Proteins in gymnosperm pollination drops.

by

Natalie Annastasia Prior
B.Sc., University of Victoria, 2008
M.Sc., University of Victoria, 2010

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

DOCTOR OF PHILOSOPHY

in the Department of Biology

© Natalie Annastasia Prior, 2014
University of Victoria

All rights reserved. This dissertation may not be reproduced in whole or in part, by photocopy or other means, without the permission of the author.

Supervisory Committee

Proteins in gymnosperm pollination drops.

by

Natalie Annastasia Prior
B.Sc., University of Victoria, 2008
M.Sc., University of Victoria, 2010

Supervisory Committee

Dr. Patrick von Aderkas, Supervisor
(Department of Biology)

Dr. Gerry Allen, Departmental Member
(Department of Biology)

Dr. Ben Koop, Departmental Member
(Department of Biology)

Dr. Terry Pearson, Outside Member
(Department of Biochemistry and Microbiology)

Abstract

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

Dr. Gerry Allen, Departmental Member
(Department of Biology)

Dr. Ben Koop, Departmental Member
(Department of Biology)

Dr. Terry Pearson, Outside Member
(Department of Biochemistry and Microbiology)

Most gymnosperms produce a pollination drop that captures and transports pollen into the ovule. Pollination drops have other functions. These include influencing pollen germination and pollen tube growth, defending the ovule from pathogens and providing a food reward in insect-pollinated gymnosperms. Mineral and organic molecules, including proteins, are responsible for these additional functions. To date, pollination drops from a handful of conifers and one non-conifer gymnosperm, *Welwitschia mirabilis*, have been subjected to proteomic analysis. In the present study, tandem mass spectrometry was used to detect proteins in all gymnosperm lineages: cycads (*Ceratozamia hildae*, *Cycas rumphii*, *Zamia furfuracea*); Gnetales (*Ephedra compacta*, *E. distachya*, *E. foeminea*, *E. likiangensis*, *E. minuta, E. monosperma*, *E. trifurca*; *Gnetum gnemon*; *Welwitschia mirabilis*); *Ginkgo biloba*; conifers (*Taxus* x *media*). PEAKS 6 DB (Bioinformatics Solutions, Waterloo, ON, Canada) was used to make protein identifications. Proteins were detected in all gymnosperm species analyzed. The numbers of proteins identified varied between samples as follows: one protein in *Welwitschia* female; nine proteins in *Cycas rumphii*; 13 proteins on average in *Ephedra* spp.; 17 proteins in *Gnetum gnemon*; 38 proteins on average in *Zamia furfuracea*; 57 proteins in *Ginkgo biloba*; 61 proteins in *Ceratozamia hildae*; 63 in *Taxus* x *media*; 138 proteins in *Welwitschia* male. The types of proteins identified varied widely. Proteins involved in

carbohydrate modification, e.g. galactosidase, chitinase, glycosyl hydrolase, glucosidase, were present in most gymnosperms. Similarly, defence proteins, e.g. reduction-oxidation proteins, lipid-transfer proteins and thaumatin-like proteins, were identified in many gymnosperms. Gymnosperms that develop a deep pollen chamber as the nucellus degrades, e.g., cycads, *Ginkgo*, *Ephedra*, generally contained higher proportions of proteins localized to intracellular spaces. These proteins represent the pollination drop degradome. Gymnosperms that either lack a pollen chamber, e.g. *Taxus*, or have a shallow pollen chamber, e.g. *Gnetum*, had greater proportions of extracellular proteins. These proteins represent the pollination drop secretome. Our proteomic analyses support the hypothesis that the pollination drops of all extant gymnosperms constitute complex reproductive secretions.

Table of Contents

Supervisory Committee ii
Abstract iii
Table of Contents v
List of Tables vii
List of Figures x
Acknowledgments xii
Chapter 1: Introduction 1
  Introduction 1
  Defining gymnosperms 3
  Pollination drops 6
    Historical overview 6
    Pollination drops as part of pollination mechanisms 7
    Drop secretion and retraction 9
    Drop constituents and their possible functions 12
  Current knowledge of reproduction in angiosperms compared to gymnosperms 16
    Proteomics of plant reproduction 18
  Proteomic analysis of gymnosperm pollination drops 19
Chapter 2: Degradome and secretome of pollination drops of *Ephedra* 21
  Introduction 21
  Materials and methods 26
    Sample collection 26
    1D SDS PAGE 27
    LC-MS/MS analysis 27
    Data analysis parameters 28
    Scanning electron microscopy 29
  Results 29
    Comparative study of seven *Ephedra* species 29
    Comparative study of *Ephedra monosperma* drops from three dates 42
  Discussion 44
Chapter 3: Using a custom transcriptome-derived database to identify proteins in gymnosperm pollination drops 51
  Introduction 51
    The challenge of peptide and protein identification in non-model organisms 51
  Methods 54
    Pollination drop collection 54
    Proteomics 55
Chapter 4: Proteins in cycad pollination drops 60
  Introduction 60
    Pollination in cycads 60
    Study species 64
    Zooidogamy and pollen tube function 65
  Methods 66
    Collection of pollination drops from cycads 66
  Results 68

Discussion 88
Chapter 5: Proteins in the pollination drops of *Ginkgo biloba* 101
Introduction 101
Methods 103
Pollination drop collection in *Ginkgo biloba* 103
Proteomics 103
Results 103
Discussion 111
Chapter 6: Proteins in the pollination drops of *Gnetum gnemon* and *Welwitschia mirabilis* 118
Introduction 118
Methods 123
Pollination drop collection from *Gnetum gnemon* 123
Pollination drop collection from *Welwitschia mirabilis* 123
Proteomics 123
Results 124
Discussion 144
Chapter 7: Proteins in the pollination drops of *Taxus* x *media* 151
Introduction 151
Methods 153
Collections of drops from *Taxus* x *media* 153
Proteomics 153
Results 154
Discussion 154
Chapter 8: Perspectives 165
Overview of pollination drop proteins in gymnosperm lineages 165
Using a boutique database for pollination drop proteomics 175
Challenges of pollination drop collection 177
Future studies of pollination drops and their proteins 178
Discerning communication between pollen and ovule 178
The need for functional characterization and quantitation of pollination drop proteins 180
Final reflection 182
Literature cited 184
Appendix 205
Appendix 1: Species and sources of transcriptome data for Gymno_DB 205
Appendix 2: Proteins and peptides detected in *Ceratozamia hildae* by PEAKS 6 DB plus SPIDER 210

List of Tables

Table 1. Degradome proteins found in pollination drops of *Ephedra* spp. 32

Table 2. Secretome proteins found in pollination drops of *Ephedra* spp. Proteins that could also be considered degradome are marked with an asterisk 33

Table 3. Peptide sequences and identities of pollination drop proteins found in *Ephedra* spp. Degradome proteins are indicated by a black line in the right margin 34

Table 4. *Ephedra monosperma* pollination drop proteins from three collection dates 43

Table 5. Proteins identified in *Zamia furfuracea* (2011) pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model 69

Table 6. Proteins identified in *Zamia furfuracea* (2012) pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model 73

Table 7. Proteins identified in *Ceratozamia hildae* pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO

Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model 77

Table 8. Proteins identified in *Cycas rumphii* pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model 82

Table 9. Comparative view of protein-types in four samples of cycad pollination drops (*Ceratozamia hildae*, two samples of *Zamia furfuracea* collected in 2011 and 2012, *Cycas rumphii*) 86

Table 10. Proteins identified in *Ginkgo biloba* pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model 104

Table 11. Proteins identified in *Welwitschia mirabilis* female pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant

BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model 125

Table 12. Proteins identified in *Welwitschia mirabilis* sterile ovule (male) pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model 126

Table 13. Proteins identified in *Gnetum gnemon* fertile ovule (female) pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model 142

Table 14. Proteins identified in *Taxus* x *media* pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model 155

Table 15. Overview of proteins detected in gymnosperm pollination drops 166

List of Figures

Figure 1. Gymnosperm pollination drops. A. Short-shoot of *Gingko biloba* L. with ovulate stalks during pollination drop production. B. Pollination drop secreted from the ovule of *G. biloba*. C. Pollination drop exuded by the ovule of *Taxus* x *media* Rehd. D. Post-pollination pre-fertilization drops secreted from ovules of *Pseudotsuga menziesii* (Mirb.) Franco on a single scale removed from a cone. E. Pollination drop exuded from an ovule of *Larix* x *marschlinsii* Coaz. F. Cones of *Chamaecyparis lawsoniana* (A. Murray) Parl., each with several ovules secreting pollination drops. G. Pollination drops at the tips of micropyles extending from two ovules of a female *Ephedra monosperma* C.A.Meyer cone. H. Female cone of *Welwitschia mirabilis* Hook. f. with many long micropylar tubes bearing pollination drops. I. Male cone of *W. mirabilis* with central sterile ovule. J. Pollination drop secreted from sterile ovule of *W. mirabilis* male cone. (Photo credits: Julia Gill A, B, D, E; Dr. Steven O'Leary C; Andrea Coulter F; Dr. Stefan Little G, I, J; Dr. Chad Husby H) 2

Figure 2. Schematic representations of gymnosperm ovules at the time of pollination drop secretion. *Taxus canadensis* Marsh (Tc) does not have a pollen chamber. Other gymnosperms vary in the depth of their pollen chambers (indicated by arrows) from small depressions in *Pinus contorta* Douglas ex. Louden (Pc) and *Ephedra foeminea* Forssk (Ef) to substantial chambers in *Picea sitchensis* (Bong.) Carr. (Ps) and *Ginkgo biloba* L. (Gb). Ovular silhouettes are modified from sections (abbreviated species in brackets) published in Dupler 1920 (Tc), Owens et al. 2005 (Pc), Rydin et al. 2010 (Ef), Owens and Blake 1984 (Ps) and Douglas et al. 2007 (Gb) 10

Figure 3. Scanning electron micrographs of *Ephedra monosperma* ovules. a. An open micropyle. b. A pollination drop partially exuded from the micropyle. c. A pollination drop fully exuded from the micropyle. Bar = 500 µm 22

Figure 4. *Ephedra* ovules. a. Ovule of *E. compacta* with pollination drop. b. *E. monosperma* with an insect feeding on the pollination drop 23

Figure 5. 1D SDS-PAGE of proteins at native concentrations in *Ephedra* pollination drops. Lanes from left to right: molecular weight ladder (kDa), 1. *E. distachya*, 2. *E. distachya*, 3. *E. foeminea*, 4. *E. minuta*, 5. *E. likiangensis*, 6. *E. monosperma* 30

Figure 6. 1D SDS-PAGE of native concentrations of proteins in pollination drops of three gymnosperms: Lane 1. *Larix* x *marschlinsii*, Lane 2. *E. monosperma*, Lane 3. *Gingko biloba.* Figure is only to show number of bands and relative band intensity 31

Figure 7. Pollination drops in cycads. A. *Cycas rumphii* Miq. female plant. B. Female cone of *C. rumphii* during receptivity to pollination. C. Pollination drop on an ovule of *C. rumphii*. D. Receptive female cone of *Ceratozamia kuesteriana* Regel. E. Interior of a

female *C. hildae* G.P. Landry & M.C. Wilson cone bearing ovules with pollination drops. F. *C. hildae* megasporophyll with two ovules, one secreting a pollination drop. G. Pollination drop exuded from the micropyle of *C. kuesteriana* 63

Acknowledgments

I wish to thank everyone who supported this work: Dr. Patrick von Aderkas, Dr. Lisheng Kong, Dr. Stefan Little and all members of the von Aderkas lab; Ian Boyes for writing Linux scripts; Derek Smith, Darryl Hardie, Jun Han, Jason Serpa and the staff at the UVic Genome BC Proteomics Centre; Dr. M. Patrick Griffith, Dr. Chad Husby, Dr. Michael Calonje, Tracy Majellan, Judy Kay, Arantza Strader and the Montgomery Botanical Center team; William Tang for sharing his invaluable knowledge of pollination drop collection in cycads; Mrs. Elaine Spears for allowing us to collect *Zamia* cones from her garden; Dr. Alan Meerow for providing access to *Zamia* cones at the United States Department of Agriculture Subtropical Horticulture Research Station; Dr. Cary Pirone for helpful discussions about pollination drop proteomics; Dr. Dennis W. Stevenson, Dr. Gane Ka-Shu Wong and all contributors to the One KP Project for providing access to their gymnosperm transcriptome datasets; Julia Gill for many hours of collection help; Kelly Sendall and the staff at the Royal British Columbia Museum; Brynn Porter, Steffi Ickert-Bond, Catarina Rydin, Kristina Bolinder, Anders Rydberg, Judy Jernstedt and Israel Loera-Carrizales for allowing me to analyze their pollination drop samples; UVic Department of Biology and the Centre for Forest Biology.

Chapter 1: Introduction

A portion of the subsection "Drop constituents and their possible functions" was published in the book: Little, S., Prior, N.A., Pirone, C. and von Aderkas, P. 2014. Pollen- ovule interactions in gymnosperms. In: Ramawat, K.G., Mérillon, J.M. and K.R. Shivanna (eds.). Reproductive Biology of Plants. CRC Press, Boca Raton, London, New York. 97-117.

Introduction

Modern biochemical tools should be used to decipher pollen-ovule interactions in gymnosperms. Most gymnosperms secrete a small drop of liquid from the ovule at some point during reproduction (Figure 1; Singh 1978). This pollination drop primarily serves to capture and transport pollen into the ovule. Previous work has determined that pollination drops contain a variety of inorganic and organic components, suggesting that they may play a dynamic role in pollen-ovule interactions (Gelbart and von Aderkas 2002). Proteomic analyses of conifer pollination drops have shown that they contain proteins (Poulis et al. 2005; O'Leary et al. 2007; Wagner et al. 2007). To date, the pollination drop of only one non-coniferous gymnosperm species, *Welwitschia mirabilis* Hook.f., has been analyzed using proteomics techniques (Wagner et al. 2007). Proteomic analyses of the remaining gymnosperm groups will allow for comparison of their pollination drop proteomes. This will help us to interpret the potential roles of pollination drop proteins in the four living gymnosperm lineages, and help us to evaluate the possibility of conserved functions for pollination drop proteins through seed-plant evolution.

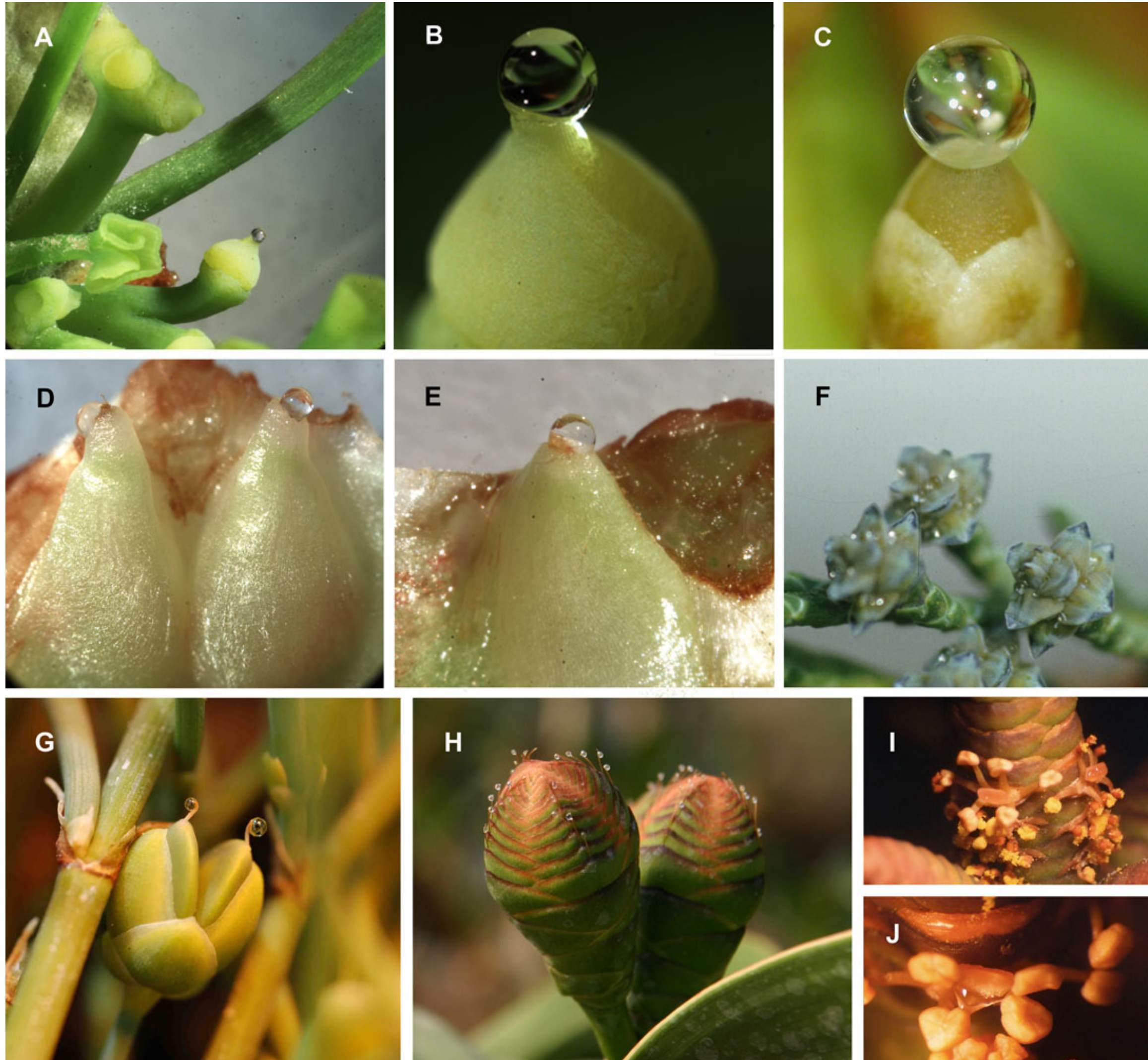

Figure 1. Gymnosperm pollination drops. **A**. Short-shoot of *Gingko biloba* L. with ovulate stalks during pollination drop production. **B**. Pollination drop secreted from the ovule of *G. biloba*. **C**. Pollination drop exuded by the ovule of *Taxus* x *media* Rehd. **D**. Post-pollination pre-fertilization drops secreted from ovules of *Pseudotsuga menziesii* (Mirb.) Franco on a single scale removed from a cone. **E**. Pollination drop exuded from an ovule of *Larix* x *marschlinsii* Coaz. **F**. Cones of *Chamaecyparis lawsoniana* (A. Murray) Parl., each with several ovules secreting pollination drops. **G**. Pollination drops at the tips of micropyles extending from two ovules of a female *Ephedra monosperma* C.A.Meyer cone. **H**. Female cone of *Welwitschia mirabilis* Hook. f. with many long micropylar tubes bearing pollination drops. **I**. Male cone of *W. mirabilis* with central sterile ovule. **J**. Pollination drop secreted from sterile ovule of *W. mirabilis* male cone. (Photo credits: Julia Gill A, B, D, E; Dr. Steven O'Leary C; Andrea Coulter F; Dr. Stefan Little G, I, J; Dr. Chad Husby H)

Defining gymnosperms

The term *gymnosperm* groups together plants that have exposed ovules around the time of pollination, as opposed to the term *angiosperm*, which groups together plants that have ovules enclosed within carpels at pollination (Tomlinson and Takaso 2002; Tomlinson 2012). Extant gymnosperms are placed into four groups: cycads, *Ginkgo*, Gnetales and conifers. They make up the four extant non-flowering seed-plant groups; the angiosperms make up the fifth seed-plant group. In general terms, the living gymnosperms can be thought of as cone-bearing seed plants, while the angiosperms can be thought of as flowering seed plants.

There are about 1100 species of gymnosperms (Mathews 2009). The species diversity of individual groups varies. The cycads include 10 genera and 331 species (Osborne et al. 2012). *Gingko biloba* L. is the only extant representative of its group (Royer et al. 2003). The Gnetales, which include *Gnetum*, *Ephedra* and *Welwitschia*, are represented by about 70-80 species (Rydin et al. 2010). Conifers are the most diverse, with about 670 extant species within 70 genera (Rai et al. 2008).

Together, gymnosperms have an extensive distribution. The distribution of conifers is the most widespread, reaching from the tree line of the Arctic to the tropics of Australia, and wrapping around Earth. Gnetales are more restricted: *Ephedra* requires arid conditions and is found in subtropical and warm temperate regions (Rydin et al. 2010); *Gnetum* grows in tropical rainforests (Kato et al. 1995); *Welwitschia* is restricted to the Namib Desert in Africa (Carafa et al. 1992). Although *G. biloba* is found lining city streets throughout the world, its natural distribution in China is now limited (del Tredici 2007; Tang et al. 2012). Cycads are found in tropical and subtropical regions of

the Americas, Africa, Asia and Australia, although there are some temperate species occurring in Asia (Norstog and Nicholls 1997).

Gymnosperms range widely in plant form. Most conifers are trees, e.g. those that dominate the forests of northern latitudes. *Ginkgo* also grows into a large tree. Cycads range from tall palm-like forms, to small single-stalked plants, and even include one parasitic epiphyte (Norstog and Nicholls 1997). *Ephedra* species consist of shrubs and climbers (Rydin et al. 2010). *Gnetum* species grow mostly as lianas and sometimes as trees (Biye et al. 2014). *Welwitschia* has one of the most peculiar growth forms of all plants, only having two long, ribbon-like leaves (some greater than 2.5 m) that split as they grow along the ground (Henschel and Seely 2000).

How the extant seed-plant groups are related to one another has divided botanists for decades. Uncertainty exists at many points in this phylogeny. Many studies support gymnosperms as a monophyletic group (Chaw et al. 1997; Bowe et al. 2000; Chaw et al. 2000; Soltis et al. 2002; Xi et al. 2013), while some suggest that cycads and angiosperms form their own clade that is sister to the remaining gymnosperms (Mathews et al. 2010). Some papers place cycads and *Ginkgo* together as a clade sister to the other gymnosperms (Zhong et al. 2010; Wu et al. 2013; Xi et al. 2013). Others place *Ginkgo* alone as sister to a clade formed by only conifers and the Gnetales (Bowe et al. 2000; Chaw et al. 2000). The Gnetales have been placed in a number of different positions: as the sister group to the angiosperms (Crane 1985; Doyle and Donoghue 1986); as the sister group to conifers (Chaw et al. 1997); nested within the conifers (Bowe et al. 2000; Hajibabaei et al. 2006; Wu et al. 2013). Morphological and/or molecular data have been found to affirm each of these alternate possibilities.

It is clear from the literature that great amounts of time, effort and money have been poured into the dilemma of clarifying the extant seed-plant phylogeny, and that the problem is not trivial. Two of the most recent studies to look at seed-plant evolution followed two different approaches, and resulted in a somewhat similar arrangement of the living groups. Lee et al. (2011) used a functional phylogenomic approach that incorporated 22 833 sets of orthologs from the nuclear genomes of 101 species of seed plants. In their analysis, angiosperms are sister to a clade of gymnosperms. Within the gymnosperms, the Gnetales are sister to all other gymnosperms, and cycads plus *Ginkgo* form their own clade sister to conifers. Xi et al. (2013) used a coalescent-based species tree estimation phylogenomic method that incorporated both genome-scale nuclear data and plastid data for 14 species representing the five seed-plant groups. Their analysis also suggested that angiosperms are sister to a monophyletic gymnosperm group, and that *Ginkgo* plus cycads form a clade. However, in their analysis *Ginkgo* plus cycads are sister to conifers plus Gnetales. Gnetales are nested within the conifers, either sister to pines (nuclear data) or sister to Cupressaceae (plastid data). The conclusion of both Lee et al. (2011) and Xi et al. (2013) was that extant gymnosperms comprise a monophyletic group that is sister to the angiosperms, and that *Ginkgo* plus cycads form their own clade. Their interpretations only differ from each other by the placement of the Gnetales - either nested within the conifers or sister to all other gymnosperms. The Gnetales are persistently considered difficult to place (Mathews 2009; Rydin et al. 2010; Zhong et al. 2010).

An additional consideration must be made for the many extinct gymnosperm lineages that are thought to have once existed. Fossils may provide an understanding of

stem lineages between crown groups, as well as extinct groups branching from them (Doyle 2012), whereas molecular studies can only include crown groups. Placing fossil groups amongst extant groups is not an easy task.

Pollination drops

Pollination drops are small secretions of liquid, typically between 10 - 1000 nL (Prior et al. 2013), that are exuded from the micropyle of ovules around the time of pollination. The primary function of pollination drops is to capture and transport pollen from the environment to the inside of the ovule (Gelbart and von Aderkas 2002). Pollination drops occur in representative genera of all living gymnosperm groups (Singh 1978). These secretions are thought to be of ancient origin and were likely present in early seed plants (Doyle 1945; Tomlinson 2012; Little et al. 2014). A recent phylogenetic analysis, which included both extinct and extant gymnosperm lineages, suggested that pollination drops were probably present in many extinct gymnosperm lineages (Little et al. 2014).

Historical overview

Observations of pollination drops have appeared in the literature since the mid-nineteenth century. Vaucher (1841) made the first published observations of pollination drops in conifers. Delpino (1868) and Strasburger (1871) later provided the first detailed descriptions. Observations of pollination drops in the other gymnosperm groups appeared soon after. Karsten first reported pollination drops in *Gnetum* in 1892 (Zeigler 1959); Hirase first described pollination drops of *Ginkgo* in 1896 (Tison 1911). The first observations of pollination drops in cycads were made by Weber and Ikeno in 1897 and

1898, respectively, and were mentioned in Tison (1911). Porsch (1910) first reported pollination drops in *Ephedra*.

Early researchers were curious about the content of pollination drops. In 1902, Schumann published the first biochemical description of pollination drops (Fujii 1903). He reported that *Taxus* pollination drops contained a carbohydrate likely originating from plant mucilage, and free acids such as malic acid. He did not detect simple sugars. These results were based on simple tests using Fehling's solution and litmus paper (Fujii 1903). In contrast to Schumann's results, Fujii (1903) detected glucose and sucrose, as well as calcium, malic acid and formic acid in *Taxus baccata* L. pollination drops. Fujii (1903) also believed latex was present. He identified the presence of a strong reducing agent and speculated that it may be important to the physiology of the drop.

Tison (1911) emphasized the importance of understanding the timing and function of pollination drop secretion. He explained that pollination drops form within the micropyle, overflowing to the exterior of the ovule as spheres of liquid. He found that the timing of drop formation was consistent from year to year, but showed phenological differences between species. Tison also discussed functional aspects of pollination drops. He added pollen to drops, and described the swelling of the pollen intine. He gave descriptions of two pollination mechanisms, one in which pollen sinks into a drop, and the other in which captured pollen must wait for drop retraction before reaching the ovule interior.

Pollination drops as part of pollination mechanisms

Following the work of Tison (1911), many studies focused on pollination drops as part of pollination mechanisms (Owens et al. 1980; Owens and Molder 1980; Tomlinson

1991; Tomlinson et al. 1991). Pollination mechanisms describe the biology of pollen capture and transport. Both the orientation of the ovule and the morphology of pollen play important roles. Some taxa have ovules that are oriented upwards-to-horizontal at the time of pollination. Non-saccate pollen or wettable pollen falls into the pollination drop where it sinks into the ovule, or is drawn in through the micropyle as the drop withdraws. A number of conifers have this pollination mechanism e.g. Cupressaceae and Taxaceae (Gelbart and von Aderkas 2002). The non-coniferous gymnosperms *Ginkgo* (del Tredici 2007), cycads (Singh 1978) and the gnetalean genera *Welwitschia* (Carafa et al. 1992), *Ephedra* (Singh 1978) and *Gnetum* (Kato et al. 1995) also have this pollination mechanism. In other taxa, the micropyle is oriented downwards. The pollen grains of these species typically have sacci, which increase their buoyancy. The pollen grains float up the fluid-filled micropyle to the nucellus (Leslie 2010). Variations of this mechanism occur. For example, some species have micropylar arms covered with tiny, sticky secretions called microdrops that capture pollen. A pollination drop is then secreted that sweeps off loosely attached pollen, e.g. some Pinaceae (Owens et al. 1998). Other species capture pollen on micropylar hairs that extend outwards from the ovular entrance. They fold inwards, pushing the pollen into the ovule. Weeks later these species produce a pollination drop that carries pollen to the nucellus. This occurs in *Pseudotsuga* (von Aderkas and Leary 1999) and *Larix* (Said et al. 1991). Some conifers do not produce a pollination drop at any time, e.g. all species of Araucariaceae, some species of *Tsuga* and *Abies* (Gelbart and von Aderkas 2002).

Drop secretion and retraction

The tissue origin of pollination drops and their solutes has been debated since the earliest pollination drop studies. Schumann believed pollination drops were secreted by cells around the rim of the micropyle (Fujii 1903). Fujii (1903) argued that these cells were covered in a hydrophobic cuticle and therefore could not secrete the pollination drop. Tison (1911) observed that nucellar cells had the characteristics of secretory cells; they were turgid and possessed dense cytoplasm. The nucellus is the sporophytic tissue that gives rise to a megaspore, which develops into the megagametophyte in the ovule. When Tison used dye to stain pollination drops, he found that cells of the nucellar tip became stained. The nucellus continues to be considered the most likely source of pollination drops (Singh 1978; Gelbart and von Aderkas 2002; Nepi et al. 2009). Water can be secreted from the nucellus because the upper surface either lacks a waxy cuticle or the cuticle separates from the epidermis during pollination drop production (Singh 1978). However, there have been few studies concerning the origin of the organic components of pollination drops. The existing evidence suggests that proteins originate from the nucellus. For example, O'Leary et al. (2004) showed that arabinogalactan proteins occurring in the pollination drop of *Taxus* immunolocalized to the nucellus.

The mechanism governing pollination drop secretion remains unexplained. McWilliam (1958) suggested guttation. Ziegler (1959) reported that metabolic inhibitors did not affect drop production, and therefore secretion was not a metabolic process. The deterioration of the nucellus during the formation of a pollen chamber may contribute to the pollination drop in some taxa (Figure 2; Gelbart and von Aderkas 2002). Active secretion may also be possible. Signal peptide domains that are associated with the export of proteins from cells were identified in proteins found in *Taxus* pollination drops

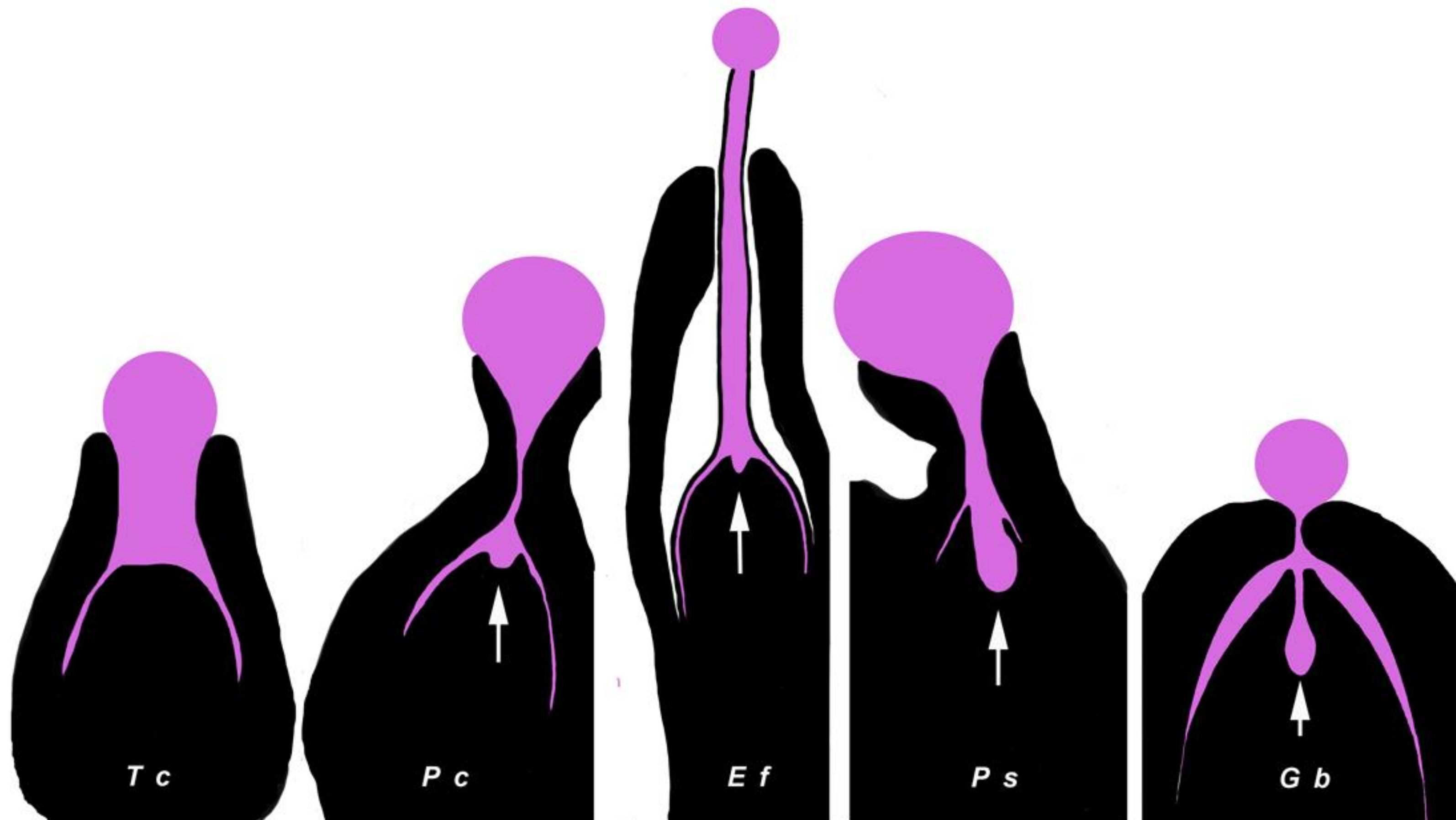

Figure 2. Schematic representations of gymnosperm ovules at the time of pollination drop secretion. *Taxus canadensis* Marsh (Tc) does not have a pollen chamber. Other gymnosperms vary in the depth of their pollen chambers (indicated by arrows) from small depressions in *Pinus contorta* Douglas ex. Louden (Pc) and *Ephedra foeminea* Forssk (Ef) to substantial chambers in *Picea sitchensis* (Bong.) Carr. (Ps) and *Ginkgo biloba* L. (Gb). Ovular silhouettes are modified from sections (abbreviated species in brackets) published in Dupler 1920 (Tc), Owens et al. 2005 (Pc), Rydin et al. 2010 (Ef), Owens and Blake 1984 (Ps) and Douglas et al. 2007 (Gb).

(O'Leary et al. 2007). Whether sugars or other solutes are actively pumped into the apoplast of the nucellus to influence the movement of water into the ovular entrance has not been investigated.

In some species, there is an approximate diurnal rhythm in the secretion and retraction of pollination drops, e.g. cycads (Tang 1987) and *Gnetum* (Kato et al. 1995). In other species, the rhythm of secretion and retraction varies over the course of receptivity to pollen. For example, Owens et al. (1980) reported a diurnal rhythm for *Chamaecyparis nootkatensis* D. Don at the beginning and end of the drop season, but reported a steady drop throughout the mid-season. In some species, there does not appear to be a rhythm, e.g. *Larix* (O'Leary and von Aderkas 2006). To support this observation, O'Leary and von Aderkas (2006) showed that there was no relation between diurnal fluctuations in xylem water potential of the tree and pollination drop secretion in the ovules of *Larix x marschlinsii* Coaz. Their analysis suggested there is an ovule-level regulation to drop secretion.

Some taxa rely on the retraction of pollination drops to move pollen to the nucellus for germination (Singh 1978). In Podocarpaceae, pollination drop secretion and retraction are part of a *pollen scavenging* mechanism. Here, the pollination drop is secreted onto a wettable surface around the ovule, where it collects pollen. The drop is then withdrawn, bringing pollen to the nucellar surface (Tomlinson et al. 1991).

Like the mechanism governing secretion, the mechanism controlling pollination drop withdrawal is also unknown. Pollination drops may be withdrawn within minutes, e.g. in pines (Doyle and O'Leary 1935). In some taxa, pollination drop retraction is stimulated by pollen capture, e.g. *Juniperus* (Mugnaini et al. 2007) and *Ginkgo* (Jin et al.

2012). Mugnaini et al. (2007) suggested that a recognition mechanism may exist between pollen grains and the ovule that causes the pollination drop to retract upon receipt of pollen. Jin et al. (2012) speculated that the presence of pollen alters the balance between active drop secretion and evaporation, causing the pollination drop to withdraw.

Drop constituents and their possible functions

The function of pollination drops likely goes beyond simply capturing and transporting pollen to the nucellus. The water contained in pollination drops plays an important role in hydrating pollen grains. Pollen is also directly exposed to the minerals and organic molecules contained within pollination drops. Since the early studies of Schuman, Tison and Fujii, additional solutes have been identified in pollination drops. Given the biochemical complexity of pollination drops, additional functions beyond pollen capture and delivery seem probable.

Pollination drops serve as the medium for pollen germination in most extant gymnosperms. Pollen germination can occur soon after pollination. In cycads, pollen germinates in the residual droplet contained in the pollen chamber, an area of degraded nucellar cells (Choi and Friedman 1991). In *Ephedra*, pollen germinates in the droplet within hours of pollen capture (Williams 2009, 2012; El-Ghazaly et al. 1998). In other cases, pollination and germination are separated by a number of weeks, yet the drop still acts as the trigger for germination. In both *Pseudotsuga* and *Larix*, pollen is captured and brought into the ovule by stigmatic hairs. Only weeks later, when the post-pollination pre-fertilization drop is released does pollen germination occur (Said et al 1991; Takaso and Owens 1996). *In vitro* studies of pollen germination support the idea that specific biochemical components are required for pollen germination (Brewbaker and Kwack

1963; Nygaard 1977; Dehoux and Pham Thi 1980). Many taxa have slower pollen tube growth rates *in vivo* versus *in vitro* (Williams 2012) which suggests that there are components in pollination drops that mediate or control germination.

Pollination drop sugars have potential roles in pollen germination. They could provide a source of energy for pollen. Monosaccharides are taken up and used by pollen during germination *in vitro* to support the growth of the pollen tube and the accumulation of polysaccharide storage molecules (Nygaard 1977). Sugars are present in the pollination drops of conifers (McWilliam 1958; Ziegler 1959), Cycadales (Tang 1987), *Gnetum* L. (Kato et al. 1995), *Welwitschia* (Carafa et al. 1992), *Ephedra* (Ziegler 1959; Bino et al. 1984a, b) and *Ginkgo* (Dr. Massimo Nepi, pers. comm.). Glucose, fructose and sucrose were identified in a number of conifers (McWilliam 1958; Ziegler 1959; Seridi-Benkaddour and Chesnoy 1988) and Cycadales (Tang 1993). Other sugars, such as mannitol (Mugnaini et al. 2007), galactose (Carafa et al. 1992), xylose and melezitose (von Aderkas et al. 2012) have also been identified. Sugars are also present in some conifers as polymers containing: arabinose, galactose, glucose, mannose, rhamnose (Seridi-Benkaddour and Chesnoy 1988). Total sugar concentrations vary between groups. For conifers, total sugar concentrations between 1- 2 % have been found (McWilliam 1958). Other gymnosperms have higher concentrations: 10 % for *Ephedra* (Ziegler 1959); 4-14 % for cycads (Tang 1993); 3-15 % for *Gnetum* (Kato et al. 1995; Nepi et al. 2009).

Total sugar concentration varies greatly between groups, and specific pollen types may have optimal osmotic conditions for germination, thus providing germination specificity in different osmotic environments. Sugar concentration and composition were

observed to be controlled by enzymes present in the pollination drop in some taxa. A functional extracellular invertase is present in *Pseudotsuga*, breaking down sucrose into glucose and fructose (von Aderkas et al. 2012) thus affecting proportions among these sugars.

Proteomic studies have revealed that pollination drops of conifers and *Welwitschia* contain a number of proteins. These include xylosidases, invertases, aspartyl proteases, peroxidases, serine-carboxypeptidases, galactosidases (Poulis et al. 2005), thaumatin-like proteins (Wagner et al. 2007; O'Leary et al. 2007), and chitinases (Poulis 2004; Wagner et al. 2007). Additional pollination drop proteins found in cupressaceous conifers include a glucanase-like protein, a glycosyl hydrolase, glucan 1,3-ß-glucosidases, a ß-D-glucan exohydrolase and subtilisin-like proteinases (Wagner et al. 2007). Several arabinogalactan proteins occur in *Taxu*s x *media* Rehder. These were discovered using immunohistology (O'Leary et al. 2004).

Pollination drop proteins likely play an active role in pollen germination. Like sugars, their presence may alter the osmotic environment of the drop (Wagner et al. 2007). If broken down to free amino acids, they may also provide a source of nutrients for germinating pollen by supplying key components for protein synthesis within pollen tubes as they grow (Zhang 1982). *In vitro*, externally supplied free amino acids have been observed to increase pollen tube growth and development (Dehoux and Pham Thi 1980). Proteases present in the pollination drop are the expected regulator of free amino acid concentrations (Poulis et al. 2005). Other active enzymes may also affect germination. Xylosidases and galactosidases could loosen the pollen cell wall by cleaving xyloglucans,

a group of hemicelluloses that support the cellulose microfibrils of the cell wall (Poulis et al. 2005). This would help prime the pollen wall for tube emergence.

Arabinogalactan proteins are abundant in the pollination drop of *T.* x *media* (O'Leary et al. 2004). Many possible roles are known for arabinogalactan proteins during reproduction in plants (reviewed by Nguema-Ona et al. 2013). These may be simple, such as to provide growing pollen tubes with an adhesive surface (Kim et al. 2002) or for nutrition (Cheung et al. 2000). Arabinogalactan proteins may also function in chemotropism to guide pollen tube growth (Wu et al. 2000) or in determining compatibility between pollen grains and carpel tissues (Cruz-Garcia et al. 2005).

Mineral components are also present in pollination drops and are known to affect pollen germination and growth. Calcium has been found in *Taxus* (Fujii 1903), *Larix* and *Pseudotsuga* (von Aderkas et al. 2012). Addition of calcium to pollen germination media was a key discovery for development of culture methods (Brewbaker and Kwack 1963). *In vitro*, calcium is required for pollen tube elongation in Norway spruce (Lazarro et al. 2005). Calcium sustains pollen viability and increases the percentage of pollen grains producing pollen tubes in *Pseudotsuga* (Fernando et al. 1997). Calcium-regulating proteins have been identified in pollen grains of *Pinus yunnanensis* Franch. (Gong et al. 1993) and *Cryptomeria japonica* D. Don (Yokota et al. 2004), suggesting an active role for calcium during pollen germination in conifers.

Components of pollination drops may also provide protection from biotic and abiotic threats. Proteomic studies have identified putative defence proteins that may prevent fungal and bacterial infection of the ovule through the open micropyle. These include glucosidases and chitinases (Wagner et al. 2007), peroxidases (Poulis et al. 2005),

and thaumatin-like proteins (O'Leary et al. 2007). Thaumatin-like proteins may also function as antifreeze proteins (O'Leary et al. 2007) in conifers of colder climates. Phenols have been detected in cycads, and may provide protection from bacteria or fungi (Tang 1987). Additionally, at higher concentrations, sugars may provide an osmotic environment that inhibits microbial growth (Little et al. 2014).

Pollination drops may also function in plant-insect interactions by attracting insect pollinators. The cycads and the Gnetales include species that rely on insects for pollination. The sugars and amino acids found in pollination drops could be sources of nutrition for insects. However, whether or not insects actively use pollination drops as a food reward is unclear. No *experimental* studies have directly tested the interactions between insects and pollination drops. Observations have been made of insects feeding on the pollination drops of *Gnetum* spp. (Kato et al. 1995) and *Welwitschia* (Wetschnig and Depisch 1999). *Ephedra* has a relatively high sugar content that is thought to attract insects in at least some species (Bino et al. 1984a,b). Terry et al. (2005) observed thrips moving towards the micropyle of the cycad *Macrozamia*, possibly to visit the pollination drop. However, whether the weevils or thrips that pollinate cycads use pollination drops as food is unknown (Tang 1987; Terry 2001).

Current knowledge of reproduction in angiosperms compared to gymnosperms

In general, less is known about the biochemistry of reproduction in the four groups of extant gymnosperms when compared to angiosperms. A key contributing factor is that most crop plants are angiosperms, and research is driven by the desire for crop improvement. The chosen model plant species, e.g. *Arabidopsis*, poplar, rice and maize, are all flowering plants. Although selected for specific reasons, such as rapid life cycle

and small genome size in the case of *Arabidopsis*, these choices neglect the evolutionary importance of the remaining seed-plant groups.

The processes of pollination and fertilization are different between angiosperms and gymnosperms. In angiosperms, pollen sticks to the stigma, germinates, and the pollen tube must grow through the style and into the ovary (Dumas and Rogowsky 2008). There, the pollen tube releases two male nuclei (gametes). One fertilizes the egg cell leading to an embryo. The other nucleus fuses with the central cell which then develops into endosperm, a nutritive tissue for the developing embryo. This is known as double-fertilization (Dumas and Rogowsky 2008). The situation is different with gymnosperms. The ovule is exposed at pollination. The pollen has a short journey; the pollen tube only grows through the nucellus and neck cells before encountering the egg. There is no true double-fertilization in gymnosperms. In place of endosperm, gymnosperms have a large amount of megagametophyte tissue that nourishes the developing embryo (Singh 1978). The processes of pollination and fertilization usually take place over hours or days in angiosperms. With the exception of *Ephedra,* the period from pollination to fertilization is usually much longer in gymnosperms, ranging from a week to over a year (Willson and Burley 1983).

Much more is understood about the processes of pollination and fertilization in angiosperms as opposed to gymnosperms. The interaction between the pollen grain and stigma (Hiscock and Allen 2008), the growth and guidance of the pollen tube in the style (Chae and Lord 2011), and the release and fusion of gametes (Chevalier et al. 2011) are all being studied at the molecular level in angiosperms (Dumas and Rogowsky 2008). Angiosperms generally have shorter lifecycles than gymnosperms, making them more

amenable to lab experiments. Model angiosperm systems, such as *Arabidopsis thaliana* (L.) Heynh. (Dumas and Rogowsky 2008) and *Zea mays* L. (Dresselhaus et al. 2011), allow detailed study of reproduction to the level of the gene (Okamoto and Kranz 2005). The roles of specific proteins and peptides in reproductive processes are also being elucidated (Chae and Lord 2011; Miernyk et al. 2011). It is clear from the sheer volume of articles available that many aspects of angiosperm reproduction are at an advanced stage of study. This is not yet the case with gymnosperm reproduction.

Proteomics of plant reproduction

The proteome is a description of all proteins present in a given sample, whether the sample is a particular species, tissue, cell or sub-cellular fraction. Proteomic analyses have been used to study reproduction in angiosperms. The majority of articles relate to seed development and germination, the male gametophyte, or flowers. Fewer articles investigate the female gametophyte, post-fertilization events, incompatibility, seed abortion and apomixis (reviewed by Miernyk et al. 2011). Although Miernyk et al. (2011) stated that relatively few proteomic studies have focused on angiosperm reproduction, citing about 70 articles, even fewer studies have focused on the proteomics of gymnosperm reproduction. A survey using Web of Science (April 8 2014) returned just fourteen articles relating to reproduction in gymnosperm genera and proteomics. These articles focused on seed development (Shi et al. 2010; Zhen and Shi 2011; Zhen et al. 2012), somatic embryogenesis (Lippert et al. 2005; Jo et al. 2014; Teyssier et al. 2014), embryo development (Balbeuena et al. 2011), pollen (Fernando 2005; Wu et al. 2008; Sheng et al. 2011) and pollination drops (Poulis et al. 2005; O'Leary et al. 2007; Wagner

et al. 2007). All but one of these articles is focused on conifers; the other gymnosperm groups have been generally ignored.

Pollination drops are well-suited to proteomic analysis (Prior et al. 2013). Their composition is not as complex as animal serum, and they do not contain over-abundant plant proteins, e.g. Rubisco, that could potentially swamp out the signal of other proteins (Miernyk et al. 2011). Pollination drops do not require complex preparation steps prior to analysis. Proteomics techniques were used to explore the pollination drop proteomes of *Pseudotsuga menziesii* Mirb. Franco (Poulis et al. 2005), *T.* x *media* (O'Leary et al. 2007), *Juniperus communis* L., *Juniperus oxycedrus* L., *Chamaecyparis lawsoniana* (A. Murray) Parl. and *W. mirabilis* (Wagner et al. 2007). Depending on the species, pollination drop samples contained few to hundreds of proteins. However, only some of the proteins present could be identified. The lack of gymnosperm sequence data available for protein identification was a limiting factor of these proteomic analyses.

Proteomic analyses of gymnosperm pollination drops

Here we present the proteomic analyses of pollination drops from representatives of all extant gymnosperm lineages, including all three Gnetalean genera (*Ephedra*, *Welwitschia* and *Gnetum*), three genera of cycads (*Zamia*, *Ceratozamia* and *Cycas*), *Ginkgo biloba* and one genus of conifer (*Taxus*). Previous proteomic analyses of pollination drops have been limited by the scarcity of gymnosperm sequence data available for use in protein identification. We used a custom protein database derived from gymnosperm transcriptomes to analyze our tandem mass spectrometry data.

Protein identifications may indicate possible roles for pollination drops beyond the capture and transport of pollen, e.g. chitinase (Coulter et al. 2012) and invertase (von

Aderkas et al. 2012) were recently shown to be functional in Douglas-fir pollination drops. Proteomics techniques allowed assessment of the complexity of the pollination drop proteomes of the extant gymnosperms and prediction of putative functions for the proteins contained within their pollination drops. Proteomics techniques were used to address the question of whether the proteins contained in gymnosperm pollination drops have conserved functions between the four extant gymnosperm lineages.

Chapter 2: Degradome and secretome of pollination drops of *Ephedra*

The following chapter is an excerpt from the paper “Degradome and secretome of pollination drops of *Ephedra*” that was accepted for publication by Botanical Review in August 2014. Co-Authors: Patrick von Aderkas[1], Natalie Prior[1], Susannah Jesse[1], Stefan Little[1,2], Tyra Cross[3], Darryl Hardie[3], Christoph Borchers[3], Robert Thornburg[4], Chen Hou[5], Alexandra Lunny[1]

[1] Centre for Forest Biology, Department of Biology, University of Victoria, Victoria BC V8W 3N2, Canada

[2] Department of Plant Sciences, Mail Stop 1,University of California, Davis, One Shields Avenue, Davis, CA 95616, USA

[3] UVic – Genome BC Proteomics Centre, University of Victoria, Victoria BC V8W 3N2, Canada

[4] Department of Biochemistry, Biophysics & Molecular Biology, Iowa State University, Ames, IA 50011, USA

[5] Department of Ecology, Environment and Plant Science, Stockholm University, SE-106 91 Stockholm, Sweden.

Introduction

In *Ephedra,* pollination drops (Figure 3) function in both pollen capture and delivery (Endress 1996). Pollen can be delivered by wind or by insects (Figure 4), but in the latter case, pollination drops also function as a nectar/reward for the pollinator (Moussel et al. 1980; Meeuse et al. 1990). A variety of insects have been recorded from *Ephedra* spp., including dipterans, as well as hymenopterans such as vespids, braconids and chalcids, but not bees (reviewed in Bino et al. 1984a,b). *Ephedra* species are not

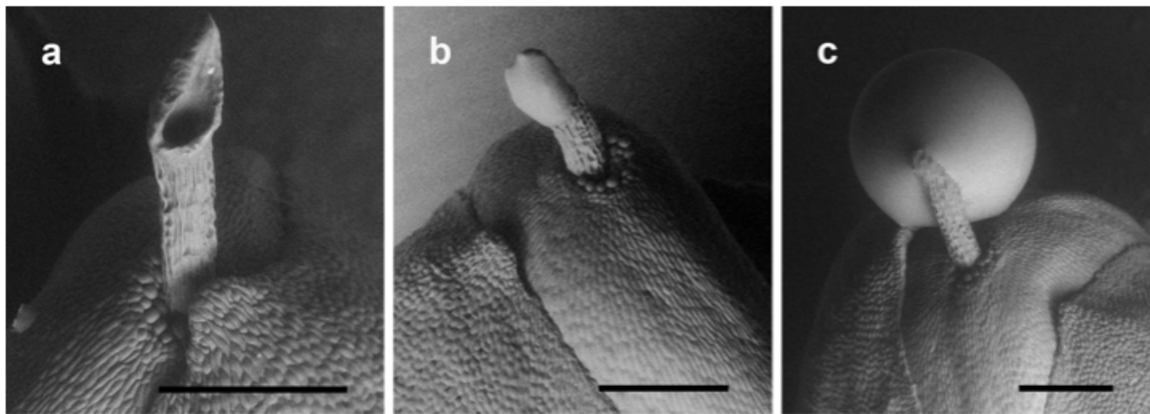

Figure 3. Scanning electron micrographs of *Ephedra monosperma* ovules. a. An open micropyle. b. A pollination drop partially exuded from the micropyle. c. A pollination drop fully exuded from the micropyle. Bar = 500 µm

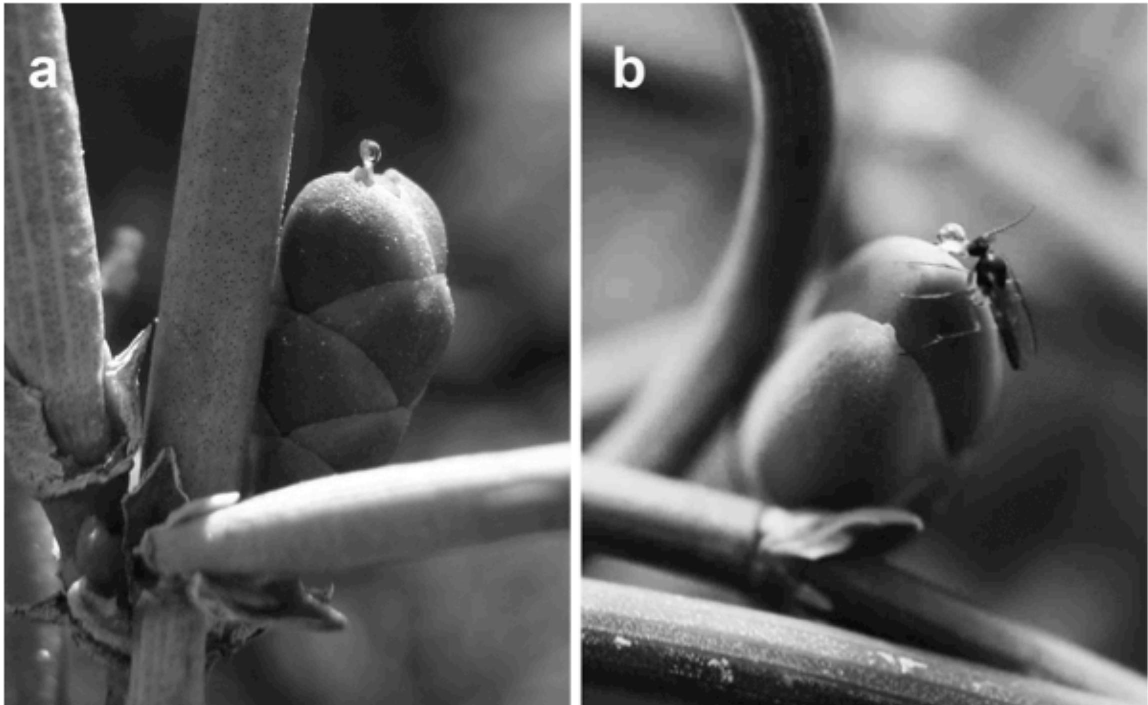

Figure 4. *Ephedra* ovules. a. Ovule of *E. compacta* with pollination drop. b. *E. monosperma* with an insect feeding on the pollination drop.

obligately insect-pollinated, as wind pollination may also occur at the same time. In this respect, *Ephedra* is similar to other gnetophytes (*Welwitschia* and *Gnetum*) (Endress 1996).

