well as *Sciadopitys verticillata*. However, detection of transcripts is infrequent and is limited to shoot, bud, male cone, and embryo tissues (Gramzow et al., 2014). The functions of these transcripts are unknown. Our dataset contains a single Type-1 MADS box transcript.

Type 2 MADS box genes, in contrast, have been extensively studied in angiosperms, where they are best known for their roles in floral organ identity. They are involved in diverse developmental processes in fruits, seeds, embryos, roots, and leaves [for reviews, see (Becker, 2003, Theissen, 2001)]. Determination of organ identity in flowers is described by the ABCDE model, with A class genes specifying sepals, A+B+E class petals, B+C+E class stamens, C class carpels, and D class ovules [for review, see (Causier et al., 2010)], and MADS-Box genes comprise most of these classes. In gymnosperms, B and C class genes appear to be involved in the specification of reproductive structures and the differentiation of male and female cones (Melzer et al., 2010). Using BLAST searches, we detected several transcripts that showed greatest similarity to Type 2 MADS-Box genes, including B-class and B-sister (Bs) transcripts, AGAMOUS (AG) and AG-like (C class), TM8, and DEFICIENS AGAMOUS-LIKE 10 (DAL10). We verified via phylogenetic analyses (S Mathews, E Kramer, C Pirone-Davies, CSIRO, Canberra, Australia, Harvard University, Cambridge, USA, The Arnold Arboretum of

Harvard University, Boston, USA, unpeel. res.) that among the BLAST hits are single homologs of B-class, Bs, and AG transcripts.

B-class genes have been detected in some conifers, *Ginkgo* and *Gnetum gnemon* (Mouradov et al., 1999, Sundstrom et al., 1999, Gramzow et al., 2014). Expression of B-class genes is largely restricted to male cones, and it appears that the role of these genes in the development of pollen bearing structures is conserved across seed plants (Sundstrom and Engstrom, 2002, Theissen and Becker, 2004). However, a recent study also detected B-class transcripts in mixed shoot tissue as well as in the female cones of *Picea abies* (Gramzow et al., 2014). Thus, the authors suggest a broad role for B-class genes in gymnosperms, but also concede that the detection of these transcripts could be due to transcriptional noise. Our detection of B-class and Bs transcripts in the ovules of *C. sinensis*, suggests that B-class genes are indeed involved in a role beyond male reproductive development. Since we collected ovules from outdoor sites where pollen was likely present, we cannot rule out completely the possibility that we detected transcripts from the pollen. However, it is unclear whether B-class transcripts are expected in mature pollen. *In situ* data indicate their presence in various pollen cone tissues, but no transcripts were detected in the pollen mother cells, and mature pollen was not analyzed (Sundstrom and Engstrom, 2002). The expression of one or more B gene homologs in ovules is not uncommon in angiosperms [for review, see (Kramer and Irish, 2000)], so it may be that this pattern is similarly common among gymnosperms.

Homologs of the C class gene *AG* have been identified in all major gymnosperm lineages (Tandre et al., 1998, Rutledge et al., 1998, Kramer et al., 2003, Winter et al., 1999), and expression of *AG* and *AG*-like has been observed in the female cones of four cycads and several conifers, the male cones of *Cryptomeria*, the shoots of *Gnetum gnemon*, non-reproductive tissues of *Picea abies*, and in the nucellus of *Taxus globoso* (Englund et al., 2011, Gramzow et al., 2014), a close relative of *C. sinensis*. Thus, the presence of these homologues in *C. sinensis* is not surprising. In flowering plants, *AG* is involved in stamen and carpel identity and in establishing determinancy of the floral meristem (Bowman et al., 1989). In gymnosperms, *AG* genes are generally involved in the development of both male and female cones (Melzer et al., 2010), and may be involved in the development of ovuliferous scales (Tandre et al., 1998) and in the transition from vegetative to reproductive identity (Carlsbecker et al., 2013).

We also detected several *TM8* transcripts as well as *DAL10*. *TM8* expression occurs in reproductive and non-reproductive tissues of diverse gymnosperms (Gramzow et al., 2014), including the ovules of *Taxus baccata* and *G. biloba* (Lovisetto et al., 2012). Although the function of *TM8* remains poorly understood, in tomato, it may be involved in controlling A class expression (Daminato et al., 2014). Given the diversity of *TM8* expression patterns across seed plants, further research into its function is needed. *DAL10* has also been found in numerous gymnosperm species and tissues, but is absent from angiosperm lineages (Carlsbecker et al., 2003). Similar to *TM8*, little is known about the function of this gene, although it appears to be involved in the shift from vegetative to reproductive buds, and its presence in developing seed and pollen cones is similar to B and C class genes (Carlsbecker et al., 2003, Carlsbecker et al., 2013). Both TM8 and DAL10 are present in a large clade sister to the C-class genes (Melzer et al., 2010).

MADS Box genes from the A class have not been detected in any gymnosperm (Melzer et al., 2010). However, the A class gene *AP2* from the AP2/ERF family of transcription factors has been detected in several conifers [e.g. (Nilsson et al., 2007)]. In *Arabidopsis*, *AP2* specifies the identity of sepals and petals, regulates C class genes, is involved in seed development, and may play a role in development of non-floral organs (Jofuku et al., 1994). Homologues of *AP2* from *Picea* likely control diverse developmental events, but also share features of their angiosperm counterparts, as Pi*AP2* has the capacity to substitute for an A class gene in *Arabidopsis* (Nilsson et al., 2007). We detected one *AP2*-related transcript.

We also detected transcripts similar to *LFY* and *ULT1*. In angiosperms *LFY* enables the transition from vegetative to floral meristems (Weigel et al., 1992), and also activates floral homeotic genes (Weigel and Meyerowitz, 1994). Members of the *LFY* lineage have been detected in non-reproductive and reproductive tissues of diverse non-flowering plants (Moyroud et al., 2010), including the nucellus of *Picea* (Carlsbecker et al., 2013). *LFY* acts in early female and male cone development (Vasquez-Lobo et al., 2007), and may regulate some ABC-type MADS-box proteins as it does in angiosperms (Moyroud et al., 2010). In flowering plants, *ULT1* controls shoot and floral meristem cell accumulation (Fletcher, 2001), and also positively regulates floral meristem determinacy, possibly through the *AG* pathway (Prunet et al., 2008). To our knowledge, *ULT1* has not been detected in gymnosperms.

