Université de Montréal

The plant ovule omics: an integrative approach for pollen–pistil interactions and pollen tube guidance studies in solanaceous species

par

Yang Liu

Département de sciences biologiques, Institut de recherche en biologie végétale

Faculté des Arts et des Sciences

Thèse présentée à la Faculté des Arts et des Sciences

en vue de l'obtention du grade de doctorat

en biologie végétale

Octobre, 2015

Université de Montréal
Faculté des études supérieures et postdoctorales

Cette thèse intitulée:

The plant ovule omics: an integrative approach for pollen–pistil interactions and pollen tube guidance studies in solanaceous species

Présentée par:
Yang Liu

a été évaluée par un jury composé des personnes suivantes :

David Morse, président-rapporteur
Daniel P. Matton, directeur de recherche
Simon Joly, membre du jury
Alice Cheung, examinatrice externe

Résumé

Chez les plantes à fleurs, l'ovaire est l'organe reproducteur femelle et il interagit de façon importante avec les gamètes mâles durant la croissance, le guidage, la réception et la rupture du tube pollinique ainsi que la fusion des gamètes. Le processus débute lorsque de nombreux gènes de l'ovule sont activés à longue distance lors de la réception du pollen sur le stigmate. Afin d'explorer les signaux provenant de l'ovule ayant un impact important sur les interactions pollen–pistil, particulièrement les molécules sécrétées impliquées dans la signalisation espèce-spécifique, l'expression génique des ovules sous forme d'ARNm ainsi et la sécrétion protéique ont été étudiées chez *Solanum chacoense*, une espèce diploïde de pomme de terre sauvage. *S. chacoense* a subi beaucoup d'hybridation interspécifique avec d'autres espèces sympatriques de solanacées, facilitant ainsi grandement l'étude des interactions pollen–ovule de façon espèce-spécifique ainsi que leur évolution. Dans ce projet, des ovules provenant de trois conditions différentes ont été comparés: des ovules matures de type sauvage, des ovules légèrement immatures, récoltés deux jours avant l'anthèse et des ovules provenant du mutant *frk1* pour lesquels le sac embryonnaire est absent. Un séquençage d'ARN à haut débit a d'abord été effectué sur les ovules de type sauvage de *S. chacoense* afin de générer un assemblage de référence comprenant 33852 séquences codantes. D'autres séquençages ont été effectués sur les trois conditions d'ovules et sur les feuilles afin de faire une analyse d'expression différentielle des gènes. En comparaison avec les ovules de type sauvage, 818 gènes sont réprimés dans les ovules du mutant *frk1*. Un sous-groupe de 284 gènes, étaient également sous-exprimés dans les ovules légèrement immatures, suggérant un rôle spécifique dans les stades tardifs de la maturation du sac embryonnaire (stade de développent FG6 à FG7) ainsi que du guidage du tube pollinique, puisque ni les ovules du mutant *frk1* ni ceux légèrement immatures ne sont capables d'attirer les tubes polliniques lors d'essais de croissance semi *in vivo*. De plus, 21% de ces gènes sont des peptides riches en cystéines (CRPs). En utilisant un transcriptome assemblé *de novo* provenant de deux proches parents de *S. chacoense*, *S. gandarillasii* et *S. tarijense*, une analyse d'orthologie a été effectuée sur ces CRPs, révélant une grande variabilité et une évolution rapide chez les solanacées. De nouveaux motifs de cystéine uniques à cette famille ont également été découverts. En comparant avec des études similaires chez *Arabidopsis*, le sac embryonnaire de *S. chacoense* montre un transcriptome fortement divergent, particulièrement en en ce qui a trait à la catégorisation fonctionnelle des gènes et de la similarité entre les gènes orthologues. De plus,

même si la glycosylation n'est pas requise lors du guidage mycropylaire du tube pollinique chez *Arabidopsis*, *Torenia* ou le maïs, des extraits d'ovules glycosylés de *S. chacoense* sont capables d'augmenter la capacité de guidage de 18%. Cette étude est donc la première à montrer une corrélation entre glycosylation et le guidage du tube pollinique par l'ovule. En complément à l'approche transcriptomique, une approche protéomique portant sur les protéine sécrétées par l'ovule (le secrétome) a été utilisée afin d'identifier des protéines impliquées dans l'interaction entre ovule et tube pollinique. Des exsudats d'ovules matures (capables d'attirer le tube pollinique) et d'ovules immatures (incapables d'attirer le tube pollinique) ont été récoltés en utilisant une nouvelle méthode d'extraction par gravité permettant de réduire efficacement les contaminants cytosoliques à moins de 1% de l'échantillon. Un total de 305 protéines sécrétées par les ovules (OSPs) ont été identifiées par spectrométrie de masse, parmi lesquelles 58% étaient spécifiques aux ovules lorsque comparées avec des données de protéines sécrétées par des tissus végétatifs. De plus, la sécrétion de 128 OSPs est augmentée dans les ovules matures par rapport aux ovules immatures. Ces 128 protéines sont donc considérées en tant que candidates potentiellement impliquées dans la maturation tardive de l'ovule et dans le guidage du tube pollinique. Cette étude a également montré que la maturation du sac embryonnaire du stade FG6 au stade FG7 influence le niveau de sécrétion de 44% du sécrétome total de l'ovule. De façon surprenante, la grande majorité (83%) de ces protéines n'est pas régulée au niveau de l'ARN, soulignant ainsi l'importance de cette approche dans l'étude du guidage du tube pollinique comme complément essentiel aux études transcriptomiques. Parmi tous les signaux sécrétés par l'ovule et reliés au guidage, obtenus à partir des approches transcriptomiques et protéomiques décrites ci-haut, nous avons spécifiquement évalué l'implication des CRPs dans le guidage du tube pollinique par l'ovule chez *S. chacoense,* vu l'implication de ce type de protéine dans les interactions pollen-pistil et le guidage du tube pollinique chez d'autres espèces. Au total, 28 CRPs étaient présentes dans les ovules capables d'attirer le tube pollinique tout en étant absentes dans les ovules incapables de l'attirer, et ce, soit au niveau de l'ARNm et/ou au niveau du sécrétome. De celles-ci, 17 CRPs ont été exprimées dans un système bactérien et purifiées en quantité suffisante pour tester le guidage. Alors que des exsudats d'ovules ont été utilisés avec succès pour attirer par chimiotactisme le tube pollinique, les candidats exprimés dans les bactéries n'ont quant à eux pas été capables d'attirer les tubes polliniques. Comme l'utilisation de systèmes d'expression hétérologue eucaryote peut permettre un meilleur repliement et une plus grande activité des protéines, les candidats restants seront de nouveau exprimés, cette fois

dans un système de levure ainsi que dans un système végétal pour produire les peptides sécrétés. Ceux-ci seront ensuite utilisés lors d'essais fonctionnels pour évaluer leur capacité à guider les tubes polliniques et ainsi isoler les attractants chimiques responsable du guidage du tube pollinique chez les solanacées comme *S. chacoense.*

Mots-clés: Interactions pollen–pistil, ovule, sac embryonnaire, maturation, guidage du tube pollinique, chimio-attractants, peptide riche en cystéine, RNA-seq, spectrométrie de masse, sécrétome, analyse d'expression différentielle

Abstract

In flowering plants, the ovary is the female reproductive organ that interacts extensively with the male gametophyte during pollen tube (PT) growth, guidance, reception, discharge and gamete fusion. The process begins when numerous ovule-expressed genes are activated when pollen lands on the stigma. To explore the ovular signals that have a great impact on successful pollen–pistil interactions, especially the secreted molecules that mediate species-specific signaling events, ovule mRNA expression and protein secretion profiles were studied in *Solanum chacoense*, a wild diploid potato species. *Solanum chacoense* has undergone extensive interspecific hybridization with sympatric solanaceous species that greatly facilitates the study of species-specific pollen–ovule interactions and evolution. In this project, three ovule conditions were studied: wild-type mature ovules, slightly immature ovules at two days before anthesis (2DBA), and *frk1* mutant ovules that lack an embryo sac (ES). RNA-seq was performed on *S. chacoense* ovules to provide a scaffold assembly comprising 33852 CDS-containing sequences, then to provide read counts for differential gene expression analyses on three ovule conditions as well as on leaf. Compared to wild-type ovules, 818 genes were downregulated in *frk1* ovules. A subset of 284 genes was concurrently under-expressed in 2DBA ovules, suggestive of their specific involvement in late stages of ES maturation (female gametophyte (FG), FG6 to FG7 developmental stage), as well as in PT guidance processes, as neither *frk1* nor 2DBA ovules attract semi *in vivo*-grown PTs. Of these 284, 21% encoded cysteine-rich peptides (CRPs). Using *de novo* assembled ovule transcriptomes of two close relatives, *S. gandarillasii* and *S. tarijense*, an orthology survey was conducted on these CRPs, revealing their highly polymorphic nature among species and rapid evolution. Interestingly, novel cysteine motifs unique to this family were also uncovered. As compared to parallel studies in *Arabidopsis*, *S. chacoense* was found to possess a highly divergent ES transcriptome, in terms of both functional categories and individual ortholog similarities. Although glycosylation is not required for micropylar guidance cues to attract PTs in *Arabidopsis*, *Torenia* or maize, glycosylated ovule extracts from *S. chacoense* showed enhanced PT guidance competency by 18%. This is the first time a positive regulation between glycosylation and ovular PT guidance has been observed. As a complement to the transcriptomic approach, a proteomic approach using secreted proteins from the ovule (secretome) was employed to identify proteins involved in pollen–pistil interactions. Ovule exudates were collected from mature ovules (PT attracting) and immature ovules at 2DBA (PT

nonattracting), using a novel tissue free-gravity extraction method (tf-GEM), which efficiently reduced the cytosolic contamination to less than 1%. Through mass spectrometry analyses, a total of 305 ovule-secreted proteins (OSPs) were identified, of which 58% were considered ovule-specific when compared to secretome studies conducted in other plant tissues. The secretion of 128 OSPs was upregulated in mature ovules vs. immature ovules. These OSPs were considered as candidate proteins involved in late ovule maturation and PT guidance. This study demonstrated that the ES maturation from FG6 to FG7 stages influenced the secretion status of 44% of ovule secretome. Surprisingly, the majority (83%) of these proteins were not regulated at the RNA level, vindicating this novel approach in the study of PT guidance as a robust complement to transcriptomic studies. Among all identified guidance-related ovular signals from the transcriptomic and proteomic approaches described above, we focused on the evaluation of the involvement of CRPs in ovular PT guidance of *S. chacoense*, due to the implication of various CRPs in pollen–pistil interactions and, especially, in PT guidance. A total of 28 CRPs were present in PT attracting ovules while being low or absent in nonattracting ovules, at the mRNA and/or protein secretion levels. Of these, 17 CRPs were expressed in bacteria and purified in sufficient amount for PT guidance assays. However, while ovule exudates were shown to induce PT chemotropism in the bead assay, refolded candidates did not show guidance competency. Since the use of eukaryotic protein expression systems might lead to better refolding and higher protein activity, the remaining candidates will be expressed in both yeast and plant-based expression systems and tested for their ability to attract PTs in a semi *in-vivo* assay, in order to lead us toward the isolation of PT guidance chemoattractants in solanaceous species like *S. chacoense*.

Keywords: pollen–pistil interactions, ovule, embryo sac, maturation, pollen tube guidance, chemoattractants, cysteine-rich peptide, RNA-seq, mass spectrometry, secretome, differential gene expression analysis

Table of Contents

List of Tables

List of Figures

List of Abbreviations

2DBA: two days before anthesis

AGL: agamous-like

AGP: arabinogalactan protein

BCIP: 5-bromo-4-chloro-3-indolyl-phosphate

BME: β-mercaptoethanol

CCG: central cell guidance

CDS: coding sequence

CNGC: cyclic nucleotide-gated ion channel

COA: coatlicue

CPM: counts per million

CRP: cysteine-rich peptide

cTPI: cytosolic triosephosphate isomerase

CV: coefficient of variation

DEFL: defensin-like

DIF1: determinant infertile1

DUF784: domain of unknown function 784

EC1: egg cell 1

ECA1: early culture abundant 1

EDTA: ethylenediaminetetraacetic acid

ES: embryo sac

ESF1: embryo surrounding factor 1

EST: expressed sequence tag

FDR: false discovery rate

FG: female gametophyte

FPLC: fast protein liquid chromatography

FRK1: fertilization related kinase 1

FW: fresh weight

GABA-T: γ-amino butyric acid transaminase

Gnd·HCl: guanidine hydrochloride

GO: gene ontology

GSP: generally secreted protein

HAP: hours after pollination

HRP: horseradish peroxidase

HSP: heat shock protein

IFR: isoflavone reductase

IPTG: isopropyl β-D-1-thiogalactopyranoside

KS: Kolmogorov-Smirnov

LC: liquid chromatography

LTP: lipid transfer protein

MAPKKK/MEKK: mitogen-activated protein kinase kinase kinase

MMC: megaspore mother cell

MS: mass spectrometry

NBT: nitro-blue tetrazolium

ORF: open reading frame

OSP: ovule-secreted protein; nOSP: non-specific OSP; oOSP: ovule-specific OSP

PELP III: class III pistil-specific extensin-like protein

PlantSecKB: plant secretome and subcellular proteome knowledgeBase

PM: Pollen Mitosis

PR: pathogenesis-related

PSK: phytosulfokine

PT: pollen tube

RALF: rapid alkalinization factor

RNase: ribonuclease

ROS: reactive oxygen species

SCA: stigma/stylar cysteine-rich adhesion

SCR: S-locus cysteine-rich ligand

SEM: scanning electron microscopy

SEP: stigma exudate protein

siRNA: small interfering RNA

SP: signal peptide

SPL/NZZ: sporocyteless/nozzle

SRK: S-locus receptor kinase

tf-GEM: tissue-free gravity-extraction method

TPST: tyrosylprotein transferase

TTS: transmitting-tissue specific

UTR: untranslated regions

VIC: vacuum-infiltration centrifugation

Zm: zea mays

Acknowledgements

First and foremost I would like to thank my dearest friends in IRBV to make me feel “one of us” and to feel that I belong here!

Heartfelt thanks to my Ph. D. supervisor, Daniel, who encouraged and supported me to go on trainings, meetings and exchange programs during my study, who invited Dr. Zahra Agharbaoui to train me for transient expression when I was stuck with that experiment, and who was dedicated to help me through difficulties in life and allowed me to grow up in this lab.

Sincere thanks to David for his positive comments on my very first class in Canada and showed me how to do research. Sincere thanks to dear Mario, who brought me THE magic chocolate to ease my tension during predoc exam!

I would like to thank my fellow Mattoneurs Audrey, Caroline, Fangwen, Valentin, Rachid for all your help in and out of the lab. I enjoy every moment spent with Audrey and Caro, for our chitchat in the cubical! I thank Fangwen, who spared no efforts in helping me through protein purification. I thank Valentin to collaborate with me on the two articles and nicely provided such a helping hand. I appreciate the company of Mathieu, Jonathan and Steve. I thank them for dropping by our lab and sharing the simple joys from every small progress during my PhD. I especially thank Mathieu for reading through my thesis and giving me so many comments! I also thank Jonathan for his corrections on the Chapter IV of my thesis. Equally important, I would like to thank Afsaneh, who took me as her family, and thank you for your precious coffee time!

I also want to acknowledge all who helped me with my experiments, including Minako, who corrected my course application, Firas and Youssef, who gave suggestion for the experiments on pollen tubes, and Louise, who helped me so cheerfully with SEM imaging.

I would like to thank China Scholarship Council to have this excellent program to support me this far to Canada, with the reference 2007220003.

I would like to, finally, thank my beloved parents for tolerating my absence during my study and thank my husband for accompanying me here so I’m not alone!

1. Introduction

1.1 Plant Reproduction

Plants reproduce both asexually and sexually. In asexual reproduction mode, offspring arise from a single organism, without the fusion of gametes. Plants reproduced in this way are genetically identical to their "mother". One type of asexual reproduction is vegetative reproduction, where offspring develop from a part of the parental tissues. Such case can be seen in tulips, where tulip bulbs are re-used year after year; or in commercial strawberries, which are reproduced by stolons or "clippings" from specific parts of the plants, which develop roots independently and create a new individual. Another type of asexual reproduction involves apomictic seed formation (Koltunow 1993), without the combination of genetic materials from both parents, as seen in dandelion reproduction. Asexual reproduction offers great advantages for farmers to grow plants of superior economical features rapidly. However, pathogens can also be introduced in asexually reproducing plants due to their lower genotypic diversity and consequently a weaker potential for rapid adaptation in case of environmental changes or pathogen attacks.

In contrast, sexual reproduction requires the fusion of the male and female gametes to generate offspring genetically different from both parents, an important feature for species adaptation to new environment in plant evolution. In sexually reproducing angiosperms, the life cycles alternate between a diploid sporophyte generation and a haploid gametophyte generation, which is realized by double fertilization.

Our laboratory studies plant sexual reproduction, using a solanaceous model species, the wild diploid potato species *Solanum chacoense*. It is closely related to agronomically important species such as *Solanum tuberosum* (potato), *Solanum lycopersicum* (tomato), *Solanum capsicum* (pepper), *Solanum melongena* (eggplant) and *Nicotiana tabacum* (tobacco). To date, the potato, tomato, pepper and *Solanum pennellii* (wild tomato) genomes have been reported (Potato Genome Consortium 2011, Tomato Genome Consortium 2012, Kim, Park et al. 2014, Bolger, Scossa et al. 2014). Expressed sequence tag (EST) library from various other solanaceous species are also available from the Sol Genomics Network (Fernandez-Pozo, Menda et al. 2015) and the Potato Genomics

Resource Database known as the Spud DB (Hirsch, Hamilton et al. 2014). These databases facilitate the search of *S. chacoense* orthologs across several different species and thus make it possible to study species-specific interactions in reproductive processes.

This introduction will take us on the courtship journey of the pollen and the pistil, with a special focus on the topic of pollen tube (PT) guidance, which involves different maternal tissues of the pistil in order to elicit chemotropic behavior of the PTs. Because of the stage-specific guidance cues along the PT path, the PTs can: 1) travel down the stigma, 2) navigate through the intercellular space of the transmitting tract, 3) emerge from the transmitting cells, 4) migrate up the funiculus and 5) target the micropyle of the ovule, with great precision (Palanivelu and Tsukamoto 2012).

At the end, PTs deliver the two sperm cells to the embryo sac (ES) to accomplish double fertilization, where one sperm cell fuses with the egg cell to form a zygote that develops into an embryo and the other sperm cell fuses with the central cell to form a triploid nucleus that will later develops into the endosperm (Hamamura, Nagahara et al. 2012). As the essence of this key process in flowering plants, the development of both female and male gametophyte will be described first.

1.1.1 Female Gametophyte Development

An ovule starts as a finger-like protrusion (nucellus) from the placenta during stage 9 of floral development (Smyth, Bowman et al. 1990). The hypodermal cell at the tip of the nucellus differentiates into a megaspore mother cell (MMC). A small RNA pathway ensures the initiation of a unique MMC at this stage through ARGONAUTE 9 (AGO 9). AGO protein is known to bind to microRNAs or small-interfering RNAs (siRNAs) and direct the cleavage of endogenous mRNA (Baumberger and Baulcombe 2005). It is expressed in the companion somatic cells of the female gametes, which restricts the specification of additional MMC in a non-cell-autonomous manner (Olmedo-Monfil, Duran-Figueroa et al. 2010).

As shown in Figure 1.1, the MMC undergoes meiosis to give rise to four haploid megaspores, where only one survives as the functional megaspore and develops into the

female gametophyte (FG). Here the functional megaspore undergoes three rounds of mitosis, resulting in a syncytial ES containing eight nuclei without cellularisation. This stage is followed by nuclei positioning and cell specification, in which the phytohormone auxin is believed to play a crucial role (Pagnussat, Alandete-Saez et al. 2009, Lituiev, Krohn et al. 2013).

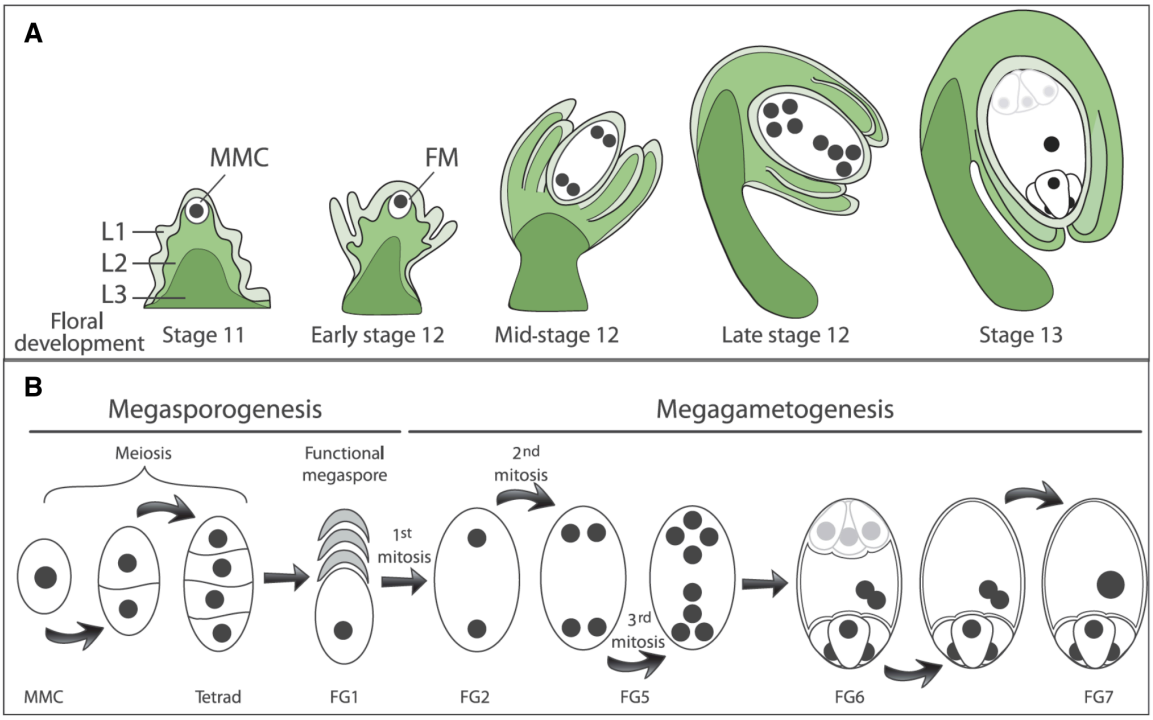


Figure 1.1 Synchronized development of ovule and the FG in *Arabidopsis*. A. The development of an ovule initiates from the protrusion of the nucellus from the placental tissue. The integuments initiate concomitantly with megasporogenesis at stage 11. As megagametogenesis progresses, the integuments elongate and continue to enclose the developing FG during mid- and late-stage 12. At stage 13, the flower is at anthesis stage. The integuments almost completely enclose the mature ES, leaving only a trail for PT entry, the micropyle. B. FG development. During megasporogenesis, a MMC undergoes two successive meiotic divisions to form a tetrad, of which only one megaspore persists as the functional megaspore (FG1). Then it undergoes three rounds of mitosis to form syncytial ES containing eight-nuclei (FG5), followed by nucleus positioning and cell differentiation (FG6). At maturity (FG7), the ES contains two synergids, one egg cell and one central cell, whereas three antipodal cells located at the chalazal end of the ES have undergone degeneration (adapted from (Chevalier, Loubert-Hudon et al. 2011)).

Recently, the role of auxin in ES cell positioning and differentiation has been highly debated. While the Sundaresan lab (Davis, California) observed an auxin gradient along the micropyle-chalaza axis in the ES of *Arabidopsis*, with the highest concentration at the

micropylar end to specify the cell fate of each gamete (Pagnussat, Alandete-Saez et al. 2009), the Grossniklaus lab (Zurich, Switzerland) has suggested otherwise (Lituiev, Krohn et al. 2013). Their mathematical modeling indicates only shallow auxin gradients can be maintained in the FG, which is not sufficient to establish proper auxin patterning. This model is supported by microscopic observation of auxin activity in confined area of the nucellar tissues adjacent to the ES, with the highest concentration close to the micropylar end. This observation was made in both maize and *Arabidopsis*, using the same auxin-sensitive reporter *DR5* (Ottenschlager, Wolff et al. 2003), and in addition, a novel degron-based reporter system, which exploits the auxin-dependent degeneration of Aux/IAA proteins by monitoring degron-GFP levels in the ovule (Lituiev, Krohn et al. 2013). Despite the discrepancies of the localization of an auxin gradient, these two groups consistently suggest an important role for auxin in cell fate specification, either through an intrinsic auxin gradient in the ES (Pagnussat, Alandete-Saez et al. 2009) or via an exterior polarized auxin activity from micropyle (highest) to chalaza (lowest) in the nucellar tissues in a non-cell-autonomous way (Lituiev, Krohn et al. 2013).

Following cell fate specification induced by an increasing auxin concentration along the micropyle-chalaza axis of the ovule, cellularised gametes acquire their own identity to form two synergid cells guarding the entrance of the ES, an egg cell, two polar nuclei that are about to fuse into a central cell and three antipodal cells at the chalazal end.

In *S. chacoense*, antipodal cells degenerate at maturity, while they may persist during fertilization and early endosperm development in *Arabidopsis* as recently observed (Song, Yuan et al. 2014), or alternatively proliferate into a cluster of cells in maize or other grasses. A mature FG consists of only three cell types in *S. chacoense*, two synergids, an egg and a central cell displayed along the micropyle-chalaza axis (Yang, Shi et al. 2010, Drews and Koltunow 2011).

The synergids are largely involved in guiding PTs toward the micropyle, as demonstrated by the identification of synergid-expressed transcription factor *MYB98* involved in filiform apparatus formation and PT guidance in *Arabidopsis* (Kasahara, Portereiko et al. 2005), and by the identification of synergid-expressed LURE-type peptides as micropylar

chemoattractants in *Torenia fournieri* (Okuda, Tsutsui et al. 2009, Kanaoka, Kawano et al. 2011) and *Arabidopsis* (Takeuchi and Higashiyama 2012).

The egg cell is flanked by two synergid cells and is located at the end of the micropyle in the ES. It receives one sperm cell to form the embryo during fertilization. The second fertilization is conducted by the central cell, which receives another sperm cell, leading to endosperm development. However, the function of the central cell is not restricted to fertilization and initiation of endosperm alone (Liu, Yan et al. 2010). Anatomically, the central cell occupies most of the volume of the ES. It can communicate with adjacent cells (synergids, egg, and antipodals), as facilitated by the discontinuous cell wall structure at the junction of the central cell and other cell types. This makes the central cell a signaling hub of the ES, as evidenced by its ability to influence PT attraction, probably via interaction with the synergids as seen in micropylar guidance mutant *ccg* (central cell guidance) (Chen, Li et al. 2007), and its ability to determine the lifespan of adjacent antipodal cells via central cell-expressed *SYCO* gene in *Arabidopsis* (Kagi, Baumann et al. 2010).

The development and function of the antipodal cells remains unclear. Since auxin signaling is present in the antipodal cells in maize (Lituiev, Krohn et al. 2013), a positive correlation has been recently established between auxin signaling and antipodal proliferation (Chettoor and Evans 2015), one more leap forward toward unveiling the role of the mysterious “accessory cells”.

Concurrent with the FG development, the integuments differentiate from L1 cell layer at the chalazal end of the primordium (Figure 1.1). Different from *S. chacoense*, the *Arabidopsis* ovule has also an outer integument, which arises from the L2/L3 cell layers (Chevalier, Batoux et al. 2005). The integuments grow and engulf the nucellus, thus separating the funiculus (a stalk that connects the ovule to the placental tissues) from the ES. Finally, when the integuments meet up at the other side of the ES, they form an opening (a path), i.e. the micropyle, for PT entry during double fertilization.

1.1.2 Male Gametophyte Development

The male gametophyte (pollen grain) forms in the anther. Male gametophyte development includes two phases, microsporogenesis and microgametogenesis, as reviewed recently (Hafidh, Fíla et al. 2015).

During microsporogenesis, a pollen microspore mother cell undergoes two meiotic divisions to produce haploid microspores in tetrads, separated by callose walls (Figure 1.2). Then distinct microspores are released. As microgametogenesis initiates, a vacuole develops and pushes the microspore nucleus against the cell wall (Yamamoto, Nishimura et al. 2003). The microspore then undergoes Pollen Mitosis I (PMI), producing a larger vegetative nucleus (haploid) and a smaller generative nucleus (haploid). Later the small generative nucleus is engulfed by the vegetative cell membrane to form a "cell-within-a-cell" structure. Finally, the vegetative cell will develop into a tube cell. The generative cell will undergo Pollen Mitosis II (PMII) to give rise to twin sperms. In *Arabidopsis*, PMII takes place before pollen grains leave the anther (Sanders, Bui et al. 1999), while in the case of *S. chacoense*, PMII takes place during PT germination (Boavida, Becker et al. 2005, O'Brien, Gray-Mitsumune et al. 2007).

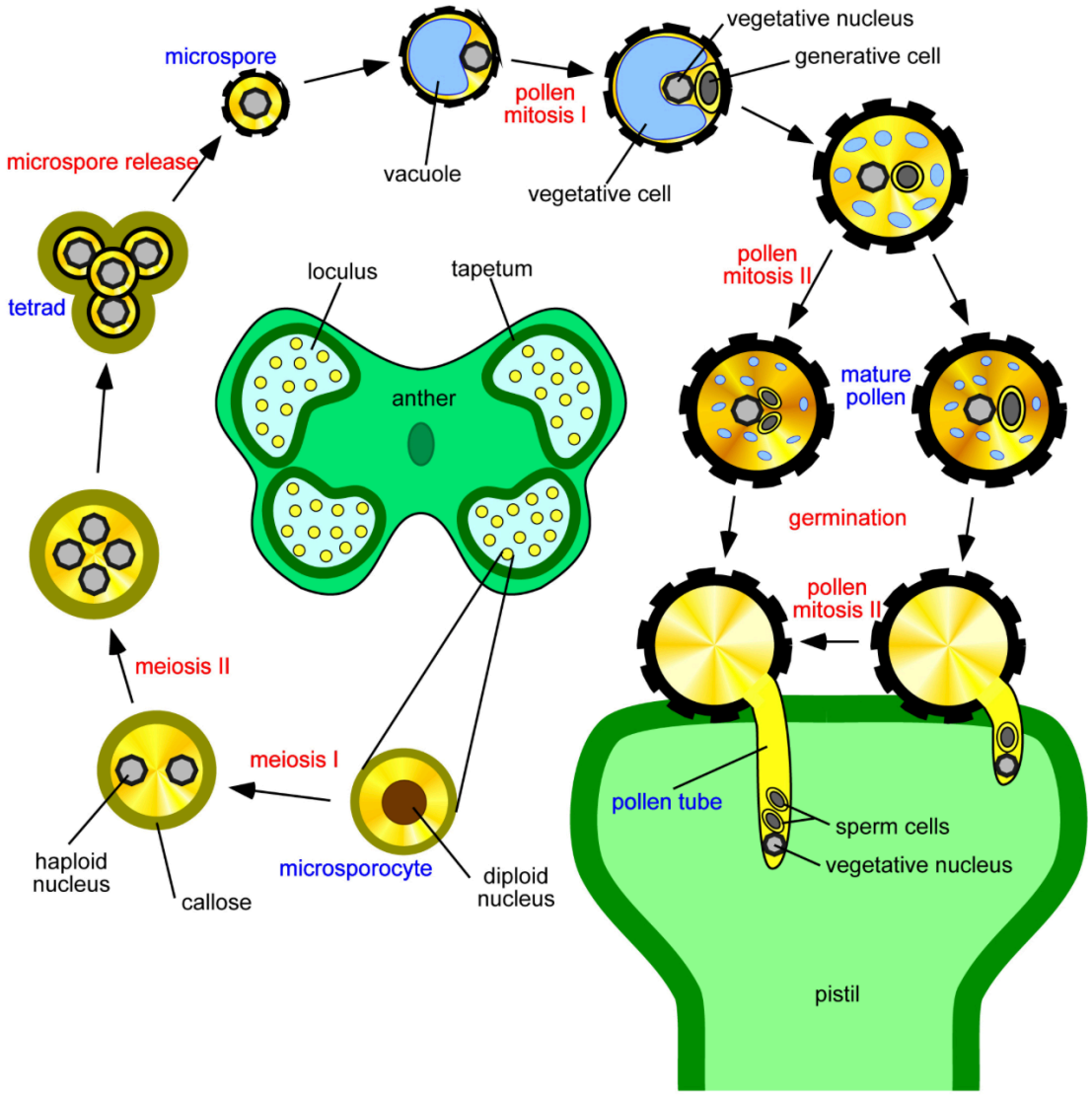


Figure 1.2 Illustration of pollen development in *Arabidopsis*. The pollen of flowering plants is produced within the anther. Pollen development comprises two successive phases, microsporogenesis and microgametogenesis. Microsporogenesis initiates when the diploid microspore mother cell undergoes two meiotic divisions to give rise to four microspores tethered by the callose wall (β-1, 3-glucan). The callose then separates the tetrad into individual microspores. In each microspore, the outer pollen wall exine develops. During microgametogenesis, a vacuole forms that takes a large volume of the cytosol. By pushing the nucleus to a peripheral position, the pollen undergoes asymmetric cell division (mitosis I), producing a larger vegetative nucleus and a smaller generative cell. The former develops into a tube cell while the latter undergoes PMII to produce two sperm cells. In *Arabidopsis*, tricellular pollen is shed, while in *S. chacoense*, the bicellular pollen is shed and PMII takes place during PT germination. This picture is adapted from laboratory of pollen biology website (http://lbp.ueb.cas.cz/field_res.htm).