In *Ephedra*, the pollination mechanism is relatively simple: non-saccate pollen is captured by a secreted pollination drop that subsequently recedes. *Ephedra* produces a relatively large drop (~ 1 µl). *Ephedra* pollen germinates rapidly and the pollen tube grows quickly, reaching the egg in 14 hours, which is much faster than with other gymnosperms (El-Ghazaly et al. 1998; Williams 2012). The pollen can even germinate while in the pollination drop outside the micropyle (Bino et al. 1984b). It would appear that the tubes do not have to be in close proximity of the nucellus to be able to grow long distances. The pollination drop with its carbohydrate and other substances is able to support long distance growth of these tubes (Bino et al. 1984b).

Ephedra pollination drops contain abundant sucrose, but are also abundant in phosphate compounds, amino acids, and polypeptides (Ziegler 1959). Ziegler (1959) showed that calcium is also present in *Ephedra*. Until this study, no proteins had been documented, although Ziegler (1959) found acid phosphatase activity in the nucellus, the sporogenous tissue that produces the pollination drop. He wrote that such nucellar proteins likely are responsible for processing cellular compounds that are secreted into the drop. We hypothesize that *Ephedra* pollination drops contain proteins, given that pollination drop proteomes have been previously described in a number of conifers and the gnetalean *Welwitschia mirabilis* (Wagner et al. 2007). To this end, we embarked on the first proteomic study of *Ephedra* pollination drops. The aim was to test for the

presence of proteins, and if present, to understand the variation in protein composition in the pollination drops of *Ephedra*.

What is different, though not unique, about the pollination drop in *Ephedra* compared to that of most conifers studied is that drop production co-occurs with nucellus tissue breakdown. A central apical portion degenerates to form a pollen chamber (Rydin et al. 2010). Pollen chambers are known from the earliest fossils of Gnetales (Rothwell and Stockey 2013). This cell degradation forms the pollen chamber where captured pollen sinks prior to germination (Moussel 1980). Pollen chambers are found in *Ephedra* (Rydin et al. 2010) and some other gymnosperms, such as cycads (Norstog and Nicholls 1997), *Gingko* (Douglas et al. 2007), *Pinus* and *Picea* (Singh 1978). In comparison, many gymnosperms do not have pollen chambers. *Taxus* has an intact nucellus, i.e. a solid dome of parenchymatous tissue that shows no sign of degeneration before or during pollination drop formation (O'Leary et al. 2004). Since *Taxus* pollination drops have proteins secreted from intact cells, it follows that ovules with cell degradation-derived pollen chambers, such as those of *Ephedra, Ginkgo* and *Pinus*, may have drops that contain proteins of two origins: those secreted from intact cells, and those released by cell lysis. The portion of proteins that originate from the degraded tissues are appropriately called the degradome.

A degradome can arise from a number of processes occurring concurrently or independently. One source of degradome already considered above is cellular debris due to senescence during pollen chamber formation (Roberts et al. 2012). A second source may be from the activity of extracellular proteases and peptidases, if present in pollination drops, that would generate breakdown products from extracellular proteins. If

this occurs then both these peptidases and proteases would be detected along with polypeptide fragments of other proteins. Degradomes may form biochemically complex networks, but these remain relatively unstudied in plants (Huesgen and Overall 2012). Some of the breakdown products may function in providing signals that regulate defence responses of living cells. Proteomics provides identification with high confidence, but proof of functionality of constituents of the degradome within the pollination drop requires further study of substrate processing. Furthermore, it must be shown that these compounds are functional *in situ*.

Here we present the results from proteomic analysis of seven species of *Ephedra*. We not only hypothesize the presence of proteins in *Ephedra* pollination drops, but we also expect that such degenerative processes in *Ephedra* at the time of pollination drop formation would influence the type of proteins present, such as protein breakdown products that accompany tissue death.

Materials and methods

Sample collection

Ephedra pollination drop samples were collected by touching the drops with a micropipette tip. Drops were expelled into an Eppendorf tube and stored at -20°C until analysis. *Ephedra likiangensis* Florin and *Ephedra minuta* Florin drops were collected by Kristina Bolinder from plants in the botanical greenhouse at Stockholm University from January 17 through February 16, 2012 and December 21 through January 10, 2012 respectively. *Ephedra foeminea* Forssk. drops were collected by Anders Ryberg in Asprovalta, Greece in July 2011. *Ephedra distachya* L. drops were collected by Kristina Bolinder in Nea Vrasna, Greece May 30 and June 2, 2011. *Ephedra trifurca* Torr. ex S.Wats. drops were collected by Dr. Steffi Ickert-Bond at the Aqua Fria River Bottom,

Maricopa County, Arizona, U.S.A. on March 17, 2012. *Ephedra monosperma* C.A.Meyer drops were collected by Dr. Stefan Little from March to April, 2011 from greenhouse-grown plants at the Orchard Park Facility, University of California at Davis. *Ephedra compacta* Rose drops were collected by Israel Loera Carrizales in Laguna de Alchichica, Puebla, Mexico from April 10 to 23, 2012. In addition, samples of *Ginkgo biloba* and *Larix* x *marschlinsii* were collected from trees growing outdoors on the campuses of University of California at Davis and University of Victoria, respectively. A separate comparative study was carried out on pollination drops of *E. monosperma* collected on three sample dates, March 9, 24 and April 10, 2011.

1D SDS PAGE

Pollination drop sample (20 µL) was mixed with 5 µL NuPage MES SDS Buffer (Life Technologies Inc., Burlington, ON) and 1 µL of 1M DDT. Samples were boiled at 99 °C for 10 min, and then loaded on to a NuPage Novex 4 – 12 % Bis-Tris precast gel. Five µL of BLUeye Prestained Protein Ladder (FroggaBio Inc., Toronto, ON) were run alongside the samples. The gel was fixed with a 40 % ethanol / 10 % acetic acid solution for 10 min, and then stained with 0.1 % G250 Coomassie Brilliant Blue overnight. The gel was then destained with 10 % acetic acid solution.

LC-MS/MS analysis

Samples were reduced with dithiothreitol (30 min at 37 °C), and cysteine sulfhydryls were alkylated with iodoacetamide (30 min at 37 °C in darkness). Trypsin (2 µg; Promega) was added to each sample, which was digested at 37 °C for 16 hr. The samples were de-salted on a Waters HLB Oasis column (Waters Corporation, Milford, MA), Speed Vac-concentrated and then stored at -80 °C prior to LC-MS analysis.

Peptide mixtures were rehydrated to 100 µL with 2 % acetonitrile/water/2 % formic acid and separated by on-line reversed phase chromatography using a Thermo Scientific EASY-nLC II system (Thermo Fisher Scientific, Bremen, Germany) with a reversed-phase pre-column Magic C-18AQ (100 µm internal diameter, 2 cm length, 5 µm, 100 Å, Michrom BioResources Inc, Auburn, CA) pre-column and a reversed phase nano-analytical column Magic C-18AQ (75 µm internal diameter, 15 cm length, 5 µm, 100 Å, Michrom BioResources Inc, Auburn, CA) both in-house prepared, at a flow rate of 300 nl/min. The chromatography system was coupled to an LTQ Orbitrap Velos mass spectrometer equipped with a Nanospray II source (Thermo Fisher Scientific). Solvents were A: 2 % acetonitrile, 0.1 % formic acid; B: 90 % acetonitrile, 0.1 % formic acid. After a 249 bar (~ 5 µL) pre-column equilibration and 249 bar (~ 8 µL) nanocolumn equilibration, samples were separated by a 90 min gradient (0 min: 5 % B; 80 min: 45 % B; 2 min: 90 % B; 8 min: 90 % B).

Data analysis parameters

Raw LCMS files were converted to Mascot Generic Format and processed with PEAKS Client 6 (Bioinformatics Sofware Inc, Waterloo, ON, Canada) with Peaks DB and Spider searches enabled against the Uniprot/Trembl and Uniprot/Swiss-Prot Allspecies taxonomy databases. Only plant species were selected. Settings were as follows: instrument type set as FT-ICR/Orbitrap; high energy CID as fragmentation mode; parent ion error tolerance 8 ppm; fragment ion error tolerance 0.03 Da; trypsin as enzyme; up to one missed cleavage allowed; carbamidomethylation as a fixed modification; deamidation and oxidation as variable modifications. Peptide spectrum match false discovery rate (FDR), peptide FDR and protein FDR all set to $\leq$ 1 %. The

quality of the spectra were verified for proteins that were identified by only a single peptide sequence.

Scanning electron microscopy

Ephedra monosperma ovules were collected from the Bev Glover Greenhouse, University of Victoria. Ovules were removed from branches and mounted on a Deben MK3 cold stage (Deben UK Ltd., Woolpit, UK) in a Hitachi S-3500N variable pressure scanning electron microscope (VP SEM) (Hitachi High Technologies Canada Inc., Etobicoke, ON, Canada). The microscope was operated at 20 kV, 50 Pa variable pressure in backscattered electron mode using a Robinson BSE detector.

Results

Comparative study of seven *Ephedra* species

All *Ephedra* pollination drops contained proteins as detected by SDS-PAGE and staining of protein bands (Figure 5). The relatively light bands of *Ephedra* proteins run at native concentrations indicate lower amounts of protein, compared to that of larch and *Ginkgo* (Figure 6). Proteins identified from liquid extractions of pollination drops can be separated into degradome and secretome proteins (Tables 1, 2). We did not include proteins that had good spectra that matched uncharacterized proteins, *e.g.* inferred proteins from *Picea sitchensis* (Bong.) Carr. cDNA, although these could be as many as a third of the high quality identities for any one species, e.g. *E. foeminea* pollination drops contained 29 proteins, of which only 20 were characterized.

The number of characterized proteins in pollination drops of *Ephedra* species ranged from 6 to 20, averaging 13.4 $\pm$ 5.3 identified proteins/species (Table 3). *Ephedra foeminea* and *E. trifurca* contained more proteins (20), compared to *E. distachya* (15),

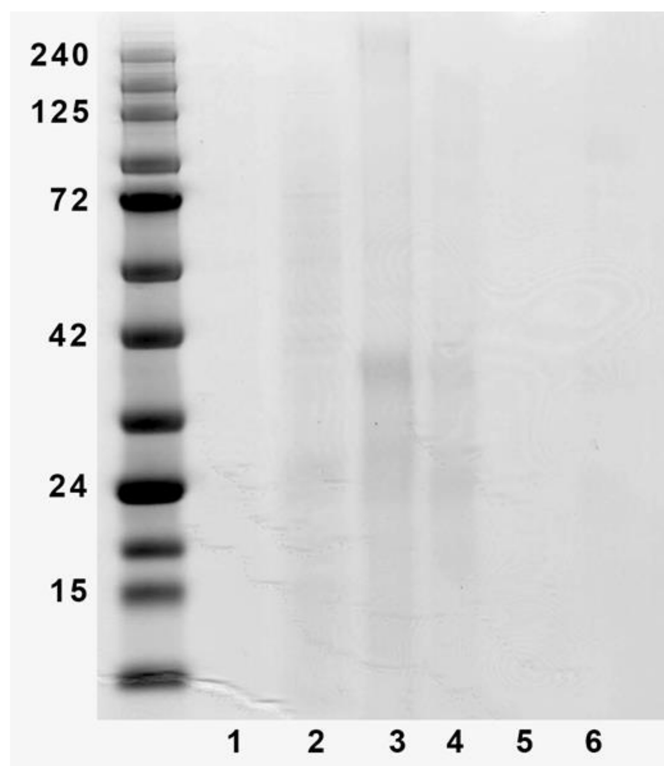

Figure 5. 1D SDS-PAGE of proteins at native concentrations in *Ephedra* pollination drops. Lanes from left to right: molecular weight ladder (kDa), 1. *E. distachya*, 2. *E. distachya*, 3. *E. foeminea*, 4. *E. minuta*, 5. *E. likiangensis*, 6. *E. monosperma*. Proteins were stained using Coomassie Brilliant Blue G-250.

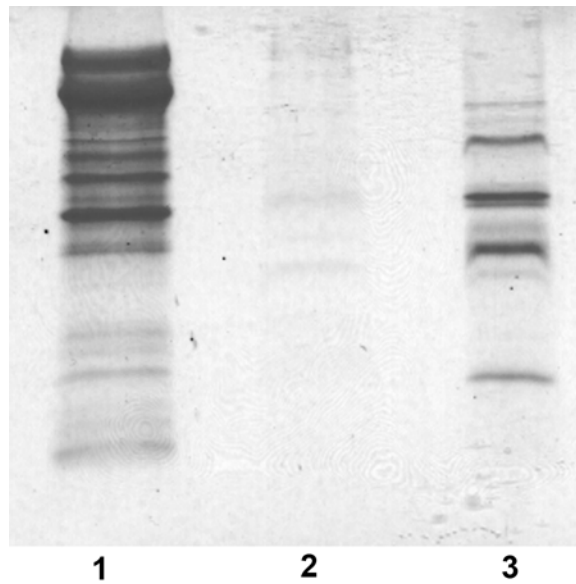

Figure 6. 1D SDS-PAGE of native concentrations of proteins in pollination drops of three gymnosperms: Lane 1. *Larix* x *marschlinsii*, Lane 2. *E. monosperma*, Lane 3. *Gingko biloba.* Figure is only to show number of bands and relative band intensity. Proteins were stained using Coomassie Brilliant Blue G-250.

Table 1. Degradome proteins found in pollination drops of *Ephedra* spp.

Protein	Species
Ubiquitins	*E. compacta* *E. foeminea* *E. likiangensis* *E. minuta* *E. trifurca*
Dessication-related protein	*E. compacta* *E. likiangensis* *E. minuta* *E. monosperma*
Cyclophilin A	*E. distachya* *E. foeminea* *E. minuta*
Elongation factor 1- α	*E. distachya* *E. foeminea* *E. trifurca*
Histones	*E. distachya* *E. foeminea* *E. trifurca*
Acyl-CoA-binding domain-containing protein 6	*E. compacta* *E. trifurca*
α -Amylase	*E. compacta* *E. likiangensis*
Calmodulin	*E. compacta* *E. distachya*
Glycosyl hydrolase	*E. foeminea* *E. trifurca*
GTP-binding nuclear protein	*E. distachya* *E. monosperma*
α -Amylase inhibitor	*E. trifurca*
Auxin response factor	*E. distachya*
Calreticulin	*E. foeminea*
Ceramidase	*E. monosperma*
Citrate synthase	*E. foeminea*
Cysteine proteinase	*E. likiangensis*
α -Gliadin	*E. trifurca*
Glycerophosphoryl diester phosphodiesterase	*E. foeminea*
Granule-bound starch synthase	*E. foeminea*
Heat shock proteins	*E. distachya*
Lactoylglutathione lyase	*E. trifurca*
Luminal-binding protein	*E. foeminea*
Profilin	*E. monosperma*
Thiol protease aleurain	*E. likiangensis*

Table 2. Secretome proteins found in pollination drops of *Ephedra* spp. Proteins that could also be considered degradome are marked with an asterisk.

Protein	Species
Xylosidases	*E. compacta* *E. foeminea* *E. minuta* *E. trifurca*
Aspartic proteinase*	*E. compacta* *E. likiangensis* *E. trifurca*
Galactosidases	*E. compacta* *E. minuta* *E. trifurca*
Peroxidase	*E. compacta* *E. likiangensis* *E. trifurca*
Serine carboxypeptidases*	*E. foeminea* *E. monosperma* *E. trifurca*
Chitinase	*E. foeminea* *E. trifurca*
Glucan endo-1,3- β -glucosidase	*E. monosperma* *E. trifurca*
Malate dehydrogenase	*E. trifurca*
Peptidase*	*E. likiangensis*
Superoxide dismutase*	*E. compacta*
Thaumatin-like protein	*E. minuta*

Table 3. Peptide sequences and identities of pollination drop proteins found in *Ephedra* spp. Degradome proteins are indicated by a black line in the right margin.

Species	Peptide amino acid sequence obtained	Protein identification
E. compacta	K.SSEEAME(sub N)DYITK.V M.GLKEEFEEY(sub H)AEK.V R.AKWDAWK.A	Acyl-CoA-binding domain-containing protein 6 OS=*Arabidopsis thaliana*
	K.EGIPPVQQR.L R.TLADYNIQK.E E.VESSDTIDNVK.A	Ubiquitin-NEDD8-like protein RUB2 OS=*Oryza sativa* subsp. *japonica*
	R.TLADYNIQK.E K.EGIPPVQQR.L	Polyubiquitin 2 OS=*Zea mays*
	R.TLADYNIQK.E E.VESSN(+.98)TIDNVK.A	Putative polyubiquitin (Fragment) OS=*Arabidopsis thaliana*
	R.NIQVVDGSNNLKAPK.G	Putative carboxyl-terminal peptidase OS=*Arabidopsis thaliana*
	R.VFDKDQNGFISAAELR.H	Calmodulin (Fragment) OS=*Pyrus communis*
	K.AVADIVINHR.C	Alpha amylase (Fragment) OS=*Cuscuta reflexa*
	L.GVESGQDAVIR.G R.TPEEILR.I	Dessication-related protein_ putative; 70055-71849 OS=*Arabidopsis thaliana*
	K.VTEQDLE(sub A)DTYNPPFK.S	Putative beta-xylosidase (Fragment) OS=*Triticum aestivum*
	R.STPEMWPDIIQK.A	Beta-galactosidase OS=*Picea sitchensis*
	R.AVVVHADPDDLGK.G	Superoxide dismutase [Cu-Zn] OS=*Pinus sylvestris*
	K.GEHTYVPVTK.K	Aspartic proteinase (Fragment) OS=*Cucumis sativus*
	R.FDNNYYK.D	Peroxidase (Fragment) OS=*Lupinus polyphyllus*
E. distachya	K.ATAGDTHLGGEDFDNR.M R.IINEPTAAAIAYGLDKK.A R.VEIIPNDQGNR.T K.NKITITNDKGR.L	Heat shock 70 kDa protein OS=*Glycine max*
	K.ATAGDTHLGGEDFDNR.M R.IINEPTAAAIAYGLDKK.A R.VEIIANDQGNR.T K.NKITITNDKGR.L	Heat shock cognate 70 kDa protein 1 OS=*Solanum lycopersicum*
	R.ELISNSSDALDKIR.F K.ADLVNNLGTIAR.S D.AIDEYAIGQLK.E R.FESLTDK.S	Heat shock protein 81-2 OS=*Arabidopsis thaliana*

	K.IGGIGTVPVGR.V N.IVVIGHVDSGK.S R.VETGVIKPG.M	Elongation factor 1-alpha OS=*Zea mays*
	F.DKDQNGFISA.A MADQLTDDQISEFK.E	Calmodulin 4 (Fragment) OS=*Daucus carota*
	FDKDGDGC(+57.02)ITTK.E	Calmodulin protein (Fragment) OS=*Pinus taeda*
	R.DNIQGITKPAIR.R	Histone H4 OS=*Solanum melongena*
	L.FEDTNLC(+57.02)AIHAK.R	Histone H3-like 1 OS=*Arabidopsis thaliana*
	R.NVIHGSDAVESAQ(sub R)K.E	Nucleoside diphosphate kinase OS=*Arabidopsis lyrata* subsp. *lyrata*
	K.AGFAGDDAPR.A	Actin-3 OS=*Glycine max*
	R.GNGTGGESIYGEK.F	Peptidyl-prolyl cis-trans isomerase OS=*Zea mays*
	R.VLQISGER.N	18.1 kDa class I heat shock protein (Fragment) OS=*Medicago sativa*
	R.VLQISGER.S	Small heat shock protein hsp10.4 (Fragment) OS=*Quercus suber*
	K.LVIVGDGGTGK.T	GTP-binding nuclear protein Ran2 OS=*Solanum lycopersicum*
	K.LVIVGDGGTGK.T	Small Ran-related GTP-binding protein OS=*Triticum aestivum*
	R.TFVKVYK.S	Auxin response factor 12 OS=*Oryza sativa* subsp. *indica*
E. foeminea	K.EALQAEVGLPVDR.N K.VVGTPAYEEM(+15.99)VR.N R.FAFSDYPELNLPER.F K.SSFDFIDGYEKPVEGR.K K.MGDGYETVR.F R.VLTVSPYYAEELISGIAR.G	Granule-bound starch synthase 1_ chloroplastic/amyloplastic OS=*Zea mays*
	R.FAFSDYPELNLPER.F K.VVGTPAYEEM(+15.99)VR.N K.EALQAEVGLPVDR.N K.MGDGYETVR.F K.SSFDFIDGYEKPVEGR.K R.VLTVSPYYAEELISGIAR.G	Granule-bound starch synthase OS=*Zea mays* subsp. *mays*
	R.EAEEFAEEDKK.V K.FELSGIPPAPR.G R.VEIESLFDGVDFSEPLTR.A K.DYFDGKEPNK.G R.LSQEEIER.M	BiP isoform A OS=*Glycine max*
	K.EAEEFAEEDKK.V R.VEIESLFDGVDFSEPLTR.A K.DYFDGKEPNK.G	Luminal-binding protein 4 OS=*Nicotiana tabacum*

	R.LSQEEIER.M	
	K.TFASGILVPK.S	Probable glycerophosphoryl diester phosphodiesterase 3 OS=*Arabidopsis thaliana*
	DNIQGITKPAIR.R R.ISGLIYEETR.G	Histone H4 (Fragment) OS=*Daucus carota*
	R.DNIQGITKPAIR.R R.ISGLIYEETR.G	Histone H4 OS=*Silene latifolia*
	K.KPEGYDDIPK.E K.LDC(+57.02)GGGYVK.L	Calreticulin OS=*Zea mays*
	K.KPEGYDDIPK.E R.FEDGWDKR.W	Calreticulin OS=*Prunus armeniaca*
	R.EIAQDFK.T	Histone H3-like 1 OS=*Arabidopsis thaliana*
	R.TLADYNIQK.E	Polyubiquitin 9 OS=*Arabidopsis thaliana*
	R.ALGLPLERPK.S	Citrate synthase OS=*Picea sitchensis*
	R.ALGLPLERPK.S	Citrate synthase 5_ mitochondrial OS=*Arabidopsis thaliana*
	R.GNGTGGESIYGEK.F	Peptidyl-prolyl cis-trans isomerase OS=*Zea mays*
	R.IGGIGTVPVGR	Elongation factor 1-alpha (Fragments) OS=*Pseudotsuga menziesii*
	K.AGLQFPVGR.I	Histone H2A OS=*Euphorbia esula*
	V.VTQ(+.98)QDLDDTYQPPFK.S	Beta-xylosidase/alpha-L-arabinofuranosidase 2 OS=*Medicago varia*
	R.VWVYSGDTDGR.V	Serine carboxypeptidase-like 32 OS=*Arabidopsis thaliana*
	R.AINSM(+15.99)ECNGGNPSAVQ(sub D)DR.V	Class IV chitinase OS=*Nepenthes alata*
	R.C(+57.02)YESYSEDPS(sub K)IVK.A	Glycosyl hydrolase-like protein (Fragment) OS=*Picea sitchensis*
E. likiangensis	K.IQDKEGIPPDQQR.L E.VESSDTIDNVK.A R.TLADYNIQK.E	Ubiquitin OS=*Triticum aestivum*
	K.YNGGIDTEEA(sub S)YPYK.G R.EDGIVSPVK.N	Thiol protease aleurain OS=*Hordeum vulgare*
	L.GVESGQDAVIR.G R.TPEEILR.I	Dessication-related protein_ putative; 70055-71849 OS=*Arabidopsis thaliana*
	K.AVADIVINHR.C	Alpha amylase (Fragment) OS=*Cuscuta reflexa*
	R.FDNNYYK.D	Peroxidase (Fragment) OS=*Lupinus polyphyllus*
	R.NIQVVDGSNNLKAPK.G	Putative carboxyl-terminal

		peptidase OS=*Arabidopsis thaliana*
	A.Q(+.98)GSGEYFTR.I	Aspartic proteinase nepenthesin-1_ putative OS=*Ricinus communis*
	K.GEHTYVPVTK.K	Aspartic proteinase (Fragment) OS=*Cucumis sativus*
	R.EDGIVSPVK.D	Cysteine proteinase OS=*Elaeis guineensis* var. *tenera*
E. minuta	R.LIFAGKQLEDGR.T K.EGIPPVQQR.L R.TLADYNIQK.E E.VESSDTIDNVKAK.I	Ubiquitin-NEDD8-like protein RUB2 OS=*Oryza sativa* subsp. *japonica*
	K.VESSDTIDNVKAK.I R.LIFAGKQLEDGR.T R.TLADYNIQK.E	Ubiquitin OS=*Musa acuminata*
	R.LIFAGKQLEDGR.T R.TLADYNIQK.E K.EGIPPVQQR.L	Polyubiquitin 2 OS=*Zea mays*
	L.GVESGQDAVIR.G R.TPEEILR.I	Dessication-related protein_ putative; 70055-71849 OS=*Arabidopsis thaliana*
	R.GNGTGGESIYGEK.F	Cyclophilin A (Fragment) OS=*Triticum aestivum*
	K.FFKGQC(+57.02)PQAYSYAK.D K.DDATSV(sub T)FTC(+57.02)PSP(sub G)TNYK.V	Thaumatin-like protein OS=*Cryptomeria japonica*
	K.GQC(+57.02)PQAYSYAK.D	Thaumatin-like protein OS=*Pinus taeda*
	R.STPEMWPDIIQK.A K.NVVFNTAK.I K.WGHLKEL.H	Beta-galactosidase OS=*Picea sitchensis*
	R.YAVNYVR.G	Beta-glucosidase_ putative OS=*Ricinus communis*
	A.VNQDSLGVQGK.K K.ALADYVHAK.G	Alpha-galactosidase OS=*Oryza sativa* subsp. *japonica*
	R.WEVPYNLLPR.E	Alpha-xylosidase OS=*Arabidopsis thaliana*
E. monosperma	K.YM(+15.99)VIQGEPGVVIR.G	Profilin-1 (Fragment) OS=*Triticum aestivum*
	K.YM(+15.99)VIQGEPGVVIR.G	Profilin OS=*Zea mays*
	L.LGVESGQDAVIR.G	Dessication-related protein_ putative; 70055-71849 OS=*Arabidopsis thaliana*
	K.LVIVGDGGTGKT.T	GTP-binding nuclear protein Ran-A1 OS=*Nicotiana tabacum*
	R.SPSAYLNNPP(sub A)EER.N	Ceramidase_ putative OS=*Ricinus communis*

	R.VWVYSGDTDGRVP.V	Serine carboxypeptidase 1 OS=*Zea mays*
	L.FNENLKPGPTG(sub S)ER.N	Glucan endo-1_3-beta-glucosidase 11 OS=*Arabidopsis thaliana*
E. trifurca	K.SSEEAME(sub N)DYITK.V M.GLKEEFEEY(sub H)AEK.V R.AKWDAWK.A	Acyl-CoA-binding domain-containing protein 6 OS=*Arabidopsis thaliana*
	K.EGIPPVQQR.L R.TLADYNIQK.E K.IQDKEGIPPDQQR.L E.VESSDTIDNVK.A	Ubiquitin-NEDD8-like protein RUB2 OS=*Oryza sativa* subsp. *japonica*
	K.EGIPPVQQR.L R.TLADYNIQK.E K.IQDKEGIPPDQQR.L L.EVESSDTIDNVK.A	Ubiquitin_ putative OS=*Ricinus communis*
	K.ITSFLDPDGWK.T K.V(sub T)VLVDNEDFLK.E	Lactoylglutathione lyase OS=*Gossypium hirsutum*
	Q.QLPQFEEIR.N	Alpha-gliadin OS=*Triticum aestivum*
	K.VTE(sub L)QDLEDTYNPPFK.S	Os11g0291000 protein OS=*Oryza sativa* subsp. *japonica*
	R.IGGIGTVPVGR	Elongation factor 1-alpha (Fragments) OS=*Pseudotsuga menziesii*
	K.EHGAQEGQAGTGAFPR.C	Alpha-amylase inhibitor 0.19 OS=*Triticum aestivum*
	K.EHGAQEGQAGTGAFPR.C	Dimeric alpha-amylase inhibitor OS=*Aegilops umbellulata*
	K.AGLQFPVGR.I	Probable histone H2A.1 OS=*Oryza sativa* subsp. *japonica*
	K.VTQ(+.98)QDLEDTYNP(sub V)PFK.S E.TMIGNYAGK.A E.WWSEALHGISDVGPGT(sub A)K.F	Beta-D-xylosidase 1 OS=*Arabidopsis thaliana*
	H.T(sub S)AITSGQGFGGTIK.A R.ELAAFFANVMHETS(sub G)GL.C S.WNYNYGAAGK.S	Class IV chitinase Chia4-Pa2 variant (Fragment) OS=*Picea abies*
	R.STPEMWPDLIQK.A A.FRTDNEPFKA.A	Beta-galactosidase OS=*Pyrus communis*
	R.STPEMWPDLIR.K	Beta-galactosidase (Fragment) OS=*Mangifera indica*
	K.MELIDAAFPLLK.G	Malate dehydrogenase OS=*Picea sitchensis*
	R.VWVYSGDTDGRVPVT.S	Serine carboxypeptidase II-3 OS=*Hordeum vulgare*
	I.GGYYDAGDNVK.F	Endoglucanase 20 OS=*Arabidopsis thaliana*
	GGYYDAGDNVK.F	Putative endo-1_4_-beta-glucanase (Fragment)

	OS=*Solanum lycopersicum*
R.FDNNYYK.D	Peroxidase (Fragment) OS=*Lupinus polyphyllus*
K.GEHTYVPVTK.K	Aspartic proteinase (Fragment) OS=*Cucumis sativus*

E. compacta (13), *E. minuta* (11), *E. likiangensis* (9), and *E. monosperma* (6). These proteins could be divided into intracellular (64 %) and extracellular proteins (36 %). The percentage of intracellular proteins ranged from 44 – 100 %: *E. likiangensis* (44 %), *E. minuta* (45 %), *E. trifurca* (50 %), *E. compacta* (54 %), *E. monosperma* (67 %), *E. foeminea* (80 %) and *E. distachya* (100 %).

In all pollination drops a variety of intracellular proteins were detected (Tables 1, 3). The most frequently detected intracellular proteins - ubiquitin and polyubiquitin - were in five species (Table 1). Dessication-related proteins were detected in four species. Cyclophilin-α, histones, and elongation factor 1-α were detected in three different species. Four of the most common proteins, *i.e.* detected in more than three or more species, were detected in drops of *E. foeminea*. However, this might be expected given that the *E. foeminea* had the most proteins of any species in this comparative analysis. *E. compacta* had three of the commonly shared proteins. The remaining proteins on Table 1 were detected one or two times only.

Extracellular proteins were less abundant than intracellular proteins (Tables 2, 3). The most commonly shared extracellular proteins were xylosidases (Table 2), which were detected in drops of four *Ephedra* species. Aspartic protease, β-galactosidase, peroxidase and serine carboxypeptidase were detected in three *Ephedra* species. The remaining seven proteins on Table 2 were detected only once or twice.

On a species level, proteins detected in drops represented a wide variety of enzymes. The proteins are either water-soluble proteins secreted into the pollination drop, or are from the water-soluble portion of plant cells: no membrane-anchored proteins were detected in any samples. *Ephedra foeminea* drops had a probable defence protein

(chitinase), two carbohydrate-modifying enzymes (β-xylosidase, glycosyl-hydrolase-like protein), and proteases (aspartic protease, serine carboxypeptidase). The largest number of proteins was associated with the cytoplasm, including histone proteins, citrate synthase, elongation-factor-1-α, cyclophilin, calreticulin, luminal-binding protein 4, a probable glycerophosphoryl diester phosphodiesterase, polyubiquitin, peptidyl-prolyl cis-trans isomerase, BIP isoform A, and granule bound starch synthase. *Ephedra trifurca* had a similar number of characterized proteins to *E. foeminea,* divided evenly between secretome and degradome. *Ephedra trifurca* had some of the same proteins as *E. foeminea* (histone, elongation-factor-1-α, ubiquitin, chitinase, β-xylosidase, aspartic proteinase, serine carboxypeptidase). The proteins found in drops of *E. trifurca* were divided evenly between degradome and secretome. *Ephedra trifurca* had defence proteins, including a chitinase and an alpha amylase inhibitor, peroxidase and endoglucanases, as well as carbohydrate-modifying enzymes, e.g. β-D-xylosidase and β-galactosidase, and a serine carboxypeptidase. Additional apoplastic enzymes, such as malate dehydrogenase, were also detected.

In drops of *E. likiangensis*, intracellular and extracellular proteins were equally present; among the symplastic proteins, ubiquitin and proteases were predominant. *Ephedra minuta* drops had abundant symplastic ubiquitins (Table 3), as well as apoplastic carbohydrate-modifying enzymes (β-xylosidase, β-glucosidase) and defence proteins (thaumatin-like proteins). Cellular proteins not normally found in the apoplast included Elongation factor 1-α, ubiquitin, acyl-CoA-binding domain-containing protein, and actin. *Ephedra compacta* had a number of ubiquitin and polyubiquitin proteins, as well as acyl-CoA-binding domain-containing protein, calmodulin, a peptidase, and α-amylase; all of

these were degradome proteins. Among the secretome proteins were β-xylosidase, β-galactosidase, superoxide dismutase, aspartic protease and peroxidase. *Ephedra monosperma* had mostly degradome proteins (profilins, desiccation-related protein, the GTP-binding protein RAN-1, and ceramidase) and had only two secretome proteins that we could detect in this initial comparative study – serine carboxypeptidase and glucan endo-1,3-β-glucosidase. *Ephedra distachya* was unique among the species sampled, because all of its 15 proteins were degradome proteins (Table 3).

Comparative study of *Ephedra monosperma* drops from three dates

We were able to get samples of *Ephedra monosperma* pollination drops from three different dates (Table 4). Thirty-two proteins were identified from these samples, more than four times the number found in *E. monosperma* sample used in the comparative study of different *Ephedra* species (Table 3). The number of proteins declined with time, with the largest number of proteins (22) found in the first sample (Mar. 9), which was not long after pollination drops began to be produced in the greenhouse. On the next two dates, progressively fewer proteins were found until only 14 proteins could be detected on the final date (Apr. 10). Four proteins were found at all three time points, including a serine carboxypeptidase-like 32 protein homologous to one found in *Arabidopsis thaliana,* a histone 4 protein homologous to one in *Pisum sativum* L., α-galactosidase and a predicted protein homologous to one in *Populus trichocarpa* Torr. & A.Gray. Fourteen proteins were detected at two time points and 14 were only found at one time. Most proteins (20/32) were degradome proteins. The exceptions were extracellular proteins, such as serine carboxypeptidase, thaumatin-like protein, acid α- and β-galactosidase, peroxidase and α–xylosidase.

Table 4. *Ephedra monosperma* pollination drop proteins from three collection dates.

Protein	Mar 9	Mar 24	Apr 10
Histone H4 OS=Pisum sativum PE=1 SV=2	x	x	x
Predicted protein OS=*Populus trichocarpa* GN=POPTRDRAFT_642406 PE=4 SV=1	x	x	x
Putative uncharacterized protein OS=*Selaginella moellendorffii* GN=SELMODRAFT_143620 PE=4 SV=1	x	x	
Putative uncharacterized protein OS=*Glycine max* PE=2 SV=1	x	x	
Acyl-CoA-binding protein (Fragment) OS=*Jatropha curcas* PE=2 SV=1	x	x	
Glycosyl hydrolase family-like protein OS=*Salvia miltiorrhiza* PE=2 SV=1	x	x	
GTP-binding nuclear protein Ran-A1 OS=Nicotiana tabacum GN=RAN-A1 PE=2 SV=1		x	x
Eukaryotic initiation factor 4A OS=Triticum aestivum PE=2 SV=1		x	x
RAS-like protein (Fragment) OS=*Arabidopsis thaliana* PE=2 SV=1		x	x
Translation initiation factor OS=*Zea mays* GN=eIF-4A PE=2 SV=1		x	x
Acid beta-fructofuranosidase OS=Solanum lycopersicum GN=TIV1 PE=2 SV=1	x		
Alpha-glucosidase OS=Hordeum vulgare PE=2 SV=1	x		
Multicystatin OS=*Helianthus annuus* GN=smc PE=2 SV=1	x		
Polyubiquitin 11 OS=Arabidopsis thaliana GN=UBQ11 PE=1 SV=1	x		
Predicted protein OS=*Populus trichocarpa* GN=POPTRDRAFT_1090916 PE=4 SV=1	x		
Endoglucanase 23 OS=Oryza sativa subsp. japonica GN=GLU12 PE=2 SV=1	x		
NtPRp27-like protein OS=*Solanum tuberosum* PE=2 SV=1		x	
Ubiquitin-like protein (Fragment) OS=*Solanum lycopersicum* GN=ubiquitin-like PE=2 SV=1		x	
Cyclophilin A (Fragment) OS=*Triticum aestivum* GN=CYP18-3 PE=3 SV=1			x
Photosystem II Q(B) protein (Fragment) OS=*Kochia scoparia* GN=psbA PE=4 SV=1			x
Alpha-galactosidase OS=Coffea arabica PE=1 SV=1	x	x	x
Serine carboxypeptidase-like 32 OS=Arabidopsis thaliana GN=SCPL32 PE=2 SV=1	x	x	x
Alpha-galactosidase OS=Oryza sativa subsp. japonica GN=Os10g0493600 PE=1 SV=1	x		x
Acid alpha galactosidase 1 OS=*Cucumis sativus* PE=2 SV=1	x		x
Alpha-xylosidase OS=*Arabidopsis lyrata* GN=ARALYDRAFT_894626 PE=4 SV=1	x	x	
Peroxidase (Fragment) OS=Lupinus polyphyllus PE=2 SV=2	x	x	
Alpha-xylosidase OS=Arabidopsis thaliana GN=XYL1 PE=1 SV=1	x	x	
Beta-galactosidase 1 OS=Oryza sativa subsp. japonica GN=Os01g0533400 PE=2 SV=1		x	x
Beta-galactosidase 8 OS=Arabidopsis thaliana GN=BGAL8 PE=2 SV=2	x		
Thaumatin-like protein OS=*Cryptomeria japonica* GN=Cry j 3.1 PE=2 SV=1	x		
Alpha-galactosidase OS=*Coffea canephora* GN=gal1 PE=2 SV=1		x	
Beta-galactosidase 9 OS=Oryza sativa subsp. japonica GN=Os06g0573600 PE=2 SV=1			x

Discussion

Pollination drops of *Ephedra* contain proteins. Although this has not been reported previously in *Ephedra*, it was expected, as all other pollination drops analyzed to date contained proteins. However, the protein profiles of this study exhibit some notable differences from those of other gymnosperms we have examined, most of which were conifers (Wagner et al. 2007). *Ephedra* spp. not only have lower concentrations of protein, judging from the light staining of the bands in the gels, but also contain fewer total proteins. In addition, the protein profiles of *Ephedra* show substantial amounts of intracellular proteins not found in conifer pollination drops. In short, *Ephedra* has a degradome, consisting of proteins, and presumably shorter peptide fragments. The most likely source of the protein degradome is from nucellar degeneration which forms the flask-shaped pollen chamber during pollination drop production, causing intracellular proteins to be added to the other pollination drop compounds. This assumption is logical, since pollen chamber formation occurs prior to and during pollination drop secretion (Rydin et al. 2010). A protein that is characteristic of this degradome is ubiquitin, which plays a major role in recycling proteins inside a cell. It is not known to function outside of the cytoplasm, *i.e.* in the apoplastic fluids of plants. Protein profiles of both degradome and secretome are composed of a few dozen proteins at most. Compared to other gymnosperms, the average number of proteins, which is about a dozen per *Ephedra* species, is slightly greater than in pollination drops of the Cupressaceae sampled to date, which range from half-a-dozen to a dozen (Wagner et al. 2007), but much less than those of pinaceous species, such as *Pseudotsuga menziesii* (Poulis et al. 2005) and *Larix* x *marschlinsii* (O'Leary et al. 2007), which have many dozens each.

In *Ephedra* pollination drops there are also proteins that are not part of the degradome. These proteins are likely formed inside cells and discharged into the apoplastic fluid by active cellular processes, and together these constitute a secretome of substances exported into pollination drops, similar to what has been found in most gymnosperms investigated using proteomics. Chitinase is an example of a protein that belongs to the secretome. In the results reported here, chitinases were present in both *E. foeminea* and *E. trifurca*. Chitinase is also found in pollination drops of another gnetophyte, *Welwitschia mirabilis*, as well as a number of conifers (Wagner et al. 2007). In Douglas-fir drops, chitinases are able to process chitin substrates *in situ*, which suggests that they are active in anti-fungal defence during sexual reproduction (Coulter et al. 2012). Should the chitinases in the pollination drop of *Ephedra* prove functional, they may also protect ovules, which like those of other gymnosperms are exposed to the elements and are, therefore, more vulnerable to wind-borne pathogens than those of angiosperms, which are enclosed within a protective ovary.

The percentage of characterized cellular versus secretory proteins in the drops ranged from 44 % to 100 %, depending on species. Other gymnosperms, such as *Juniperus*, typically have no intracellular proteins in their pollination drops (Wagner et al. 2007). The most common intracellular protein found in *Ephedra* pollination drops is ubiquitin, which is found in five of the seven species. Of the 24 intracellular proteins detected, only 10 are found in more than one species. This implies that although a degradome is universal in *Ephedra* pollination drops, its composition may widely differ among the species. To provide a better idea of variation of protein profiles, studies need

to be undertaken that focus on variation among individual plants as well as over the period of pollination drop secretion.

A measure of the variation in degradome is given by our samples of *E. monosperma* plants from the same greenhouse population over three time points from early to late in the pollination drop period. There were more identifiable proteins at the beginning of the period than at the end, which may imply that proteins initially present in drops are broken down over time. Most of the proteins were clearly intracellular proteins, e.g. GTP-binding nuclear protein RAN 1, confirming that a degradome is constantly present in the drops. Only a few proteins are found across all time points, e.g. histone 4, the majority varying widely. This was equally true for secretome and degradome profiles. *Ephedra monosperma* has as much variation over time as there is among all species of *Ephedra* (Table 3). Investigations into variation within a species are important, as they will better allow us to isolate proteins that may have biological function.

The question of function must be considered carefully. Caution must be exercised for many reasons. These drops not only capture pollen, but fungi, bacteria, viruses and dust. We have been able to show in previous studies that enzymes in the drop, in particular, chitinases and invertases are able to function *in situ*, but this work is difficult because of the small amount of liquid with which one has to work. As a consequence, it is one thing to find proteins with identities and therefore, possible functions, but it is quite another to prove that the proteins function as expected based on their sequence-based identities.

We assume that the degradome proteins, for example, ubiquitin, and histones are not functional in the drop, because they are outside the cell where they are normally

located. Cytoplasmic proteins such as ubiquitin are involved in recycling proteins and peptides targeted for breakdown *inside the cell*. Ubiquitin has not been previously found in pollination drops of members of Pinaceae in which pollen chambers are not formed and in which the nucellus does not undergo degradation at the time of drop release, e.g. *Pseudotsuga* and *Larix*. Other proteins that are strictly cytoplasmic include cyclophilin A (a plant immunophilin), which is restricted to cell organelles: its presence in the drop is likely due to cell death and subsequent leakage of cellular contents.

Focusing on two species in the comparative study, *E. foeminea* and *E. distachya*, we found that *E. foeminea* had the most detected proteins, half of which were degradome proteins, whereas *E. distachya* had only degradome proteins. Having about 50 percent degradome proteins is close to the average for the seven species that we measured. In addition to ubiquitin, just discussed, notable degradome proteins in *E. foeminea* are histones, which are normally restricted to the nucleus and involved in chromosome organization, Granule-bound starch synthase, which synthesizes amylose in the chloroplast, BIP isoform A, which is a molecular chaperone located on the endoplasmic reticulum, and immunophilins such as peptidyl-prolyl cis-trans isomerase, which are found in a number of locations within the cell. The predominance of these cytoplasmic proteins among the degradomic fraction is probably due to either their abundance in degrading cells, and/or in their slower rate of degradation compared to that of other proteins (i.e. already reduced to small peptides or amino acids). The profile of proteins detected in pollination drops of *E. distachya* consists entirely of intracellular proteins, none of which are normally found in apoplastic secretions, including: proteins involved in signal transduction, e.g. small Ran-related GTP-binding protein; calmodulin 4, which

is a regulatory protein controlled by calcium; nucleoside diphosphate kinase, which regulates metabolic pools of nucleoside diphosphates; histones, which control chromosome organization; and heat-shock proteins, which regulate a plant cell's response to stress. Recently, there have been papers that suggest a few of these proteins may function in the apoplast. For example, root border cells of angiosperms and gymnosperms (Wen et al. 2008b) upregulate gene expression that results in secretion of intracellular proteins such as DNA-bound histones, that act a trap for pathogens (Hawes et al. 2012).

In other gymnosperm pollination drops, most proteins do not appear to be related to a degradome, but are likely secreted by cells directly into the apoplast. In these cases the collective secreted protein component is known as a secretome. The secretome proteins that we have been able to identify from our analyses of various species of gymnosperms were from many classes of enzymes. We detected a variety of defence proteins, including among others, thaumatin-like protein, peroxidase, glucan-endo-β-1,3-glucanase, and superoxide dismutase. However, the proteins of the secretome are probably not all involved in defence. There are also carbohydrate-modifying enzymes such as α- and β-galactosidase proteins. In roots of peas, galactosidases operate on cell wall fragments to produce galactose, which is inhibitory to root growth (Wen et al. 2008a). All of the proteins that we have designated as part of the secretome, e.g. peroxidase, malate dehydrogenase, superoxide dismutase and thaumatin-like proteins, have been found apoplastically in other plants. Some protein classes have many members that have diverse functions, e.g. serine carboxypeptidases. These include serine carboxypeptidases that have regulatory functions both in the cytoplasm, and in the extracellular spaces. Until these proteins are shown to function *in situ* in the pollination

drop, they are, like all other enzymes included in our lists, assigned to the secretome because they or members of their class of protein have been detected in the apoplast of other plants. In our survey of *Ephedra* presented here, no proteins are common to the secretomes of all species. The number of proteins ranges among the *Ephedra* species between 2 and 10 per pollination drop/species.

We expected to find acid phosphatase in the drop, since two different laboratories have reported its presence *via* activity assays in Gnetales. Ziegler (1959) detected it in the nucellus of *Ephedra helvetica* C. A. Mey. (=*E. distachya* subsp. *helvetica*) and in the non-gnetalean *Taxus baccata* (Taxaceae), and Carafa et al. (1992) reported its presence in pollination drops of *W. mirabilis*. However, we did not detect this enzyme in any pollination drops of the seven *Ephedra* species that we analyzed using proteomics methods. We have never found it in any conifers, although a proteomic analysis of the nucellus has yet to be completed.

There are more proteins in these species than we have been able to describe. In all *Ephedra* species, there was a relatively high percentage of uncharacterized proteins. Although the mass spectra of proteins for which no identity can be assigned are of high quality, the databases against which we search this information often have insufficient depth, particularly with regards to gymnosperms. This situation should improve if, in future, databases improve. For example, genomes of *Picea abies* (L.) H. Karst. (Nystedt et al. 2013) and *Picea glauca* (Moench) Voss (Birol et al. 2013) will be useful once they are annotated. As more gymnosperms are covered, the improved depth of the databases will assist in protein identification. Molecular biologists will be able to use these

databases to make better protein identifications and to improve the prediction of functions for these proteins.

Until this study, proteins in pollination drops were considered to be functional (Nepi et al. 2009). The possibility that proteins may also be byproducts of pollen chamber formation that have been washed into the drop has never been explored. This is due to the fact that the species investigated to date did not have pollen chambers formed from nucellar breakdown. Thus the pollination drops of *Ephedra* are probably a mixture of functional and formerly functional, as well as biologically inactive proteins and/or peptides. As such, *Ephedra* differs from conifers analyzed to date, such as Pinaceae and Cupressaceae. It will be interesting to expand pollination drop analysis into *Pinus*, *Ginkgo* and cycads, all of which have pollen chambers. The low amount of protein in *Ephedra* drops suggests a less important role, if any, for these proteins during reproduction. The higher sucrose concentrations in these drops result in higher osmotic pressure in these drops, which may prevent foreign pollen from germinating (von Aderkas et al. 2012) and pathogens from establishing and growing.

Ephedra pollination drops have proteins that can be divided into those that belong to the degradome, itself a result of pollen chamber formation, and those that are exported by the cytoplasm into the drop and form an active part of the secretome that is, based on similarity to other gymnosperms, involved in carbohydrate modification, defence and other apoplastic activities.

Chapter 3: Using a custom transcriptome-derived database to identify proteins in gymnosperm pollination drops.

Introduction

Pollination drops are small and elusive. Their volumes range from 10-1000 nL, and they are only produced when plants are receptive to pollination (Prior et al. 2013). Receptivity is usually seasonal (Tang 1987; Takaso 1990; Poulis et al. 2005; del Tredici 2007), or occurs for only a brief time on individual plants (Kato et al. 1995). An ovule may only produce drops for a short number of days, and drops may only be present at certain hours of the day (Tang 1987; Kato et al. 1995; Wetschnig and Depisch 1999). One must be at the right place at the right time to collect pollination drops. This can only be determined by careful monitoring of reproductive cycles in a given geographical location. Collecting a sufficient quantity of pollination drops for biochemical analysis is therefore challenging due to their tiny volume and ephemeral nature.

Proteomics experiments on pollination drops have been ongoing for more than a decade (Poulis et al. 2005; O'Leary et al. 2007; Wagner et al. 2007). Pollination drops lend themselves well to proteomics experiments. In comparison to cellular extracts, they are relatively simple in composition. Extensive sample preparation is not required. Advances in mass spectrometry methods and analysis software have contributed to increased sensitivity and increased throughput for proteomics experiments. These methods can be applied to the study of pollination drops (Prior et al. 2013).

The challenge of peptide and protein identification in non-model organisms

The gymnosperm groups (cycads, *Ginkgo*, Gnetales, conifers) are poorly represented in publicly available databases. For example, the UniProt-SwissProt database, which is a database of high-quality, manually annotated and reviewed

sequences, contains no gymnosperms amongst its 250 best-represented species (542 782 sequence entries, UniProt Website, March 2014). A number of plant species are amongst the top 250 representatives in this database, but none of these are gymnosperms. The same holds true for the larger, automatically annotated Uniprot-TrEMBL database (54 247 468 sequence entries, Expasy Website, March 2014). This presents a problem for protein identification using traditional database searching methods (Evans et al. 2012; Champagne and Boutry 2013). These methods rely on exact mass matching between theoretical peptides generated by the *in silico* digestion of sequences in a selected database and mass information derived from the mass spectra of peptides in a given sample (Steen and Mann 2004; Martin et al. 2013). If the sequence is not in the searched database, no identification can be made of the given peptide. Although sequence data from some plant species, such as *Arabidopsis*, are well represented in public databases, there are limitations to cross-species peptide identifications. The greater the evolutionary distance between species, the greater the likelihood of sequence change, and the less likely it becomes for exact matches to occur (Champagne and Boutry 2013).

The increasing popularity of next generation RNA-sequencing experiments is making transcriptome data available for numerous non-model plant species (Martin et al. 2013). The 1000+ Plants Project (also known as OneKP or 1KP) is the most extensive plant transcriptomics project to date (OneKP website). This project aims to publish the transcriptomes of 1000 plant species, representing many evolutionarily interesting species from all clades of extant green plants. Currently, the 1KP dataset is only available upon request. The entire dataset will become publicly available when the 1KP capstone paper is published in 2015. To date, the 1KP dataset contains two orders of magnitude more

plant gene sequences than all other public databases (OneKP Capstone Wiki website). Other sources of plant transcriptomic data are already publicly available. For example, the Dendrome Forest Tree Genome Database provides free access to the transcriptomes of nine gymnosperm species (Dendrome website).