Finally, we detected *JOINTLESS*, a MADS-Box gene involved in the formation of the abscission zone in tomato (Mao et al., 2000). JOINTLESS works at least in part via the regulation of phytohormones, transcription factors involved in meristem identity, and genes involved in cell wall formation and lipid metabolism (Nakano et al., 2012). It will be interesting to explore the role of JOINTLESS, as well as other transcripts discussed here, in *Cephalotaxus*.

Our data support previous studies indicating that PD proteins are important in defense, polysaccharide metabolism, and pollen tube growth and guidance. We also detected several novel proteins that are likely involved in defense (Cup a3, Cup s, Jun r), starch degradation (alpha amylase), and callose degradation (pollen allergen CJP-38), while others are likely by-products of nucellar degradation (luminal binding protein, ceramidase, and others). The implications of the latter in PD function, if any, are unclear. Examination of the transcriptome revealed several transcripts that may be important in pollen-ovule recognition, pollination drop formation and retraction, and reproductive and developmental processes within the ovule. Together, these proteomic and transcriptomic data provide a foundation to better understand the complexities of gymnosperm reproduction and ultimately, to gain insight into seed plant evolution.

FUNDING

This work was supported by a Putnam Fellowship from the Arnold Arboretum of Harvard University to CPD; The Natural Sciences and Engineering Research Council of Canada's PGS and Discovery Grant Programs provided financial support to NAP and PvA, respectively.

INFORMATION

Supplementary Table 1 lists the proteins and peptides identified in all samples. Supplementary Tables 2 and 3 report BLAST and blast2go results, respectively, of all transcripts identified. Supplementary Table 4 states which proteins and transcripts were found in both the proteome and the transcriptome, and Supplementary Table 5 highlights transcripts that may be involved in pollen recognition, development, and pollination drop formation.

ACKNOWLEDGEMENTS

We thank Elena Kramer for her insight on ovule development and for sharing with us her data matrix of MADS-Box genes.

LITERATURE CITED

2008. *Self-incompatibility in flowering plants - evolution, diversity, and mechanisms*: Springer-Verlag Berlin Heidelberg.

Abercrombie JM, O'Meara BC, Moffatt AR, Williams JH. 2011. Developmental evolution of flowering plant pollen tube cell walls: callose synthase (CalS) gene expression patterns. *Evodevo,* **2**: 14.

Alvim FC, Carolino SM, Cascardo JC, Nunes CC, Martinez CA, Otoni WC, Fontes EP. 2001. Enhanced accumulation of BiP in transgenic plants confers tolerance to water stress. *Plant Physiology,* **126**: 1042-54.

Andrews S. FastQC A Quality Control tool for High Throughput Sequence Data.

Bassett CL, Nickerson ML, Farrell RE, Artlip TS, Ghaouth AE, Wilson CL, Wisniewski ME. 2005. Characterization of an S-locus protein kinase-like gene from peach. *Tree Physiology,* **25**: 403-411.

Becker A. 2003. The major clades of MADS-box genes and their role in the development and evolution of flowering plants. *Molecular Phylogenetics and Evolution,* **29**: 464-489.

Bender R, Klinkenberg P, Jiang Z, Bauer B, Karypis G, Nguyen N, Perera MADN, Nikolau BJ, Carter CJ. 2012. Functional genomics of nectar production in the Brassicaceae. *Flora - Morphology, Distribution, Functional Ecology of Plants,* **207**: 491-496.

Boston RS, Viitanen PV, Vierling E. 1996. Molecular chaperones and protein folding in plants. *Plant Molecular Biology,* **32**: 191-222.

Bowman JL, Smyth DR, Meyerowitz EM. 1989. Genes Directing Flower Development in Arabidopsis. *The Plant Cell,* **1**: 37-52.

Breiteneder H, Mills EN. 2005. Molecular properties of food allergens. *Journal of Allergy and Clinical Immunology,* **115**: 14-23.

Carafa AM, Carratù G, Pizzolongo P. 1992. Anatomical observations on the nucellar apex of Wellwitschia mirabilis and the chemical composition of the micropylar drop. *Sexual Plant Reproduction,* **5**: 275-279.

Carlsbecker A, Sundstrom J, Tandre K, Englund M, Kvarnheden A, Johanson U, Engstrom P. 2003. The DAL10 gene from Norway spruce (Picea abies) belongs to a potentially gymnosperm-specific subclass of MADS-box genes and is specifically active in seed cones and pollen cones. *Evolution and Development,* **5**: 551-561.

Carlsbecker A, Sundstrom JF, Englund M, Uddenberg D, Izquierdo L, Kvarnheden A, Vergara-Silva F, Engstrom P. 2013. Molecular control of normal and acrocona mutant seed cone development in Norway spruce (Picea abies) and the evolution of conifer ovule-bearing organs. *New Phytol,* **200**: 261-75.

Causier B, Schwarz-Sommer Z, Davies B. 2010. Floral organ identity: 20 years of ABCs. *Semin Cell Dev Biol,* **21**: 73-9.

Chang S, Puryear J, Cairney J. 1993. A simple and efficient method for isolating RNA from pine trees. *Plant Molecular Biology Reporter,* **11**: 113-116.

Conesa A, Gotz S, Garcia-Gomez JM, Terol J, Talon M, Robles M. 2005. Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics,* **21**: 3674-6.

Cortegano I, Civantos E, Aceituno E, del Moral A, Lopez E, Lombardero M, del Pozo V, Lahoz C. 2004. Cloning and expression of a major allergen from Cupressus arizonica pollen, Cup a 3, a PR-5 protein expressed under polluted environment. *Allergy,* **59**: 485-90.

Coulter A, Poulis BAD, von Aderkas P. 2012. Pollination drops as dynamic apoplastic secretions. *Flora - Morphology, Distribution, Functional Ecology of Plants,* **207**: 482-490.

Daminato M, Masiero S, Resentini F, Lovisetto A, Casadoro G. 2014. Characterization of TM8, a MADS-box gene expressed in tomato flowers. *BMC Plant Biology,* **14**.

Dogra PD. 1967. Seed sterility and disturbances in embryogeny in conifers with particular reference to seed testing and tree breeding in Pinaceae. *Studia Forestalia Suecica,* **45**: 4-28.

Doyle J, O' Leary M. 1935. Pollination in Tsuga, Cedrus, Pseudotsuga, and Larix. With 2 plates. *The Scientific Proceedings of the Royal Dublin Society,* **XXI**: 191-204.