1.2 The PT Path In the Pistil of *S. chacoense*

In general, the PT path is partitioned into five distinct phases prior to the double fertilization (Figure 1.3) (Swanson, Edlund et al. 2004).

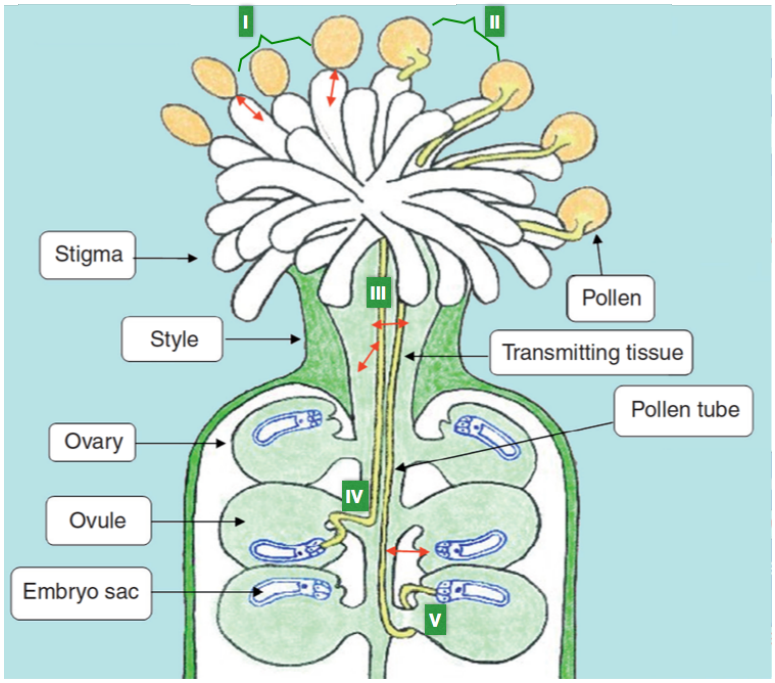


Figure 1.3 The PT pathway in *Arabidopsis*. In phase I, pollen grain adheres to and hydrates on the stigma. In phase II, pollen grain germinates and penetrates the papilla cell wall of the stigma. In phase III, PT enters the transmitting tract of the style. In phase IV, PT exits the transmitting tract and approaches the ovules. In phase V, PT targets the micropyle of the ovule for fertilization (adapted from (Feijó 2010)).

In phase I, *S. chacoense* pollen grains adhere to and hydrate on the wet stigma. Pollen hydration depends on the ability to sense a water gradient at the site of the contact zone between the pollen and the stigma, which is conferred through a lipid matrix derived from either the pollen coat or the stigma, which later determines the polarity of PT germination (Lush, Spurck et al. 2000). The hydrated pollen grain then extends a PT towards the aqueous interface.

In phase II, germinating PTs penetrate the papilla cell wall of the stigma. Once in the solid style of *S. chacoense*, PTs enter phase III, where they navigate through the

intercellular matrix of the transmitting tract. Growing through the style is considered vital for PTs to acquire the competency to respond to guidance cues emitted from the FG to accomplish fertilization (Palanivelu and Preuss 2006).

Okuda and colleagues (Okuda, Suzuki et al. 2013) reported that only when a PT elongates in the style for a sufficient length and for a sufficient time, can it acquire the ability to react to and be attracted to LURE attractants secreted from the synergid cells of the ovule in *Torenia fournieri.* In this experiment, *Torenia* PTs not only have to travel through 15 mm of a cut style (out of a total length of ~18 mm), they also need to grow an additional 7 h on the medium after exiting the cut end of the style to acquire this competency. Consistently, Qin and colleagues (Qin, Leydon et al. 2009) demonstrated that PT penetration of the stigma and style elicits a distinct gene expression profile, as opposed to that in pollen grains or *in vitro* grown PTs. In this comparison, the functional categories that associated with signal transduction, PT growth and transcription were over-represented in semi *in vivo* PTs (grow through the style), which have provided the molecular basis for this competency control.

Interestingly, a parallel can be drawn between PT-transmitting tract cell-interactions and spermatozoa-oviduct interactions in mammals, where in both cases the male gametes need to travel through a female reproductive tract to become capacitated. Studies conducted in human and mice showed that only 10% of the sperm populations are chemotactically responsive to the female reproductive organ (Eisenbach and Giojalas 2006). Moreover, the gene expression profile of the epithelial cells of the oviduct changes significantly upon spermatozoa arrival (Holt and Fazeli 2010), which is quite similar to the observation made in plants during PT elongation in phase III (Qin, Leydon et al. 2009).

In phase IV, the PTs exit the transmitting tract from the ovarian end of the style and start approaching the ovules. In *S. chacoense*, PT path splits along the ovary septa and merges with the placental epithelium, as described previously in the related solanaceous species *Solanum lycopersicum* (Webb and Williams 1988) and *Nicotiana alata* (Cornish, Pettitt

et al. 1987). At the final stage (stage V), PTs target the micropylar opening of each ovule to accomplish double fertilization (Hamamura, Nagahara et al. 2012).

However, between PTs entry to the ES and the onset of double fertilization, PTs go through two sequential stages, i.e. growth arrest and pollen discharge. First, PT growth arrest is achieved via interaction with the filiform apparatus-localized receptor-like kinase, *feronia* (*fer*) in the ovule (Escobar-Restrepo, Huck et al. 2007). Acting in the *fer*-mediated PT reception pathway, two endoplasmic reticulum-localized proteins TURAN and EVAN that are expressed both in pollen grains and ovules, mediate PT reception as their mutants displayed similar PT overgrowth phenotype reminiscent of the *fer* phenotype (Lindner, Kessler et al. 2015). Despite the fact that *TURAN* and *EVAN* were shown to affect pollen development and PT integrity, the *fer*-phenotype is traced back to a defect in the FG since no mutant PTs were formed in reciprocal crosses with wild-type plants (Lindner, Kessler et al. 2015). *TURAN* and *EVAN* encode proteins involved in the *N*-glycosylation pathway, suggestive of a "dual recognition system" in PT reception, where both the protein backbone and the glycosyl residues on the surface of the male and FG are required to be recognized for successful PT reception.

Following PT growth arrest, *fer*-dependent accumulation of reactive oxygen species (ROS) at the entrance of the FG induces PT rupture (Duan, Kita et al. 2014), where PTs discharge their content to fuse with the egg cell and the central cell.

Live cell imaging of double fertilization reveals three distinct sperm cell behavioral steps during PT discharge in *Arabidopsis*. Initially, the PT bursts to deliver two sperm cells to the ES within seconds. Sperm cells remain in a stationary phase for minutes before the final gamete fusion for double fertilization. Interestingly, the two sperm cells are functionally equivalent to fertilize the central cell or the egg cell (Hamamura, Saito et al. 2011).

In maize, PT discharge is mediated by ZmES4 (Zea mays embryo sac 4), a defensin-like (DEFL) CRP derived from the synergid cells, via the opening of potassium channel on the PTs. Fusion of the two female gametes with the sperm cells is found mediated by EC1 (egg cell 1) peptides. At this moment, PTs have finished their journey in the pistil.

Embryogenesis is initiated to switch the life cycle of a plant into the sporophyte generation.

1.3 Why Studying PT Guidance In *S. chacoense*?

1.3.1 The Significance of PT Guidance In Plant Reproduction

In general, PT guidance occurs all along the PT path in the pistil, which allows PTs to travel down the stigma, exit the transmitting tract of the style, and enter the micropyle to achieve double fertilization. These path finding events require intimate cell signaling between the PTs and various female tissues, including stigmatic papilla cells, stylar transmitting tract cells, nucellar cells and finally the female gametes (Palanivelu and Tsukamoto 2012), which makes PT guidance an excellent model to study cell-cell communication, and which will allow different guidance mechanisms to be revealed.

Taking another view, a main goal of plant sexual reproduction, especially for crop plants, is to overcome breeding barriers so that the gene pool of one species can be expanded, allowing biodiversity and "new" species to be raised. Breeding barriers can be derived from a large number of prezygotic and/or postzygotic isolation. By studying gametophytic PT guidance, Márton and coworkers (Marton, Fastner et al. 2012) used genetic engineering to express maize micropylar guidance attractant ZmEA1 in the synergid cells of *Arabidopsis thaliana*. The proper secretion of ZmEA1 peptide enabled *Arabidopsis* ovules to guide maize PTs to enter the micropylar opening of the ovule, which is the first step to overcome the reproductive barrier between unrelated plant families, at the level of gametophytic guidance. The barrier induced by PT guidance represents only a small part of complex, multi-mechanism breeding barriers in plant. Yet, by studying PT guidance, an increased hybrid vigor arising from crossbreeding of much divergent families or species can be anticipated.

1.3.2 PT Guidance In *S. chacoense* and Species-Specific Pollen−Pistil Interactions

PT guidance has been extensively studied in the model species *Arabidopsis* and maize, from which the genome had been annotated. Alternatively, *Torenia* (*Linderniaceae*) was

chosen as a model for its ovule anatomy since its ES protrudes from the ovule, enabling live cell imaging of PT attraction to the synergid cells directly under light microscopy (Higashiyama, Kuroiwa et al. 1998), and the destruction of individual protruding cells by laser cell ablation (Higashiyama, Yabe et al. 2001). This anatomical advantage also allowed cDNA sequencing from individually isolated synergid cells, thus providing the molecular basis for dissecting the complex network of PT–ovule interaction (Okuda, Tsutsui et al. 2009).

In this project, *S. chacoense* was chosen as our model species in the study of pollen−pistil interaction and PT guidance. *Solanum chacoense* belongs to the Solanaceae family, one of the most important families to human beings for its great value as food sources (potato and tomato), drugs (tobacco) and ornamentals (petunia) (Knapp, Bohs et al. 2004). *Solanum chacoense* is native to South America and is one of the most widely distributed wild potato species. It occurs sympatrically with various other *Solanum* species, thus allowing large opportunities for natural hybridization. However, the number of hybrids found so far is still limited, compared to the total number of species in this family (3000-4000 species for the Solanaceae, almost half for genus *Solanum*) (Knapp, Bohs et al. 2004). We thus hypothesize that species-specific attracting molecules may be produced by different solanaceous species to direct PT guidance. Such species-specific molecules could lead to reproductive isolation for sympatrically localized *Solanum* species.

In fact, such species-specific interactions are common in pollen–pistil interactions (Swanson, Edlund et al. 2004). For example, pollen adhesion to the stigma in *Arabidopsis* (dry stigma) depends on the exine wall component sporopollenin, a highly stable mixed polymer containing long-chain fatty acids and phenolics (Thom, Grote et al. 1998). Detergent wash assays demonstrated that *Arabidopsis* pollen barely binds to the stigma of monocot plants, while they exhibit significant binding to some dicot plants from various genera, such as *Arabidopsis*, *Brassica*, *Platanus*, *Ambrosia*, etc. More specifically, *Arabidopsis* pollens bind to an *Arabidopsi*s stigma with greater affinity than to its close relative *Brassica campestris* (Zinkl, Zwiebel et al. 1999), suggestive of a species-specificity in pollen-stigma interaction according to their taxonomic distance. Different functional groups may exist in sporopollenin polymers from different plant taxa,

providing the basis for this selective pollen–pistil interaction (Domínguez, Mercado et al. 1999).

Likewise, studies on the self-incompatibility mechanism during early pollen–pistil interaction revealed that different families have adopted different mechanisms to reject self-pollen. In the Solanaceae, Scrophulariaceae and Rosaceae, the pistil produces S-RNases to degrade self-PT RNA in order to stop self-PT in the style (McClure and Franklin-Tong 2006). In *Papaveraceae*, the S-haplotype specific recognition between stigmatic PrsS (*Papaver rhoeas* stigma S) (Foote, Ride et al. 1994) and the pollen expressed PrpS (*Papaver rhoeas* pollen S) (Wheeler, de Graaf et al. 2009) triggers a Ca^{2+} increase and downstream responses to reject self-pollen. Different from the above-mentioned mechanisms where pollen rejection is determined by the haploid pollen S-haplotype, pollen rejection in *Brassicaceae* is determined by the diploid anther S-haplotype. In the *Brassicaceae*, the male determinant is S-locus cysteine-rich ligand (SCR/SP11) (Takayama, Shiba et al. 2000) and the female determinant is S-locus receptor kinase (SRK) (Stein, Howlett et al. 1991, Goring and Rothstein 1992). The two determinants co-evolve rapidly and are highly polymorphic between species (Boggs, Dwyer et al. 2009).

Along the PT path in the pistil, species-specificity can be found in various other phases, such as in PT guidance, reception and discharge. First, species-specific PT attraction has been demonstrated in *Torenia* (Kanaoka, Kawano et al. 2011), *Arabidopsis* (Takeuchi and Higashiyama 2012) and maize (Marton, Fastner et al. 2012), where the chemoattractants displayed a much compromised attraction efficiency toward PTs of a different species. Second, during PT reception, interspecific crosses using pollens from *Arabidopsis lyrata* and *Cardamine flexuosa* on *Arabidopsis thaliana* stigmas resulted in a PT overgrowth phenotype, emphasizing the importance of species-specific recognition in PT reception (Escobar-Restrepo, Huck et al. 2007). Third, in PT discharge, ZmES4, a defensin like (DEFL) protein, mediates PT burst in a species-specific manner in maize (Amien, Kliwer et al. 2010).

So far, several guidance cues have been uncovered from both the sporophytic (stigma, style, ovule integuments) and the gametophytic tissues of the pistil to guide directional PT growth in *Arabidopsis*, lily, maize and *Torenia*. However, PT guidance in solanaceous species is largely unknown. By using *S. chacoense* as our model species, and taking advantage of nine other *Solanum* species currently grown in the greenhouse, including *S. bulbocastanum*, *S. commersonii*, *S. microdontum*, *S. pinnatisectum*, *S. tuberosum*, *S. tarijense*, *S. gandarillasii*, *S. boliviense*, and *S. megistacrolobum*, we expect to study these species-specific signaling molecules involved in PT guidance and explore their possible involvement in creating a breeding barrier and thus in driving their speciation and evolution.

1.4 PT Guidance In the Style Is Governed By Sporophytic Tissues

Before reaching the ES to undergo double fertilization, PTs experience numerous intimate contacts with the pistil sporophytic tissues, including the stigma, the transmitting tract and its extracellular matrix in the style, as well as the nucellus and integuments in the ovary. This takes place through dynamic vesicle exocytosis and endocytosis and intensive signaling at the clear zone (the extreme tip of the tube) (Zonia and Munnik 2008). The sporophytic tissues synthesize and secrete molecules, which not only promote a robust tube growth, but also assist the PTs in growth directionality from the stigma straight to the ovarian end of the pistil. This stage is termed sporophytic guidance. So far, chemotactic proteins involved in sporophytic guidance have been identified from the stigma, style and ovary, respectively.

1.4.1 Sporophytic Guidance Down the Stigma

In lily, stigma exudates attract *in vitro*-grown PTs, as shown below (Figure 1.4). By use of biochemical methods including cation exchange, gel filtration and HPLC, plantacyanin (9.9 kDa), a member of the blue copper protein family, was identified from the active fraction and named chemocyanin. *In vitro* bioassay indicates that chemocyanin induces PT chemotropism (Kim, Mollet et al. 2003). This attraction is enhanced by the presence

of two other molecules, including a 9.0 kDa SCA (stigma/stylar cysteine-rich adhesion) (Park, Jauh et al. 2000) and a larger stylar pectic polysaccharide (Mollet, Park et al. 2000).

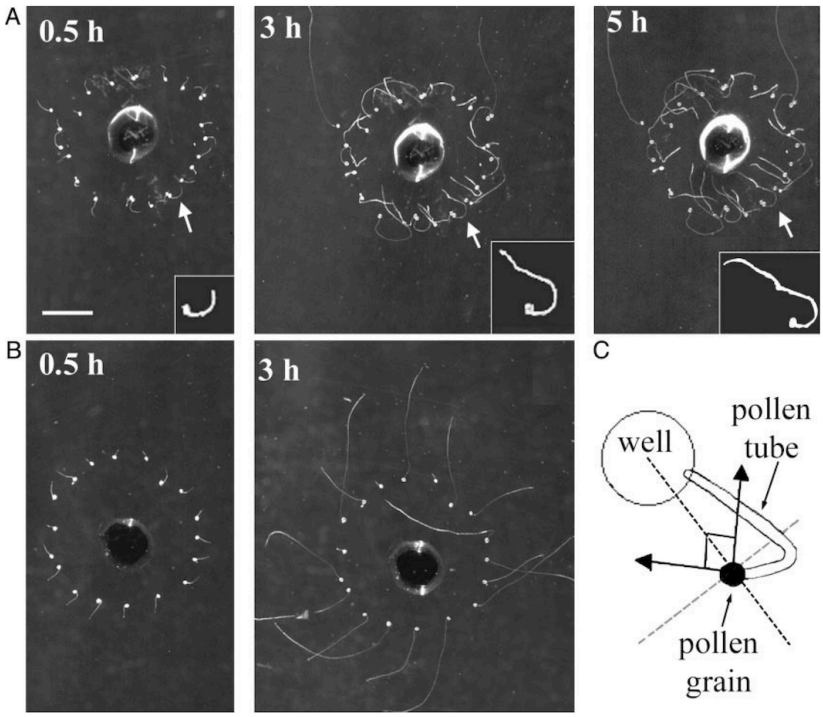


Figure 1.4 *in vitro* PT attraction assay showing PT reorientation over time (A and B) and corresponding quantitation method (C) in lily. A, PT chemotropism was induced after application of stigma exudates (1.5 μg/μl) in the well 40 min before time zero. At this time point, individual pollen grains were placed around the wells. Germinated PTs were manually oriented away from the well before the time course imaging. B, no chemotropism was induced by water control in the well. C, determination of a positive chemotropism marked by PT reorientation of at least 90°. Bar = 2 mm. (Adapted from (Kim, Mollet et al. 2003)).

Positively charged SCA interacts with negatively charged pectin in the transmitting tract of the pistil to create an active extracellular matrix for the PT to attach. SCA is a lipid transfer-like protein secreted from both the pollen and the transmitting tract tissues. It contains eight cysteine residues (Park, Jauh et al. 2000) and is part of the DEFL group. Therefore, chemocyanin is general considered to be involved in adhesion-mediated PT guidance.

In *Arabidopsis*, only one gene encodes plantacyanin, which shares 86.8% sequence similarity to lily chemocyanin at the amino acid level (Kim, Mollet et al. 2003). This cell wall protein is secreted to the extracellular matrix of the transmitting tract and is found in a gradient of increasing concentration from the stigma to the ovary and the mature ES (Dong, Kim et al. 2005). Over-expression of plantacyanin caused wild-type PTs to dwell on the stigma and to turn around without penetrating the papillar cells, as shown in Figure 1.5. An increased plantacyanin concentration was detected in the stigma of overexpression lines, leading to a disrupted PT guidance. The role of *Arabidopsis* plantacyanin reinforced the chemotropic effect of lily chemocyanin and also confirmed the contribution of sporophytic factors to PT guidance.

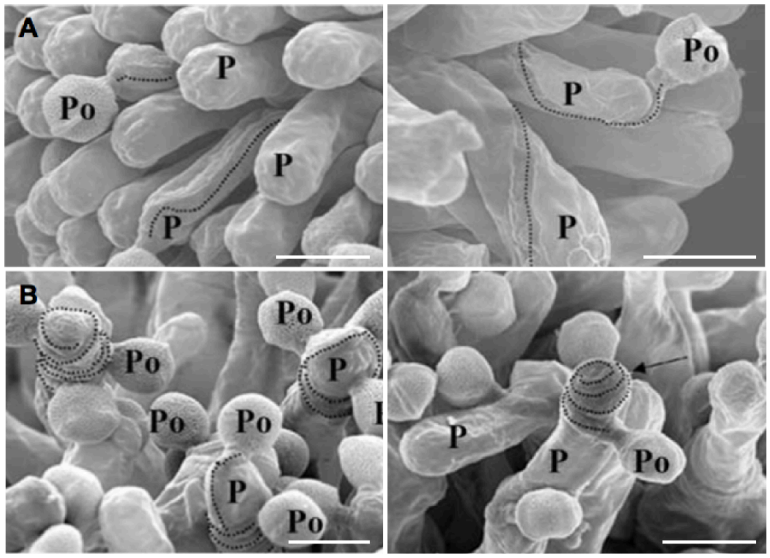


Figure 1.5 Scanning Electron Microscopy (SEM) images of wild-type *Arabidopsis* pollens on (A) wild-type stigmas and (B) plantacyanin over-expression stigmas *OXP12* Dotted line indicates the trajectory of PT path after it penetrates the papilla cell wall. The arrow in B indicates that the PT made many turns on the papilla cell, displaying aberrant tube growth on the stigma. P, papilla cell; Po, pollen grain. Bar = 20 μm. (Adapted from (Dong, Kim et al. 2005)).

1.4.2 Sporophytic Guidance From Stigma to Ovary

The transmitting tissue-specific (TTS) glycoprotein, was identified in the extracellular matrix of the style in tobacco (Wang, Wu et al. 1993). TTS was shown to stimulate PT growth *in vitro*, while transgenic plants expressing significantly lower levels of *TTS* mRNA and protein exhibited a reduced PT growth rate. Besides its role as a growth stimulant, TTS was also shown to attract PTs in a semi *in vivo* assay (Cheung, Wang et al. 1995). TTS belongs to the large arabinogalactan protein (AGP) family. The glycosylation status of TTS is required for PT guidance in the style. PTs that grew through the style were attracted by native TTS but not by a chemically deglycosylated version (Cheung, Wang et al. 1995). In the pistil, TTS works as a vehicle that carries sugar moieties, adheres to PT surface, becomes incorporated into the PT cell wall and finally deglycosylated by the PTs. The glycosylation level of TTS increases along the style, with lower levels of glycosylation in the stigma and higher level in the lower part of the pistil, as indicated by the molecular weight of TTS on a SDS-PAGE gel (Wu, Wang et al. 1995). Concurrently, the increasing level of TTS glycosylation is also associated with an increasing level of TTS protein acidity, which may contribute to PT guidance from the stigma to the ovary.

Characterization of tobacco TTS homolog in *Nicotiana alata* (NaTTS) revealed that they share similar features as a PT stimulant and attractant (Wu, Wong et al. 2000). In addition, NaTTS proteins are part of a protein family, with various degrees of sugar modification on the protein backbone. This variability significantly affects the function of TTS since purified low-molecular weight NaTTS proteins did not show any activity in stimulating PT growth or in PT guidance (Sommer-Knudsen, Lush et al. 1998).

1.4.3 Sporophytic Guidance To the Micropylar Region

Gamma-amino butyric acid (γ-amino butyric acid, GABA) is a signaling factor that plays critical roles in neural development of mammals (Pontes, Zhang et al. 2013). In plants, GABA stimulates PT growth (Palanivelu, Brass et al. 2003). A proper amount of GABA stimulates tube growth *in vitro*, while higher concentrations are inhibitory. GABA levels are regulated by *POP2* (pollen pistil interaction 2) in *Arabidopsis*, which encodes a

transaminase that degrades GABA. By examination of the abundance of GABA in the pistil, an increasing GABA gradient was detected from the stigma to the inner integument, with the highest GABA level in the micropylar region of the ovule. This is likely due to limiting amount of POP2 proteins in high GABA cells, since anti-POP2 antibodies showed staining complementary to the GABA localization in the ovule. Since the GABA gradient coincides with the direction of PT growth, GABA has been considered as a sporophytic factor that guides PTs toward the micropyle region of the ovule.

POP2 is ubiquitously expressed in both the PT and the pistil, and interestingly, only when both were knocked down could a guidance defect be detected. Such defects included PT growth arrest or random growth on the surface of the ovule (Palanivelu, Brass et al. 2003). The *pop2* mutant displays PT elongation defects in the transmitting tract (Renault, El Amrani et al. 2011), which is likely to be associated with modulation of putative Ca^{2+}-permeable channels on the plasma membrane of the PTs by GABA, as shown in tobacco (Yu, Zou et al. 2014).

To sum up, when PTs navigate through the pistil to approach the ES, they are guided by different attracting molecules. On the stigma, PT directional growth is influenced by chemocyanin, a small basic protein in the plant cell wall that belongs to the phytocyanin family of blue copper proteins (Kim, Mollet et al. 2003). Next, when PTs have travelled further into the style, guidance is taken over by TTS, a heavily glycosylated AGP, which becomes deglycosylated as PTs elongate toward the ovarian end of the style (Cheung, Wang et al. 1995, Wu, Wang et al. 1995). Meanwhile, a non-protein amino acid, GABA, may be involved in guiding the PTs toward the micropylar region of the ovule, where highest GABA levels are detected.

The distinct nature of different guidance cues isolated so far suggests that different guidance mechanisms may act. Along the PT path, a new mechanism is likely to take over when PTs migrate into its effective range. With the presence of each new guidance cue, PTs are likely to break free from the previous attraction and be precisely guided under the new influence, until they achieve double fertilization.

1.5 PT Guidance In the Ovary Is Controlled By Both Gametophytic and Sporophytic Tissues

Before PTs exit the bottom-end of the style, sporophytic stylar tissues navigate all the PTs toward the ovary. Once inside the ovary, a precise one-to-one guidance is achieved between the PT and the ovule, a stage described as ovular PT guidance (Higashiyama and Takeuchi 2015). Multi-steps control checkpoints are known to act in ovular guidance, as described in Figure 1.6.

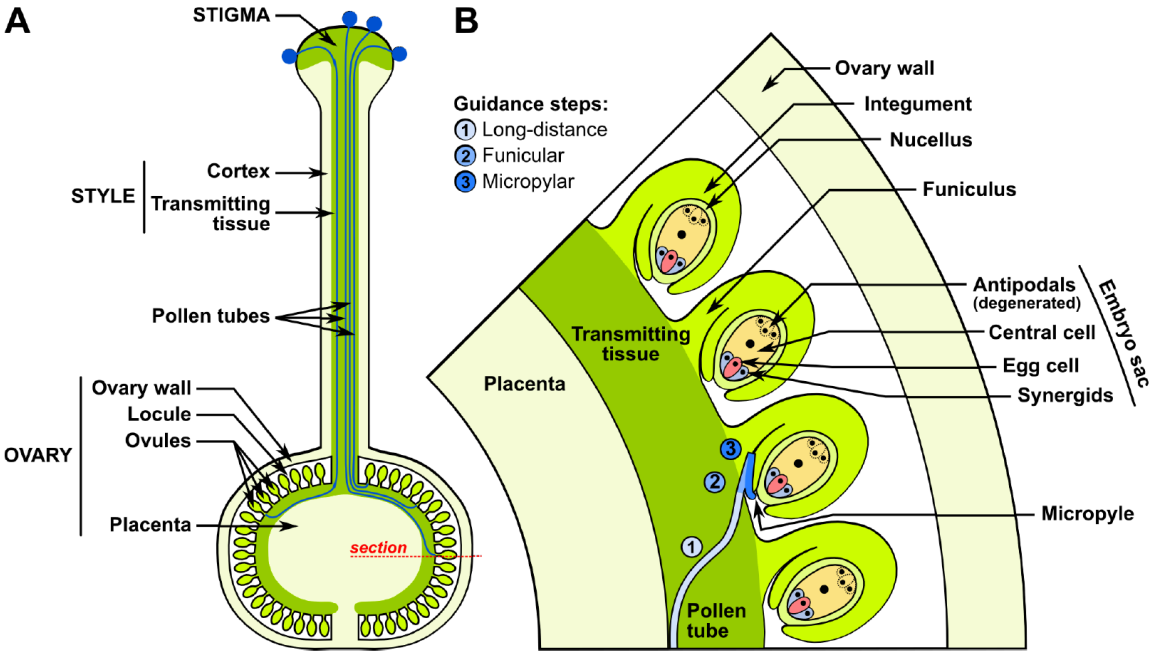


Figure 1.6 Schematic depiction of PT growth in *S. chacoense* pistil. A. Longitudinal section of the pistil. Dark green indicates the PT pathway, a continuous cell tract originating from the stigma surface, extending through the style down to the placental epithelium, where it connects to the ovules. At the mature stage of an ovule, the ES consists of an egg cell, two synergid cells and a central cell, embedded in one layer of integument in *S. chacoense*. B. Close-up of an ovary transverse section. PT typically approaches the micropyle of an ovule through three guidance steps. Step 1 (long-distance guidance), PT emerges from the intercellular space of the transmitting tract to the placental surface; step 2 (funicular guidance), the PT climbs up the funiculus; and step 3 (micropylar guidance), PT navigates toward and into the micropyle. Noticeably, the funiculus in *S. chacoense* has a much shorter stalk compared to *Arabidopsis*, which largely reduces the distance between the micropyle and the placental surface, possibly rendering PTs that emerge onto the placental epidermis immediately competent for micropylar guidance (Liu, Joly et al. 2015).

1.5.1 Long-Distance Guidance Is Governed by the Ovule

A long-distance guidance mechanism was first suggested years ago to describe ovular signals that direct PTs to emerge from the intercellular space of the transmitting tract cells toward the ovules, preferentially at the upper part of the ovary (Hulskamp, Schneitz et al. 1995). It was observed that in wild-type *Arabidopsis* ovules, 40% of the PTs emerged from the transmitting tract at the position of the first ovule. In the *bel1* and *54D12* mutants (Ray, Robinson-Beers et al. 1994), which are impaired in ES development, this preferentiality is compromised, but a slight preference was still observed. However, in the ES-devoid mutants *sin1* and *47H4* (Ray, Lang et al. 1996), PT showed equal probability to emerge anywhere along the entire length of the transmitting tract. This shows there is an ovule-derived long-distance guidance (transmitting tract toward ovules), as opposed to micropylar short-distance guidance. Long-distance guidance is supported by the transmitting tract-trapped phenotype observed in the *Arabidopsis chx21*/*chx23 double* mutant. *CHX21* and *CHX23* are redundant cation/proton exchangers expressed in the PT. Mutant *chx21*/*chx23* PTs do not respond to ovular cues and remain in the transmitting tract (Lu, Chanroj et al. 2011).

1.5.2 Funicular Guidance Is Under the Influence of Ovular Sporophytes

The second step involves funicular guidance, where PTs adhere to and migrate up the funiculus, a stalk-like structure that attaches the ovule to the placenta. The source of funicular guidance cues can be derived from the funiculus itself, as seen in single ovule semi *in vivo* assays, where ovules without a funiculus attracted PTs at an efficiency reduced from 54 to 35% (Palanivelu and Preuss 2006). However, little is known about this signal, due to the difficulty in obtaining sufficient amount of funiculi from the ovules without tissue injury. Alternatively, some evidence suggests that funicular guidance may be associated with ovule sporophytic tissues in general, including the funiculus, the integuments and the nucelluar tissues (Lausser, Kliwer et al. 2010), as observed in the sporophytic mutants *pop2* (Palanivelu, Brass et al. 2003), *siz1-2* (Ling, Zhang et al. 2012) and *pdil2*-1 (Wang, Boavida et al. 2008). In these mutants, a portion of PTs loses their way soon after exiting the transmitting tract tissue and does not migrate up the funiculus.

Since funicular guidance occurs before micropylar guidance within a short period of time, it is difficult to distinguish the two stages. The first article to clarify a funicular guidance mutant is the study of *Arabidopsis mpk*3/*mpk*6 mutant (Guan, Lu et al. 2014). Here the majority of mutant PTs exited the transmitting tract, but either got lost in the septum or took much longer time to eventually target the ovule, suggesting a guidance defect *in vivo*. Interestingly, when they conducted single ovule semi *in vivo* assay on PT growth medium, it was discovered that micropylar guidance was not affected in this mutant since PTs made sharp turn to target the micropyle precisely on the medium. This experimental set-up emphasizes the distinct nature of different signals controlling funicular and micropylar guidance, at least in *Arabidopsis*.

Anatomically, compared to *Arabidopsis*, *S. chacoense* ovules have a highly reduced funiculus. PTs that emerge from the placental epidermis may have direct access to the micropyle, thus casting doubt on the role of funicular guidance in this species.