The availability of plant transcriptomic data has opened the door to deeper proteomic analysis in non-model plants (Brautigam et al. 2008; Lopez-Casado et al. 2012; Champagne and Boutry 2013; Martin et al. 2013; Seo et al. 2013). RNA-seq data can be translated into amino acid sequence to create a custom database suitable for proteomic search software (Evans et al. 2012; Lopez-Casado et al. 2012). Further, the quality of RNA-seq databases has been shown to be of comparable quality to databases built over time. Lopez-Casado et al. (2012) demonstrated that there was little difference in the quantitative or qualitative proteomic data obtained from searching a database derived from RNA-seq data compared to a database built from over 10 years' worth of EST data.

There are additional methods that can improve peptide identification in proteomics experiments of non-model organisms. For example, de novo sequencing and homology searching can be incorporated into the analysis of mass spec data. De novo peptide sequencing works by deriving sequence information directly from mass spec data, without relying on a sequence database. De novo sequencing can discover additional peptides not found in databases, or can be combined with database search algorithms to add support to database derived identifications (Ma and Johnson 2012; Zhang et al. 2012). Proteomics-specific homology search tools can also improve cross-species protein identifications. These algorithms allow detection of peptides with

conservative amino acid substitutions in peptide sequences which otherwise match to database sequences. Some programs, such as SPIDER, are more sophisticated in that they consider both substitutions and common de novo sequencing errors (Ma and Johnson 2012).

After the completion of the *Ephedra* pollination drop data analyses and submission of the results to Botanical Review, our group gained access to the extensive gymnosperm transcriptome dataset created by the 1KP Project, in addition to other private and public resources described below. The methods of mass spectrometry and data analysis described in this chapter pertain to the remaining gymnosperm pollination drop samples: *Zamia furfuracea*, *Ceratozamia hildae*, *Cycas rumphii*, *Ginkgo biloba*, *Gnetum gnemon*, *Welwitschia mirabilis* and *Taxus* x *media*. To identify proteins in the pollination drops of these gymnosperms, we created a custom protein database from gymnosperm RNA-seq data. We analyzed our pollination drop mass spec data using software that combines both de novo peptide sequencing and homology searching with a traditional database searching algorithm. We chose these methods to maximize the number of peptides identified in our samples, thereby improving the quality of the resulting protein identifications.

Methods

Pollination drop collection

In general, pollination drops were collected using either a 10 µL micropipette tip or a 10 µL glass capillary tube that had been pulled over a flame to create a fine tip. Drops were drawn into the tubes or tips by capillary action. Samples were expelled into 1.5 mL microcentrifuge tubes or 1.5 mL Eppendorf tubes. Great care was taken to avoid keratin contamination during collection. Drops were stored at - 20 ˚C until mass

spectrometry analysis. Details of pollination drop collections pertaining to individual species are provided in the Methods sections of the following chapters.

Proteomics

Creation of the Gymno_DB database

Gymnosperm RNA-Seq data were collected from the following resources: unpublished data from the NSERC Strategic Grant *Megastigmus* and Conifers: The Biology of Invasion (1 species, courtesy of Ian Boyes, Dr. Stefan Little, Dr. Patrick von Aderkas); unpublished data from the New York Plant Genomics Consortium NSF grant: IOS-0922738 (three species, courtesy of Dr. Dennis W. Stevenson); published data from the Dendrome Project (nine species, Dendrome website); unpublished data from the 1KP project (84 species, OneKP website). Information pertaining to the species, tissue types and sources is provided in Appendix 1. A total of 97 RNA-Seq datasets representing 91 species (six duplicate species) were included in the final database. Representatives from all of the extant gymnosperm clades were included.

A Python script (provided by Ian Boyes) was used to break-up sequence data from the 1KP project at each unknown base (at each "n"). This was necessary to make the sequence data compatible with the proteomics program PEAKS 6 (Bioinformatics Solutions Inc., Waterloo, Canada). The assemblies collected from the other sources did not require this adjustment.

The program EMBOSS GetORF (MRC Rosalind Franklin Centre for Genomics Research, Wellcome Trust Genome Campus, Cambridge, UK) was used to predict all possible open reading frames, and to translate the transcript data into amino acid sequence. The following qualifiers were used: Table 1 (Standard with alternative initiation codons); Find 1 (Translation of regions between START and STOP codons).

All of the resulting opening reading frames were concatenated into one FASTA file along with the Common Repository of Adventitious Proteins database version 2012.01.01 (cRAP, The Global Proteome Machine Organization). The final database contained 62 148 544 entries. Each entry consisted of a unique header that identified the source species and the transcript name, followed by the amino acid sequence of the predicted open reading frame. This database is hereafter referred to as Gymno_DB. The Gymno_DB database was verified by PEAKS 6 and uploaded by Derek Smith to the PEAKS Online 6 server located at the University of Victoria Genome British Columbia Proteomics Centre.

Mass spectrometry analysis

Mass spectrometry was carried out by the UVic Genome BC Proteomics Centre.

Sample preparation

Pollination drop samples were first reduced with dithiothreitol for 30 min at 37˚C, and then alkylated with iodoacetamide at 37˚C in darkness for 30 min. Samples were then digested with 2 µg trypsin (Promega, Madison, WI, USA) at 37˚C for 16 hr. A Waters HLB Oasis column (Milford, MA, USA) was used to desalt the samples. The samples were concentrated using a Speed Vac (Thermo Fisher Scientific, Bremen, Germany) and then stored at -80 ˚C until analysis.

Liquid chromatography – tandem mass spectrometry analysis (LC-MS/MS)

A 100 µL sample of the peptide mixture was rehydrated with 2 % acetonitrile/water/2 % formic acid. Samples were separated by on-line reversed-phase chromatography using a Thermo Scientific EASY-nLC II system (Thermo Fisher Scientific) with an in-house prepared reversed-phase pre-column Magic C-18AQ (100

μm in diameter, 2 cm length, 5 μm, 100Å, Michrom BioResources Inc., Auburn, CA) and an in-house prepared reversed-phase nano-analytical column Magic C-18AQ (75 μm in diameter, 15 cm length, 5 μm, 100Å, Michrom BioResources Inc.). The flow rate was 300 nl/min. An LTQ Orbitrap Velos mass spectrometer with a Nanospray II source (Thermo Fisher Scientific) was coupled to the chromatography system. Two solvents were used: solvent A consisted of 2 % acetonitrile and 0.1 % formic acid; solvent B consisted of 90 % acetonitrile and 0.1 % formic acid. A 249 bar (~5 μL) pre-column equilibration and a 249 bar (~8 μL) nano-column equilibration were performed. The samples were then separated by a 90 min gradient (0 min: 5 % solvent B; 80 min: 45 % solvent B; 2 min: 90 % solvent B; 8 min: 90 % solvent B).

The following parameters were used on the LTQ Orbitrap Velos (Thermo Fisher Scientific): nano-electrospray ion source with spray voltage 2.2 kV; capillary temperature 225 ˚C; survey MS1 scan *m/z* range 400-2000 profile mode; resolution 60 000 @ 400 *m/z* with AGC target 1E6; one microscan with maximum inject time 200 ms; Lock mass Siloxane 445.120024 for internal calibration with preview mode for FTMS scans: ON; injection waveforms: ON; monoisotopic precursor selection: ON; rejection of charge state:1. The five most intense ions (charge state 2-4) exceeding 5000 counts were selected for a 7500 resolution high collision dissociation (HCD) (FT MSMS fragmentation scans 2-6). Detection was in profile mode. Dynamic exclusion settings were: repeat count: 2; repeat duration: 15 sec; exclusion list size: 500; exclusion duration: 60 sec with a 10 ppm mass window. The HCD activation isolation window was: 2 Da; AGC target: 1E5; maximum inject time: 500 ms; activation time: 0.1 ms; activation Q: 0.250; HCD collision energy: 30 %.

Protein Identification: Data analysis parameters

Raw LC-MS/MS data were converted into Mascot Generic Format (MGF) using Proteome Discoverer 1.4 (Thermo Fisher Scientific). MGF files were searched using PEAKS 6 (Bioinformatics Software Inc.) against our in-house database Gymno_DB. The PEAKS software suite includes a database search program called PEAKS DB that is able to incorporate traditional database searching and de novo sequencing. It also includes an optional homology search tool called SPIDER. PEAKS DB + SPIDER searches were performed using the following parameters: parent mass error tolerance: 8.0 ppm; fragment mass error tolerance: 0.03 Da; precursor mass search type: monoisotopic; enzyme: trypsin; max-missed cleavages: 1; non-specific cleavage: none; fixed modifications: carbamidomethylation 57.02; variable modifications: deamidation (NQ) 0.98, oxidation (M) 15.99; max variable PTM per peptide: 3.

PEAKS 6 software includes a results validation program. It uses a decoy-fusion method to calculate the false discovery rates (FDR) at the peptide spectrum match (PSM), peptide and protein levels (Zhang et al. 2012). All results in this study were filtered with an FDR of $\leq$ 1 % at the PSM, peptide and protein levels. This is the filter stringency recommended by PEAKS (PEAKS website). In addition, we required at least one unique peptide per protein identification.

Protein identification: Annotation of predicted open reading frames from Gymno_DB

Since the predicted open reading frames contained in Gymno_DB are not annotated with protein names, cellular locations or functional information, BLASTP (BLAST 2.2.28+, National Centre for Biotechnology Information) was used to annotate the predicted open reading frames (proteins) identified as good matches to pollination drop mass spec data. PEAKS 6 was used to generate FASTA files containing the amino

acid sequences corresponding to the full-length open reading frames of all good hits for each LC-MS/MS run. Each FASTA file was blasted against the *Arabidopsis* Information Resource (TAIR 10) protein database accessed online from ftp://ftp.arabidopsis.org/home/tair/Proteins/TAIR10_protein_lists/. BLASTP was run with standard parameters using the best-hit algorithm. These BLASTP searches used computing resources provided by WestGrid and Compute/Calcul Canada (Hermes and Hungabee). The BLASTP searches provided TAIR10 gene model numbers. BLASTP hits were filtered at a threshold of e value ≤ e-5. Descriptive names and Gene Ontology (GO) annotations for the TAIR10 gene models (Berardini et al. 2004) were collected from the file ATH_GO_GOSLIM.txt.gz (03/09/2013) accessible on the website ftp://ftp.arabidopsis.org/home/tair/Ontologies/Gene_Ontology/.

Chapter 4: Proteins in cycad pollination drops.

Introduction

Cycads are a distinct lineage amongst the five extant seed-plant groups (Stevenson 1992). The earliest known cycad fossils date to the Lower Permian age of China (Martinez et al. 2012). A period of peak diversity and abundance occurred during the Jurassic-Cretaceous periods (199.6 to 65.5 Mya), followed by an apparent loss of diversity possibly due to extinctions (Crisp and Cook 2011; Nagalingum et al. 2011). The diversity of extant species is thought to be largely the result of a more recent radiation that occurred within the past 12 million years (Crisp and Cook 2011; Nagalingum et al. 2011; Salas-Leiva et al. 2013). About 300 species of cycads in 10 genera (Osborne et al. 2012) grow in tropical, subtropical and warm temperate regions today. Due to anthropogenic causes, many cycad taxa are now rare in the wild (Nagalingum et al. 2011). Rare cycads are targeted by plant poachers looking to profit from the sale of unusual specimens (IUCN website). The majority of cycad species (63 %) are now listed in the IUCN Red List of Threatened Species (IUCN Red List 2013).

Pollination in cycads

Cycads are dioecious, bearing male and female cones (strobili) on separate plants. Male cones (microsporangiate strobili) are constructed of sporophylls arranged on a central axis. Microsporangia occur on the abaxial surface where pollen is formed and released (Norstog and Nicholls 1997). Female cones (megasporangiate strobili) occur in two forms, differing between the Zamiaceae and Cycadaceae. Zamiaceae produce cones with peltate (shield-like) megasporophylls arranged around a central axis (Norstog and Nicholls 1997). Ovules, which are attached to the megasporophylls, are oriented with

their micropyles towards the cone axis (Tang 1993). Megasporophylls are held tightly together, closing off the cone interior from the environment. Only when ovules become receptive at pollination do the megasporophylls separate (Tang 1987). After pollination, these fissures close and the cone interior is essentially sealed from the environment once again. The female reproductive structures of *Cycas* spp. (Cycadaceae) are described as indeterminate strobili (Stevenson 1990). They consist of whorls of leaf-shaped megasporophylls that surround and cover the stem apex. The megasporophylls are held tightly together and thus resemble a cone (Norstog and Nicholls 1997). At receptivity, the cones of some species 'bloom' and the ovules, attached to the now recurved megasporophylls, are exposed to the surrounding environment. In other species, the cone-shape persists until seed is shed. In all species, a new flush of vegetative leaves eventually emerges from the stem apex (Norstog and Nicholls 1997).

Until the 1980s, it was accepted that cycads were wind-pollinated, despite the fact that many insects had been observed in close association with cycads since the early 1900s (reviewed by Tang 1987). In the 1980s, Norstog et al. (1986) and Tang (1987) used exclusion experiments, in which either pollen grains borne by wind or by insect were prevented from reaching cones, to show that *Zamia furfuracea* L. f. and *Z. pumila* L. required beetles (*Rhopalotria mollis* (Sharp, 1890), and *Pharaxonotha zamiae* (Blake, 1928) and *Rhopalotria slossoni* (Schaeffer, 1905) respectively) for successful pollination. Since then, exclusion experiments on other cycads have demonstrated similar insect associations: *Stangeria* (Proches and Johnson 2009); *Lepidozamia* (Hall et al. 2004); *Bowenia* (Wilson 2002); *Encephalartos* (Donaldson 1997); *Macrozamia* (Terry 2001). It is thought that *Cycas* spp., which have exposed

ovules, are at least facultatively dependent on insects (Kono and Tobe 2007). However, wind cannot be ruled out completely (Kono and Tobe 2007). It is now accepted that the interactions between cycads and their insect pollinators are complex. Cycads may provide a place for feeding, mating and pupation (Tang 2004). The plants may influence insect behaviour by complex cues, e.g. production of volatiles and heat within their cones (Seymour et al. 2004; Terry et al. 2004; Terry et al. 2007).

During the period of pollen receptivity, pollination drops appear on the micropylar tips of cycad ovules (Figure 7). Pollen is captured and transported into the ovule by these drops. Pollination drops have been observed in at least six genera of cycads: *Cycas* (Ikeno 1898; Tang 1993, 1995), *Zamia* (Webber 1897; Tang 1987, 1993, 1995), *Dioon*, *Macrozamia*, *Encephalartos*, and *Ceratozamia* (Tang 1993, 1995). They range in volume from 0.01 to 0.77 µL, depending on the diameter of the micropyle in a given species (Tang 1987). The ovules of receptive *Cycas* cones are oriented away from the stem axis, and their pollination drops are exposed to the environment. The ovules of Zamiaceae face the cone axis. Their pollination drops are hidden inside the cone.

Tang (1987, 1993) completed the only biochemical analyses of cycad drops to date. He found sugars in *Z. pumila*, *Ceratozamia robusta* Miq., *Dioon spinulosum* Dyer ex Eichl., and *Encephalartos ferox* G. Bertol at concentrations ranging from 5 - 12 %. Further analysis identified fructose, glucose and sucrose in drops of *Z. pumila*, and sucrose and glucose in those of *C. robusta*. Free amino acids were also found. In *Z. pumila*, Tang (1987) reported proline, alanine and glycine, ranging from 0.2-1.6 µL/mL. In *C. robusta*, proline, alanine, glycine, cysteine, glutamine, leucine,

Figure 7. Pollination drops in cycads. **A**. *Cycas rumphii* Miq.female plant. **B**. Female cone of *C. rumphii* during receptivity to pollination. **C**. Pollination drop on an ovule of *C. rumphii*. **D**. Receptive female cone of *Ceratozamia kuesteriana* Regel. **E**. Interior of a female *C. hildae* G.P. Landry & M.C. Wilson cone bearing ovules with pollination drops. **F**. *C. hildae* megasporophyll with two ovules, one secreting a pollination drop. **G**. Pollination drop exuded from the micropyle of *C. kuesteriana*.

lysine, methionine, phenylalanine, serine and tyrosine were present at 1.2 µL/mL. Phenols were also present in *C. robusta* drops.

Study species

We hypothesized that like all conifers and gnetalean species analyzed so far, pollination drops of cycads would also contain proteins. We collected pollination drops for proteomic analysis from three species of cycads: *Zamia furfuracea*, *Ceratozamia hildae* G. P. Landry & M. C. Wilson, and *Cycas rumphii* Miq.

Zamia furfuracea is endemic to Mexico. It is a popular landscape plant in warm climates, where it is commonly known as the cardboard palm. In Florida, *Z. furfuracea* is pollinated by the weevil *Rhopalotria mollis* (Sharp, 1890) (Norstog et al. 1986). The association of *R. mollis* and *Z. furfuracea* is well-studied. *Rhopalotria mollis* swarm male cones and feed on the starch-filled sporophylls. Mating, oviposition and pupation all occur in male cones. Weevils visit female cones, carrying pollen along with them, but do not appear to feed or reproduce within female cones. Norstog and Fawcett (1989) suggested that the size, shape and odour of female cones are similar enough to those of male cones that weevils may be attracted to visit. The lack of starch in megasporophylls and higher levels of toxins present in female cones may prevent weevils from settling. It is unknown if weevils feed on, or are attracted to, the pollination drops in female cones.

Ceratozamia hildae is also native to Mexico. Although common in the horticultural trade, where it is sold as the bamboo cycad, it is nearly extinct in the wild (Norstog and Nicholls 1997). No information specific to pollination of *C. hildae* is available, but pollination events in other *Ceratozamia* spp. have been observed. Vovides (1991) reported that languriid beetles associate with *C. mexicana* Brongn at pollination.

Sánchez-Rotonda et al. (1995) conducted exclusion experiments and found that *Pharaxonotha* sp. beetles were responsible for pollination of *C. mexicana* in the wild.

Natural populations of *Cycas rumphii* occur on the islands of Indonesia and Papua New Guinea (IUCN Red List of Threatened Species website). In cultivation, *C. rumphii* is known as King Sago palm. Anecdotal evidence suggests that pollination in *C. rumphii* is at least facultatively reliant on insect pollinators, e.g. there are reports that *C. rumphii* has low seed yield in botanical gardens despite male and female plants being located near to one another (Norstog and Nicholls 1997). Other *Cycas* spp. may have both wind- and insect-pollination, e.g. *C. micronesica* K.D. Hill is pollinated by wind, *Anatrachyntis* moths and nitidulid beetles (Terry et al. 2009); *C. revoluta* Thunb. is pollinated by wind and nitidulid beetles (Kono and Tobe 2007).

Zooidogamy and pollen tube function

Cycad pollen is drawn into the ovule through the micropyle as the pollination drop retracts, and comes to rest in the pollen chamber, a pit formed in the nucellus by degradation of nucellar cells (Singh 1978; Choi and Friedman 1991; Norstog and Nicholls 1997). Pollen germinates in the remnants of the pollination drop (Choi and Friedman 1991). The pollen tube cell expands and emerges from the sulcus of the pollen grain (Pettitt 1982). The pollen exine is not shed. Within a week, the pollen tube breaches the nucellar cells of the ovule (Choi and Friedman 1991). It does not grow towards the archegonia, but grows into the subepidermal layers of the nucellus. At first this growth is intercellular, but eventually the pollen tube acts as a haustorium that grows into cells and gathers nutrients to support development of motile gametes within the pollen grain (Choi and Friedman 1991). Many months later, the proximal end of the pollen tube (the part

nearest to the original pollen grain) grows through the remaining cells of the nucellar inner-epidermis that separate the pollen chamber from the archegonial chamber, and releases motile gametes that swim to the archegonia and fertilize the egg cell (Norstog and Nicholls 1997).

These events are typical of species with sperm-containing pollen. Cycads, along with *Ginkgo*, are the only extant zooidogamous seed plants. Their pollen tubes function primarily in nutrient acquisition, rather than delivery of gametes like those of conifers, Gnetales and all angiosperms, which are siphonogamous. It is thought that zooidogamy in cycads and *Ginkgo* must represent an ancestral remnant of land-plant reproduction (Chamberlain 1935).

Here, we report the results of proteomic analyses on pollination drops of three species of cycads: *Z. furfuracea*, *Ceratozamia hildae* and *Cycas rumphii.* Cycads present a unique combination of pollination drop-related reproductive characteristics, including haustorial pollen tubes, insect-pollination, and pollen chambers. We hypothesize that their characteristic reproductive secretions will be different from those of conifers or Gnetales, which are gymnosperms with different pollination and fertilization mechanisms. We used our proteomic surveys to detect evidence of potential interactions between pollination drops, pollen and the environment during pollination.

Methods

Collection of pollination drops from cycads

Pollination drops were collected from the extensive cycad collection at the Montgomery Botanical Center in Coral Gables, Florida. Receptive *Ceratozamia hildae* cones were collected May 11 2011. An additional accession was collected from William Tang's garden in Miami. Receptivity in *C. hildae* is characterized by the separation of

sporophylls at the base of the cone, and the slight exposure of the pink-coloured ovules. Receptive *Z. furfuracea* cones were collected mainly from the Montgomery Botanical Center July 15-19 2011 and July 23-August 7 2012. Additional accessions were collected from the campus of Florida International University, the grounds of the United States Department of Agriculture Subtropical Horticulture Research Station and the garden of Mrs. Elaine Spears in Miami. Receptivity in *Z. furfuracea* is indicated by a vertical fissure between sporophylls, and horizontal fissures at the top and bottom of the cone. Cones were collected between 6 am and 10 am.

Megasporophylls of *Z. furfuracea* and *C. hildae* were snapped off of the central cone axis and placed onto moistened paper towel set between two plastic plates in order to maintain a humid environment. Drops were collected using a 10 μL glass capillary tube that was drawn over a flame to create a fine tip. Care was taken to reduce keratin contamination. Drops were collected throughout the day as they appeared on the ovules for up to two days after cone dissection. Collections were pooled between days for *Z. furfuracea* and *C. hildae.*

Drops were collected from *Cycas rumphii* on May 16 2011. Receptive *C. rumphii* cones displayed recurved megasporophylls that felt loose to the touch. Since the ovules of *C. rumphii* are exposed at pollination, drops could be collected directly from the cone. Drops were collected with a 10 μL glass capillary tube that was drawn to a fine tip. Collections were made between 6 am and 8 am.

Samples were expelled into 1.5 mL microcentrifuge tubes and stored at -20 °C until analysis. Proteomic analyses were carried out at the University of Victoria Genome BC Proteomics Centre (details of proteomics methods are presented in Chapter 3).

Results

Proteins were detected in the pollination drops of all three species of cycads analyzed: *Z. furfuracea* (Tables 5, pgs 69-72; 6, pgs 73-76), *Ceratozamia hildae* (Table 7, pgs 77-81) and *Cycas rumphii* (Table 8, pgs 82-83). The program PEAKS 6 reports a number of proteins together as a protein group when its algorithm determines that there are two or more proteins that explain a group of peptides equally well. For the purposes of discussion, one representative protein was selected from each group to be included in the results tables (by highest PEAKS –log10P score, and then by lowest BLAST e value). Careful scrutiny of numerous protein groups in this study revealed that all proteins in a group unwaveringly identify equivalent proteins – in most cases all Gymno_DB transcripts in a protein group BLASTed to the same TAIR gene model. The complete dataset, including sequence data for each equivalent protein, is maintained in electronic format but is unsuitable for paper format. From this point onwards, protein groups will be referred to as proteins. The proteins reported do not include hits to common contaminants.

The number of proteins detected differed among the three species. *Ceratozamia hildae* had 69 proteins (Table 7). Of those, 61 were annotated with good quality BLAST hits (e value < -5) to the TAIR10 database. The remaining eight did not have a significant match to TAIR10. A list of all proteins and peptides identified for *C. hildae* is provided as an example (Appendix 2). *Z. furfuracea* drops had fewer proteins than those of *C. hildae*. Similar numbers of proteins were detected in collections from two separate sample years. The 2011 sample contained 43 proteins, 40 of which had good TAIR10 matches (Table 5), while the 2012 sample had 41 protein identifications with 37

Table 5. Proteins identified in *Zamia furfuracea* (2011) pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model.

Accession	-10lgP	Total Peptides	Unique Peptides	TAIR10 Gene Model	TAIR10 Description	BLASTP e value	Gene Ontology Biological Process	Gene Ontology Cellular Component
picea_abies_isotig11461_30	236.15	13	1	AT3G54420.1	homolog of carrot EP3-3 chitinase	1e-96	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region; cell wall
Dougfir-megastigmus-comp75096_c0_seq1_5	233.63	13	8					
scaffold-JBND-2003172-Pinus_ponderosa_sub2_5	232.92	13	1	AT2G43590.1	Chitinase family protein	6e-49	cell wall macromolecule catabolic process; carbohydrate metabolic process	extracellular region
scaffold-GNQG-2006729-Encephalartos_barteri_single_18	227.91	14	13	AT3G54420.1	homolog of carrot EP3-3 chitinase	1e-85	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region; cell wall
scaffold-GNQG-2014022-Encephalartos_barteri_sub2_17	185.99	8	7	AT5G08370.1	alpha-galactosidase 2	0	carbohydrate metabolic process	extracellular region
scaffold-GNQG-2009885-Encephalartos_barteri_single_23	177.53	9	8	AT1G20160.1	Subtilisin-like serine endopeptidase family protein	0	negative regulation of catalytic activity; proteolysis	apoplast; extracellular region
Pinus_taeda_isotig10246_40	175.96	8	8	AT3G54420.1	homolog of carrot EP3-3 chitinase	4e-76	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region; cell wall

Cycas-rumphii-NODE_6535_length_896_cov_3.635045_11	173.81	6	6					
scaffold-AIGO-2003467-Chamaecyparis_lawsoniana_single_10	172.02	8	7	AT4G11650.1	osmotin 34	1e-79	response to salt stress	extracellular region
Pinus_taeda_isotig17261_8	163.29	5	5	AT2G43590.1	Chitinase family protein	1e-23	cell wall macromolecule catabolic process; carbohydrate metabolic process	extracellular region
scaffold-GNQG-2083102-Encephalartos_barteri_single_11	156.31	5	5	AT2G43610.1	Chitinase family protein	5e-91	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region
scaffold-GNQG-2085284-Encephalartos_barteri_single_32	133.69	3	3	AT5G50260.1	Cysteine proteinases superfamily protein	2e-150	proteolysis	extracellular region
Dougfir-megastigmus-comp587304_c1_seq5_261	122.34	3	3	AT4G02930.1	GTP binding Elongation factor Tu family protein	0	translational elongation	mitochondrion; intracellular; cell wall
scaffold-GKCZ-2009082-Diselma_archeri_single_14	119.95	2	2	AT2G03200.1	Eukaryotic aspartyl protease family protein	3e-144	proteolysis	extracellular region
Pinus_taeda_isotig01953_43	119.76	3	2	AT4G11650.1	osmotin 34	9e-79	response to salt stress	extracellular region
scaffold-GNQG-2013990-Encephalartos_barteri_sub2_11	119.12	3	3	AT3G14240.1	Subtilase family protein	1e-70	proteolysis; negative regulation of catalytic activity	
Gnetum_gnom_isotig10747_19	112.54	3	3	AT3G12500.1	basic chitinase	4e-145	cell wall macromolecule catabolic process; response to cadmium ion; carbohydrate metabolic process	vacuolar membrane; cytosol; extracellular region; vacuole
scaffold-ETCJ-2058429-Pilgerodendron_uviferum_single_28	111.36	2	1	AT3G54420.1	homolog of carrot EP3-3 chitinase	4e-93	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region; cell wall
scaffold-GNQG-2014246-Encephalartos_barteri_sub2_5	110.8	2	2	AT2G46750.1	D-arabinono-1,4-lactone oxidase family protein	7e-39		membrane; extracellular region

scaffold-GNQG-2009884-Encephalartos_barteri_single_24	103.37	3	2	AT1G20160.1	Subtilisin-like serine endopeptidase family protein	0	negative regulation of catalytic activity; proteolysis	apoplast; extracellular region
scaffold-BTTS-2012067-Austrotaxus_spicata_single_10	101.84	2	2	AT4G11650.1	osmotin 34	2e-88	response to salt stress	extracellular region
scaffold-EGLZ-2040097-Prumnopitys_andina_single_55	99.86	1	1	AT1G05590.1	beta-hexosaminidase 2	0	carbohydrate metabolic process	extracellular region
scaffold-CDFR-2000584-Manoao_colensoi_single_14	98.62	2	1	AT5G08370.1	alpha-galactosidase 2	0	carbohydrate metabolic process	extracellular region
scaffold-EGLZ-2040320-Prumnopitys_andina_single_69	98.28	2	1	AT1G20160.1	Subtilisin-like serine endopeptidase family protein	0	negative regulation of catalytic activity; proteolysis	apoplast; extracellular region
scaffold-GNQG-2083444-Encephalartos_barteri_single_20	95.78	2	2	AT5G44400.1	FAD-binding Berberine family protein	2e-96	oxidation-reduction process	plasmodesma; cytoplasm; cell wall
Pinus_taeda_isotig20310_55	89.73	2	1	AT3G13790.1	Glycosyl hydrolases family 32 protein	0	carbohydrate metabolic process	extracellular region; cell wall
scaffold-FRPM-2003461-Calocedrus_decurrens_sub2_22	89.6	2	1	AT1G56120.1	Leucine-rich repeat transmembrane protein kinase	1e-96	protein phosphorylation	plasma membrane
scaffold-JDQB-2003689-Neocallitropsis_pancheri_single_9	88.75	2	1	AT4G11650.1	osmotin 34	2e-69	response to salt stress	extracellular region
scaffold-GNQG-2014464-Encephalartos_barteri_single_13	87.47	2	2	AT3G13790.1	Glycosyl hydrolases family 32 protein	7e-135	carbohydrate metabolic process	extracellular region; cell wall
scaffold-AREG-2012104-Nothotsuga_longibracteata_single_57	86.21	1	1	AT3G13790.1	Glycosyl hydrolases family 32 protein	0	carbohydrate metabolic process	extracellular region; cell wall
scaffold-GNQG-2074230-Encephalartos_barteri_single_12	84.66	2	2	AT5G64570.1	beta-D-xylosidase 4	1e-45	carbohydrate metabolic process	cell wall; apoplast; extracellular region
scaffold-GNQG-2014246-Encephalartos_barteri_sub1_6	84.08	2	2	AT5G56490.1	D-arabinono-1,4-lactone oxidase family protein	3e-52	oxidation-reduction process	extracellular region; membrane
scaffold-GNQG-2074716-Encephalartos_barteri_single_7	81.19	2	2					
scaffold-EGLZ-2039129-Prumnopitys_andina_single_40	79.75	2	2	AT2G46740.1	D-arabinono-1,4-lactone oxidase family protein	2e-143		extracellular region; cell wall; membrane

scaffold-GNQG-2006048-Encephalartos_barteri_single_92	76.03	1	1	AT5G12950.1	Putative glycosyl hydrolase of unknown function (DUF1680)	0		vacuole; plant-type cell wall; extracellular region
scaffold-EGLZ-2002349-Prumnopitys_andina_single_12	74.77	2	2	AT3G54420.1	homolog of carrot EP3-3 chitinase	2e-91	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region; cell wall
Pseudotsuga_menz_isotig22337_24	73.7	1	1	AT2G03200.1	Eukaryotic aspartyl protease family protein	9e-107	proteolysis	extracellular region
scaffold-CDFR-2064691-Manoao_colensoi_single_56	72.9	1	1	AT5G67360.1	Subtilase family protein	0	negative regulation of catalytic activity; proteolysis	apoplast; extracellular region; cell wall; plant-type cell wall
scaffold-AIGO-2001928-Chamaecyparis_lawsoniana_single_10	72.37	1	1	AT3G12500.1	basic chitinase	4e-70	cell wall macromolecule catabolic process; response to cadmium ion; carbohydrate metabolic process	vacuolar membrane; cytosol; extracellular region; vacuole
scaffold-DSXO-2069887-Cryptomeria_japonica_single_16	71.81	2	2	AT1G75030.1	thaumatin-like protein 3	4e-67		extracellular region
Ginkgo_Contig62498_7	65.68	1	1	AT2G43590.1	Chitinase family protein	2e-38	cell wall macromolecule catabolic process; carbohydrate metabolic process	extracellular region
scaffold-EGLZ-2040499-Prumnopitys_andina_single_75	61.58	1	1	AT5G55480.1	SHV3-like 1	0	glycerol metabolic process; lipid metabolic process	plasmodesma; anchored to plasma membrane; plasma membrane
scaffold-JDQB-2072413-Neocallitropsis_pancheri_single_28	58.6	1	1	AT5G64570.1	beta-D-xylosidase 4	0	carbohydrate metabolic process	cell wall; apoplast; extracellular region

Table 6. Proteins identified in *Zamia furfuracea* (2012) pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model.

Accession	-10lgP	Total Peptides	Unique Peptides	TAIR10 Gene Model	TAIR10 Description	BLASTP e value	Gene Ontology Biological Process	Gene Ontology Cellular Component
Ginkgo_Contig71315_26	294.57	3	3	AT1G13950.1	eukaryotic elongation factor 5A-1	2e-91	positive regulation of translational termination; positive regulation of translational elongation; peptidyl-lysine modification to hypusine; translational frameshifting	cytoplasm
Ginkgo_Contig76970_32	293.77	3	3	AT1G13950.1	eukaryotic elongation factor 5A-1	3e-96	positive regulation of translational termination; positive regulation of translational elongation; peptidyl-lysine modification to hypusine; translational frameshifting	cytoplasm
scaffold-GNQG-2000199-Encephalartos_barteri_sub2_11	280.83	11	9	AT3G12500.1	basic chitinase	1e-133	cell wall macromolecule catabolic process; response to cadmium ion; carbohydrate metabolic process	vacuolar membrane; cytosol; extracellular region; vacuole
scaffold-GNQG-2009884-Encephalartos_barteri_single_24	257.41	7	7	AT1G20160.1	Subtilisin-like serine endopeptidase family protein	0	negative regulation of catalytic activity; proteolysis	apoplast; extracellular region
Ginkgo_Contig77135_24	237.73	2	2	AT1G70890.1	MLP-like protein 43	2e-05	defense response	chloroplast

scaffold-GNQG-2086986-Encephalartos_barteri_single_30	233.74	5	5	AT5G20950.1	Glycosyl hydrolase family protein	0	carbohydrate metabolic process	cell wall; membrane; extracellular region; plasmodesma; plant-type cell wall
scaffold-GNQG-2009885-Encephalartos_barteri_single_23	215.48	10	10	AT1G20160.1	Subtilisin-like serine endopeptidase family protein	0	negative regulation of catalytic activity; proteolysis	apoplast; extracellular region
scaffold-GNQG-2083444-Encephalartos_barteri_single_20	203.8	3	3	AT5G44400.1	FAD-binding Berberine family protein	2e-96	oxidation-reduction process	plasmodesma; cytoplasm; cell wall
Ginkgo_Contig67748_9	200	1	1	AT2G39050.1	hydroxyproline-rich glycoprotein family protein	1e-56		nucleus
Ginkgo_Contig76756_14	200	1	1	AT5G40370.2	Glutaredoxin family protein	1e-23	cell redox homeostasis	Golgi apparatus; cytoplasm; plasma membrane
scaffold-GNQG-2014022-Encephalartos_barteri_sub2_17	193.55	5	5	AT5G08370.1	alpha-galactosidase 2	0	carbohydrate metabolic process	extracellular region
Gnetum_gnom_isotig10747_19	179.08	3	3	AT3G12500.1	basic chitinase	4e-145	cell wall macromolecule catabolic process; response to cadmium ion; carbohydrate metabolic process	vacuolar membrane; cytosol; extracellular region; vacuole
scaffold-GGEA-2013202-Cedrus_libani_sub1_10	177.53	1	1	AT1G13950.1	eukaryotic elongation factor 5A-1	4e-98	positive regulation of translational termination; positive regulation of translational elongation; peptidyl-lysine modification to hypusine; translational frameshifting	cytoplasm
Podocarpus_macro_isotig12653_12	166.61	1	1	AT5G60390.1	GTP binding Elongation factor Tu family protein	0	translational elongation	cytoplasm; vacuole; plasma membrane; plasmodesma

scaffold-GNQG-2014246-Encephalartos_barteri_sub2_5	152.2	3	3	AT2G46750.1	D-arabinono-1,4-lactone oxidase family protein	7e-39		membrane; extracellular region
Ginkgo_Contig77189_25	151.26	1	1	AT2G16600.1	rotamase CYP 3	8e-108	protein folding	plasma membrane; Golgi apparatus; chloroplast; cytosol; cytoplasm; plasmodesma
scaffold-GNQG-2074230-Encephalartos_barteri_single_12	149.49	2	2	AT5G64570.1	beta-D-xylosidase 4	1e-45	carbohydrate metabolic process	cell wall; apoplast; extracellular region
Pinus_taeda_isotig17261_8	144.04	3	3	AT2G43590.1	Chitinase family protein	1e-23	cell wall macromolecule catabolic process; carbohydrate metabolic process	extracellular region
scaffold-GNQG-2014464-Encephalartos_barteri_single_13	141.86	2	2	AT3G13790.1	Glycosyl hydrolases family 32 protein	7e-135	carbohydrate metabolic process	extracellular region; cell wall
scaffold-FHST-2063328-Taxodium_distichum_single_31	134.98	4	1	AT3G12500.1	basic chitinase	6e-110	cell wall macromolecule catabolic process; response to cadmium ion; carbohydrate metabolic process	vacuolar membrane; cytosol; extracellular region; vacuole
scaffold-IZGN-2072338-Dacrydium_balansae_single_33	132.36	1	1	AT1G05590.1	beta-hexosaminidase 2	4e-161	carbohydrate metabolic process	extracellular region
scaffold-JRNA-2004820-Phyllocladus_hypophyllus_single_88	115.21	2	2	AT5G64570.1	beta-D-xylosidase 4	0	carbohydrate metabolic process	cell wall; apoplast; extracellular region
scaffold-GNQG-2000309-Encephalartos_barteri_single_16	114.63	1	1	AT3G51030.1	thioredoxin H-type 1	2e-47	glycerol ether metabolic process; cell redox homeostasis	cytoplasm
scaffold-GNQG-2083102-Encephalartos_barteri_single_11	107.9	1	1	AT2G43610.1	Chitinase family protein	5e-91	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region
scaffold-GKCZ-2009082-Diselma_archeri_single_14	103.71	2	2	AT2G03200.1	Eukaryotic aspartyl protease family protein	3e-144	proteolysis	extracellular region

scaffold-GNQG-2014246-Encephalartos_barteri_sub1_6	101.92	1	1	AT5G56490.1	D-arabinono-1,4-lactone oxidase family protein	3e-52	oxidation-reduction process	extracellular region; membrane
scaffold-GGEA-2067891-Cedrus_libani_single_37	98.07	1	1	AT1G53280.1	Class I glutamine amidotransferase-like superfamily protein	4e-167		chloroplast stroma; chloroplast
scaffold-HQOM-2001973-Torreya_nucifera_sub1_13	96.31	2	1	AT3G12500.1	basic chitinase	5e-141	cell wall macromolecule catabolic process; response to cadmium ion; carbohydrate metabolic process	vacuolar membrane; cytosol; extracellular region; vacuole
scaffold-GGEA-2068436-Cedrus_libani_single_84	96.16	2	2	AT2G28470.1	beta-galactosidase 8	0	carbohydrate metabolic process	extracellular region; cell wall
scaffold-BTTS-2079521-Austrotaxus_spicata_single_61	93.52	2	2	AT1G20160.1	Subtilisin-like serine endopeptidase family protein	0	negative regulation of catalytic activity; proteolysis	apoplast; extracellular region
picea_abies_isotig06000_43	74.73	1	1	AT4G30610.1	alpha/beta-Hydrolases superfamily protein	4e-38	proteolysis	extracellular region
scaffold-IFLI-2003267-Callitris_gracilis-March_single_17	72.45	1	1	AT5G13980.1	Glycosyl hydrolase family 38 protein	3e-170	carbohydrate metabolic process; mannose metabolic process	extracellular region; vacuolar membrane; plant-type cell wall; vacuole; apoplast; cell wall
scaffold-GTHK-2008154-Gnetum_montanum_single_37	65.05	1	1	AT1G14410.1	ssDNA-binding transcriptional regulator	2e-61		nucleoid; chloroplast
scaffold-GNQG-2014697-Encephalartos_barteri_single_51	60.05	1	1	AT2G46750.1	D-arabinono-1,4-lactone oxidase family protein	4e-150		membrane; extracellular region
scaffold-ETCJ-2064873-Pilgerodendron_uviferum_single_64	59.67	1	1	AT1G31870.1		4e-99		nucleus
scaffold-ESYX-2002314-Cunninghamia_lanceolata-branch_apex_with_needles_single_23	52.1	1	1	AT3G13790.1	Glycosyl hydrolases family 32 protein	0	carbohydrate metabolic process	extracellular region; cell wall
Pinus_taeda_isotig33731_28	51.93	1	1	AT5G67360.1	Subtilase family protein	0	negative regulation of catalytic activity; proteolysis	apoplast; extracellular region; cell wall; plant-type cell wall

Table 7. Proteins identified in *Ceratozamia hildae* pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model.

Accession	-10lgP	Total Peptides	Unique Peptides	TAIR10 Gene Model	TAIR10 Description	BLASTP e value	Gene Ontology Biological Process	Gene Ontology Cellular Component
scaffold-GNQG-2013496-Encephalartos_barteri_single_52	254.31	7	4	AT4G22010.1	SKU5 similar 4	0	oxidation-reduction process	plasmodesma; extracellular region; plant-type cell wall; membrane
scaffold-GNQG-2008016-Encephalartos_barteri_single_19	236.46	9	6	AT3G61490.1	Pectin lyase-like superfamily protein	0	carbohydrate metabolic process	
scaffold-GNQG-2014022-Encephalartos_barteri_sub2_17	232.38	8	8	AT5G08370.1	alpha-galactosidase 2	0	carbohydrate metabolic process	extracellular region
scaffold-GNQG-2011552-Encephalartos_barteri_single_14	222.46	6	3	AT5G10560.1	Glycosyl hydrolase family protein	1e-171	carbohydrate metabolic process	vacuole; extracellular region; vacuolar membrane
scaffold-GNQG-2015735-Encephalartos_barteri_single_28	198.03	7	7	AT2G28470.1	beta-galactosidase 8	0	carbohydrate metabolic process	extracellular region; cell wall
scaffold-AWQB-2000726-Picea_engelmanii_single_70	196.61	6	2	AT1G76160.1	SKU5 similar 5	0	oxidation-reduction process	plasmodesma; cell wall; apoplast; extracellular region; plant-type cell wall
scaffold-GNQG-2006873-Encephalartos_barteri_single_24	194.36	7	7	AT2G27920.1	serine carboxypeptidase-like 51	2e-104	proteolysis	extracellular region
scaffold-GJTI-2006440-Cephalotaxus_harringtonia_single_31	183.22	6	1					
scaffold-GNQG-2084473-Encephalartos_barteri_single_34	183	4	4	AT5G06480.1	Immunoglobulin E-set superfamily protein	4e-42		chloroplast
scaffold-GNQG-2086382-Encephalartos_barteri_single_12	153.9	3	3	AT2G45470.1	FASCICLIN-like arabinogalactan protein 8	1e-126		plant-type cell wall; plasma membrane; anchored to plasma membrane; apoplast

scaffold-IOVS-2056527-Pseudotsuga_menziesii_single_18	148.24	3	3	AT4G02290.1	glycosyl hydrolase 9B13	0	carbohydrate metabolic process	extracellular region
scaffold-GNQG-2013749-Encephalartos_barteri_single_15	143.3	4	4	AT1G13130.1	Cellulase (glycosyl hydrolase family 5) protein	0	carbohydrate metabolic process	chloroplast
Ginkgo_Contig28425_14	140.75	4	4	AT4G25980.1	Peroxidase superfamily protein	3e-132	response to oxidative stress; oxidation-reduction process	extracellular region
scaffold-AUDE-2008057-Widdringtonia_cedarbergensis_single_49	136.31	3	3	AT5G65760.1	Serine carboxypeptidase S28 family protein	0	proteolysis	vacuole; chloroplast
scaffold-HILW-2113257-Acmopyle_pancheri_single_45	133.53	2	2	AT3G45310.1	Cysteine proteinases superfamily protein	2e-172	proteolysis	extracellular region
scaffold-EFMS-2081589-Torreya_taxifolia_single_25	133.35	2	1	AT3G19620.1	Glycosyl hydrolase family protein	0	carbohydrate metabolic process	cell wall; extracellular region
scaffold-IFLI-2000858-Callitris_gracilis-March_single_27	131.29	2	1	AT1G76160.1	SKU5 similar 5	3e-128	oxidation-reduction process	plasmodesma; cell wall; apoplast; extracellular region; plant-type cell wall
scaffold-GNQG-2014884-Encephalartos_barteri_sub2_26	128.36	2	1	AT3G19620.1	Glycosyl hydrolase family protein	2e-157	carbohydrate metabolic process	cell wall; extracellular region
scaffold-FMWZ-2012745-Dacrycarpus_compactus_single_21	126.05	2	2	AT3G18080.1	B-S glucosidase 44	0	carbohydrate metabolic process	plant-type cell wall; cytosolic ribosome; cell wall; chloroplast
picea_abies_isotig06000_43	124.58	3	3	AT4G30610.1	alpha/beta-Hydrolases superfamily protein	4e-38	proteolysis	extracellular region
scaffold-ETCJ-2059606-Pilgerodendron_uviferum_single_11	120.59	2	1	AT5G44360.1	FAD-binding Berberine family protein	5e-97	oxidation-reduction process	cytoplasm
scaffold-GAMH-2009786-Tsuga_heterophylla_single_29	120.41	2	1	AT5G10560.1	Glycosyl hydrolase family protein	0	carbohydrate metabolic process	vacuole; extracellular region; vacuolar membrane
scaffold-GNQG-2081815-Encephalartos_barteri_single_4	117.81	2	2					
scaffold-GNQG-2010926-Encephalartos_barteri_single_38	117.64	2	1	AT5G63810.1	beta-galactosidase 10	0	carbohydrate metabolic process	cell wall; extracellular region; plant-type cell wall
scaffold-JDQB-2002660-Neocallitropsis_pancheri_sub1_10	107.45	3	2	AT4G22010.1	SKU5 similar 4	3e-136	oxidation-reduction process	plasmodesma; extracellular region; plant-type cell wall; membrane
Ginkgo_Contig47182_59	105.76	2	1	AT1G30760.1	FAD-binding Berberine family protein	1e-143	oxidation-reduction process	cytoplasm

scaffold-AWQB-2000846-Picea_engelmanii_sub1_26	105.25	2	1	AT1G68560.1	alpha-xylosidase 1	0	carbohydrate metabolic process	chloroplast; plant-type cell wall; plasmodesma; extracellular region; cell wall
scaffold-AQFM-2079970-Pseudolarix_amabilis_single_25	104.32	2	2	AT3G12500.1	basic chitinase	1e-148	cell wall macromolecule catabolic process; response to cadmium ion; carbohydrate metabolic process	vacuolar membrane; cytosol; extracellular region; vacuole
Pinus_taeda_contig32485_14	102.67	2	2	AT1G20160.1	Subtilisin-like serine endopeptidase family protein	6e-116	negative regulation of catalytic activity; proteolysis	apoplast; extracellular region
Ginkgo_Contig35216_33	97.92	2	1	AT5G63810.1	beta-galactosidase 10	0	carbohydrate metabolic process	cell wall; extracellular region; plant-type cell wall
scaffold-ESYX-2016725-Cunninghamia_lanceolata-branch_apex_with_needles_single_67	97.7	2	1	AT1G68560.1	alpha-xylosidase 1	0	carbohydrate metabolic process	chloroplast; plant-type cell wall; plasmodesma; extracellular region; cell wall
Pseudotsuga_menz_isotig03151_34	97.62	2	1	AT5G61250.2	glucuronidase 1	7e-106		extracellular region; membrane
Ginkgo_Contig33358_18	97.59	2	2	AT4G09740.1	glycosyl hydrolase 9B14	0	carbohydrate metabolic process	extracellular region
Dougfir-megastigmus-comp462464_c0_seq1_21	97.15	1	1	AT3G62110.1	Pectin lyase-like superfamily protein	0	carbohydrate metabolic process	vacuole; extracellular region
Pinus_taeda_isotig07390_32	95.34	2	1	AT5G50260.1	Cysteine proteinases superfamily protein	9e-133	proteolysis	extracellular region
Dougfir-megastigmus-comp535867_c0_seq3_4	90.46	1	1					
scaffold-GAMH-2006106-Tsuga_heterophylla_single_17	90.38	1	1	AT2G32300.1	uclacyanin 1	7e-27		plasma membrane
scaffold-GNQG-2083101-Encephalartos_barteri_single_20	90.05	1	1	AT2G38540.1	lipid transfer protein 1	2e-24	lipid transport	chloroplast thylakoid; extracellular region; apoplast
scaffold-DSXO-2059923-Cryptomeria_japonica_single_3	83.82	2	2	AT2G03200.1	Eukaryotic aspartyl protease family protein	3e-19	proteolysis	extracellular region
Ginkgo_Contig72727_21	82.01	1	1	AT5G64570.1	beta-D-xylosidase 4	0	carbohydrate metabolic process	cell wall; apoplast; extracellular region
Ginkgo_Contig47067_103	80.73	1	1	AT5G12950.1	Putative glycosyl hydrolase of unknown function (DUF1680)	0		vacuole; plant-type cell wall; extracellular region

scaffold-JRNA-2006442-Phyllocladus_hypohyllus_single_62	77.82	1	1	AT3G62110.1	Pectin lyase-like superfamily protein	0	carbohydrate metabolic process	vacuole; extracellular region
scaffold-GKCZ-2007628-Diselma_archeri_single_10	76.47	1	1	AT4G30610.1	alpha/beta-Hydrolases superfamily protein	2e-27	proteolysis	extracellular region
scaffold-GJTI-2002248-Cephalotaxus_harringtonia_single_8	75.33	1	1	AT4G09600.1	GAST1 protein homolog 3	7e-25		extracellular region
scaffold-EGLZ-2010263-Prumnopitys_andina_single_25	73.47	1	1	AT1G78060.1	Glycosyl hydrolase family protein	0	carbohydrate metabolic process	chloroplast; apoplast; cell wall; plasmodesma; extracellular region
Pinus_taeda_isotig17300_19	73.01	1	1	AT1G02850.4	beta glucosidase 11	1e-170	carbohydrate metabolic process	extracellular region
scaffold-IOVS-2040036-Pseudotsuga_menziesii_single_7	68.73	1	1	AT5G60360.3	aleurain-like protease	2e-22	proteolysis	extracellular region
scaffold-BBDD-2075959-Microstrobos_fitzgeraldii_single_30	65.45	1	1	AT4G25980.1	Peroxidase superfamily protein	2e-131	response to oxidative stress; oxidation-reduction process	extracellular region
scaffold-GNQG-2086713-Encephalartos_barteri_single_19	65.28	1	1	AT5G64570.1	beta-D-xylosidase 4	0	carbohydrate metabolic process	cell wall; apoplast; extracellular region
scaffold-GNQG-2004817-Encephalartos_barteri_single_28	64.85	1	1	AT2G27920.1	serine carboxypeptidase-like 51	5e-78	proteolysis	extracellular region
scaffold-IIOL-2000981-Pinus_parviflora_sub2_24	64.81	1	1	AT5G61460.1	P-loop containing nucleoside triphosphate hydrolases superfamily protein	0		nucleus
scaffold-HQOM-2013044-Torreya_nucifera_sub1_11	63.61	1	1	AT5G50260.1	Cysteine proteinases superfamily protein	2e-32	proteolysis	extracellular region
scaffold-ESYX-2074066-Cunninghamia_lanceolata-branch_apex_with_needles_single_8	62.53	1	1					
scaffold-CDFR-2059046-Manoao_colensoi_single_6	62.5	1	1	AT4G11650.1	osmotin 34	5e-70	response to salt stress	extracellular region
scaffold-GKCZ-2007555-Diselma_archeri_sub2_8	61.53	1	1	AT3G24480.1	Leucine-rich repeat (LRR) family protein	7e-118		cell wall; plasmodesma; extracellular region
Gnetum_gnom_isotig07407_83	59.66	1	1	AT4G36360.1	beta-galactosidase 3	0	carbohydrate metabolic process	extracellular region
Dougfir-megastigmus-comp560917_c0_seq1_10	58.25	1	1					
scaffold-ACWS-2003997-Arucaria_sp._single_13	57.83	1	1	AT4G35350.1	xylem cysteine peptidase 1	2e-119	proteolysis	nucleus; extracellular region

scaffold-IFLI-2135760-Callitris_gracilis-March_single_3	57.79	1	1	AT3G02130.1	receptor-like protein kinase 2	9e-24	protein phosphorylation	plasma membrane
scaffold-GGEA-2007079-Cedrus_libani_single_16	57.11	1	1	AT1G20160.1	Subtilisin-like serine endopeptidase family protein	7e-54	negative regulation of catalytic activity; proteolysis	apoplast; extracellular region
scaffold-FMWZ-2049673-Dacrycarpus_compactus_single_5	56.05	1	1	AT2G40370.1	laccase 5	2e-52	lignin catabolic process; oxidation-reduction process	apoplast; extracellular region
Pinus_lamb_isotig19184_50	54.45	1	1	AT4G37930.1	serine transhydroxymethyltransferase 1	0	response to cadmium ion	chloroplast; mitochondrion; chloroplast thylakoid; chloroplast stroma; cytosolic ribosome; apoplast; membrane; nucleus; plasma membrane
Ginkgo_Contig55712_69	54.19	1	1					
scaffold-CGDN-2000025-Tetraclinis_sp._single_65	52.85	1	1	AT1G20160.1	Subtilisin-like serine endopeptidase family protein	0	negative regulation of catalytic activity; proteolysis	apoplast; extracellular region
scaffold-IAJW-2026102-Amentotaxus_argotaenia_single_17	51.97	1	1					
scaffold-FMWZ-2002490-Dacrycarpus_compactus_single_28	50.9	1	1	AT5G50400.1	purple acid phosphatase 27	0		extracellular region
scaffold-IFLI-2026620-Callitris_gracilis-March_single_1	47.69	1	1					
scaffold-AIGO-2072006-Chamaecyparis_lawsoniana_single_34	46.35	1	1	AT1G68750.1	phosphoenolpyruvate carboxylase 4	0	tricarboxylic acid cycle; carbon fixation	cytoplasm
scaffold-GNQG-2016467-Encephalartos_barteri_single_82	46.11	1	1	AT5G55480.1	SHV3-like 1	0	glycerol metabolic process; lipid metabolic process	plasmodesma; anchored to plasma membrane; plasma membrane

Table 8. Proteins identified in *Cycas rumphii* pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model.