Dwyer KG, Kanakachari MK, Mahosky DI, Acciai J, Kudish BI, Miller JE, Nasrallah ME, Nasrallah JB. 1994. A superfamily of S locus-related sequences in Arabidopsis: diverse structures and expression patterns. *The Plant Cell,* **6**: 1829-1843.

Elias JE, Haas W, Faherty BK, Gygi SP. 2005. Comparative evaluation of mass spectrometry platforms used in large-scale proteomics investigations. *Nature Methods,* **2**: 667-75.

Englund M, Carlsbecker A, Engstrom P, Vergara-Silva F. 2011. Morphological "primary homology" and expression of AG-subfamily MADS-box genes in pines, podocarps, and yews. *Evol Dev,* **13**: 171-81.

Fernando DD, Long SM, Sniezko RA. 2005. Sexual reproduction and crossing barriers in white pines: the case between Pinus lambertiana (sugar pine) and P. monticola (western white pine). *Tree Genetics & Genomes,* **1**: 143-150.

Fernando DD, Quinn CR, Brenner ED, Owens JN. 2010. Male gametophyte development and evolution in extant gymnosperms. *International Journal of Plant Developmental Biology,* **4**: 47-63.

Fletcher JC. 2001. The ULTRAPETALA gene controls shoot and floral meristem size in Arabidopsis. *Development,* **128**: 1323-1333.

Franklin-Tong VE. 2008. Self-incompatibility in flowering plants. Berlin, Germany: Springer-Verlag Berlin Heidelberg.

Galili G, Sengupta-Gopalan C, Ceriotti A. 1998. The endoplasmic reticulum of plant cells and its role in protein maturation and biogenesis of oil bodies. *Plant Molecular Biology,* **38**: 1-29.

Gelbart G, von Aderkas P. 2002. Ovular secretions as part of pollination mechanisms in conifers. *Annals of Forest Science,* **59**: 345-357.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nature Biotechnology,* **29**: 644-52.

Gramzow L, Theissen G. 2010. A hitchhiker's guide to the MADS world of plants. *Genome Biology,* **11**: 214-225.

Gramzow L, Weilandt L, Theissen G. 2014. MADS goes genomic in conifers: towards determining the ancestral set of MADS-box genes in seed plants. *Ann Bot,* **114**: 1407-29.

Gross-Hardt R, Lenhard M, Laux T. 2002. WUSCHEL signaling functions in interregional communication during Arabidopsis ovule development. *Genes and Development,* **16**: 1129-38.

Grover A. 2012. Plant Chitinases: Genetic Diversity and Physiological Roles. *Critical Reviews in Plant Sciences,* **31**: 57-73.

Ham BM, Yang F, Jayachandran H, Jaitly N, Monroe ME, Gritsenko MA, Livesay EA, Zhao R, Purvine SO, Orton D, Adkins JN, Camp DG, 2nd, Rossie S, Smith RD. 2008. The influence of sample preparation and replicate analyses on HeLa Cell phosphoproteome coverage. *Journal of Proteome Research,* **7**: 2215-21.

Han Y, Ma B, Zhang K. 2005. SPIDER: software for protein identification from sequence tags with de novo sequencing error. *Journal of Bioinformatics and Computational Biology,* **3**: 697-716.

Hon WC, Griffith M, Mlynarz A, Kwok YC, Yang DS. 1995. Antifreeze proteins in winter rye are similar to pathogenesis-related proteins. *Plant Physiology,* **109**: 879-89.

Hruba P, Honys D, Twell D, Capkova V, Tupy J. 2005. Expression of beta-galactosidase and beta-xylosidase genes during microspore and pollen development. *Planta,* **220**: 931-40.

Huang N, Stebbins GL, Rodriguez RL. 1992. Classification and evolution of alpha-amylase genes in plants. *Proceedings of the National Academy of Sciences U S A,* **89**: 7526-7530.

Igic B, Lande R, Kohn JR. 2008. Loss of Self ‐ Incompatibility and Its Evolutionary Consequences. *International Journal of Plant Sciences,* **169**: 93-104.

Iwano M, Takayama S. 2012. Self/non-self discrimination in angiosperm self-incompatibility. *Curr Opin Plant Biol,* **15**: 78-83.

Jain SK. 1976. The evolution of inbreeding in plants. *Annual Review of Ecology and Systematics,* **7**: 469-495.

Jelitto-Van Dooren EP, Vidal S, Denecke J. 1999. Anticipating endoplasmic reticulum stress. A novel early response before pathogenesis-related gene induction. *Plant Cell,* **11**: 1935-44.

Jofuku KD, den Boer BGW, Montagu MV, Okamuro JK. 1994. Control of Arabidopsis flower and seed development by the homeotic gene APETALA2. *The Plant Cell,* **6**: 1211-1225.

Joshi NA, Fass JN. 2011. Sickle: A sliding-window, adaptive, quality-based trimming tool for FastQ files

Kemp BP, Doughty J. 2007. S cysteine-rich (SCR) binding domain analysis of the Brassica self-incompatibility S-locus receptor kinase. *New Phytologist,* **175**: 619-29.

Kormatuk A. 1999. Self-incompatibility in silver fir (Abies alba). *Biologia Bratislava,* **54**: 101-105.

Kosinski G. 1986. Megagametogenesis, fertilization, and embryo development in Larix decidua. *Canadian Journal of Forest Research,* **16**: 1301-1309.

Kotake T, Li YQ, Takahashi M, Sakurai N. 2000. Characterization and function of wall-bound exo-β-glucanases of Lilium longiflorum pollen tubes. *Sexual Plant Reproduction,* **13**: 1-9.

Kramer EM, Irish VF. 2000. Evolution of the Petal and Stamen Developmental Programs: Evidence from Comparative Studies of the Lower Eudicots and Basal Angiosperms. *International Journal of Plant Sciences,* **161**: S29-S40.

Kramer EM, Jaramillo MA, Di Stilio VS. 2003. Patterns of gene duplication and functional evolution during the diversification of the AGAMOUS subfamily of MADS Box genes in angiosperms. *Genetics,* **166**: 1011-1023.

Laux T, Mayer KFX, Berger J, Jurgens G. 1996. The WUSCHEL gene is required for shoot and floral meristem integrity in Arabidopsis. *Development,* **122**: 87-96.