1.5.3 Micropylar Guidance Is Under the Influence of the FG

Besides long-distance and funicular guidance, the third step in ovular guidance is micropylar guidance, which targets PTs to the micropyle of the ovule within close range of about 100-150 μm (Higashiyama, Kuroiwa et al. 1998, Palanivelu and Preuss 2006, Stewman, Jones-Rhoades et al. 2010). This phase has been well studied since the first identification of micropylar chemoattractants in maize (Marton, Cordts et al. 2005), *Torenia fournieri* (Okuda, Tsutsui et al. 2009) and *Arabidopsis* (Takeuchi and Higashiyama 2012).

1.5.3.1 Synergid Cells Are the Sources of Micropylar PT Guidance

The synergid cells were reported to be the source of micropylar attractants emitted from the ES (Higashiyama, Yabe et al. 2001). *Torenia fournieri*, which bears a naked ES protruding from the micropyle of the ovule, was used to investigate the role of synergid cells in guidance. In their experimental setup, two synergids, the egg cell and the central cell were selectively laser-ablated from the ovule and placed in front of the PTs that grew out of a style to test if the tubes were attracted by the ovules. It was observed that the attraction was not affected when the egg cell or the central cell was ablated, suggesting

that they were not directly involved in guidance. However, the removal of one synergid cell partially impaired attraction and the removal of both completely abolished the attraction, confirming the synergid cells are the source for these chemotactic signals.

In the junction of two synergid cells lies the filiform apparatus, a thickened structure originating from intensive cell wall ingrowth of the synergids. It largely increases the surface area in this region and is thought to mediate the transport of molecules into and out of the synergid cells (Huang and Russell 1992). In *Torenia*, entry of the PT to the synergid cell is through the filiform apparatus (Higashiyama 2002). In *Arabidopsis*, however, the PT grows beyond the filiform apparatus to enter the receptive synergid cell at a more distant site to burst its content (Leshem, Johnson et al. 2013).

1.5.3.2 Identified Micropylar PT Attractants

To date, two types of micropylar attractants have been identified. Maize employs a transmembrane domain-containing peptide to guide PTs entry to the micropyle, while *Torenia* and *Arabidopsis* both employ LURE-type attractants, which are cysteine-rich peptides (CRPs), belonging to defensin-like (DEFL) CRP subfamily (Silverstein, Moskal et al. 2007).

The first micropylar attractant was identified from model monocot species maize (Marton, Cordts et al. 2005). ZmEA1 (Zea mays EGG APPARATUS1) is a small protein of 94 amino acids. It is expressed exclusively in the synergid cells and the egg cell. This gene encodes a transmembrane domain, which is thought to be cleaved from the mature peptide through proteolysis for its secretion. Interestingly, careful examination of the translation start site of ZmEA1, together with the detection of the Kozak consensus sequence (RCCAUGG, R is a purine) near the third methionine (out of four) upstream of the predicted transmembrane domain, led to the discovery of a smaller version of ZmEA1. It contains 76 amino acids and bears a signal peptide for secretion (Germain, Chevalier et al. 2006). After cleavage, a mature ZmEA1 contains only 49 amino acids.

ZmEA1 is spread from the egg apparatus to the nucellar cell wall toward the micropylar opening to attract PTs and is progressively down regulated after fertilization (Marton,

Cordts et al. 2005). Semi *in vivo* assays indicate that mature ZmEA1 attracts maize PTs (Marton, Fastner et al. 2012). Upon contact, PTs internalize these peptides. By degradation of internalized ZmEA1, PTs remain sensitive to surrounding ZmEA1 peptides for their continuous path finding to the micropyle. It was also found that ZmEA1 peptides interact with maize PTs in a species-specific manner, since they did not attract PTs from the maize relative *Tripsacum dactyloides* (Uebler, Dresselhaus et al. 2013). Such species-specificity of ZmEA1 may result from its polymorphism among different species, although sequence similarity revealed that the C-terminal of ZmEA1 is well conserved among various species from both monocot and dicot species, even from basal angiosperms species (Gray-Mitsumune and Matton 2006).

Different from ZmEA1, the LURE-type attractants are CRPs. In *Torenia fournieri*, *TfLURE1* and *TfLURE2* were isolated from a cDNA library made from 20 isolated synergid cells (Okuda, Tsutsui et al. 2009). Following random sequencing of 2112 ESTs, CRPs were found to be the most abundant ESTs, with *TfLURE1* and *TfLURE2* being the most expressed. TfLURE1 and TfLURE2 were shown to attract approaching PTs in close range of 40 μm, a distance corresponding to short-range micropylar guidance. TfLURE1 and TfLURE2 have the same cysteine backbone and share 62.1% amino acid sequence similarity. They function independently to attract PTs, indicating partial functional redundancy in guidance.

Through amplification by gene-specific and degenerate primers, a *TfLURE1* ortholog in *Torenia concolor*, a close species that possesses a similar floral structure as *T. fournieri*, was identified as the micropylar chemoattractant and named *TcLURE1* (Kanaoka, Kawano et al. 2011). TcLURE1 is highly similar to TfLURE1, with only eight amino acids substitution in TcLURE1. These eight substitutions in TcLURE1 resulted in a weaker PT–ovule recognition and attraction toward *T. fournieri* PTs, which may become a speciation driving force.

Using a different approach, Takeuchi and colleagues (Takeuchi and Higashiyama 2012) isolated another LURE-type micropylar attractants cysteine-rich peptide 810_1 (CRP 810_1) in *A. thaliana* and *A. lyrata* through a comparative phylogenetic analysis of more

than 300 DEFL genes in *Arabidopsis*. In this survey, CRP 810_1 peptides form a sole species-specific cluster in *A. thaliana*, which is clearly separated from its close relative *A. lyrata*, whereas other *DEFL* homologs between these two species are intermingled in the clustering. These polymorphic molecules were shown to attract PTs within the range of micropylar guidance and renamed AtLUREs and AlLUREs. Consistently, all LURE-type attractants in *Torenia* and *Arabidopsis* contain six cysteines, with cysteine arrangement slightly different in different families.

1.5.3.3 Characteristics of Micropylar Attractants

Based on current knowledge of micropylar attractants isolated from *Torenia*, *Arabidopsis* and maize, certain characteristics are found in common. First, all these attractants possess a signal peptide. Mature peptides of these attractants function outside of an ES to attract a PT, as shown in Figure 1.7. Immunostaining with anti-AtLURE1 antibodies detected LURE peptide secretion beyond the micropylar region of the ovule to the funiculus (Takeuchi and Higashiyama 2012) and even on the surface of the septum (Higashiyama and Takeuchi 2015). Similarly, ZmEA1 in maize is found secreted to the micropylar region of the ovules (Marton, Cordts et al. 2005).

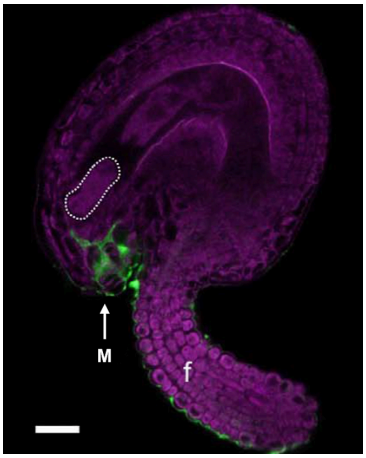


Figure 1.7 Localization of AtLURE1 peptides in the ovule of *A. thaliana*. An ovule was immunostained with anti-AtLURE1 (CRP810_1.2) antibodies. Green Alexa Fluor fluorescence was observed at the micropyle and funicular surface using confocal scanning microscopy. Magenta indicates autofluorescence of the ovule. The synergid cell is delineated by the dashed line. Scale bar = 20 μm. (Adapted from (Takeuchi and Higashiyama 2012)).

The second characteristic of micropylar guidance cues is developmental regulation. In single ovule semi *in vivo* assays, ovules at 32 h before anthesis did not show attraction toward PTs. The guidance competency was gradually increased to 18% by ovules at 16 to 24 h before anthesis and to 49% by anthesis ovules (Palanivelu and Preuss 2006), which revealed the developmental regulation feature of micropylar attractants.

This developmental regulation feature is likely to associate with the maturation stage of the attractants itself. It was observed that only at maturity (anthesis stage), the attractants in maize ZmEA1 are correctly localized to the synergid cells and the egg cell region, whereas premature ZmEA1 was detected in the filiform apparatus of young ovules (Marton, Cordts et al. 2005).

The third characteristic is species-specificity. Prior to the isolation of micropylar guidance cues, experiments were designed to investigate species-preferentiality of micropylar attraction in *Torenia* (Higashiyama, Inatsugi et al. 2006) and *Arabidopsis* (Palanivelu and Preuss 2006). Both studies demonstrate that PTs are significantly attracted to ovules of their own species and rarely target the ovules in other close relatives. In agreement with these results obtained from single ovule semi *in vivo* assays, bead assays using purified attractants from different species in both *Torenia* (Kanaoka, Kawano et al. 2011) and *Arabidopsis* (Takeuchi and Higashiyama 2012) confirmed the species-specific feature of these attracting molecules. This species-specificity may form a pre-zygotic barrier in plant reproduction and thus becomes a speciation driving force.

1.5.3.4 Key Genes Affecting Micropylar Guidance

The myb98 mutant characterized in *Arabidopsis* is defective in micropylar guidance (Kasahara, Portereiko et al. 2005). This gene encodes a transcription factor expressed in synergid cells and affects both PT guidance and the formation of the filiform apparatus. Homozygous *myb98* mutant showed no obvious sporophytic defects. However, 48 genes encoding CRPs, including CRP810 (AtLURE1), CRP3700, CRP3730 and CRP3740, are downregulated in the *myb98* mutant (Punwani, Rabiger et al. 2008), with five genes expressed in the filiform apparatus (Punwani, Rabiger et al. 2007). *MYB98* is thus an important regulator in PT guidance and its promoter has been used to develop a synergid

cell-specific marker for research in PT guidance (Denninger, Bleckmann et al. 2014).

By comparing the expression profile of wild-type *Arabidopsis* ovules to *myb98* mutant ovules and *determinate infertile 1* (*dif1*) mutant ovules (devoid of an ES), 382 genes were found to be downregulated in *dif1* ovules. Of these, 78% had a N-terminal signal peptide for protein secretion, defining a group of extracellular molecules that could function in ES development and/or interaction between the ES and the PT (Jones-Rhoades, Borevitz et al. 2007). Interestingly, 40 genes out of the 382 encoded DUF784 (domain of unknown function 784) family proteins that belong to the CRP superfamily, and 37 of these 40 genes required *MYB98* for expression. Four of these DUF784 genes are exclusively expressed in the synergid cells (Jones-Rhoades, Borevitz et al. 2007). These small, secreted proteins were proposed to be potential mediators of the last stages of ES–PT interaction. Taking into consideration the functional redundancy that *DUF784* genes might have, the expression level of 18 DUF784 genes were simultaneously decreased by RNAi. As a result, 30% of transgenic ovules were not fertilized by wild-type PTs *in vivo* (Chen, Du et al. 2010), indicating their possible involvement in PT guidance.

1.5.3.5 Micropylar Guidance Is Orchestrated By the Egg Cell and the Central Cell

In the ES, communication between the synergid cells, the central cell and the egg cell forms a complex signaling network. Although the synergid cells have been demonstrated as the source of micropylar guidance cues, an intact central cell and egg cell also contribute to the production of active chemoattractants in the synergid cells.

The first piece of evidence can be seen in the maize micropylar guidance attractant ZmEA1, which is localized to both the synergid cells and the egg cell (Marton, Cordts et al. 2005). Similarly, GEX3, a protein involved in micropylar PT guidance, is expressed in the egg cell and pollen (Alandete-Saez, Ron et al. 2008). Reciprocal crosses suggested that the FG was mostly responsible for the reduced seed set since half of the ovules were not fertilized when using *GEX3* antisense lines as the female. Self-crossed pistils in *GEX3* antisense lines showed that, even though 62% of the PTs reached the funiculus of the ovules, all failed in front of the micropyle, pointing to a micropylar guidance defect.

Besides the egg cell, the central cell was also shown to influence micropylar guidance. The *CCG* (*central cell guidance*) gene is exclusively expressed in the central cell. The *ccg* mutant plants are defective in micropylar PT guidance (Chen, Li et al. 2007). In *ccg* hemizygous (Ds/-) mutants, only half of the ovules are penetrated by a PT, compared to 98% in wild-type ovules *in vivo*. Complementation with the *CCG* gene using a central cell specific promoter recovered seed set from 50% to 77%. In the *ccg* mutant, the ES is morphologically normal and the cell fate of other cell types in the ES are correctly specified (Chen, Li et al. 2007). *CCG* encodes a N-terminal zinc ß-ribbon domain, which is functionally interchangeable with the N-terminal of the transcription factor II B (TFIIB). It is reasonable to speculate that *CCG* works as a transcription factor to regulate downstream genes involved in PT guidance. However, it is important to note that the 75kD *CCG* encodes a nuclear protein that is not secreted to the synergid cells, thus suggesting a central cell/synergids communication route. Indeed, the morphology of the central cell opens such likelihood. The central cell occupies most of the volume in the ES, closely adjacent to the antipodal cells and the egg apparatus at the chalazal and micropyle end. It was observed that the cell wall of the central cell is discontinuous around the synergids and egg cell, which could facilitate the communication between these three cell types (Coimbra and Salema 1999). Moreover, the mature central cell cytoplasm harbors an extended endoplasmic reticulum network that could deliver vesicles to the membrane and transmit information to adjacent cells to attract PTs (Liu, Yan et al. 2010).

The *MAA1* and *MAA3* genes also support the involvement of the central cell in PT guidance (Shimizu and Okada 2000, Shimizu, Ito et al. 2008). In the *maa1* (*magatama1*) and *maa3* mutants, development of the FG is delayed and the two polar nuclei do not fuse. In these mutants, PTs emerge from the transmitting tissue of the style, migrate up the funiculus, but fail to approach the micropyle, pointing to a micropylar guidance defect. Here, the fusion of the polar nuclei (central cell maturation) was suggested to be vital for the secretion of the attractants.

Collectively, these studies suggest that both the egg cell and the central cell have an impact on PT attraction, either by triggering a signal to the synergid cells to release the

attractants, *e.g.* at the time of central cell maturation, or by synthesizing the attractants itself and therefore directly involved in guidance.

1.5.4 PTs Slow Down for Fine-Tuning Before Targeting the Micropyle

Stewman and colleagues (Stewman, Jones-Rhoades et al. 2010) investigated PT behavior in a single ovule semi *in vivo* assay in *Arabidopsis*. It was observed that PTs slow down near the micropyle. Furthermore, this reduced growth rate is associated with an increased PT ability to accurately target the micropyle. The studies on nitric oxide (NO) may provide some explanations for this (Domingos, Prado et al. 2015). Gaseous NO was first detected in the PTs of lily and was proposed as a negative factor in reorienting PT growth (Prado, Porterfield et al. 2004). NO is synthesized in the peroxisomes of the PTs from the sub-apical region down the shank. An external point source of NO can penetrate the apical region of the PT and disturb its NO balance, which causes PT reorientation within minutes. On the ovule side, NO signal was also detected all over the ovules, however, leaving the micropylar region devoid of NO, and thus allowing for PT entry (Prado, Colaco et al. 2008). The production of NO in the edges of the micropyle is likely to slow down approaching PTs and therefore, a better targeting efficiency is achieved in micropylar guidance. In consistence with this explanation, *Arabidopsis nos1* (*NO synthase1*) mutant is defective in NO production and seed set (Guo, Okamoto et al. 2003).

1.6 *semi in vivo* Systems to Study Ovular PT Guidance

In the study of ovular PT guidance, different types of assays have been developed. To study short-range micropylar guidance, Higashiyama and colleagues (Higashiyama, Kuroiwa et al. 1998) co-cultivated PTs that grow out of the style with single ovules dissected from the ovary of *T. fournieri* to observe PT attraction to the ES. This system has also been adapted to *Arabidopsis* (Palanivelu and Preuss 2006). In the single ovule assay system, the style is cut from the pistil and placed on solid pollen germination medium, which renders PTs competent to react to ovular signals for attraction (Qin, Leydon et al. 2009, Okuda, Suzuki et al. 2013). Single ovules were excised from the ovary and placed at a distance accessible to emerging PTs from the style (Figure 1.8A). Since PTs grow randomly on the medium, only the ones that grow within ~ 100 μm of an

unfertilized ovule (the micropyle) are observed in this assay (Palanivelu and Preuss 2006).

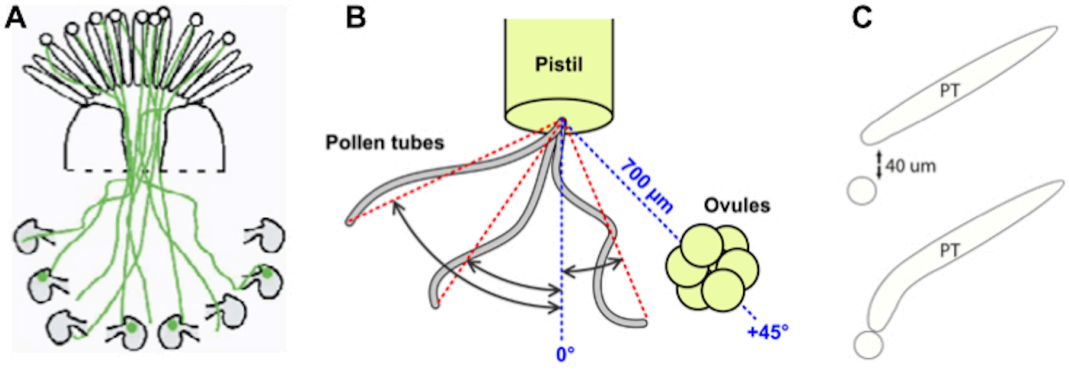


Figure 1.8 Diagram of different assay systems to evaluate PT attraction. A. In the single ovule assay, PTs emerge from the cut end of the pistil, travel across the agarose medium before entering the excised ovules. A green dot inside the ovules describes successful PT targeting and indicates the area where PT bursts. (Adapted from (Palanivelu and Preuss 2006)). B. The semi *in vivo* assay system in *S. chacoense*. Ovule clusters connected by placental tissue are positioned ~ 700 µm away from the ovarian end of the style (equivalent to the radius of the ovary), with a 45° angle. C. In the bead assay, proteins are encapsulated into a gelatin bead and positioned along the path of a growing PT (~ 40 µm).

In *S. chacoense*, a slightly different semi *in vivo* assay was used to study PT guidance. Because of their densely packed ovules and very short funiculus, *S. chacoense* ovule clusters (5 to 10 ovules) are dissected and placed on one side of a pollinated style (Figure 1.8B), at ~ 700 µm from the cut style end, a distance equivalent to the radius of the ovary (Lafleur, Kapfer et al. 2015). These two assays differ in terms of the effective attraction range, which results in completely different PT behavior in response to ovular signals. In the single ovule assay, PTs usually make a sharp turn toward the micropyle of a single ovule within 100 µm, while in our semi *in vivo* assay, PTs gradually change their trajectory toward ovule clusters and this tendency can be observed right after PTs emerge from the style, suggestive of a long-distance guidance activity, as opposed to the short-distance micropylar guidance.

In parallel to the ovule-based semi *in vivo* assays, a protein-based assay, the bead assay, was also designed (Okuda, Tsutsui et al. 2009). In the bead assay (Figure 1.8C), a

purified protein or proteins (candidate chemoattractant) are encapsulated into a gelatinous matrix to form microbeads that can be precisely positioned along the path of a PT (~ 40 μm away) in order to determine their chemoattracting potential. It has been widely used in validating PT chemoattractant interactions in *Torenia*, *Arabidopsis* and maize.

On the basis of the above-mentioned bioassays, several on-chip devices have also been designed, which mimic the *in vivo* environment of PT targeting toward the ovules and showed a quite similar PT targeting efficiency compared to their macroscopic counterparts (Yetisen, Jiang et al. 2011, Horade, Kanaoka et al. 2013, Sato, Sugimoto et al. 2015). For these miniaturized devices, specialized chips are manufactured, which carry narrow channels to steer *semi in vivo*-grown PTs toward two reservoirs, containing either plant tissues (ovules) or liquid solution (attractants). These on-chip assaying systems facilitate researchers to place the ovules or putative attractants to the precise location over multiple repetitions. More importantly, the microchannels where PTs elongate will allow the formation of a gradient of chemicals that diffuse from two reservoirs. Such gradient gives PTs a distinct choice to grow toward either left or right, in order to screen for true attractants.

1.7 Objectives

The main focus of my PhD project is to study ovular signals, more specifically, embryo sac (ES)-expressed genes that play important roles in pollen–pistil interactions in solanaceous species, using the model plant *S. chacoense*. This study is achieved through transcriptomic and proteomic (secretomic) comparative analyses using:

(1) Wild-type ovules that contain a mature ES at the FG7 stage;
(2) ES-devoid mutant ovules (*frk1*) where the ES is stopped at the functional megaspore stage (FG1) or that entirely lack an ES;
(3) Slightly immature wild-type ovules at two days before anthesis (2DBA), which contain a developing ES at the FG6 stage (unfused polar nuclei) while the overall morphology resembles the mature ones.

Compared to wild-type mature ovules, the use of *frk1* ovules defines ES-expressed genes and the use of 2DBA ovules defines ES-expressed genes that are specifically involved in late stages of ES maturation during the FG6 to FG7 transition.

Both transcriptomic and secretomic approaches were used in this study. Chapter II describes the transcriptomic approach that studied ES-genes at the mRNA level. Chapter III describes the secretomic approach that specifically targets secreted proteins emanating from the ES.

In addition, as compared to wild-type mature ovules that were shown to attract semi *in vivo* grown PTs, neither *frk1* nor 2DBA ovules attract PTs. These two comparative analyses thus generated two candidate gene pools that are potentially involved in ovular PT guidance in *S. chacoense*. Thus, Chapter IV describes ovular PT guidance chemoattractant candidate selection, expression, purification and their functional assays. It is postulated that these chemoattractants evolved among species that could serve as interspecific breeding barriers and become a driving force for speciation. By studying the structure/function relationship of these chemoattractants, it will become clearer how pollen–pistil interactions affect speciation and the role of chemoattractants in interspecific breeding barriers.

The following detailed objectives are given for each chapter.

Chapter II. Ovule Transcriptome Profiling:

a. Construction of a scaffold assembly for *S. chacoense*.

b. Construction of a scaffold assembly for two *S. chacoense* close relatives, e.g., *S. gandarillasii* and *S. tarijense*.

c. Identification of ES-expressed genes in *S. chacoense*.

d. Identification of ES maturation-associated genes during FG6 to FG7 transition.

e. Studying the evolution of ES-expressed CRPs in solanaceous species.

f. Determination of candidates genes involved in ovular guidance.

g. Elucidation of the relationship between glycosylation and ovular PT guidance.

Chapter III. Ovule Secretome Profiling:

a. Profiling secreted proteins in the ovular locule.

c. Identification of ES maturation-related secreted proteins (FG6 to FG7).

d. Determination of candidate proteins involved in ovular PT guidance.

e. Studying the correlation between protein abundance and their mRNA level.

Chapter IV. Isolation of Chemoattractants in Ovular PT Guidance:

a. Identification of ovular PT guidance chemoattractants in *S. chacoense*.

b. Identification of their orthologs in other solanaceous species and determination of their role in interspecific breeding barriers.

2. Gene expression during late stages of embryo sac development: a critical building block for successful pollen–pistil interactions

Yang Liu[*], Valentin Joly[*], Mohammed Sabar[†], and Daniel P. Matton[§]

Institut de Recherche en Biologie Végétale, Département de Sciences Biologiques, Université de Montréal, 4101 rue Sherbrooke est, Montréal, Québec H1X 2B2, Canada

Author contributions: [*]Y.L. and V.J. contributed equally to the work. Y.L. carried out all of the experiments, except for PT attraction assays, which was performed by Mohammed Sabar. V.J. performed bioinformatics analyses and generated tables and figures from these analyses. Y.L. interpreted the results and led the writing of this article. D.P.M. oversaw this project and gave critical corrections and improvement to the manuscript.

[†]Present address: Institut des sciences de la santé et de la vie, La Cité collégiale, 801 Aviation Parkway, Ottawa, ON K1K 4R3, Canada; Department of Biology, McGill University, 1205 ave Docteur Penfield, Montreal, QC H3A 1B1, Canada

[§]To whom correspondence should be addressed.

In preparation for: *The Journal of Experimental Botany*

Abstract

The Solanaceae family, especially the genus *Solanum*, contains several economically important crop species, such as potato, tomato and eggplant. *Solanum* itself harbors almost half of the species in this family, with around 1500 species including all the wild potato species found in the Western hemisphere. With widespread interspecific hybridization, *Solanum* species are an ideal model for the study of sexual reproduction and breeding barriers. In this study, we assembled a catalog of 33852 CDS-containing contigs for *Solanum chacoense* through whole transcriptome shotgun sequencing **(**RNA-seq). Gene expression profiles were compared between wild-type mature ovules, slightly immature ovules two days before anthesis, and ovules from the MAPKKK *frk1* mutant that produces embryo sac (ES)-less ovules, in order to define ES-expressed genes and genes involved in late ovule maturation steps relevant to pollen tube guidance. Eight hundred and eighteen (818) genes were found downregulated in *frk1* ovules, and identified as ES-dependent genes. A subset of 284 genes was concurrently under-expressed in immature ovules and thus could be directly associated with late stages of ES maturation. Of these 284 genes, 75% were ovule-enriched when compared to leaf expression levels and 21% encoded cysteine-rich peptides (CRPs), many of which showed unique cysteine motifs not yet observed in other species. Compared to parallel studies conducted in *Arabidopsis thaliana* using similar ES-devoid mutants, *S. chacoense* was found to possess a highly divergent ES transcriptome from that of *A. thaliana*, in terms of both functional categories and limited orthologous similarities. Interestingly, of the 59 CRPs enriched in *S. chacoense* ovules, 60% were predicted to be glycosylated. Although micropylar guidance proteins that attract pollen tubes in *Torenia*, maize and *Arabidopsis* are not glycosylated proteins, our study showed that the ovule glycoprotein fraction enhanced guidance efficiency in *S. chacoense*, emphasizing its highly divergent ovule transcriptome from various plant species.

Keywords: female gametophyte, embryo sac maturation, pollen tube guidance, glycosylation, ovule, Solanaceae

2.1 Introduction

The plant female reproductive organ, the ovule, consists of the female gametophyte (FG), integument(s) and nucellus. So far, plant ovules have been shown to control a wide range of pollen–pistil interactions through various key signaling molecules, including pollen tube (PT) elongation, competency control, guidance, growth arrest and discharge, as well as gamete fusion and embryonic development (Long 2006, Dresselhaus and Franklin-Tong 2013, Bleckmann, Alter et al. 2014). Interestingly, though immotile, ovules can also communicate with the male gametophyte at a distance, as observed from specific genes being induced or enhanced in the ovule by pollination events even before PTs are in their vicinity (Dong, Kvarnheden et al. 1998, Lantin, O'Brien et al. 1999).

The FG is a major ovular signaling hub, where extensive cell-cell communication events take place to effect fertilization, leading to seed set (Sundaresan and Alandete-Saez 2010). Embryo sac (ES)-expressed genes have been identified through expression profile comparisons between wild-type and mutants ovules in the dicot species *Arabidopsis thaliana* (Yu, Hogan et al. 2005, Johnston, Meier et al. 2007, Jones-Rhoades, Borevitz et al. 2007, Steffen, Kang et al. 2007, Sanchez-Leon, Arteaga-Vazquez et al. 2012) and in monocot species like maize (Yang, Kaur et al. 2006, Wang, Wang et al. 2014) and rice (Ohnishi, Takanashi et al. 2011, Kubo, Fujita et al. 2013) using enzymatic maceration and micro-dissection techniques. Simultaneously, the ES-dependent and ovular sporophyte-expressed genes are being uncovered, revealing a complex gametophyte-to-sporophyte crosstalk (Johnston, Meier et al. 2007, Jones-Rhoades, Borevitz et al. 2007, Armenta-Medina, Huanca-Mamani et al. 2013). Genes expressed in individual cell types of the FG have also been defined in *Arabidopsis* (Wuest, Vijverberg et al. 2010). With only one dicot species *Arabidopsis* surveyed for ES-expressed genes, and a recent discovery of an intriguing expansion of some specific gene families in solanaceous species that function in ES development, like the protein kinases from the FRK family (Gray-Mitsumune, O'Brien et al. 2006, Daigle and Matton 2015), a comparative analysis of *Arabidopsis* ES-expressed genes with another important dicot species, *S. chacoense*, is indispensible and informative from an evolutionary perspective.

Here, we aim to study ES-expressed genes in *S. chacoense*, a wild diploid potato species native to South America. Among the 196 wild potato species (*Solanum* sect. *Petota*), *S. chacoense* is one of the most widespread species and is distributed from the South-Eastern regions of Peru to central Argentina (Miller and Spooner 1996). Numerous close relatives, such as *S. gandarillasii* and *S. tarijense*, occur sympatrically with *S. chacoense*, thus providing great opportunity for extensive hybridization (Hijmans, Spooner et al. 2002). For this reason, *S. chacoense* was the first wild potato species investigated for introgressive hybridization (Hawkes 1962) and over the years it has become a model species in the study of self-incompatibility and sexual reproduction (Luu, Qin et al. 2000, Gray-Mitsumune, O'Brien et al. 2006, Soulard, Boivin et al. 2014, Lafleur, Kapfer et al. 2015). By sequencing and annotating the transcriptome of *S. chacoense*, a wealth of data can be gained on this model species and will be an asset to the Solanaceae research community. Together with the genome annotation of its close relatives of tomato (Tomato Genome Consortium 2012) and potato (Potato Genome Consortium 2011), and the expressed sequence tag (EST) databases of various other solanaceous species available from Sol Genomic Network (Fernandez-Pozo, Menda et al. 2015), species-specific ovular signals can be broadly studied.

In this study, we assembled a catalog of 33852 coding sequence (CDS)-containing contigs for *S. chacoense*. Differential gene expression analyses were performed between wild-type mature ovules and two other ovule conditions: anthesis stage ES-devoid ovules from the *fertilization-related kinase 1* (*frk1*) mutant background and slightly immature ovules obtained two days before anthesis (2DBA). In strongly affected *Scfrk1* mutant lines, functional megaspores do not progress beyond the FG1 stage and 94% of the mutant ovules entirely lack an ES (Lafleur, Kapfer et al. 2015). In wild-type plants at the 2DBA stage, most of the ovules are at the FG6 stage (Chevalier, Loubert-Hudon et al. 2013). During the last maturation step from FG6 to FG7, the polar nuclei fuse and the filiform apparatus arising from synergids cell wall ingrowth matures (Sundaresan and Alandete-Saez 2010). This maturation step was shown to be critical to the secretion of chemoattracting proteins involved in PT guidance (Liu, Joly et al. 2015). Comparison between these three ovule conditions shall lead to a better understanding of the gene

regulatory networks involved in late ovule maturation steps as well as to uncover a series of ES-specific genes that are vital to successful pollen-ovule interactions.

2.2 Material and Methods

2.2.1 RNA Purification and Sequencing

S. chacoense Bitt. individuals were greenhouse grown at ~25°C under long-day conditions (16 h light/8 h dark). Plant tissues from three ovule conditions (anthesis, 2DBA, *frk1*), as well as from mature dry pollen and from *in vitro* grown PTs were collected over the same period of time (2010). Total RNAs were extracted using TRIzol reagent (Life technologies). Ovaries were dissected to remove the pericarp and flash-frozen in liquid nitrogen upon harvest. PTs were obtained through incubation in BK liquid medium for 5 h (Brewbaker and Kwack 1963).

RNA-seq was performed using two next-generation sequencing platforms. First, the 454 GS-FLX Titanium platform was used to sequence the ovules and PTs, with each sample occupying one sequencing plate. cDNA libraries were constructed with the rapid library preparation kit (Roche) after an mRNA enrichment step performed with Dynabeads Oligo$(dT)_{25}$. Second, the Illumina HiSeq 2000 platform was used to sequence ovules at anthesis stage only, which occupied one lane. The cDNA library was prepared using the TruSeq cDNA preparation kit (Illumina). For *S. gandarillasii* and *S. tarijense*, RNAs from anthesis ovules were sequenced likewise using the Illumina Hiseq 2000 platform.

2.2.2 Sequence Assembly and Curation

For *S. chacoense*, 454 reads were assembled into 53524 contigs using Newbler the package (Margulies, Egholm et al. 2005). Illumina reads were assembled into 47138 contigs using the Trinity package (Grabherr, Haas et al. 2011). A sequence curation protocol was applied to the two primary assemblies, following a protocol described previously with modifications (Liu, Joly et al. 2015). In brief, contigs from both assemblies were concatenated and first blasted against the NCBI's refseq_rna database using BLASTn to detect and split chimeric sequences into parts. Resulting sequences were then blasted against the NCBI's refseq_protein using BLASTp. Frameshifted

contigs were resolved by addition of one or two Ns at the appropriate position in the sequence. Due to the large volume of RNA-seq data, contig groups were then generated based on all-versus-all BLASTn comparison. Contigs displaying high level of similarity, i.e. ≥ 95% sequence identity over ≥ 80% sequence coverage, were grouped and aligned with the MAFFT program (Katoh and Standley 2013). A consensus sequence was extracted from each alignment, constituting a hybrid assembly, on which we performed the above-mentioned chimera and frameshift curation steps again and obtained a final dataset of 45211 consensus contigs. For *S. gandarillasii* and *S. tarijense*, Illumina reads were assembled using the Trinity package.