Accession	-10lgP	Total Peptides	Unique Peptides	TAIR10 Gene Model	TAIR10 Description	BLASTP e value	Gene Ontology Biological Process	Gene Ontology Cellular Component
Cycas-rumphii-NODE_22354_length_721_cov_71.540916_19	379.13	24	24					
scaffold-GNQG-2062526-Encephalartos_barteri_single_3	216.1	5	5					
Cycas-rumphii-NODE_22380_length_515_cov_12.669903_13	191.42	4	4	AT2G37870.1	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	3e-33	lipid transport	extracellular region; plasmodesma
Cycas-rumphii-NODE_22379_length_705_cov_13.903546_4	159.66	2	2	AT1G22690.1	Gibberellin-regulated family protein	2e-20		extracellular region
Cycas-rumphii-NODE_16812_length_350_cov_12.334286_1	149.7	2	2					
Dougfir-megastigmus-comp573804_c1_seq2_17	142.55	3	3	AT4G05050.1	ubiquitin 11	2e-105		
Cycas-rumphii-NODE_13037_length_72_cov_210.458328_2	120.41	1	1					
scaffold-HILW-2111690-Acmopyle_pancheri_single_10	118.95	1	1	AT5G24090.1	chitinase A	3e-118	carbohydrate metabolic process	extracellular region
scaffold-GAMH-2056163-Tsuga_heterophylla_single_20	100.43	1	1	AT2G47010.1		0		

scaffold-GNQG-2081815-Encephalartos_barteri_single_4	91.47	1	1					
scaffold-GJTI-2007921-Cephalotaxus_harringtonia_single_88	76.07	1	1	AT1G07990.1	SIT4 phosphatase-associated family protein	0		
scaffold-GKCZ-2031492-Diselma_archeri_single_4	68.47	1	1					
scaffold-JRNA-2006714-Phyllocladus_hypohyllus_single_16	67.28	1	1	AT5G05340.1	Peroxidase superfamily protein	4e-82	oxidation-reduction process; response to oxidative stress	extracellular region; apoplast; Golgi apparatus; cytosol; cell wall
scaffold-JBND-2065583-Pinus_ponderosa_single_4	64.95	1	1	AT5G59310.1	lipid transfer protein 4	7e-25	lipid transport	extracellular region
Pinus_taeda_isotig22456_12	64.73	1	1	AT5G05340.1	Peroxidase superfamily protein	9e-144	oxidation-reduction process; response to oxidative stress	extracellular region; apoplast; Golgi apparatus; cytosol; cell wall

good TAIR10 hits (Table 6). The *Cycas rumphii* sample had the smallest number of proteins at 15, with nine having good TAIR10 matches (Table 8).

The TAIR10 database includes Gene Ontology (GO) annotations in three categories: Molecular Function, Biological Process and Cellular Component. GO annotations use controlled vocabularies to describe genes and gene products (TAIR website). For example, Carbohydrate Metabolic Process and Proteolysis are controlled terms used in the Biological Process category. It is important to note that not all TAIR 10 gene models include GO annotations in all categories. Additionally, GO terms are not mutually exclusive; for example, a given protein may contain annotations for both intracellular and extracellular spaces. The GO annotations relating to Biological Process and Cellular Component for the proteins identified in cycad pollination drops are included in Tables 5-8.

Of those proteins with an acceptable BLAST match to the TAIR10 database, the majority contained an annotation for Cellular Component referring to the extracellular space: 82 % for *Ceratozamia hildae*; 67 % for *Cycas rumphii*; 93 % for *Z. furfuracea* (2011); 68 % for *Z. furfuracea* (2012). A smaller portion had an annotation for categories referring to the intracellular space (including GO terms for membrane, plasmodesma, vacuole): 44 % for *Ceratozamia hildae*; 33 % for *Cycas rumphii*; 23 % for *Z. furfuracea* (2011); 59 % for *Z. furfuracea* (2012). Few had an annotation specifically to the plasma membrane: 8 % for *Ceratozamia hildae*; 0 % for *Cycas rumphii*; 5 % *for Z. furfuracea* (2011); 8 % for *Z. furfuracea* (2012).

The best represented GO term for Biological Process in each cycad sample was Carbohydrate Metabolic Process: 38% *Ceratozamia hildae*; 11 % *Cycas rumphii*; 48 %

Z. furfuracea (2011); 41 % *Z. furfuracea* (2012). In all samples, except for *C. rumphii,* this was followed by Proteolysis: 23 % *Ceratozamia hildae*, 20 % *Z. furfuracea* (2011), 16 % *Z. furfuracea* (2012); 0 % *Cycas rumphii*. Except for *C. rumphii*, the third most common Biological Process was Oxidation-Reduction Process: 15% *Ceratozamia hildae*, 5 % *Z. furfuracea* (2011), 5 % *Z. furfuracea* (2012) and 22 % *Cycas rumphii*.

To provide a comparative overview between species and samples, a summary table is presented (Table 9) in which related proteins are grouped together by name/description. A number of carbohydrate-modifying proteins were present in the samples. All samples contained at least one chitinase. Three samples contained galactosidases and xylosidases. Two samples contained beta-hexosaminidases. Beta-glucosidases, a cellulase and pectin lyase-like proteins were each contained in separate samples. Three samples contained one or more proteins classified as glycosyl hydrolase family proteins, a broad category of carbohydrate-modifying proteins within which a number of the proteins listed above are grouped. Further probing of the TAIR accessions (gene models) under glycosyl hydrolase revealed that both *Z. furfuracea* samples contained glycosyl hydrolase family 32 proteins with descriptions that imply invertase activity, and one *Z. furfuracea* (2012) sample contained glycosyl hydrolase family 38 proteins with mannosidase activity.

A number of different types of proteolytic enzymes were detected within the drop samples. Three samples contained subtilisin-like serine endopeptidase family proteins and aspartyl protease family proteins. Two groups contained proteins identified from the cysteine proteinases superfamily, and one group additionally contained an aleurain-like protease with cysteine proteinase-like activity. Alpha/beta-hydrolase superfamily proteins

Table 9. Comparative view of protein types in four samples of cycad pollination drops (*Ceratozamia hildae*, two samples of *Zamia furfuracea* collected in 2011 and 2012, *Cycas rumphii*).

	Ceratozamia hildae	*Zamia furfuracea* 2011	*Zamia furfuracea* 2012	*Cycas rumphii*
Carbohydrate metabolic process:				
chitinases	X	X	X	X
galactosidases	X	X	X	
xylosidases	X	X	X	
glycosyl hydrolases	X	X	X	
beta-hexosaminidase		X	X	
cellulase (glycosyl hydrolase family 5) protein	X			
beta glucosidase	X			
pectin lyase-like superfamily protein	X			
Proteolysis:				
aspartyl protease	X	X	X	
subtilisin-like serine endopeptidase family protein	X	X	X	
cysteine proteinases superfamily protein	X	X		
alpha/beta-Hydrolases superfamily protein	X		X	
aleurain-like protease	X			
serine carboxypeptidases	X			
Oxidation-reduction process:				
FAD-binding Berberine family protein	X	X	X	
peroxidase superfamily protein	X			X
D-arabinono-1,4-lactone oxidase family protein		X	X	
laccase 5	X			
SKU5 similar	X			
Lipid transport:				
lipid transfer protein	X			X
Various extracellular function:				
thaumatin-like proteins / osmotin	X	X		
leucine-rich repeat (LRR) family protein	X			
GAST1 protein homolog 3	X			
purple acid phosphatase 27	X			
glucuronidase 1	X			
FASCICLIN-like arabinogalactan protein 8	X			
gibberellin-regulated family protein				X
Various intracellular function:				
SHV3-like 1	X	X		
GTP binding Elongation factor Tu family protein		X	X	
receptor-like protein kinase 2	X			
serine transhydroxymethyltransferase 1	X			
phosphoenolpyruvate carboxylase 4	X			
Immunoglobulin E-set superfamily protein	X			
uclacyanin 1	X			
P-loop containing nucleoside triphosphate hydrolases superfamily protein	X			
leucine-rich repeat transmembrane protein kinase		X		
Class I glutamine amidotransferase-like superfamily protein			X	
ssDNA-binding transcriptional regulator			X	
hydroxyproline-rich glycoprotein family protein			X	
glutaredoxin family protein			X	
MLP-like protein 43			X	
thioredoxin H-type 1			X	
eukaryotic elongation factor 5A-1			X	
rotamase CYP 3			X	
ubiquitin				X
SIT4 phosphatase-associated family protein				X

with serine-type carboxypeptidase activity were detected in two samples, and annotations specifically to serine carboxypeptidase-like 51 proteins and serine carboxypeptidase S28 family protein were found in the *Ceratozamia hildae* sample.

Five groups of proteins with GO biological process annotations to oxidation-reduction processes were detected. Three samples contained FAD-binding berberine family proteins; two contained peroxidase family proteins; two contained D-arabinono-1,4-lactone oxidase family proteins; one sample contained SKU5 similar proteins; one sample contained laccase family proteins.

A number of other GO molecular process categories were represented in the drop protein dataset. Proteins involved in lipid transport (lipid transfer proteins) were found in two samples; proteins involved in protein phosphorylation (receptor-like protein kinases and leucine-rich repeat transmembrane protein kinases) were found in separate samples. Glutaredoxin family proteins and thioredoxin H-type 1 proteins involved in cell redox homeostasis were found in *Z. furfuracea* 2012. Other categories represented include: defence response (MLP-like protein 43, *Z. furfuracea* 2012); glycerol metabolic process/lipid metabolic process (SHV3-like 1 protein, *C. hildae* and *Z. furfuracea* 2011); positive regulation of translational termination (eukaryotic elongation factor proteins, *Z. furfuracea* 2012); protein folding (rotomase CYP 3, *Z. furfuracea* 2012); response to cadmium ion (serine transhydroxymethyltransferase 1, *C. hildae*); translational elongation (GTP binding elongation factor Tu family protein, *Z. furfuracea* 2011 and 2012); tricarboxylic acid cycle (phosphoenolpyruvate carboxylase 4, *C. hildae*).

Amongst the proteins without GO biological process annotations, some were annotated to the extracellular space (thaumatin-like proteins/osmotins; leucine-rich repeat

family protein; GAST1 protein homolog 3; purple acid phosphatase 27; glucuronidase 1; fascilin-like arabinogalactan protein 8; gibberellin-regulated family protein) or the intracellular space (immunoglobulin E-set superfamily protein; uclacyanin 1; P-loop containing nucleoside triphosphate hydrolase; SIT4 phosphatase-associated family; ubiquitin; Class I glutamin amidotransferase-like superfamily protein; ssDNA-binding transcriptional regulator; hydroxyproline-rich glycoprotein family protein).

Discussion

Proteomic analyses revealed that proteins are present in the pollination drops of the three species of cycads analyzed: *Cycas rumphii*, *Z. furfuracea*, and *Ceratozamia hildae*. Cycads have a unique set of reproductive characteristics relating to pollen capture and germination. Many species require insect pollinators, and when pollen germinates, the initial growth of the pollen tube is not towards the egg-containing archegonia, but into the subepidermal layer of the nucellus where it absorbs nutrients to support development of motile gametes (Choi and Friedman 1991). Additionally, cycads undoubtedly form their own clade in the extant seed plants. Given the uniqueness of cycads, are the proteins in their pollination drops functionally equivalent to those identified in conifer or *Ephedra* pollination drops?

The gene ontology (GO) annotations for Cellular Component strongly suggest that like conifers and *Ephedra*, cycads have a drop secretome, a set of proteins that is secreted into the pollination drop. Between 67 % and 93 % of proteins were annotated to the extracellular space depending on the sample. Secreted proteins in pollination drops probably originate from the cells of the nucellus, the female sporogenous tissue. O'Leary et al. (2004) showed that arabinogalactan proteins localized to the nucellus in hybrid yew,

and that their peak production coincided with drop secretion. There is also evidence that some proteins may be released from the ovule integument (O'Leary 2004; Patrick von Aderkas, pers. com.). Localization of specific proteins of interest would be necessary to confirm the exact location of secretion in cycad species.

A substantial portion of drop proteins (23 % to 59 % depending on species) contained GO annotations linking them to intracellular locations. Although GO annotations may be derived from experimental (physical) evidence, many are inferred from computational pipelines (Lamesch et al.2012). It is possible that some of the proteins with intracellular annotations are in fact extracellular, but lack the signal peptide required for a predicted extracellular location. Such nontraditional secretion pathways have been observed in plants. Leaderless secretory proteins may explain greater than 50 % of secretome proteins in some plant systems (Agrawal et al. 2010). It is therefore possible that some of the proteins annotated to the intracellular space in cycad drops are in fact working in the extracellular space.

A different interpretation is that the presence of typically intracellular proteins is due to degradation of the nucellus tissue. Cycads, in common with other gymnosperms such as the Gnetales, *Ginkgo* and some conifers (e.g. *Cephalotaxus* spp.), have pollen chambers that form synchronously with pollination drop production (Singh 1978). As the nucellar cells degrade to form this cavity, their contents are washed into the pollination drop. Proteins that enter the drop this way should be referred to as degradome proteins. Degradome proteins are present in abundance in *Ephedra* drops (Chapter 2), but have not been identified in conifers to date. Whether or not degradome proteins can retain

intracellular functions in the extracellular chemical environment of pollination drops is unknown.

Many proteins involved in carbohydrate metabolic processes were identified in cycad pollination drops. These include enzymes previously identified in conifers and *Ephedra* spp. that were predicted to affect cell wall restructuring during pollen germination and pollen tube growth. Examples include alpha- and beta-galactosidases, beta-glucosidases and xylosidases. In addition, glucanases were found, which are not present in *Ephedra* spp. Poulis et al. (2005) suggested that enzymes, such as xylosidases and galactosidases may be capable of cleaving xyloglucan components which are present in the pollen cell wall of Douglas-fir. By loosening cell walls, these enzymes may contribute to successful pollen germination. Wagner et al. (2007) suggested that beta-glucosidases in *Juniperus oxycedrus* and *Chamaecyparis lawsoniana* drops may enhance the plasticity of pollen tube tip wall. Glucanases are associated with the digestion of cellulose microfibrils and the release of cell wall xyloglucans (Ohmiya et al. 2000). This may contribute to the loosening of cell walls for growth (Ohmiya et al. 2000).

We detected three additional typical cell wall-modifying enzymes in cycad drops that had not been reported previously in conifers or *Ephedra*: a pectin lyase-like protein (*C. hildae*), a glycosyl hydrolase with mannosidase activity (*Z. furfuracea* 2012) and beta-hexosaminidase (both *Z. furfuracea* samples). Pectin lyases degrade the polysaccharide pectin found in primary cell walls (Cao 2012) and act to loosen cell walls (Prakash and Prathapasenan 1990). Mannosidases release galactomannan from cell walls (Reid and Meier 1973). Beta-hexosamindases contribute to cell wall dissolution by degrading proteoglycans (Ghosh et al. 2011).

The pollen intine and tube cell walls of cycads consist of similar structural components as those of conifer cell walls: pectins, beta-glucans, cellulose, callose and AGPs (Yatomi et al. 2002). If drop enzymes play a role in cell wall loosening, as suggested for Douglas-fir (Poulis et al. 2005), it makes sense that cycads and conifers would have a similar complement of cell wall-modifying proteins. However, unlike Douglas-fir which sheds its exine immediately upon contact with the pollination drop, thereby exposing the intine to drop enzymes, cycad pollen does not shed its exine (Pettitt 1982). Instead, the cycad pollen tube pushes through at the sulcus. This might mean that in cycads, the action of cell wall degrading enzymes in cell wall loosening is delayed until after pollen tube emergence. O'Leary (2004) also proposed that cell wall degrading enzymes may work to loosen the structure surrounding nucellar cells to reduce resistance for pollen tubes as they grow between cells. In *Z. furfuracea* initial growth of the pollen tube is indeed intercellular (Choi and Friedman 1991).

Another hypothesized role for pollination drop proteins is to mobilize nutrients for the emerging pollen tube (Poulis et al. 2005; Wagner et al. 2007). This is a possible role for glycosyl hydrolase family enzymes, which may release sugars from the cell wall (Lee et al. 2007). Both *Z. furfuracea* samples contained glycosyl hydrolases from family 32 with invertase activity. Invertases convert sucrose to fructose and glucose. Invertases have been identified in Douglas fir pollination drops (Poulis et al. 2005) and have been shown to be functional in vitro (von Aderkas et al. 2012). Besides making monosaccharides available for uptake, invertases may influence the osmotic balance of pollination drops. Von Aderkas et al. (2012) suggested this may even create a specific environment favouring homospecific pollen. In cycads, Tang (1987, 1993) found that

sucrose and glucose were present in the pollination drops of *Z. pumila* and *C. robusta*, and that fructose was also present in *Z. pumila*. It may be that these extracellular sugars are regulated by extracellular enzymes in cycads, as shown for invertase in Douglas-fir (von Aderkas et al. 2012). In vitro or in situ assays would be necessary to demonstrate the function of invertase in regulating drop sugars in cycads.

Proteases are also present in cycad pollination drop samples. Proteases cleave proteins. Three of the four general categories of plant proteases (van der Hoorn 2008) are represented in *C. rumphii* and *Z. furfuracea* drops: cysteine-type, serine-type and aspartic-type proteases. Only metalloproteases are not present. No proteases of any kind were detected in *C. rumphii*. A putative function for proteases suggested by Poulis et al. (2005) and Wagner et al. (2007) was to cleave proteins to make amino acids available for uptake by pollen. Serine-carboxypeptidase-like protein and aspartyl protease were identified in *Pseudotsuga* (Poulis et al. 2005), and subtilisin-like proteinases were identified in *Juniperus* and *Chamaecyparis* (Wagner et al. 2007). Aspartic proteinase, cysteine proteinase, and serine carboxypeptidase were found in *Ephedra* spp. (Chapter 2). However, proteases are a diverse group of proteins. There are more than 800 proteases in *Arabidopsis* (van der Hoorn 2008), and although it is known that proteases have diverse biological functions, they are relatively poorly understood. Specific substrates are only known for a handful of proteases (Tsiatsiani et al. 2012). Proteases can control biological processes by controlling the fate of specific proteins (van der Hoorn 2008). Extracellular proteases in pollination drops may have very precise functions if they are eventually found to have specific protein substrates.

At least one chitinase was identified in each of the cycad pollination drop samples. Chitinases have been detected in the pollination drops of *Juniperus* spp. (Wagner et al. 2007), *Pseudotsuga menziesii* (Poulis et al. 2005), *Welwitschia* (Wagner et al. 2007) and *Ephedra* spp. Chitinolytic activity was detected using biochemical and in-gel assays of Douglas-fir pollination drops (Coulter et al. 2012). This strongly supports a role in defence against fungal pathogens in pollination drops, which is probably also the function of chitinase in cycads. Chitinases hydrolyze chitin in fungal cell walls, sometimes working with beta-1,3-glucanases (Grover 2012). They may also have lysozyme activity, and inhibit bacterial growth (Majeau et al. 1990). Chitinases are hypothesized to mediate defence through chitin-signaling, when chitin oligomers that are created upon hydrolysis elicit up-regulation of defence systems (Shibuya and Minami 2001; Grover 2012; Boller 1995). Chitinases are present as constitutive members of apoplastic fluids in plants, but they also accumulate within plant vacuoles (Collinge et al. 1993). When cells are breached, intracellular chitinase is released; although now in the apoplast, the chitinase is still functional (Mauch and Staehelin 1989). Both intracellular and extracellular chitinases are present in cycad drops. This could potentially be an example of an intracellular protein with an extracellular function in the pollination drop.

Thaumatin-like proteins (TLPs) were identified in *Taxus* (O'Leary et al. 2007), *Juniperus* spp. (Wagner et al. 2007), *Chamaecyparis* (Wagner et al. 2007), and *Ephedra*. TLPs and/or osmotins (a type of TLP; Liu et al. 2010) were detected in *Ceratozamia hildae* and *Z. furfuracea* drops. These proteins are members of the Pathogenesis-Related 5 (PR5) group of proteins. TLPs can work in conjunction with endochitinases to stop fungal infection (Garcia-Casado et al. 2000). TLPs themselves

may have glucanase-like activity that disturbs cell wall formation in fungi (Zareie et al. 2002). They may also inhibit fungal xylanases from acting on plant cell walls (Fierens et al. 2007). In conifer pollination drops, TLPs have been suggested as antifungal agents (O'Leary et al. 2007; Wagner et al. 2007). TLPs have also been proposed as antifreeze proteins in conifer pollination drops (O'Leary et al. 2007, Wagner et al. 2007, Poulis et al. 2005), through an ice-binding mechanism, which prevents ice crystal formation (Griffith et al. 1997). Given that cycads are primarily tropical and sub-tropical species, this seems an unlikely function in cycad pollination drops.

Enzymes involved in reduction-oxidation processes were found in cycad pollination drops. Proteins involved in reduction-oxidation reactions have been studied in other types of plant reproductive secretions. For example, redox reactions are known to play an important role in the defence of tobacco nectar from incoming pathogens (Carter and Thornburg 2004; Park and Thornburg 2009). Robert Thornburg and colleagues have determined that a suite of nectarins (proteins in nectar) work together to create an extremely high level of hydrogen peroxide in nectar (Park and Thornburg 2009). Two enzymes produce hydrogen peroxide directly: Nectarin I, a superoxide dismutase, and Nectarin V, a flavin-containing berberine bridge protein with glucose oxidase activity. Nectarin III is a monodehydroascorbate reductase that produces ascorbate. Together with a membrane bound NADPH oxidase, these enzymes create an antimicrobial environment with levels of hydrogen peroxide reaching 4 mM (Park and Thornburg 2009). Hydrogen peroxide levels in conifer pollination drops were found to range from 1-10 µM (Wagner 2007), which is 1000-fold lower than the 4 mM concentration reported in tobacco nectar (Park and Thornburg 2009). However, Wagner (2007) noted that the low level of

hydrogen peroxide detected could be the result of sample degradation before analysis. Hydrogen peroxide levels have not been reported for cycads. It is intriguing that three of four cycad pollination drop samples contained FAD-binding berberine family proteins, possibly similar to Nectarin V. Two samples contained D-arabinono-1,4-lactone oxidase family proteins. In yeast, this protein catalyses the last step of L-ascorbic acid formation (Huh et al. 1994) – perhaps in a manner similar to Nectarin III, which produces ascorbate (Park and Thornburg 2009). Of the other reduction-oxidation proteins present in cycad pollination drops, both peroxidase (Kawano 2003) and laccase (Mayer and Staples 2002) are known defence proteins. Peroxidase was previously identified in *P. menziesii* (Poulis et al. 2005) and *Ephedra* pollination drops (see Chapter 2). Extracellular peroxidases can both produce reactive oxygen species or degrade them depending on the conditions (Kawano 2003). It is possible that the role of reduction-oxidation reactions in pollination drops is the sames as in nectar – to provide an antimicrobial environment.

Lipid transfer proteins (LTPs) were detected in *C. hildae* and *Cycas rumphii*. An LTP was previously identified in larch pollination drops (O'Leary 2004). LTPs are non-specific and can bind many types of hydrophobic and amphiphilic molecules (Edstam et al. 2013). They make up a large family of extracellular proteins, and although they are not well-understood at present, they are suspected to have a broad range of functions. LTPs might be involved in wax, suberin and sporopollenin deposition, including in the pollen wall (Edstam et al. 2013). They might be involved in pollen tube adhesion to the stigma and style in *Lillium longiflorum* Thunb. (Park et al. 2000). It is possible that in cycads, LTPs are present on the surface of the nucellus and help pollen to adhere. There is also evidence that LTPs have a defence function. In *Arabidopsis*, LTPs might function

in long-distance signaling during pathogen attack (Maldonado et al. 2002). They may also act directly, disrupting fungal (Regente et al. 2005) and bacterial membranes (Nielsen et al. 1996). Since the pollination drop is open to environmental pathogens, defence roles for LTP proteins are also plausible.

A small portion of the proteins present in cycad pollination drops had annotations specifically to the plasma membrane (0 % to 8 % depending on the sample). It is possible that these proteins washed into the drop from the nucellar surface. These proteins may provide clues to the types of interactions between the nucellar surface and drop components. Two such proteins are the receptor-like protein kinase 2 in C*eratozamia hildae* and the leucine-rich repeat transmembrane protein kinase in *Z. furfuracea*. These are members of a large family of proteins known as receptor-like kinases (RLKs). The *Arabidopsis* genome contains more than 600 RLK genes, and as such they have diverse functions in plant growth, development and response to stress (Mizuno et al. 2007). Leucine-rich repeat receptor-like kinases (of which RPK2 is one example) detect exogenous signals including plant hormones and pathogens (Mizuno et al. 2007). Although it seems unlikely that proteins that are normally membrane bound would be functional within the drop, their presence does provide indirect evidence that protein kinases may be present in membranes in contact with the pollination drop. Recently, extracellular ATP was found in the pollination drops of a diverse range of gymnosperm taxa (Rachel Wong, pers. comm.). The detection of receptor-like protein kinases provides some evidence that the membranes in contact with the pollination drop may in fact be interacting with its components.

The extracellular proteins GAST1 and the related gibberellin-regulated family protein were detected in *C. hildae* and *Cycas rumphii*, respectively. GAST1 is a gene known to be regulated by gibberellic acid and abscisic acid, but displays no phenotype when over- or under-expressed in transgenic tomato plants (Shi and Olszewski 1998). A function beyond developmental regulation has not been elucidated for this gene (Herzog et al. 1995). The presence or effect of plant hormones has not been investigated in the context of pollination drops.

Some of the proteins detected in cycad drops fall under the umbrella term of arabinogalactan proteins (Seifert and Roberts 2007): fascilin-like arabinogalactan protein 8; uclacyanin 1; lipid transfer proteins 1 and 4; SKU5 similar 4 and 5 proteins; SHV3-like 1 protein with glycerophosphodiesterase activity. Arabinogalactan proteins are glycoproteins that are modified by complex carbohydrates made up mostly of galactan and arabinose residues. Arabinogalactans are thought to be involved in many developmental processes (cell division, programmed cell death, pattern formation, growth), but the actual mechanism of their action is not understood (Seifert and Roberts 2007). They are often present in plant reproductive tissues, e.g. angiosperm stigmas, and are thought to interact with pollen by acting as an adhesive (Clarke et al. 1979), influencing self-recognition reactions (Lind et al. 1994; Wu et al. 2000), and defining a path along which the pollen tube can grow (Cheung et al. 1995). O’Leary et al. (2004) identified arabinogalactan proteins in yew which were localized to the nucellus. Arabinogalactan production in the nucellus peaked during drop production. O’Leary et al. (2004) proposed that AGPs might be involved in conifer pollen tube guidance and growth. A similar role may exist for the possible AGP proteins in cycads.

Like *Ephedra* spp., it appears that cycads have a degradome present within their pollination drops. Numerous proteins contained annotations to the intracellular space (Tables 5-8). Proteins annotated to the cytoplasm could be explained as proteins released from nucellar cells upon their degradation to form the pollen chamber, as suggested for *Ephedra* spp. Whether any of these intracellular proteins are functional within the pollination drop is unknown. At a minimum, proteins affect the osmotic balance of the drop, which may influence pollen germination (Wagner et al. 2007).

Zamia furfuracea pollination drop samples from two consecutive years were analyzed. Although there was substantial overlap in the types of proteins detected, the results lists were not identical for the two samples (Tables 5, 6, 9). One possible biological explanation for this variation could be that the pollination drop samples were collected at different points during the course of pollination. Perhaps certain proteins are secreted at specific time points. If this is true, it could explain the slight differences in the proteomes of the two samples. Our samples were pooled over multiple days, but this may not have equalized variation between the samples.

There was a difference in the number of proteins identified in *C. rumphii* compared to *Ceratozamia hildae* and *Z. furfuracea*, with *Cycas rumphii* having fewer proteins present in the pollination drop. More than one factor could explain this. It could be a result of evolution in reproductive mechanisms – *Cycas* is the only genus in the Cycadaceae, the group considered the sister to all other extant cycads (Zgurski et al. 2008). *Cycas* spp. also have a unique “cone” morphology, and wind might play a more important role in their pollination mechanism than for Zamiaceae. Perhaps the lower complexity of the *C. rumphii* drop proteome is associated with its unique phylogenetic

position or reproductive biology. There is also the possibility of sampling differences. The *C. rumphii* samples were collected outdoors, and the sample could be influenced by relative humidity which can be extremely high during the rainy season in Florida. Moisture or dew could have diluted the sample. We have no knowledge of the total protein content of any cycad drops. Due to the small volume of the drops and the sporadic nature of their appearance, our total sample volumes were at best ~15 µL each. Even though we started with similar volumes of sample for proteomic analysis, if the *C. rumphii* sample was more dilute than the others, this could provide a technical explanation for the lower number of proteins detected. Of the protein types that *C. rumphii* shares with the other species, half are involved in defence (chitinase, peroxidase, lipid-transfer proteins), suggesting a role in defence is important in all cycad pollination drops. When cycad cones are receptive, the micropyle is a potential entry point for not only wind-borne, but also insect-borne bacterial and fungal pathogens.

The protein content of cycad pollination drops could have a role in attracting insect pollinators. It is unknown whether insect pollinators use pollination drops as a food source (Tang 1987; Terry 2001). Thrips have been observed moving towards the micropyle of *Macrozamia* possibly to visit the pollination drops (Terry et al. 2005). Cycad pollination drops have comparable concentrations and types of sugars to some flower nectars (Nepi et al. 2009). They also contain free amino acids, which are known attractants of insect pollinators in flowering plants (Nepi et al. 2009). To date, there are no published reports focused on the relationship between pollination drops and insect pollinators in cycads. It would be challenging to observe insect behavior inside female

cones of Zamiaceae because of their construction. This would likely require fibre-optic photography within the receptive cone.

This study represents the first proteomic analyses of cycad pollination drops. In terms of predicted function, the types of proteins found in cycad drops were similar to those found in conifers and *Ephedra*: a complement of carbohydrate-modifying enzymes that likely modify pollen intine and tube cell wall components; a cocktail of defence proteins; a group of proteolytic enzymes likely involved in nutrient mobilization; arabinogalactan proteins possibly involved in pollen tube guidance. Additional secreted proteins were also present, such as oxidation-reduction proteins, as well as potential receptor-like kinases, which have not been detected previously in any pollination drops. Cycads, like *Ephedra,* have a protein degradome that probably originates during the formation of the pollen chamber. Despite cycads being insect-pollinated, zooidogamous and an independent clade of seed plants, the overall secretome and degradome composition of cycads is similar to conifers and *Ephedra* spp.

Chapter 5: Proteins in the pollination drops of *Ginkgo biloba.*

Introduction

Ginkgo biloba holds an important place in our understanding of seed-plant evolution, being the only living representative of the once diverse order Ginkgoales (Christianson and Jernstedt 2009). *Ginkgo biloba* is considered a living fossil because seemingly few morphological changes have occurred over the past 100 million years (Little et al. 2013 and references within). Remnant wild populations of *G. biloba* are thought to exist in the Dalou Mountains of southwestern China (Tang et al. 2012), but most extant *G. biloba* trees are cultivated.

Ginkgo biloba is dioecious, with separate male and female trees. On male trees, pollen sacs are arranged into catkins (Christianson and Jernstedt 2009) that are borne in the axils of bud scales and leaves of short-shoots (del Tredici 2007). Female reproductive structures of *G. biloba* also develop in axillary positions on short-shoots. Ovules are borne in pairs on short stalks (del Tredici 2007). *G. biloba* females produce neither an ovule bract, nor a distinct cone (Douglas et al. 2007).

Ginkgo biloba is wind-pollinated. Pollen is captured by pollination drops, which are completely exposed to the environment (del Tredici 2007). Jin et al. (2012) estimated the volume of *G. biloba* pollination drops to be ~ 100 nL. Little information is published about the organic or mineral content of *G. biloba* pollination drops. Del Tredici (2007) mentioned that the drops are mucilaginous. Friedman (1987) referred to an unpublished observation by Irene Baker that *Ginkgo biloba* drops contain large amounts of sugars. I am not aware of any other published accounts of the contents of *G. biloba* drops.

Pollen is drawn into the ovule as the drop retracts. It settles into a pollen chamber previously formed by the degradation of nucellar cells (Friedman 1987). A recent study by Jin et al. (2012) suggested that the mechanism of drop withdrawal in *G. biloba* operates on the balance between active secretion and passive evaporation. Jin et al. (2012) found that viable pollen may alter the balance between secretion and evaporation, and stimulate drop withdrawal. Non-viable pollen and inter-generic pollen had varying degrees of effect on drop withdrawal.

Pollen germination and pollen tube growth have a unique pattern in *G. biloba*. Pollen germination begins about one week after pollination (Friedman 1987). The pollen grain swells, and the tube cell pushes through the sulcus of the exine. The pollen tube grows, unbranched, between cells of the nucellus for some distance. Then the pollen tube changes its mode of growth. Numerous slender branches grow from the pollen tube tip into the intercellular spaces of the nucellus. Friedman (1987) speculated that the primary function of the branched pollen tube is nutrient uptake. The proximal end of the pollen tube (the original pollen grain) swells, and two multiflagellate sperm develop within. The proximal end of the pollen tube eventually bursts, and the sperm swim through the fluid-filled archegonial chamber to fertilize the egg cell. The entire process, from pollination to fertilization, takes approximately five to six months (Friedman 1987).

We set out to investigate the proteome of *G. biloba* pollination drops. We hypothesized that the protein content of *G. biloba* pollination drops would be similar to that of the cycads. *G. biloba* and cycads are both distinct lineages, but they do share reproductive characteristics. Both are zooidogamous and have haustorial pollen tubes, traits which have traditionally been thought of as having a deep evolutionary connection

(Chamberlain 1935). In addition, both have substantial pollen chambers at the time of pollination drop production, leading us to predict that *G. biloba* pollination drops would also have a degradome component. Here I report the results of proteomic analysis of *G. biloba* pollination drops.

Methods

Pollination drop collection in *Ginkgo biloba*

Pollination drops of *G. biloba* were collected by Dr. Stefan Little in April 2012. The ovules of *G. biloba* are exposed, allowing pollination drops to be collected directly from the tree. A 10 µL filter-pipette tip was used to collect drops from *G. biloba* trees growing on the north side of Haring Hall, located at the University of California at Davis. Collections were made between 6 am and 10 am. *Ginkgo biloba* drops appeared to crystalize over the period of collection.

Proteomics

Methods for the proteomic analysis of *G. biloba* pollination drops were presented in Chapter 3.

Results

There were 66 proteins detected in the *G. biloba* pollination drop sample. Of those, 57 had acceptable hits to the TAIR 10 database (Table 10). Approximately equal numbers were annotated to the intracellular and extracellular space, 32 and 34 proteins respectively. Ten proteins had an annotation specifically to the plasma membrane. GO annotations are not mutually exclusive.

The best-represented GO category for Biological Process was Translation/Translation elongation; 18 % of the proteins that had an acceptable match to

Table 10. Proteins identified in *Ginkgo biloba* pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model.

Accession	-10lgP	Total Peptides	Unique Peptides	TAIR10 Gene Model	TAIR10 Description	BLASTP e value	Gene Ontology Biological Process	Gene Ontology Cellular Component
Ginkgo_Contig57949_45	338.59	30	29	AT2G03200.1	Eukaryotic aspartyl protease family protein	1e-123	proteolysis	extracellular region
Ginkgo_Contig17660_5	271.67	14	14					
Ginkgo_Contig23254_29	250.94	11	11	AT5G52300.1	CAP160 protein	3e-06		cytoplasm
Ginkgo_Contig51048_31	213.18	6	6	AT2G36640.1	embryonic cell protein 63	4e-27		
Ginkgo_k25ctg8139545_7	204.85	7	7	AT5G44310.2	Late embryogenesis abundant protein (LEA) family protein	1e-10		
scaffold-AUDE-2060069-Widdringtonia_cedarbergensis_single_14	198.58	7	2	AT2G03200.1	Eukaryotic aspartyl protease family protein	3e-143	proteolysis	extracellular region
Ginkgo_Contig62116_21	175.17	6	6					
Podocarpus_macro_isotig11681_35	173.94	6	1	AT5G60390.1	GTP binding Elongation factor Tu family protein	0	translational elongation	cytoplasm; vacuole; plasma membrane; plasmodesma

Pinus_taeda_isotig29049_39	159.01	3	3	AT4G20360.1	RAB GTPase homolog E1B	0	translational elongation	chloroplast stroma; chloroplast envelope; apoplast; chloroplast thylakoid; chloroplast; membrane; chloroplast thylakoid membrane; nucleolus; intracellular; nucleoid; nucleus
scaffold-DZQM-2011929-Pinus_radiata_sub2_45	155.66	6	1	AT5G60390.1	GTP binding Elongation factor Tu family protein	0	translational elongation	cytoplasm; vacuole; plasma membrane; plasmodesma
Sciadopitys_verti_isotig00670_23	152.52	4	1	AT2G03200.1	Eukaryotic aspartyl protease family protein	2e-137	proteolysis	extracellular region
Ginkgo_Contig74012_25	150.14	3	1	AT5G08370.1	alpha-galactosidase 2	2e-118	carbohydrate metabolic process	extracellular region
Ginkgo_Contig49707_38	134.99	4	4					
Ginkgo_Contig73991_34	134.28	2	2	AT4G02930.1	GTP binding Elongation factor Tu family protein	0	translational elongation	mitochondrion; intracellular; cell wall
Ginkgo_Contig76923_35	130.67	3	3	AT1G76940.1	RNA-binding (RRM/RBD/RNP motifs) family protein	6e-51		
Ginkgo_Contig52403_17	130.31	2	2	AT4G11650.1	osmotin 34	1e-69	response to salt stress	extracellular region
Dougfir-megastigmus-comp573804_c1_seq2_17	129.38	3	3	AT4G05050.1	ubiquitin 11	2e-105		
Sciadopitys_verti_isotig11199_16	128.12	3	3	AT3G12580.1	heat shock protein 70	0	response to cadmium ion	plasma membrane; vacuolar membrane; mitochondrion; cell wall; cytosol; Golgi apparatus

scaffold-CDFR-2006382-Manoao_colensoi_single_14	127.84	3	3	ATMG01190.1	ATP synthase subunit 1	0	ATP synthesis coupled proton transport; ATP hydrolysis coupled proton transport; ATP metabolic process; proton transport	proton-transporting two-sector ATPase complex; proton-transporting ATP synthase complex, catalytic core F(1); cell wall; mitochondrion; proton-transporting two-sector ATPase complex, catalytic domain
Dougfir-megastigmus-comp585996_c0_seq1_41	122.52	3	2	AT5G08680.1	ATP synthase alpha/beta family protein	0	ATP hydrolysis coupled proton transport; ATP biosynthetic process; ATP catabolic process; ATP synthesis coupled proton transport; proton transport; ATP metabolic process; response to cadmium ion	mitochondrial proton-transporting ATP synthase complex, catalytic core F(1); proton-transporting ATP synthase complex, catalytic core F(1); mitochondrion; proton-transporting two-sector ATPase complex, catalytic domain; membrane; proton-transporting two-sector ATPase complex
Ginkgo_Contig61103_7	117.84	2	2	AT4G33720.1	CAP (Cysteine-rich secretory proteins, Antigen 5, and Pathogenesis-related 1 protein) superfamily protein	5e-60		extracellular region
Ginkgo_Contig76656_8	116.46	1	1					
Dougfir-megastigmus-comp551308_c0_seq1_8	113.99	2	2	AT1G74030.1	enolase 1	7e-99	glycolysis	phosphopyruvate hydratase complex; chloroplast stroma; chloroplast

Ginkgo_Contig71819_24	111.43	2	2	AT4G02930.1	GTP binding Elongation factor Tu family protein	0	translational elongation	mitochondrion; intracellular; cell wall
Ginkgo_Contig73569_13	105.38	2	2					
picea_abies_isotig07279_9	97.04	1	1	AT2G29560.1	cytosolic enolase	1e-91	glycolysis	phosphopyruvate hydratase complex; cytosol
scaffold-HOUF-2112961-Cryptomeria_japonica-leaf_single_25	96.32	1	1	AT2G31060.3	elongation factor family protein	0		chloroplast
Pseudotsuga_menz_isotig20519_40	94.84	1	1	AT2G03200.1	Eukaryotic aspartyl protease family protein	8e-141	proteolysis	extracellular region
Ginkgo_Contig76589_25	91.81	1	1					
Dougfir-megastigmus-comp588241_c0_seq1_11	91.01	2	1	AT5G60390.1	GTP binding Elongation factor Tu family protein	0	translational elongation	cytoplasm; vacuole; plasma membrane; plasmodesma
scaffold-AREG-2011042-Nothotsuga_longibracteata_single_12	89.08	1	1	AT2G26870.1	non-specific phospholipase C2	0		plasma membrane
Dougfir-megastigmus-comp4395363_c0_seq1_4	87.09	1	1	AT4G02930.1	GTP binding Elongation factor Tu family protein	2e-10	translational elongation	mitochondrion; intracellular; cell wall
scaffold-GNQG-2086713-Encephalartos_barteri_single_19	83.41	1	1	AT5G64570.1	beta-D-xylosidase 4	0	carbohydrate metabolic process	cell wall; apoplast; extracellular region
Ginkgo_Contig76315_11	83.27	1	1	AT2G31610.1	Ribosomal protein S3 family protein	2e-96	translation; response to salt stress	small ribosomal subunit; vacuolar membrane; vacuole; intracellular; cytosolic ribosome; chloroplast; cytoplasm; plasmodesma; membrane; cytosol; ribosome
Ginkgo_Contig11187_15	83.22	1	1	AT2G03200.1	Eukaryotic aspartyl protease family protein	1e-134	proteolysis	extracellular region
Ginkgo_Contig70013_35	82.63	1	1	AT5G12950.1	Putative glycosyl hydrolase of unknown function (DUF1680)	0		vacuole; plant-type cell wall; extracellular region

scaffold-EGLZ-2000421-Prumnopitys_andina_single_13	81.3	1	1	AT3G17390.1	S-adenosylmethionine synthetase family protein	0	S-adenosylmethionine biosynthetic process	nucleolus; plasma membrane; cytoplasm; cell wall; membrane; plasmodesma
Dougfir-megastigmus-comp542103_c0_seq1_14	80.87	1	1	AT2G21430.1	Papain family cysteine protease	2e-178	proteolysis	extracellular region
scaffold-GKCZ-2008328-Diselma_archeri_single_37	84.72	2	2	AT1G69830.1	alpha-amylase-like 3	2e-175	carbohydrate metabolic process	chloroplast stroma; chloroplast; extracellular region
Ginkgo_Contig53208_22	79.28	1	1	AT1G13130.1	Cellulase (glycosyl hydrolase family 5) protein	4e-173	carbohydrate metabolic process	chloroplast
Ginkgo_Contig71643_18	78.34	1	1	AT5G55180.2	O-Glycosyl hydrolases family 17 protein	3e-32	carbohydrate metabolic process	plasma membrane
scaffold-GKCZ-2011432-Diselma_archeri_single_58	78.15	1	1	AT3G53950.1	glyoxal oxidase-related protein	0		extracellular region
Ginkgo_Contig38871_4	76.52	1	1					
scaffold-IOVS-2014849-Pseudotsuga_menziesii_single_10	74.84	1	1	AT3G03220.1	expansin A13	3e-105	plant-type cell wall organization	extracellular region
Ginkgo_Contig77070_30	74.63	1	1	AT2G32910.1	DCD (Development and Cell Death) domain protein	2e-24		cytoplasm
Ginkgo_Contig19162_37	70.48	1	1	AT3G54400.1	Eukaryotic aspartyl protease family protein	2e-98	proteolysis	apoplast; extracellular region; cell wall; plant-type cell wall; chloroplast
Dougfir-megastigmus-comp584823_c4_seq3_146	69.33	1	1	AT5G63680.1	Pyruvate kinase family protein	1e-132	glycolysis; response to cadmium ion	cytosol; plasma membrane; cytoplasm
Ginkgo_Contig72342_24	69.08	1	1	AT4G25000.1	alpha-amylase-like	1e-54	carbohydrate metabolic process	extracellular region
scaffold-BBDD-2076070-Microstrobos_fitzgeraldii_single_12	67.99	1	1	AT4G16260.1	Glycosyl hydrolase superfamily protein	2e-100	response to salt stress	cell wall; vacuolar membrane
Ginkgo_Contig76726_22	67.2	1	1	AT5G15270.2	RNA-binding KH domain-containing protein	8e-116		nucleus

Podocarpus_macro_isotig04452_20	64.08	1	1	AT1G63660.1	GMP synthase (glutamine-hydrolyzing), putative / glutamine amidotransferase, putative	0	GMP biosynthetic process; purine nucleotide biosynthetic process; asparagine biosynthetic process	cytosol; cytoplasm
Sciadopitys_verti_isotig08537_31	61.7	1	1	AT3G52580.1	Ribosomal protein S11 family protein	9e-90	translation	intracellular; ribosome; cytoplasm
scaffold-BTTS-2079163-Austrotaxus_spicata_single_14	61.61	1	1	AT3G12780.1	phosphoglycerate kinase 1	0	glycolysis; response to cadmium ion	chloroplast stroma; chloroplast; cell wall; chloroplast envelope; thylakoid; cytosol; membrane; nucleus; apoplast
scaffold-CGDN-2011024-Tetraclinis_sp._sub1_13	58.8	1	1	AT2G45470.1	FASCICLIN-like arabinogalactan protein 8	5e-123		plant-type cell wall; plasma membrane; anchored to plasma membrane; apoplast
Dougfir-megastigmus-comp578170_c0_seq12_4	57.62	1	1					
scaffold-FHST-2064055-Taxodium_distichum_single_37	55.37	1	1	AT1G56340.1	calreticulin 1a	0	response to salt stress; response to cadmium ion; protein folding	plasmodesma; chloroplast; vacuolar membrane
Metasequoia_Contig8207_18	55.26	1	1	AT4G11650.1	osmotin 34	2e-69	response to salt stress	extracellular region
scaffold-EGLZ-2039750-Prumnopitys_andina_single_20	55.03	1	1					
Ginkgo_Contig47693_63	54.93	1	1	AT1G65590.1	beta-hexosaminidase 3	0	carbohydrate metabolic process	plant-type cell wall; cell wall
Dougfir-megastigmus-comp583960_c1_seq2_733	53.69	1	1	AT4G11120.1	translation elongation factor Ts (EF-Ts), putative	1e-35	translational elongation	intracellular; mitochondrion
Ginkgo_Contig30218_32	52.3	1	1	AT3G62760.1	Glutathione S-transferase family protein	1e-76		cytoplasm

scaffold-GNQG-2086106-Encephalartos_barteri_single_44	51.97	1	1	AT1G26560.1	beta glucosidase 40	4e-162	carbohydrate metabolic process	apoplast; chloroplast; extracellular region
scaffold-GJTI-2057049-Cephalotaxus_harringtonia_single_9	50.21	1	1	AT4G11650.1	osmotin 34	2e-76	response to salt stress	extracellular region
scaffold-GTHK-2004686-Gnetum_montanum_sub1_32	50.11	1	1	AT3G52930.1	Aldolase superfamily protein	0	glycolysis; response to cadmium ion; response to salt stress	plasma membrane; nucleolus; plasmodesma; cell wall; chloroplast; apoplast; cytosol; vacuolar membrane; cytoplasm
Podocarpus_macro_isotig11599_26	50.06	1	1	AT5G55190.1	RAN GTPase 3	5e-157	GTP catabolic process; nucleocytoplasmic transport; small GTPase mediated signal transduction; intracellular protein transport; signal transduction	Golgi apparatus; plasmodesma; nucleus
Ginkgo_Contig17796_22	47.81	1	1	AT2G37870.1	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	1e-12	lipid transport	extracellular region; plasmodesma

the TAIR10 database had an annotation to this category. About 14 % had annotations to carbohydrate metabolic process, 12 % of proteins had an annotation to proteolysis, and 9 % to glycolysis. Other categories included response to salt stress (12 %), response to cadmium ion (11 %) and ATP metabolic process (4 %). Many other GO categories were represented (Table 10). However, a number of proteins did not contain GO annotations for Biological Process.

Discussion

Ginkgo biloba is the only extant representative of the ginkgoalean lineage. Proteins have been detected in pollination drops of multiple species of conifers (Poulis et al. 2005; O'Leary et al. 2007; Wagner et al. 2007), *Ephedra* (Chapter 2) and cycads (Chapter 4). The discovery of proteins in *Ginkgo* pollination drops confirms that proteins are present in the pollination drops of all extant gymnosperm lineages. This means that all gymnosperm lineages produce complex reproductive secretions.

We hypothesized that the pollination drop proteome of *G. biloba* would be similar to those of the cycads, since *G. biloba* and the cycads share reproductive characteristics that are unique in seed plants - both are zooidogamous and both have haustorial pollen tubes. Our proteomic analysis indicated that *G. biloba* pollination drops contained a mixture of proteins annotated to both extracellular and intracellular locations (Table 10), similar to what was detected in cycad pollination drop proteomes (Chapter 4). The extracellular proteins identified in *G. biloba* are likely secreted into the pollination drop. The proteins with GO annotations to intracellular locations (Table 10) are probably the products of tissue degradation in the nucellus during pollen chamber formation. *G. biloba* has a deep pollen chamber at the time of pollination drop secretion. As in the cycads and

Ephedra, these intracellular proteins should be considered part of the pollination drop degradome.

The carbohydrate-modifying proteins detected in *G. biloba* pollination drops were comparable to those found in cycads. Galactosidase, xylosidase, cellulase, hexosaminidase, glucosidase and glycosyl hydrolase were identified in both *G. biloba* and cycad drops. These types of proteins are implicated in cell wall reorganization during cell growth (Ohmiya et al. 2000; Minic 2008; Ghosh et al. 2011). Their suggested roles in pollination drops are to aid in pollen germination and pollen tube growth (Poulis et al. 2005; Wagner et al. 2007), and to cause cell wall loosening in the nucellus, thereby allowing the pollen tube to grow more easily through the intercellular space (O'Leary 2004).

An additional protein involved in carbohydrate modification, alpha-amylase, was detected in the *G. biloba* sample. This protein had annotations to both the chloroplast and to the extracellular space. Alpha-amylases are involved in the breakdown of starch. Traditionally, alpha-amylases were assumed to have a role in the diurnal pattern of starch degradation in plant plastids. However, *Arabidopsis* does not require alpha-amylases for starch degradation in living cells (Yu et al. 2005). Their role in the process is probably minor (Delatte et al. 2006). Alpha-amylases do have a well-established role in starch degradation in the dead endosperm cells of germinating seeds of cereal plants (Doyle et al. 2007). Secreted alpha-amylases that degrade starch in other types of dead plant cells are also found in *Arabidopsis*, pea and tobacco (reviewed in Doyle et al. 2007). Doyle et al. (2007) speculated that the glucose released from starch of dead cells is probably made available to surrounding living cells. A potential role for alpha-amylase in the pollination

drop could be to liberate glucose from starch in degrading nucellar cells, and to make glucose available for uptake by the growing pollen tube.

Ginkgo biloba pollination drops share other extracellular proteins with cycads. A lipid-transfer protein was identified in *G. biloba*. Lipid-transfer proteins may be involved in pollen adhesion (Park et al. 2000), wax deposition (Edstam et al. 2013), or pathogen defence (Maldonado et al. 2002; Regente et al. 2005). A number of proteins detected in *G. biloba* pollination drops matched to osmotin, a type of thaumatin-like protein. Thaumatin-like proteins have been proposed as anti-fungal agents within pollination drops (O'Leary et al. 2007; Wagner et al. 2007). One protein detected in *G. biloba* matched to a fascilin-like arabinogalactan protein (AGP). O'Leary et al. (2004) proposed that AGPs might be involved in conifer pollen tube guidance and growth. Finally, proteolytic enzymes, namely aspartyl protease and cysteine protease, were present in *G. biloba*. Proteases may release amino acids for uptake by the growing pollen tube (Poulis 2004).

Ginkgo biloba pollination drops appeared to lack a strong complement of oxidation-reduction enzymes, as was present in cycad pollination drops. There were no proteins with GO annotations for the oxidation-reduction process in *G. biloba*. However, one protein, a glyoxal oxidase-related protein, is likely to be an oxidative enzyme. Zhou et al. (2007) speculated that in *Vitis pseudoreticulata* W.T. Wang, glyoxal oxidase is involved in defence through the production of hydrogen peroxide. Transient over-expression of a glyoxal oxidase gene in *V. pseudoreticulata* resulted in plants that resisted infection by the fungus *Erysiphe necator* (Schwein.) Burrill (Guan et al. 2011). Guan et al. (2011) suggested that extracellular hydrogen peroxide produced by glyoxal oxidase

may cause the up-regulation of defence genes. The glyoxal oxidase-like protein detected in *G. biloba* drops was annotated to the extracellular space. It is possible that glyoxal oxidase may, similarly, act as an antimicrobial agent in *G. biloba* pollination drops.