Leslie AB, Beaulieu JM, Rai HS, Crane PR, Donoghue MJ, Mathews S. 2012. Hemisphere-scale differences in conifer evolutionary dynamics. *Proceediings of the National Academy of Sciences U S A,* **109**: 16217-21.

Levy A, Erlanger M, Rosenthal M, Epel BL. 2007. A plasmodesmata-associated beta-1,3-glucanase in Arabidopsis. *Plant Journal,* **49**: 669-82.

Lin IW, Sosso D, Chen LQ, Gase K, Kim SG, Kessler D, Klinkenberg PM, Gorder MK, Hou BH, Qu XQ, Carter CJ, Baldwin IT, Frommer WB. 2014. Nectar secretion requires sucrose phosphate synthases and the sugar transporter SWEET9. *Nature,* **508**: 546-9.

Liu JJ, Sturrock R, Ekramoddoullah AK. 2010. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. *Plant Cell Reports,* **29**: 419-36.

Lohse M, Bolger AM, Nagel A, Fernie AR, Lunn JE, Stitt M, Usadel B. 2012. RobiNA: a user-friendly, integrated software solution for RNA-Seq-based transcriptomics. *Nucleic Acids Research,* **40**: W622-7.

Lovisetto A, Guzzo F, Tadiello A, Toffali K, Favretto A, Casadoro G. 2012. Molecular analyses of MADS-box genes trace back to Gymnosperms the invention of fleshy fruits. *Mol Biol Evol,* **29**: 409-19.

Mao L, Begum D, Chuang H, Budlman MA, Szymkowiak EJ, Irish EE, Wing RA. 2000. JOINTLESS is a MADS-box gene controlling tomato flower abscission zone development. *Nature,* **406**: 910-913.

Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBNet,* **17**: 10-12.

Mathews S, Kramer EM. 2012. The evolution of reproductive structures in seed plants: a re-examination based on insights from developmental genetics. *New Phytologist,* **194**: 910-23.

McWilliam JR. 1959. Interspecific Incompatibility in Pinus. *American Journal of Botany,* **46**: 425-433.

Melzer R, Wang YQ, Theissen G. 2010. The naked and the dead: the ABCs of gymnosperm reproduction and the origin of the angiosperm flower. *Semin Cell Dev Biol,* **21**: 118-28.

Mouradov A, Hamdorf B, Teasdale RD, Kim JT, Winter K, Theissen G. 1999. A DEF/GLO-like MADS-Box Gene From a Gymnosperm: Pinus radiata Contains an Ortholog of Angiosperm B Class Floral Homeotic Genes. *Developmental Genetics,* **25**: 245-252.

Moyroud E, Kusters E, Monniaux M, Koes R, Parcy F. 2010. LEAFY blossoms. *Trends Plant Sci,* **15**: 346-52.

Mugnaini S, Nepi M, Guarnieri M, Piotto B, Pacini E. 2007. Pollination drop in Juniperus communis: response to deposited material. *Annals of Botany,* **100**: 1475-81.

Nakano T, Kimbara J, Fujisawa M, Kitagawa M, Ihashi N, Maeda H, Kasumi T, Ito Y. 2012. MACROCALYX and JOINTLESS interact in the transcriptional regulation of tomato fruit abscission zone development. *Plant Physiol,* **158**: 439-50.

Nardmann J, Reisewitz P, Werr W. 2009. Discrete shoot and root stem cell-promoting WUS/WOX5 functions are an evolutionary innovation of angiosperms. *Molecular Biology and Evolution,* **26**: 1745-55.

Nepi M, von Aderkas P, Wagner R, Mugnaini S, Coulter A, Pacini E. 2009. Nectar and pollination drops: how different are they? *Annals of Botany,* **104**: 205-19.

Nilsson L, Carlsbecker A, Sundas-Larsson A, Vahala T. 2007. APETALA2 like genes from Picea abies show functional similarities to their Arabidopsis homologues. *Planta,* **225**: 589-602.

O'Leary SJB, Joseph C, von Aderkas P. 2004. Origin of arabinogalactan proteins in the pollination drop of Taxus × media. *Austrian Journal of Forest Science,* **121**: 35-46.

O'Leary SJB, Poulis BAD, von Aderkas P. 2007. The identification of two thaumatin-like proteins (TLPs) in the pollination drop of hybrid yew that may play a role in pathogen defence during pollen collection. *Tree Physiology,* **27**: 1649-1659.

O`Leary SJB, von Aderkas P. 2005. Postpollination drop production in hybrid larch is not related to the diurnal pattern of xylem water potential. *Trees,* **20**: 61-66.

Owens JN, Bennett J, L'Hirondelle S. 2005. Pollination and cone morphology affect cone and seed production in lodgepole pine seed orchards. *Canadian Journal of Forest Research,* **35**: 383-400.

Owens JN, Blake MD. 1983. Pollen morphology and development of the pollination mechanism in Tsuga heterophylla and T. mertensiana. *Canadian Journal of Botany,* **61**: 3041-3049.

Owens JN, Colangeli AM, Morris SJ. 1990. Factors affecting seed set in Douglas-fir (Pseudotsuga menziesii). *Canadian Journal of Botany,* **69**: 229-238.

Owens JN, Simmons SJ. 1987. The pollination mechanism of Engelmann spruce (Picea engelmanii). *Canadian Journal of Botany,* **65**: 1439-1450.

Passardi F, Cosio C, Penel C, Dunand C. 2005. Peroxidases have more functions than a Swiss army knife. *Plant Cell Reports,* **24**: 255-65.

Pata MO, Hannun YA, Ng CK. 2010. Plant sphingolipids: decoding the enigma of the Sphinx. *New Phytologist,* **185**: 611-30.

Poulis BAD, O'Leary SJB, Haddow JD, 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.

Prior N, Little SA, Pirone C, Gill JE, Smith D, Han J, Hardie D, O'Leary SJ, Wagner RE, Cross T, Coulter A, Borchers C, Olafson RW, von Aderkas P. 2013. Application of proteomics to the study of pollination drops. *Applications in Plant Science,* **1**.

Prunet N, Morel P, Thierry AM, Eshed Y, Bowman JL, Negrutiu I, Trehin C. 2008. REBELOTE, SQUINT, and ULTRAPETALA1 function redundantly in the temporal regulation of floral meristem termination in Arabidopsis thaliana. *Plant Cell,* **20**: 901-19.