2.2.3 Gene Expression Analysis

The cDNA libraries made from three biological replicates of anthesis, 2DBA, *frk1* ovules and leaf were barcoded and sequenced on the Illumina Hiseq platform, with each replicate occupying $1/4^{th}$ of a lane. The expression level of contigs was assessed by mapping reads to the hybrid assembly with the Bowtie aligner (Langmead, Trapnell et al. 2009). The RSEM program was used to estimate the abundance of contigs and to generate a matrix of raw read counts (Li and Dewey 2011). The edgeR package was used to compute adjusted fold-changes between conditions and the corresponding *p*-values (Robinson, McCarthy et al. 2010).

Semi-quantitative RT-PCR validation of differential gene expression data was performed on selected genes using 70 ng of total RNA from anthesis, 2DBA, *frk1* ovules and leaf in biological triplicates using the M-MLV reverse transcriptase kit (Invitrogen). The cDNA templates were amplified for 26 cycles for all genes tested. Ubiquitin was used as the internal control. Primer sequences used for RT-PCR analyses are listed in Table S2.1.

2.2.4 Aniline Blue Staining

Wild-type ovules (G4 genotype, $S_{12}S_{14}$ self-incompatibility alleles) were pollinated with compatible pollens (V22 genotype, $S_{11}S_{13}$ self-incompatibility alleles). Forty hours after pollination, pistils were fixed, squashed and stained following a protocol described previously (Matton, Maes et al. 1997). The pericarps were carefully removed before

squashing the pistils and staining. PT behavior near the ovule was observed with an Axio Imager.M1 microscope equipped with an Axiocam HRc camera (Zeiss).

2.2.5 Scanning Electron Microscopy (SEM)

Fertilized pistils were fixed, dehydrated and critical-point-dried as described previously (Gray-Mitsumune, O'Brien et al. 2006). Pistils were mounted on double-taped SEM stubs and the pericarps were dissected to reveal fertilization events prior to coating. Pictures were acquired with JEOL JSM-35 SEM.

2.2.6 Protein Extraction

Ovules at anthesis from *S. chacoense* were collected (pericarp removed), placed on dry ice and stored at −80 °C until use. For each extraction, 300 mg of ovules were used. For total protein extraction, samples were ground on ice with a mortar and pestle in liquid nitrogen until a fine powder was obtained. The powder was resuspended in a buffer containing 10 mM Tris·HCl (pH 7.2), 0.5 mM L-cysteine and 0.01% (v/v) Triton X-100. The cold ovule paste was transferred to a 1.5 ml tube and ground again with a plastic pestle on ice until an aqueous extract was obtained. The homogenate was centrifuged for 15 min at 13000 *g* and 4 °C. The supernatant was flash-frozen until used. For soluble protein extraction, Triton X-100 was omitted from the buffer. For glycoprotein purification, Con A Sepharose 4B beads (GE Healthcare) were used. Frozen ovules were ground on ice and total proteins were extracted in ConA binding buffer, containing 10 mM Tris·HCl (pH 7.2), 0.5 M NaCl and 0.5 mM L-cysteine. Five hundred (500) µL bed volume of ConA beads were packed into a column and equilibrated with 10 mM Tris·HCl (pH 7.2). After loading the total protein extracts, the beads were washed with 10 mM Tris·HCl (pH 7.2) and eluted with 1 M methyl-glucopyranoside in the wash buffer. Eluents were dialyzed against 10 mM Tris·HCl (pH 7.2) before application in PT guidance assay.

2.2.7 semi *in vivo* PT guidance Assay

Two types of PT guidance assays were used. When fresh ovules were used to attract semi *in vivo*-grown PTs, the assays were performed as described previously (Liu, Joly et al.

2015). In brief, 24 hours after compatible pollination of the *S. chacoense* G4 (♀) by the V22 (♂) genotypes, styles were excised and placed on a BK solid medium (Brewbaker and Kwack 1963), allowing PTs to exit the ovarian end of the style roughly 30 h after pollination. PT attraction toward different ovule clusters (anthesis) from *S.chacoense*, *S. tarijensi* and *S. gandarillasii* was evaluated in a single-choice assay system (Lafleur, Kapfer et al. 2015) The turning angle of each PT was measured with ImageJ (http://imagej.nih.gov/ij). Angle distribution was compared between conditions using a Kolmogorov–Smirnov (KS) statistical test ($n > 100$). When protein extracts were used to attract PTs, the attraction was evaluated in a two-choice assay system (Lafleur, Kapfer et al. 2015). On the one side, protein extracts (1 µl) was applied on BK agarose medium at 24 h after pollination using a thin glass capillary inserted into a 0.5-10 µl pipet tip, at a distance of 700 µm away from the end of the style, with a 45° angle. On the other side, protein extraction buffer (1 µl) were used as negative control. Drop applications were performed under a stereomicroscope. A Student's *t*-test was performed to evaluate the significance of PT attraction toward each protein extract (n = 150). Guidance response was observed in bright field with a Discovery V12 stereomicroscope equipped with an AxioCam HRc camera (Zeiss).

2.2.8 *in silico* Predictions

Bioinformatic predictions were performed for all protein sequences deduced from the hybrid assembly. Gene Ontology (GO) annotations were extracted using Blast2GO (Conesa and Gotz 2008). For GO enrichment analysis, the number of sequences associated with each GO-term was compared between datasets. GO-terms that had different numbers of associated sequences in two datasets (above 2-fold), and that had at least 3 associated sequences were reported as enriched, following a Fisher's exact test (*p*-value ≤ 0.05). Pfam motif search was performed using Pfam public database (http://pfam.sanger.ac.uk/). Protein secretion and glycosylation status was predicted using SignalP (Petersen, Brunak et al. 2011), SecretomeP (Bendtsen, Jensen et al. 2004), NetNGlyc and NetOGlyc (http://www.cbs.dtu.dk/services/).

2.2.9 Estimates of Diversifying Positive Selection

The *S. chacoense* transcriptome was queried for genes encoding CRPs using the KAPPA program (Joly and Matton 2015). CRP orthologs in other solanaceous species were retrieved using the reciprocal best BLAST hits method and aligned codon-by-codon. Positive selection estimates were conducted based on the ratio of divergence at nonsynonymous and synonymous sites (*dN/dS*) in each alignment. Sequence similarity networks were generated using Cytoscape version 3.2.1.

2.2.10 Accession numbers

All raw sequence data used in this study have been submitted to the NCBI sequence read archive (http://www.ncbi.nlm.nih.gov/sra) under BioProject PRJNA299204, with accession numbers SRX1370955, SRX1370956, SRX1370981, SRX1370994, SRX1371065, SRX1371085, SRX1371099, SRX1371104-1106, SRX1371110-1115 and SRX1371203-1205. Assembled sequences from the curated hybrid assembly (>200 bp) have been submitted to the Transcriptome Shotgun Assembly Database in GenBank with the accession numbers GEDG01000001-GEDG01042873.

2.3 Results

2.3.1 Data Curation

Through trancriptome curation, 13972 chimeras and 757 frameshifted contigs stemming from both individual *de novo* assemblies as well as from the effort of merging them were detected and curated. Curated sequences accounted for 13.9% and 0.8%, respectively, of the two assemblies. The resulting 454/Illumina hybrid assembly comprised 45211 sequences. Of these, 33852 sequences have their CDS delimitated either through comparison to their best BLAST hit in the NCBI's reference sequence (RefSeq) database (34.0%) or through theoretical prediction for contigs longer than 180 bp (66.0%). These CDS-containing contigs were used to perform various transcriptomic analyses in this study.

2.3.2 Validation of the Assembly Through Known Ovule-Expressed Genes

To evaluate the quality of the hybrid assembly, nine genes previously demonstrated by RNA gel blot analyses and *in-situ* hybridizations to be expressed in *S. chacoense* ovules during pre- or post-fertilization events were surveyed. Various localizations were reported for these genes, including the egg and synergid cells (Lafleur, Kapfer et al. 2015), ES (Germain, Gray-Mitsumune et al. 2013), micropylar region (Lantin, O'Brien et al. 1999, Lagace, Chantha et al. 2003), integument (Chantha, Emerald et al. 2006, Gray-Mitsumune, O'Brien et al. 2006, Germain, Gray-Mitsumune et al. 2008) and placental epidermis (O'Brien, Chantha et al. 2005). All marker genes tested were correctly assembled. By contrast, the *rapid alkalinization factor 3* (*ScRALF3*), whose expression decreases concomitantly with megagametogenesis progression and becomes totally absent in flowers at anthesis (Chevalier, Loubert-Hudon et al. 2013), was not found in the assembly, as expected to be absent in the anthesis ovule transcriptome. These results are summarized in Table S2.2.

2.3.3 Gene Expression Validation

The consistency of gene expression data obtained from three biological replicates for each condition was evaluated. As shown in Figure S2.1, the expression level for each gene derived from individual RNA-seq runs is highly reproducible across replicates ($R^2 >$ 0.9), except for replicate #3 of anthesis ovules, which exhibits weak correlation with replicate #1 and #2 ($R^2 < 0.6$). This replicate was thus not included in later analyses.

Next, the accuracy of gene expression data was validated by RT-PCR using selected genes that showed different expression patterns in anthesis, 2DBA, *frk1* ovules and leaf. Figure 2.1 reveals a strong correlation between RT-PCR results and the corresponding gene expression predictions, attesting that the digital gene expression data is an authentic representation of gene expression levels for *S. chacoense* ovule transcriptome.

Figure 2.1 RT-PCR validation of gene expression levels derived from RNA-seq data. Nine genes displaying different expression patterns in ovules at anthesis (A), 2DBA (D), *frk1* (F) and leaf (L) were randomly chosen. Ubiquitin (ScOT_16634) was used as the internal control. The expression level fold-changes obtained from differential gene expression analysis for each gene is shown as a heat map, as compared to the expression level in anthesis ovules.

Verification of the accurate assembly of marker genes and the gene expression data collectively confirmed the high quality of the hybrid assembly and the sufficient depth of RNA-seq used for differential gene expression analyses, enabling us to carry out comparative transcriptomic analyses with high confidence.

2.3.4 Functional Annotation of *S. chacoense* Ovule Transcriptome

Previously, gene expression during *S. chacoense* embryogenesis was analyzed using a 7.7 K cDNA microarray (Tebbji, Nantel et al. 2010). As the first next-generation sequencing-based transcriptome study in this species, functional annotations were performed globally using Blast2GO program based on amino acid sequences extracted through CDS delimitation. GO-terms were successfully extracted from 72.4% of genes. Number of proteins associated with each GO-term is indicated under biological process, molecular function and cellular component (Figure S2.2A–C). With a particular interest in plant reproduction and ovule development, we observed that proteins associated with "developmental process" (GO: 0032502), "reproduction" (GO:0000003) and "growth" (GO:0040007) accounted for 8.0% of all annotated genes, as shown in Figure S2.2A. Noticeably, during the BLAST step, only the most informative blast hit was retained to describe each gene. Table S2.3 provides an overall description of the basic characteristics for these proteins, including their known functions, GO annotations, Pfam family and

domain information as well as gene expression levels among different ovule conditions in *S. chacoense*.

2.3.5 Identification of ES-dependent genes

To identify ES-expressed genes in *S. chacoense*, gene expression profiles were compared between wild-type and *Scfrk1* ovules, which entirely lack an ES. As shown in Figure S2.3A, *Scfrk1* ovules exhibit a highly distinguishable expression profile, in which 818 genes were significantly downregulated as compared to wild-type ovules, under a 2-fold cut-off (p-value $\leq$ 0.05). Among these 818 ES-dependent genes, 312 genes were abundantly expressed in anthesis ovules as compared to leaf under a 5-fold cut-off (Figure 2.2 and Table S2.4).

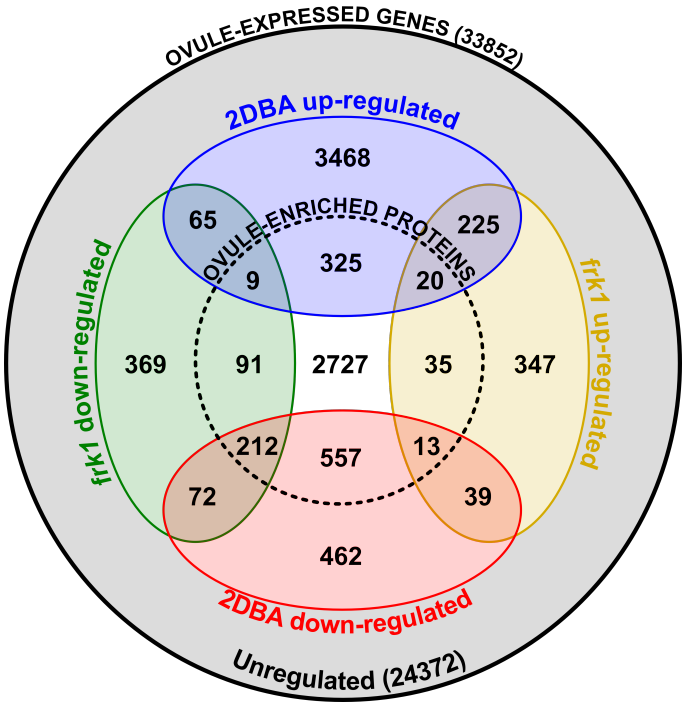


Figure 2.2 Euler diagram showing overlaps between gene sets differentially regulated in *frk1*, 2DBA ovules and leaf, as compared to anthesis ovules. Differentially regulated genes are defined based on a 2-fold cut-off and *p*-value ≤ 0.05 in the comparison between three ovule conditions. Ovule-enriched genes refer to the ones upregulated in anthesis ovules relative to leaf by 5-fold cut-off and *p*-value ≤ 0.05, as indicated by the dotted concentric circle.

GO enrichment analysis was performed for *Scfrk1*-downregulated genes against *Scfrk1*-unregulated genes to determine if certain functional categories are enriched in the ES-dependent gene list. As summarized in Table S2.5, genes associated with "regulation of protein localization to cell surface" were exclusively present in the former, indicating this GO-term (GO:2000008) was significantly enriched in the ES (2-fold cut-off, *p*-value ≤ 0.05). These genes (ScOT_32048, ScOT_36982 and ScOT_36991) encode for egg cell-secreted EC1-like peptides as recognized by their two signature motifs (S1 and S2) and are required for gamete fusion prior to double fertilization (Sprunck, Rademacher et al. 2012). Other GO-terms associated with polysaccharide catabolic process (GO:0000272), cell wall modification (GO:0042545) and regulation of hydrolase activity (GO:0051336) were also significantly enriched in ES-dependent genes.

2.3.6 Developmentally Regulated Genes During ES Transition From FG6 to FG7 Stage

In 2DBA immature ovules, though the overall size is not significantly different from a mature ovule, the majority of the ovules are at FG6 stage. To identify genes essential to the ES transition from FG6 to FG7, gene expression profiles were compared between mature and slightly immature ovules at 2DBA (Figure S2.3B), in which 1355 genes were upregulated in mature ovules compared to 2DBA ovules, thus being potentially associated with polar nuclei fusion, synergid cell maturation and antipodal cells degeneration. Of these, 782 genes were ovule-enriched as seen in their abundant expression in anthesis ovules as compared to leaf as determined by a 5-fold threshold (Figure 2.2 and Table S2.4).

Among developmentally regulated genes in *S. chacoense* ovules, *ScOT_03461* encodes an ortholog to the cyclic nucleotide-gated ion channel 7 and 8 as determined by the phylogenetic distances between ScOT_03461 and 20 members of the CNGC family in *Arabidopsis* (data not shown). Both AtCNGC7 and 8 are essential to PT growth. Both are highly expressed in pollen and, to a lessor extent, in flowers (Tunc-Ozdemir, Rato et al. 2013). As expected from its expression in *A. thaliana*, ScOT_03461 was also weakly expressed in *S. chacoense* mature ovules. Nonetheless, it was 5-fold more expressed in mature vs. 2DBA ovules and absent in leaf, suggesting a FG localization, with its low expression in mature ovules possibly associated with the dilution effect from ovule sporophytic tissues. Another gene expressed during the 2DBA to anthesis transition is *ScOT_26781*. It encodes a ScRALF2 peptide, which was previously isolated from a fertilized ovule cDNA library in *S. chacoense* (Germain, Chevalier et al. 2005). Up-regulation of *ScOT_26781* from 2DBA ovules to mature ones suggests its involvement in late stages of ES maturation.

GO enrichment analysis revealed that genes associated with isoflavonoid biosynthesis process (GO:0009717) were enriched in developmentally regulated gene list, as compared to unregulated genes during the 2DBA to anthesis transition (Table S2.6), suggestive of their special involvement in ovule development. *ScOT_09522* encodes an

isoflavone reductase (IFR), which shares 96.2% of nucleotide identity to the IFR-like *CP100* gene in *S. tuberosum*, whose expression is significantly enhanced upon pollination (van Eldik, Ruiter et al. 1997). In accordance with the developmental regulation feature of ScOT_09522, interestingly, an IFR-like protein OSP130 sharing 67% of amino acid identity to ScOT_09522, was previously found in *S. chacoense* ovule secretome, which was upregulated in anthesis ovule exudates relative to the ones in 2DBA ovules (Liu, Joly et al. 2015).

According to differential gene expression analyses conducted between three ovule conditions, a subset of 284 genes was found abundantly expressed in wild-type mature ovules, while concurrently under-represented in *Scfrk1* and 2DBA ovules, as summarized in Figure 2.2 and Table S2.4. Of these, 74.6% were ovule-enriched when compared to their expression in leaf. Due to the absence of an ES in *frk1* ovules and a FG6 stage ES in slightly immature ovules 2DBA, these overlapping 284 genes were potentially expressed in the ES and associated with late ES maturation (FG6 to FG7), a stage crucial to successful pollen–pistil interactions. Previously, we have shown that neither *Scfrk1* nor 2DBA ovules attract semi *in vivo* grown PTs (Lafleur, Kapfer et al. 2015, Liu, Joly et al. 2015). Consistently, *in vivo* pollination confirmed such guidance defect in *Scfrk1* mutant, where wild-type PTs grew randomly on *Scfrk1* ovules using all available surfaces without entering the micropyle, as shown in both aniline blue staining and SEM images (Figure 2.3). As a result, this gene pool is not only important for ES development but also provides candidates mediating PT guidance processes.

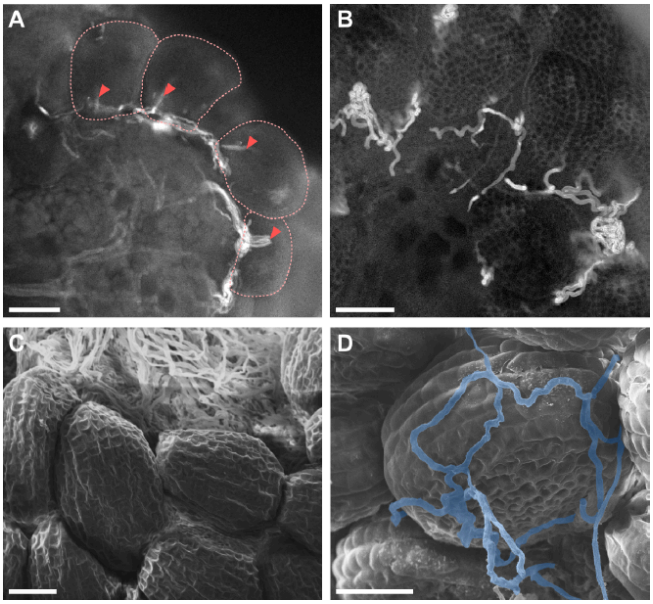


Figure 2.3 PT guidance in wild-type and *frk1* plants. Upper panels are squashes of pollinated pistils (pericarp removed). PTs were stained with aniline blue. (A) PT behavior near wild-type ovules. Dotted red lines indicate the contour of each single ovule. Red arrowhead indicates PTs that turned sharply to target the ovule for double fertilization. (B) PT behavior near *frk1* ovules. PTs migrated randomly on the surface of ovules. Lower panels are SEM images of *in vivo* pollinated pistils. (C) PT behavior near wild-type ovules showing PTs elongated along the placental epithelium, with the micropyle facing the placental side of the ovule. (D) PT behavior near *frk1* ovules. PTs that elongated randomly on the surface of the ovule were colored in blue. Scale bar = 100 μm.

2.3.7 Solanaceous Species Encode Highly Diversified Cysteine-Rich Peptides (CRPs)

CRPs are signaling molecules implicated in numerous pollen–pistil interactions (Higashiyama 2010, Marshall, Costa et al. 2011) and are particularly well represented in the FG of the ovule (Jones-Rhoades, Borevitz et al. 2007, Punwani, Rabiger et al. 2008). In the *S. chacoense* transcriptome, 59 ovule-enriched CRPs were identified, and were abundantly expressed in anthesis ovules when compared to leaf tissues (5-fold cut-off, *p*-value $\leq$ 0.05). All these CRPs bore a N-terminal signal peptide. Through sequence alignment, ScCRPs are classified into various subgroups, including defensin-like (DEFL) proteins, lipid-transfer proteins (LTP), early culture abundant 1 (ECA1) gametogenesis-related proteins, EC1, RALF, and thionin-like proteins, *etc*. Interestingly, novel CRP motifs that were not identified in any previously defined CRP subgroups were also uncovered, accounting for 22% of the total number of CRPs identified. Table S2.7 summarized the classification of ScCRPs and extracted the cysteine spacing pattern for each subgroup. According to differential gene expression analyses, 41% of ovule-enriched ScCRPs were simultaneously under-represented in ES-devoid *Scfrk1* and slightly immature ovules at 2DBA, suggestive of their special involvement in ES maturation and PT guidance processes.

Some CRPs have been shown to be highly polymorphic among species. By calculating the rate of non-synonymous substitutions vs. synonymous substitutions (*dN/dS*) between orthologous CRPs, several CRPs under positive selection have been identified, indicating rapid evolution among species (Kanaoka, Kawano et al. 2011, Takeuchi and Higashiyama 2012). To uncover species-specific CRPs in solanaceous species, an orthology survey was conducted in close relatives *S. gandarillasii* and *S. tarijense* (here denoted as Sgan and Star, respectively). For 59 ScCRPs, reciprocal best BLAST hits method identified 27 SganCRP orthologs and 28 StarCRP orthologs in the *S. gandarillasii* and *S. tarijense de novo* assembled ovule transcriptome, respectively (Table S2.7). Interestingly, 50% of the StarCRP orthologs were under positive selection with a *dN/dS* ratio above 1, whereas none of the SganCRP orthologs was. Interestingly, *S. gandarillasii* is phylogenetically closer to *S. chacoense* than *S. tarijense* (Rodríguez and

Spooner 2009). This is also vindicated by the absence of PT guidance from *S. chacoense* PTs toward *S. tarijense* ovules, whereas *S. gandarillasii* ovules can attract *S. chacoense* PTs (Figure 2.4). These results undoubtedly suggest a positive corelation between species-specificity of PT guidance and their taxonomic distance.

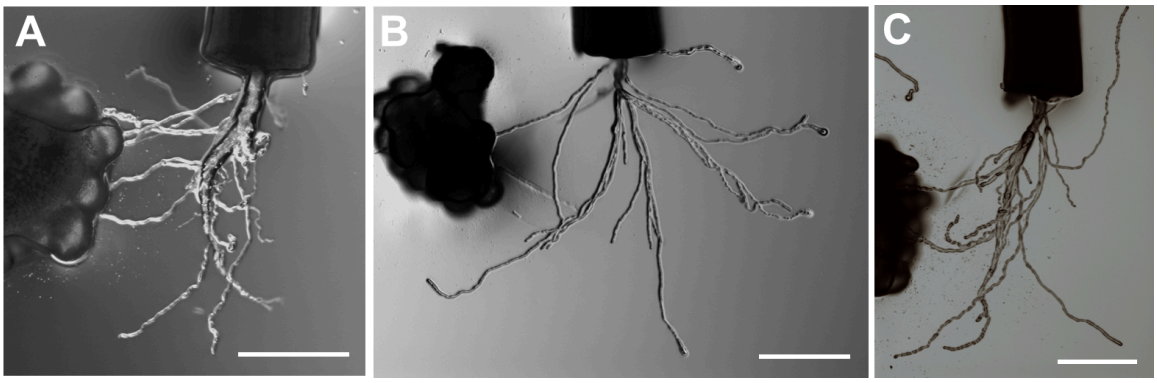


Figure 2.4 Chemotropism of *S. chacoense* PTs toward (A) *S. chacoense* ovule clusters (anthesis) (B) *S. tarijense* ovule clusters (anthesis) and (C) *S. gandarillasii* ovule clusters (anthesis). Scale bars = 400 μm. (C) Schematic depiction of semi *in vivo* PT guidance assay in *S. chacoense*. The angles of PT curvature toward ovules from different species were compared by the Kolmogorov–Smirnov test (KS). Results obtained from KS test indicated *S. tarijense* ovules induced significantly different PT behavior (p = 0.00) as compared to *S. chacoense* ovules, whereas *S. gandarillasii* ovules induced similar PT attracting behavior to *S. chacoense* ovules (p = 0.69).

Furthermore, the divergence of CRPs between *S. chacoense* and S. *tarijense* also indicates a rapid evolution of CRPs in solanaceous species ovules. Figure 2.5 exemplified the most represented, ovule-enriched CRP subgroups in solanaceous species, classified based on their cysteine patterns, sequence similarity and gene expression levels in ovules and leaf.

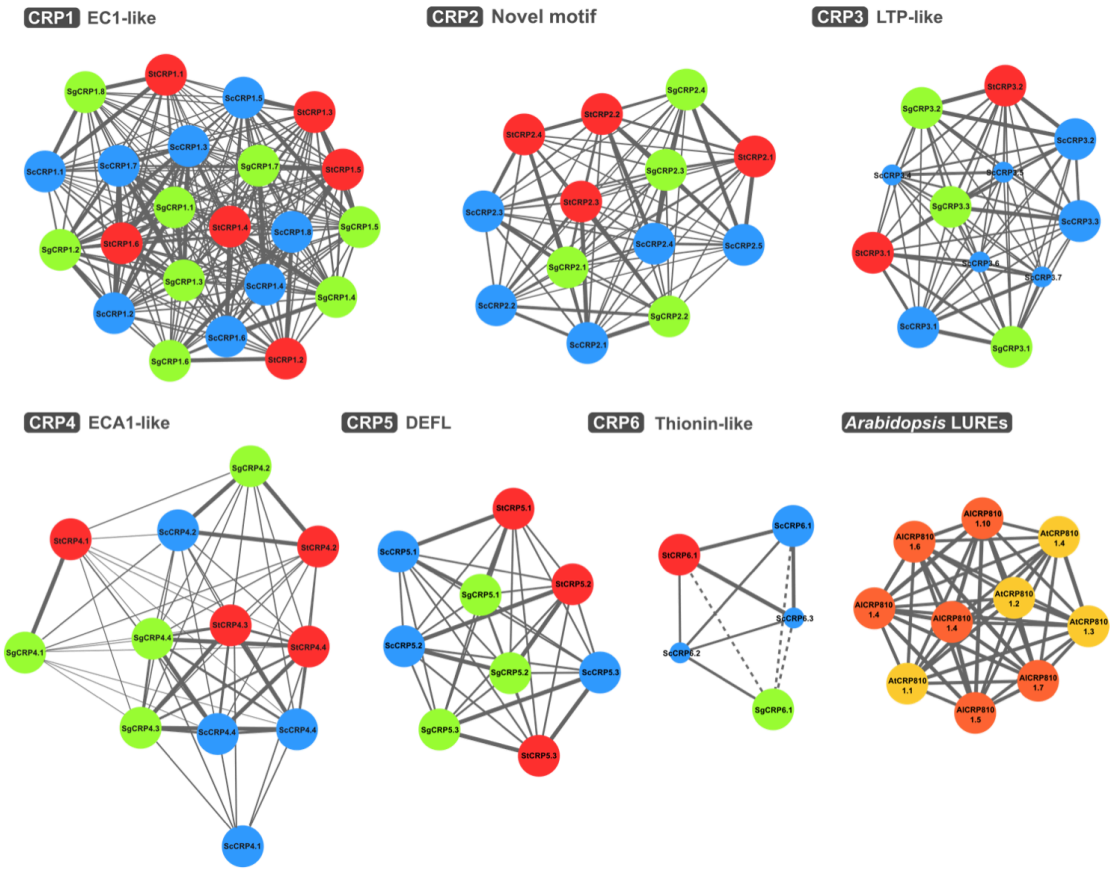


Figure 2.5 Classification of ovule-enriched CRP subgroups in the ovule transcriptome of *S. chacoense* and their orthologs in *S. gandarillasii* and *S. tarijensi*. Divergent LURE-peptides in *Arabidopsis* were used as outgroups. Each node represents a CRP, belonging to *S. chacoense* (blue), *S. gandarillasii* (green), *S. tarijensi* (red), *A. thaliana* (light orange) and *A. lyrata* (dark orange). Larger node size indicates the downregulation of corresponding CRP in both *frk1* and 2DBA ovules (2-fold cut-off, p-value ≤ 0.05) and leaf (5-fold cut-off, p-value ≤ 0.05). Smaller node size indicates the downregulation in leaf only. Edge length and width indicates amino acid sequence identity between two CRPs.

2.3.8 Divergent Molecules Are Recruited From the ES of *S. chacoense* and *A. thaliana* To Accomplish Reproductive Processes

So far, 113 genes in *Arabidopsis* have been experimentally confirmed for their cell-specific expression pattern in the FG of the ovule (Drews, Wang et al. 2011). Their putative orthologs were queried in the *S. chacoense* transcriptome, more specifically in ovule-expressed genes (Table S2.8). Using an *E*-value cut-off of e^{-10}, 49 of the 113 *Arabidopsis* genes shared putative orthologs in *S. chacoense* ovules, of which 55% showed expression level of anthesis ovules/leaf ≥ 2-fold and 24% showed expression level of anthesis ovules/leaf ≥ 5-fold (p-value ≤ 0.05), indicating ovule-enriched expression of known *Arabidopsis* ES genes in *S. chacoense* ovules. Among the 64 genes with no close matches, 40 (63%) encoded highly polymorphic CRPs. We then applied a higher stringency cut-off (*E*-value e^{-10} with a minimum of 50% amino acid sequence identity and 50% coverage) to this orthology search; putative orthologs were retrieved from only 23 ES-expressed genes (20%) in *Arabidopsis* (Table S2.8). This result suggests that a portion of the ES-expressed genes is conserved between *S. chacoense* and *A. thaliana* as judged by their high sequence similarity. However, the majority of ES-expressed genes are largely divergent in two species and such difference was not restricted to the polymorphic CRPs family.

In order to compare the two FG profiles on a broader scale, ES-expressed candidates in *S. chacoense* were compared to their well studied counterparts in *Arabidopsis* using equivalent ES-devoid mutant ovules, including *determinant infertile1* (*dif1*) (Jones-Rhoades, Borevitz et al. 2007, Steffen, Kang et al. 2007), *coatlicue* (*coa*) (Johnston, Meier et al. 2007) and *sporocyteless/nozzle* (*spl/nzz*) (Yu, Hogan et al. 2005, Sanchez-Leon, Arteaga-Vazquez et al. 2012) background. Additionally, genes expressed in individual cell types of the ES in *Arabidopsis* were also included (Wuest, Vijverberg et al. 2010).

To allow divergent sequences to be identified in the orthology search, genes displaying a minimum identity of 30% based on a minimum coverage of 60% in the alignment are reported in Table S2.9. Ninety (90) genes were collectively identified in the FG profile of

the two species, accounting for 8.3% of ES-dependent genes in *S. chacoense*. These included a leucine rich-repeat receptor-like kinase ERECTA-like 1 (ERL1) (ScOT_32446) involved in signal transduction during organ growth and flower development (Shpak, Berthiaume et al. 2004), a MADS box transcription factor AGL80 (ScOT_40704) required for central cell differentiation (Portereiko, Lloyd et al. 2006, Bemer, Wolters-Arts et al. 2008, Steffen, Kang et al. 2008), and a γ-amino butyric acid (GABA) transaminase 2 (ScOT_10034) required for sporophytic PT guidance by creating an increasing gradient of GABA from the stigma to the micropylar region of the ovules (Palanivelu, Brass et al. 2003). However, the *Arabidopsis* GABA transaminase 2 (POP2) contains a mitochondria-targeting sequence, whereas ScOT_10034 is a leaderless secretory protein as predicted by SecretomeP. Previously ScOT_10034 protein was also detected in the *S. chacoense* ovule secretome (Liu, Joly et al. 2015).