An expansin was detected in the *G. biloba* pollination drop. Expansins have not been identified in any other pollination drops to date. Alpha- and beta-expansins are extracellular proteins involved in cell wall expansion. The mechanism by which they cause cell expansion is not understood (Lipchinsky 2013). They may work by altering hydrogen bonds and Van der Waals forces between cell wall components, rather than altering covalent bonds (Lipchinsky 2013). Beta-expansins are known to diffuse from the pollen of *Triticale* (x *Triticosecale* Wittmack), where they may loosen the cell walls of the stigma tissue, allowing the pollen tube to grow more easily between cells (Zaidi et al. 2012). Valdivia et al. (2009) demonstrated that under-expression of beta-expansins in maize pollen resulted in pollen tubes that could not easily penetrate the stigma. It is plausible that the alpha-expansin protein in *G. biloba* pollination drops could play a role in either pollen tube cell wall expansion or in preparation of the nucellar tissue for pollen tube penetration.

A substantial portion of the proteins in *G. biloba* pollination drops likely originate from nucellar cells that are in the process of degradation, thereby contributing to pollen chamber formation. Ten proteins involved in Translation or Translation Elongation were detected in the *G. biloba* pollination drops. These proteins have GTP-binding and/or GTPase activity. Two ribosomal proteins were also detected amongst this group. These proteins are probably involved in protein translation within the cell. A few of the proteins identified had annotations referring to the cell wall or apoplast, however the precise

function of such enzymes in the cell wall is not well-described in the literature. It is also likely that the glycolytic enzymes present in the pollination drop are a result of pollen chamber formation. Glycolysis occurs in the plastids and cytosol to produce ATP by the oxidation of hexose sugars (Plaxton 1996). A number of other proteins with annotations to the intracellular space can be included in the protein degradome (Table 10), e.g. CAP160 protein, RNA-binding KH domain-containing protein, and calreticulin 1a.

Ubiquitin proteins have well-established intracellular roles and could be considered degradome proteins. As part of the ubiquitin-proteasome system in plants, ubiquitin is involved in the regulation of numerous plant processes. These include plant development and cell division, as well as response to hormonal signals and abiotic/biotic stressors (reviewed by Sadanandom et al. 2012). Ubiquitin is conjugated to acceptor proteins, which are subsequently degraded by the 26S proteasome.

Ubiquitin proteins may also function in the apoplast, where they may influence the uptake of other proteins. A role for extracellular ubiquitin has been proposed for pollen tube adhesion in lily. Kim et al. (2006) reported that free ubiquitin is endocytosed at the tips of lily pollen tubes in vitro. They found that the addition of free ubiquitin enhanced the uptake of stigma/stylar cysteine-rich adhesion (SCA) protein, a type of lipid-transfer protein known to be involved in pollen tube adhesion. As mentioned earlier, a lipid-transfer protein was identified in *G. biloba* pollination drops. A BLASTp search of the open reading frame sequence relating to the lipid-transfer protein in our *G. biloba* sample against the NCBI nr database, restricted to the lily family, returns matches to SCA proteins in lily. However, the e-values of the returned proteins are too high to be considered a significant match (0.54 – 0.63). Investigation of a possible role for lipid-

transfer proteins and free ubiquitin in *G. biloba* pollen tubes would certainly be interesting to undertake.

Heat shock proteins are another example of proteins with well-known intracellular functions, and possible extracellular functions. Heat shock proteins act as molecular chaperones. They ensure the correct folding of proteins. Extracellular heat shock proteins have been reported in animals with a role in immune signaling (Schmitt et al. 2007; de Maio and Vazquez 2013). Whether functional extracellular heat shock proteins also occur in plants is not evident in the literature.

A few proteins were annotated to the plasma membrane only, e.g. non-specific phospholipase C2. Plant phospholipases are involved in a broad range of plant physiological processes, from lipid metabolism to cell signaling (reviewed by Chen et al. 2011; Pokotylo et al. 2013). Phospholipase C cleaves diacylglycerols (DAG) from the hydrophilic headgroups of membrane phospholipids (Pokotylo et al. 2013). In signaling pathways, DAG can act as a secondary messenger. It is also a precursor for storage and structural lipids (Chen et al. 2011). Non-specific phospholipase C is up-regulated during phosphate deprivation (Nakamura et al. 2005), as well as by the application of the plant hormones brassinolide and auxin (Wimalasekera et al. 2010). This implies possible roles in both stress response and cell signaling (Chen et al. 2011; Pokotylo et al. 2013). Phospholipases from another distinct group, the phosphatidylinositol-specific phospholipases, have been implicated in signaling during pollen germination in lily (Pan et al. 2005) and in regulating pollen tube tip elongation in *Petunia* (Dowd et al. 2006). However, roles have not been elucidated for non-specific phospholipases, such as the non-specific phospholipase C2 found in *G. biloba* pollination drops.

Our proteomic analysis indicated that the proteome of *G. biloba* pollination drops contained secretome proteins similar to those found previously in the cycads. Both *G. biloba* and the cycads contained similar types of carbohydrate-modifying enzymes, as well as other extracellular proteins, e.g. osmotins, lipid-transfer proteins, proteases, arabinogalactan proteins. Many of these secretome proteins were also previously detected in conifers and *Ephedra*. Notably, however, *G. biloba* pollination drops lacked the variety of reduction-oxidation enzymes that were present in the cycads.

Ginkgo biloba pollination drops also contained a mixture of proteins with annotations to intracellular locations. Like cycads and *Ephedra*, *G. biloba* develops a deep pollen chamber during pollination drop production. As in cycads and *Ephedra*, the intracellular proteins of *G. biloba* are best explained as degradome proteins that are washed into the pollination drops from the degrading nucellus. In *G. biloba* pollination drops, the degradome was dominated by proteins involved in translation and glycolysis.

Our results demonstrate that the pollination drops of *Ginkgo*, cycads, *Ephedra* and conifers share similar types of proteins, e.g. carbohydrate modifying enzymes, proteases, lipid transfer proteins and thaumatin-like proteins, despite the diversity of pollination mechanisms that they display. These results support the hypothesis presented by Wagner et al. (2007), which suggested that there may be a suite of conserved proteins in the pollination drops of gymnosperms.

Chapter 6: Proteins in the pollination drops of *Gnetum gnemon* and *Welwitschia mirabilis*.

Introduction

The placement of the Gnetales in extant seed-plant phylogeny is fraught with controversy. Evidence for sister relationships to the conifers (Chaw et al. 1997), to the Cupressaceae within the conifers (Wu et al. 2013), to the Pinaceae within the conifers (Bowe et al. 2000; Hajibabaei et al. 2006), or to the angiosperms (Doyle and Donoghue 1986) are only some of the molecularly- or morphologically-supported hypotheses. Three extant genera make up the Gnetales: *Ephedra*, *Gnetum* and *Welwitschia*. These are grouped together because they share common traits, such as the presence of vessel elements in their wood and the compound structure of their micro- and mega-strobili (Gifford and Foster 1989; Fernando et al. 2010). They are, however, thought to be the survivors of a once much more diverse clade, due to the distinct morphological and ecological characteristics of each genus (Rydin and Friis 2005).

Two gnetalean genera will be discussed here, *Gnetum* and *Welwitschia*. The genus *Gnetum* consists of around 40 species. Their distribution is limited to tropical and sub-tropical climates mostly in Asia, but also north-western South America and Africa (Biye et al. 2014). The majority of *Gnetum* species are lianas (woody vines) but a few tree species also occur (Biye et al. 2014). *Welwitschia mirabilis* makes up a monospecific genus. It is endemic to the Namib Desert in Africa. It has been referred to as the “plant monster” (Martens 1978) or “botanical octopus” (Tijmens 1968) in the literature because of its unusual morphology. It only ever produces two foliage leaves which grow meters long and split as they grow (Henschel and Seely 2000).

Both *Gnetum* and *Welwitschia* are dioecious. Female cones occur on separate plants from male cones. Female cones of *Gnetum* have a paired bract, followed by several whorls that form circular collars along the cone axis (Endress 1996). Ovules are formed within the axillary position of the collars (Takaso and Bouman 1986). There may be from two to six ovules per collar (Endress 1996). Male *Gnetum* cones are more complex. They are also made up of a series of collars. However, not only are there whorls of microsporangia, but each collar may also contain a whorl of sterile ovules (Endress 1996). Female *Welwitschia* cones are made up of decussate pairs of bracts containing an ovule within the axil of each bract (Endress 1996). Ovules have a long micropylar tube that extends outward from the bract (Carafa et al. 1992). Female cones can have upwards of 50 ovules per cone (Endress 1996). The male plants of *W. mirabilis* also have cones consisting of decussate pairs of bracts. A flower-like reproductive complex is found in the axil of the bracts. The reproductive complex is made of six microsporangia that surround one sterile ovule (Haycraft and Carmichael 2001). These sterile ovules differ from those on the female plants in shape, as well as by having a flattened micropylar disk that stops pollen from entering the micropyle (Endress 1996).

Pollination drops are secreted by fertile ovules on female cones, and by sterile ovules on male cones in both *Gnetum gnemon* (Carmichael and Friedman 1996) and *W. mirabilis* (Wetschnig and Depisch 1999). Pollination drops of *G. gnemon* are small: 0.164 µL on sterile ovules and 0.190 µL on fertile ovules (Kato et al. 1995). There is some discrepancy in the literature over the timing of pollination drop secretion in these genera. Van der Pijl (1953) observed that pollination drops of *G. gnemon* appeared in the morning and disappeared over time, whereas Kato et al. (1995) observed that pollination

drops of both female and male cones appeared in the evening. Both agreed that humidity affected pollination drop production in *G. gnemon*. In *W. mirabilis*, Wetschnig and Depisch (1999) found that the pattern of pollination drop production differed between fertile and sterile ovules. Pollination drops on female cones appeared at noon, peaked in production around 2 pm and were resorbed or collected by insects by 5pm. Male drops appeared at 11 am and were not resorbed (Wetschnig and Depisch 1999). In contrast, Carafa et al. (1992) did not observe drop retraction of female drops of *W. mirabilis*. Unlike other researchers, their observations were made in greenhouse conditions (Wetschnig and Depisch 1999).

Biochemical analyses of both *G. gnemon* and *W. mirabilis* pollination drops have revealed diverse classes of compounds. Kato et al. (1995) reported the presence of sugars in *G. gnemon* ranging between 3 - 13 %. Murch and Tomlinson (2011, abstract only) conducted a metabolomic analysis that showed a complex array of metabolites are present in both the male and female drops of *G. gnemon*. Their abstract states that the male drops were more complex than the female drops in terms of the number and types of putative metabolites. The signals detected in male drops were consistent with the presence of polyphenols, indoles and flavonoids, and male drops contained a wider array of amino acids. Carafa et al. (1992) conducted a biochemical analysis of female pollination drops in *W. mirabilis*. Sugars were present in the form of glucose, fructose, galactose, and uronic acid. They also detected the amino acids serine, tyrosine and valine. Although Carafa et al. (1992) tested for the presence of a number of proteins, they only found trace amounts of an acid phosphatase in the drop. Wagner et al. (2007) took a proteomics approach and were able to identify a chitinase in the pollination drops of

female *Welwitschia* cones. Other proteins were present in SDS-PAGE gels, but were unidentifiable due to the incompleteness of public database coverage of gymnosperms at the time.

Both *G. gnemon* and *W. mirabilis* are insect-pollinated. Kato et al. (1995) observed nocturnal moths from the families Pyralidae and Geometridae visiting *G. gnemon*. The moths unfurled their proboscides and consumed the pollination drop. Pollen became attached to their proboscides and antennae. Kato et al. (1995) also observed that a strong odour was emitted from both male and female cones. Wetschnig and Depisch (1999) suggested flies are the main pollinator of *W. mirabilis*. Flies were observed harvesting pollination drops secreted from both male and female cones. Pollen grains attached to their bodies were deposited into the drops. Bees were also observed on both male and female cones, and pollen-laden wasps were seen visiting female cones. Wetschnig and Depisch (1999) eliminated wind as a pollination vector because their experiments showed that the sticky pollen rarely traveled more than a couple of meters away from the plants.

In *Gnetum*, pollen is drawn into the ovule by the pollination drop and lands on the nucellus (van der Pijl 1953; Carmichael and Friedman 1996). A shallow pollen chamber is present at pollination (van der Pijl 1953; Martens 1971 in Singh 1978) and may be a source of some drop constituents (van der Pijl 1953). Within about four days of pollination, the inaperturate pollen grain (Yao et al. 2004) sheds its exine and germinates (Carmichael and Friedman 1996). The pollen tube grows by an intercellular path through the nucellus towards the female gametophyte (Carmichael and Friedman 1996). *Gnetum* and *W. mirabilis* do not develop archegonia containing egg cells. Rather, the female gametophyte is coenocytic at fertilization and any of the nuclei present may function as

an egg cell (Carmichael and Friedman 1996). In *Gnetum*, two sperm are released from the pollen tube into the female gametophyte. Each sperm fuses with a female nucleus to form a zygote. Multiple pollen tubes may penetrate the female gametophyte, and each may release two sperm cells. Ultimately, one embryo out-competes the rest (Carmichael and Friedman 1996).

In *W. mirabilis*, pollen settles into a deep pollen chamber formed from degenerating nucellar cells (Carafa et al. 1992). Pollen germination in *W. mirabilis* is different from *Gnetum*. The pollen grain exine splits at the distal sulcus, but the exine is not completely shed (Rydin and Friis 2005). The male gametophyte swells into a spherical shape, and the pollen tube grows out from this point (Rydin and Friis 2005). The pollen tube grows through the nucellus towards the female gametophyte. The female gametophyte of *W. mirabilis* is peculiar in that cells at the micropylar end extend upwards through the nucellar tissue forming prothallial tubes (Martens 1974 in Singh 1978). The pollen tube grows into the nucellus to meet the tip of the prothallial tubes, where the male gametes are discharged (Singh 1978).

Here we use proteomics to analyze the pollination drops from the fertile ovules of *G. gnemon* and *W. mirabilis*, as well as pollination drops collected from the sterile ovules of male cones of *W. mirabilis*. Previously (Chapter 2), we reported the proteome content of seven species of *Ephedra*. We found that in addition to secretome proteins, a strong degradome was present in *Ephedra*. The proteome of *W. mirabilis* was analyzed by Wagner et al. (2007) but only one protein, a chitinase, was identified. We set out to test the hypothesis that pollination drops from the fertile ovules of all gnetalean genera would have similar protein profiles, given that they are, together, considered a unique lineage.

Second, we hypothesized that the pollination drops collected from fertile and sterile ovules would have different protein profiles, since the ovule of the male is essentially non-functional in a reproductive capacity.

Methods

Pollination drop collection from *Gnetum gnemon*

Pollination drops were collected from the fertile ovules of female *G. gnemon* plants by Dr. Dennis W. Stevenson at the Nolan Glasshouse in the New York Botanical Garden. Drops were collected with a 10 µL pipette tip directly from the ovules between 2 and 4 pm. Samples were expelled into 1.5 mL Eppendorf tubes and stored at -20 °C until proteomic analysis.

Pollination drop collection from *Welwitschia mirabilis*

Pollination drops were collected from fertile ovules on female plants and sterile ovules on male plants of *W. mirabilis* by Dr. Stefan Little in May 2011. Plants were located at the Sciences Laboratory Building Greenhouse and the Botanical Conservatory at the University of California at Davis. Pollination drop collections took place in the morning. Pollination drops are exposed on *W. mirabilis* plants, and so drops were collected directly from the plants using 10 or 20 µL filter-pipette tips. The pollination drops of *W. mirabilis* were viscous. Samples were expelled into 1.5 mL Eppendorf tubes and stored at – 20 ˚C until proteomic analysis.

Proteomics

Details of the proteomic analyses of *W. mirabilis* and *G. gnemon* pollination drops were presented in Chapter 3.

Results

Of the three samples investigated, *W. mirabilis* pollination drops from fertile ovules had the fewest number of proteins (Table 11, pg 125). Six proteins were identified, but only one had an acceptable BLASTp match to the TAIR10 database. Further probing of the protein hits by running the amino acid sequences of the corresponding open reading frames through BLASTp against the NCBI non-redundant database (all species taxonomy) did not result in any additional acceptable BLASTp matches. The pollination drop sample from the sterile ovule of *W. mirabilis* had 139 proteins, which is the greatest number of proteins detected in the Gnetales to date (Table 12, pgs 126-141). Of those, 138 had acceptable BLASTp matches to the TAIR 10 database. The *G. gnemon* sample had an intermediate number; 24 proteins were detected, 17 of which had good BLASTp hits to TAIR 10 (Table 13, pgs 142-143).

The single identifiable protein in the female *W. mirabilis* sample was HOPZ-ACTIVATED RESISTANCE 1, a protein with a GO annotation to the nucleus and a GO annotation for Biological Process to defence response (Table 11). Of the 138 proteins identified with acceptable TAIR 10 matches in the *W. mirabilis* male sample, the majority, 130 (94 %), had GO annotations to intracellular locations (Table 12). There were 52 proteins (38 %) with annotations to extracellular locations. Similarly, there were 51 proteins (37 %) with annotations specifically to the plasma membrane. GO annotations are not mutually exclusive.

According to the GO annotations for Biological Process, 17 % of the proteins in the male *W. mirabilis* sample were involved in protein translation; 14 % were involved in oxidation-reduction processes; 12 % were involved in protein folding (Table 12). Other categories included: glycolysis 4 %; intracellular transport 4 %; carbohydrate metabolic

Table 11. Proteins identified in *Welwitschia mirabilis* fertile ovule (female cone) pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model.

Accession	-10lgP	Total Peptides	Unique Peptides	TAIR10 Gene Model	TAIR10 Description	BLASTP e value	Gene Ontology Biological Process	Gene Ontology Cellular Component
scaffold-GTHK-2000777-Gnetum_montanum_single_42	98.44	2	2					
scaffold-FHST-2012955-Taxodium_distichum_single_27	57.26	1	1					
scaffold-JDQB-2000365-Neocallitropsis_pancheri_sub2_34	56.68	1	1	AT3G50950.2	HOPZ-ACTIVATED RESISTANCE 1	3e-36	defense response	nucleus
scaffold-FMWZ-2034159-Dacrycarpus_compactus_single_4	53.22	1	1					
Dougfir-megastigmus-comp524843_c6_seq1_7	46.79	1	1					
scaffold-CDFR-2065652-Manoao_colensoi_single_67	45.83	1	1					

Table 12. Proteins identified in *Welwitschia mirabilis* sterile ovule (male cone) pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model.

Accession	-10lgP	Total Peptides	Unique Peptides	TAIR10 Gene Model	TAIR10 Description	BLASTP e value	Gene Ontology Biological Process	Gene Ontology Cellular Component
scaffold-GTHK-2013894-Gnetum_montanum_sub1_22	261.81	8	1	AT5G42020.1	Heat shock protein 70 (Hsp 70) family protein	0	response to cadmium ion	nucleus; vacuolar membrane; cell wall; chloroplast; nucleolus; plasma membrane; vacuole; plasmodesma; membrane
scaffold-CGDN-2003586-Tetraclinis_sp._sub1_50	247.97	7	1	AT5G42020.1	Heat shock protein 70 (Hsp 70) family protein	0	response to cadmium ion	nucleus; vacuolar membrane; cell wall; chloroplast; nucleolus; plasma membrane; vacuole; plasmodesma; membrane
Gnetum_gnom_isotig01036_18	243.33	6	6	AT3G08590.1	Phosphoglycerate mutase, 2,3-bisphosphoglycerate-independent	0	metabolic process; glucose catabolic process; response to cadmium ion	cytoplasm; cytosol; plasmodesma; apoplast

Sciadopitys_verti_isotig10884_21	234.6	8	6	AT5G02500.1	heat shock cognate protein 70-1	0	response to cadmium ion	membrane; plasmodesma; cytosolic ribosome; Golgi apparatus; vacuolar membrane; apoplast; cell wall; nucleolus; nucleus; plasma membrane; chloroplast
scaffold-FMWZ-2059145-Dacrycarpus_compactus_single_23	228.09	6	1	AT5G42020.1	Heat shock protein 70 (Hsp 70) family protein	0	response to cadmium ion	nucleus; vacuolar membrane; cell wall; chloroplast; nucleolus; plasma membrane; vacuole; plasmodesma; membrane
scaffold-GTHK-2008718-Gnetum_montanum_sub2_13	226.1	8	1	AT5G08680.1	ATP synthase alpha/beta family protein	0	ATP hydrolysis coupled proton transport; ATP biosynthetic process; ATP catabolic process; ATP synthesis coupled proton transport; proton transport; ATP metabolic process; response to cadmium ion	mitochondrial proton-transporting ATP synthase complex, catalytic core F(1); proton-transporting ATP synthase complex, catalytic core F(1); mitochondrion; proton-transporting two-sector ATPase complex, catalytic domain; membrane; proton-transporting two-sector ATPase complex
scaffold-GTHK-2056904-Gnetum_montanum_single_14	208.13	6	1	AT2G36530.1	Enolase	0	glycolysis; response to salt stress; response to cadmium ion	phosphopyruvate hydratase complex; chloroplast; plasmodesma; plasma membrane; cytosol; cytoplasm; apoplast; membrane
Taxus_bac_isotig03689_56	201.89	4	1	AT5G56000.1	HEAT SHOCK PROTEIN 81.4	0	protein folding	vacuolar membrane; chloroplast stroma; Golgi apparatus; cytoplasm; plasma membrane; apoplast; cell wall

scaffold-GTHK-2004667-Gnetum_montanum_single_45	201.76	7	1	AT3G12780.1	phosphoglycerate kinase 1	0	glycolysis; response to cadmium ion	chloroplast stroma; chloroplast; cell wall; chloroplast envelope; thylakoid; cytosol; membrane; nucleus; apoplast
scaffold-GTHK-2014932-Gnetum_montanum_sub1_14	199.26	3	3	AT3G02520.1	general regulatory factor 7	1e-129		plasma membrane; chloroplast; Golgi apparatus; plasmodesma; cytoplasm
scaffold-CGDN-2007304-Tetraclinis_sp._sub2_12	195.98	7	3	AT5G60390.1	GTP binding Elongation factor Tu family protein	0	translational elongation	cytoplasm; vacuole; plasma membrane; plasmodesma
scaffold-GTHK-2010284-Gnetum_montanum_single_50	192.74	5	1	AT1G56070.1	Ribosomal protein S5/Elongation factor G/III/V family protein	0		nucleolus; plasma membrane; vacuolar membrane; membrane; cytoplasm; cytosol; chloroplast; plasmodesma
scaffold-CGDN-2003149-Tetraclinis_sp._single_15	188.88	4	1	AT5G16970.1	alkenal reductase	1e-94	response to cadmium ion	cytosol; cytoplasm
Sequoia_semp_isotig05393_40	177.69	3	3	AT5G09810.1	actin 7	0		chloroplast stroma; chloroplast envelope; cell wall; nucleolus; plasma membrane; plasmodesma; cytoplasm; cytosol
scaffold-GGEA-2012180-Cedrus_libani_single_7	174.94	5	4	AT1G68850.1	Peroxidase superfamily protein	3e-81	oxidation-reduction process; response to oxidative stress	extracellular region
Ginkgo_Contig59576_27	169.97	3	3	AT1G72730.1	DEA(D/H)-box RNA helicase family protein	1e-157		vacuolar membrane; cytoplasm
scaffold-FMWZ-2006561-Dacrycarpus_compactus_sub2_26	169.93	4	4	AT3G55610.1	delta 1-pyrroline-5-carboxylate synthase 2	0		mitochondrion; cytoplasm; plasmodesma; cytosol
scaffold-GTHK-2057331-Gnetum_montanum_single_25	169.15	2	2	AT1G78900.1	vacuolar ATP synthase subunit A	0	response to salt stress; proton transport; ATP metabolic process; ATP hydrolysis coupled proton transport	vacuole; chloroplast; Golgi apparatus; cell wall; plasmodesma; plasma membrane; plant-type vacuole; chloroplast envelope; vacuolar membrane; membrane; mitochondrion

scaffold-FHST-2063930-Taxodium_distichum_single_38	164.89	3	2	ATMG01190.1	ATP synthase subunit 1	0	ATP synthesis coupled proton transport; ATP hydrolysis coupled proton transport; ATP metabolic process; proton transport	proton-transporting two-sector ATPase complex; proton-transporting ATP synthase complex, catalytic core F(1); cell wall; mitochondrion; proton-transporting two-sector ATPase complex, catalytic domain
Gnetum_gnom_isotig03749_19	164.73	2	2	AT5G03340.1	ATPase, AAA-type, CDC48 protein	0		Golgi apparatus; cytosol; cell wall; nucleus
scaffold-GNQG-2083787-Encephalartos_barteri_single_17	163.48	2	1	AT5G42020.1	Heat shock protein 70 (Hsp 70) family protein	2e-120	response to cadmium ion	nucleus; vacuolar membrane; cell wall; chloroplast; nucleolus; plasma membrane; vacuole; plasmodesma; membrane
scaffold-IOVS-2008275-Pseudotsuga_menziesii_single_8	162.92	3	3	AT5G65730.1	xyloglucan endotransglucosylase/hydrolase 6	1e-27	cellular glucan metabolic process; carbohydrate metabolic process	cell wall; extracellular region; apoplast
scaffold-GAMH-2008637-Tsuga_heterophylla_single_43	161.15	5	1	AT5G60390.1	GTP binding Elongation factor Tu family protein	0	translational elongation	cytoplasm; vacuole; plasma membrane; plasmodesma
Gnetum_gnom_isotig10086_53	157.41	2	2	AT4G24190.1	Chaperone protein htpG family protein	0	protein folding; response to cadmium ion	vacuolar membrane; plasmodesma; nucleus; vacuole; membrane; plasma membrane; chloroplast
scaffold-IFLI-2145561-Callitris_gracilis-March_single_27	156.64	5	1	AT5G08570.1	Pyruvate kinase family protein	0	glycolysis	cytosol; cytoplasm
Taxus_bac_isotig11662_24	154.05	3	2	AT1G55490.1	chaperonin 60 beta	0	cellular protein metabolic process; protein folding; protein refolding	cytosolic ribosome; chloroplast stroma; chloroplast envelope; membrane; apoplast; plasma membrane; cytoplasm; chloroplast
scaffold-IIOL-2076512-Pinus_parviflora_single_13	152.16	2	1	AT2G22480.1	phosphofructokinase 5	7e-144	glycolysis	6-phosphofructokinase complex; chloroplast

scaffold-GGEA-2068291-Cedrus_libani_single_56	150.36	2	2	AT3G06720.1	importin alpha isoform 1	0	protein import into nucleus	nucleus; cytoplasm; nucleolus; cell wall; cytosol
Pseudotsuga_menz_isotig11281_39	149.79	2	1	AT5G19770.1	tubulin alpha-3	0	GTP catabolic process; microtubule-based movement; microtubule-based process; protein polymerization	plasma membrane; plasmodesma; cytoplasm; cell wall
scaffold-GTHK-2015333-Gnetum_montanum_single_67	149.58	4	1	AT1G70730.3	Phosphoglucomutase/phosphomannomutase family protein	0	carbohydrate metabolic process	cytoplasm; cytosol
scaffold-EGLZ-2010503-Prumnopitys_andina_single_17	149.39	4	1	AT5G16970.1	alkenal reductase	2e-96	response to cadmium ion	cytosol; cytoplasm
scaffold-CGDN-2019601-Tetraclinis_sp._single_3	147.5	3	3	AT5G42020.1	Heat shock protein 70 (Hsp 70) family protein	7e-11	response to cadmium ion	nucleus; vacuolar membrane; cell wall; chloroplast; nucleolus; plasma membrane; vacuole; plasmodesma; membrane
Sciadopitys_verti_isotig08139_8	145.63	2	2	AT5G35530.1	Ribosomal protein S3 family protein	8e-149	response to salt stress; translation	cytoplasm; plasmodesma; cytosolic ribosome; cytosol; nucleolus; small ribosomal subunit; ribosome; membrane; intracellular
Gnetum_gnom_isotig00037_25	145.3	4	4	AT4G39230.1	NmrA-like negative transcriptional regulator family protein	1e-93	response to cadmium ion	cytoplasm
Pseudotsuga_menz_isotig04598_44	141.68	2	1	AT1G50010.1	tubulin alpha-2 chain	0	GTP catabolic process; protein polymerization; microtubule-based process; microtubule-based movement; response to salt stress	cell wall; microtubule; membrane; cytoplasm; protein complex

Ginkgo_Contig67500_70	138.84	2	1	AT2G07698.1	ATPase, F1 complex, alpha subunit protein	0	ATP hydrolysis coupled proton transport; ATP metabolic process; ATP synthesis coupled proton transport; proton transport	plasma membrane; mitochondrion; vacuolar membrane; membrane; chloroplast envelope; chloroplast; proton-transporting ATP synthase complex, catalytic core F(1); proton-transporting two-sector ATPase complex; vacuole; proton-transporting two-sector ATPase complex, catalytic domain; nucleolus
Gnetum_gnom_isotig06648_34	138.35	2	2	AT5G59240.1	Ribosomal protein S8e family protein	5e-117	translation	cytosolic ribosome; membrane; cytoplasm
scaffold-GNQG-2086502-Encephalartos_barteri_single_49	136.86	2	2	AT5G20890.1	TCP-1/cpn60 chaperonin family protein	0	protein folding; cellular protein metabolic process	cell wall; anchored to plasma membrane; cytoplasm; cytosol
Pinus_taeda_isotig25862_23	136.53	3	3	AT4G25050.2	acyl carrier protein 4	9e-36		chloroplast
Sequoia_semp_isotig12955_8	135.56	2	2	AT4G02080.1	secretion-associated RAS super family 2	3e-134	intracellular protein transport	plasma membrane
scaffold-ACWS-2067405-Arucaria_sp._single_19	132.48	3	3	AT4G25740.1	RNA binding Plectin/S10 domain-containing protein	2e-72		cytosolic ribosome; cytosol; membrane
scaffold-GTHK-2056450-Gnetum_montanum_single_13	126.6	2	1	AT1G09210.1	calreticulin 1b	0	response to salt stress; protein folding	vacuole; chloroplast
Pinus_lamb_isotig24551_17	130.78	2	1	AT1G65930.1	cytosolic NADP+-dependent isocitrate dehydrogenase	0	isocitrate metabolic process; oxidation-reduction process; response to salt stress; response to cadmium ion	apoplast; chloroplast stroma; chloroplast; plasmodesma; cytosol; plasma membrane
Gnetum_gnom_isotig10498_47	123.97	2	1	AT1G55490.1	chaperonin 60 beta	0	cellular protein metabolic process; protein folding; protein refolding	cytosolic ribosome; chloroplast stroma; chloroplast envelope; membrane; apoplast; plasma membrane; cytoplasm; chloroplast
scaffold-CGDN-2006923-Tetraclinis_sp._single_47	123.64	3	3	AT2G21130.1	Cyclophilin-like peptidyl-prolyl cis-trans isomerase family protein	1e-101	protein folding	cytoplasm; cytosol; plasma membrane

Taxus_bac_isotig08127_17	123.45	2	1	AT4G09320.1	Nucleoside diphosphate kinase family protein	3e-93	nucleoside diphosphate phosphorylation; response to salt stress; UTP biosynthetic process; CTP biosynthetic process; GTP biosynthetic process; response to cadmium ion	vacuolar membrane; cytoplasm; chloroplast; plasma membrane; plasmodesma; vacuole; apoplast; cytosol
Sciadopitys_verti_isotig07658_20	123.25	1	1	AT5G59850.1	Ribosomal protein S8 family protein	1e-88	translation	intracellular; ribosome; chloroplast; cell wall; membrane
Gnetum_gnom_isotig04644_15	120.64	2	1	AT1G34030.1	Ribosomal protein S13/S18 family	3e-102	translation	vacuolar membrane; cytosolic ribosome; intracellular; membrane; plasmodesma; plasma membrane; cytoplasm; cell wall; ribosome
scaffold-GTHK-2055300-Gnetum_montanum_single_21	120.27	2	2	AT5G09500.1	Ribosomal protein S19 family protein	5e-86	translation	small ribosomal subunit; cytoplasm; ribosome
Gnetum_gnom_isotig09735_41	119.92	1	1	AT3G03250.1	UDP-GLUCOSE PYROPHOSPHORYLASE 1	0	metabolic process; response to salt stress	plasma membrane; cytoplasm; cytosol
scaffold-FRPM-2008332-Calocedrus_decurrens_single_44	117.39	2	1	AT1G64190.1	6-phosphogluconate dehydrogenase family protein	0	response to salt stress; pentose-phosphate shunt; oxidation-reduction process	membrane; cytoplasm; chloroplast stroma; cytosol; chloroplast
scaffold-AQFM-2082975-Pseudolarix_amabilis_single_60	116.84	2	2	AT2G28000.1	chaperonin-60alpha	0	cellular protein metabolic process; protein refolding; protein folding	cytosolic ribosome; thylakoid; chloroplast envelope; chloroplast stroma; chloroplast; apoplast; mitochondrion; cytoplasm; membrane
Gnetum_gnom_isotig10182_14	115.67	1	1	AT1G16300.1	glyceraldehyde-3-phosphate dehydrogenase of plastid 2	0	oxidation-reduction process; glucose metabolic process	chloroplast

Gnetum_gnom_isotig10720_7	115.17	1	1	AT3G09640.1	ascorbate peroxidase 2	2e-140	response to oxidative stress; oxidation-reduction process	cytoplasm
scaffold-GTHK-2056801-Gnetum_montanum_single_45	115.03	1	1	AT1G68010.2	hydroxypyruvate reductase	0		
scaffold-GTHK-2057168-Gnetum_montanum_single_54	114.35	2	1	AT3G03960.1	TCP-1/cpn60 chaperonin family protein	0	protein folding; cellular protein metabolic process	cytosol; membrane; cytoplasm; plasmodesma
Cycas-rumphii-NODE_6213_length_496_cov_0.554435_8	114.28	2	1	AT1G26320.1	Zinc-binding dehydrogenase family protein	3e-57	oxidation-reduction process	cytoplasm
scaffold-IIOL-2016225-Pinus_parviflora_sub2_23	113.94	2	2	AT3G52990.1	Pyruvate kinase family protein	1e-137	glycolysis; response to cadmium ion	membrane; cytosol
scaffold-GNQG-2003545-Encephalartos_barteri_sub2_68	112.71	3	1	AT1G56070.1	Ribosomal protein S5/Elongation factor G/III/V family protein	0		nucleolus; plasma membrane; vacuolar membrane; membrane; cytoplasm; cytosol; chloroplast; plasmodesma
Sciadopitys_verti_isotig08815_19	111.09	2	1	AT1G65930.1	cytosolic NADP+-dependent isocitrate dehydrogenase	0	isocitrate metabolic process; oxidation-reduction process; response to salt stress; response to cadmium ion	apoplast; chloroplast stroma; chloroplast; plasmodesma; cytosol; plasma membrane
scaffold-IIOL-2001767-Pinus_parviflora_sub2_15	110.83	1	1	AT3G53430.1	Ribosomal protein L11 family protein	8e-105	translation	membrane; ribosome; cytoplasm
scaffold-AQFM-2080247-Pseudolarix_amabilis_single_34	110.13	1	1	AT4G00680.1	actin depolymerizing factor 8	4e-75		intracellular; cytoplasm
scaffold-JBND-2069565-Pinus_ponderosa_single_26	107.76	2	2	AT1G08830.1	copper/zinc superoxide dismutase 1	3.00E-81		cytosol; cytoplasm
Pinus_taeda_isotig11825_43	105.65	2	1	AT1G56340.1	calreticulin 1a	0	response to salt stress; response to cadmium ion; protein folding	plasmodesma; chloroplast; vacuolar membrane
picea_abies_isotig08479_27	105.05	2	2	AT3G23990.1	heat shock protein 60	0	protein folding; cellular protein metabolic process; protein refolding; response to cadmium ion	cytosolic ribosome; vacuolar membrane; mitochondrion

scaffold-JRNA-2060886-Phyllocladus_hypohyllus_single_5	103.39	1	1	AT4G00100.1	ribosomal protein S13A	6e-94	translation	ribosome; nucleolus; intracellular; membrane; chloroplast
picea_abies_isotig04270_6	102.89	1	1	AT2G37270.1	ribosomal protein 5B	1e-130	translation	vacuolar membrane; small ribosomal subunit; cytosol; membrane; chloroplast; cytosolic ribosome; intracellular; plasma membrane; cell wall; ribosome
Pseudotsuga_menz_isotig16021_21	102.41	1	1	AT5G45775.2	Ribosomal L5P family protein	8e-125	translation	vacuole; ribosome; intracellular; cytoplasm
scaffold-GTHK-2015292-Gnetum_montanum_sub2_30	101.97	1	1	AT4G31990.3	aspartate aminotransferase 5	6e-158	biosynthetic process; response to cadmium ion; cellular amino acid metabolic process	apoplast; chloroplast
scaffold-AQFM-2082250-Pseudolarix_amabilis_single_12	101.45	1	1	AT1G21750.1	PDI-like 1-1	0	glycerol ether metabolic process; cell redox homeostasis; response to salt stress	vacuole; thylakoid; plant-type cell wall; membrane; chloroplast; endoplasmic reticulum
scaffold-AIGO-2068573-Chamaecyparis_lawsoniana_single_12	99.85	1	1	AT5G56450.1	Mitochondrial substrate carrier family protein	3e-158	transmembrane transport; transport	mitochondrion; mitochondrial inner membrane
Dougfir-megastigmus-comp568853_c0_seq1_31	98.63	1	1	AT1G04410.1	Lactate/malate dehydrogenase family protein	0	carbohydrate metabolic process; response to salt stress; malate metabolic process; oxidation-reduction process; cellular carbohydrate metabolic process; response to cadmium ion	vacuolar membrane; chloroplast stroma; plasma membrane; vacuole; nucleus; chloroplast; membrane; cytoplasm; plasmodesma; apoplast
scaffold-GAMH-2010422-Tsuga_heterophylla_single_91	98.43	1	1	AT1G52360.2	Coatomer, beta' subunit	0	intracellular protein transport; vesicle-mediated transport	membrane coat

Ginkgo_Contig27736_11	97.6	1	1	AT4G37930.1	serine transhydroxymethyltransferase 1	1e-168	response to cadmium ion	chloroplast; mitochondrion; chloroplast thylakoid; chloroplast stroma; cytosolic ribosome; apoplast; membrane; nucleus; plasma membrane
Pinus_taeda_isotig34138_16	96.4	1	1	AT5G51890.1	Peroxidase superfamily protein	1e-70	oxidation-reduction process; response to oxidative stress	extracellular region
scaffold-BBDD-2008784-Microstrobos_fitzgeraldii_single_51	95.35	2	1	AT5G13490.1	ADP/ATP carrier 2	0	transmembrane transport; transport	mitochondrion; chloroplast; vacuolar membrane; chloroplast envelope; membrane
Gnetum_gnom_isotig03834_19	94.38	1	1	AT3G05590.1	ribosomal protein L18	2e-111	translation	chloroplast; cytoplasm; vacuole; vacuolar membrane; intracellular; plasma membrane; ribosome
scaffold-IAJW-2147976-Amentotaxus_argotaenia_single_24	94.24	1	1	AT3G16640.1	translationally controlled tumor protein	5e-89	response to cadmium ion; defense response to bacterium	plasmodesma; vacuolar membrane; cytoplasm; nucleus; cytosol; chloroplast; plasma membrane; Golgi apparatus; thylakoid; apoplast
scaffold-FHST-2003197-Taxodium_distichum_single_28	94.1	1	1	AT5G39740.1	ribosomal protein L5 B	3e-174	translation	nucleolus; plasma membrane; vacuole; cytosolic ribosome; ribosome; intracellular; cytoplasm; cytosol
scaffold-GTHK-2057477-Gnetum_montanum_single_91	93.22	1	1	AT4G31480.1	Coatomer, beta subunit	0	intracellular protein transport; vesicle-mediated transport	membrane coat; nucleus; cytoplasm; COPI vesicle coat; plasmodesma; cytosol
scaffold-JBND-2010831-Pinus_ponderosa_sub1_2	92.31	3	3	AT2G21660.1	cold, circadian rhythm, and rna binding 2	2e-41	response to cadmium ion; response to salt stress	peroxisome; chloroplast; cytosol

Metasequoia_Contig7367_23	91.96	2	2	AT5G15200.1	Ribosomal protein S4	6e-124	translation	small ribosomal subunit; chloroplast; cytosolic ribosome; intracellular; nucleolus; cell wall; mitochondrion; plasmodesma; vacuolar membrane; membrane; cytosol
scaffold-IZGN-2006995-Dacrydium_balansae_single_30	91.86	1	1	AT1G35720.1	annexin 1	1e-130	response to cadmium ion	plasma membrane; plasmodesma; vacuole; apoplast; nucleus; vacuolar membrane; chloroplast stroma; chloroplast; mitochondrion; thylakoid; cell wall
Pinus_lamb_isotig10342_15	91.74	1	1	AT1G65980.1	thioredoxin-dependent peroxidase 1	9e-89	response to cadmium ion	membrane; cytoplasm; plasma membrane; cytosol; chloroplast
Dougfir-megastigmus-comp570212_c0_seq1_11	90.68	2	1	AT4G09320.1	Nucleoside diphosphate kinase family protein	1e-92	nucleoside diphosphate phosphorylation; response to salt stress; UTP biosynthetic process; CTP biosynthetic process; GTP biosynthetic process; response to cadmium ion	vacuolar membrane; cytoplasm; chloroplast; plasma membrane; plasmodesma; vacuole; apoplast; cytosol
scaffold-JDQB-2000171-Neocallitropsis_pancheri_single_17	89.25	2	2	AT2G24270.4	aldehyde dehydrogenase 11A3	0	oxidation-reduction process; metabolic process	cytoplasm
Sciadopitys_verti_isotig10683_29	89.08	1	1	AT1G48630.1	receptor for activated C kinase 1B	0		cytosolic ribosome; cytosol; cytoplasm
Podocarpus_macro_isotig09394_13	88.84	2	1	AT3G03960.1	TCP-1/cpn60 chaperonin family protein	4e-80	protein folding; cellular protein metabolic process	cytosol; membrane; cytoplasm; plasmodesma
scaffold-GTHK-2003109-Gnetum_montanum_single_10	87.62	1	1	AT3G12130.1	KH domain-containing protein / zinc finger (CCCH type) family protein	4e-06		nucleus
scaffold-ACWS-2013189-Arucaria_sp._sub2_35	87.38	1	1	AT1G64980.1	Nucleotide-diphospho-sugar transferases superfamily protein	7e-129		nucleus; cytosol

scaffold-AQFM-2081022-Pseudolarix_amabilis_single_11	86.89	2	2	AT5G39740.1	ribosomal protein L5 B	1e-166	translation	nucleolus; plasma membrane; vacuole; cytosolic ribosome; ribosome; intracellular; cytoplasm; cytosol
scaffold-HQOM-2013154-Torreya_nucifera_single_16	86.07	1	1	AT5G05010.1	clathrin adaptor complexes medium subunit family protein	0	intracellular protein transport; transport; vesicle-mediated transport	cytosol; clathrin adaptor complex; membrane; plasmodesma
scaffold-GKCZ-2004624-Diselma_archeri_sub1_10	85.98	1	1	AT3G23990.1	heat shock protein 60	0	protein folding; cellular protein metabolic process; protein refolding; response to cadmium ion	cytosolic ribosome; vacuolar membrane; mitochondrion
scaffold-GTHK-2057502-Gnetum_montanum_single_97	85.41	1	1	AT4G11420.1	eukaryotic translation initiation factor 3A	0		plasma membrane; cytosol; cytoplasm
Ginkgo_Contig74004_2	84.77	1	1	AT5G56600.1	profilin 3	9e-29		chloroplast; actin cytoskeleton
scaffold-GTHK-2056113-Gnetum_montanum_single_30	83.2	1	1	AT5G63400.1	adenylate kinase 1	1e-144	nucleobase-containing compound metabolic process; nucleotide phosphorylation; response to cadmium ion	cytoplasm; mitochondrion; vacuolar membrane
scaffold-IOVS-2003855-Pseudotsuga_menziesii_single_13	83.12	2	1	AT5G58400.1	Peroxidase superfamily protein	2e-37	response to oxidative stress; oxidation-reduction process	extracellular region
scaffold-HILW-2010734-Acmopyle_pancheri_sub2_16	83.01	1	1	AT3G29360.1	UDP-glucose 6-dehydrogenase family protein	0		cytosol; cytoplasm
Pinus_taeda_isotig17431_12	82.12	2	1	AT5G06720.1	peroxidase 2	2e-119	oxidation-reduction process; response to oxidative stress	extracellular region; Golgi apparatus
scaffold-HOUF-2002024-Cryptomeria_japonica-leaf_sub2_9	79.98	1	1	AT5G50850.1	Transketolase family protein	0	defense response to bacterium; metabolic process	vacuolar membrane; mitochondrion; nucleolus

Taxus_bac_isotig05468_11	79.9	1	1	AT3G62870.1	Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein	3e-160	ribosome biogenesis	cytosolic ribosome; cytoplasm; membrane; vacuolar membrane; cytosol; ribonucleoprotein complex
scaffold-HILW-2114528-Acmopyle_pancheri_single_70	76.72	1	1	AT4G23100.1	glutamate-cysteine ligase	0	glutathione biosynthetic process; response to cadmium ion	chloroplast stroma; chloroplast
scaffold-ACWS-2065367-Arucaria_sp._single_10	76.54	1	1	AT3G05560.1	Ribosomal L22e protein family	1e-54	translation	cytosolic ribosome; nucleolus; plasma membrane; ribosome; cytosol; intracellular; plasmodesma; cytoplasm
Podocarpus_macro_isotig08803_14	76.1	1	1	AT1G31812.1	acyl-CoA-binding protein 6	1e-37		cytosol; cytoplasm; plasma membrane
scaffold-GGEA-2068351-Cedrus_libani_single_50	75.65	1	1	AT2G28000.1	chaperonin-60alpha	0	cellular protein metabolic process; protein refolding; protein folding	cytosolic ribosome; thylakoid; chloroplast envelope; chloroplast stroma; chloroplast; apoplast; mitochondrion; cytoplasm; membrane
Ginkgo_Contig69164_58	75.43	1	1	AT3G29360.1	UDP-glucose 6-dehydrogenase family protein	0		cytosol; cytoplasm
Gnetum_gnom_isotig12850_13	74.29	1	1	AT4G16160.2	Mitochondrial import inner membrane translocase subunit Tim17/Tim22/Tim23 family protein	2e-36		
Gnetum_gnom_isotig06697_10	72.95	1	1	AT1G77710.1		1e-54		
Ginkgo_Contig77008_45	72.55	1	1	AT4G01850.1	S-adenosylmethionine synthetase 2	0	S-adenosylmethionine biosynthetic process	cell wall; cytoplasm; nucleolus; plasmodesma
Pseudotsuga_menz_isotig03589_31	72.39	1	1	AT5G16990.1	Zinc-binding dehydrogenase family protein	2e-154	oxidation-reduction process	plasma membrane; cytoplasm; cytosol
scaffold-GTHK-2055122-Gnetum_montanum_single_20	72.27	1	1	AT1G31812.1	acyl-CoA-binding protein 6	8e-39		cytosol; cytoplasm; plasma membrane

Cycas-rumphii-NODE_16297_length_510_cov_6.872549_11	71.68	1	1	AT3G47370.2	Ribosomal protein S10p/S20e family protein	2e-73	translation	plasmodesma; cell wall; cytosolic ribosome; small ribosomal subunit; intracellular; membrane; cytosol; nucleolus; chloroplast; ribosome
scaffold-GKCZ-2006482-Diselma_archeri_sub1_1	70.06	1	1	AT5G60670.1	Ribosomal protein L11 family protein	7e-19	translation	ribosome; cytosol; cytoplasm
picea_abies_isotig06588_15	69.8	1	1	AT4G11600.1	glutathione peroxidase 6	4e-93	response to salt stress; oxidation-reduction process; response to oxidative stress; response to cadmium ion	mitochondrion; cytosol; plasma membrane; chloroplast
scaffold-GTHK-2003605-Gnetum_montanum_sub2_22	66.99	1	1	AT3G09630.1	Ribosomal protein L4/L1 family	6e-165	translation	cytosolic ribosome; chloroplast; nucleolus; membrane; plasmodesma; vacuole; plasma membrane; cytosol; cell wall; ribosome; cytoplasm
scaffold-BBDD-2007880-Microstrobos_fitzgeraldii_single_6	66.33	1	1	AT5G62300.1	Ribosomal protein S10p/S20e family protein	3e-71	translation	intracellular; small ribosomal subunit; cell wall; ribosome; chloroplast
scaffold-IIOL-2076738-Pinus_parviflora_single_9	65.7	1	1	AT1G32470.1	Single hybrid motif superfamily protein	2e-68	glycine catabolic process; glycine decarboxylation via glycine cleavage system	mitochondrion; glycine cleavage complex; chloroplast
scaffold-AREG-2068507-Nothotsuga_longibracteata_single_48	65.58	1	1	AT5G07350.2	TUDOR-SN protein 1	0	gene silencing by RNA	cytosol
scaffold-EGLZ-2037573-Prumnopitys_andina_single_30	65.15	1	1	AT3G55440.1	triosephosphate isomerase	7e-143	response to salt stress; response to cadmium ion; metabolic process	chloroplast stroma; cytosol; cytoplasm; vacuolar membrane; vacuole; chloroplast; plasma membrane; cell wall; apoplast; plasmodesma

Ginkgo_Contig76583_27	64.83	1	1	AT5G02500.1	heat shock cognate protein 70-1	3e-05	response to cadmium ion	membrane; plasmodesma; cytosolic ribosome; Golgi apparatus; vacuolar membrane; apoplast; cell wall; nucleolus; nucleus; plasma membrane; chloroplast
scaffold-EFMS-2024243-Torreya_taxifolia_single_4	63.98	1	1	AT5G09590.1	mitochondrial HSO70 2	2e-13	response to salt stress; response to cadmium ion; protein folding	cell wall; vacuolar membrane; mitochondrion; chloroplast
Gnetum_gnom_isotig06431_20	63.89	1	1	AT3G44100.1	MD-2-related lipid recognition domain-containing protein	4e-41		extracellular region; vacuole; cell wall
Pinus_lamb_isotig27842_17	61.83	1	1	AT4G34200.1	D-3-phosphoglycerate dehydrogenase	0	L-serine biosynthetic process; oxidation-reduction process; metabolic process	membrane; chloroplast; chloroplast stroma; mitochondrion
Dougfir-megastigmus-comp75096_c0_seq1_5	61.42	1	1					
picea_abies_isotig07141_45	60.61	1	1	AT4G02930.1	GTP binding Elongation factor Tu family protein	0	translational elongation	mitochondrion; intracellular; cell wall
scaffold-BBDD-2078298-Microstrobos_fitzgeraldii_single_83	59.09	1	1	AT2G04030.1	Chaperone protein htpG family protein	0		chloroplast stroma; chloroplast; chloroplast envelope; vacuolar membrane; mitochondrion
Ginkgo_Contig76692_25	55.75	1	1	AT5G09590.1	mitochondrial HSO70 2	0	response to salt stress; response to cadmium ion; protein folding	cell wall; vacuolar membrane; mitochondrion; chloroplast
scaffold-FRPM-2002333-Calocedrus_decurrens_single_28	55.18	1	1	AT3G11130.1	Clathrin, heavy chain	0	intracellular protein transport; vesicle-mediated transport	clathrin coat of trans-Golgi network vesicle; vacuole; vacuolar membrane; clathrin coat of coated pit; plasma membrane; Golgi apparatus; plasmodesma
Ginkgo_Contig74291_4	54.89	1	1	AT5G17310.2	UDP-glucose pyrophosphorylase 2	9e-50	response to salt stress; response to cadmium ion; metabolic process	plasma membrane; cytosol; cytoplasm

scaffold-ACWS-2012455-Arucaria_sp._single_14	52.09	1	1	AT4G13930.1	serine hydroxymethyltransferase 4	0	response to cadmium ion	membrane; plasmodesma; cytosol; cytoplasm; plasma membrane
scaffold-ESYX-2013010-Cunninghamia_lanceolata-branch_apex_with_needles_single_61	51.75	1	1	AT5G42740.1	Sugar isomerase (SIS) family protein	0	response to cadmium ion; glycolysis; gluconeogenesis	cytoplasm; cytosol
scaffold-GJTI-2060992-Cephalotaxus_harringtonia_single_20	51.57	1	1	AT3G62120.1	Class II aaRS and biotin synthetases superfamily protein	0	tRNA aminoacylation for protein translation; prolyl-tRNA aminoacylation	cytoplasm; cytosol; plasmodesma; membrane
scaffold-ACWS-2068983-Arucaria_sp._single_13	50.76	1	1	AT2G47470.1	thioredoxin family protein	2e-171	cell redox homeostasis; glycerol ether metabolic process; response to cadmium ion	endoplasmic reticulum; vacuolar membrane; extracellular region
Dougfir-megastigmus-comp585917_c1_seq8_57	50.25	1	1	AT2G33150.1	peroxisomal 3-ketoacyl-CoA thiolase 3	0	metabolic process	chloroplast; nucleolus; vacuolar membrane; membrane; peroxisome
scaffold-AUDE-2001556-Widdringtonia_cedarbergensis_single_5	49.88	1	1	AT2G21730.1	cinnamyl alcohol dehydrogenase homolog 2	5e-34	oxidation-reduction process	Cytoplasm
Sciadopitys_verti_isotig09797_23	48.56	1	1	AT1G29880.1	glycyl-tRNA synthetase / glycine--tRNA ligase	0	glycyl-tRNA aminoacylation; response to cadmium ion; tRNA aminoacylation for protein translation	cytoplasm; cytosol
scaffold-GTHK-2001751-Gnetum_montanum_sub1_59	47.48	1	1	AT3G63400.1	Cyclophilin-like peptidyl-prolyl cis-trans isomerase family protein	5e-84	protein folding	Cytoplasm
Ginkgo_Contig73925_29	47.33	1	1	AT3G51800.2	metallopeptidase M24 family protein	5e-120	cellular process; proteolysis	nucleus; nucleolus; plasma membrane
scaffold-GAMH-2056169-Tsuga_heterophylla_single_29	46.44	1	1	AT4G30000.2	Dihydropterin pyrophosphokinase / Dihydropteroate synthase	0	pteridine-containing compound metabolic process; cellular metabolic process; folic acid-containing compound biosynthetic	mitochondrion

Table 13. Proteins identified in *Gnetum gnemon* fertile ovule (female) pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model.