Rai HS, Reeves PA, Peakall R, Olmstead RG, Graham SW. 2008. Inference of higher-order conifer relationships from a multi-locus plastid data set. *Botany,* **86**: 658-669.

Rejon JD, Delalande F, Schaeffer-Reiss C, Carapito C, Zienkiewicz K, de Dios Alche J, Rodriguez-Garcia MI, Van Dorsselaer A, Castro AJ. 2013. Proteomics profiling reveals novel proteins and functions of the plant stigma exudate. *Journal of Experimental Botany,* **64**: 5695-705.

Runions CJ, Owens JN. 1998. Evidence of pre-zygotic self-incompatibility in a conifer. In: Rudall PJ, Owens SJ, eds. *Reproductive Biology*: Royal Botanic Gardens, Kew.

Rutledge R, Regan S, Nicolas O, Fobert P, Cote C, Bosnich W, Kauffeldt C, Sunohara G, Seguin A, Stewart D. 1998. Characterization of an AGAMOUS homologue from the conifer black spruce (Picea mariana) that produces floral homeotic conversions when expressed in Arabidopsis. *The Plant Journal,* **15**: 625-634.

Sela-Buurlage MB, Ponstein AS, Bres-Vloemans SA, Melchers LS, Van Den Elzen P, Cornelissen B. 1993. Only specific tobacco (Nicotiana tabacum) chitinases and [beta]-1,3-glucanases exhibit antifungal activity. *Plant Physiology,* **101**: 857-863.

Seridi-Benkaddour R, Chesnoy L. 1988. Secretion and Composition of the Pollination Drop in Cephalotaxus drupaceae (Gymnosperm, Cephalotaxeae). In: Cresti M, Gori P, Pacini E, eds. *Sexual Reproduction in Higher Plants*. Berlin: Springer Verlag.

Shiu SH, Bleecker AB. 2003. Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. *Plant Physiology,* **132**: 530-43.

Shiu SH, Bleeker AB. 2001. Plant receptor-like kinase gene family: diversity, function, and signaling. *Science Signaling,* **2001**: re22.

Singh H. 1978. *Embryology of gymnosperms.* Berlin: Gerbrüder Borntraeger.

Stebbins GL. 1957. Self fertlization and population variability in the higher plants. *The American Naturalist,* **91**: 337-354.

Stein JC, Howlett B, Boyes DC, Nasrallah ME, Nasrallah JB. 1991. Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of Brassica oleracea. *Proceedings of the National Academy of Sciences U S A,* **88**: 8816-8820.

Sundstrom J, Carlsbecker A, Svensson ME, Svenson M, Johanson U, Theissen G, Engstrom P. 1999. MADS-Box genes active in developing pollen cones of Norway Spruce (Picea abies) are homologous to the B-Class floral homeotic genes in angiosperms. *Developmental Genetics,* **25**: 253-266.

Sundstrom J, Engstrom P. 2002. Conifer reproductive development involves B-type MADS- box genes with distinct and different activities in male organ primordia. *The Plant Journal,* **31**: 161-169.

Takaso T, Owens JN. 1995. Pollination drop and microdrop secretions in Cedrus. *International Journal of Plant Sciences,* **156**: 640-649.

Takayama S, Isogai A. 2005. Self-incompatibility in plants. *Annual Review of Plant Biology,* **56**: 467-489.

Tandre K, Svenson M, Svensson ME, Engstrom P. 1998. Conservation of gene structure and activity in the regulation of reproductive organ development of conifers and angiosperms. *The Plant Journal,* **15**: 615-623.

Theissen G. 2001. Development of floral organ identity: stories from the MADS house. *Current Opinion in Plant Biology* **4**: 75-85.

Theissen G, Becker A. 2004. Gymnosperm orthologues of class B floral homeotic genes and their impact on understanding flower origin. *Critical Reviews in Plant Sciences,* **23**: 129-148.

Tomlinson PB, Braggins JE, Rattenbury JA. 1991. Pollination drop in relation to cone morphology in Podocarpaceae: A novel reproductive mechanism. *American Journal of Botany,* **78**: 1289-1303.

Tomlinson PB, Braggins JE, Rattenbury JA. 1997. Contrasted pollen capture mechanisms in Phyllocladaceae and certain Podocarpaceae (Coniferales). *American Journal of Botany,* **84**: 214-223.

van Loon LC, Rep M, Pieterse CM. 2006. Significance of inducible defense-related proteins in infected plants. *Annual Review of Phytopathology,* **44**: 135-62.

Vasquez-Lobo A, Carlsbecker A, Vergara-Silva F, Alvarez-Buyila ER, Pineroa D, Engstrom P. 2007. Characterization of the expression patterns of

LEAFY/FLORICAULA and NEEDLY orthologues in female and male cones of the conifer genera Picea, Podocarpus, and Taxus: implications for current evo-devo hypotheses for gymnosperms. *Evolution and Development,* **9**: 446-459.

Vogel C, Marcotte EM. 2012. Insights into the regulation of protein abundance from proteomic and transcriptomic analyses. *Nature Reviews Genetics,* **13**: 227-32.

von Aderkas P, Nepi M, Rise M, Buffi F, Guarnieri M, Coulter A, Gill K, Lan P, Rzemieniak S, Pacini E. 2012. Post-pollination prefertilization drops affect germination rates of heterospecific pollen in larch and Douglas-fir. *Sexual Plant Reproduction,* **25**: 215-25.

von Aderkas P, Prior NA, Gagnon S, Little S, Cross T, Hardie D, Borchers C, Thornburg R, Hou C, Lunny A. 2015. Degradome and secretome of pollination drops of Ephedra. *Botanical Review,* **81**: 1-27.

Wagner RE, Mugnaini S, Sniezko R, Hardie D, Poulis B, Nepi M, Pacini E, von Aderkas P. 2007. Proteomic evaluation of gymnosperm pollination drop proteins indicates highly conserved and complex biological functions. *Sexual Plant Reproduction,* **20**: 181-189.

Weigel D, Meyerowitz EM. 1994. The ABCs of floral homeotic genes. *Cell,* **78**: 203-209.

Weigel G, Alvarez J, Smyth DR, Yanofsky MF, Meyerowitz EM. 1992. LEAFY controls floral meristem identity in Arabidopsis. *Cell,* **69**: 843-859.