GO enrichment analysis between the two FG profiles revealed that GO-terms associated to double fertilization processes were consistently enriched in the ES of *S. chacoense* and *A. thaliana* (Figure S2.4). Several GO-terms were specifically enriched in the ES of *S. chacoense* only, while being significantly under-represented in the ES of *A. thaliana*, such as glycoside metabolism (GO:0016137), intercellular transport (GO:0010496), activation of MAPK activity (GO:0000187) and γ-amino butyric acid metabolism (GO:0009448). By contrast, some *Arabidopsis* ES-enriched GO-terms were under-represented in *S. chacoense* as well. These include defense response (GO:0009814), gametophyte development (GO:0048229) and organ morpgogenesis (GO:0009887), *etc*. Such difference may stem from the extensive duplication events in *Arabidopsis* genome (Blanc, Barakat et al. 2000). The resulting multigene families, usually accompanied by functional redundancy within each family, significantly enriched certain GO-terms in the FG of *Arabidopsis*. In all, the GO-terms reported in this heat map reveal a large difference of the functional categories between *S. chacoense* and *A. thaliana*, which corresponds well to the high sequence divergence discovered in these two species (Table S2.8 and S2.9).

2.3.9 The Glycoprotein Fraction of Total Ovule Extracts Enhances PT Attraction Efficiency

Post-translational modifications modulate multiple aspects of proteins activity, including half-life and cellular location, as well as interactivity behavior towards other proteins. For secreted proteins, *N*- and *O*-linked glycosylation are common co- and post-translational modifications that participate in several biological processes, such as ligand binding affinity (Haweker, Rips et al. 2010), protein stability (Inoue, Ikawa et al. 2008) and protein folding (Frank, Kaulfurst-Soboll et al. 2008). Of the 59 CRPs enriched in *S. chacoense* ovules, 35 (59%) were predicted to harbor potential *N*- and/or *O*-glycosylation sites (Table S2.7). Since these CRPs were candidates associated to PT guidance processes, and the majority of them being predicted as glycosylated proteins, the relationship between glycosylation and PT guidance was examined.

Protein extracts were obtained from ovules at anthesis stage, which were previously shown to attract growing PTs in *S. chacoense* (Lafleur, Kapfer et al. 2015). Three ovule extracts were tested: total proteins, including both soluble and membrane proteins; soluble proteins; and the glycoprotein fraction purified by Concanavalin A affinity chromatography. Here, PT attraction was evaluated in a two-choice assay system, where ovule extracts and negative control (buffer only) were applied at equal distance from PT exiting point from the style.

Although all three fractions attracted PTs in a two-choice semi *in vivo* assay ($p < 0.001$), attraction percentage increased concomitantly with the enrichment of the glycoprotein fraction by 18%, and was also accompanied by a decrease in the percentage of neutrally responding PTs, as summarized in Table 2.1. Figure 2.6A−B showed PT attraction by the glycoprotein fraction, as contrasted by the neutral response when using extraction buffer in Figure 2.6C.

Table 2.1 Statistics of two-choice PT attraction assay

Category	Total protein	Soluble protein	Glycoprotein
Toward extract	76/151 (50%)	84/149 (56%)	101/148 (68%)
Toward buffer	15/151 (10%)	18/149 (12%)	7/148 (5%)
Neutral	60/151 (40%)	47/149 (32%)	40/148 (27%)
p-value	1.62E-03	8.55E-06	8.064E-05

The guidance response of semi *in vivo* grown PTs was tested in the two-choice assay system, where ovule extracts and negative control (buffer only) were deposited on the PT growth medium at equal distance to the exiting point of PTs. Toward extract category sums the total number of styles that have the majority of PTs grow toward protein extracts. Neutral sums the styles that have the majority of PTs grow symmetrically toward both sides. *p*-value was calculated by Student's *t*-test between "toward extract" and "toward buffer" populations, based on 4 independent semi *in vivo* assays for each protein extract.

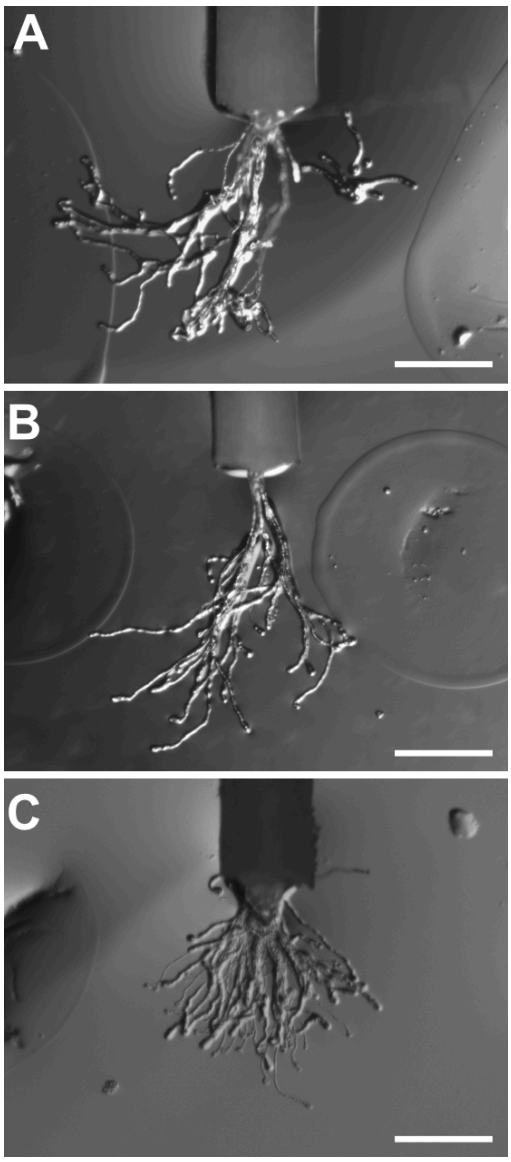


Figure 2.6 Glycoprotein fraction from total ovule extracts attracts PTs in the two-choice assay system. (A, B) Glycoprotein fraction was deposited on the left side of the medium while the extraction buffer used during glycoprotein purification was deposited on the right side, at equal distance to the exiting point of PTs. Protein drops left a circular print on each side of the style. (C) Extraction buffer was deposited on both sides of the style. Scale bar = 400 μm.

2.4 Discussion

2.4.1 All Features In One Hybrid Assembly: Long Reads, Coverage, Optimized Accuracy and Extracted Consensus From Contig Groups to Facilitate High Throughput Bioinformatic Analysis

The *S. chacoense* ovule transcriptome was sequenced using two distinct next-generation sequencing RNA-seq platforms. The 454 RNA-seq technology generated long reads with an average length of 400 nt. In parallel, Illumina HiSeq technology generated short reads of 100 nt. Considering the size of the whole potato genome (844 Mb) (Potato Genome Consortium 2011) and the anthesis ovule trancriptome generated by Illumina sequencing (29 Gb), the sequencing depth can be estimated at 34-fold using the Lander/Waterman equation (Lander and Waterman 1988). The deep sequencing coverage thus provides a massive data amount to characterize and annotate a comprehensive solanaceous ovule transcriptome.

The Hybrid 454/Illumina assembly was previously shown to outperform individual 454 and Illumina assemblies (Hornett and Wheat 2012). Therefore two primary assemblies were merged and curated simultaneously following a curation protocol designed to correct chimeras and frameshifts originating from *de novo* assembly (Liu, Joly et al. 2015). Merging of two next-generation sequencing RNA-seq assemblies enabled us to exploit both the long reads from 454 data and the sequencing depth from Illumina data. For example, of the 59 ovule-enriched CRPs (Table S2.7), 11 CRPs were acquired through 454 RNA-seq and 14 CRPs were acquired through Illumina RNA-seq, while the majority of 34 CRPs were obtained from the hybrid assembly, demonstrating its robust performance over individual ones.

Following this protocol, 15% of the assembled contigs were detected as chimeric or frameshifted sequences. These were curated to produce a high-quality ovule transcriptome, eliminating chimeras and fixing frameshifts. Contig groups were then created. Contig sequences sharing high sequence similarity (≥ 95% sequence identity over ≥ 80% sequence coverage) were clustered and a consensus sequence was extracted

from each contig group. This largely condensed the massive data volume and enabled the generation of a concise CDS sequence database for high-throughput annotations and bioinformatic predictions (Table S2.3). Additionally, this curation pipeline (Joly and Matton, unpublished) can be extended readily to other transcriptome studies, laying a solid ground for analyzing next-generation sequencing data in non-model species.

2.4.2 Solanaceae/Brassicaceae Split Resulted In Highly Divergent Reproductive Biology

Following an exhaustive orthology search for conserved ES-dependent genes in *S. chacoense* and *A. thaliana*, it was shown that only 8.3% of the ES-dependent gene list from *S. chacoense* were detected in *A. thaliana*, based on a minimum identity of 30% and coverage of 60% on the query sequence (Table S2.9). Furthermore, the heat map created to present the GO enrichment results between these two FG-expressed genes profiles revealed that various GO-terms associated with a vast array of biological processes were differentially enriched in *S. chacoense* compared to *A. thaliana* (Figure S2.4), thus supporting the high divergence found in the ES-gene pool.

This result can be partially explained by the prevalence of highly polymorphic CRPs in the FG. By investigating the CRPs in *S. chacoense* ovules, 46 CRPs comprising six CRP subgroups were found, including DEFL, LTP, ECA1, EC1, RALF, and thionin-like subgroups. An additional 13 CRPs contained novel CRP motifs unique to solanaceous species (Figure 2.5 and Table S2.7).

In addition, such a difference between the FG profiles from these two organisms could be related to the pleotropic effects stemming from the use of different mutants that produce ES-less ovules (*FRK1*, *DIF1*, *SPL*, *COA*) for genetic subtraction. This observation was supported by the rather limited overlaps between FG-expressed genes in *Arabidopsis* alone across multiple studies, in which only 10 genes (out of 2729) were detected among all four data sets (Sanchez-Leon, Arteaga-Vazquez et al. 2012). Moreover, different sequencing platforms were adopted in these studies, including ATH1 microarray (Johnston, Meier et al. 2007, Steffen, Kang et al. 2007), tiling array (Jones-Rhoades,

Borevitz et al. 2007), and massively parallel signature sequencing (Sanchez-Leon, Arteaga-Vazquez et al. 2012) for *Arabidopsis* and 454 pyrosequencing and Illumina platforms in the current study.

Yet, another fundamental reason for the divergence of ES-genes in the two organisms may be traced back to their evolutionary history. *S. chacoense* and *A. thaliana* belong to two different families, i.e. the Solanaceae and Brassicaceae, respectively that diverged around 112-156 million years ago (Yang, Lai et al. 1999). Consistent with this, 40 CRPs encoding domain of unknown function (DUF784) in *A. thaliana* were downregulated in *dif1 ovules* and 37 of them required the expression of the *MYB98*, a key transcription factor regulating synergid cell differentiation and PT guidance (Jones-Rhoades, Borevitz et al. 2007). In contrast, only one DUF784 protein (ScOT_28214) was downregulated in *Scfrk1* ovules and the *MYB98* transcription factor was not detected in three *Solanum* species (*S. chacoense*, *S. tuberosum* and *S. lycopersicum*).

Apart from FG profiles, other major differences exist between the two lineages. For example, in order to reject self-pollination, Brassicaceae species use a sporophytic self-incompatibility system that involves the recognition of the male determinant S-locus cysteine-rich protein (SCR) (Takayama, Shiba et al. 2000) and the female determinant S-locus receptor kinase (SRK) (Kachroo, Nasrallah et al. 2002). In contrast, Solanaceae self-incompatibility is of the gametophytic type, which involves the pistil-produced allele-specific S-RNases that enter the PTs, and a counteracting array of pollen-expressed F-box proteins (McClure and Franklin-Tong 2006). Similarly, a recent investigation on the mitogen-activated protein kinase kinase kinase (MAPKKK, or MEKK) family revealed an expansion of this family in *S. chacoense* (*FRK1-6*) as well as in other Solanaceae species like potato and tomato, compared to *A. thaliana* (*MPKKK19, 20 and 21*) (Daigle and Matton, 2015). So far, *ScFRK1* (Lafleur, Kapfer et al. 2015), *ScFRK2* (Gray-Mitsumune, O'Brien et al. 2006, O'Brien, Gray-Mitsumune et al. 2007) and *ScFRK3* (Daigle and Matton, unpublished) have been shown to be involved in pollen and ovule development, while no reproductive roles have yet been associated with their closest counterparts in *Arabidopsis*, suggesting a clear functional divergence. Collectively, these data indicate that the gene population in the FG of *S. chacoense* is

highly divergent from that of *A. thaliana* and suggest *S. chacoense* may employ a solanaceous-specific reproductive biology.

2.4.3 Glycoprotein Enrichment Enhances PT Attraction

A chemotactic transmitting tissue-specific (TTS) glycoprotein that attracts semi *in vivo* grown PTs and that supports PT growth *in vivo* by providing an increasing gradient of glycosylation from the stigma to the ovarian end of the style was previously isolated from the style of *Nicotiana tabacum* (Cheung, Wang et al. 1995, Wu, Wang et al. 1995). Similarly, the TTS from *Nicotiana alata* induces chemotropism of semi *in vivo* grown PTs (Wu, Wong et al. 2000). TTS belongs to the arabinogalactan protein (AGP) family that is hydroxyproline-rich and massively *O*-glycosylated on hydroxyproline residues. Using specific monoclonal antibodies against the carbohydrate moiety, AGPs were detected in the synergid cell, the filiform apparatus and the micropylar region of the ovule in *Arabidopsis*, a localization closely associated with PT guidance (Coimbra, Almeida et al. 2007). Glycosylated-TTS proteins are key molecules that induce PT chemotropism in the pre-ovular stage. Yet, a relationship between glycosylation and ovular PT guidance has not been observed. Our results strongly suggests that increasing purification of the glycoprotein fraction from anthesis ovules modified PT chemotactic behavior and increased the attraction efficiency as compared to non-enriched glycoprotein fractions (Figure 2.6 and Table 2.1).

Recently, two genes involved in protein *N*-glycosylation in the endoplasmic reticulum, namely *TURAN* and *EVAN*, were shown to be required for successful PT reception in *Arabidopsis* (Lindner, Kessler et al. 2015). This study showed the importance of *N*-glycosylation in PT reception and thus suggests a dual recognition system, where the interaction between both protein backbone and glycosyl residues of the male and female gametophytes is required for PT reception. This dual recognition can be adopted during ovular PT guidance by recruiting glycoproteins as attracting molecules. Though no glycosylation modification has been reported for micropylar guidance attractants in *Torenia* (Okuda, Tsutsui et al. 2009, Kanaoka, Kawano et al. 2011), *Arabidopsis* (Takeuchi and Higashiyama 2012) or maize (Marton, Cordts et al. 2005), this post-

translational modification could be adopted by other organisms during micropylar guidance or participate in other guidance mechanisms, e.g. in long-distance guidance.

2.5 Supplementary Data

Figure S2.1 Gene expression correlations between three biological replicates in anthesis ovules (blue), frk1 ovules (orange), 2DBA ovules (red) and leaf (green) obtained from RNA-seq data. Each dot represents the expression level of a gene measured in counts per million reads (CPM). The square of Pearson correlation coefficient is shown in each scatter plot.

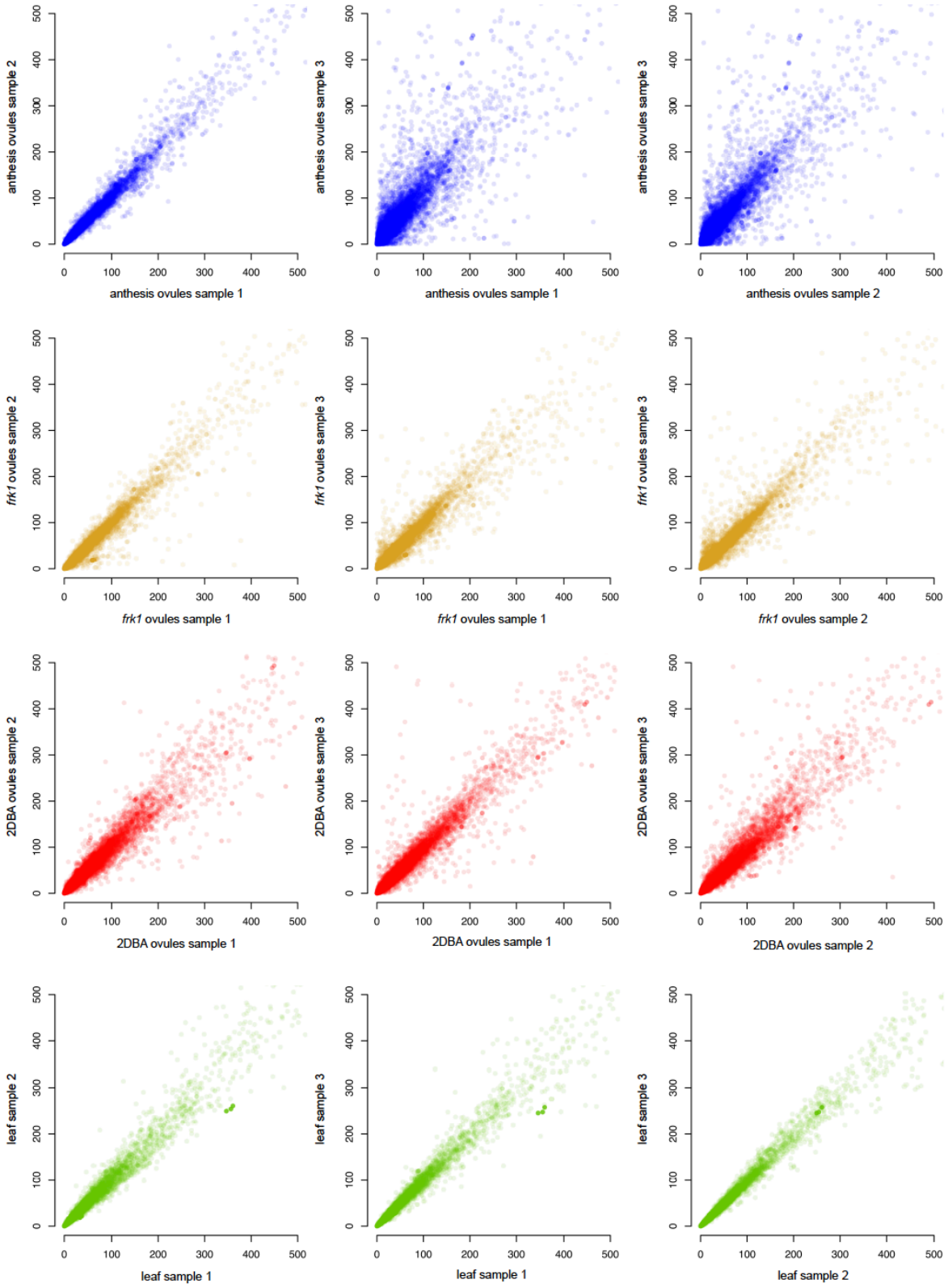

anthesis ovules sample 2
anthesis ovules sample 1
anthesis ovules sample 3
anthesis ovules sample 1
anthesis ovules sample 3
anthesis ovules sample 2
frk1 ovules sample 2
frk1 ovules sample 1
frk1 ovules sample 3
frk1 ovules sample 1
frk1 ovules sample 3
frk1 ovules sample 2
2DBA ovules sample 2
2DBA ovules sample 1
2DBA ovules sample 3
2DBA ovules sample 1
2DBA ovules sample 3
2DBA ovules sample 2
leaf sample 2
leaf sample 1
leaf sample 3
leaf sample 1
leaf sample 3
leaf sample 2
0
100
200
300
400
500

Figure S2.2 Bar plots showing the number of proteins associated with GO terms. (A) Biological process (B) Molecular function (C) Cellular component

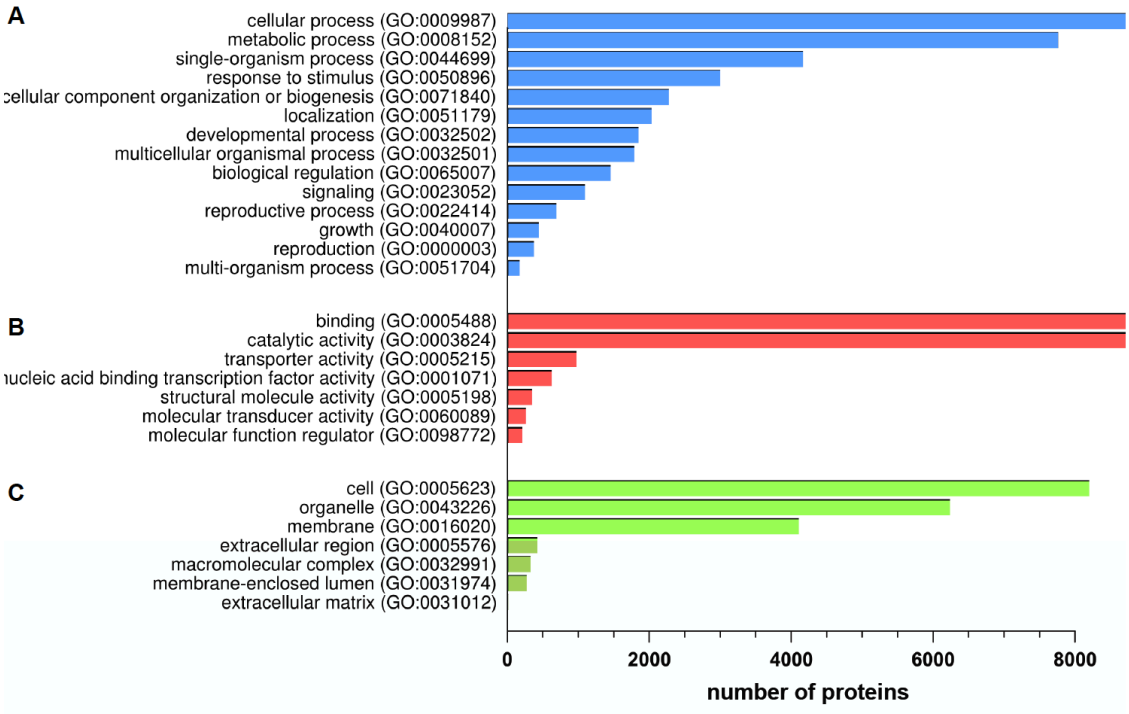


Figure S2.3 Comparison of gene expression profiles between different ovule conditions by MA plots. (A) Anthesis vs. *frk1* ovules (B) Anthesis vs. 2DBA ovules. Each dot represents a gene in the ovule transcriptome. The x-axis indicates the expression level of each gene in anthesis ovules. The y-axis indicates the corresponding fold-changes. Red and blue dots represent significantly up and downregulated genes in each comparison (*p*-value ≤ 0.05, |fold-change| ≥2). Dotted lines mark the fold-change threshold (2-fold) in differential gene expression analyses. Dark gray dots represent significantly regulated genes (*p*-value ≤ 0.05, 1≤ |fold-change| ≤2). Light gray dots represent unregulated genes (*p*-value ≥ 0.05) between two expression profiles.

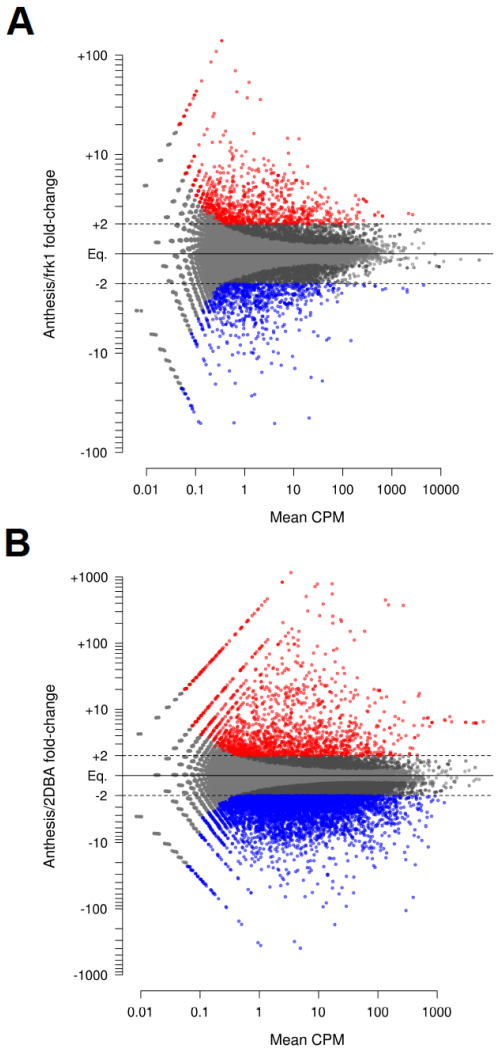
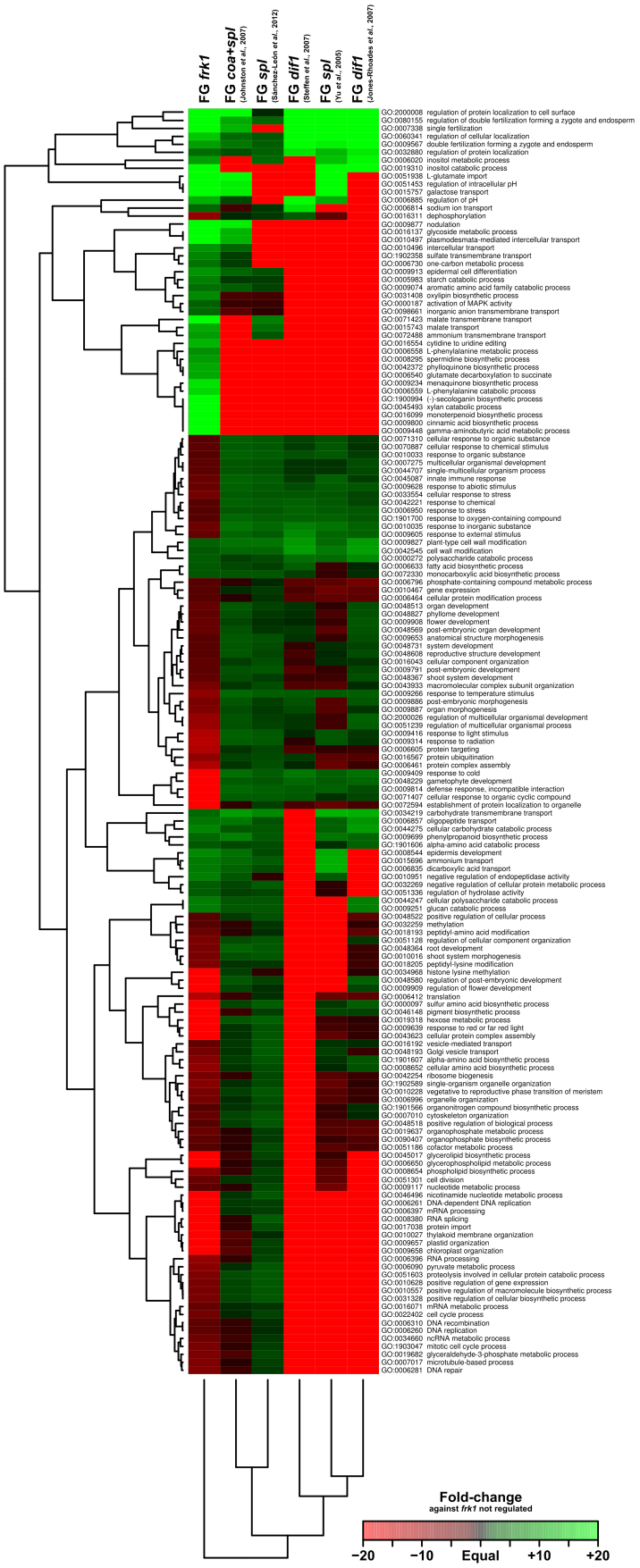


Figure S2.4 GO enrichment analysis of ES-dependent genes in *S. chacoense* (*frk1* mutant background) and *A. thaliana* (*dif1*, *spl*, and *coa* mutant background), compared to unregulated genes in *frk1* ovules. All GO-terms that are significantly regulated (2-fold cut-off, p-value ≤ 0.05) in the ES of *S. chacoense* were included in the heat map.


FG frk1
FG coa+spl (Johnston et al., 2007)
FG spl (Sánchez-León et al., 2012)
FG dif1 (Steffen et al., 2007)
FG spl (Yu et al., 2005)
FG dif1 (Jones-Rhoades et al., 2007)
GO:2000008 regulation of protein localization to cell surface
GO:0080155 regulation of double fertilization forming a zygote and endosperm
GO:0007338 single fertilization
GO:0060341 regulation of cellular localization
GO:0009567 double fertilization forming a zygote and endosperm
GO:0032880 regulation of protein localization
GO:0006020 inositol metabolic process
GO:0019310 inositol catabolic process
GO:0051938 L-glutamate import
GO:0051453 regulation of intracellular pH
GO:0015757 galactose transport
GO:0006885 regulation of pH
GO:0006814 sodium ion transport
GO:0016311 dephosphorylation
GO:0009877 nodulation
GO:0016137 glycoside metabolic process
GO:0010497 plasmodesmata-mediated intercellular transport
GO:0010496 intercellular transport
GO:1902358 sulfate transmembrane transport
GO:0006730 one-carbon metabolic process
GO:0009913 epidermal cell differentiation
GO:0005983 starch catabolic process
GO:0009074 aromatic amino acid family catabolic process
GO:0031408 oxylipin biosynthetic process
GO:0000187 activation of MAPK activity
GO:0098661 inorganic anion transmembrane transport
GO:0071423 malate transmembrane transport
GO:0015743 malate transport
GO:0072488 ammonium transmembrane transport
GO:0016554 cytidine to uridine editing
GO:0006558 L-phenylalanine metabolic process
GO:0008295 spermidine biosynthetic process
GO:0042372 phylloquinone biosynthetic process
GO:0006540 glutamate decarboxylation to succinate
GO:0009234 menaquinone biosynthetic process
GO:0006559 L-phenylalanine catabolic process
GO:1900994 (-)-secologanin biosynthetic process
GO:0045493 xylan catabolic process
GO:0016099 monoterpenoid biosynthetic process
GO:0009800 cinnamic acid biosynthetic process
GO:0009448 gamma-aminobutyric acid metabolic process
GO:0071310 cellular response to organic substance
GO:0070887 cellular response to chemical stimulus
GO:0010033 response to organic substance
GO:0007275 multicellular organismal development
GO:0044707 single-multicellular organism process
GO:0045087 innate immune response
GO:0009628 response to abiotic stimulus
GO:0033554 cellular response to stress
GO:0042221 response to chemical
GO:0006950 response to stress
GO:1901700 response to oxygen-containing compound
GO:0010035 response to inorganic substance
GO:0009605 response to external stimulus
GO:0009827 plant-type cell wall modification
GO:0042545 cell wall modification
GO:0000272 polysaccharide catabolic process
GO:0006633 fatty acid biosynthetic process
GO:0072330 monocarboxylic acid biosynthetic process
GO:0006796 phosphate-containing compound metabolic process
GO:0010467 gene expression
GO:0006464 cellular protein modification process
GO:0048513 organ development
GO:0048827 phyllome development
GO:0009908 flower development
GO:0048569 post-embryonic organ development
GO:0009653 anatomical structure morphogenesis
GO:0048731 system development
GO:0048608 reproductive structure development
GO:0016043 cellular component organization
GO:0009791 post-embryonic development
GO:0048367 shoot system development
GO:0043933 macromolecular complex subunit organization
GO:0009266 response to temperature stimulus
GO:0009886 post-embryonic morphogenesis
GO:0009887 organ morphogenesis
GO:2000026 regulation of multicellular organismal development
GO:0051239 regulation of multicellular organismal process
GO:0009416 response to light stimulus
GO:0009314 response to radiation
GO:0006605 protein targeting
GO:0016567 protein ubiquitination
GO:0006461 protein complex assembly
GO:0009409 response to cold
GO:0048229 gametophyte development
GO:0009814 defense response, incompatible interaction
GO:0071407 cellular response to organic cyclic compound
GO:0072594 establishment of protein localization to organelle
GO:0034219 carbohydrate transmembrane transport
GO:0006857 oligopeptide transport
GO:0044275 cellular carbohydrate catabolic process
GO:0009699 phenylpropanoid biosynthetic process
GO:1901606 alpha-amino acid catabolic process
GO:0008544 epidermis development
GO:0015696 ammonium transport
GO:0006835 dicarboxylic acid transport
GO:0010951 negative regulation of endopeptidase activity
GO:0032269 negative regulation of cellular protein metabolic process
GO:0051336 regulation of hydrolase activity
GO:0044247 cellular polysaccharide catabolic process
GO:0009251 glucan catabolic process
GO:0048522 positive regulation of cellular process
GO:0032259 methylation
GO:0018193 peptidyl-amino acid modification
GO:0051128 regulation of cellular component organization
GO:0048364 root development
GO:0010016 shoot system morphogenesis
GO:0018205 peptidyl-lysine modification
GO:0034968 histone lysine methylation
GO:0048580 regulation of post-embryonic development
GO:0009909 regulation of flower development
GO:0006412 translation
GO:0000097 sulfur amino acid biosynthetic process
GO:0046148 pigment biosynthetic process
GO:0019318 hexose metabolic process
GO:0009639 response to red or far red light
GO:0043623 cellular protein complex assembly
GO:0016192 vesicle-mediated transport
GO:0048193 Golgi vesicle transport
GO:1901607 alpha-amino acid biosynthetic process
GO:0008652 cellular amino acid biosynthetic process
GO:0042254 ribosome biogenesis
GO:1902589 single-organism organelle organization
GO:0010228 vegetative to reproductive phase transition of meristem
GO:0006996 organelle organization
GO:1901566 organonitrogen compound biosynthetic process
GO:0007010 cytoskeleton organization
GO:0048518 positive regulation of biological process
GO:0019637 organophosphate metabolic process
GO:0090407 organophosphate biosynthetic process
GO:0051186 cofactor metabolic process
GO:0045017 glycerolipid biosynthetic process
GO:0006650 glycerophospholipid metabolic process
GO:0008654 phospholipid biosynthetic process
GO:0051301 cell division
GO:0009117 nucleotide metabolic process
GO:0046496 nicotinamide nucleotide metabolic process
GO:0006261 DNA-dependent DNA replication
GO:0006397 mRNA processing
GO:0008380 RNA splicing
GO:0017038 protein import
GO:0010027 thylakoid membrane organization
GO:0009657 plastid organization
GO:0009658 chloroplast organization
GO:0006396 RNA processing
GO:0006090 pyruvate metabolic process
GO:0051603 proteolysis involved in cellular protein catabolic process
GO:0010628 positive regulation of gene expression
GO:0010557 positive regulation of macromolecule biosynthetic process
GO:0031328 positive regulation of cellular biosynthetic process
GO:0016071 mRNA metabolic process
GO:0022402 cell cycle process
GO:0006310 DNA recombination
GO:0006260 DNA replication
GO:0034660 ncRNA metabolic process
GO:1903047 mitotic cell cycle process
GO:0019682 glyceraldehyde-3-phosphate metabolic process
GO:0007017 microtubule-based process
GO:0006281 DNA repair
Fold-change
against frk1 not regulated
−20
−10
Equal
+10
+20

The supplementary tables are available via a link from:
https://dl.dropboxusercontent.com/u/8321715/Supplementary_Tables_ChapterII.zip

Table S2.1 RT-PCR primers used to validate gene expression predictions obtained from RNA-seq data

Table S2.2 Validation of the assembly through known ovule-expressed genes in *S. chacoense*

Table S2.3 Functional annotation of *S. chacoense* transcritome

Table S2.4 Summary of differentially expressed genes between three ovule conditions and leaf

Table S2.5 GO enrichment analysis on gene sets between ES-dependent (*frk1*-downregulated) and *frk1*-unregulated genes

Table S2.6 GO enrichment analysis on gene sets between two ovule developmental stages

Table S2.7 Classification of CRPs in *S. chacoense* and their orthology survey in other solanaceous species

Table S2.8 Orthology survey of validated ES-expressed genes of *A. thaliana* in *S. chacoense* ovule transcriptome

Table S2.9 Comparison of ES-expressed genes between *A. thaliana* and *S. chacoense*

2.6 Acknowledgments

This work was supported by FRQNT (Fonds de Recherche du Québec – Nature et technologies). Y.L. is the recipient of a Ph. D. fellowship from the China Scholarship Council.