Accession	-10lgP	Total Peptides	Unique Peptides	TAIR10 Gene Model	TAIR10 Description	BLASTP e value	Gene Ontology Biological Process	Gene Ontology Cellular Component
Gnetum_gnom_isotig02210_36	512.25	227	92	AT5G44400.1	FAD-binding Berberine family protein	9e-115	oxidation-reduction process	plasmodesma; cytoplasm; cell wall
Gnetum_gnom_isotig02209_36	509.6	200	65	AT5G44400.1	FAD-binding Berberine family protein	7e-115	oxidation-reduction process	plasmodesma; cytoplasm; cell wall
Gnetum_gnom_isotig04192_5	341.39	25	25					
Gnetum_gnom_isotig02404_22	293.92	16	16	AT3G57270.1	beta-1,3-glucanase 1	6e-81	carbohydrate metabolic process	extracellular region
Gnetum_gnom_isotig12272_26	247.89	9	9	AT2G43590.1	Chitinase family protein	8e-87	cell wall macromolecule catabolic process; carbohydrate metabolic process	extracellular region
Gnetum_gnom_isotig07000_20	238.57	11	11	AT4G11650.1	osmotin 34	8e-79	response to salt stress	extracellular region
Gnetum_gnom_isotig12955_15	238.41	10	10	AT2G21740.1	Protein of unknown function (DUF1278)	1e-07		extracellular region
Gnetum_gnom_isotig09998_39	188.47	6	6	AT5G67360.1	Subtilase family protein	0	negative regulation of catalytic activity; proteolysis	apoplast; extracellular region; cell wall; plant-type cell wall
scaffold-GTHK-2010671-Gnetum_montanum_single_94	170.21	5	5	AT4G34260.1	1,2-alpha-L-fucosidases	0		extracellular region; apoplast
Gnetum_gnom_isotig08558_8	147.14	5	4	AT5G08370.1	alpha-galactosidase 2	1e-117	carbohydrate metabolic process	extracellular region
Gnetum_gnom_isotig03842_23	129.91	2	2	AT3G52500.1	Eukaryotic aspartyl protease family protein	8e-81	proteolysis	membrane; extracellular region; cell wall; plant-type cell wall

Gnetum_gnom_isotig04737_12	90.49	1	1	AT4G30880.1	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein	1e-15		extracellular region
scaffold-ACWS-2010562-Arucaria_sp._sub2_25	86.89	2	1	AT5G08380.1	alpha-galactosidase 1	5e-105	carbohydrate metabolic process	apoplast; extracellular region; cell wall
scaffold-GNQG-2014362-Encephalartos_barteri_sub1_18	65.55	1	1	AT5G24318.1	O-Glycosyl hydrolases family 17 protein	2e-173	carbohydrate metabolic process	extracellular region
scaffold-GTHK-2010417-Gnetum_montanum_single_19	64.38	1	1	AT3G12490.2	cystatin B	3e-29		cytosol; extracellular region
scaffold-GTHK-2007440-Gnetum_montanum_single_9	62.79	1	1	AT5G05340.1	Peroxidase superfamily protein	4e-119	oxidation-reduction process; response to oxidative stress	extracellular region; apoplast; Golgi apparatus; cytosol; cell wall
scaffold-BUWV-2045555-Platycladus_orientalis_single_11	55.78	1	1	AT1G30760.1	FAD-binding Berberine family protein	2e-44	oxidation-reduction process	cytoplasm
scaffold-CGDN-2063804-Tetraclinis_sp._single_2	52.7	1	1					
scaffold-HOUF-2011124-Cryptomeria_japonica-leaf_sub1_9	51.88	1	1	AT5G23200.1		1e-49		
scaffold-ESYX-2015512-Cunninghamia_lanceolata-branch_apex_with_needles_sub1_26	51.58	1	1					
scaffold-DZQM-2005963-Pinus_radiata_single_22	50.75	1	1					
Ginkgo_Contig45341_40	50.03	1	1					
scaffold-AQFM-2006076-Pseudolarix_amabilis_sub2_6	46.97	1	1					
scaffold-GAMH-2006940-Tsuga_heterophylla_sub1_4	46.89	1	1					

process 2 %. Additionally, 30 % also had an annotation to response to cadmium ion, and 14 % response to salt stress. Many other GO categories for Biological Process are listed in Table 12.

In *G. gnemon*, 15 of the 17 proteins with good TAIR 10 matches had GO annotations to the extracellular space (Table 13). Just six proteins (35 %) had annotations to the intracellular space. None of the proteins had an annotation specifically to the plasma membrane. The best-represented Biological Process categories were quite different for *G. gnemon*, as compared to *W. mirabilis*. Carbohydrate metabolic process was the best-represented category for *G. gnemon* (29 %), followed by oxidation-reduction process (24 %), and then proteolysis (12 %). One protein had an annotation to response to salt stress.

Discussion

Proteins were detected in the pollination drops of *G. gnemon* and *W. mirabilis*. Along with our analysis of *Ephedra* spp. (Chapter 2), this study demonstrates for the first time that proteins are present in the pollination drops of all three extant gnetalean genera. The numbers of proteins and the types of proteins varied between the three genera. In addition to producing pollination drops on fertile ovules of female plants, many gnetalean species also produce pollination drops from sterile ovules on male plants. This is the first proteomic analysis of male pollination drops. Our results showed that male *Welwitschia* pollination drops contain numerous proteins and that their proteome is unlike the proteome of any other gymnosperm pollination drop to date.

The proteome of *G. gnemon* pollination drops consisted mostly of proteins annotated to the extracellular space (Table 13). Based on these annotations, these proteins

should be considered part of the pollination drop secretome. Fewer proteins were annotated to intracellular locations and should be considered degradome proteins. *Gnetum* has a shallow pollen chamber at the time of pollination drop secretion (van der Pijl 1953; Singh 1978). A relatively small proportion of nucellus degrades to form the pollen chamber. This may explain the higher proportion of secretome proteins to degradome proteins.

The number of proteins identified in *G. gnemon* pollination drops was similar to those reported for *Ephedra* spp. (Chapter 2), which had between 6 and 20 proteins depending on species. However, the *Ephedra* samples were analyzed using a different bioinformatics strategy, and we may expect more proteins to be identified if analyzed by the current method. The *G. gnemon* sample had fewer proteins than *Ginkgo*, *Zamia* and *Ceratozamia*, but more proteins than *Cycas*.

The secretome proteins found in *G. gnemon* pollination drops were similar to those found in *Ephedra*, *Ginkgo* and the cycads. A complement of carbohydrate-modifying enzymes was present, e.g. glucanase, chitinase, galactosidase and glycosyl hydrolase. *G. gnemon* contained an additional enzyme, a fucosidase, which may also be involved in modifying carbohydrate components of the cell wall (Günl et al. 2011). Proteins involved in oxidation-reduction processes were found, e.g. peroxidase and FAD-binding berberine family protein. In addition, osmotin, lipid-transfer, aspartyl protease and subtilase proteins were all identified in the *G. gnemon* pollination drop. All of these proteins have been detected in one, or more, of the other gymnosperm groups.

An extracellular cystatin protein was detected in *G. gnemon* pollination drops. Plant cystatins are protease inhibitors. They function in defence against pathogens (Popovic et

al. 2013) and insect herbivores (Irie et al. 1996). They also control the degradation activity of plant proteases, for example during storage and release of protein in seeds (Szewinska et al. 2013). Cystatins may also play a role in regulating programmed cell death, by controlling the activity of associated proteases (Solomon et al. 1999; Belenghi et al. 2003). Therefore, the following activities could be interpreted as putative functions for cystatin in *G. gnemon* ovules during pollination drop secretion: defence from pathogens; regulation of programmed cell death in the nucellus; control of protein degradation to supply amino acids for pollen tube growth.

Pollination drops collected from fertile ovules of *W. mirabilis* contained far fewer proteins than were detected in drops from sterile ovules of *W. mirabilis*. We detected six proteins in the female drop, only one of which matched to the TAIR10 database. A further BLASTp search using the NCBI nr database did not lead to any additional information on the five unidentified proteins. Furthermore, we did not detect the chitinase that Wagner et al. (2007) identified in female *W. mirabilis* drops. Two peptide sequences from the chitinase were provided by Wagner (2007). A search for these sequences in the lists of peptides and de novo peptide tags for our *W. mirabilis* female sample did not result in any matches. Similarly, we did not find the acid phosphatase detected by Carafa et al. (1992). The only protein identified in our female *W. mirabilis* pollination drop sample was HOPZ-ACTIVATED RESISTANCE 1. In *Arabidopsis*, HOPZ-ACTIVATED RESISTANCE 1 is a protein required to recognize the type III secreted effector protein (T3SE) HopZ1a from the bacteria *Pseudomonas syringae* (Van Hall, 1904) (Lewis et al. 2010). Lewis et al. (2010) deduced that HOPZ-ACTIVATED RESISTANCE 1 is required to stop infection by *P. syringae*.

Multiple factors could explain the lack of proteins found in *W. mirabilis* female drops. The total protein content of the drop could be relatively low, making proteins difficult to detect. A one dimensional SDS-PAGE gel by Wagner et al. (2007) indicated that the total protein present in *W. mirabilis* was lower than in conifers. This could explain why we did not detect chitinase (Wagner et al. 2007) or acid phosphatase (Carafa et al. 1992). It could also explain why we did not detect degradome proteins, despite the fact that *W. mirabilis* has a deep pollen chamber at the time of drop secretion. On the other hand, we know that *Ephedra* and *Gnetum* drops also show weak bands when run on SDS-PAGE gels in our lab, and multiple proteins were detected in both *Ephedra* and *G. gnemon* drops.

Another explanation could be that our custom database, Gymno_DB, lacked the sequence information for the proteins present in the *W. mirabilis* drops, and therefore the proteins were not identifiable. However, this explanation seems unlikely since numerous proteins were detected in the drop from the sterile ovule of the male plant. The few proteins detected in the female *W. mirabilis* pollination drop might reflect the actual content of the drop. Perhaps *W. mirabilis* pollination drops do lack a strong secretome, and do not have a strong complement of antimicrobial proteins. There is some evidence to support this. In their natural habitat, *W. mirabilis* ovules are commonly infected by the fungus *Aspergillus niger* (van Tieghem 1867). In a survey by Whitaker et al. (2008), 80 % of seeds were infected. Fungal spores are thought to gain access to the ovule through the pollination drop. In fact, infection of the ovules peaked during pollination drop production (Whitaker et al. 2008).

Pollination drops from the sterile ovules of *W. mirabilis* contained twenty-fold

more types of proteins than drops from fertile ovules. *W. mirabilis* “male” pollination drops lacked the types of proteins that have been attributed to the secretome of other groups. For example, *W. mirabilis* did not contain the carbohydrate modifying proteins chitinase or galactosidase, both of which were found in *Ephedra* (Chapter 2) and *G. gnemon* samples. In the category of oxidation-reduction, the only potentially extracellular protein common to all gnetaleans was peroxidase. Overall, there was little overlap between the putative secreted proteins of the *W. mirabilis* male pollination drop and the pollination drops of the fertile ovules of the other gnetalean groups. The sterile male drop appeared to lack a distinct secretome.

The proteome of *W. mirabilis* male pollination drops was dominated by proteins involved in intracellular processes, e.g. translation, protein folding, glycolysis, intracellular protein transport, etc. (Table 12). In fact, the sample was overrun with proteins attributable to the degradome, more so than the fertile ovules of the other gymnosperm groups. Perhaps the nucellus of the *W. mirabilis* sterile ovule degrades to a greater extent than in the fertile ovules of *Ephedra*, *Ginkgo* or the cycads. Unfortunately, sterile ovules of *W. mirabilis* were not sectioned at the time of drop production, and to my knowledge, no publication exists with a section of a sterile ovule. Coulter and Chamberlain (1910) published a line drawing of a cross-section through a *W. mirabilis* male cone. It only shows a small dip in the apex of the nucellus, leaving ambiguity about the existence of a pollen chamber. There was no accompanying description in the text. The sterile ovule of *G. gnemon* was investigated by Haycraft and Carmichael (2001). It was not clear from the cross section that they presented whether a pollen chamber was present. A study on the anatomy of the sterile ovule of *W. mirabilis* would be warranted

to determine the extent of nucellar degradation during pollen chamber formation.

A functional difference between the pollination drops of fertile and sterile ovules might be reflected in the organic composition of the drops. Male drops may function as nectar, attracting insects to a nutritious food reward, and ensuring that pollen is dispersed. Female drops may function in capture, transport and germination of pollen. This hypothesis has been proposed for *G. gnemon* (P.B. Tomlinson, pers. comm.). A difference in the metabolome of pollination drops of fertile and sterile ovules was observed in *G. gnemon* (P.B. Tomlinson, pers. comm.). Perhaps the same could be true for *W. mirabilis*. The behaviour of the drops on fertile and sterile cones is different. Wetschnig and Depisch (1999) observed that the pollination drop of the sterile ovule is not retracted, whereas the drop of the fertile ovule is secreted and retracted on a daily basis. This suggests that the drop of the sterile ovule functions as nectar, whereas the drop of the fertile cone functions in pollen transport. If the nucellus degrades to a greater degree in the sterile ovule, it could release more protein to create a more nutritious and attractive reward for insects.

There might also be a technical explanation for the abundance of proteins in the male *W. mirabilis* pollination drop. It is possible that pollen contaminated the pollination drop of the sterile ovule. If this were true, many of the proteins identified could then be attributed to pollen proteins, rather than drop proteins. A future experiment should include removal of pollen bearing microsporangia that surround the sterile ovule before drop production begins.

Our proteomic studies demonstrated that there is substantial variation in the proteomes of gnetalean pollination drops. The pollination drops of the fertile ovule of

G. gnemon contained many secretome proteins, but few degradome proteins. *Gnetum gnemon* shared secretome proteins in common with *Ephedra*. In contrast, the female drop of *W. mirabilis* contained only a handful of proteins, just one of which could be identified. The *W. mirabilis* male drop was also unique, dominated by its degradome. It shared few secretome proteins with the other gnetaleans, or even other gymnosperm groups. One could argue that the stark contrast in protein content could indicate functional differences between male and female pollination drops, where the male acts as nectar to attract pollinators. Further analyses would be required to test this hypothesis. It would be interesting to know whether drops from the sterile ovules of *G. gnemon* and *Ephedra* are also dominated by degradome proteins. To this end, with the help of Professor P.B. Tomlinson of Harvard University, we collected drops from the sterile ovules of *G. gnemon* sampled from the National Tropical Botanical Garden’s Kampong property in Coconut Grove, Miami, Florida in February 2014. The drops are currently being analyzed.

Chapter 7: Proteins in the pollination drops of *Taxus* x *media.*

Introduction

Taxus spp. are dioecious, with the only known exception being *Taxus canadensis* Marshall (Anderson and Owens 2000). On female trees, single ovules are produced in leaf axils. Ovules are completely exposed to the environment at pollination. As seeds mature, fleshy red arils develop from the ovulate structures (Anderson and Owens 1999). On male plants, microstrobili are produced in leaf axils. They consist of several microsporophylls, with pollen-bearing microsporangia, arranged around a central axis (Anderson and Owens 2000).

Like all conifers, *Taxus* is wind-pollinated. At pollen receptivity, a pollination drop is secreted from the ovule (Vaucher 1841; Fujii 1903; Zeigler 1959; Anderson and Owens 2000). The orientation of *Taxus* ovules is haphazard; their orientation ranges from upright to inverted. When pollen enters the pollination drop, it either sheds its exine rapidly and sinks into the ovule (Xing et al. 2000), or is carried in through the micropyle as the drop retracts (Anderson and Owens 2000). Pollination drops can be produced by an individual ovule for up to two weeks, but will retract and will not return if pollinated (Anderson and Owens 2000).

The biochemistry of *Taxus* pollination drops has been well-studied. Inorganic phosphate (Ziegler 1959) and calcium (Fujii 1903) are present in *Taxus* pollination drops. Sucrose, fructose and glucose were detected by Ziegler (1959) and Seridi-Benkaddour and Chesnoy (1988), along with malic acid, citric acid (Ziegler 1959) and galacturonic acid (Seridi-Benkaddour and Chesnoy 1988). Ziegler (1959) also detected a range of free amino acids and one 6 amino acid peptide. Using 2D SDS-PAGE gels, O'Leary et al.

(2004, 2007) demonstrated that approximately 20 proteins were present in the pollination drops of *Taxus* x *media*. O'Leary et al. (2007) identified the two most abundant proteins to be an acidic thaumatin-like protein (Txm-TLPa) and a basic thaumatin-like protein (Txm-TLPb). A beta-1,3-endoglucanase was also identified in the pollination drop (O'Leary et al. 2007). Arabinogalactan proteins appeared in the nucellus of *T.* x *media* during pollination drop production (O'Leary et al. 2004).

Taxus ovules do not form a pollen chamber during pollination drop production (Dupler 1917; Chesnoy 1993). Only the outermost cells of the nucellus lose their cytoplasm, forming a ragged edge (Dupler 1917; Chesnoy 1993). This is different from the cycads, *Ginkgo*, and the Gnetales. O'Leary et al. (2004, 2007) observed that in *T.* x *media*, the nucellus remained undifferentiated during drop production. The tissue of the ovule stayed intact. A cross-section of the ovule showed that the nucellus was not degraded during pollination drop secretion (O'Leary et al. 2004).

Taxus pollen grains hydrate quickly upon contact with the pollination drop (Xing et al. 2000) or in vitro solutions (Anderson and Owens 2000). In *Taxus canadensis* Marsh, pollen settles on the surface of the nucellus (Dupler 1917). Elongation of the pollen grain results in the rupture of the pollen exine (Dupler 1917). The pollen tube then emerges, and quickly grows into the nucellus. Pollen tubes are able to reach the megagametophyte within 10 days (Dupler 1917). The diameter of the pollen tube expands as it grows (Dupler 1917). When the pollen tube reaches the micropylar end of the megagametophyte, it spreads extensively (Anderson and Owens 1999). The pollen tube may even branch over the apex of the megagametophyte (Dupler 1917). There is

about a one month delay between pollination and fertilization (Dupler 1917; Anderson and Owens 1999).

Our sample species, *T.* x *media*, is a hybrid between *Taxus baccata* (English yew) and *Taxus cuspidata* Siebold & Zucc. (Japanese yew). It is a common horticultural plant, favoured because it is both ornamental and cold hardy. *T.* x *media* is the same species used by O'Leary et al. (2004, 2007) to study the proteome of pollination drops. O'Leary et al. (2007) only identified the three most abundant proteins in *Taxus* x *media* pollination drops. Proteomics methods have become more sensitive since O'Leary et al. (2007) published their study. Here we present the results of proteomic analysis of *T.* x *media* pollination drops. We hypothesized that the drops would contain a strong secretome because *Taxus* spp. do not have a pollen chamber at the time of drop production.

Methods

Collection of pollination drops from *Taxus* x *media*

Pollination drops were collected from *T.* x *media* shrubs growing on the campus of the University of Victoria by Brynn Porter in April 2008. Twigs of about 20 cm were cut from the hedge and placed into a plastic container lined with moistened paper towels. A lid was placed over the container to create a humid environment. Drops were collected with a 10 µL pipette tip directly from the exposed ovules. Samples were expelled into a 1.5 mL Eppendorf tube and stored at -20 ˚C until analysis.

Proteomics

Mass spectrometry and protein identification were conducted as described in Chapter 3.

Results

We detected 81 proteins in the pollination drop of *T.* x *media* (Table 14). Acceptable BLASTp matches to the TAIR 10 database were found for 63 of the proteins. Of these, most of the proteins were annotated to the extracellular space (57 proteins, 90 %). Fewer proteins were annotated to the intracellular space (21 proteins, 33 %), and even fewer were annotated specifically to the plasma membrane (6 proteins, 10 %).

Of those proteins that had good BLASTp matches to the TAIR 10 database, the best-represented GO categories for Biological Process were carbohydrate metabolic process (24 proteins, 38 %), proteolysis (6 proteins, 10 %), lipid transport (6 proteins, 10 %) and oxidation-reduction process (5 proteins, 8 %). Fifteen proteins (24 %) also had an annotation to response to salt stress.

Discussion

The proteome of *T.* x *media* pollination drops is mainly composed of secretome proteins. The majority of proteins were annotated to the extracellular space. The nucellus of *T.* x *media* remains intact at the time of pollination drop production, and no pollen chamber is formed. This explains why fewer proteins were annotated to the intracellular space.

The types of proteins identified in *T.* x *media* pollination drops were like those found in the secretome of previously studied conifers, cycads, *Ginkgo*, *Gnetum* and *Ephedra*. All of the types of proteins involved in carbohydrate-modification have been found in at least one other gymnosperm group. The same is true for those proteins involved in oxidation-reduction processes, with the addition of glucose-methanol-choline oxidoreductase family protein. Beyond a putative role in oxidation-reduction processes, little is known about the function of this protein in plants (Jiang et al. 2008). *T.* x *media*

Table 14. Proteins identified in *Taxus* x *media* pollination drops. Accession is the Gymno_DB transcript name; -10lgP is the PEAKS 6 protein score; Total Peptides is the number of peptides matched to the translated transcript by PEAKS6; Unique Peptides is the number of peptides only found to match the given transcript; TAIR 10 Gene Model is the BLASTp result from running the transcript amino acid sequence against the TAIR10 database; TAIR10 Description is the name assigned to the gene model; BLASTp e value is the e value for the BLAST result (cutoff < e-5); GO Biological Process and GO Cellular Component give all annotations for that category linked to the given gene model; Blank spaces occur where there were no significant BLASTp matches to the TAIR10 database, or where no description or Gene Ontology information was linked to the gene model.

Accession	-10lgP	Total Peptides	Unique Peptides	TAIR10 Gene Model	TAIR10 Description	BLASTP e value	Gene Ontology Biological Process	Gene Ontology Cellular Component
Taxus_bac_isotig02974_42	464.33	46	40	AT4G11650.1	osmotin 34	2e-90	response to salt stress	extracellular region
Taxus_bac_isotig12595_5	439.5	53	15	AT5G59310.1	lipid transfer protein 4	4e-16	lipid transport	extracellular region
Taxus_bac_isotig03058_45	432.59	50	12	AT5G59320.1	lipid transfer protein 3	1e-14	lipid transport	extracellular region; apoplast
Taxus_bac_isotig08792_8	414.18	28	18	AT4G16260.1	Glycosyl hydrolase superfamily protein	2e-98	response to salt stress	cell wall; vacuolar membrane
Taxus_bac_isotig13443_12	396.08	18	16	AT5G59320.1	lipid transfer protein 3	2e-15	lipid transport	extracellular region; apoplast
Taxus_bac_isotig01203_27	383.46	29	27					
Taxus_bac_isotig06303_19	368.25	17	14	AT4G11650.1	osmotin 34	1e-85	response to salt stress	extracellular region
Taxus_bac_isotig06402_16	333.29	13	8	AT2G43610.1	Chitinase family protein	5e-48	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region
scaffold-BTTS-2074994-Austrotaxus_spicata_single_19	326.4	14	8	AT2G43610.1	Chitinase family protein	9e-73	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region
scaffold-BTTS-2012738-Austrotaxus_spicata_single_20	311.33	9	4					
scaffold-BTTS-2005471-Austrotaxus_spicata_sub2_26	306.53	12	1	AT4G16260.1	Glycosyl hydrolase superfamily protein	3e-91	response to salt stress	cell wall; vacuolar membrane
scaffold-BTTS-2006150-Austrotaxus_spicata_sub2_25	293.52	15	7	AT4G16260.1	Glycosyl hydrolase superfamily protein	1e-88	response to salt stress	cell wall; vacuolar membrane

scaffold-HQOM-2001812-Torreya_nucifera_sub3_6	260.75	5	1					
scaffold-IAJW-2147927-Amentotaxus_argotaenia_single_28	260.59	8	2	AT2G43610.1	Chitinase family protein	6e-70	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region
scaffold-CGDN-2004325-Tetraclinis_sp._single_31	253.08	8	1					
scaffold-BTTS-2013220-Austrotaxus_spicata_single_29	251.89	9	6	AT3G12500.1	basic chitinase	4e-82	cell wall macromolecule catabolic process; response to cadmium ion; carbohydrate metabolic process	vacuolar membrane; cytosol; extracellular region; vacuole
scaffold-GJTI-2011543-Cephalotaxus_harringtonia_single_24	246.46	8	3	AT5G24090.1	chitinase A	1e-126	carbohydrate metabolic process	extracellular region
scaffold-EFMS-2015782-Torreya_taxifolia_single_22	244.43	7	1	AT5G24090.1	chitinase A	3e-131	carbohydrate metabolic process	extracellular region
scaffold-FHST-2009045-Taxodium_distichum_single_10	232.38	4	3	AT5G24090.1	chitinase A	2e-128	carbohydrate metabolic process	extracellular region
Sciadopitys_verti_isotig14279_7	223.45	3	2					
scaffold-BTTS-2007224-Austrotaxus_spicata_single_8	213.57	7	7	AT2G15220.1	Plant basic secretory protein (BSP) family protein	1e-77		extracellular region
scaffold-CGDN-2068122-Tetraclinis_sp._single_9	213.3	8	1	AT4G16260.1	Glycosyl hydrolase superfamily protein	3e-93	response to salt stress	cell wall; vacuolar membrane
scaffold-HQOM-2051691-Torreya_nucifera_single_3	210.2	5	1	AT2G43610.1	Chitinase family protein	6e-11	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region
scaffold-IFLI-2004730-Callitris_gracilis-March_sub2_35	198.36	5	5	AT5G08370.1	alpha-galactosidase 2	4e-131	carbohydrate metabolic process	extracellular region
scaffold-FHST-2064047-Taxodium_distichum_single_19	196.78	4	1					
scaffold-BTTS-2012067-Austrotaxus_spicata_single_10	193.35	6	1	AT4G11650.1	osmotin 34	2e-88	response to salt stress	extracellular region
Taxus_bac_isotig10446_3	186.51	4	3	AT5G59310.1	lipid transfer protein 4	1e-16	lipid transport	extracellular region

scaffold-BTTS-2002927-Austrotaxus_spicata_single_17	179.33	5	3	AT5G59320.1	lipid transfer protein 3	1e-14	lipid transport	extracellular region; apoplast
scaffold-BTTS-2078726-Austrotaxus_spicata_single_19	176.71	2	2	AT2G46750.1	D-arabinono-1,4-lactone oxidase family protein	2e-130		membrane; extracellular region
Metasequoia_Contig8430_8	174.49	3	1					
scaffold-ACWS-2059033-Arucaria_sp._single_4	168.86	3	1	AT2G43610.1	Chitinase family protein	2e-44	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region
Taxus_bac_isotig07804_30	166.99	4	4	AT1G12570.1	Glucose-methanol-choline (GMC) oxidoreductase family protein	4e-96	alcohol metabolic process; oxidation-reduction process	extracellular region
Sciadopitys_verti_isotig03206_13	165.43	4	3					
scaffold-IAJW-2003925-Amentotaxus_argotaenia_sub1_14	161.71	3	1					
scaffold-EFMS-2078668-Torreya_taxifolia_single_25	161.56	4	2	AT2G43610.1	Chitinase family protein	3e-61	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region
scaffold-GAMH-2007447-Tsuga_heterophylla_single_8	153.22	3	1	AT4G16260.1	Glycosyl hydrolase superfamily protein	5e-102	response to salt stress	cell wall; vacuolar membrane
scaffold-HBGV-2001436-Sequoia_sempervirens_single_19	150.02	3	2	AT4G11650.1	osmotin 34	9e-91	response to salt stress	extracellular region
scaffold-GJTI-2009856-Cephalotaxus_harringtonia_single_32	149.18	4	4	AT5G08370.1	alpha-galactosidase 2	2e-156	carbohydrate metabolic process	extracellular region
scaffold-BTTS-2003862-Austrotaxus_spicata_single_5	148.32	3	3	AT3G54420.1	homolog of carrot EP3-3 chitinase	2e-27	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region; cell wall
Ginkgo_Contig76263_98	146.98	3	1	AT3G57270.1	beta-1,3-glucanase 1	2e-81	carbohydrate metabolic process	extracellular region
scaffold-AIGO-2008228-Chamaecyparis_lawsoniana_single_7	146.92	3	1	AT2G43610.1	Chitinase family protein	2e-84	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region

scaffold-BTTS-2008340-Austrotaxus_spicata_single_10	142.54	3	1	AT4G11650.1	osmotin 34	3e-80	response to salt stress	extracellular region
Taxus_bac_isotig02719_46	141.2	3	2	AT2G16230.1	O-Glycosyl hydrolases family 17 protein	2e-169	carbohydrate metabolic process	plasma membrane
scaffold-CGDN-2009822-Tetraclinis_sp._single_9	138.16	3	1	AT2G43610.1	Chitinase family protein	8e-94	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region
scaffold-JRNA-2063739-Phyllocladus_hypohyllus_single_7	133.45	3	3	AT1G03220.1	Eukaryotic aspartyl protease family protein	3e-41	response to salt stress; proteolysis	membrane; cell wall; extracellular region; plasmodesma; Golgi apparatus; plant-type cell wall
scaffold-BTTS-2004894-Austrotaxus_spicata_single_12	122.83	2	2	AT3G54400.1	Eukaryotic aspartyl protease family protein	8e-82	proteolysis	apoplast; extracellular region; cell wall; plant-type cell wall; chloroplast
scaffold-HOUF-2110787-Cryptomeria_japonica-leaf_single_23	114.93	1	1	AT1G26560.1	beta glucosidase 40	0	carbohydrate metabolic process	apoplast; chloroplast; extracellular region
scaffold-ETCJ-2000669-Pilgerodendron_uviferum_single_6	114.36	2	1	AT1G19320.1	Pathogenesis-related thaumatin superfamily protein	2e-35		extracellular region
scaffold-GKCZ-2009296-Diselma_archeri_single_22	111.07	2	2					
Taxus_bac_isotig02434_26	110.54	2	2	AT5G24090.1	chitinase A	6e-121	carbohydrate metabolic process	extracellular region
Taxus_bac_isotig10309_39	104.86	2	2	AT1G25510.1	Eukaryotic aspartyl protease family protein	1e-164	proteolysis	extracellular region
Pinus_lamb_isotig05096_12	104.04	1	1	AT4G35880.1	Eukaryotic aspartyl protease family protein	9e-72	proteolysis	extracellular region
scaffold-EFMS-2016181-Torreya_taxifolia_single_18	101.08	1	1	AT5G05390.1	laccase 12	0	lignin catabolic process; oxidation-reduction process	extracellular region; apoplast
scaffold-FHST-2001656-Taxodium_distichum_sub2_15	96.63	2	1	AT2G16230.1	O-Glycosyl hydrolases family 17 protein	3e-175	carbohydrate metabolic process	plasma membrane
scaffold-FMWZ-2045597-Dacrycarpus_compactus_single_8	95.9	2	2					
scaffold-BUWV-2002713-Platycladus_orientalis_single_25	90.28	1	1	AT5G08370.1	alpha-galactosidase 2	0	carbohydrate metabolic process	extracellular region

scaffold-FMWZ-2007637-Dacrycarpus_compactus_sub2_44	87.06	1	1	AT2G46750.1	D-arabinono-1,4-lactone oxidase family protein	5e-136		membrane; extracellular region
scaffold-AIGO-2004099-Chamaecyparis_lawsoniana_single_55	85.4	1	1	AT1G61180.2	LRR and NB-ARC domains-containing disease resistance protein	1e-44	defense response	plasma membrane
scaffold-AUDE-2053773-Widdringtonia_cedarbergensis_single_4	85.03	1	1	AT5G59310.1	lipid transfer protein 4	5e-15	lipid transport	extracellular region
Taxus_bac_isotig03174_14	83.86	1	1	AT2G30870.1	glutathione S-transferase PHI 10	2e-15	response to cadmium ion	plasma membrane; chloroplast; vacuole; apoplast; cytoplasm; cytosol; cell wall
scaffold-ESYX-2009068-Cunninghamia_lanceolata-branch_apex_with_needles_sub1_15	83.74	1	1	AT5G05340.1	Peroxidase superfamily protein	5e-77	oxidation-reduction process; response to oxidative stress	extracellular region; apoplast; Golgi apparatus; cytosol; cell wall
scaffold-IZGN-2005986-Dacrydium_balansae_sub2_9	83.04	1	1	AT1G07890.3	ascorbate peroxidase 1	3e-146	oxidation-reduction process; response to oxidative stress; response to salt stress; response to cadmium ion	cytosol; plasma membrane; Golgi apparatus; cytoplasm; chloroplast; plasmodesma; cell wall
Metasequoia_Contig19227_9	82.1	1	1					
Taxus_bac_isotig09284_57	81.69	1	1	AT1G28110.1	serine carboxypeptidase-like 45	0	proteolysis	plasmodesma; extracellular region
scaffold-JRNA-2010446-Phyllocladus_hypohyllus_sub1_10	79	1	1					
Pinus_taeda_isotig43905_56	78.96	1	1	AT3G07320.1	O-Glycosyl hydrolases family 17 protein	4e-177	carbohydrate metabolic process	plant-type cell wall; plasmodesma; plasma membrane
Dougfir-megastigmus-comp578137_c2_seq16_112	75.27	1	1	AT1G55860.1	ubiquitin-protein ligase 1	0	cellular protein modification process	plasmodesma; intracellular; nucleus; membrane
scaffold-IAJW-2023654-Amentotaxus_argotaenia_single_31	74.55	1	1	AT1G03220.1	Eukaryotic aspartyl protease family protein	5e-99	response to salt stress; proteolysis	membrane; cell wall; extracellular region; plasmodesma; Golgi apparatus; plant-type cell wall
scaffold-JRNA-2002487-Phyllocladus_hypohyllus_single_27	73.45	1	1	AT2G15220.1	Plant basic secretory protein (BSP) family protein	4e-78		extracellular region

scaffold-HOUF-2011572-Cryptomeria_japonica-leaf_single_26	72.63	1	1	AT2G43610.1	Chitinase family protein	2e-84	carbohydrate metabolic process; cell wall macromolecule catabolic process	extracellular region
scaffold-JDQB-2010505-Neocallitropsis_pancheri_single_41	72.24	1	1	AT5G08370.1	alpha-galactosidase 2	6e-136	carbohydrate metabolic process	extracellular region
scaffold-GJTI-2048118-Cephalotaxus_harringtonia_single_5	68.37	1	1					
Podocarpus_macro_isotig11391_12	67.86	1	1	AT1G71695.1	Peroxidase superfamily protein	8e-107	oxidation-reduction process; response to oxidative stress	extracellular region; vacuole; plasmodesma; cell wall; membrane
scaffold-HQOM-2133998-Torreya_nucifera_single_4	61.91	1	1	AT3G04720.1	pathogenesis-related 4	3e-61	response to salt stress; defense response to fungus; cell wall macromolecule catabolic process; defense response to bacterium	extracellular region
scaffold-IIOL-2013061-Pinus_parviflora_single_16	61.56	1	1	AT4G11650.1	osmotin 34	6e-76	response to salt stress	extracellular region
scaffold-IOVS-2056627-Pseudotsuga_menziesii_single_18	60.63	1	1					
scaffold-GTHK-2011493-Gnetum_montanum_sub1_11	59.15	1	1					
scaffold-GNQG-2014602-Encephalartos_barteri_single_31	58.96	1	1	AT1G67930.1	Golgi transport complex protein-related	0		chloroplast; cytosol
Metasequoia_Contig18413_10	58.86	1	1					
scaffold-ESYX-2015051-Cunninghamia_lanceolata-branch_apex_with_needles_sub2_84	57.57	1	1	AT2G46020.2	transcription regulatory protein SNF2, putative	0		nucleus
scaffold-DSXO-2075894-Cryptomeria_japonica_single_34	56.74	1	1					

pollination drops contained lipid transport proteins, proteolytic enzymes, osmotins, and the D-arabinono-1,4-lactone oxidase family protein, all of which have been identified in other gymnosperm pollination drops.

O'Leary (2004) detected two thaumatin-like proteins, one acidic and one basic, as well as an endo-glucanase in the pollination drops of *T.* x *media*. We identified multiple thaumatin-like proteins (osmotins) and one beta-1,3-glucanase in our *T.* x *media* sample. A BLASTp search revealed that the highest scoring protein in our sample showed high homology to the acidic thaumatin-like protein identified by O'Leary (2004). O'Leary et al. (2004, 2007) hypothesized that a critical function of drop proteins is to defend pollination drops from wind-borne pathogens, such as fungi and bacteria. Pollination drops in yew are openly exposed to the environment, and they contain nutrients, such as sugars and amino acids, that should create a suitable growing medium for pathogens. O'Leary et al. (2007) pointed out that little pathogenic contamination had been observed within the ovules of numerous samples that they had taken.

The results of our proteomic analysis add support to the defence hypothesis presented by O'Leary et al. (2004, 2007). Many of the proteins in *T.* x *media* pollination drops may function as defence enzymes. Chitinases, glucanases, peroxidases, thaumatin-like proteins (osmotin) and lipid-transfer proteins have all been found to exhibit antimicrobial activities. Additional putative defence proteins were also found in T x *media* pollination drops. Plant basic secretory proteins are thought to function in defence, although they are not well-characterized (Kuwabara et al. 1999; Herrmann et al. 2012). PR-4 proteins are induced by both pathogens and other elicitors such as ethylene and O_3 (Guevara-Morato et al. 2010). They have been found in the extrafloral nectar of

Acacia, where a putative anti-fungal role has been proposed (Gonzalez-Teuber et al. 2009). Glutathione S-transferases are a large family of proteins that respond to both biotic and abiotic stressors. They protect plant cells from pathogens, toxins and oxidative stress (Wagner et al. 2002; Soranzo et al. 2004). One of the proteins detected, LRR and NB-ARC domains-containing disease resistance protein, may be involved in detecting plant pathogens at the interface of the apoplast and plasma membrane. Proteins containing LRR and NB-ARC domains are part of a group of resistance family (R) proteins. They contribute to the innate immune system of plants (Jupe et al. 2013) and function in the recognition of plant pathogens (van Ooijen et al. 2008).

Proteomics methods have become more sensitive since O'Leary et al. (2004, 2007) published their proteomic analysis of *T.* x *media* pollination drops. O'Leary et al. (2004, 2007) used 2D SDS-PAGE gels combined with mass spectrometry to identify proteins. We were able to perform a direct trypsin digest on the pollination drop sample, and to run this sample through the LTQ Orbitrap Velos (Thermo Fisher Scientific), a highly sensitive instrument. Our analysis detected ~80 proteins. This means that we detected 80 proteins that had at least one unique peptide, and that matched to different open reading frames in our Gymno_DB database. We detected about four times the number of proteins predicted by O'Leary et al. (2007). However, in our sample, multiple proteins matched to the same, or similar, TAIR 10 gene models or protein types. For example, 15 proteins were annotated as chitinases and 6 proteins were annotated as osmotin 34. It was a requirement of our filter settings in PEAKS 6 for each identified protein to contain at least one unique peptide. Each protein should therefore contain at least one peptide which differs by at least one amino acid compared to all others in the

list. However, confirmation of the number of isoforms of a given type of protein would require further investigation. O'Leary et al. (2007) suggested that the presence of multiple TLPs with slightly different substrate specificities may confer increased resistance to fungal pathogens within the pollination drop. This could also be true for the apparent multiplicity of protein types found here in the *T.* x *media* pollination drop proteome. Perhaps there are multiple forms of protein types within the pollination drop that interact with slightly different substrates.

O'Leary et al. (2004) reported that an abundance of arabinogalactan proteins (AGPs) were present in *T*. x *media* pollination drops. Our proteomic analysis did not detect an abundance of AGPs. Only one type protein identified in our *T.* x *media* sample, the lipid transfer proteins, could be considered as potential AGPs (Seifert and Roberts 2007). There may be a technical explanation for the lack of AGPs detected. Studies aiming to identify the core protein of AGPs usually follow a specific protocol to remove the attached carbohydrate moieties, e.g. a deglycosylation step (Gleeson et al. 1989; Showalter 2001). We did not tailor our proteomics methods specifically to detect AGPs.

A small number of proteins had GO annotations only to the intracellular space: transcription regulatory protein SNF2; Golgi transport complex protein-related; ubiquitin-protein ligase 1. Although the nucellus of *T.* x *media* does not degrade to form a pollen chamber, it is likely that these proteins have escaped from intracellular compartments. Dupler (1917) and Chesnoy (1993) observed that in *Taxus* spp., the cells of the outermost layer of the nucellus degraded during drop production. If a similar situation occurs in *T.* x *media*, it could explain the presence of these proteins. I could not find any examples of putative extracellular functions for these proteins in the available literature.

The pollination drop proteome of *T.* x *media* is composed mainly of secretome proteins. The majority of proteins present in *T*. x *media* pollination drops have been detected previously in other gymnosperm pollination drops. Many of the protein types detected in *T*. x *media* pollination drops are known to have defensive functions in other plant systems. Our proteomic analysis further supports the hypothesis proposed by O'Leary et al. (2007), who suggested defence from wind-borne microbes was a primary function of pollination drop proteins in *T.* x *media*.

Chapter 8: Perspectives

Overview of pollination drop proteins in gymnosperm lineages

Proteomic analyses revealed that proteins are present in the pollination drops of all gymnosperm taxa sampled: *Ephedra* spp., *Gnetum gnemon*, *Welwitschia mirabilis*, cycads, conifers and *Ginkgo biloba*. These taxa represent all living gymnosperm lineages. Our results, therefore, support the hypothesis that the pollination drops of all extant gymnosperm lineages are complex reproductive secretions. Proteins involved in carbohydrate modification were detected in all groups (Table 15). The ubiquity of these proteins supports the hypothesis put forth by O'Leary (2004) and Poulis (2004). They speculated that carbohydrate-modifying proteins contained in pollination drops influence pollen germination, pollen tube growth and pollen tube penetration of the nucellus by interacting with components of the pollen and nucellar cell walls. Along with proteolytic enzymes, they may support pollen tube growth by liberating nutrients for uptake by the growing pollen tube. From our analyses, it seems that the presence of carbohydrate modifying proteins may be a conserved feature of pollination drops. Our results also support the hypothesis, presented by Wagner (2007), that defence from invading pathogens is likely a conserved function of pollination drop proteins amongst the gymnosperm lineages. Potential defence proteins, e.g. reduction-oxidation proteins, lipid-transfer proteins and thaumatin-like proteins, amongst others, were found in most groups (Table 15).

There was a trend connecting the degree of nucellar degradation at the time of drop production, and the types of proteins detected in the pollination drop. *Taxus* x *media*, which does not form a pollen chamber, had fewer proteins that are typically

Table 15. Overview of proteins detected in gymnosperm pollination drops.

	Cycad spp.	*Ginkgo biloba*	*Ephedra* spp.	*Gnetum gnemon*	*Welwitschia mirabilis* female	*Taxus* x *media*	Conifers (previous)
Carbohydrate metabolic process:							
galactosidases	x	x	x	x		x	x[b]
chitinases	x		x	x	x[c]	x	x[bd]
glycosyl hydrolases	x	x	x	x		x	x[d]
beta glucosidase	x	x	x			x	x[d]
glucanases/cellulases	x	x		x		x	x[d]
xylosidases	x	x	x				x[b]
beta-hexosaminidases	x	x					
invertases	x						x[b]
pectin lyase-like superfamily proteins	x						
fucosidases				x			
xyloglucan endotransglycosylases							x[c]
Oxidation-reduction process/Cell redox homeostasis:							
peroxidases	x		x	x		x	x[b]
FAD-binding Berberine family proteins	x			x			
D-arabinono-1,4-lactone oxidase family proteins	x					x	
laccases	x					x	
glutaredoxin family proteins	x						
SKU5 similar proteins	x						
thioredoxin H-type 1 proteins	x						
glucose-methanol-choline (GMC) oxidoreductases						x	
Proteolysis:							
aspartyl proteases	x	x	x	x		x	x[b]
serine carboxypeptidases	x		x			x	x[b]
cysteine proteinases	x	x	x				
subtilisin-like serine endopeptidases	x			x			x[d]
aleurain-like proteases	x		x				
alpha/beta-Hydrolases superfamily proteins	x						

Defense response:							
MLP-like proteins	X						
HOPZ-ACTIVATED RESISTANCE 1 proteins					X		
LRR and NB-ARC domains containing disease resistance proteins					X		
PR-4 proteins						X	
Translation/Translational elongation:							
Elongation Factor proteins	X	X	X				
RAB GTPase homolog E1B proteins		X					
Ribosomal protein S11 family proteins		X					
Ribosomal protein S3 family proteins		X					
Glycolysis:							
aldolase superfamily proteins		X					
cytosolic enolases		X					
phosphoglycerate kinase 1 proteins		X					
pyruvate kinase family proteins		X					
ATP metabolic process:							
ATP synthase alpha/beta family proteins		X					
ATP synthase subunit 1 proteins		X					
Protein phosphorylation:							
receptor-like protein kinase 2	X						
leucine-rich repeat transmembrane protein kinase	X						
Plant-type cell wall organization:							
expansins		X					
Lipid transport:							
lipid transfer proteins	X	X		X		X	X[c]

Various extracellular function:							
thaumatin-like proteins / osmotin	X	X	X	X		X	X[bcd]
FASCICLIN-like arabinogalactan protein 8	X	X					
acid phosphatase	X				X[a]		
glutathione S-transferase		X				X	
GAST1 protein homolog 3	X						
gibberellin-regulated family protein	X						
glucuronidase 1	X						
leucine-rich repeat (LRR) family protein	X						
malate dehydrogenase			X				
superoxide dismutase			X				
plant basic secretory protein						X	
Various intracellular function:							
ubiquitin	X	X	X				
amylase		X	X				
calreticulin		X	X				
heat shock proteins		X	X				
Class I glutamine amidotransferase-like superfamily protein	X						
hydroxyproline-rich glycoprotein family protein	X						
immunoglobulin E-set superfamily protein	X						
phosphoenolpyruvate carboxylase 4	X						
P-loop containing nucleoside triphosphate hydrolases superfamily protein	X						
rotamase CYP 3	X						
serine transhydroxymethyltransferase 1	X						
SHV3-like 1	X						
SIT4 phosphatase-associated family protein	X						
ssDNA-binding transcriptional regulator	X						
uclacyanin 1	X						
CAP (Cysteine-rich secretory proteins, Antigen 5, and Pathogenesis-related 1 protein) superfamily protein		X					
DCD (Development and Cell Death) domain protein		X					

embryonic cell protein 63	X				
glyoxal oxidase-related protein	X				
GMP synthase (glutamine-hydrolyzing), putative / glutamine amidotransferase, putative	X				
late embryogenesis abundant protein (LEA) family protein	X				
non-specific phospholipase C2	X				
RAN GTPase 3	X				
RNA-binding (RRM/RBD/RNP motifs) family protein	X				
RNA-binding KH domain-containing protein	X				
S-adenosylmethionine synthetase family protein	X				
acyl-CoA-binding domain-containing protein 6		X			
alpha-Amylase inhibitor		X			
alpha-Gliadin		X			
auxin response factor		X			
calmodulin		X			
ceramidase		X			
citrate synthase		X			
cyclophilin A		X			
dessication-related protein		X			
glycerophosphoryl diester phosphodiesterase		X			
granule-bound starch synthase		X			
GTP-binding nuclear proteins		X			
histones		X			
lactoylglutathione lyase		X			
luminal-binding protein		X			
profilin		X			
cystatin			X		
golgi transport complex protein-related					X
transcription regulatory protein SNF2					X
ubiquitin-protein ligase					X

[a] Carafa et al. 1992; [b] Poulis et al. 2005; [c] O'Leary 2004; [d] Wagner et al. 2007

regarded as intracellular than extracellular. Similarly, *G. gnemon*, which has a shallow pollen chamber, had fewer proteins annotated to the intracellular space than to the extracellular space. Furthermore, *G. biloba*, cycads and *Ephedra*, which have deeper pollen chambers, had greater proportions of degradome proteins. *Welwitschia mirabilis* was an exception. Although known to have a deep pollen chamber during drop production, it had few detectable or identifiable proteins present.

We did not observe any obvious trends between protein types and pollination syndromes, e.g. wind- vs. insect- pollination, or pollen behaviour, e.g. zooidogamy vs. siphonogamy. However, recent work by Dr. Massimo Nepi and colleagues at the Università di Siena, Italy (pers. comm.) has revealed that the non-protein amino acid content of pollination drops may differ between insect- and wind- pollinated lineages. Preliminary results show that conifers, which are wind-pollinated, do not contain λ-amino butyric acid (GABA), but the insect-pollinated Gnetales do. In plants, GABA may mediate defence or stress responses. It may also mediate interactions with animal pollinators. GABA affects the physiology of insect muscles and nervous systems, and has phagostimulatory effects (reviewed by Nepi 2014). Dr. Massimo Nepi and his students are now conducting experiments to determine whether GABA and other non-protein amino acids affect insect-pollinator behaviour. They are also keen to determine whether levels of non-protein amino acids correlate with carbohydrate composition and concentration in pollination drops, as carbohydrates might also influence insect-pollinator behaviour.

Whether the types of proteins detected in gymnosperm pollination drops reflect an evolutionary pattern was unclear. This raises several questions: Do the proportions of

degradome versus secretome proteins in different genera tell an evolutionary story? Traditionally, the cycads and *Ginkgo* are considered as ‘ancient’ lineages (Chamberlain 1935). The pollination drop proteomes of *Ginkgo* and the cycads had substantial proportions of degradome proteins. Do greater proportions of degradome proteins represent an ancestral reproductive mechanism? *Ephedra* also contained substantial proportions of degradome proteins. *Gnetum*, which is also part of the Gnetales, contained a lower proportion of degradome proteins than *Ephedra*. We detected only one degradome protein in *Welwitschia* female pollination drops – a surprising result given that *Welwitschia* is known to have a deep pollen chamber. Does the variation in composition of gnetalean pollination drop proteomes reflect shifts in their reproductive mechanisms? Does this mean that the proteomes of *Gnetum* and *Welwitschia* are more derived than *Ephedra*? To date, no substantial degradome has been detected in any conifer species (Poulis 2005; O’Leary et al. 2007; Wagner et al. 2007). We detected fewer intracellular than extracellular proteins in *Taxus* (Chapter 7). Do the pollination drop proteomes of conifers also reflect a more derived reproductive mechanism? None of the conifers analyzed to date form pollen chambers, which likely explains the absence of degradome proteins. However, some conifers do produce pollen chambers, e.g. *Cephalotaxus*. Our lab is currently analyzing *Cephalotaxus* pollination drops in collaboration with Dr. Cary Pirone of Harvard University. A consensus of the phylogeny of seed-plant groups may come with the completed analyses of large sequencing projects, such as the OneKP Project. Once available, mapping the reproductive characteristics of gymnosperms along with their pollination drop proteomes onto the phylogeny would more concretely reveal evolutionary patterns.

Complex reproductive secretions also occur in the fifth group of extant seed plants, the angiosperms. These include placental fluids (Willemse and Vletter 1995), stylar exudates (Park et al. 2000), stigmatic exudates (Rejon et al. 2013) and nectars (Nepi et al. 2009; Seo et al. 2013). Unlike pollination drops, these liquids are not secreted by the nucellus. In angiosperms, the nucellus is typically only a few cells thick (Nepi et al. 2009). Even though these secretions do not have the same tissue origin as pollination drops, they do share similar functions to those proposed for pollination drops. Stigmatic exudate is involved in capture, adhesion, hydration, nutrition and chemotropism of pollen and pollen tubes, as well as defence from pathogens (Rejon et al. 2013). Stylar fluid supports and directs pollen tube growth (Park et al. 2000). Floral nectar is also considered a reproductive secretion, but it is not secreted onto the reproductive tract of the pistil. Nectar is secreted by specialized organs called nectaries. Its main function is to attract insect-pollinators. However, components within the nectar may function in pathogen defence (Thornburg et al. 2003).

Proteins are found in angiosperm reproductive secretions (Willemse and Vletter 1995; Park et al. 2000; Sang et al. 2012; Chalivendra et al. 2013; Rejon et al. 2013; Seo et al. 2103). There is an overlap between the types of proteins found in angiosperm reproductive secretions and the types of proteins that we detected in gymnosperm pollination drops. Carbohydrate modifying enzymes were found in recent proteomic analyses of floral nectar of *Nicotiana attenuata* Torr. ex S. Watson (Seo et al. 2013) and stigmatic exudates of *Nicotiana tabacum* L. (Sang et al. 2012), *Solanum pennellii* Correll (Chalivendra et al. 2013), *Lilium longiflorum* and *Olea europapea* L. (Rejon et al. 2013). These included galactosidases (Sang et al. 2012; Rejon et al. 2013; Seo et al. 2013),

glucosidases (Sang et al. 2012; Chalivendra et al. 2013; Rejon et al. 2013; Seo et al. 2013), glucanases (Chalivendra et al. 2013; Rejon et al. 2013) and xylosidases (Rejon et al. 2013; Seo et al. 2013). Rejon et al. (2013) suggested that in stigmatic exudates these types of proteins regulate pollen tube growth by degrading cell wall polysaccharides and by breaking down carbohydrate components for uptake into pollen tubes. These are similar putative functions to those suggested for the carbohydrate modifying proteins in pollination drops. Nectar also contains carbohydrate modifying enzymes, but unlike pollination drops and stigmatic exudates, nectar does not have direct contact with pollen. Alternative functions have been suggested for some of these proteins. Nepi (2011) speculated that by breaking down oligosaccharides released from the cell wall by pathogens, the xylosidase found in the nectar of *Curcurbita pepo* L. may play a role in the reduction of pathogen virulence. Released oligosaccharides would otherwise elicit increased damage by the pathogen. Zha et al. (2012) reported that alpha-galactosidase was the most abundant protein in tobacco nectar. Tobacco ovules are immersed in nectar during their early development. Zha et al. (2012) suggested that alpha-galactosidase affects cell wall restructuring of ovules during this period of development.