Winter K, Becker A, Munster T, Kim JT, Saedler H, Theissen G. 1999. MADS-box genes reveal that gnetophytes are more closely related to conifers than to flowering plants. *Proceedings of the National Academy of Sciences U S A,* **96**: 7342–7347.

Xing S, Li M, Liu P. 2013. Evolution of S-domain receptor-like kinases in land plants and origination of S-locus receptor kinases in Brassicaceae. *BMC Evolutionary Biology,* **13**: 1-11.

Xuan YH, Hu YB, Chen LQ, Sosso D, Ducat DC, Hou BH, Frommer WB. 2013. Functional role of oligomerization for bacterial and plant SWEET sugar transporter family. *Proceedings of the National Academy of Sciences U S A,* **110**: E3685-94.

Yatomi R, Nakamura S, Nakamura N. 2002. Immunochemical and cytochemical detection of wall components of germinated pollen of gymnosperms. *Grana,* **41**: 21-28.

Ziegler H. 1959. Über die Zusammensetzung des “Bestäubungstropfens” und den Mechanismus seiner Sekretion. *Planta,* **52**: 587-599.

TABLES

Table 1. Proteins previously identified in the pollination drops of gymnosperms (adapted from Coulter et al., 2012 and von Aderkas et al., 2015). Plant names are abbreviated as follows: Chae. laws.= *Chaemecyparis lawsonia*; Jun. com. = *Juniperus communis*; Jun. oxy = *Juniperus oxycedrus*; Pseud. menz. = *Pseudotsuga menziesii*, Tax. X med. = *Taxus* X *media*; Wel. mir. = *Welwitschia mirabilis*; Eph. com = *Ephedra communis*; Eph. foe. = *Ephedra foeminea*; Eph. min. = *Ephedra minuta*; Eph. tri. = *Ephedra trifurca*; Eph. lik. = *Ephedra likiangensis*; Eph. mon. = *Ephedra monosperma*; Eph. com. = *Ephedra compacta*.

	Chae. .laws.	***Jun. com.***	***Jun. oxy.***	***Pseu. men.***	***Tax. X med.***	***Wel. mir.***	***Eph. com.***	***Eph. .foe.***	***Eph. min.***	***Eph. tri.***	***Eph. lik***	***Eph. mon.***	***Eph. com.***

β-D-glucan exohydrolase	x												
glucan 1,3-β-glucosidase (or precursor)	x	x								x		x	
subtilisin-like proteinase	x												
glycosyl hydrolase		x											
subtilisin-like proteinase		x											
glucanase-like protein			x										
chitinase			x			x		x		x			
thaumatin-like protein	x	x	x						x				
xylosidase				x			x	x	x	x			x
galactosidase				x			x		x	x			x
peroxidase				x			x			x	x		x
invertase				x									
aspartyl protease				x			x			x	x		x
serine carboxypeptidase (-like) protein				x				x		x		x	
arabinogalactan protein					x								
malate dehydrogenase										x			
peptidase											x		
superoxide dismutase							x						

Table 2. Proteins identified in the pollination drops of *C. koreana* (outdoor), *C. sinensis* (indoor) and *C. sinensis* (outdoor). Orange indicates proteins identified in all three samples, green indicates proteins identified in *C. koreana* (outdoor) and *C. sinensis* (outdoor), and pink indicates

proteins identified in *C. koreana* (indoor) and *C. koreana* (outdoor). Bold indicates proteins that were detected in previous studies.

***C. sinensis* (outdoor)**	***C. koreana* (outdoor)**	***C. koreana* (indoor)**
Class IV **chitinase**	Class IV **chitinase**	Class IV **chitinase**
Luminal-binding protein (BiP)	Luminal-binding protein (BiP)	Luminal-binding protein (BiP)
Peroxidase	**Peroxidase**	**Peroxidase**
Pollen allergen CJP38	Pollen allergen CJP38	Pollen allergen CJP38
Thaumatin-like protein	**Thaumatin-like protein**	**Thaumatin-like protein**
Alpha-galactosidase	Alpha-**galactosidase**	Acidic endochitinase
Alpha-amylase type B isozyme	Alpha-amylase type B isozyme	Neutral ceramidase
Beta-galactosidase	Beta-**galactosidase**	Alkaline leaf **peroxidase**
Putative class I chitinase	Class I **chitinase**	Alpha-**galactosidase**
Cup a 3 protein	Cup a 3 protein	Beta-1_3-**glucanase**_ basic
Pathogenesis-related protein	Pathogenesis-related protein	**Chitinase** 5
PR5 allergen Cup s	PR5 allergen Jun r	Class III **peroxidase**
Alpha-1_4-glucan-protein synthase	Acidic endochitinase	Class IV endochitinase
Alpha-amylase isozyme	Neutral ceramidase	Endo-beta-1 3-**glucanase**
Ascorbate peroxidase	**Acidic thaumatin-like protein**	Endochitinase
Aspartate aminotransferase	Allergen Ara h 1	**Glycoside hydrolase family**
Calmodulin	Allergen Arah3/Arah4	Mannosylglycoprotein endo-beta-mannosidase
Cytosolic glyceraldehyde-3-phosphate dehydrogenase	Alpha-amylase	Secretory peroxidase
Elongation factor 1-alpha	Ara h 1 allergen	**Serine carboxypeptidase-like**
Elongation factor	Arachin	
Enolase	Conarachin	
Eukaryotic translation elongation factor	Endo-beta-mannanase	
Glutathione-S-transferase	**Glucan 1_3-beta-glucosidase**	
Glyceraldehyde 3-phosphate dehydrogenase	**Glucan** endo-**1_3-beta-glucosidase**	
Heat shock protein 70	Glutathione reductase	
Histone H2A variant 1	Gly1	
Isocitrate dehydrogenase [NADP]	Glycinin	
Monodehydroascorbate reductase	Iso-Ara h3	
Probable histone H2A variant 1	Malate dehydrogenase	
Triosephosphate isomerase	Polygalacturonase	
UDP-glucose:protein transglucosylase-like protein SlUPTG	Storage protein	
	Zinc finger protein_ putative	

Supplementary Table 1. Peptides and proteins identified in the pollination drops of *C. koreana* (outdoor), *C. sinensis* (indoor) and *C. sinensis* (outdoor).

Supplementary Table 2. BLAST results (nr, SwissProt, and trEMBL) for the ovule transcriptome of *C. sinensis*.