3. The Plant Ovule Secretome: A Different View Toward Pollen–Pistil interactions

Yang Liu[*], Valentin Joly[*], Sonia Dorion, Jean Rivoal, and Daniel P. Matton[§]

Institut de Recherche en Biologie Végétale, Département de Sciences Biologiques, Université de Montréal, 4101 rue Sherbrooke est, Montréal, Québec H1X 2B2, Canada

Author contributions: [*]Y.L. and V.J. contributed equally to this work. Y.L., V.J., and D.P.M. conceived the study and wrote the manuscript. Y.L. and V.J. performed all of the experimental work and bioinformatic analyses, except for purity assessment, which was performed by S.D. and J.R.

[§]To whom correspondence should be addressed.

Published in: *Journal of proteome research*

J. Proteome Res., ***2015***, *14 (11), pp 4763–4775*.

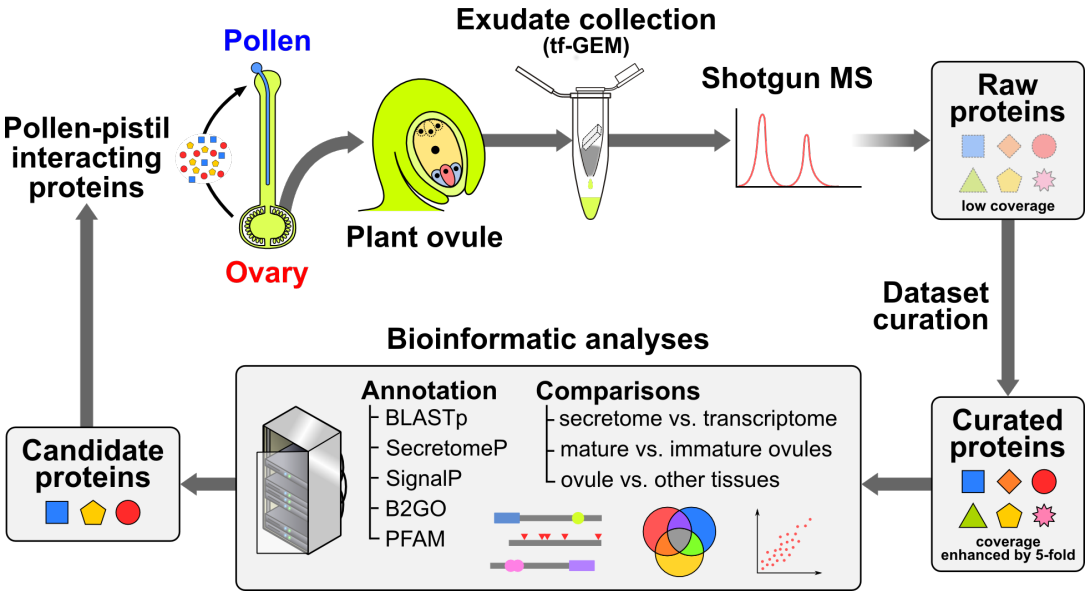
DOI:10.1021/acs.jproteome.5b00618.

Abstract

During plant sexual reproduction, continuous exchange of signals between the pollen and the pistil (stigma, style, and ovary) plays important roles in pollen recognition and selection, establishing breeding barriers and, ultimately, leading to optimal seed set. After navigating through the stigma and the style, pollen tubes (PTs) reach their final destination, the ovule. This ultimate step is also regulated by numerous signals emanating from the embryo sac (ES) of the ovule. These signals encompass a wide variety of molecules, but species-specificity of the pollen−ovule interaction relies mainly on secreted proteins and their receptors. Isolation of candidate genes involved in pollen–pistil interactions has mainly relied on transcriptomic approaches, overlooking potential post-transcriptional regulation. To address this issue, ovule exudates were collected from the wild potato species *Solanum chacoense* using a tissue-free gravity-extraction method (tf-GEM). Combined RNA-seq and mass spectrometry-based proteomics led to the identification of 305 secreted proteins, of which 58% were ovule-specific. Comparative analyses using mature ovules (attracting PTs) and immature ovules (not attracting PTs) revealed that the last maturation step of ES development affected almost half of the ovule secretome. Of 128 upregulated proteins in anthesis stage, 106 were not regulated at the mRNA level, emphasizing the importance of post-transcriptional regulation in reproductive development.

Keywords: pollen–pistil interactions, ovule exudates, secretome, chemoattraction, pollen tube guidance, gel-free nano-LC−MS/MS, label-free quantification

Graphical Abstract



3.1 Introduction

Communication through extracellular signals starts with the synthesis and secretion of signaling molecules from a signaling cell(s). Some molecules may remain bound to the signaling cell and only affect cells in its vicinity. Such a case can be seen in the attachment of sperm to egg cells during fertilization in mammals via the interaction between sperm-tethered membrane protein Izumo (Inoue, Ikawa et al. 2005) and its sole surface receptor Juno (Bianchi, Doe et al. 2014) on the egg cell to achieve proper gamete binding. In most cases, however, signal molecules are secreted, followed by their active transport or diffusion, to where they exert their function. Detection of the signals by the target cell then relies on specific receptor proteins that relay this information inside the cell, where a second signaling phase starts, leading to appropriate downstream responses. One way to address the compendium of secreted proteins that could play roles in extracellular signaling is to analyze the secretome of a specific tissue or organ under various developmental conditions or following biotic or abiotic stresses. The secretome refers to the proteins and peptides that are secreted out of the plasma membrane and, more specifically in plants, to the cell wall and apoplastic fluid. So far, secretomic studies have been reported for some tissues in angiosperms including root cap, seeds, seedlings, fruit pericarps, leaves, and stems *in planta,* as well as in cell suspension cultures (Wen, VanEtten et al. 2007, Goulet, Goulet et al. 2010, Pandey, Rajamani et al. 2010, Gupta, Wardhan et al. 2011, Song, Zhang et al. 2011, Zhou, Bokhari et al. 2011, Floerl, Majcherczyk et al. 2012, Regente, Pinedo et al. 2012, Konozy, Rogniaux et al. 2013). Secretomes of reproductive tissues have also been conducted in tobacco, maize, lily, and olive tree stigmas (Sang, Xu et al. 2012, Rejon, Delalande et al. 2013, Rejon, Delalande et al. 2014) as well as in PTs (Hafidh, Potesil et al. 2014). In gymnosperm species, the secretome of pollination drops was also reported (Prior, Little et al. 2013). These originate from ovular secretion that fills the micropyle of the ovule to enhance pollen capture and facilitate pollen germination.

Depending on the biological context, secreted proteins participate in a wide range of processes, directly or indirectly affecting their neighboring cellular environment. For instance, analysis of the root tip secretome revealed numerous proteins released to the

extracellular space that protect root tips from fungal infection (Wen, VanEtten et al. 2007). To resist low-temperature stress, seedlings were reported to secrete antifreeze molecules in response to low-temperature stress (Gupta and Deswal 2012). Besides protecting against microbial infections and abiotic stresses, plants also use exudates to distinguish dissimilar individuals from their own siblings in the context of plant competition (Badri, De-la-Peña et al. 2012).

Here, we report the first secretome study of a plant ovule, an organ deeply enclosed in the carpel, the female reproductive organ consisting of the stigma, style, and ovary. From pollen landing on the stigma until fertilization, PTs are believed to be precisely guided to target each available ovule (Higashiyama, Kuroiwa et al. 2003). In solanaceous species, the PT path is a continuous tract of specialized cells running from the stigma surface through the stylar transmitting tract that ultimately splits along the ovary septa and merges with the placental epithelium (Figure 3.1A), as showed in *Solanum lycopersicum* (Webb and Williams 1988) and *Nicotiana alata* (Cornish, Pettitt et al. 1987). In the preovular guidance stage, pollen grains germinate and elongate extracellularly in the transmitting tract of the style. After entering the ovary, ovular guidance cues take over to eventually steer the PT toward the micropyle (Figure 3.1B−D), a small opening at the ovule surface through which the PT penetrates to effect double fertilization (Dresselhaus and Franklin-Tong 2013, Higashiyama and Takeuchi 2015). Ovular guidance can be divided into three distinct steps, as depicted in Figure 3.1B. A long-distance activity was suggested to describe ovular signals that change PT trajectory from the intercellular space of the transmitting cells to the septum, associated with a tendency for PTs to emerge from the ovary proximal position to the stylar end (Hulskamp, Schneitz et al. 1995). This phase is illustrated by the phenotype observed in the *Arabidopsis* mutant for the pollen-expressed cation/proton exchangers *CHX21* and *CHX23*. In this mutant, PTs do not respond to ovular cues and fail to exit the transmitting tract (Lu, Chanroj et al. 2011). The second step involves funicular guidance, where PTs adhere to and migrate up the funiculus, a stalk-like structure that attaches the ovule to the placenta, and is followed by a third step, described as micropylar guidance, that targets the PT at close range, within ~100−150 μm to the micropyle (Higashiyama, Kuroiwa et al. 1998, Palanivelu and Preuss 2006, Stewman, Jones-Rhoades et al. 2010). The latter two steps can be further

distinguished on the basis of the phenotype observed in the *Arabidopsis mpk*3/*mpk*6 mutant (Guan, Lu et al. 2014). Here, the majority of mutant PTs exit the transmitting tract, but they either get lost in the septum or take a much longer time to eventually target the ovule, exhibiting a funicular guidance defect *in vivo*. Interestingly, in the *mpk*3/*mpk*6 mutant, micropylar guidance is not affected, as demonstrated by a semi *in vivo* guidance assay, emphasizing the distinct nature of different signals controlling funicular and micropylar guidance, at least in *Arabidopsis*.

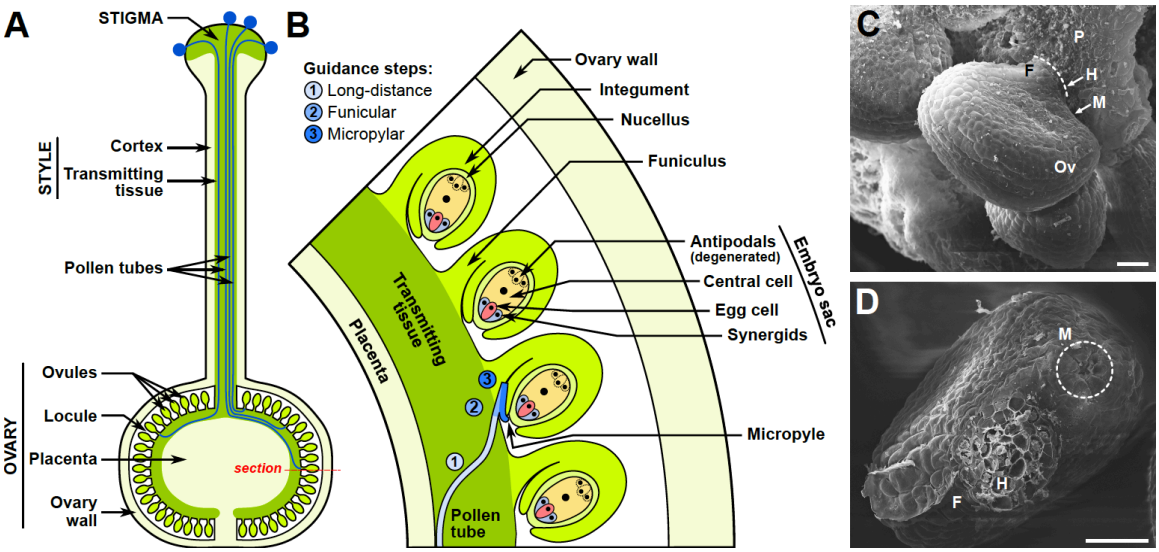


Figure 3.1 Schematic depiction of PT growth in the *S. chacoense* pistil. (A) Longitudinal section of the pistil. Dark green indicates the PT pathway, a continuous cell tract originating from the stigma surface, extending through the style down to the placental epithelium, where it connects to the ovules. At the mature stage of an ovule, the ES consists of an egg cell, two synergid cells, and a central cell, embedded in one layer of integument in *S. chacoense*. (B) Close-up of an ovary transverse section. PT typically approaches the micropyle of an ovule through three guidance steps. Step 1 (long-distance guidance), PT emerges from the intercellular space of the transmitting tract to the placental surface; step 2 (funicular guidance), PT climbs up the funiculus; and step 3 (micropylar guidance), PT navigates toward and into the micropyle. (C) Scanning Electron Microscopy (SEM) image of *S. chacoense* wild-type ovules connected to the placental tissue. The micropylar opening is hidden, facing the placental side. Dotted lines trace the position of the hilum. Noticeably, the particular morphology of *S. chacoense* ovules brings the micropyle in the immediate vicinity of the placental surface, rendering the micropyle directly accessible to emerging PTs. (D) SEM image of a *S. chacoense* ovule detached from the placenta. The dashed circle marks the micropylar opening. F, funiculus; H, hilum; M, micropyle; Ov, ovule; P, placenta. Scale bar = 30 μm.

Among the multistep control points of PT guidance, long-distance guidance activity, although shown to be governed by the ovule, is largely unexplored (Hulskamp, Schneitz et al. 1995). Funicular guidance appears to be associated with the ovule sporophytic tissues (Lausser, Kliwer et al. 2010), as demonstrated by the *pop2* (Palanivelu, Brass et al. 2003), *siz1-2* (Ling, Zhang et al. 2012), and *pdil2*-1 (Wang, Boavida et al. 2008) mutants, where PTs migrate up the funiculus but become lost near the micropyle. Micropylar guidance is controlled by the female gametophyte (Higashiyama, Yabe et al. 2001). Such short-distance attractants have been identified in *Torenia* (Okuda, Tsutsui et al. 2009, Kanaoka, Kawano et al. 2011) and maize (Marton, Cordts et al. 2005) through transcriptomic analysis and in *Arabidopsis* (Takeuchi and Higashiyama 2012) through comparative phylogenetic studies. Collectively, these results suggest that the ovule is an active player in male−female gametophyte communication, guiding PTs to navigate through the ovarian mucilage to achieve double fertilization (Chevalier, Loubert-Hudon et al. 2013, Dresselhaus and Franklin-Tong 2013).

Here, we present the ovule secretome as a different approach to study pollen–pistil interactions. By examining the protein composition of this bioactive environment, we anticipate gaining more insight into pollen−ovule interaction processes, involving PT growth stimulation, competency acquisition, guidance and PT repulsion. To this aim, comparative proteomic analyses were conducted on mature attracting ovules, as well as immature ovules, unable to attract PTs, as shown in semi *in vivo* assays.

3.2 Material and Methods

3.2.1 Plant Materials and Growth Conditions

Solanum chacoense Bitt. individuals were greenhouse-grown under long-day conditions (16 h light/8 h dark). Two genotypes, G4 ($S_{12}S_{14}$ self-incompatibility alleles) and V22 ($S_{11}S_{13}$), were used to perform compatible pollinations: G4 (♀) × V22 (♂).

3.2.2 semi *in vivo* PT Guidance Assay for *S. chacoense*

Semi *in vivo* guidance assays were performed as described previously (Lafleur, Kapfer et al. 2015). In brief, 24 h after pollination (HAP), styles were excised and placed on BK

solid medium (Brewbaker and Kwack 1963), allowing PTs to grow properly when exiting from the style around 30 HAP. Clusters of 5−10 ovules connected by placental tissue were positioned ~700 µm away from the end of the style (equivalent to the radius of the ovary), with a 45° angle. Two developmental stages were tested: mature ovules from flowers at anthesis (ES stage FG7) and immature ovules from 6 to 7 mm flower buds taken two days before anthesis (2DBA, ES stage FG6). Guidance response of emerging PTs to ovule clusters was observed in bright field with an Axio Observer.Z1 microscope equipped with an AxioCam HRm camera (Zeiss). The turning angle of each PT was measured with ImageJ (http://imagej.nih.gov/ij). Angle distribution was compared between conditions using a two-sample Kolmogorov-Smirnov (KS) statistical test ($n > 80$).

3.2.3 Scanning Electron Microscopy (SEM)

S. chacoense G4 ovaries were hand-dissected to remove the pericarp (ovary wall). A group of ovules connected by placental tissues was fixed, dehydrated, and critical-point-dried as described previously (Gray-Mitsumune, O'Brien et al. 2006). Since *S. chacoense* ovules align tightly around the placental tissues, some ovules from the clusters were removed under a stereomicroscope in order to leave enough space for single-ovule observation. The tissues were coated with gold−palladium and viewed in a JEOL JSM-35 SEM.

3.2.4 Collection of Ovule Exudates

The pericarp was removed from the ovary. Ovule clusters connected by placental tissues were aligned on a 0.5% agarose strip (0.5 × 1 cm) and incubated in a humidified chamber for 24 h in darkness at room temperature. Ovules were then taken out with a needle, and the gel strip containing ovule exudates was placed in a 0.5 ml centrifuge tube perforated at the bottom. A small amount of glass wool was placed under the gel piece to allow the flow-through of ovule exudates to the collection tube and to retain the gel matrix in place during centrifugation. The device was then inserted in a 1.5 ml tube and centrifuged 5 min at 5000 rpm. Exudates collected in the flow-through were centrifuged for 10 min at

14400 rpm and 4°C to remove residual impurities. The supernatant was flash frozen in liquid nitrogen and stored at −80 °C.

3.2.5 Total Protein Extraction

For western blot analyses and enzymatic assays, three biological replicates of whole ovules (50 −70 mg fresh weight, FW) were collected from flowers at anthesis, placed on dry ice and stored at – 80 °C until use. Samples were then ground on ice with a pestle in a buffer containing 30 mM Tris·HCl pH 7.5, 5 mM $MgCl_2$, 100 mM KCl, 1 mM EDTA, 1 mM EGTA, 0.1% (v/v) Triton X-100, 10% (v/v) glycerol, 5 mM DTT, 5% (w/v) insoluble PVPP, 5 mM ε-amino caproic acid, 1 mM benzamidine, 1 µg/ml leupeptin and 2 mM PMSF, with a 3:1 ratio (milliliters of extraction buffer per gram of FW). Homogenates were centrifuged 15 min at 12000 *g* at 4 °C. Supernatants were further centrifuged for 5 min. Clarified supernatants were used immediately for enzyme activity measurement. For immunoblot analysis, an aliquot of the supernatant was immediately heat-denatured in SDS sample buffer and kept frozen at –20 °C until use. Proteins were quantified with the Bradford method using bovine serum albumin as a standard (Bradford 1976).

3.2.6 Enzymatic Assay

Three biological replicates of ovule exudates were obtained for enzymatic assays. Lyophilized samples were solubilized in a buffer containing 30 mM Tris·HCl pH 7.5, 5 mM $MgCl_2$, 1 mM EDTA, 1 mM EGTA, 5 mM DTT, 5 mM ε-amino caproic acid, 1 mM benzamidine, 1 µg/ml leupeptin, and 2 mM PMSF. The samples were centrifuged for 5 min at 12000 *g* and 4 °C to eliminate possible insoluble materials. Supernatants from the three samples were pooled in a final volume of 20 µL and used immediately for enzymatic activity assays. Triosephosphate isomerase (TPI) activity assays were performed according to a protocol described previously (Dorion, Parveen et al. 2005). One unit (U) of enzyme activity corresponds to the appearance of the reaction product at a rate of 1 $\mu mol \cdot min^{-1}$. Activity was tested in two independent experiments. Intracellular contamination of exudates was assessed as the percentage of TPI activity present in the exudates over that of the total ovule extract.

3.2.7 SDS-PAGE and Immunoblot Analysis

Three independent ovule samples were obtained to perform western blots. SDS-PAGE analysis was performed on 15% acrylamide gels with 9.0, 3.0, 1.0, 0.3 and 0.1 µg of proteins from ovules at anthesis and 9.0 µg of proteins from ovule exudates. Proteins were then transferred onto a nitrocellulose membrane at 70 V for 60 min. Membranes were incubated with an anti-cytosolic TPI (cTPI) antibody (1/500 dilution) for 1 h at room temperature (Dorion, Parveen et al. 2005). Polypeptides were detected using goat anti-rabbit IgG secondary antibodies (1/10000 dilution) conjugated to alkaline phosphatase (Promega). The reaction was visualized using BCIP (5-bromo-4-chloro-3-indolyl-phosphate) and NBT (nitro-blue tetrazolium) and was allowed to develop for 20 min at room temperature.

3.2.8 RNA Sequencing and *de novo* Assembly

Total RNA was extracted using the TRIzol reagent (Invitrogen) as recommended by the manufacturer. Two next-generation sequencing platforms were used to generate the ovule transcriptome in parallel. First, the 454 GS-FLX Titanium platform was used to perform RNA-seq on ovules at anthesis and at 2DBA. cDNA libraries were constructed for each condition with the rapid library preparation kit (Roche) after an mRNA enrichment step performed with Dynabeads Oligo$(dT)_{25}$ (Invitrogen). Reads were *de novo* assembled with Newbler software (Roche) (Margulies, Egholm et al. 2005). A second RNA-seq was performed with the Illumina HiSeq 2000 platform, on the anthesis sample only. The cDNA library was prepared using the TruSeq cDNA preparation kit (Illumina). A *de novo* assembly was then performed using Trinity software (Grabherr, Haas et al. 2011). Both assemblies were used to generate a reference ovule transcriptome for protein identification.

3.2.9 Quantification of RNA Expression

In order to assess differential gene expression for ovule-secreted proteins, three biological replicates of mRNA samples from anthesis and 2DBA ovules were sequenced on the Illumina platform as described above. Bowtie 2 software was used to align reads to all contigs encoding an ovule-secreted protein (Langmead, Trapnell et al. 2009). RSEM (Li

and Dewey 2011) and edgeR (Robinson, McCarthy et al. 2010) were used to compute expression fold-changes between conditions. Transcripts with a fold-change below –2.0 or above 2.0 ($p \leq 0.05$) were considered to be differentially regulated.

3.2.10 Mass Spectrometry (MS)

One microgram (1 µg) of lyophilized exudates from three biological replicates was trypsin-digested for 8 h at 37 °C. Samples were desalted with Ziptips (Millipore) and separated on a reversed-phase column (150 µm i.d. × 150 mm) equipped with a precolumn (0.3 mm × 5 mm) using a 56 min gradient from 10 to 60% (v/v) acetonitrile in 0.2% (v/v) formic acid and a 600 nL/min flow rate on a 2D-nanoLC system (Eksigent). Liquid chromatography (LC) was connected to LTQ-Orbitrap Elite mass spectrometer (Thermo Fisher Scientific). Full-scan MS spectra were acquired in the Orbitrap with a mass resolution of 60000. The mass window for precursor ion selection was set to 2 m/z. Each full MS spectrum was followed by 12 MS/MS scans, where the 12 most abundant ions above a threshold of 10000 counts with charge ≥ 2 were subjected to collision-induced dissociation in the linear ion trap at the Institute for Research in Immunology and Cancer (IRIC, Université de Montréal). The dynamic exclusion was set to 45 s. The collision energy was set to 35 with an activation Q of 0.25.

3.2.11 Protein Identification and Quantification

Spectral processing and peak list generation was performed using Mascot Distiller v2.5.1 (Matrix Science). The data were searched with the Mascot search engine, v2.3.01, against concatenated forward and reversed six-frame translations of all 454 and Illumina contigs containing 100662 forward sequences. The reversed sequences served as decoys, as opposed to the forward sequences in the target database. Tolerance was set at 15 ppm for precursor and 0.5 Da for fragment ions. Variable modifications were specified, including carbamidomethylation of cysteine, oxidation of methionine, deamidation, and phosphorylation of serine, threonine, and tyrosine residues. The false discovery rate (FDR) was calculated as the ratio of decoy matches versus target matches. Initially, proteins at 5% FDR with at least 2 peptides detected were reported. We applied a

curation step to these contigs, as explained in Figure S3.1. After curation, proteins at 3% FDR with at least 6 assigned peptides were reported in the ovule secretome.

Label-free protein quantification was performed using ProteoProfile program (http://www.thibault.iric.ca/ProteoProfile/). Mascot peptide identifications were matched to MS peak intensity extracted from the aligned MS raw data files (tolerances set to m/z: 15 ppm and retention times: 1 min). Only peptides with a minimum intensity value of 10000 counts were analyzed. Normalization was performed across different replicates and conditions for peptide intensities during each LC−MS run to adjust the median of their logarithms to zero. Protein intensities were summed using the median of all associated peptides. Only proteins defined by at least two quantified peptides were reported in ovule secretome. The mean coefficient of variation (CV) of protein intensity for ovule secretome at anthesis was 0.372 and 56.3% of proteins were below this mean (Figure S3.2). Proteins showing a fold-change below –2.0 or above 2.0 between conditions with a *p*-value below 0.05 following a Student's *t*-test were reported as being differentially secreted.

3.2.12 *in silico* Predictions

Protein sequences were analyzed *in silico* for secretion predictions with the signalP 4.0 (Petersen, Brunak et al. 2011) and SecretomeP 1.0 (Bendtsen, Jensen et al. 2004) (score ≥ 0.5) programs. Sequences were also blasted against the nr database using BLASTp with default parameters (Altschul, Gish et al. 1990). The best informative BLAST hits were retained. The Blast2GO program was used to annotate the ovule secretome (Conesa, Gotz et al. 2005). Protein families and domains information were retrieved from PFAM (http://pfam.sanger.ac.uk/).

3.2.13 Estimates of Diversifying Positive Selection

Positive selection estimates were conducted on cysteine-rich peptides (CRPs) in the ovule secretome based on the ratio of divergence at nonsynonymous and synonymous sites (*dN/dS*). For each CRP investigated, closest orthologs in 10 solanaceous species were retrieved using the reciprocal best BLAST hits method and aligned codon-by-codon. The codeml program from the PAML suite (Yang 2007) was run to compute likelihoods of

the M7 (neutralist) and M8 (selectionist) models using the phylogeny described by Goldberg, Kohn et al (Goldberg, Kohn et al. 2010). A likelihood ratio test was finally performed to determine if positive selection was acting on the aligned sequences.

3.3 Results

3.3.1 PTs Are Attracted By Mature Ovules

The chemotropic effect from ovules at two developmental stages was examined. Figure 3.2A,B shows typical PT behavior toward each ovule condition. Most PTs grew directly toward mature ovule clusters, whereas 2DBA ovule clusters showed lesser or no attraction**.** On the basis of the angles defined in Figure 3.2C for this assay, a Kolmogorov-Smirnov (KS) test was employed to statistically distinguish random growth from PT attraction by ovules (Figure 3.2D, $n > 80$). KS test results indicated that immature ovules exhibited significantly different PT behavior ($p = 0.02$) compared to that of ovules at anthesis stage.

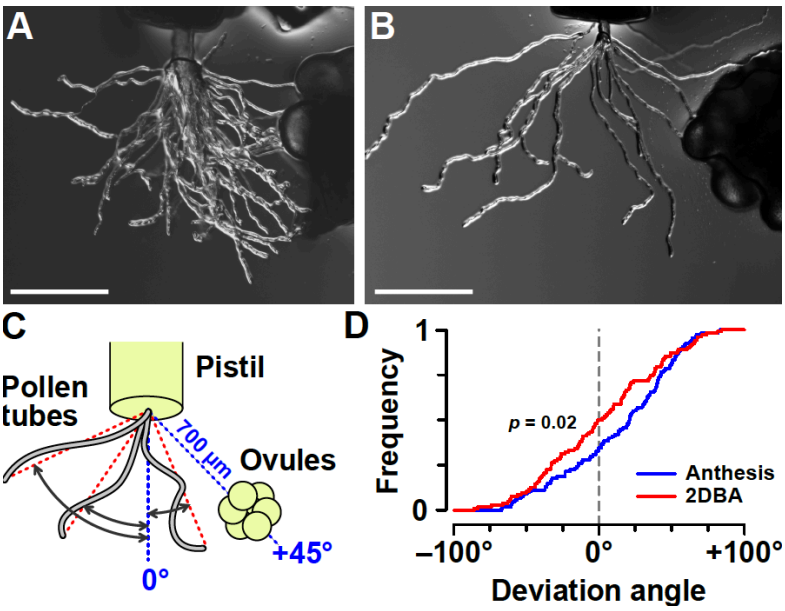


Figure 3.2 Chemotropism of PTs toward (A) anthesis ovule clusters and (B) immature ovule clusters (2DBA) in *S. chacoense*. Scale bars = 400 µm. (C) Schematic depiction of semi *in vivo* PT guidance assay in *S. chacoense*. Double-headed arrows define the angle of each PT in this assay. (D) The angles of PT curvature between two developmental stages were compared by the KS test. Results obtained from KS test indicated immature ovules exhibited significantly different PT behavior ($p = 0.02$) as compared to that of ovules at anthesis stage.