Like pollination drops, stigmatic exudates and nectar are exposed to the environment, and to wind- or insect- borne pathogens. Both stigmatic exudates and floral nectar have proteins usually associated with plant defence. Some of the putative defence proteins found in these secretions are similar to those found in pollination drops, e.g. chitinases (Sang et al. 2012; Chalivendra et al. 2013; Rejon et al. 2013; Seo et al 2013), lipid transfer proteins (Chalivendra et al. 2013; Rejon et al. 2013) and thaumatin-like proteins (Sang et al. 2012; Chalivendra et al. 2013; Rejon et al. 2013; Seo et al. 2013).

Additional putative defence proteins have been described in angiosperm reproductive secretions that we did not detect in pollination drops, e.g. RNases (Hillwig et al. 2010), alliin lyase (Peumans et al. 1997), mannose-binding lectin (Peumans et al. 1997), fructokinase (Hillwig et al. 2011). The best studied nectar defence proteins are the nectarins of tobacco, which are part of the Carter-Thornburg nectar redox cycle (Carter and Thornburg 2004). These proteins work together to elevate the level of hydrogen peroxide in nectar to a level that is toxic to microbes.

Some flowering plants also produce nectar in extrafloral nectaries. The purpose of extrafloral nectar is to protect the plant from insect herbivores by attracting insect carnivores (Gonzalez-Teuber et al. 2009). Protein composition of extrafloral nectars bears some similarity to that of pollination drops. Carbohydrate modifying enzymes, e.g. invertases (Gonzalez-Teuber et al. 2009), glucosidases (Gonzalez-Teuber et al. 2010), glucanases (Gonzalez-Teuber et al. 2010; Escalante et al. 2012), and defence proteins, e.g. chitinases (Gonzalez-Teuber et al. 2009, 2010; Escalante et al. 2012), glycoside hydrolases (Gonzalez-Teuber et al. 2009, 2010), thaumatin-like proteins (Gonzalez-Teuber et al. 2009; Escalante et al. 2012), PR proteins (Gonzalez-Teuber et al. 2009; Escalante et al. 2012), and peroxidases (Gonzalez-Teuber et al. 2009), have all been detected in extrafloral nectars. In fact, the pollination drop proteome is similar to other non-reproductive plant secretions. The apoplastic proteomes of *Arabidopsis* (Albenne et al. 2013), *Vitis vinifera* L. (Delaunois et al. 2013) and *Solanum lycopersicum* L. (Konozy et al. 2013) contain similar types of carbohydrate modifying proteins and defence proteins as those found in pollination drops.

One aspect that makes gymnosperm pollination drop proteomes different from the proteomes of all other plant apoplastic fluids is their source – the nucellus. The nucellus is a specialized reproductive tissue that almost certainly has a unique set of expressed genes differentiating it from other tissues involved in secretion to the apoplast, e.g. leaf, root, stigma, style. The pollination drop is the environment in which most gymnosperm pollen germinates and pollen tubes initially grow. Experimental evidence suggests that pollination drops somehow mediate and control pollen germination and pollen tube growth (von Aderkas et al. 2012; Williams 2012). It is logical to conclude that pollination drop proteins could be involved in these reproductive processes. Therefore, even though there is overlap in the types of proteins present in plant apoplastic fluids and pollination drops, the functions of these proteins need to be considered within the context of the tissue that secretes them. It is my opinion that the similarities between the proteomes of pollination drops and those of angiosperm reproductive secretions and plant apoplastic fluids highlights the need for further experimental exploration of the functions of pollination drop proteins. Only functional analysis of pollination drop proteins will be able to determine their unique effects on gymnosperm pollen-ovule interactions.

Using a boutique database for pollination drop proteomics

Creating a custom database from transcriptome data improved the number and quality of protein identifications in our proteomic analysis. Trial searches showed that even the use of a smaller boutique database, which included 16 species, provided these benefits as compared to the UniProt-SwissProt or UniProt-TrEMBL databases (data not presented). By comparing the results from searches using the UniProt databases versus this small gymnosperm database, we determined that we were not losing any protein

identifications, only gaining them. Our results were improved even though we did not have exact species matches for all of our samples.

There is one caveat to our approach that should be acknowledged. Some of the available transcriptome datasets were derived from non-ovule tissues, such as leaves, shoots and branch apices. Any proteins whose genes were expressed in the nucellus but not in the tissues used to create the transcriptome datasets may have been missed in the proteomic analyses. The ultimate custom database would be created from the same species and tissues, or better yet, the same individual and tissues, as the protein sample. We have recently created a Douglas-fir nucellus database. Preliminary results show a dramatic increase in the number of peptide spectrum matches between Douglas-fir pollination drops and this custom Douglas-fir nucellar database as compared to the Uniprot databases. Custom databases based on transcriptome data will likely become commonplace as the cost of acquiring transcriptome data continues to decrease.

A common rule of thumb in proteomics analyses is to require at least two unique peptides per protein identification. In this study, I chose to include all protein identifications with at least one unique peptide on the condition that they met stringent false discovery rate criteria. Each protein was required to have false discovery rates of ≤ 1 % at the levels of peptide spectrum matches, peptides and proteins. These false discovery rate filters were recommended by the manufacturers of the software program PEAKS 6. I chose to include these proteins for two reasons. First, it is still debated in the literature whether the two peptide minimum rule is, in fact, a good rule. Some researchers have determined that many true protein identifications are lost due to this rule (Nesvizhskii 2010). Second, one goal of this study was to provide sequence-level

information about pollen-ovule interactions in gymnosperms. This is a neglected area of plant research. I intend to upload all proteomic data from this project to a public data repository. By searching my complete protein dataset, a researcher may find evidence that their protein of interest exists in the reproductive tracts of these key seed-plant lineages. As databases and annotations improve, other researchers can retrieve more useful information from our sets of LC-MS/MS and proteomic data.

Challenges of pollination drop collection

Three factors affect successful collection: size of individual drops; availability and accessibility of cones; timing of drop production. Some gymnosperms lend themselves better to collection than others. For example, *Ephedra monosperma* can be easily propagated in a glasshouse, and can be induced to produce cones by applying a 12-hour daylight cycle (Dr. Stefan Little, pers. comm.). The pollination drops of this species are relatively large, measuring ~1 µL each. On the other hand, the cycads can be a difficult group to collect from. In the Zamiaceae, ovules face the interior of the cone. The only way to access pollination drops is by destructive sampling. Pollination drop production is not synchronous in all ovules of the cone, and the pollination drops can be small, e.g. ~ 0.02 µL in *Z. furfuracea*. After dissecting dozens of receptive cones and pooling our collections, our samples did not exceed perhaps 15 µL in total volume. The other gymnosperm groups have their own sampling peculiarities. *Ginkgo biloba* is a common boulevard tree in urban areas. However, few female plants are grown because of the putrid odour of their fleshy ovules. *Ginkgo biloba* matures into a large tree. Accessing ovules requires either pruning for research purposes (Arnold Arboretum, Dr. Cary Pirone, pers. comm.) or rooftop accessibility (University of California at Davis, Dr. Stefan Little,

pers. comm.). Our samples of *G. gnemon* and *W. mirabilis* were both collected in glasshouse facilities. Timing and rarity of these plants appear to be the greatest challenges in collecting drops from these species. The ease of collecting samples should be considered when planning future pollination drop studies.

Future studies on pollination drops and their proteins

Discerning communication between pollen and ovule

Little is understood about the molecular mechanisms of pollen-ovule interactions in gymnosperms. A number of phenomena suggest that chemical signals exist between pollen and the tissues of the ovule. Pollination drops are secreted and retracted, and at some point, pollination drops cease to be produced. In *Z. furfuracea*, I observed that pollination drop production is not synchronous within the cone, or even between the two ovules on one sporophyll. Drops are secreted and retracted, and they may or may not reappear after they have been removed. In some taxa, the ovule is able to recognize that it has been pollinated. I often observed in *Z. furfuracea* that of two ovules on one sporophyll, one may be producing drops, but the other had presumably been pollinated and was no longer producing drops. The post-pollination ovule was noticeably enlarged. Pollination is also a requirement for further development of the ovule in other gymnosperm taxa (Nakao et al. 1998; Owens et al. 2005; Ortiz et al. 1998).

There is no mechanistic explanation for the control of secretion, retraction or cessation of pollination drops. Experimental observations have been made regarding the effect of pollination on drop secretion and retraction (Jin et al. 2012; Mugnaini et al. 2007). Mugnaini et al. (2007) found that the degree of drop retraction was dependent upon the type of particles that entered the drop. Viable conspecific pollen caused immediate and complete drop withdrawal, whereas partial drop withdrawal resulted from

the addition of silica or other biological material. An increase in drop volume occurred when silica dust of a larger diameter (63 – 200 µm) was applied. In *Ginkgo biloba* pollination drops, conspecific pollen also induced drop withdrawal (Jin et al. 2012). Mugnaini et al. (2007) suggested that a molecular recognition mechanism to control pollination drop retraction may exist in addition to a non-specific particle-size-dependent mechanism. It is possible that proteins could play a role in a recognition mechanism. They could take part in signalling, e.g. cleaving other proteins or polysaccharides to create chemical signalling molecules, or they could act as receptors in the plasma membrane of the nucellus.

In the zooidogamous taxa, the cycads and *G. biloba,* the initial growth of the pollen tube is directed into the subepidermal layer of the nucellus, rather than towards the archegonia as in most other groups. It is unknown how the pollen tube is directed this way. There is a continuum between the apoplastic space and the pollination drop. It is possible that proteins could be localized in the nucellar apoplast in such a way as to guide the initial growth of the pollen tube. Arabinogalactan proteins have been implicated in pollen tube guidance and their expression is concomitant with pollen receptivity in a number of plant species (Wu et al. 2000; Coimbra and Duarte 2003; Losada and Herrero 2012; Suarez et al. 2013). O'Leary et al. (2004) observed that arabinogalactan protein expression peaked during pollination drop production in *Taxus* x *media*. We detected possible arabinogalactan proteins in cycad pollination drops. Cycad ovules might provide a good system for studying arabinogalactans during reproduction because of their large size. The development of the pollen tube is distinct in cycads, having an initial phase where the pollen tube grows into the nucellus away from the archegonia, and a later

phase where the proximal end of the pollen tube grows down into the archegonial chamber (Choi and Friedman 1991). If arabinogalactans play a role in guiding pollen tube growth (Coimbra and Duarte 2003), then presumably, there would be a corresponding pattern of expression in the nucellar tissues during and after pollination drop secretion. Various reagents (e.g. Yariv reagent) and antibodies are available to detect arabinogalactan proteins. Application of these reagents over the time course of pollen tube growth might reveal such a pattern.

The need for functional characterization and quantitation of pollination drop proteins

To date, the activities of only two pollination drop proteins have been demonstrated, a chitinase (Coulter et al. 2012) and an invertase (von Aderkas et al. 2012), both in Douglas-fir. The data generated in our proteomic study could provide starting points for investigating the activities of additional proteins in pollination drops. There is no doubt that functional characterization of pollination drop enzymes would be challenging. It is expensive and difficult to collect large volumes of pollination drops! Chitinase and invertase activity assays were conducted directly on pollination drop samples (Coulter et al. 2012; von Aderkas et al. 2012). Alternatively, it might be possible to clone the genes encoding pollination drop proteins from nucellus tissue by using degenerate primers for later heterologous expression. Another approach that is currently being explored is to use proteins produced in liquid culture derived from nucellus tissue (A. Coulter, University of Victoria, pers. comm.).

Carbohydrate modifying proteins and defence proteins occurred in the pollination drops of all gymnosperm lineages. Carbohydrate modifying proteins have been hypothesized to affect pollen germination and pollen tube growth. There is one simple

preliminary study that should be conducted to help determine if any proteins do, in fact, have an effect on pollen cell walls. The proteins in a sample of pollination drops should be removed or denatured, and then pollen germination should be compared in samples with or without proteins present. The general effect of proteins on pollen germination could then be evaluated. This could also serve as a preliminary test of the effect of proteins on microbial growth in the drop. It would be ideal to use fresh pollination drops. The challenge, of course, is collecting an adequate volume of sample to conduct this experiment.

No quantitative studies of pollination drop proteins have been conducted. Poulis (2004), O'Leary (2004) and Wagner (2007) used 1D and 2D gels in their proteomics approaches, which gave an indication of relative protein quantities. O'Leary identified the two most abundant proteins based on the intensity of staining in a 2D gel. However, no studies using quantitative proteomic methods, such as iTRAQ (isobaric Tags for Relative and Absolute Quantitation), have been completed. Using the iTRAQ system, it is possible to make quantitative comparisons between the proteins contained in up to eight samples (Applied Biomics website). One useful approach would be to conduct a quantitative study of pollination drop proteins over the course of pollen receptivity. Quantitative comparisons at different time points might allow us to relate the expression of particular proteins to reproductive events between the pollen and ovule. For example, we could determine protein content during pollen chamber formation, before pollination, at the time of pollen capture, and after pollination. An additional step would be to localize proteins of interest in the nucellus and ovule over the period of pollen

receptivity. These experiments would help to build a model of the events occurring in the ovule during pollination and the initial stages of pollen tube growth.

Extracting proteins from the nucellus during drop production would add another layer to our understanding of pollen-ovule interactions. Douglas-fir would be the best species to use, given that we have prepared a protein database from the nucellus transcriptome (transcriptome created by Ian Boyes). Ideally, protein extraction methods would be developed to discern apoplast, membrane and intracellular proteins in the nucellus. We would expect the nucellar apoplast proteome to be similar to the drop proteome. We might also expect to see drop proteins dominate the secretory cells of the nucellus. A nucellar proteome dataset would also allow us to look for membrane-bound receptors or other proteins that might interact with drop components, e.g. extracellular ATP.

Final reflection

From our proteomic analyses of pollination drops, we can conclude that all extant gymnosperms have complex reproductive secretions containing proteins. There is currently no consensus regarding the relationships between extant seed-plant groups. For this reason, we can only infer that the composition of reproductive secretions might have been complex as far as these extant groups reach into the evolution of seed plants, or perhaps further, if pollination drops existed before the divergence of these lineages. Little et al. (2014) recently mapped the presence of pollination drops on a phylogeny that included both extant and extinct seed-plant lineages. Evidence of liquids such as pollination drops is almost non-existent in the fossil record (Rothwell 1977). However, the presence of saccate pollen can be used as a proxy since saccate pollen is associated

with pollination mechanisms incorporating pollination drops. Little et al. (2014) concluded that pollination drops were most likely present in many extinct gymnosperms.

The study presented here had a broad scope – to detect, to identify and to compare the types of proteins present in pollination drops of different gymnosperms. The next step is to look at single protein types in much greater depth. Peptide sequence information derived from proteomic analysis should be used as the starting point to determine the full length sequences of transcripts or genes in a given species. Full length sequences could be found in transcriptome datasets where available, or confirmed through gene cloning and sequencing. With this information, phylogenetic methods could be used to reveal the evolutionary story behind specific proteins involved in gymnosperm, or even seed-plant, reproduction.

In general, research into the biochemistry of gymnosperms has lagged behind angiosperms partly due to the lack of available sequence data. This restriction will fade as large scale transcriptome projects continue to be made publicly available, e.g. Dendrome Project and OneKP Project, as the cost of transcriptomics decreases, and as clean versions of gymnosperm genomes are released, e.g. *Picea* (Birol et al. 2013; Nystedt et al. 2013) and *Pinus* (Zimin et al. 2014). This study demonstrates the value of transcriptome data for proteomics in non-model organisms. Protein identifications made through database search tools will only improve as annotations improve in databases. Studying the four extant gymnosperm groups using the latest “omics” and biochemical methods is no longer out of reach. Studies of this kind will certainly lead to important evolutionary conclusions regarding reproduction in seed plants.

Literature Cited

Agrawal, G.K., Jwa, N.S., Lebrun, M.H., Job, D. and Rakwal, R. 2010. Plant secretome: unlocking secrets of the secreted proteins. Proteomics 10: 799-827.

Albenne, C., Canut, H. and Jamet, E. 2013. Plant cell wall proteomics: the leadership of *Arabidopsis thaliana*. Frontiers in Plant Science 4: 111.

Anderson, E.D. and Owens, J.N. 1999. Megagametophyte development, fertilization, and cytoplasmic inheritance in *Taxus brevifolia*. International Journal of Plant Sciences 160: 459-469.

Anderson, E.D. and Owens, J.N. 2000. Microsporogenesis, pollination, pollen germination and male gametophyte development in *Taxus brevifolia*. Annals of Botany 86: 1033-1042.

Applied Biomics website (accessed June 23 2014): http://www.appliedbiomics.com/Services/itraq.html

Balbuena, T.S., Jo, L., Pieruzzi, F.P., Dias, L.L.C., Silveira, V., Santa-Catarina, C., Junqueira, M., Thelen, J.J., Shevchenko, A. and Floh, E.I.S. 2011. Differential proteome analysis of mature and germinated embryos of *Araucaria angustifolia*. Phytochemistry 72: 302-311.

Belenghi, B., Acconcia, F., Trovato, M., Perazzolii, M., Bocedi, A., Polticelli, F., Ascenzi, P. and Delledonne, M. 2003. AtCYS1, a cystatin from *Arabidopsis thaliana*, suppresses hypersensitive cell death. European Journal of Biochemistry 270: 2593-2604.

Berardini, T.Z., Mundodi, S., Reiser, R., Huala, E., Garcia-Hernandez, M., Zhang, P., Mueller, L.M., Yoon, J., Doyle, A., Lander, G., Moseyko, N., Yoo, D., Xu, I., Zoeckler, B., Montoya, M., Miller, N., Weems, D. and Rhee, S.Y. 2004. Functional annotation of the *Arabidopsis* genome using controlled vocabularies. Plant Physiology 135:1-11.

Bino, R.J., Dafni, A. and Meeuse, A.D.J. 1984a. Entomophily in the dioecious gymnosperm *Ephedra aphylla* Forsk (=E. alte C.A. Mey), with some notes on *Ephedra campylopoda* C.A. Mey. 1. Aspects of the entomophilous syndrome. Proceedings of the Koninklijke Nederlandse Akademie Van Wetenschappen, Series C, Biological and Medical Sciences 87: 1-13.

Bino, R.J., Devente, N. and Meeuse, A.D.J. 1984b. Entomophily in the dioecious gymnosperm *Ephedra aphylla* Forsk (=E. alte C.A. Mey), with some notes on *Ephedra campylopoda* C.A. Mey. 2. pollination droplets, nectaries, and nectarial secretion in *Ephedra*. Proceedings of the Koninklijke Nederlandse Akademie Van Wetenschappen, Series C, Biological and Medical Sciences 87: 15-24

Birol, I., Raymond, A. and Jackman, S.D. (and 20 other co-authors). 2013. Assembling the 20 Gb white spruce (*Picea glauca*) genome from whole-genome shotgun sequencing data. Bioinformatics 29: 1492-1497.

Biye, E.H., Balkwill, K. and Cron, G.V. 2014. A clarification of *Gnetum* L. (Gnetaceae) in Africa and the description of two new species. Plant Systematics and Evolution 300: 263-272.

Boller, T. 1995. Chemoperception of microbial signals of plant cells. Annual Review of Plant Physiology and Plant Molecular biology 46:189-214.

Bowe, L.M., Coat, G. and de Pamphilis, C.W. 2000. Phylogeny of seed plants based on all three genomic compartments: extant gymnosperms are monophyletic and Gnetales' closest relatives are conifers. Proceedings of the National Academy of Sciences of the United States of America 97: 4092-4097.

Brautigam, A., Shrestha, R.P., Whitten, D., Wilkerson, C.G., Carr, K.M., Froehlich, J.E. and Weber, A.P.M. 2008. Low-coverage massively parallel pyrosequencing of cDNAs enables proteomics in non-model species: Comparison of a species-specific database generated by pyrosequencing with databases from related species for proteome analysis. Journal of Biotechnology 136: 44-53.

Brewbaker, J.L. and Kwack, B.H. 1963. The essential role of calcium ion in pollen germination and pollen tube growth. American Journal of Botany 50: 859-865.

Cao, J. 2012. The pectin lyases in *Arabidopsis thaliana*: evolution, selection and expression profiles. PLoS ONE 7: e46944.

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

Carmichael, J.S. and Friedman, W.E. 1996. Double fertilization in *Gnetum gnemon* (Gnetaceae): its bearing on the evolution of sexual reproduction within the Gnetales and the Anthophyte clade. American Journal of Botany 83: 767-780.

Carter, C. and Thornburg, R.W. 2004. Is the nectar redox cycle a floral defense against microbial attack? TRENDS in Plant Science 9: 320-324.

Chae, K. and Lord, E.M. 2011. Pollen tube growth and guidance: roles of small, secreted proteins. Annals of Botany 108: 627-636.

Chalivendra, S.C., Lopez-Casado, G., Kumar, A., Kassenbrock, A.R., Royer, S., Tovar-Mendez, A., Covey, P.A., Dempsey, L.A., Randle, A.M., Stack, S.M., Rose, J.K.C., McClure, B. and Bedinger, P.A. 2013. Developmental onset of reproductive barriers and associated proteome changes in stigma/styles of *Solanum pennellii*. Journal of Experimental Botany 64: 265-279.

Chamberlain, C.J. 1935. Gymnosperms: structure and evolution. University of Chicago Press, Chicago, Illinois, USA.

Champagne, A. and Boutry, M. 2013. Proteomics of nonmodel plant species. Proteomics 13: 663-673.

Chaw, S.M., Parkinson, C.L., Cheng, Y., Vincent, T.M. and Palmer, J.D. 2000. Seed plant phylogeny inferred from all three plant genomes: monophyly of extant gymnosperms and origin of gnetales from conifers. Proceedings of the National Academy of Sciences of the United States of America 97: 4086-4091.

Chaw, S.M., Zharkikh, A., Sung, H.M., Lau, T.C. and Li, W.H. 1997. Molecular phylogeny of extant gymnosperms and seed plant evolution: analysis of nuclear 18S rRNA sequences. Molecular Biology and Evolution 14: 56-68.

Chen, G., Snyder, C.L., Greer, M.S. and Weselake, R.J. 2011. Biology and biochemistry of plant phospholipases. Critical Reviews in Plant Sciences 30: 239-258.

Chesnoy, L. 1993. Les secretions dans la pollinisation des Gymnospermes. Acta Botanica Gallica 140:145-156.

Cheung, A.Y., Wang, H. and Wu, H.M. 1995. A floral transmitting tissue-specific glycoprotein attracts pollen tubes and stimulates their growth. Cell 82: 383-393.

Cheung, A.Y., Wu, H.M., di Stilio, V., Glaven, R., Chen, C., Wong, E., Ogdahl, J. and Estavillo, A. 2000. Pollen-pistil interactions in *Nicotiana tabacum*. Annals of Botany 85: 29-37.

Chevalier, E., Loubert-Hudon, A., Zimmerman, E.L. and Matton, D.P. 2011. Cell-cell communication and signaling pathways within the ovule: from its inception to fertilization. New Phytologist 192: 13-28.

Choi, J.S. and Friedman, W.E. 1991. Development of the pollen tube of *Zamia furfuracea* (Zamiaceae) and its evolutionary implications. American Journal of Botany 78: 544-560.

Christianson, M.L. and Jernstedt, J.A. 2009. Reproductive short-shoots of *Ginkgo biloba*: a quantitative analysis of the disposition of axillary structures. American Journal of Botany 96: 1957-1966.

Clarke, A., Gleeson, P., Harrison, S. and Knox, R.B. 1979. Pollen-stigma interactions: identification and characterization of surface components with recognition potential. Proclamations of the National Academy of Science USA 76: 3358-3362.

Coimbra, S. and Duarte, C. 2003. Arabinogalactan proteins may facilitate the movement of pollen tubes from the stigma to the ovules in *Actinidia deliciosa* and *Amaranthus hypochondriacus*. Euphytica 133: 171-178.

Collinge, D.B., Kragh, K.M., Mikkelsen, J.D., Nielsen, K.K., Rasmussen, U. and Vad, K. 1993. Plant chitinases. The Plant Journal 3: 31-40.

Coulter, A., Poulis, B.A.D. and P. von Aderkas. 2012. Pollination drops as dynamic apoplastic secretions. Flora 207: 482-490.

Coulter, J.M. and Chamberlain, C.J. 1910. Morphology of gymnosperms. University of Chicago Press, Chicago, USA.

Crane, P. 1985. Phylogenetic analysis of seed plants and the origin of angiosperms. Annals of the Missouri Botanical Garden 72: 716-793.

Crisp, M.D. and Cook, L.G. 2011. Cenozoic extinctions account for the low diversity of extant gymnosperms compared with angiosperms. New Phytologist 192: 997-1009.

Cruz-Garcia, F., Hancock, C.N., Kim, D. and McClure, B. 2005. Stylar glycoproteins bind to S-RNase *in vitro.* The Plant Journal 42: 295-304.

de Maio, A. and Vazquez, D. 2013. Extracellular heat shock proteins: a new location, a new function. Shock 40: 239-246.

Dehoux, E. and Pham Thi, A.T. 1980. Influence de quelques acides aminés libres de l'ovule sur la croissance et le développement cellulaire in vitro du tube pollinique chez *Juniperus communis* (Cupressaceae). Physiologia Plantarum 50: 6-10.

del Tredici, P. 2007. The phenology of sexual reproduction in *Ginkgo biloba*: ecological and evolutionary implications. Botanical Review 73: 267-278.

Delatte, T., Umhang, M., Trevisan, M., Eicke, S., Thorneycroft, D., Smith, S.M. and Zeeman, S.C. 2006. Evidence for distinct mechanisms of starch granule breakdown in plants. Journal of Biological Chemistry 281: 12050-12059.

Delaunois, B., Colby, T., Belloy, N., Conreux, A., Harzen, A., Baillieul, F., Clement, C., Schmidt, J., Jeandet, P. and Cordelier, S. 2013. Large-scale proteomic analysis of the grapevine leaf apoplastic fluid reveals mainly stress-related proteins and cell wall modifying enzymes. BMC Plant Biology 13: 24.

Delpino, F. 1868. Osservazione sulla dicogamia nel regno vegetale. Atti della società italiana di scienze naturali XI: 265-331.

Dendrome website accessed April 2013: dendrome.ucdavis.edu/treegenes/transcriptome/transcr_summary.php

Donaldson, J.S. 1997. Is there a floral parasite mutualism in cycad pollination? The pollination biology of *Encephalartos villosus* (Zamiaceae). American Journal of Botany 84: 1398-1406.

Douglas, A.W., Stevenson, D.W. and Little, D.P. 2007. Ovule development in *Ginkgo biloba* L., with emphasis on the collar and nucellus. International Journal of Plant Science 168: 1207-1236.

Dowd, P.E., Coursol, S., Skirpan, A.L., Kao, T.H. and Gilroy, S. 2006. *Petunia* phospholipase C1 is involved in pollen tube growth. The Plant Cell 18: 1438-1453.

Doyle, E.A., Lane, A.M., Sides, J.M., Mudgett, M.B. and Monroe, J.D. 2007. An alpha-amylase (At4g25000) in *Arabidopsis* leaves is secreted and induced by biotic and abiotic stress. Plant, Cell and Environment 20: 388-398.

Doyle, J. 1945. Developmental lines in pollination mechanisms in the Coniferales. Scientific Proceedings of the Royal Dublin Society Series A 24: 43–62

Doyle, J. and O'Leary, M. 1935. Pollination in *Pinus*. Scientific Proceedings, Royal Dublin Society 21: 181-190.

Doyle, J.A. 2012. Molecular and fossil evidence on the origin of angiosperms. Annual Review of Earth and Planetary Sciences: 40: 301-326.

Doyle, J.A. and Donoghue, M.J. 1986. Seed plant phylogeny and the origin of angiosperms: an experimental cladistics approach. Botanical Review 52: 321-431.

Dresselhaus, T., Lausser, A. and Marton, M.L. 2011. Using maize as a model to study pollen tube growth and guidance, cross-incompatibility and sperm delivery in grasses. Annals of Botany 108: 727-737.

Dumas, C. and Rogowsky, P. 2008. Fertilization and early seed formation. Comptes Rendus Biologies 221: 715-725.

Dupler, A.W. 1917. The gametophytes of *Taxus canadensis* Marsh. Botanical Gazette 64: 115-136.

Dupler, A.W. 1920. Ovuliferous structures in *Taxus canadensis*. Botanical Gazette 69: 492-520.

Edstam, M.M., Blomqvist, K., Eklof, A., Wennergren, U. and Edqvist, J. 2013. Coexpression patterns indicate that GPI-anchored non-specific lipid transfer proteins are involved in accumulation of cuticular wax, suberin and sporopollenin. Plant Molecular Biology 83: 625-649.

El-Ghazaly, G., Rowley, J. and Hesse, M. 1998. Polarity, aperture condition and germination in pollen grains of *Ephedra* (Gnetales). Plant Systematics and Evolution 213: 217-231.

Endress, P.K. 1996. Structure and function of female and bisexual organ complexes in Gnetales. International Journal of Plant Sciences 157: S113-S125.

Escalante-Perez, M., Jaborsky, M., Reinders, J., Kurzai, O., Hedrich, R. and Ache, P. 2012. Poplar extrafloral nectar is protected against plant and human pathogenic fungus. Molecular Plant 5: 1157-1159.

Evans, V.C., Barker, G., Heesom, K.J., Fan, J., Bessant, C. and Matthews, D.A. 2012. De novo derivation of proteomes from transcriptomes for transcript and protein identification. Nature Methods 9: 1207-1214

Expasy Website, UniProt-SwissProt Statistics accessed March 2014: http://web.expasy.org/docs/relnotes/relstat.html

Fernando, D.D. 2005. Characterization of pollen tube development in *Pinus strobus* (Eastern white pine) through proteomic analysis of differentially expressed proteins. Proteomics 5: 4917-4926.

Fernando, D.D., Owens, J.N., von Aderkas, P. and Takaso, T. 1997. In vitro pollen tube growth and penetration of female gametophyte in Douglas fir (*Pseudotsuga menziesii*). Sexual Plant Reproduction 10: 209-216.

Fernando, D.D., Quinn, C.R., Brenner, E.D. and Owens, J.N. 2010. Male gametophyte development and evolution in extant gymnosperms. International Journal of Plant Developmental Biology 4: 47-63.

Fierens, E., Rombouts, S., Gebruers, K., Goesaert, H., Brijs, K., Beaugrand, J., Volckaert, G., van Campenhout, S., Proost, P., Courtin, C.M. and Delcour, J.A. 2007. TLXI, a novel type of xylanase inhibitor from wheat (*Triticum aestivum*) belonging to the thaumatin family. Biochemistry Journal 403: 583-591.

Friedman, W.E. 1987. Growth and development of the male gametophyte of *Ginkgo biloba* within the ovule (in vivo). American Journal of Botany 74: 1797-1815.

Fujii, K. 1903. Über die Bestäubungstropfen der Gymnospermen, Berichte der Deutschen Botanischen Gesellschaft 21: 211-217.

Garcia-Casado, G., Collada, C., Allona, I., Soto, A. Casado, R., Rodriguez-Cerezo, E., Gomez, L. and Aragoncillo, C. 2000. Characterization of an apoplastic basic thaumatin-like protein from recalcitrant chesnut seeds. Physiologia Plantarum 110: 172-180,

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

Ghosh, S., Meli, V.S., Kumar, A., Thakur, A., Chakraborty, N., Chakraborty, S. and Datta, A. 2011. The N-glycan processing enzymes α-mannosidase and β-D-N-acetylhexosaminidase are involved in ripening-associated softening in the non-climacteric fruits of *Capsicum*. Journal of Experimental Botany 62: 571-582.

Gifford, E.M. and Foster, A.S. 1989. Morphology and Evolution of Vascular Plants 3rd edition. W.H. Freeman and Company, New York, USA.

Gleeson, P.A., McNamara, M., Wettenhall, R.E.H., Stone, B.A. and Fincher, G.B. 1989. Characterization of the hydroxyproline-rich protein of an arabinogalactan-protein secreted from suspension-cultured *Lolium multiforum* (Italian ryegrass) endosperm cells. Biochemical Journal 264:857-862.

Gong, M., Yang, Z.H. and Tsao, T.H. 1993. Isolation and characterization of calmodulin and a novel calcium-binding protein calpollenin from *Pinus yunnanensis* pollen. Plant Science 89: 5-12.

Gonzalez-Teuber, M., Eilmus, S., Muck, A., Svatos, A. and Heil, M. 2009. Pathogenesis-related proteins protect extrafloral nectar from microbial infestation. The Plant Journal 58: 464-473.

Gonzalez-Teuber, M., Pozo, M.J., Muck, A., Svatos, A., Adame-Alvarez, R.M. and Heil, M. 2010. Glucanases and chitinases as causal agents in the protection of *Acacia* extrafloral nectar from infestation by phytopathogens. Plant Physiology 152: 1705-1715.

Griffith, M., Antikainen, M., Hon, W.C., Pihakaski-Maunsbach, K., Yu, X.M., Chun, J.U. and Yang, D.S.C. 1997. Antifreeze proteins in winter rye. Physiologia Plantarum 100: 327-332.

Grover, A. 2012. Plant chitinases: genetic diversity and physiological roles. Critical Reviews in Plant Sciences 31: 57-73.

Guan, X., Zhao, H., Xu, Y. and Wang, Y. 2011. Transient expression of glyoxal oxidase from the Chinese wild grape *Vitis pseudoreticulata* can suppress powdery mildew in a susceptible genotype. Protoplasma 248: 415-423.

Guevara-Morato, M.A., de Lacoba, M.G., Garcia-Luque, I. and Serra, M.T. 2010. Characterization of a pathogenesis-related protein 4 (PR-4) induced in *Capsicum chinense* L^3 plants with duel RNase and DNase activities. Journal of Experimental Botany 61: 3259-3271.

Günl, M., Neumetzler, L., Kraemer, F., de Souza, A., Schultink, A., Pena, M., York, W.S. and Pauly, M. 2011. AXY8 encodes an alpha-fucosidase, underscoring the importance of apoplastic metabolism on the fine structure of *Arabidopsis* cell wall polysaccharides. The Plant Cell 23: 4025-4040.

Hajibabaei, M., Xia, J. and Drouin, G. 2006. Seed plant phylogeny: gnetophytes are derived conifers and a sister group to Pinaceae. Molecular Phylogenetics and Evolution 40: 208-217.

Hall, J.A., Walter, G.H., Bergstrom, D.M. and Machin, P. 2004. Pollination ecology of the Australian cycad *Lepidozamia peroffskyana* (Zamiaceae). Australian Journal of Botany 52: 333-343.

Hawes, M.C., Curlando-Rivera, G., Xiong, Z. and Kessler, J.O. 2012. Roles of border cells in plant defense and regulation of rhizosphere microbial populations by extracellular DNA 'trapping'. Plant Soil 355: 1-16.

Haycraft, C.J. and Carmichael, J.S. 2001. Development of sterile ovules of bisexual cones of *Gnetum gnemon* (Gnetaceae). American Journal of Botany 88: 1326-1330.

Henschel, J.R., and Seely, M.K. 2000. Long-term growth patterns of *Welwitschia mirabilis*, a long-lived plant of the Namib Desert. Plant Ecology 150: 7-26.

Herrmann, A., Konig, S., Lechtenberg, M., Sehlbach, M., Vakhrushev, S.Y., Peter-Katalinic, J. and Hensel, A. 2012. Proteoglycans from *Boswellia serrata* Roxb. and *B. carteri* Birdw. and identification of a proteolytic plant basic secretory protein. Glycobiology 22: 1424-1439.

Herzog, M., Dorne, A.M. and Grellet, F. 1995. GASA, a gibberellin-regulated gene family from *Arabidopsis thaliana* related to the tomato GAST1 gene. Plant Molecular Biology 27: 743-752

Hillwig, M.S., Kanobe, C., Thornburg, R.W. and MacIntosh, G.C. 2011. Identification of S-RNase and peroxidase in petunia nectar. Journal of Plant Physiology 168: 734-738.

Hillwig, M.S., Liu, X., Liu, G., Thornburg, R.W. and MacIntosh, G.C. 2010. *Petunia* nectar proteins have ribonuclease activity. Journal of Experimental Botany 61: 2951-2965.

Hiscock, S.J. and Allen, A.M. 2008. Diverse cell signaling pathways regulate pollen-stigma interactions: the search for consensus. New Phytologist 179: 286-317.

Huesgen, P.I. and Overall, C.M. 2012. N- and C-terminal degradomics: new approaches to reveal biological roles for plant proteases from substrate identification. Physiologia Plantarum 145: 5-17.

Huh, W.K., Kim, S.T., Yang, K.S., Seok, Y.J., Hah, Y.C. and Kang, S.O. 1994. Characterization of D-arabinono-1,4-lactone oxidase from *Candida albicans* ATCC 10231. European Journal of Biochemistry 225: 1073-1079.

Ikeno, S. 1898. Untersuchungen ueber die entwickelung der geschlechtsorgane und den vorgang der befruchtung bei *Cycas revoluta*. The Journal of the College of Science, Imperial University of Tokyo, Japan 12: 151-214.

Irie, K., Hosoyama, H., Takeuchi,T., Iwabuci, K., Watanabe, H., Abe, M., Abe, K. and Arai, S. 1996. Transgenic rice established to express corn cystatin exhibits strong inhibitory activity against insect gut proteinases. Plant Molecular Biology 30: 149-157.

IUCN Red List 2013.Website: www.iucnredlist.org/about/summary-statistics#Tables_3_4

IUCN Red List of Threatened Species website: http://www.iucnredlist.org/details/42081/0

IUCN website: http://www.iucn.org

Jiang, W., Chu, S.H., Piao, R., Chin, J.H., Jin, Y.M., Lee, J., Qiao, Y., Han, L., Piao, Z. and Koh, H.J. 2008. Fine mapping and candidate gene analysis of *hwh1* and *hwh2*, a set of complementary genes controlling hybrid breakdown in rice. Theoretical and Applied Genetics 116: 1117-1127.

Jin, B., Zhang, L., Lu, Y., Wang, D., Jiang, X.X., Zhang, M. and Wang, L. 2012. The mechanism of pollination drop withdrawal in *Ginkgo biloba* L. BMC Plant Biology 12: 1-9.

Jo, L., dos Santos, A.L.W., Bueno, C.A., Barbosa, H.R. and Floh, E.I.S. 2014. Proteomic analysis and polyamines, ethylene and reactive oxygen species levels of *Araucaria angustifolia* (Brazilian pine) embryogenic cultures with different embryogenic potential. Tree Physiology 34: 94-104.

Jupe, F., Witek, K., Verweij, W., Sliwka, J., Pritchard, L., Ehterington, G.J., Maclean, D., Cock, P.J., Leggett, R.M., Bryan, G.J., Cardle, L., Hein, I. and Jones, J.D.G. 2013. Resistance gene enrichment sequencing (RenSeq) enables reannotation of the NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. The Plant Journal 76: 530-544.

Kato, M., Inoue, T. and Nagamitsu, T. 1995. Pollination biology of *Gnetum* (Gnetaceae) in a lowland mixed dipterocarp forest in Sarawak. American Journal of Botany 82: 862-868.

Kawano, T. 2003. Roles of the reactive oxygen species-generating peroxidase reactions in plant defense and growth induction. Plant Cell Reports 21: 829-837.

Kim, J.E., Jeong, H.W., Nam, J.O., Lee, B.H., Choi, J.Y., Park, R.W., Park, J.Y. and Kim, I.S. 2002. Identification of motifs in the fascilin domains of the transforming growth factor-beta-induced matrix protein beta-ig-h3 that interact with alpha-v-beta-5 integrin. Journal of Biological Chemistry 277: 46159-46165.

Kim, S.T., Zhang, K., Dong, J. and Lord, E.M. 2006. Exogenous free ubiquitin enhances lily pollen tube adhesion to an in vitro stylar matrix and may facilitate endocytosis of SCA. Plant Physiology 142: 1397-1411.

Kono, M. and Tobe, H. 2007. Is *Cycas revoluta* (Cycadaceae) wind- or insect-pollinated? American Journal of Botany 94: 847-855.

Konozy, E.H.E., Rogniaux, H., Causse, M. and Faurobert, M. 2013. Proteomic analysis of tomato (*Solanum lycopersicum*) secretome. Journal of Plant Research 126: 251-266.

Kuwabara, C., Arakawa, K. and Yoshida, S. 1999. Abscisic acid-induced secretory proteins in suspension-cultured cells of winter wheat. Plant Cell Physiology 40: 184-191.

Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, 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. 2012. The *Arabidopsis* Information Resource (TAIR): improved gene annotation and new tools. Nucleic Acids Research 40: D1202-D1210.

Lazzaro, M.D., L. Cardenas, A.P. Bhatt, C.D. Justus, M.S. Phillips, Holdaway-Clarke, T.L. and Hepler, P.K. 2005. Calcium gradients in conifer pollen tubes; dynamic properties differ from those seen in angiosperms. Journal of Experimental Botany 56: 2619-2628.

Lee, E.J., Matsumura, Y., Soga, K., Hoson, T. and Koizumi, N. 2007. Glycosyl hydrolases of cell wall are induced by sugar starvation in *Arabidopsis*. Plant Cell Physiology 48:405-413.

Lee, E.K., Cibrian-Jaramillo, A., Kolokotronis, S.O., Katari, M.S., Stamatakis, A., Ott, M., Chiu, J.C., Little, D.P., Stevenson, D.Wm., McCombie, W.R., Martienssen, R.A., Coruzzi, G. and DeSalle, R. 2011. A functional phylogenomic view of the seed plants. PLoS Genetics 7: e1002411.

Leslie, A.B. 2010. Flotation preferentially selects saccate pollen during conifer pollination. New Phytologist 188: 273-279.

Lewis, J.D., Wu, R., Guttman, D.S. and Desveaux, D. 2010. Allele-specific virulence attenuation of the *Pseudomonas syringae* HopZ1a Type III Effector via the *Arabidopsis* ZAR1 resistance protein. PLoS Genetics 6: e1000894.

Lind, J.L., Bacic, A., Clarke, A.E. and Anderson, M.A. 1994. A style-specific hydroxyproline-rich glycoprotein with properties of both extensins and arabinogalactan proteins. The Plant Journal 6: 491-502

Lipchinsky, A. 2013. How do expansins control plant growth? A model for cell wall loosening via defect migration in cellulose microfibrils. Acta Physiologia Plantarum 35: 3277-3284.

Lippert, D., Zhuang, J., Ralph, S., Ellis, D.E., Gilbert, M., Olafson, R., Ritland, K., Ellis, B., Douglas, C.J. and Bohlmann, J. 2005. Proteome analysis of early somatic embryogenesis in *Picea glauca.* Proteomics 5: 461-473.

Little, S.A., Jacobs, B., McKechnie, S.J., Cooper, R.L., Christianson, M.L. and Jernstedt, J.A. 2013. Branch architecture in *Ginkgo biloba*: wood anatomy and long shoot-short shoot interactions. American Journal of Botany 100: 1923-1935.

Little, S.A., Prior, N.A., Pirone, C. and von Aderkas, P. 2014. Pollen- ovule interactions in gymnosperms. In: Ramawat, K.G., Mérillon, J.M. and K.R. Shivanna (eds.).

Reproductive Biology of Plants. CRC Press, Boca Raton, London, New York. 97-117.

Liu, J.J., Sturrock, R. and Ekramoddoullah, K.M. 2010. The superfamily of thaumatin-like proteins: its origin, evolution, and expression towards biological function. Plant Cell Reports 29: 419-436.

Lopez-Casado, G., Covey, P.A., Bedinger, P.A., Mueller, L.A., Thannhauser, T.W., Zhang, S., Fei, Z., Giovannoni, J.J. and Rose, J.KC. 2012. Enabling proteomic studies with RNA-Seq: the proteome of tomato pollen as a test case. Proteomics 12: 761-774.

Losada, J.M. and Herrero, M. 2012. Arabinogalactan-protein secretion is associated with the acquisition of stigmatic receptivity in the apple flower. Annals of Botany 110: 573-584.

Ma, B. and Johnson, R. 2012. De novo sequencing and homology searching. Molecular and Cellular Proteomics 11: O111.014902.

Majeau, N., Trudel, J. and Asselin, A. 1990. Diversity of cucumber chitinase isoforms and characterization of one seed basic chitinase with lysozyme activity. Plant Science 68: 9-16.

Maldonado, A.M., Doerner, P., Dixon, R.A., Lamb, C.J. and Cameron, R.K. 2002. A putative lipid transfer protein involved in systemic resistance signaling in *Arabidopsis*. Nature 419: 399-403.

Martens, P. 1978. Some cytologic and ontogenetic particularities of a monster plant *Welwitschia-mirabilis*. Bulletin de la Classe des Sciences Academie Royale de Belgique 5E Serie 64: 337-339.

Martin, L.B.B., Fei, Z., Giovannoni, J.J. and Rose, J.K.C. 2013. Catalyzing plant science research with RNA-seq. Frontiers in Plant Science 4: 1-10.

Martinez, L.C.A., Artabe, A.E. and Bodnar, J. 2012. A new cycad stem from the Cretaceous in Argentina and its phylogenetic relationships with other Cycadales. Botanical Journal of the Linnean Society 170: 436-458.

Mathews, S. 2009. Phylogenetic relationships among seed plants: persistent questions and the limits of molecular data. American Journal of Botany 96: 228-236.

Mathews, S., Clements, M.D. and Beilstein, M.A. 2010. A duplicate gene rooting of seed plants and the phylogenetic position of flowering plants. Philosophical Transactions of the Royal Society: Biological Sciences 365: 383-395.

Mauch, F. and Staehelin, L.A. 1989. Functional implications of the subcellular localization of ethylene-induced chitinase and β-1,3-glucanase in bean leaves. The Plant Cell 1: 447-457.

Mayer, A.M. and Staples, R.C. 2002. Laccase: new functions for an old enzyme. Phytochemistry 60: 551-565.

McWilliam, J.R. 1958. The role of the micropyle in the pollination of *Pinus*. Botanical Gazette 120: 109-117.

Meeuse, A.D.J., de Meijer, A.H., Mohr, O.W.P and Wellinga, S.M. 1990. Entomophily in the dioecious gymnosperm *E. aphylla* Forsk. (= *E. alte* C. A. Mey) with some notes on *E. campylopoda* C. A. Mey. III. Further anthecological studies and relative importance of entomophily. Israel Journal of Botany 39: 113-123.

Miernyk, J.A., Pretova, A., Oldedilla, A., Klubicova, K., Obert, B. and Hajduch, M. 2011. Using proteomics to study sexual reproduction in angiosperms. Sexual Plant Reproduction 24: 9-22.

Minic, Z. 2008. Physiological roles of plant glycoside hydrolases. Planta 227: 723-740.

Mizuno, S., Osakabe, Y., Maruyama, K., Ito, T., Osakabe, K., Sato, T., Shinozaki, K. and Yamaguchi-Shinozaki, K. 2007. Receptor-like protein kinase 2 (RPK 2) is a novel factor controlling anther development in *Arabidopsis thaliana*. The Plant Journal 50: 751-766.

Moussel, B. 1980. Gouttelette receptrice du pollen et pollinisation chez l'*Ephedra distachya* L.: observations sur le vivant et en microscopies photonique et electronique. Revue de Cytologie et de Biologie Végétales – le Botaniste 3: 65-89.

Mugnaini, S., Nepi, M., Guarnieri, M, Piotto, B. and Pacini, E. 2007. Pollination drop in *Juniperus communis*: response to deposited material. Annals of Botany 100: 1475-1481.

Murch, S. and Tomlinson, P.B. 2011. Chemical composition of pollination drops in *Gnetum gnemon* (Gnetales) suggests gender differences. Abstract for Botany 2011, St. Louis, MO, July 9-13, 2011.

Nagalingum, N.S., Marshall, C.R., Quental, T.B., Rai, H.S., Little, D.P. and Mathews, S. 2011. Recent synchronous radiation of a living fossil. Science 334:796-799.

Nakamura, Y., Awai, K., Masuda, T., Yoshioka, Y., Takamiya, K. and Ohta, H. 2005. A novel phosphatidylcholine-hydrolyzing phospholipase C induced by phosphate starvation in *Arabidopsis*. Journal of Biological Chemistry 280: 7469-7476.

Nakao, Y., Tateishi, A., Kawase, K., Ogata, T., Shiozaki, S. and Horiuchi, S. 1998. Seed set of *Ginkgo biloba* L. as related to pollination and its optimum pollination time. Journal of the Japanese Society of Horticultural Science 67: 753-758.

Nepi, M. 2014. Beyond nectar sweetness: the hidden ecological role of non-protein amino acids in nectar. Journal of Ecology 102: 108-115.

Nepi, M., Bini, L., Bianchi, L., Puglia, M., Abate, M. and Cai, G. 2011. Xylan-degrading enzymes in male and female flower nectar of *Curcurbita* pepo. Annals of Botany 108: 521-527.

Nepi, M., von Aderkas, P., Wagner, R., Mugnaini, S., Coulter, A. and Pacini, E. 2009. Nectar and pollination drops: how different are they? Annals of Botany 104: 205-219.

Nesvizhskii, A. I. 2010. A survey of computational methods and error rate estimation procedures for peptide and protein identification in shotgun proteomics. Journal of Proteomics 73: 2092-2123.

Nguema-Ona, E., Coimbra, S., Vicre-Gibouin, M., Mollet, J.C. and Driouich, A. 2012. Arabinogalactan proteins in root and pollen-tube cells: distribution and functional aspects. Annals of Botany 110: 383-404.

Nielsen, K.K., Nielsen, J.E., Madrid, S.M. and Mikkelsen, J.D. 1996. New antifungal proteins from sugar beet (*Beta vulgaris* L.) showing homology to non-specific lipid transfer proteins. Plant Molecular Biology 31: 539-552.

Norstog, K.J. and Fawcett, P.K.S. 1989. Insect-cycad symbiosis and its relation to the pollination of *Zamia furfuracea* (Zamiaceae) by *Rhopalotria mollis* (Curculionidae). American Journal of Botany 76: 1380-1394.

Norstog, K.J. and Nicholls, T.J. 1997. The biology of the cycads. Cornell University Press, Ithaca, New York, USA.

Norstog, K.J., Stevenson, D.W.M. and Niklas, K.J. 1986. The role of beetles in the pollination of *Zamia furfuracea* L. fil. (Zamiaceae). Biotropica 18: 300-306.

Nygaard, P. 1977. Utilization of exogenous carbohydrates for tube growth and starch synthesis in pine pollen suspension cultures. Physiologia Plantarum 39: 206-210.

Nystedt, B., Street, N.R., and Wetterborn, A. (and 53 other co-authors) 2013. The Norway spruce genome sequence and conifer genome evolution. Nature 497: 579-584.

O'Leary, S.J.B. 2004. Proteins in the ovular secretions of conifers. PhD Dissertation, University of Victoria, Canada.

O'Leary, S.J.B., Chani, J. and von Aderkas, P. 2004. Origin of arabinogalactan proteins in the pollination drop of *Taxus x media*. Austrian Journal of Forest Science 121: 35-46.

O'Leary, S.J.B., Poulis, B.A.D and 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 defense during pollen collection. Tree Physiology 27: 1649-1659.

O'Leary, S.J.B and von Aderkas, P. 2006. Postpollination drop production in hybrid larch is not related to the diurnal pattern of xylem water potential. Trees 20: 61-66.

Ohmiya, Y., Samejima, M., Shiroishi, M., Amano, Y., Kanda, T., Sakai, F. and Hayashi, T. 2000. Evidence that endo-1,4-β-glucanases act on cellulose in suspension-cultured poplar cells. The Plant Journal 24:147-158.

Okamoto, T. and Kranz, E. 2005. In vitro fertilization – a tool to dissect cell specification from a higher plant zygote. Current Science 89: 1861-1869.

One KP Capstone Wiki website accessed March 2014: https://pods.iplantcollaborative.org/wiki/display/iptol/OneKP+Capstone+Wiki

One KP website accessed March 2014: sites.google.com/a/ualberta.ca/onekp/

Ortiz, P.L., Arista, M. and Talavera, S. 1998. Low reproductive success in two subspecies of *Juniperus oxycedrus* L. International Journal of Plant Science 159: 843-847.

Osborne, R., Calonje, M., Hill, K.D., Stanberg, L. and Stevenson, D.W. 2012. The world list of cycads. Memoirs of the New York Botanical Garden 106: 480-510.

Owens, J.N., Bennett, J. and 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, J.N. and Blake, M.D. 1984. The pollination mechanism of Sitka spruce (*Picea sitchensis*). Canadian Journal of Botany 62: 1136-1148.

Owens, J.N. and Molder, M. 1980. Sexual reproduction in western red cedar (*Thuja plicata*). Canadian Journal of Botany 58: 1376-1392.

Owens, J.N., Simpson, S.J., and Molder, M. 1980. The pollination mechanism in yellow cypress (*Chamaecyparis nootkatensis*). Canadian Journal of Forest Research 10: 564-572.

Owens, J.N., Takaso, T. and Runions, C.J. 1998. Pollination in conifers. Trends in Plant Science 3: 479-485.

Pan, Y.Y., Wang, X., Ma, L.G. and Sun, D.Y. 2005. Characterization of phosphatidylinositol-specific phospholipase C (PI-PLC) from *Lilium daviddi* pollen. Plant Cell Physiology 46: 1657-1665.

Park, S.Y., Jauh, G.Y., Mollet, J.C., Eckard, K.J., Nothnagel, E.A., Walling, L.L.and Lord, E.M. 2000. A lipid transfer-like protein is necessary for lily pollen tube adhesion to an in vitro stylar matrix. The Plant Cell 12: 151-163.

Park, S. and Thornburg, R.W. 2009. Biochemistry of Nectar Proteins. Journal of Plant Biology 52: 27-34.

PEAKS website accessed March 2014: http://www.bioinfor.com/doc/peaks6/peaks6.pdf

Pettitt, J.M. 1982. Ultrastructural and immunocytochemical demonstration of gametophytic proteins in the pollen tube wall of the primitive gymnosperm *Cycas*. Journal of Cell Science 57:189-213.