Supplementary Table 3. Blast2go annotations for the ovule transcriptome of *C. sinensis.*

Supplementary Table 4. List of transcripts and proteins found in both the transcriptome and proteome datasets.

Supplementary Table 5. Transcripts identified in the ovules of *C. sinensis* with the putative functions of sugar transport, pollen-ovule interactions, and development.

Figure 1. *Cephalotaxus koreana* A. Branch with male cones before pollen shedding. B. Female cones with pollination drops.

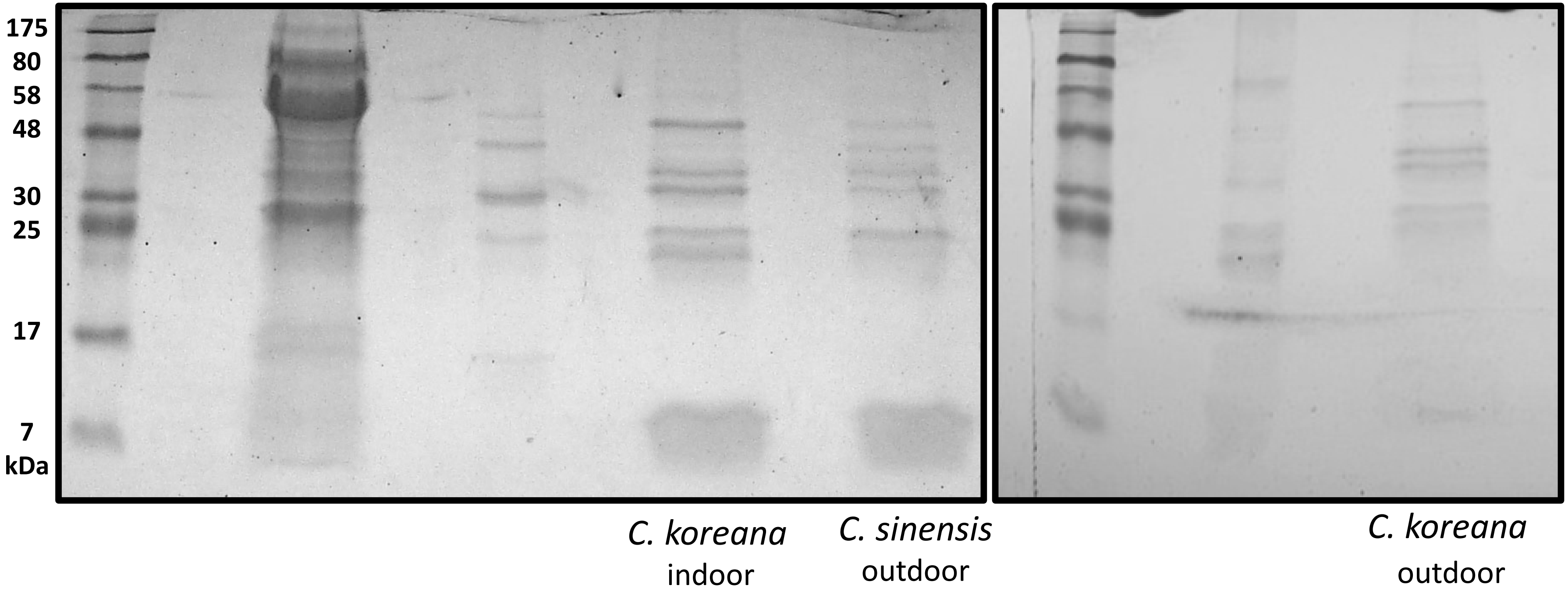
Figure 2. SDS-PAGE of pollination drops from *C. koreana* (outdoor), *C. sinensis* (indoor) and *C. sinensis* (outdoor).

Figure 3. Histogram of Trinity-assembled transcript sizes.

Figure 4. Histogram of Gene Ontology classifications, subdivided into biological processes, cellular components, and molecular functions. The right y-axis = number of genes in a category. The left y-axis = the percentage of genes in a category.