3.3.2 Secretome Protein Isolation Through A Modified Tissue-Free Gravity-Extraction Method (tf-GEM)

In order to isolate proteins that could be involved in PT−ovule interactions, ovule exudates were collected. Two methods are commonly used for *in planta* protein collection, namely the vacuum-infiltration centrifugation (VIC) method (Terry and Bonner 1980) and the gravity-extraction-method (GEM) (Jung, Jeong et al. 2008). In the VIC method, a vacuum is applied to infiltrate buffer into the extracellular space and then the apoplastic fluid is harvested by centrifugation. In an effort to minimize cell damage during vacuum infiltration, the GEM was developed, where plant tissues are centrifuged directly at low speed to extract apoplastic fluid without bathing extracellular space in extraction buffer/H_2O in advance. Due to the fragility of the ovule, a GEM-based system was specifically tailored for ovule secretome studies and named tf-GEM for tissue free-GEM. A similar approach was also recently used to study PT secretome (Hafidh, Potesil et al. 2014). The workflow is schematically described in Figure 3.3. Briefly, a gel-based medium is used as a support for the tissue considered. The sample is placed on a gel, allowing the exudates to soak in the solid medium. Following incubation, ovules are removed, and the supporting gel is spun to recover the exudates. It should also be noted that both VIC and GEM methods aim to collect all apoplastic proteins including cell wall proteins, whereas the tf-GEM herein, by removal of ovule clusters before spinning, is intended to collect secreted proteins that diffuse from the apoplastic space.

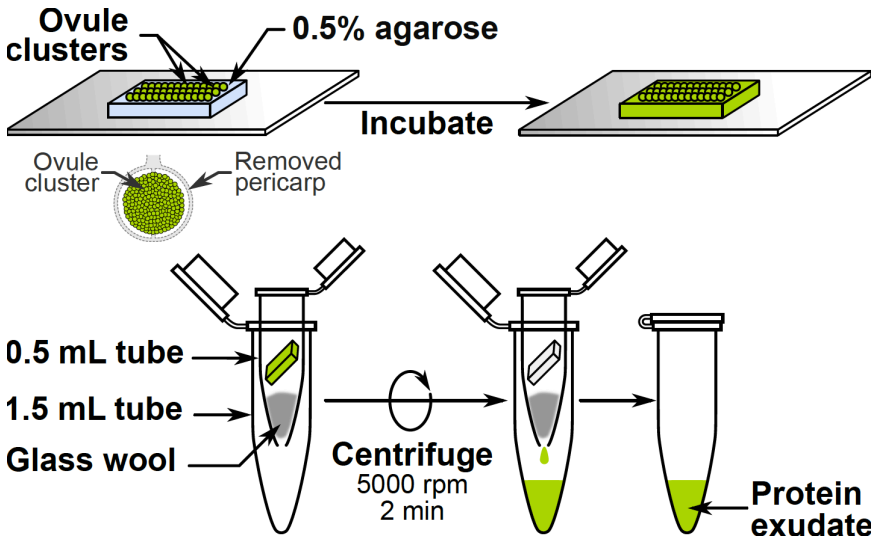


Figure 3.3 Tissue-free GEM system workflow for ovule secretome isolation. After removal of the pericarp, ovule clusters are incubated on an agarose matrix. Following an overnight incubation, ovules are removed from the gel matrix under a stereomicroscope. The matrix is laid on top of glass wool in a small perforated tube, and exudates are collected in a larger tube by centrifugation. Collected protein exudates are flash frozen in liquid nitrogen and stored at –80 °C until use.

3.3.3 Purity Assessment

To monitor if dissection caused contamination of the ovule exudates due to intracellular protein leakage, presence and activity of a control protein, triose phosphate isomerase (TPI), was tested. First, a western blot was performed using an antibody specific for cytosolic TPI (cTPI), the main isoform of TPI present in plant tissues (Dorion, Parveen et al. 2005). This experiment revealed no trace of cTPI in ovule exudates, whereas a clear band was observed when using total ovule proteins as control (Figure 3.4). Second, a highly sensitive enzymatic assay assessing total TPI activity in the exudates (Dorion, Parveen et al. 2005) showed that intracellular contamination was limited to 0.79% (Table 3.1). These results confirm that tf-GEM is a noninvasive technique that minimizes cytosolic content leakage.

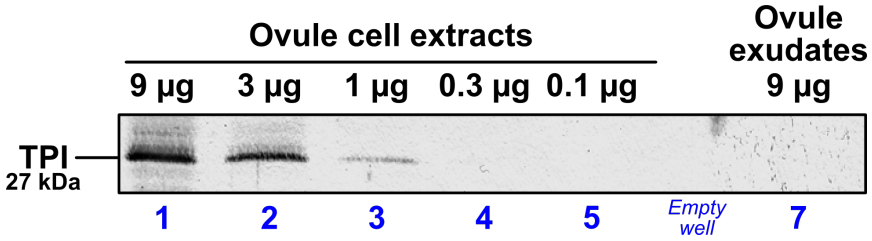


Figure 3.4 Ovule exudates' purity assessment. Western blot analyses of the cytosolic TPI isoform were performed from whole ovule extracts and ovule exudates to detect cytosolic leakage stemming from tissue wounding. Lanes 1−5, dilution series (9, 3, 1, 0.3, and 0.1 µg) of total ovule proteins. Lane 7, ovule exudates (9 µg) obtained from the tf-GEM protein collection.

Table 3.1 Assessment of TPI Enzymatic Activity in Whole Ovules Extracts and Ovule Exudates from *S. chacoense*

TPI enzymatic activity (U·mg^{-1} prot)		**cytosolic contamination**
whole ovules	exudates	
12.928 ± 1.675	0.102 ± 0.013	0.79%

Data (±) indicate standard deviation of the means. Contamination was calculated from TPI activity in ovule exudates divided by the one from whole ovule extracts.

3.3.4 Curation of Transcriptomic Data Drastically Increased Coverage for Protein Identification

Maximization of protein identification by MS requires a high-quality reference proteome, which can be a major hurdle when using a nonmodel species like *S. chacoense*. Although annotated genomes are available for other solanaceous species such as potato (*S. tuberosum*) (Potato Genome Consortium 2011) and tomato (*S. lycopersicum*) (Tomato Genome Consortium 2012), reproductive proteins are often found to be highly divergent (e.g., LURE-type PT attractants or self-incompatibility S-RNases). Thus, the usage of an annotated genome derived from a related species in MS may increase the risk of missing species-specific proteins. Therefore, we used transcriptomic data from *S. chacoense* to

build our own reference proteome. First, RNA-seq assemblies from both 454 and Illumina platforms were translated into six frames and concatenated into an initial reference proteome. On the basis of this database, Mascot assigned 363 proteins (FDR=5%) to the ovule secretome. However, the large size of the RNA-seq database, the presence of untranslated regions (UTR), and the possibility of undesired peptide assignment to the junctions of concatenated amino acid sequences are factors that can increase the risk of false peptide discovery. In this study, a data curation protocol was designed to maximize and improve protein identification, as illustrated in Figure S3.1. We applied the curation protocol to the 363 initially identified proteins. These contigs were queried against refseq_rna and refseq_protein databases from NCBI with the BLASTn and BLASTx programs, respectively, in order to delimitate open reading frames (ORFs) and to resolve frameshifted and chimeric contigs. Following this protocol, 362 curated proteins were obtained. On the basis of this refined reference proteome (362 proteins), Mascot assigned 305 proteins (FDR=3%) to the ovule secretome. On average, 17 peptides (ranging from 6 to 106 peptides) were assigned to each protein, compared to 4 peptides for the former search. Accordingly, by using all identified peptides, the average protein coverage was increased from 13.6 to 67.3%, as calculated by Protein Coverage Summarizer (http://omics.pnl.gov/). This 5-fold increase in protein coverage increased the reliability of protein identification and demonstrated the feasibility and necessity of our curation protocol to MS sequencing, using transcriptomic data as reference.

3.3.5 Ovule Secretome Annotation

In all, 305 proteins were reported for the ovule secretome, with at least 6 peptides identified and a protein FDR of 3%. These are hereafter designated as ovule-secreted proteins (OSPs). Table S3.1 provides a catalog of these 305 proteins, with descriptions and annotations as predicted from bioinformatic analyses. Although a noninvasive method was used that did not require centrifugation or bathing the sample for exudates collection, it is noteworthy that the number of proteins recovered is consistent with other plant secretomic studies (Cheng, Blackburn et al. 2009, Cho, Chen et al. 2009, Hafidh, Potesil et al. 2014). Biological process GO (gene ontology) categories were used to

globally describe the ovule secretome (Figure S3.3). The most represented GO categories are associated with metabolism (GO:0008152; 26%), cell growth and/or maintenance (GO:0009987; 19%), single-organism process (GO:0044699; 21%), and physiological response to stimulus (GO:0050896; 8%). To a lesser extent, proteins that are associated with developmental process, growth, signaling and reproduction account for 8%. In parallel, PFAM domain and family information were queried. Proteins with the same domain or family accession were pooled and displayed in Table S3.2.

Following Blast2GO annotation of biological processes, we examined protein secretion patterns of OSPs to study their predicted cellular localization. Protein secretion can proceed through different pathways. The best known is the classical ER−Golgi secretory pathway, where proteins require an N-terminal signal peptide (SP) for protein sorting. Lesser known are proteins secreted independently of the ER−Golgi pathway. A growing body of evidence demonstrates the importance of this unconventional protein secretion pathway, i.e., the production of leaderless secretory proteins, which could account for more than 50% of plant secretomes (Agrawal, Jwa et al. 2010). In the ovule secretome, 64 (21%) proteins were predicted to possess a SP. Some 86 proteins (28%) were predicted as leaderless secretory proteins by SecretomeP while not being predicted by SignalP. Similarly, SP-containing proteins also accounted for 20% of root cap secretome (Wen, VanEtten et al. 2007) and of the semi *in vivo*-PT secretome (Hafidh, Potesil et al. 2014).

3.3.6 Specificity of the Ovule Secretome

In order to sort out proteins that are unique to the ovule secretome, OSPs were separated into ovule-specific OSPs (oOSPs) and nonspecific OSPs (nOSP), by comparing them to the generally secreted proteins (GSPs) described in the PlantSecKB curated databases (Lum, Meinken et al. 2014), and stigma exudate proteins (SEPs) from lily (*Lilium longiflorum*), olive (*Olea europaea*) (Rejon, Delalande et al. 2013, Rejon, Delalande et al. 2014) and tobacco (*Nicotiana tabacum*) (Sang, Xu et al. 2012). SEPs were treated separately from GSPs since stigma exudates are active in *in vitro* PT chemotropism (Kim, Mollet et al. 2003) and might share a certain level of similarity with ovule exudates. The ovule secretome was then queried against the two aforementioned subproteomes with

BLASTp. Putative orthologs based on best hits with an e-value below 1×10^{-10}, a minimum alignment coverage of 80% and a minimum amino acid identity of 50%, are reported in Table S3.3.

As shown in Figure 3.5, 116 (38%) and 52 (17%) OSPs have a putative ortholog in the GSP and SEP databases, respectively, with an overlap of 41 proteins (13%) in all three profiles. OSPs shown to have a putative ortholog detected in either GSPs or SEPs were named nonspecific OSPs (nOSPs). The existence of nOSPs indicates that common factors are secreted into the extracellular matrix, regardless of the tissues examined. Such factors include pathogenesis-related (PR) proteins, like chitinase, β-1,3-glucanase, and peroxidases. Besides PR proteins, peptidases, heat shock proteins HSP70 and HSP90, and chaperonin 60 are also released to the extracellular environment. These proteins are mainly involved in plant defense against pathogens and stress responses (Jung, Jeong et al. 2008). A self-incompatibility ribonuclease (S-RNase, OSP286) was also found to be common among all three secretome profiles, although its putative ortholog in SEP database is only 32% identical to OSP286 due to the highly divergent nature of S-RNases (Matton, Maes et al. 1997).

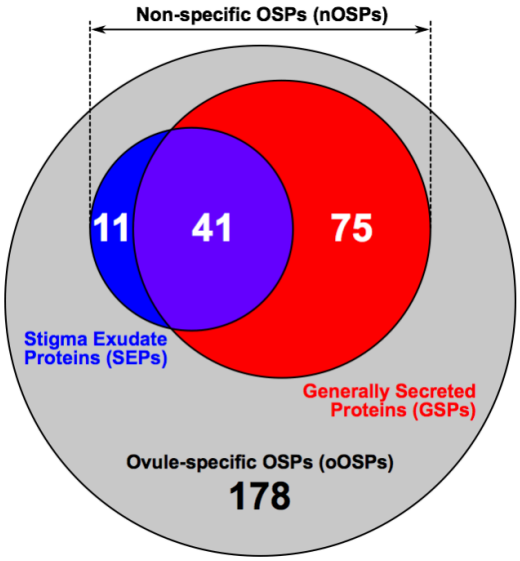


Figure 3.5 Diagram illustrating the number of oOSPs and nOSPs in the ovule secretome as compared to GSPs obtained from PlantSecKB database and SEPs collected from lily, olive, and tobacco stigma exudates. The gray area represents oOSPs.

In contrast to nOSPs, 178 OSPs did not have any ortholog in other known secretomes. These oOSPs accounted for 58% of the ovule secretome. Notably, these include a γ-amino butyric acid-transaminase (GABA-T), OSP131, which was previously shown to be involved in sporophytic PT guidance in *Arabidopsis* (Palanivelu, Brass et al. 2003). Four members of the CRP family were also detected in the ovule secretome (OSP21, 123, 227, 305), belonging to lipid-transfer proteins (LTP), early culture abundant 1 (ECA1) gametogenesis-related family, and thionin-like CRP subgroups.

3.3.7 Label-Free Quantification Revealed Differentially Secreted Proteins Between Mature and Immature Ovules

As expected from the PT guidance assay, ovules at anthesis exhibited a distinct protein profile from that of nonattracting ovules (2DBA) in terms of secretion abundance. Compared to the 2DBA secretome, 128 and 7 proteins were, respectively, up- and downregulated at anthesis, whereas the abundance of the majority of proteins (56%) remained unchanged (Figure 3.6). Morphologically, the overall size of the ovule is not significantly different between 2DBA ovules and mature ones, but ES development is not yet completed in 2DBA ovules (Chevalier, Loubert-Hudon et al. 2013). At this stage, polar nuclei fusion in the central cell has not yet occurred while antipodal cells degeneration has not yet completed. Although synergid cells and egg cell are fully cellularized by 2DBA, vacuolar development inside the egg apparatus has just initiated (Chevalier, Loubert-Hudon et al. 2013). Since cell wall ingrowth at the micropylar apex of the synergid cells starts after the formation of a fully vacuolated, pear-shaped egg apparatus, the development of the filiform apparatus, although initiated, is most probably incomplete. Accordingly, immature ovules at 2DBA did not attract PTs (Figure 3.2). Indeed, synergids and the filiform apparatus had been previously shown to be involved in PTs attractant secretion (Higashiyama, Yabe et al. 2001, Kasahara, Portereiko et al. 2005).

On the basis of this result, we observe that, although deeply embedded in the ovule nucellus, the maturation step arising from stage FG6 to FG7 (mature ES) affected almost

half (44%) of the ovule secretome. The secretion status of ovule-expressed proteins is thus closely associated with late ES development.

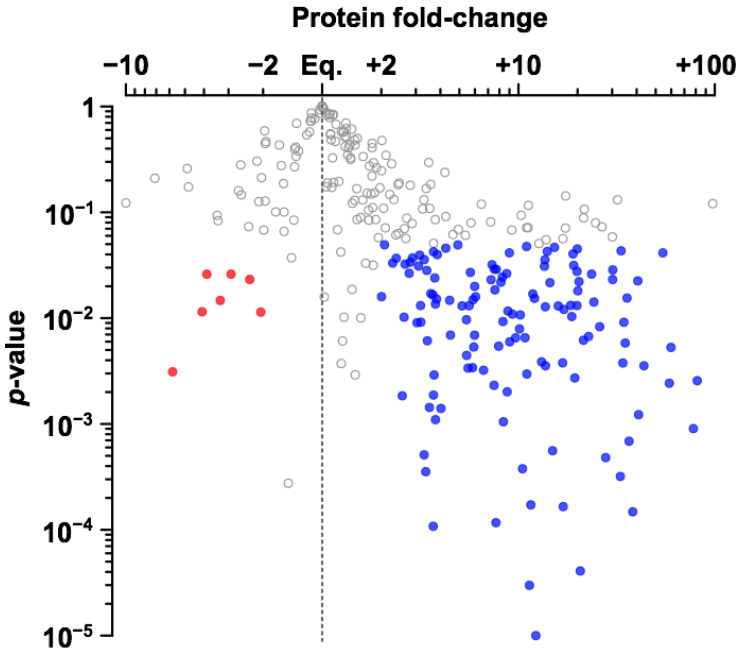


Figure 3.6 Protein secretion ratio in anthesis vs 2DBA ovules. Each dot represents a secreted protein in ovule secretome. Protein secretion ratios with a value < 1.0 were inverted to avoid fractional number. A minus symbol (–) was added to these values to indicate a negative regulation. Blue dots highlight OSPs that have a secretion fold-change of anthesis/2DBA > 2.0 and a *p*-value < 0.05. Red dots highlight OSPs that have a secretion fold-change of anthesis/2DBA < – 2.0 and a *p*-value < 0.05.

To extend our secretomic study, we conducted differential gene expression analyses for the 305 OSPs. As shown in Figure 3.7, there is no correlation between gene expression and secretion abundance. Of the 128 upregulated proteins at anthesis (14-fold on average), 83% (106) were not regulated at the mRNA level (1-fold on average), emphasizing the importance of post-transcriptional regulation in reproductive development and the necessity of conducting proteomic analysis in parallel to transcriptomic studies.

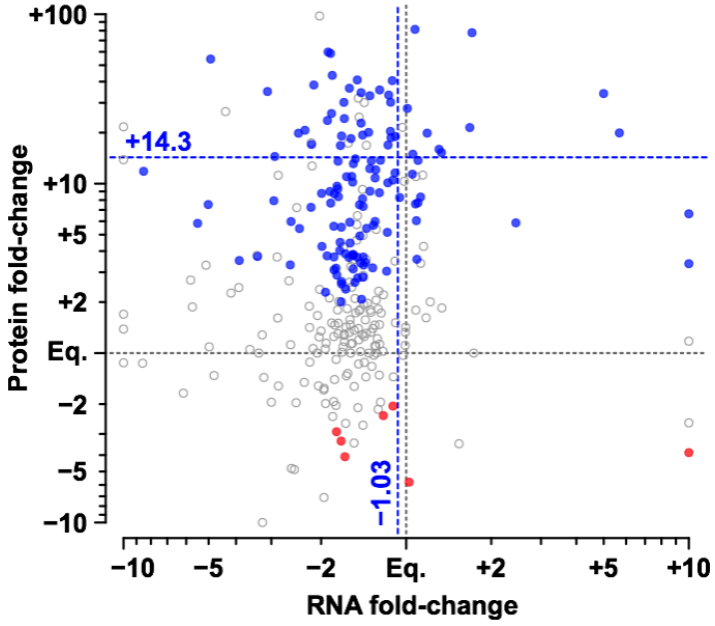


Figure 3.7 Fold-change correlations between protein secretion abundance and gene expression during the 2DBA to anthesis transition. Each dot represents a secreted protein in the ovule secretome. Blue dots highlight OSPs that are upregulated at anthesis at the protein level (fold-change > 2.0, p-value < 0.05). Red dots highlight OSPs that are downregulated at anthesis. Mean RNA (−1.03) and protein (+14.3) fold-changes from upregulated OSPs are represented as dotted lines.

3.4 Discussion

3.4.1 tf-GEM, A New Approach for Secretomic Studies

In this study, we established tf-GEM as a simple and efficient method to isolate tissue exudates with minimal cytosolic contamination. First, as shown in Figure 3.3, instead of severing individual ovules from the placenta, ovule clusters were placed on the agarose strip, preventing placental cells from directly touching the gel, thus minimizing cell leakage from wounded plant tissues. Second, unlike the VIC and GEM methods, tf-GEM does not require centrifugation of plant tissues that can lead to cell breakage. The

noninvasive nature of the tf-GEM is also substantiated by its very low cytosolic contamination levels (Figure 3.4 and Table 3.1).

3.4.2 A Long-Distance Guidance semi *in vivo* Assay for PT Attraction

The guidance mechanism operating behind our semi *in vivo* system differs from that in previous models designed for *Torenia* (Higashiyama, Kuroiwa et al. 1998) and *Arabidopsis* (Palanivelu and Preuss 2006). While these aimed at observing micropylar guidance of a single ovule within a short distance between a PT and an ovule (~100−150 μm), the current assay investigates guidance capacity of ovule clusters from a longer distance at ~700 μm. These two types of semi *in vivo* assays can be distinguished based on different PT behaviors in each system. In the single-ovule semi *in vivo* assay, PTs make an abrupt turn toward the micropyle only when they are near an ovule. In our assay, PTs gradually change their trajectory toward the ovule clusters, and this attraction can be seen as soon as PTs exit the bottom-end of the style. Although it is quite possible that short-range diffusible signals are also secreted onto the medium when ovule exudates are collected, the underlying mechanism behind our semi *in vivo* assay may point to a long-distance guidance mechanism.

3.4.3 The Ovule Secretome Defines a Microenvironment for Pollen–Pistil interactions Before Fertilization

In the ovular guidance stage, as soon as PTs exit the transmitting tract, they are in direct contact with the locular fluid produced by the placental tissues in the ovary. The carbohydrate composition of this fluid was studied previously in the angiosperm *Gasteria verrucosa* and includes fructose, sucrose, and glucose as the main components (Willemse and Franssen-Verheijen 1988). These carbohydrates are important for creating proper osmolarity for PT growth. In contrast, the protein composition had not been addressed. Profiling of the ovule secretome in *S. chacoense* provides the first description of the protein composition of the locular fluid, more specifically of ovular origin, based on the large difference observed between the mature vs. immature ovule secretomes. Our data suggests that proteins secreted in the locular fluid create a dynamic microenvironment that functions not only in PT growth and guidance but also in controlling late self-

incompatibility, stress tolerance, and regulation of unconventional secretion from the ovules, to name a few. Here, we highlighted the importance of a few OSPs based on the different functional categories with which they are associated in pollen–pistil interactions.

3.4.3.1 Intra- and Interspecific Incompatibility

S-RNases are involved in gametophytic self-incompatibility in the Solanaceae, Scrophulariaceae and Rosaceae, preventing self-fertilization and inbreeding (McClure, Cruz-Garcia et al. 2011). BLASTp search assigned two putative orthologous S-RNases to OSP286: one from tobacco stigma exudates (S_{15}-RNase) (Sang, Xu et al. 2012) in the SEP database and one from potato style (S_2-RNase) (Kaufmann, Salamini et al. 1991) in the GSP database. OSP286 corresponds to the S_{14}-RNase of *S. chacoense* (O'Brien, Kapfer et al. 2002). Although S-RNases are mostly expressed in the stylar transmitting tract where they exert their major function, i.e., self-PT growth cessation, *in situ* hybridization revealed that they are also expressed in the ovule integument and in the epidermal cells of the placenta that are continuous with the transmitting tract tissue (Matton, Bertrand et al. 1998), a localization consistent with the ovule secretome. Furthermore, accumulation of OSP286 is developmentally regulated, being 11-fold more abundant at anthesis compared to 2DBA (Table S3.1), consistent with the accumulation of the S-RNases and HT-B modifier genes in the style where peak accumulation occurs at the anthesis stage (O'Brien, Kapfer et al. 2002). We thus postulate that ovular-secreted S-RNases may act as a late checkpoint for self-PTs that would have survived their journey through the style. This is not uncommon since ribonuclease activity can be quite variable between different S-alleles and even from style to style for the same allele.

OSP30 is similar to PELPIII, a class III pistil-specific extensin-like protein with arabinogalactan protein (AGP) properties initially isolated from tobacco pistil (Goldman, Pezzotti et al. 1992), where it was found to be expressed in the transmitting tract of the style and further translocated in the growing PT cell wall (Bosch, Knudsen et al. 2001, de Graaf, Knuiman et al. 2003). More recently, PELPIII was also shown to be required for proper rejection of interspecific pollen (Eberle, Anderson et al. 2013). Compared to

PELP III, OSP30 is proline-rich (16%) at the N-terminus, shares more than 50% identity with PELPIII at the C-terminus, and is predicted to be heavily *O*-glycosylated (Steentoft, Vakhrushev et al. 2013). AGPs have been detected in the ES of *A. thaliana* ovules (Coimbra, Almeida et al. 2007) as well as in other reproductive tissues (Pereira, Masiero et al. 2014). A member of the AGP family, the transmitting tissue-specific (TTS) protein is a PT growth simulant that also induces chemotactic behavior of semi *in vivo* grown PTs in tobacco (Cheung, Wang et al. 1995) and *Nicotiana alata* (Wu, Wong et al. 2000). OSP30 is developmentally regulated, with a 7-fold increase during ES maturation between the 2DBA and anthesis stages. Apart from rejecting interspecific PTs, OSP30 might be involved in PT directional growth through its interaction with PTs.

3.4.3.2 PT Guidance

OSP 131 is a γ-amino butyric acid-transaminase (GABA-T), closely related to the *A. thaliana* GABA-T POP2, with 73.2% amino acid sequence identity. The *Arabidopsis pop2* mutant is defective in GABA catabolism and accumulates GABA. In *pop2* mutant, GABA accumulation leads to defective PT elongation in the transmitting tract *via* putative Ca^{2+}-permeable channels on the PT plasma membrane (Renault, El Amrani et al. 2011, Yu, Zou et al. 2014). In the pistil, POP2 degrades GABA and contributes to the formation of a GABA gradient from the stigma (lowest) to the inner integument (highest) and was shown to be important in leading the PTs to the micropyle (Palanivelu, Brass et al. 2003). However, compared to POP2, which bears a N-terminal mitochondrial targeting sequence, OSP131 lacks such a sequence, as predicted by both TargetP (Emanuelsson, Nielsen et al. 2000) and PSORT (Nakai and Horton 1999). In contrast, OSP131 is predicted as a leaderless secretory protein by SecretomeP. Since *POP2* is a single-copy gene in *Arabidopsis*, the *S. chacoense* ovule transcriptome was queried for other isoforms of OSP131 that might be targeted to the mitochondria. A mitochondrial isoform of GABA-T sharing 68.5% amino acid identity with OSP131 was found, but is not detected in the ovule secretome. Two other GABA-T isoforms sharing 79 and 93% amino acid sequence identity with OSP131 were detected in the transcriptome and predicted as leaderless secretory proteins. However, they were also not detected in the ovule exudates. This suggests that in *S. chacoense* ovules OSP131 is delivered through

the unconventional protein secretion pathway and might function in the extracellular matrix of the ovary to mediate a GABA gradient. Interestingly, solanaceous species like potato and tomato have both mitochondria-targeted and leaderless secretory protein GABA-T, whereas *Arabidopsis* has only the mitochondrial isoform (data not shown).

Besides GABA-T, four CRPs are also present in the ovule secretome. CRPs are of particular interest since short-distance guidance signaling involves several small secreted CRPs. The significance of diverse CRPs in pollen–pistil interaction has been recently reviewed in detail (Higashiyama 2010, Chae and Lord 2011, Marshall, Costa et al. 2011). According to cysteine motifs classified previously from plants (Silverstein, Moskal et al. 2007), OSP21 (8-cys) belongs to the LTP subgroup of CRPs. OSP227 (6-cys) falls into the ECA1 gametogenesis-related family, whereas OSP305 (12-cys) is predicted to be a thionin-like protein. These three CRPs harbor a SP followed by a mature peptide of roughly 90 amino acids. Another thionin-like CRP, OSP123, is larger in size, with a predicted mature peptide of 133 amino acids and a peculiar organization of 18 cysteines comprising three blocks of six cysteines, of which the first two blocks have an identical cysteine pattern, sharing 84% sequence identity. Each block is preceded by a basic amino acid residue (lysine or arginine), suggesting a possibility of proteolytic processing through the presence of monobasic cleavage sites (Rholam and Fahy 2009), similar to what is observed in the BTH6 barley thionin (Plattner, Gruber et al. 2015).

An orthology survey was conducted on these secreted CRPs in 10 related solanaceous species (Table S3.4). Data indicated that OSP21 (LTP) and OSP227 (ECA1) were both under positive selection (dN/dS >1). No reciprocal best-BLAST hit was assigned to thionin-like OSP123 and OSP305, likely due to their high divergence. Such divergent CRPs could be involved in reproductive processes in a species-specific manner. In particular, the secretion of OSP227 is regulated in the mature ES and marked as an upregulated OSP in the ovule secretome (Table S3.1). In *Arabidopsis*, 119 ECA1 genes remain functionally unknown (Sprunck, Hackenberg et al. 2014), although subcellular localization of several ECA1 genes was associated with the synergid cells (Jones-Rhoades, Borevitz et al. 2007, Steffen, Kang et al. 2007). The putative ortholog of OSP227 was found only in *Solanum tarijense*, a very closely related species to *S.*

chacoense, among 10 other solanaceous species, suggesting that OSP227 is highly divergent and could function in a species-specific manner, as expected for short-distance guidance signals (Higashiyama, Inatsugi et al. 2006, Palanivelu and Preuss 2006).

3.4.3.3 Stress Tolerance

OSP130 is predicted to be an isoflavone reductase-like (IFR-like) protein, one of the leaderless secretory proteins. OSP130 is 66% identical (81% similar) to the pistil-expressed and pollination-enhanced CP100, an IFR-like protein from *S. tuberosum* (van Eldik, Ruiter et al. 1997). Such IFR-like proteins in maize and *Arabidopsis* were shown to be involved in defense against oxidative stress (Babiychuk, Kushnir et al. 1995, Petrucco, Bolchi et al. 1996). Transgenic plants overexpressing rice IFR-like gene exhibited growing tolerance to reactive oxygen species (ROS) in leaves and suspension-cultured cells (Kim, Kim et al. 2010). In the context of ES development, the central cell was shown to be the main source of ROS before pollination. ROS were also shown to affect the central cell fate and were detected in synergid cells after pollination (Martin, Fiol et al. 2013). Furthermore, high levels of ROS under the control of the receptor-kinase FERONIA (FER) were observed at the female gametophyte entrance, where ROS were shown to mediate PT rupture in order to release its sperm cells (Duan, Kita et al. 2014). Considering OSP130 secretion is developmentally upregulated by 6-fold in anthesis ovules compared to that in 2DBA (Table S3.1), it will be interesting to examine the localization of OSP130 in the ovule and explore its potential role in ROS regulation, ES development, and PT guidance.

3.4.3.4 Unconventional Secretion

Four nOSPs were annotated as 14-3-3 proteins. The extracellular localization of 14-3-3's was confirmed by *in situ* immunolocalization in root border cells of pea and maize (Wen, VanEtten et al. 2007). 14-3-3's are encoded by a gene family containing multi-isoforms (Wu, Rooney et al. 1997, DeLille, Sehnke et al. 2001). One characteristic of 14-3-3's is their highly conserved core and divergent N- and C-terminus sequences, which possibly explains their organelle-specific functions (Chung, Sehnke et al. 1999). 14-3-3 proteins are generally involved in signal transduction. They can modulate protein function through

interaction with their phosphorylated binding partners and are therefore considered to be phospho-sensors. 14-3-3 proteins were also reported to stimulate unconventional secretion (Carreno, Goni et al. 2005). The presence of four distinct 14-3-3 proteins in our data set may relate to the high percentage of leaderless secretory proteins found in the ovule secretome. On the other hand, phosphorylation is one of the fastest responses to relay a stimulus to the cellular machinery. This post-translational modification could be employed by PTs in response to chemoattractant cues in order to promptly change their growth direction. It is possible that ovule-emitted 14-3-3's are used to modulate signal transduction of phosphorylated proteins in PTs once they are internalized. Of all of the 14-3-3 family OSPs, only the secretion of OSP23 is upregulated at anthesis, suggesting that it may be involved in reproductive-related processes.

3.4.4 ES Developmental Stage Influences Secretion Status of 44% of the Ovule Secretome

The developmental stage of the ES between FG6 and FG7 stages influenced the secretion of 44% of total OSPs. Thus, the late transition from an ES that has unfused polar nuclei, immature filiform apparatus and not fully degenerated antipodal cells to a fully mature ES leads to a drastic change in protein secretion. Since ovules at 2DBA are defective in PT attraction (Figure 3.2), these 128 OSPs could be involved in PT directional growth and guidance and thus provided a candidate pool for the study of isolating PT attractants. These OSPs may be directly secreted from the ES or secreted from the sporophytic tissue of the ovule, mainly the integument. Of major interest, this comparative analysis between mature and immature ovules revealed that, during the last maturation step of ES development, 106 of the 128 upregulated proteins are not regulated at the mRNA level (Figures 3.6 and 3.7), suggesting a late translation of the mRNAs that had accumulated at the 2DBA stage and/or the selective secretion of already-made proteins at anthesis stage, emphasizing the importance of post-transcriptional regulation in reproductive development.

3.5 Conclusion

The study of ovule secretome in *S. chacoense* reveals that various ovular signals are secreted to the ovarian locules, which form an active microenvironment to interact with approaching PTs. Comparative proteomic analysis allowed us to identify 128 secreted proteins potentially involved in pollen−ovule interactions, including PT guidance processes that are closely associated with the ES transition from the penultimate stage onward. Of these 128 upregulated OSPs from anthesis stage ovules, the majority was found to be unregulated at the mRNA level. This lack of correlation in gene expression and protein secretion suggests a strict regulation exerted from the ES over ovular signal secretion, vindicating this novel approach in the study of pollen–pistil interactions as a robust alternative to transcriptomic studies.