Peumans, W.J., Smeets, K., van Nerum, K., van Leuven, F. and van Damme, E.J.M. 1997. Lectin and alliinase are the predominant proteins in nectar from leek (*Allium porrum* L.) flowers. Planta 201: 298-302.

Plaxton, W.C. 1996. The organization and regulation of plant glycolysis. Annual Review of Plant Physiology and Plant Molecular Biology 47:185-214.

Pokotylo, I., Pejchar, P., Potocky, M., Kocourkova, D., Krckova, Z., Ruelland, E., Kravets, V. and Martinec, J. 2013. The plant non-specific phospholipase C gene family: novel competitors in lipid signaling. Progress in Lipid Research 52: 62-79.

Popovic, M., Andjelkovic, U., Burazer, L., Lindner, B., Petersen, A. and Gavrovic-Jankulovic, M. 2013. Biochemical and immunological characterization of a recombinantly-produced antifungal proteinase inhibitor from green kiwifruit (*Actinidia deliciosa*). Phytochemistry 94: 53-59.

Porsch, O. 1910. *Ephedra campylopoda* CA Mey, eine entomophile Gymnosperme. Berichte der Deutschen Botanischen Gesellschaft 28: 404-412.

Poulis, B.A.D. 2004. Safe sex in Douglas-fir. Ph.D. Dissertation, University of Victoria, Canada.

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 Science 166: 733-739.

Prakash, L. and Prathapasenan, G. 1990. NaCl- and gibberellic acid-induced changes in the content of auxin and the activities of cellulase and pectin lyase during leaf growth in rice (*Oryza sativa*). Annals of Botany 65: 251-257.

Prior, N., Little, S.A., Pirone, C., Gill, J.E., Smith, D., Han, J., Hardie, D., O'Leary, S.J.B., Wagner, R.E., Cross, T., Coulter, A., Borchers, C., Olafson, R.W. and von Aderkas, P. 2013. Application of proteomics to the study of pollination drops. Applications in Plant Sciences 1: 1300008

Proches, S. and Johnson, S.D. 2009. Beetle pollination of the fruit-scented cones of the South African cycad *Stangeria eriopus*. American Journal of Botany 96: 1722-1730.

Rai, H.S., Reeves, P.A., Peakall, R., Olmstead, R.G. and Graham, S.W. 2008. Inference of higher-order conifer relationships from a multi-locus plastid data set. Botany 86: 658-669.

Regente, M.C., Giudici, A.M., Villalain, J. and de la Canal, L. 2005. The cytotoxic properties of a plant lipid transfer protein involve membrane permeabilization of target cells. Letters in Applied Microbiology 40: 183-189.

Reid, J.S.G. and Meier, H. 1973. Enzymatic activities and galactomannan mobilisation in germinating seeds of fenugreek (*Trigonella foenum-graecum* L. Leguminosae). Planta 112: 301-308.

Rejon, J.D., Delalande, F., Schaeffer-Reiss, C., Carapito, C., Zienkiewicz, K., de Dios Alche, J., Rodriguez-Garcia, M.I., van Dorsselaer, A. and Castro, A.J. 2013. Proteomics profiling reveals novel proteins and functions of the plant stigma exudate. Journal of Experimental Botany 64: 5696-5705.

Roberts, I.N., Caputo, C., Criado, M.V. and Funk, C. 2012. Senescence-associated proteins in plants. Physiologia Plantarum 145: 130-139.

Rothwell, G.W. 1977. Evidence for pollination-drop mechanism in Paleozoic pteridosperms. Science 198: 1251-1252.

Rothwell, G.W. and Stockey, R.A. 2013. Evolution and phylogeny of gnetophytes: evidence from the anatomically preserved seed cone *Protoephridites eamesii* sp. nov. and the seeds of several Bennettitalean seed. International Journal of Plant Sciences 174: 511-529.

Royer, D.L., Hickey, L.J. and Wing, S.L. 2003. Ecological conservatism in the "living fossil" *Ginkgo*. Paleobiology 29: 84-104.

Rydin, C. and Friis, E.M. 2005. Pollen germination in *Welwitschia mirabilis* Hook. f.: differences between the polyplicate pollen producing genera of the Gnetales. Grana 44: 137-141.

Rydin, C., Khodabandeh, A. and Endress, P.K. 2010. The female reproductive unit of *Ephedra* (Gnetales): comparative morphology and evolutionary perspectives. Botanical Journal of the Linnean Society 163: 387-430.

Sadanandom, A., Bailey, M., Ewan, R., Lee, J. and Nelis, S. 2012. The ubiquitin-proteasome system: central modifier of plant signaling. New Phytologist 196: 13-28.

Said, C., Villar, M. and Zandonella, P. 1991. Ovule receptivity and pollen viability in Japanese larch (*Larix leptolepis* Gord.). Silvae Genetica 40: 1-6.

Salas-Leiva, D.E., Meerow, A.W., Calonje, M., Griffith, M.P., Francisco-Ortega, J., Nakamura, K., Stevenson, D.W., Lewis, C.E. and Namoff, S. 2013. Phylogeny of the cycads based on multiple single-copy nuclear genes: congruence of concatenated parsimony, likelihood and species tree inference methods. Annals of Botany 112: 1263-1278.

Sánchez-Rotonda, G., Vazquez-Torres, M. and Sanchez-Tinoco, M.Y. 1995. Entomofilia en una poblacion natural de *Ceratozamia mexicana* Brongn. (Zamiaceae). La ciencia y el hombre 7: 95-110.

Sang, Y.L., Xu, M., Ma, F.F., Chen, H., Xu, X.H., Gao, X.Q. and Zhang, X.S. 2012. Comparative proteomic analysis reveals similar and distinct features of proteins in dry and wet stigmas. Proteomics 12: 1983-1998.

Schmitt, E., Gehrmann, M., Brunet, M., Multhoff, G. and Garrido, C. 2007. Intracellular and extracellular functions of heat shock proteins: repercussions in cancer therapy. Journal of Leukocyte Biology 81: 15-27.

Seifert, G.J. and Roberts, K. 2007. The biology of arabinogalactan proteins. Annual Review of Plant Biology 58: 137-161.

Seo, P.J., Wielsch, N., Kessler, D., Svatos, A., Park, C.M., Baldwin, I.T. and Kim, S.G. 2013. Natural variation in floral nectar proteins of two *Nicotiana attenuata* accessions. BMC Plant Biology 13: 101

Seridi-Benkaddour, R. and Chesnoy, L. 1988. Secretion and composition of the pollination drop in the *Cephalotaxus drupacea* (Gymnosperm, Cephalotaxaceae). In: Cresti, M., Gori, P., and Pacini, E. (eds.). Sexual Reproduction in Higher Plants. Springer-Verlag, Berlin, Germany. 345-350.

Seymour, R.S., Terry, I. and Roemer, R.B. 2004. Respiration and thermogenesis by cones of the Australian cycad *Macrozamia machinii*. Functional Ecology 18: 925-930.

Sheng, X.Y., Dong, X.L., Zhang, S.S., Jiang, L.P., Tan, L.L. and Li, X. 2011. Unequal distribution of ubiquitinated proteins during *Pinus bungeana* pollen development. Trees – Structure and Function 25: 407-414.

Shi, J., Zhen, Y. and Zheng, R.H. 2010. Proteome profiling of early seed development in *Cunninghamia lanceolata* (Lamb.) Hook. Journal of Experimental Botany 61: 2367-2381.

Shi, L., and Olszewski, N.E. 1998. Gibberellin and abscisic regulate GAST1 expression at the level of transcription. Plant Molecular Biology 38: 1053-1060.

Shibuya, N. and Minami, E. 2001. Oligosaccharide signalling for defence responses in plant. Physiological and Molecular Plant Pathology 59: 223-233.

Showalter, A.M. 2001. Arabinogalactan-proteins: structure, expression and function. Cellular and Molecular Life Sciences 58: 1399-1417.

Singh, H. 1978. Embryology of Gymnosperms. Gebruder Borntraeger, Berlin-Stuttgart, Germany.

Solomon, M., Belenghi, B., Delledonne, M., Menachem, E. and Levine, A. 1999. The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. The Plant Cell 11: 431-443.

Soltis, D.E., Soltis, P.S. and Zanis, M.J. 2002. Phylogeny of seed plants based on evidence from eight genes. American Journal of Botany 89: 1670-1681.

Soranzo, N., Gorla, M.S., Mizzi, L., De Toma, G. and Frova, C. 2004. Organization and structural evolution of the rice glutathione *S*-transferase gene family. Molecular Genetics and Genomics 271: 511-521.

Steen, H. and Mann, M. 2004. The abc's (and xyz's) of peptide sequencing. Nature Reviews: Molecular and Cellular Biology 5: 699-711.

Stevenson, D.W. 1990. Morphology and systematics of the Cycadales. Memoirs of the New York Botanical Garden 57: 8-55.

Stevenson, D.W. 1992. A formal classification of the extant cycads. Brittonia 44: 220-223.

Strasburger, E. 1871. Die Bestäubung der Gymnospermen. Jenaische Zeitschrift für Medizin und Naturwissenschaft 6: 249-262.

Suarez, C., Zienkiewicz, A., Castro, A.J., Zienkiewicz, K., Majewska-Sawka, A. and Rodriguez-Garcia, M.I. 2013. Cellular localization and levels of pectins and arabinogalactan proteins in olive (*Olea europaea* L.) pistil tissues during development: implications for pollen-pistil interaction. Planta 237: 305-319.

Szewinska, J., Prabucka, B., Krawczyk, M., Mielecki, M. and Bielawski, W. 2013. The participation of phytocystatin TrcC-4 in the activity regulation of EP8, the main prolamin degrading cysteine endopeptidase in triticale seeds. Plant Growth Regulation 69: 131-137.

TAIR website (accessed May 2014): http://www.arabidopsis.org/portals/genAnnotation/functional_annotation/go.jsp

Takaso, T. 1990. "Pollination Drop" time at the Arnold Arboretum. Arnoldia 50: 2-7.

Takaso, T. and Bouman, F. 1986. Ovule and seed ontogeny in *Gnetum gnemon* L. Botanical Magazine of Tokyo 99: 241-166.

Takaso, T. and Owens, J.N.. 1996. Postpollination-prezygotic ovular secretions into the micropylar canal in *Pseudotsuga menziesii*. Journal of Plant Research 109: 147-180.

Tang, C.Q., Yang, Y., Ohsawa, M., Yi, S.R., Momohara, A., Su, W.H., Wang, H.C., Zhang, Z.Y., Peng, M.C. and Wu, Z.L. 2012. Evidence for the persistence of wild *Ginkgo biloba* (Ginkgoaceae) populations in the Dalou Mountains, southwestern China. American Journal of Botany 99: 1408-1414.

Tang, W. 1987. Insect pollination in the cycad *Zamia pumila* (Zamiaceae). American Journal of Botany 74: 90-99.

Tang, W. 1993. Nectar-like secretions in female cones of cycads. Cycad Newsletter 16: 10-13.

Tang, W. 1995. Pollination drops in female cycad cones. Palms and cycads 48: 20-21.

Tang, W. 2004. Cycad insects and pollination. Vistas in Paleobotany and Plant Morphology, Professor D.D. Pant Memorial Volume: 383-394.

Terry, I. 2001. Thrips and weevils as dual, specialist pollinators of the Australian cycad *Macrozamia communis* (Zamiaceae). International Journal of Plant Sciences 162: 1293-1305.

Terry, I., Moore, C.J., Walter, G.H., Forster, P.I., Roemer, R.B., Donaldson, J.D. and Machin, P.J. 2004. Association of cone thermogenesis and volatiles with pollinator specificity in *Macrozamia* cycads. Plant Systematics and Evolution 243: 233-247.

Terry, I., Roe, M., Tang, W. and Marler, T.E. 2009. Cone insects and putative pollen vectors of the endangered cycad, *Cycas micronesica*. Micronesica 41: 83-99.

Terry, I., Walter, G.H., Moore, C., Roemer, R. and Hull, C. 2007. Odor-mediated push-pull pollination in cycads. Science 318: 70.

Terry, L.I., Walter, G.H., Donaldson, J.S, Snow, E., Forster, P.I. and Machin, P.J. 2005. Pollination of Australian *Macrozamia* cycads (Zamiaceae): effectiveness and behaviour of specialist vectors in a dependent mutualism. American Journal of Botany 92: 931-940.

Teyssier, C., Maury, S., Beaufour, M., Grondin, C., Delaunay, A., le Mette, C., Ader, K., Cadene, M., Label, P. and Lelu-Walter, M.A. 2014. In search of markers for somatic embryo maturation in hybrid larch (*Larix* x *eurolepis*): global DNA methylation and proteomic analyses. Physiologia Plantarum 150: 271-291.

Thornburg, R.W., Carter, C., Powell, A., Mittler, R., Rizhsky, L. and Horner, H.T. 2003. A major function of the tobacco floral nectary is defense against microbial attack. Plant Systematics and Evolution 238: 211-218.

Tijmens, W.J. 1968. *Welwitschia* - a botanical octopus. Horticulture 46: 18.

Tison, P.A. 1911. Remarques sur les gouttelettes collectrices des ovules des conifères. Mémoires de la Société Linnéenne de Normandie 24: 51-66.

Tomlinson, P.B. 1991. Pollen scavenging. National Geographic Research and Exploration 7: 188-195.

Tomlinson, P.B. 2012. Rescuing Robert Brown - the origins of angio-ovuly in seed cones of conifers. Botanical Review 78: 310-334.

Tomlinson, P.B., Braggins, J.E. and Rattenbury, J.A. 1991. Pollination drop in relation to cone morphology in Podocarpaceae: a novel reproduction mechanism. American Journal of Botany 78: 1289-1303.

Tomlinson, P.B. and Takaso, T. 2002. Seed cone structure in conifers in relation to development and pollination: a biological approach. Canadian Journal of Botany 2002: 1250-1273.

Tsiatsiani, L., Gevaert, K. and van Breusegem, F. 2012. Natural substrates of plant proteases: how can protease degradomics extend our knowledge? Physiologia Plantarum 145: 28-40.

UniProt Website, TrEMBL Statistics accessed March 2014: http://www.ebi.ac.uk/uniprot/TrEMBLstats

Valdivia, E.R., Stephenson, A.G., Durachko, D.M. and Cosgrove, D. 2009. Class B beta-expansins are needed for pollen separation and stigma penetration. Sexual Plant Reproduction 22: 141-152.

van der Hoorn, R.A.L. 2008. Plant proteases: from phenotypes to molecular mechanisms. Annual Review of Plant Biology 59: 191-223.

van der Pijl, L. 1953. On the flower biology of some plants from Java with general remarks on fly-traps (species of *Annona*, *Artocarpus*, *Typhonium*, *Gnetum*, *Arisaema* and *Abroma*). Annales Bogorienses 1: 77-99.

van Ooijen, G., Mayr, G., Kasiem, M.M.A., Albrecht, M., Cornelissen, B.J.C. and Takken, F.L.W. 2008. Structure-function analysis of the NB-ARC domain of plant disease resistance proteins. Journal of Experimental Botany 59: 1383-1397.

Vaucher, J.P. 1841. Histoire physiologique des plantes d'Europe. Vol. 4. Marc Aurel Frères, Paris, France.

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

von Aderkas, P., Nepi, M., Rise, M., Buffi, F., Guarnieri, M., Coulter, A., Gill, K., Lan, P., Rzemieniak, S. and Pacini. E. 2012. Post-pollination prefertilization drops affect germination rates of heterospecific pollen in larch and Douglas-fir. Sexual Plant Reproduction 25: 215-225.

Vovides, A.P. 1991. Insect symbionts of some Mexican cycads in their natural habitat. Biotropica 23: 102-104.

Wagner, R.E. 2007. Gymnosperm pollination drop proteins and their relation to function and phylogeny. M.Sc. Thesis. University of Victoria, Canada.

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

Wagner, U., Edwards, R., Dixon, D.P. and Mauch, F. 2002. Probing the diversity of the *Arabidopsis* glutathione *S*-transferase gene family. 2002. Plant Molecular Biology 49: 515-532.

Webber, H.J. 1897. Structures occurring in the pollen tube of *Zamia*. Botanical Gazette 23: 453-459.

Wen, F., Celoy, R., Price, I., Ebola, J.J. and Hawes, M.C. 2008a. Identification and characterization of a rhizosphere β-galactosidase from *Pisum sativum* L. Plant Soil 304: 133-144.

Wen, F., Woo, H.H., Pierson, E.A., Eldhuset, T., Fossdal, G.C., Nagy, N.E. and Hawes, M.C. 2008b. Synchronous elicitation of development in root caps induces transient gene expression changes common to legume and gymnosperm species. Plant Molecular Biology Reports 27: 58-68.

Wetschnig, W. and Depisch, B. 1999. Pollination biology of *Welwitschia mirabilis* Hook. f. (*Welwitschia,* Gnetopsida). Phyton 39: 167-183.

Whitaker, C., Pammenter, N.W. and Berjak, P. 2008. Infection of the cones and seeds of *Welwitschia mirabil*is by *Aspergillus niger* var. phoenicis in the Namib-Naukluft Park. South African Journal of Botany 74: 41-50.

Willemse, M.T.M. and Vletter, A. 1995. Appearance and interaction of pollen and pistil pathway proteins in *Gasteria verrucosa* (Mill.) H. Duval. Sexual Plant Reproduction 8: 161-167.

Williams, J.H. 2009. *Amborella trichopoda* (Amborellaceae) and the evolutionary developmental origins of the angiosperm progamic phase. American Journal of Botany 96: 144-165.

Williams, J.H. 2012. Pollen tube growth rates and the diversification of flowering plant reproductive cycles. International Journal of Plant Science 173: 649-661.

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

Wilson, G.W. 2002. Insect pollination in the cycad genus *Bowenia* Hook. ex Hook. f. (Stangeriaceae). Biotropica 34:438-441

Wimalasekera, R., Pejchar, P., Holk, A., Martinec, J. and Scherer, G.F.E. 2010. Plant phosphatidylcholine-hydrolyzing phospholipases C NPC3 and NPC4 with roles in root development and brassinolide signaling in *Arabidopsis thaliana*. Molecular Plant 3: 610-625.

Wu, C.S., Chaw, S.M. and Huang, Y. 2013. Chloroplast phylogenomics indicates that *Ginkgo biloba* is sister to cycads. Genome Biology and Evolution 5: 243-254.

Wu, H.M., Wong, E., Ogdahl, J. and Cheung, A.Y. 2000. A pollen tube growth-promoting arabinogalactan protein from *Nicotiana alata* is similar to the tobacco TTS protein. The Plant Journal 22: 165-176.

Wu, X.Q., Chen, R., Zheng, M.Z., Chen, Y.M., Teng, N.J., Samaj, J., Baluska, F. and Lin, J.X. 2008. Integrative proteomic and cytological analysis of the effects of extracellular Ca(2+) influx on *Pinus bungeana* pollen tube development. Journal of Proteome Research 10: 4299-4312.

Xi, Z., Rest, J.S. and Davis, C.C. 2013. Phylogenomics and coalescent analyses resolve extant seed plant relationships. PLOS ONE 8: 1-11.

Xing, S.P., Chen, Z.K., Hu, Y.X., Zhou, F. and Lin, J.X. 2000. Ovule development, formation of pollination drop and pollination process in *Taxus chinensis* (Taxaceae). Acta Botanica Sinica 42: 126 – 132.

Yao, Y.F., Xi, Y.Z., Geng, B.Y. and Li, C.S. 2004. The exine ultrastructure of pollen grains in *Gnetum* (Gnetaceae) from China and its bearing on the relationship with the ANITA group. Botanical Journal of the Linnean Society 146: 415-425.

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

Yokota, E., Ohmori, T. and Muto, S. 2004. 21-kDa polypeptide, a low-molecular-weight cyclophilin, is released from pollen of higher plants into the extracellular medium in vitro. Planta 218: 1008-1018.

Yu, T.S., Zeeman, S.C., Thorneycroft, D., Fulton, D.C., Dunstan, H., Lue, W.L., Hegemann, B., Tung, S.Y., Umemoto, T., Chapple, A., Tsai, D.L., Wang, S.M., Smith, A.M., Chen, J. and Smith, S.M. 2005. Alpha-amylase is not required for breakdown of transitory starch in *Arabidopsis* leaves. Journal of Biological Chemistry 280: 9773-9779.

Zaidi, M.A., O'Leary, S., Wu, S., Gleddie, S., Eudes, F., Laroche, A. and Robert, L.S. 2012. A molecular and proteomic investigation of proteins rapidly released from triticale pollen upon hydration. Plant Molecular Biology 79: 101-121.

Zareie, R., Melanson, D.L. and Murphy, P.J. 2002. Isolation of fungal cell wall degrading proteins from barley (*Hordeum vulgare* L.) leaves infected with *Rhynchosporium secalis*. Molecular Plant-Microbe Interactions 15: 1031-1039.

Zgurski, J.M., Rai, H.S., Fai, Q.M., Bogler, D.J., Francisco-Ortega, J. and Graham, S.W. 2008. How well do we understand the overall backbone of cycad phylogeny? New insights from a large, multigene plastid data set. Molecular Phylogenetics and Evolution 47: 1232-1237.

Zha, H.G., Flowers, V.L., Yang, M., Chen, L.Y. and Sun, H. 2012. Acidic alpha-galactosidase is the most abundant nectarin in floral nectar of common tobacco (*Nicotiana tabacum*). Annals of Botany 109: 735-745.

Zhang, H.Q., Croes, A.F. and Linskens, H.F. 1982. Protein-synthesis in germinating pollen of *Petunia* - role of proline. Planta 154: 199-203.

Zhang, J., Xin, L., Shan, B.Z., Chen, W.W., Xie, M.J., Yuen, D., Zhang, W.M., Zhang, Z.F., Lajoie, G.A. and Ma, B. 2012. PEAKS DB: de novo sequencing assisted database search for sensitive and accurate peptide identification. Molecular and Cellular Proteomics 11: M111.010587.

Zhen, Y. and Shi, J. 2011. Evaluation of sample extraction methods for proteomic analysis of coniferous seeds. Acta Physiologiae Plantarum 33: 1623-1630.

Zhen, Y., Zhao, Z.Z., Zheng, R.H. and Shi, J.S. 2012. Proteomic analysis of early seed development in *Pinus massoniana* L. Plant physiology and Biochemistry 54: 97-104.

Zhong, B., Yonezawa, T., Zhong, Y. and Hasegawa, M. 2010. The position of Gnetales among seed plants: Overcoming pitfalls of chloroplast phylogenomics. Molecular Biology and Evolution 27: 2855-2863.

Zhou, B.J., Wang, X.P. and Wang, Y.J. 2007. cDNA cloning, expression, protein purification, and characterization of a novel glyoxal oxidase related gene from *Vitis pseudoreticulata*. Biologia Plantarum 51: 458-466.

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

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

Appendix 1. Sources and species of transcriptomic data used to create the Gymno_DB database.

From: 1KP (downloaded October 16 2013)
https://sites.google.com/a/ualberta.ca/onekp/
Illumina

Clade	Family	Species	Tissue Type	Sample Provider	RNA Extractor
Conifers	Araucariaceae	*Agathis robusta*	leaf	D. W. Stevenson	D.W. Stevenson
Conifers	Araucariaceae	*Araucaria rulei*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Araucariaceae	*Araucaria sp.*	young Leaf	D. W. Stevenson	D.W. Stevenson
Conifers	Araucariaceae	*Wollemia nobilis*	leaves	S. Graham	S. Graham
Conifers	Cephalotaxaceae	*Amentotaxus argotaenia*	leaf	D. W. Stevenson	D. W. Stevenson
Conifers	Cephalotaxaceae	*Cephalotaxus harringtonia*	leaves	S. Graham	S. Graham
Conifers	Cephalotaxaceae	*Cephalotaxus harringtonia*	branch apex including needles	M. Deyholos	M. Deyholos
Conifers	Cupressaceae	*Athrotaxis cupressoides*	leaf	S. Graham	L. DeGironimo
Conifers	Cupressaceae	*Austrocedrus chilensis*	leaves	S. Graham	S. Graham
Conifers	Cupressaceae	*Callitris gracilis*		A. Lowe	BGI
Conifers	Cupressaceae	*Callitris macleayana*		A. Lowe	BGI
Conifers	Cupressaceae	*Calocedrus decurrens*	leaves	S. Graham	S. Graham
Conifers	Cupressaceae	*Chamaecyparis lawsoniana*	leaf	D. W. Stevenson	BGI
Conifers	Cupressaceae	*Cryptomeria japonica*	branch apex including needles and cones	M. Deyholos	M. Deyholos
Conifers	Cupressaceae	*Cryptomeria japonica*	leaf	D. W. Stevenson	BGI
Conifers	Cupressaceae	*Cunninghamia lanceolata*	branch apex incl. needles	M. Deyholos	M. Deyholos
Conifers	Cupressaceae	*Cunninghamia lanceolata*	young shoot	S. Graham	BGI

Conifers	Cupressaceae	*Cupressus dupreziana*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Cupressaceae	*Diselma archeri*	leaf	D. W. Stevenson	D. W. Stevenson
Conifers	Cupressaceae	*Fokienia hodginsii*	leaf	D. W. Stevenson	D. W. Stevenson
Conifers	Cupressaceae	*Glyptostrobus pensilis*	young shoot	S. Graham	BGI
Conifers	Cupressaceae	*Juniperus scopulorum*	young shoot	S. Graham	BGI
Conifers	Cupressaceae	*Metasequoia glyptostroboides*	leaves	S. Graham	S. Graham
Conifers	Cupressaceae	*Microbiota decussata*	leaves	S. Graham	S. Graham
Conifers	Cupressaceae	*Neocallitropsis pancheri*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Cupressaceae	*Papuacedrus papuana*	leaf	D. W. Stevenson	D. W. Stevenson
Conifers	Cupressaceae	*Pilgerodendron uviferum*	leaves	S. Graham	S. Graham
Conifers	Cupressaceae	*Platycladus orientalis*	leaves	S. Graham	S. Graham
Conifers	Cupressaceae	*Sequoia sempervirens*	branch apex including needles	M. Deyholos	M. Deyholos
Conifers	Cupressaceae	*Sequoiadendron giganteum Glaucum*	leaf	D. W. Stevenson	BGI
Conifers	Cupressaceae	*Taiwania cryptomerioides*	leaves	S. Graham	S. Graham
Conifers	Cupressaceae	*Taxodium distichum*	leaves	S. Graham	S. Graham
Conifers	Cupressaceae	*Tetraclinis sp.*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Cupressaceae	*Thuja plicata*	young shoot	S. Graham	BGI
Conifers	Cupressaceae	*Thujopsis dolabrata*	young shoot	S. Graham	BGI
Conifers	Cupressaceae	*Widdringtonia cedarbergensis*	young leaf	D. W. Stevenson	L. DeGironimo
Conifers	Pinaceae	*Abies lasiocarpa*	leaf	D. W. Stevenson	BGI
Conifers	Pinaceae	*Cathaya argyrophylla*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Pinaceae	*Cedrus libani*	young shoot	S. Graham	BGI
Conifers	Pinaceae	*Keteleeria evelyniana*	young shoot	J. Leebens-Mack	J. Leebens-Mack
Conifers	Pinaceae	*Larix speciosa*	young shoot	S. Graham	S. Graham
Conifers	Pinaceae	*Nothotsuga longibracteata*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Pinaceae	*Picea engelmannii*	leaf	D. W. Stevenson	BGI
Conifers	Pinaceae	*Pinus jeffreyi*	stems and leaves	N. Stewart	N. Stewart

Conifers	Pinaceae	*Pinus parviflora*	leaf	D. W. Stevenson	D. W. Stevenson
Conifers	Pinaceae	*Pinus ponderosa*	stems and leaves	N. Stewart	N. Stewart
Conifers	Pinaceae	*Pinus radiata*	leaf	D. W. Stevenson	D. W. Stevenson
Conifers	Pinaceae	*Pseudolarix amabilis*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Pinaceae	*Pseudotsuga wilsoniana*	young shoot	S. Graham	BGI
Conifers	Pinaceae	*Tsuga heterophylla*	leaf	D. W. Stevenson	BGI
Conifers	Podocarpaceae	*Acmopyle pancheri*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Podocarpaceae	*Dacrycarpus compactus*	leaf	D. W. Stevenson	D. W. Stevenson
Conifers	Podocarpaceae	*Dacrydium balansae*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Podocarpaceae	*Falcatifolium taxoides*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Podocarpaceae	*Falcatifolium taxoides*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Podocarpaceae	*Halocarpus bidwillii*	leaves	S. Graham	S. Graham
Conifers	Podocarpaceae	*Lagarostrobos franklinii*	young end shoots	M. Ruhsam	M. Ruhsam
Conifers	Podocarpaceae	*Manoao colensoi*	leaf	D. W. Stevenson	BGI
Conifers	Podocarpaceae	*Microcachrys tetragona*	leaves	S. Graham	S. Graham
Conifers	Podocarpaceae	*Microstrobos fitzgeraldii*	young shoot	D. W. Stevenson	BGI
Conifers	Podocarpaceae	*Nageia nagi*	leaf	D. W. Stevenson	BGI
Conifers	Podocarpaceae	*Parasitaxus usta*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Podocarpaceae	*Phyllocladus hypohyllus*	leaf	D. W. Stevenson	D. Stevenson
Conifers	Podocarpaceae	*Podocarpus coriaceus*	leaf	P. Thomas	D. W. Stevenson
Conifers	Podocarpaceae	*Podocarpus rubens*	leaf	P. Thomas	D. W. Stevenson
Conifers	Podocarpaceae	*Prumnopitys andina*	young shoot	S. Graham	BGI
Conifers	Podocarpaceae	*Retrophyllum minus*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Podocarpaceae	*Saxegothaea conspicua*	young shoot	D. W. Stevenson	BGI
Conifers	Podocarpaceae	*Sundacarpus amarus*		D. Soltis	D. Soltis
Conifers	Sciadopityaceae	*Sciadopitys verticillata*	young shoot	S. Graham	BGI

Conifers	Taxaceae	*Austrotaxus spicata*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Taxaceae	*Pseudotaxus chienii*		J. Leebens-Mack	J. Leebens-Mack
Conifers	Taxaceae	*Taxus baccata*	mature leaves and small woody branch	M. Deyholos	M. Deyholos
Conifers	Taxaceae	*Taxus cuspidata*	branch apex including needles	M. Deyholos	M. Deyholos
Conifers	Taxaceae	*Torreya nucifera*	young shoot	S. Graham	BGI
Conifers	Taxaceae	*Torreya taxifolia*	young shoots	J. Leebens-Mack	J. Leebens-Mack
Cycadales	Cycadaceae	*Cycas micholitzii*	leaves	Tao Chen	BGI
Cycadales	Stangeriaceae	*Stangeria eriopus*	young Leaf	D. W. Stevenson	D.W. Stevenson
Cycadales	Zamiaceae	*Dioon edule*	leaf	D. W. Stevenson	BGI
Cycadales	Zamiaceae	*Encephalartos barteri*		D. W. Stevenson	BGI
Ginkgoales	Ginkgoaceae	*Ginkgo biloba*	developing shoots	D. W. Stevenson	M. Deyholos
Gnetales	Ephedraceae	*Ephedra sinica*	shoot	M. Deyholos	M. Deyholos
Gnetales	Gnetaceae	*Gnetum montanum*	leaves	Tao Chen	BGI
Gnetales	Welwitschiaceae	*Welwitschia mirabilis*		J. Leebens-Mack	J. Leebens-Mack

From: unpublished data from the New York Plant Genomics Consortium. NSF grant: IOS-0922738 (Downloaded April 2013)

Illumina

Clade	Family	Species	Tissue Type
Conifers	Cupressaceae	*Metasequoia*	young ovules, seeds, young leaves
Ginkgoales	Ginkgoaceae	*Ginkgo biloba*	young ovules, seeds, young leaves
Cycadales	Cycadaceae	*Cycas rumphii*	leaves

From: Dendrome (downloaded April 30 2013)
454 Assemblies
http://dendrome.ucdavis.edu/treegenes/transcriptome/transcr_summary.php
http://loblolly.ucdavis.edu/bipod/ftp/Transcriptome_Data/454/

Clade	Family	Species
Gnetales	Gnetaceae	*Gnetum gnemon*
Conifers	Pinaceae	*Picea abies*
Conifers	Pinaceae	*Pinus lambertiana*
Conifers	Pinaceae	*Pinus taeda*
Conifers	Podocarpaceae	*Podocarpus macrophyllus*
Conifers	Pinaceae	*Pseudotsuga menziesii*
Conifers	Sciadopityaceae	*Sciadopitys verticillata*
Conifers	Cupressaceae	*Sequoia sempervirens*
Conifers	Taxaceae	*Taxus baccata*

From: Ian Boyes / Stefan Little: Douglas fir assembly (downloaded April 2013)

Clade	Family	Species	Tissue Type	Sample Provider	RNA Extractor
Conifers	Pinaceae	*Pseudotsuga menziesii*	megagametophyte, nucellus, cone scale complexes	Ian Boyes	Ian Boyes

Appendix 2. Proteins and peptides detected in *Ceratozamia hildae* by PEAKS 6 DB plus SPIDER. Proteins are identified by accession name and are listed in alphabetical order (accession refers to the unique header given to transcripts in Gymno_DB). For each protein, a -10lgP score > 20 is significant. For individual peptides, a -10lgP score > 45.9 is considered significant. Although PEAKS 6 only uses peptides exceeding the score threshold to determine which proteins are significant, all of the peptides detected by PEAKS6 DB searches are reported in this table.

Accession	-10lgP
Dougfir-megastigmus-comp462464_c0_seq1_21	97.15
G.P(sub S)QESRDWPLIDPLPSYGR.G	82.01
K.IAGYFHGIQ(+.98)K.A	30.27
Dougfir-megastigmus-comp535867_c0_seq3_4	90.46
H.MLKE(sub S)SRFN(+.98)N(+.98)PFNC(+57.02)R.R	50.59
R.FN(+.98)N(+.98)PFNC(+57.02)R.R	40.39
R.FN(+.98)N(+.98)PFNC(+57.02)G(sub R)R.G	39.61
H.M(+15.99)LKE(sub S)SRFN(+.98)N(+.98)PFNC(+57.02)R.R	26.54
K.R(sub S)Q(sub S)RFN(+.98)N(+.98)PFNC(+57.02)R.R	25.88
Dougfir-megastigmus-comp560917_c0_seq1_10	58.25
M.YSQYIIK.L	58.25
Ginkgo_Contig28425_14	140.75
R.GFDVIENAK.K	69.82
R.GFDVIENAKK.Q	68.49
G.FYSS(sub P)TCPDAESVVR.A	63.93
Y.SS(sub P)TC(+57.02)PDAESVVR.A	61.46
Ginkgo_Contig33358_18	97.59
R.SDYSHLEPT.T	67.74
S.EFSWDNK.F	59.69
K.MSYMVGFGSKYPTQLHHR.G	13.79
Ginkgo_Contig35216_33	97.92
R.YWPLPSSPYDDC(+57.02)PSK.C	73.42
K.SLIIDGQR.K	49
K.FASFGTPQGTC(+57.02)GAFHEGTC(+57.02)HGIQSSSTIEK.A	5.03
Ginkgo_Contig47067_103	80.73
T.LIRDLEFN(+.98)HSTSFT.A	71.82
R.FGVLLFTLFALLYC(+57.02)FLC(+57.02)KEC(+57.02)TN(+.98)VPSQ(+.98)LSSHTFR.Y	17.81
Ginkgo_Contig47182_59	105.76
K.D(sub S)MGEDVFWALR.G	75.88
K.ISPL(sub S)ETPFPHR.A	59.75
K.LINQTFPELGMVSQDC(+57.02)EEMSWIESVAYLQR.V	13.33
Ginkgo_Contig55712_69	54.19
R.LFDPPLR.S	54.19
Ginkgo_Contig72727_21	82.01
Q.DLEDTFQPPFK.S	82.01
K.SC(+57.02)VIDGQVASVM(+15.99)C(+57.02)SYNK.V	10.47
Gnetum_gnom_isotig07407_83	59.66
K.YVPGISFR.T	51.33
R.TDNEPFKMAMQSFTK.K	16.67
K.GQ(+.98)AWVN(+.98)GQSIGRYWPANIAQGPC(+57.02)SDTC(+57.02)N(+.98)YK.G	13.53
picea_abies_isotig06000_43	124.58
R.GAGHEVPAFQPSR.A	70.16
R.GAGHEVPS(sub A)FQPA(sub S)R.A	67.52
K.SFLAGKPLP.G	62
Pinus_lamb_isotig19184_50	54.45
E.AFRLDPKK.W	54.45
R.TVAAFATR.V	10.98
Pinus_taeda_contig32485_14	102.67
R.DFQ(+.98)SNVVLGNGK.I	75.51
T.T(sub S)IDRDFQ(+.98)SNVVLGNGK.I	54.33
Pinus_taeda_isotig07390_32	95.34
K.GVFTGECGTELDHGVVAV.G	64.82
K.GVFTGEC(+57.02)GTELDHGVVAV.G	63.13
K.GVFTGECGTELDHQ(sub G)V.V	61.05
R.GTQSGSFMYQNSDNLPESIDWR.E	10.78
G.ECGTELDHGVVAV.G	7.86
K.GVFTGECGTELDHGVVAV.G	64.82
K.GVFTGEC(+57.02)GTELDHGVVAV.G	63.13
K.GVFTGECGTELDHQ(sub G)V.V	61.05

R.GTQSGSFMYQNSDNLPESIDWR.E	10.78
G.ECGTELDHGVVAV.G	7.86
Pinus_taeda_isotig17300_19	73.01
S.EALNDQFR.M	73.01
R.M(+15.99)DLHQEVLQYVLAATRN(+.98)GSDIR.G	10.67
Peptide	-10lgP
Pseudotsuga_menz_isotig03151_34	97.62
K.AWELGNELSSE(sub S)GIVA.S	65.57
K.AWELGNELSE(sub S)SGIVA.S	64.11
K.RWMSNHAVEGPWDSSNAR.D	7.33
scaffold-ACWS-2003997-Arucaria_sp._single_13	57.83
S.WGENGYIR.L	57.83
scaffold-AIGO-2072006-Chamaecyparis_lawsoniana_single_34	46.35
R.FD(sub G)SWMGGDR.D	46.35
MSNLLC(+57.02)DQPK.F	8.08
scaffold-AQFM-2079970-Pseudolarix_amabilis_single_25	104.32
Q.ISWNYNYGPAGR.A	73.33
D.GINN(+.98)PDIVAR(sub N)DPTVSFK.T	61.99
R.ELSAFFGQTSHETTGGWPTAPDGPYAWGYC(+57.02)FK.E	9.26
scaffold-AUDE-2008057-Widdringtonia_cedarbergensis_single_49	136.31
A.E(sub S)FLMPLPAYPIR.E	73.99
N.GLWDPWSGGGVLR.N	72.09
A.G(sub S)FLMPLPAYPIR.E	64.02
K.DWLYEYYQ(+.98)EKGLYSFN(+.98)I	19.77
scaffold-AWQB-2000726-Picea_engelmanii_single_70	196.61
L.TFDNAGMWNLR.S	84.49
K.LVEVEGSHTVQN.V	84.47
K.LKLVEVEGSHTVQ.N	81.79
K.LVEVEGSHTVQ.N	74.22
T.ASAARPNPQGSYHYG.N	68.11
F.DNAGMWNLR.S	62.68
K.LAN(+.98)SAPFIN(+.98)GKQ(+.98)R.Y	11
scaffold-AWQB-2000846-Picea_engelmanii_sub1_26	105.25
R.STFVGSGSYAAH.W	67.9
K.ATWEDLR.Y	58.13
K.HDGVPFM(+15.99)GQVWPGAVYFPDFLNPK.T	24.86
scaffold-BBDD-2075959-Microstrobos_fitzgeraldii_single_30	65.45
L.SNGPS(sub Q)YEVPTGR.R	55.9
R.GSELQFDNNFLQ(+.98)N(+.98)VR.E	19.12
scaffold-CDFR-2059046-Manoao_colensoi_single_6	62.5
C.SFDQNGR.G	50.32
V.DINSQ(sub I)C(+57.02)PAQ(sub E)LK.V	24.36
V.DINSA(sub I)C(+57.02)PAELK.V	9.77
R.LGC(+57.02)SFDQNGR.G	5.18
scaffold-CGDN-2000025-Tetraclinis_sp._single_65	52.85
R.TVTNVSPN(+.98)S(sub G)ESTYK.V	52.85
K.GILVVC(+57.02)SAGNNGPDSR.S	5.81
scaffold-DSXO-2059923-Cryptomeria_japonica_single_3	83.82
L.FGDAADLR.G	57.56
M.LFGDAADLR.G	52.51
scaffold-EFMS-2081589-Torreya_taxifolia_single_25	133.35
Q.DLEDTYQPPFK.S	94.67
I.HYTP(sub K)TPEDAA(sub V)AVALK.A	58.03
K.LGFC(+57.02)N(+.98)TSMPFN(+.98)AR.A	28.99
scaffold-EGLZ-2010263-Prumnopitys_andina_single_25	73.47
R.GQETPGEDPLVASK.Y	64.2
K.LN(+.98)ESDIN(+.98)RALYN(+.98)LFSVR.M	18.55
scaffold-ESYX-2016725-Cunninghamia_lanceolata-branch_apex_with_needles_single_67	97.7
K.S(sub K)TDTYGPDIPR.L	68.63
K.ATWEDLR.R	58.13
K.DFTLDPINYPEDK.L	9.74
scaffold-ESYX-2074066-Cunninghamia_lanceolata-branch_apex_with_needles_single_8	62.53
M.SYGVNFGK.S	62.53
scaffold-ETCJ-2059606-Pilgerodendron_uviferum_single_11	120.59
K.GMGEDVFWAIR.G	82.65
K.D(sub G)MGEDVFWAIR.G	75.88
scaffold-FMWZ-2002490-Dacrycarpus_compactus_single_28	50.9
R.LVEDIDNIDIL(sub V)FH.I	50.9
K.SSDGKVYDQ(+.98)FWITR.E	13.93
scaffold-FMWZ-2012745-Dacrycarpus_compactus_single_21	126.05
L.GYDNGIFAPAR.C	84.52
T.AGNSATEPYIVAH.N	83.05

K.ATDYQMDWNAGFAYAR.N	9.14
scaffold-FMWZ-2049673-Dacrycarpus_compactus_single_5	56.05
R.DFPDNPPVR.F	56.05
scaffold-GAMH-2006106-Tsuga_heterophylla_single_17	90.38
K.TFVVGDKL.V	70.24
K.TFVVGDKLL(sub V)F.T	25.59
MGVNYTDWASGKTFVVGDK.L	22.03
scaffold-GAMH-2009786-Tsuga_heterophylla_single_29	120.41
K.AGMDIN(+.98)CGTYLLR.N	85.1
L.WVGYPGEAGGQALAEIIFGDYNPGGR.L	70.62
R.LFTVMSWQ(+.98)MR.A	5.52
scaffold-GGEA-2007079-Cedrus_libani_single_16	57.11
L.TQNIIKPDISAPGVN.I	57.11
K.SSSASIASIC(+57.02)SLDSLDPSKAEGK.V	11.15
scaffold-GJTI-2002248-Cephalotaxus_harringtonia_single_8	75.33
R.C(+57.02)NC(+57.02)VPPGTYGNK.E	75.33
R.AC(+57.02)GTC(+57.02)C(+57.02)ARC(+57.02)N(+.98)C(+57.02)VPPGTYGN(+.98)K.E	7.27
R.AC(+57.02)GTC(+57.02)C(+57.02)ARC(+57.02)N(+.98)C(+57.02)VPPGTYGNK.E	5.42
scaffold-GJTI-2006440-Cephalotaxus_harringtonia_single_31	183.22
K.VYMAG(sub S)PSPWFFK.N	81.96
K.GNYQGPGLWIK.K	80.94
Y.MAG(sub S)PSPWFFK.N	70.75
M.AG(sub S)PSPWFFK.N	63.89
A.SPSPWFFK.N	60.3
T.DGSLS(sub E)GVQGTLDKFK.D	55.08
scaffold-GKCZ-2007555-Diselma_archeri_sub2_8	61.53
T.VFDVSFNK.L	61.53
K.IPDSIC(+57.02)ALPNLENFTYSYNFFSGEPPNC(+57.02)LALPSR.G	6.5
scaffold-GKCZ-2007628-Diselma_archeri_single_10	76.47
L.IYSGDVDAVVPVTS(sub A)TR.Y	76.47
MFPIYR.R	10.86
scaffold-GNQG-2004817-Encephalartos_barteri_single_28	64.85
T.ALQ(+.98)NDFMK.D	64.85
K.N(+.98)SREN(+.98)ETDFEM(+15.99)DK.F	6.03
scaffold-GNQG-2006873-Encephalartos_barteri_single_24	194.36
R.SGTSDGSM(sub E)Q(sub E)WGYIEVRP.K	77.89
R.LEAADEDNA(sub I)NSIALK.I	73.72
R.SGTSDGSEEWGYIEVRP.K	73.68
R.SGTP(sub S)DGSEQ(sub E)WGYIEVRP.K	66.64
R.LEAADEDG(sub N)INSIAQ(sub L)K.I	65.76
G.NFQ(sub E)EVGPV(sub L)DTQLN(sub K)PR.Q	63.63
A.QTDAEVL(sub T)QDLLR(sub S).F	54.94
E.AADEDG(sub N)INSIAQ(sub L)K.I	27.52
V.AQTDAEVL(sub T)QDLLR(sub S).F	16.16
MQQ(+.98)SNM(+15.99)KR.S	15.43
scaffold-GNQG-2008016-Encephalartos_barteri_single_19	236.46
R.IENC(+57.02)YVSTGDDVVAIK.S	89.39
L.HSPNTDGIDPDSSSYVR.I	87.28
D.GIDPDSSSYVR.I	84.75
K.SGWDEYGIAYGKPSEH.I	83.02
L.GSQDLEDWPLIPPLPSYGR.G	79.01
G.SQDLEDWPLIPPLPSYGR.G	75.79
C.YVSTGDDVVAIK.S	74.24
Q.DLEDWPLIPPLPSYGR.G	73.39
T.GFYGSHPDDKYNP.N	55.9
R.DIHVN(+.98)GM(+15.99)TM(+15.99)VNMK.W	9.75
scaffold-GNQG-2010926-Encephalartos_barteri_single_38	117.64
D.ENEPFLS(sub G)NGDLPVLIVESK.G	93.14
K.SLIIDGQR.K	49
scaffold-GNQG-2011552-Encephalartos_barteri_single_14	222.46
G.YAQTSEDAVADVLK.A	103.72
Q.DLEDTYQPPFK.S	94.67
R.GFQGEDN(sub T)QMEPSR.L	84.3
R.GQETPGEDPLVSSK.Y	81.69
T.SEDAVADVLK.A	74.59
S.NLFA(sub S)VQMRLGLFDGDPK.H	47.75
scaffold-GNQG-2013496-Encephalartos_barteri_single_52	254.31
K.NSWQDGVFGTTCPIPPGR.N	129.98
K.NSWQDGVFGTTC(+57.02)PIPPGR.N	101.35
K.LKLVEVEGSHTVQ.S	81.79
R.NFVEIVFQNNER.T	79.75
Q.APPTGFEWSY(sub L)NQAR.S	72.4

T.ASAARPNPQGSYHYG.M	68.11
N.SRPMIPVPFPTPDGDFSIL.I	63.15
F.VVGMG(sub E)AGQWTPASR.K	56.41
S.F(sub Y)VRPDTPLK.L	36.25
K.AAGGFGAIKV(sub I).N	18.08
scaffold-GNQG-2013749-Encephalartos_barteri_single_15	143.3
F.MFTDPSLGGTTVR.Q	70.94
R.WIVDE(sub G)N(+.98)GNRVK.L	64.72
R.GAFVITPNR.T	56.54
R.WIVDGNGE(sub N)R.V	54.19
R.LLDLPLVEA(sub V)YK.E	38.02
R.GAFVITPN(+.98)R.T	11.68
scaffold-GNQG-2014022-Encephalartos_barteri_sub2_17	232.38
C.EWGDDNPALWAPQVGNSWR.T	91.67
D.DC(+57.02)WGEQNRDSQGNLVAR.S	84.41
N.IDDC(+57.02)WGEQNRDSQGNLVAR.S	83.51
K.YINIDDC(+57.02)WGEQNR.D	78.76
K.TMPGSLGHED(sub E)QDAR(sub K).T	67.64
N.IDDC(+57.02)WGEQNR.D	64.35
N.DVGLSPGAVVR.A	62.63
K.YINIDDC(+57.02)WGEQNRDSQGNLVAR.S	59.6
K.IDENIIR.R	40.65
K.IDENIIRR.T	35.78
N.GLGVRPPMG.W	24.65
scaffold-GNQG-2014884-Encephalartos_barteri_sub2_26	128.36
Q.DLEDTYQPPFK.S	94.67
R.DRIDLLLPGQQQK.L	67.36
MDNWEGVDRFHFDAR.V	7.02
scaffold-GNQG-2015735-Encephalartos_barteri_single_28	198.03
F.EEIGGDPTQISFTTR.S	86.39
K.LCEGALISSDS(sub A)K.H	77.67
A.QIENEYGNIDSAYGAAAK.S	71.03
E.M(+15.99)WADLIQK.S	68.05
E.NEYGNIDSAYGAAAK.S	66.36
K.GEELGFYR.G	65.92
R.L(sub V)PSRPAEDLAFAVAR.F	55.06
scaffold-GNQG-2016467-Encephalartos_barteri_single_82	46.11
K.DGIPIC(+57.02)R.E	46.11
scaffold-GNQG-2081815-Encephalartos_barteri_single_4	117.81
K.C(+57.02)DC(+57.02)TPPIPTR.G	79.47
K.C(+57.02)DC(+57.02)TPPV(sub I)PTR.G	65.3
C.DC(+57.02)TPPV(sub I)PTR.G	17.07
scaffold-GNQG-2083101-Encephalartos_barteri_single_20	90.05
K.LPGLC(+57.02)GVS(sub N)VGVPISASTNC(+57.02)AAIH	68.2
R.SIGINFSK.A	43.71
scaffold-GNQG-2084473-Encephalartos_barteri_single_34	183
N.SQSLPGFTPPGSYK.L	95.87
T.NSQSLPGFTPPGSYK.L	88.08
K.SVDVDPNPVVR.G	74.08
Y.S(sub A)GDN(sub K)FNYCDGEGS(sub N)YPVQVK.S	57.02
K.EGSFLLTNSQSLPGFTPPGSYK.L	20.66
scaffold-GNQG-2086382-Encephalartos_barteri_single_12	153.9
R.VSSSGDTITLNTGVVK.S	94.02
K.VGFGPAAPGSK.L	60.06
K.ILN(+.98)GLPDF(sub Y)SSFNR(sub S).Y	56.45
L.A(sub V)VDNAAM(sub V)SALTG(sub S)K.N	30.91
L.DYFDPA(sub S)KLH.D	16.53
S.WGENGYLR.M	57.83
K.HGVNPPNPGPTPPSPVKPPNWC(+57.02)DDYSSC(+57.02)PESNTC(+57.02)C(+57.02)C(+57.02)VFSIGR.D	8.97
scaffold-GNQG-2086713-Encephalartos_barteri_single_19	65.28
M.GYPGQ(+.98)AGGQ(sub D)ALADVIFGR.Y	65.28
K.SC(+57.02)VLDGQVASVM(+15.99)C(+57.02)SYNK.V	10.47
scaffold-HILW-2113257-Acmopyle_pancheri_single_45	133.53
K.YNGGLDTEESYPYTAK.D	102.01
A.GAFN(+.98)NFGC(+57.02)NGGLPSH(sub Q).A	63.05
scaffold-HQOM-2013044-Torreya_nucifera_sub1_11	63.61
S.EWGEDGYIR.M	55.39
K.N(+.98)SWGSEWGEDGYIRMQR.A	16.43
scaffold-IAJW-2026102-Amentotaxus_argotaenia_single_17	51.97
R.SVAFEVVK.V	51.97
scaffold-IFLI-2000858-Callitris_gracilis-March_single_27	131.29
K.LKLVEVEGSHTVQ.N	81.79

K.ALGPPDA(sub G)VLINGR.G	71.62
K.ALGPPDA(sub G)VLIN(+.98)GR.G	58.03
M.IGDWYKR.N	41.08
scaffold-IFLI-2026620-Callitris_gracilis-March_single_1	47.69
MLSGSC(+57.02)DDFLKR.G	47.69
scaffold-IFLI-2135760-Callitris_gracilis-March_single_3	57.79
T.VFN(+.98)VSFNK.L	57.79
M.Q(+.98)GSLPSQFGK.L	14.6
scaffold-IIOL-2000981-Pinus_parviflora_sub2_24	64.81
K.IFEYESEASKIP(sub S)R(sub E).L	54.45
R.DELHLLVEHFNVDVENPC(+57.02)VIMTQ(+.98)DKSR.E	20.71
scaffold-IOVS-2040036-Pseudotsuga_menziesii_single_7	68.73
R.ETGIVSPVKN(+.98)Q	68.73
scaffold-IOVS-2056527-Pseudotsuga_menziesii_single_18	148.24
Q.HVDLVGGYYDAGDNVK.F	89.52
D.LVGGYYDAGDNVK.F	78.46
R.RDSALSDGSSQQ(sub H)VDLVGGYYDAGDNVK.F	58.45
R.LIHTAM(+15.99)RVFDFADK.H	14.28
scaffold-JDQB-2002660-Neocallitropsis_pancheri_sub1_10	107.45
N.S(sub G)VSFVR(sub Q)PDTPLK.L	61.72
V.SFVR(sub Q)PDTPLK.L	58.73
S.FVR(sub Q)PDTPLK.L	49.1
K.LAN(+.98)SAPFIN(+.98)GKQ(+.98)R.Y	11
scaffold-JRNA-2006442-Phyllocladus_hypophyllus_single_62	77.82
L.GSQDSS(sub E)DWPVIA(sub D)PLPSYGR.G	77.82
R.VGIC(+57.02)FTGFYGDHPDDR.Y	13.41