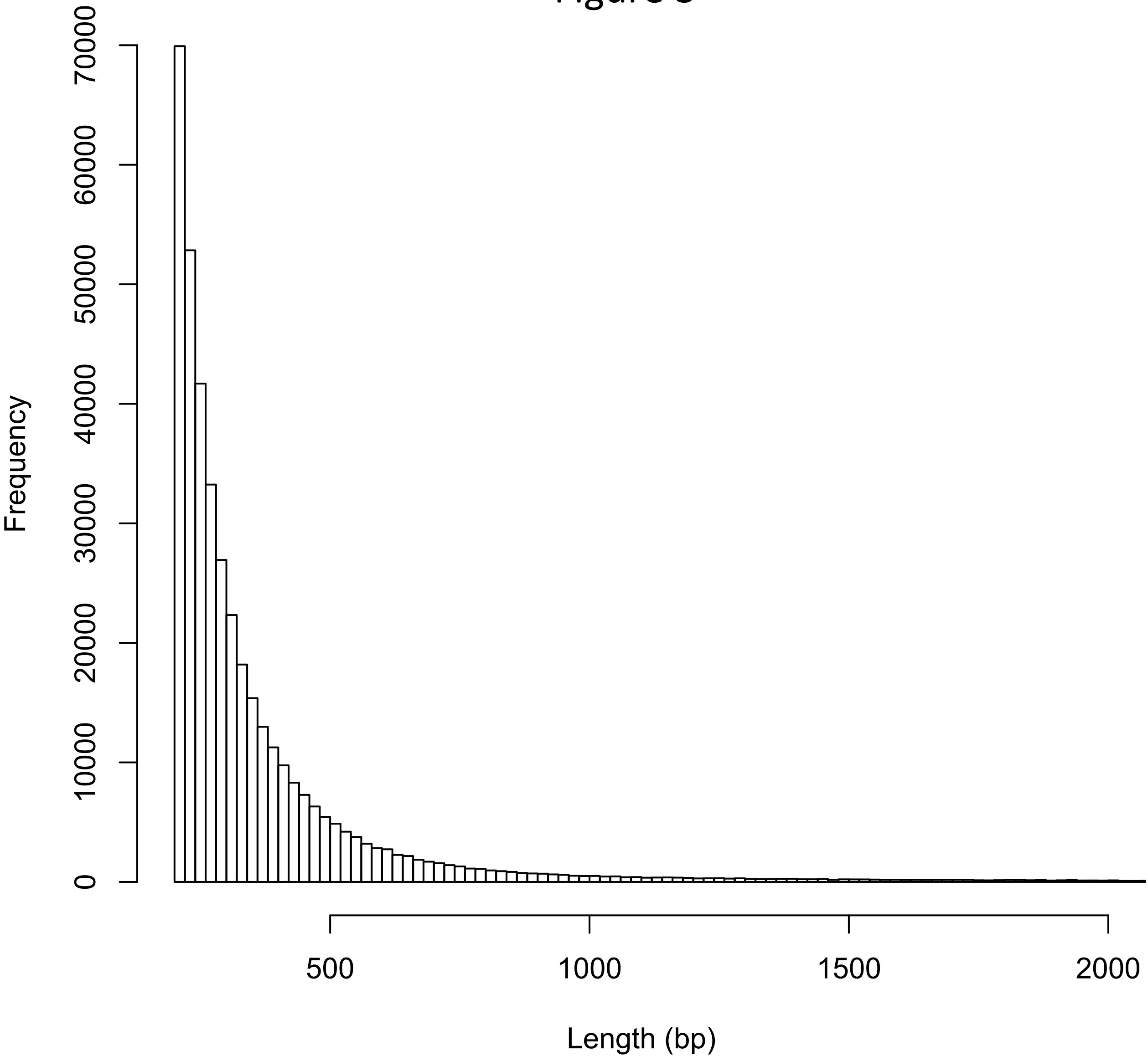
Figure 1

Figure 2


175
80
58
48
30
25
17
7
kDa
C. koreana
indoor
C. sinensis
outdoor
C. koreana
outdoor

Histogram of Transcript Length

Figure 3



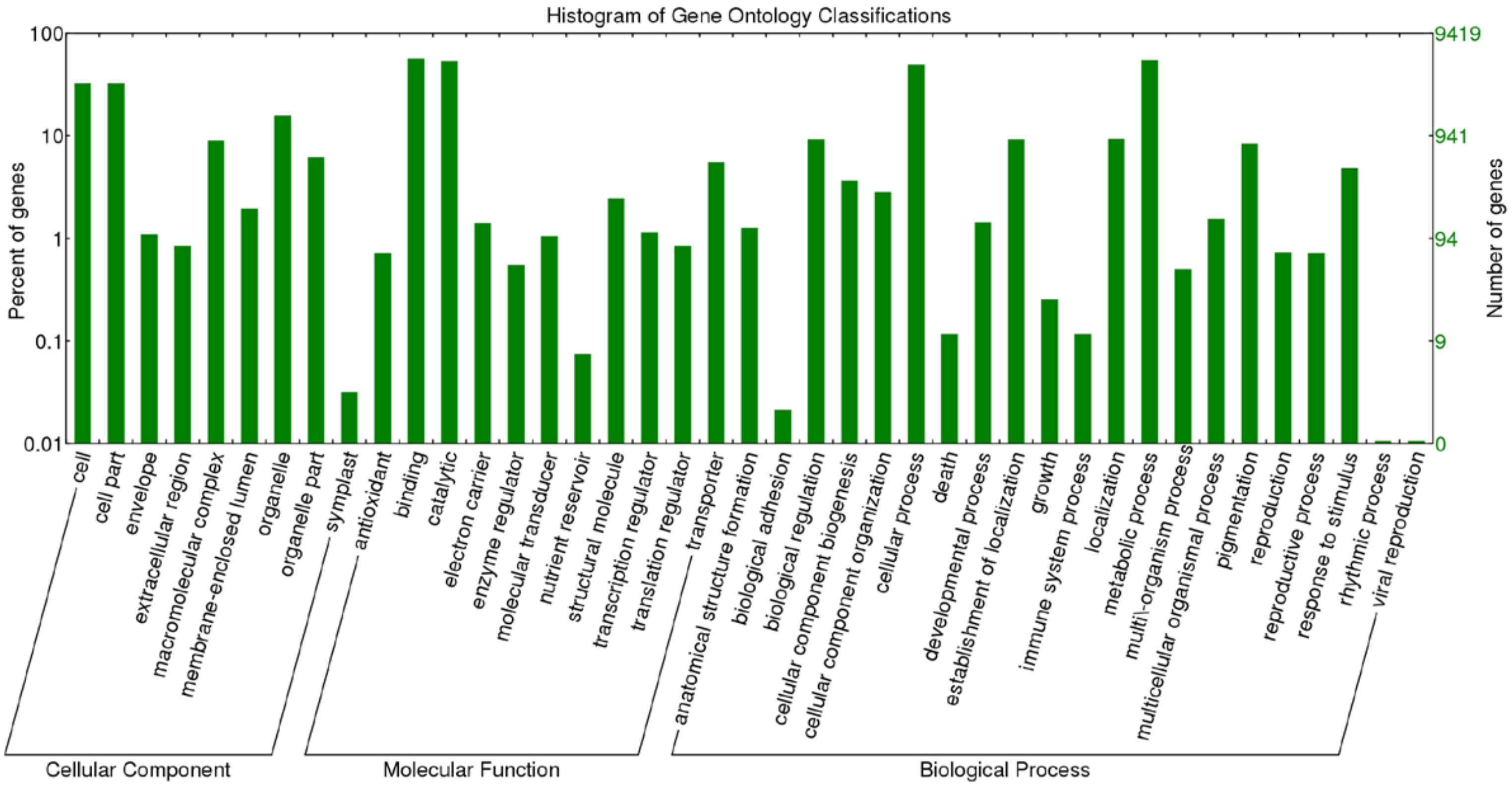

Histogram of Gene Ontology Classifications
Percent of genes
100
10
1
0.1
0.01
Number of genes
9419
941
94
9
0
cell
cell part
envelope
extracellular region
macromolecular complex
membrane-enclosed lumen
organelle
organelle part
symplast
antioxidant
binding
catalytic
electron carrier
enzyme regulator
molecular transducer
nutrient reservoir
structural molecule
transcription regulator
translation regulator
transporter
anatomical structure formation
biological adhesion
biological regulation
cellular component biogenesis
cellular component organization
cellular process
death
developmental process
establishment of localization
growth
immune system process
localization
metabolic process
multi\-organism process
multicellular organismal process
pigmentation
reproduction
reproductive process
response to stimulus
rhythmic process
viral reproduction
Cellular Component
Molecular Function
Biological Process

Figure 4