3.6 Supplementary data

Figure S3.1 Protein identification using customized protein database derived from RNA-seq data and its optimization through contig curation protocol. Ovule RNAs were sequenced in parallel using 454 and Illumina next-generation sequencing platforms. Two *de novo* assembled ovule transcriptome were generated, comprising 100662 sequences in total. MS/MS spectra obtained from ovule exudations were searched against six-frame translations of the concatenated database. Contig curation was applied to the initially identified 363 proteins to fix frameshifts and chimera stemming from RNA-seq or assembly (module contig curation). Protein sequences in the ovule secretome were deduced after curation (module protein prediction).

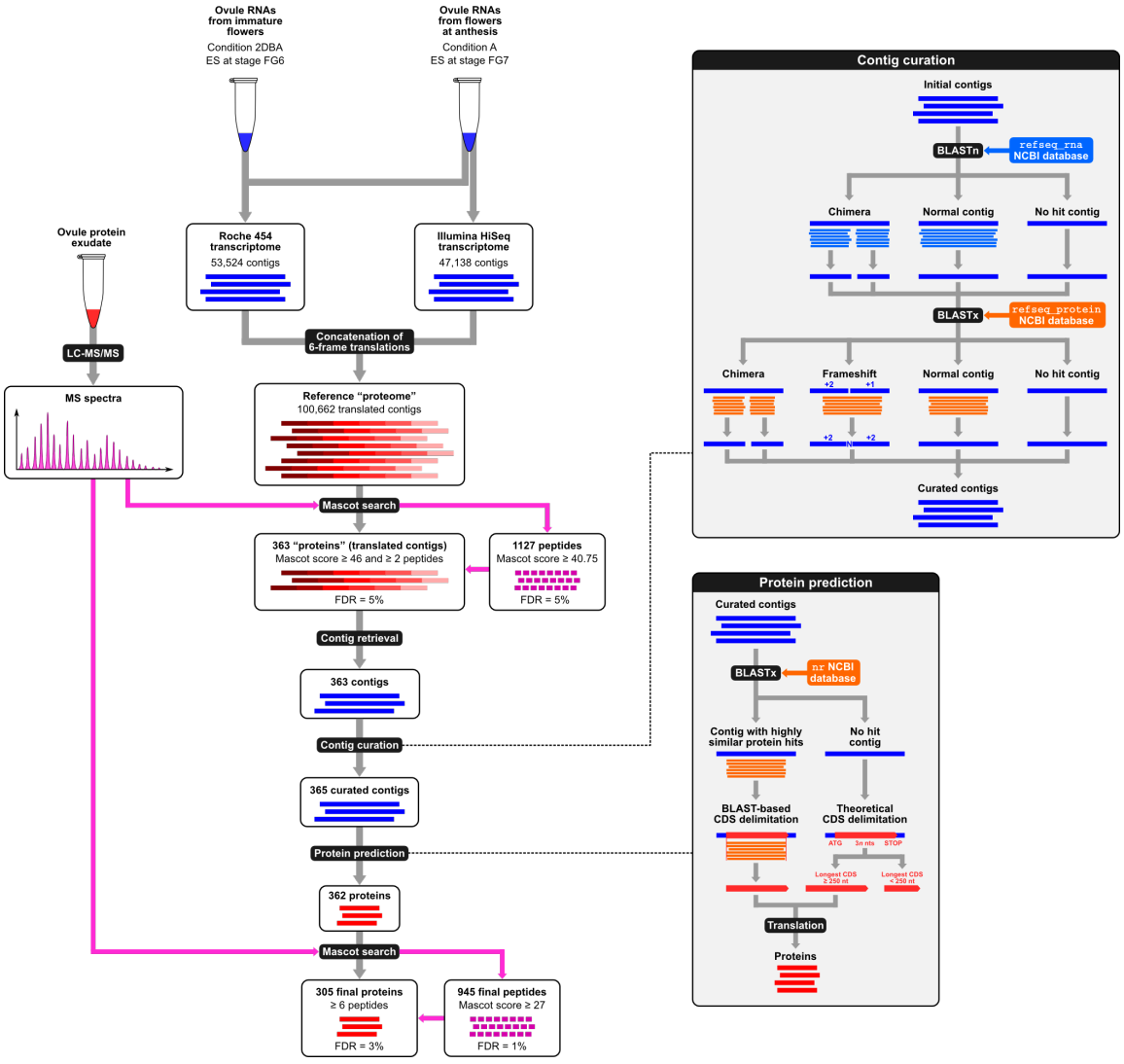


Figure S3.2 Reproducibility of label-free quantification of ovule secretome based on three biological replicates. (A) Peptide intensity coefficient of variation (CV) vs. the number of relative peptides in the reference condition (anthesis). (B) Protein intensity CV vs. the number of relative proteins. Dotted lines mark 50% of the peptide and protein population, respectively.

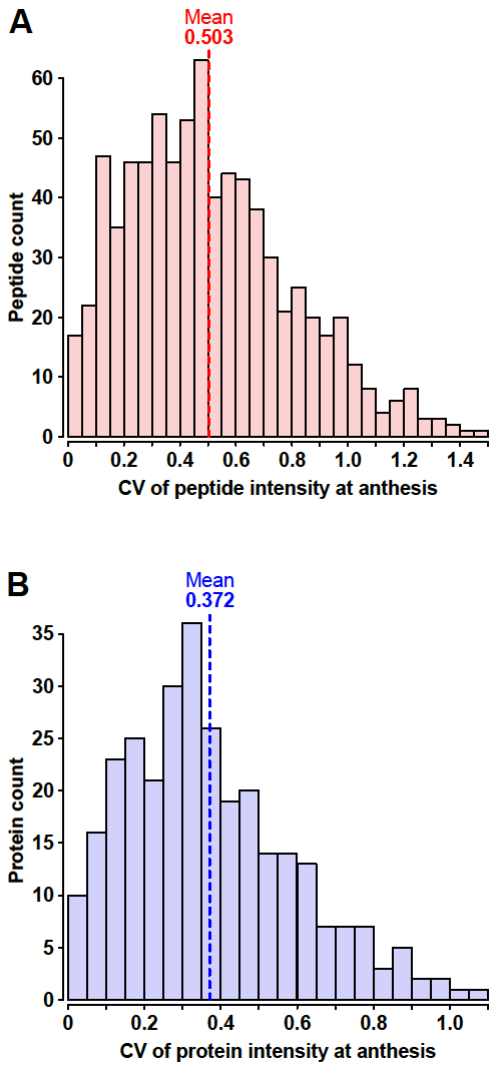
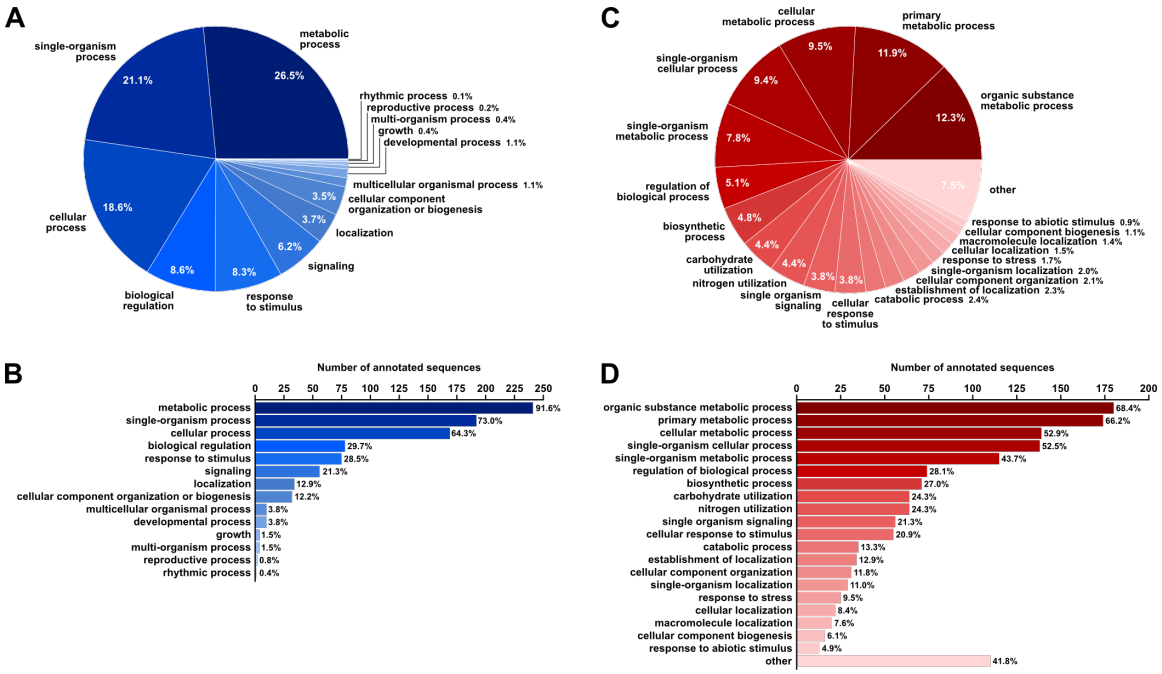


Figure S3.3 GO classification of ovule secretome in *S. chacoense*. (A) A pie chart showing GO classification under biological process. (B) A bar plot displaying the proportion of sequences associated with each GO-term under biological process. (C) A pie chart showing GO classification under molecular function. (D) A bar plot displaying the proportion of sequences associated with each GO-term under molecular function.



The supplementary tables are available on the ACS Publications website at DOI: 10.1021/acs.jproteome.5b00618.

Table S3.1 Main characteristics of ovule secretome in *S. chacoense.*

Table S3.2 PFAM domain and family annotation associated with each secreted protein in ovule secretome.

Table S3.3 Determination of OSP specificity by comparison of ovule secretome to other secretomic datasets.

Table S3.4 Ortholog survey of novel ScCRPs detected from the ovule secretome against other solaneceous species.

3.7 Acknowledgments

This work was supported by the Fonds de Recherche du Québec–Nature et technologies (FRQNT: 2011-PR-148467) and by the Natural Sciences and Engineering Research Council of Canada (NSERC: RGPIN-2014-03883).

Y.L. is the recipient of a Ph.D. fellowship from the China Scholarship Council.

4. Selection, Purification and Functional Assays of Candidate Proteins Involved In Ovular Pollen Tube Guidance In *S. chacoense*

Yang Liu, Fangwen Bai and Daniel P. Matton[§]

Institut de Recherche en Biologie Végétale, Département de Sciences Biologiques, Université de Montréal, 4101 rue Sherbrooke est, Montréal, Québec H1X 2B2, Canada

Author contributions: Y.L. performed all of the experimental work, except for protein expression experiments using the pBAD/gIII vector, which was performed by F.B. Y.L. wrote the manuscript and D.P.M. gave critical corrections.

[§]To whom correspondence should be addressed.

This manuscript is not intended for publication.

Abstract

In previous chapters, comparative transcriptomic and secretomic approaches were used to identify ovular signals involved in PT guidance of *S. chacoense*, using PT attracting ovules (anthesis) vs. nonattracting ovules (slightly immature ovules at 2DBA and ES-devoid mutant *frk1*). Differentially regulated molecules include CRP family proteins, which are small secreted signaling peptides that have been reported to be abundantly expressed in the ES and participate in a vast array of pollen–pistil interactions, including self-incompatibility, PT adhesion to the stylar matrix, intracellular signaling in PT growth, PT guidance and gamete fusion. This study focuses on the evaluation of these CRPs for their possible involvement in *S. chacoense* ovular guidance. To this end, 28 ovular-expressed CRPs were selected based on strict criteria. Of these, 18 CRPs have been previously identified in an initial comparative transcriptomic analysis using individual 454 and Illumina assemblies. These candidates were expressed in a bacterial expression system, purified, refolded and tested for their attraction activity toward semi *in vivo* grown PTs. The other 10 CRPs were most recently acquired upon transcriptome curation of the hybrid 454/Illumina assembly (Chapter II) and from the ovule secretome study (Chapter III), and are not included in the current screening. As a positive control for PT attraction functional assay (bead assay), anthesis ovule exudates were collected and were shown to induce chemotactic behavior of semi *in vivo* PTs (43%). However, purified candidate proteins did not show attraction activity. In order to overcome the difficulties during protein refolding and to achieve a bioactive form of recombinant proteins, both an *Agrobacterium rhizogene*-mediated plant expression system using *S. chacoense* roots and a *Pichia pastoris* yeast expression system intended for expressing secreted proteins with possible glycosylation modification were proposed to express new candidates. Meanwhile, to optimize PT behavior in the functional assay, the growth medium for *S. chacoense* PTs can be improved to sustain longer PT elongation after exiting the style, which will allow sufficient time for PTs to become competent before encountering purified candidates.

Keywords: pollen tube guidance, chemoattractant, cysteine-rich peptide, differential gene expression analysis, protein refolding, semi *in vivo* assay

4.1 Introduction

In flowering plants, double fertilization relies on successful pollen–pistil interactions for precise delivery of sperm cells to the ovule. To date, various small secreted peptides have been reported to play pivotal roles in cell-cell communication of pollen–pistil interactions, such as self-pollen recognition and rejection during self-incompatibility, PT growth stimulation, PT guidance, PT growth cessation and discharge (Higashiyama 2010).

For example, phytosulfokine (PSK), a disulfated pentapeptide which was first reported to be functioning as a growth factor in roots and shoots in *Asparagus officinalis* (Matsubayashi and Sakagami 1996), was also recently reported to be involved in sexual reproduction in *Arabidopsis* (Sauter 2015, Stuhrwohldt, Dahlke et al. 2015). Indeed, the knockout of two PSK receptors in a *pskr1-3 pskr2-1* double mutant causes a significant reduction of seed production. That reduced seed set was traced back to both maternal and paternal defects in PSK signaling which resulted in 20% of ovules that failed to attract a PT. Tyrosine sulfation of PSK, which is carried out by tyrosylprotein transferase (TPST) in *Arabidopsis* (Komori, Amano et al. 2009), was demonstrated to be required for its activity as a growth factor, also seems to be necessary for its function in sexual reproduction, as the TPST knockout mutant *tpst-1* displays a phenotype similar to the *pskr1-3 pskr2-1* double mutant.

Apart from PSK, a large number of reports have addressed the major involvement of CRPs in pollen–pistil interactions (Chae and Lord 2011, Marshall, Costa et al. 2011). CRPs constitute a large protein family that shares certain characteristics, i.e. small size (less than 160 amino acid residues), the presence of a signal peptide at the N-terminus and a divergent charged or polar mature peptide with conserved 4-16 cysteine residues (Silverstein, Moskal et al. 2007, Marshall, Costa et al. 2011). Based on a motif study of CRP family proteins, Silverstein, Moskal Jr. et al. identified 825 CRPs in *Arabidopsis* and 598 in rice, which was estimated to account for 2 to 3% of the total number of genes in each species (Silverstein, Moskal et al. 2007).

Within the CRP super family, the three largest subgroups are the defensin and defensin-like (DEFL) proteins, the lipid transfer proteins (LTP) and the early culture abundant 1 (ECA1) gametogenesis-related proteins (Zhou, Silverstein et al. 2013). Their specific involvement in various stages of pollen–pistil interactions has been well characterized.

Of the DEFL subgroup, S-locus cysteine-rich (SCR) peptide is the male determinant of sporophytic self-incompatibility in the *Brassicaceae* (Takayama, Shiba et al. 2000). LURE-type peptides are micropylar PT chemoattractants in *Torenia* (Okuda, Tsutsui et al. 2009, Kanaoka, Kawano et al. 2011) and *Arabidopsis* (Takeuchi and Higashiyama 2012), while *Zea mays* embryo sac 4 (ZmES4) peptide mediates PT burst in maize (Amien, Kliwer et al. 2010).

Of the LTP subgroup, the stigma/stylar cysteine-rich adhesion (SCA) protein is required for adhesion of PTs to the stylar transmitting tract epidermis in the hollow style of lily (Park, Jauh et al. 2000). SCA is secreted from both the pollen and the transmitting tract tissues. Positively charged SCA interacts with negatively charged pectin in the transmitting tract of the pistil to create an active extracellular matrix for the PT to attach (Mollet, Park et al. 2000).

The study of SCA attracted the attention to LTPs in flowering plants. Twelve SCA-like LTPs were found in *Arabidopsis* (Chae, Gonong et al. 2010). These LTPs are expressed diversely in different floral reproductive tissues, indicating that these SCA-like LTPs may have evolved different roles in plant growth and reproduction as extracellular signaling molecules, instead of the conventional intracellular lipid transfer function. For example, SCA-like LTP5 (At3g51600) is expressed in both PT and the pistil transmitting tract tissues (Chae, Gonong et al. 2010). The *ltp5-1* T-DNA insertion line exhibited delayed PT growth and decreased fertilization. Reciprocal crosses revealed that this mutant resulted in partial sterility in both male and female organs. In addition to LTP5, LTP3 (At5g59320) and LTP6 (At3g08770) are also SCA-like proteins. Both LTP3 and LTP6 are expressed specifically in the ovule, suggestive of a potential role in reproductive processes.

Lastly, 119 ECA1 members were found in *Arabidopsis* and many are preferentially expressed in the synergid cell of the ovule (Sprunck, Hackenberg et al. 2014). Although their functions remain largely unexplored, a subclade of ECA1, the egg cell-secreted EC1 peptides were shown to trigger sperm cell activation in *Arabidopsis* by redistributing the potential gamete fusogen HAPLESS 2/GENERATIVE CELL SPECIFIC 1 to the sperm cell surface (Sprunck, Rademacher et al. 2012).

In the quest for isolating ovular guidance cues from *S. chacoense*, we focused on investigating the role of ovule-expressed CRPs, based on prior knowledge of the involvement of CRPs in a

vast array of pollen–pistil interactions, and especially, in micropylar PT guidance in *Torenia* (Okuda, Tsutsui et al. 2009, Kanaoka, Kawano et al. 2011) and *Arabidopsis* (Takeuchi and Higashiyama 2012).

Previous research on the *S. chacoense* ovule transcriptome (Liu, Joly et al. in preparation) and secretome (Liu, Joly et al. 2015) identified a total of 59 CRPs in this reproductive organ, all of which showed ovule-enriched expression as compared to leaf (5X cut-off, p-value ≤ 0.05; Table S2.7 and S3.1). On the basis of this gene list, candidate chemoattractants involved in ovule PT guidance of *S. chacoense* was selected. Here, a candidate selection strategy is described, together with their expression, purification, and the functional assays to validate their role in PT guidance. Meanwhile, the pitfalls incurred during the screening are analyzed and future perspectives of this ongoing project were discussed.

4.2 Material and Methods

4.2.1 RT-PCR

Total RNAs from three ovule conditions (anthesis, 2DBA and *frk1*) and leaf was extracted using the TRIzol reagent (Life technologies). The ovary pericarp was removed before the extraction step. RT-PCR was performed using 70 ng of total RNAs from these samples using the M-MLV reverse transcriptase kit (Invitrogen). Twenty-six cycles were conducted to amplify candidate genes and *UBIQUITIN* using Crimson Taq polymerase (NEB). RT-PCR products were visualized using a 2% agarose gel. Primer sequences used in RT-PCR are listed in Table S4.1. Similarly, total RNAs from infiltrated leaves (see below) were extracted and RT-PCR was performed using 200 ng of total RNAs amplified for 30 PCR cycles.

4.2.2 Transient Expression by Agro-Infiltration of *Nicotiana benthamiana Leaves*

The binary vector pGWB8 was used to overexpress recombinant proteins with a polyhistidine tag at the C-terminus (Nakagawa, Kurose et al. 2007). Agrobacterium-mediated transient expression in *Nicotiana benthamiana* was performed using *Agrobacterium tumefaciens* strain EHA105 following a protocol as described (Kane, Agharbaoui et al. 2007), with modifications. The p19 suppressor for post-transcriptional gene silencing was used to enhance transient

expression of recombinant proteins (Voinnet, Rivas et al. 2003). In brief, the constructs and the p19 vector were transformed into EHA105 strain individually using electroporation. Agrobacteria carrying the gene constructs were grown in LB medium supplemented with 50 μg ml^{-1} kanamycin and 50 μg ml^{-1} rifampicin. Agrobacteria carrying the p19 vector were cultured in LB medium containing 5 μg ml^{-1} tetracycline and 50 μg ml^{-1} rifampicin. The agrobacteria culture was placed immediately on ice when the OD_{600} reached 0.7. Equal volume of the construct suspension and p19 suspension were mixed. To sediment agrobacteria, the mixture was centrifuged at 4,000 g for 15 min and then resuspended in equal volume of MMA infiltration medium containing 10 mM $MgCl_2$, 10 mM MES (pH 5.6) and 200 μM acetosyringone (freshly added). This suspension was incubated at room temperature for 3 h. Infiltration was done on the lower side of *N. benthamiana* leaves of 3 to 4 weeks old plants with a 1 ml syringe. After infiltration, plants were kept in dark condition for 24 h. Infiltrated leaves were harvested 2 d after infiltration.

4.2.3 Protein Extraction and Western Blot

Western blots were performed to confirm protein expression in *N. benthamiana* leaves. Two leaves were ground in liquid nitrogen and solubilized with 1 ml extraction buffer containing 50 mM Tris·HCl (pH 8.0), 500 mM NaCl and 20 mM imidazole supplied with protease inhibitor cocktail tablets (Roche). Protein extracts were swirled gently at 4 °C for 30 min and then centrifuged at 14,400 rpm for 30 min at 4 °C to remove cell debris. The supernatant (20 μg) was used for western blot. SDS-PAGE analysis was performed on 15% acrylamide gels. Proteins were transferred onto a nitrocellulose membrane at 90 V for 45 min at 4 °C. Membranes were incubated with an anti-polyhistidine antibody (1/10,000 dilution) for 3 h at room temperature (Life Technology). Proteins were detected using goat anti-mouse IgG secondary antibodies (1/5,000 dilution) conjugated with horseradish peroxidase (HRP) (Life Technology). The reaction was visualized using HRP substrate of luminol reagent catalyzed by peroxide solution (Millipore).

4.2.4 Expression and Purification of Recombinant Peptides in *E. coli*

Candidate genes were cloned in the expression vector pET28b carrying a polyhistidine tag at the C-terminus. Recombinant peptides were expressed in *Escherichia coli* Rosetta-gami 2 strain or

BL21 strain. The expression was induced by 1 mM isopropyl β-D-1-thiogalactopyranoside (IPTG) for 5 h at 28 °C. The protein purification strategy was adapted from the LURE-type CRP purification protocol (Okuda, Suzuki et al. 2013).

For recombinant denatured proteins that aggregated as inclusion bodies, 1000 ml of *E. coli* culture was centrifuged and resuspended in 30 ml suspension buffer containing 50 mM Tris·HCl (pH 8.0), 50 mM NaCl and 1 mM ethylenediaminetetraacetic acid (EDTA). A French press was used to lyse bacterial cells. The inclusion bodies were collected by centrifugation at 14,000 rpm for 30 min at 4 °C. The inclusion bodies were washed twice in a solution containing 0.5% Triton X-100 (V/V) and 1 mM EDTA, and in 1 M NaCl to remove genomic DNA released from the breakdown of remaining cells. The inclusion bodies were solubilized in a solution containing 6 M guanidine hydrochloride (Gnd·HCl), 50 mM Tris·HCl (pH 8.0), 500 mM NaCl, 20 mM imidazole and 1 mM freshly added β-mercaptoethanol (BME) at room temperature for 3 h without shaking. Solubilized proteins were centrifuged at 20,000 rpm for 30 min at 4°C and ultra-filtrated (pore size 0.22 μm) to remove precipitates.

Proteins were purified on HisTrap FF columns (GE Healthcare). On-column refolding was performed using a linear gradient of 6 − 0 M Gnd·HCl buffer containing 50 mM Tris·HCl (pH 8.0), 500 mM NaCl, 20mM imidazole and 1 mM BME. Protease inhibitor cocktail tablets (Roche) were added to 0 M Gnd·HCl refolding buffer. Refolding was performed at a flow rate of 0.5 ml min^{-1} for a total of 800 min by fast protein liquid chromatography (FPLC) (ÄKTA Avant; GE Healthcare). Purified proteins were eluted in a solution containing 50 mM Tris·HCl (pH 8.0), 500 mM NaCl, 500 mM imidazole and 1 mM BME supplied with protease inhibitors cocktail (Roche). After elution, 1 μl of eluate was taken from each fraction for verification on SDS-PAGE. Fractions containing the desired protein were pooled and concentrated using Amicon ultra centrifugal filters at 3 K cut-off (Millipore). The concentrated 200−250 μl protein solutions were dialyzed against a redox buffer containing 100 mM L-arginine, 1 mM oxidized glutathione, 10 mM reduced glutathione and 50 mM Tris·HCl (pH 8.0) at 4 °C for 3 d in order to properly fold the disulfide bonds. Finally, this redox buffer was exchanged into the final assay buffer containing 50 mM Tris·HCl (pH 8.0) using Amicon 3 K filters. Purified proteins were centrifuged at 14,400 rpm for 30 min at 4 °C to remove precipitates. For assay-ready purified proteins, the concentration was evaluated on a 2100 Bioanalyzer instrument (Agilent).

For recombinant proteins expressed as native protein from the supernatant, purification was performed by FPLC (ÄKTAxpress; GE Healthcare) at 4 °C. Similarly, 1000 ml of *E. coli* culture was centrifuged and resuspended in 30 ml suspension buffer containing 50 mM Tris·HCl (pH 8.0), 500 mM NaCl and 20 mM imidazole supplied with protease inhibitors cocktail (Roche). Bacterial cells were lysed using pre-cooled French press. Cell lysate was centrifuged immediately at 14,400 rpm for 30 min at 4 °C to remove cell debris and impurities. Supernatants were loaded onto the HisTrap FF column, washed with the same suspension buffer and eluted in a buffer containing 50 mM Tris·HCl (pH 8.0), 500 mM NaCl and 500 mM imidazole supplied with protease inhibitor cocktail (Roche).

4.2.5 Bead Assay

PT attraction by candidate proteins was evaluated by the bead assay, following a protocol described by Okuda, Tsutsui et al (Okuda, Tsutsui et al. 2009). As indicated in this protocol, pre-conditioned medium was prepared using 10 whole ovaries from *S. chacoense* flowers at anthesis stage submerged in 1 ml fresh BK liquid medium overnight. Guidance response of elongating PTs to purified proteins was observed in bright field with an Axio Observer.Z1 microscope equipped with an AxioCam HRm camera (Zeiss).

4.3 Results

4.3.1 Candidate Selection Strategy

As described previously in Chapter II, RNA-seq of *S. chacoense* ovules identified 59 CRPs, all of which showed ovule-enriched expression as compared to leaf (5-fold cut-off, p-value ≤ 0.05) (Table S2.7). Of these, 24 CRPs were 2-fold more abundantly expressed (p-value ≤ 0.05) in attracting ovules (anthesis) vs. nonattracting ovules (*frk1* and 2DBA), which provides a first pool of candidates.

Another four CRPs (ScCRP3.1, 4.1, 6.1, 15.1) were detected in the ovule secretome through mass spectrometry (Chapter III). Of these, ScCRP4.1 was 2-fold more abundantly secreted (p-value ≤ 0.05) from attracting ovules (anthesis) vs. nonattracting ovules (2DBA) (Table S3.1), whereas ScCRP3.1, 6.1 and 15.1 were not significantly regulated at the protein or RNA levels. Due to their presence in anthesis ovule exudates, a bioactive microenvironment in which PTs

were precisely targeted to the micropyle of the ovule, these four CRPs were added to the candidate list obtained from transcriptome analysis. Table 4.1 summarized the main characteristics of these 28 candidate CRPs.

Of the 28 candidates, 18 genes were obtained through previous comparative transcriptomic studies of the same samples (A, 2DBA, and *frk1* ovules), using the 454 and Illumina assemblies individually. The other 10 candidates were recently acquired through an optimized hybrid 454/Illumina assembly from the transcriptomic study (Liu, Joly et al. in preparation), as well as through the ovule secretomic study (Liu, Joly et al. 2015). This chapter will focus on the expression, purification and functional assays of the primary 18 candidates.

Table 4.1 Summary of candidate genes for ovular PT guidance in *S. chacoense*

Candidate ID	CRP subgroup	Mature length(aa)[a]	Cys. count	Selection criteria[b]				RNA Expression level[c]			Secretion
				1	2	3	4	Anth/*frk1*	Anth/2DBA	Anth/Leaf	Anth/2DBA
ScCRP1.1	EC1-like	205	7	•			•	4.5***	20.8***	2254.0***	
ScCRP1.2	EC1-like	115	8	•			•	5.0***	64.3***	461.8***	
ScCRP1.3	EC1-like	118	7	•			•	5.1***	29.5***	5824.2***	
ScCRP1.4	EC1-like	103	6	•			•	3.9***	33.5***	2416.4***	
ScCRP1.5	EC1-like	124	6	•			•	3.6***	23.0***	1245.0***	
ScCRP1.6	EC1-like	104	6	•			•	4.2***	89.0***	400.9***	
ScCRP1.7	EC1-like	118	7	•			•	5.3***	22.7***	1121.7***	
ScCRP1.8	EC1-like	87	6	•			•	3.4***	3.0***	491.3***	
ScCRP2.1	Novel ScCRP (4-cysteine motif)	60	6	•			•	7.6***	inf.***	1227.0***	
ScCRP2.2	Novel ScCRP (4-cysteine motif)	59	6	•			•	8.0***	inf.***	295.9***	
ScCRP2.3	Novel ScCRP (4-cysteine motif)	56	6	•			•	6.7***	60.0***	1080.7***	
ScCRP2.4	Novel ScCRP (4-cysteine motif)	NA	7	•			•	6.3***	inf.***	448.5***	
ScCRP2.5	Novel ScCRP (4-cysteine motif)	49	6	•			•	7.3***	inf.***	529.6***	
ScCRP4.2	ECA1 gametogenesis related-like	135	7	•			•	4.7***	24.2***	4348.3***	
ScCRP4.3	ECA1 gametogenesis related-like	158	6	•			•	5.5***	inf.***	2353.1***	
ScCRP4.4	ECA1 gametogenesis related-like	NA	6	•			•	4.3***	56.0***	1009.3***	
ScCRP5.1	DEFL	60	6	•			•	4.7***	93.0***	1667.1***	
ScCRP5.2	DEFL	55	6	•			•	6.0***	inf.***	747.8***	
ScCRP5.3	DEFL	NA	7	•			•	5.1***	inf.***	740.3***	
ScCRP7.1	S-protein homologue precursor-like	NA	7	•			•	3.0***	5.1***	2565.2***	
ScCRP8.1	Snakin-like	89	13	•			•	3.1***	62.0***	1123.4***	
ScCRP9.1	DEFL (TfLURE CRP motif)	NA	10	•			•	3.7***	22.0***	390.9***	
ScCRP11.1	DEFL	NA	8	•			•	4.0***	16.0***	41.7***	
ScCRPx.66	Novel ScCRP (6-cysteine motif)	43	6	•			•	2.2***	9.4***	105.6***	
ScCRP3.1	LTP-like	89	9			•	•	1.9***	18.6***	8357.3***	1.2
ScCRP4.1	ECA1 gametogenesis related-like	86	7		•	•	•	0.8*	5.3***	13.8***	5.9*
ScCRP6.1	Thionin-like	89	12			•	•	1.4*	0.5***	1188.3***	-4.9
ScCRP15.1	Thionin-like	133	18			•	•	0.9	1.5**	682.2***	11.2

a. NA indicates incomplete sequence
b. Criterion 1. More abundantly expressed in attracting ovules (anthesis) vs. nonattracting ovules (*frk1* and 2DBA) by 2-fold cut-off and *p*-value ≤ 0.05 (RNA-seq); Criterion 2. More abundantly secreted from attracting ovules (anthesis) vs. nonattracting ovules (2DBA) by 2-fold cut-off and *p*-value ≤ 0.05 (mass spectrometry, MS); Criterion 3. Present in anthesis ovule secretome (MS); Criterion 4. Ovule-enriched expression as compared to leaf by 5-fold cut-off, *p*-value ≤ 0.05) (RNA-seq).
c. *p*-value ≤ 0.001^{***} ; *p*-value ≤ 0.01^{**} ; *p*-value ≤ 0.1^{*}; inf. refers to infinity values (division by zero)

4.3.2 RT-PCR Confirmation of Candidate Genes Expression In Attracting Ovules With Low or Absent RNA Levels In Nonattracting Ovules

Prior to expression and purification of the 18 candidate genes, RT-PCRs were performed to validate their relative abundance in attracting ovules (anthesis) vs. nonattracting ovules (2DBA and *frk1*). As shown in Figure 4.1, RT-PCR indicated that these selected genes were highly expressed in anthesis ovules, as opposed to low or zero expression in nonattracting ovules, consistent with the differential gene expression levels predicted by RNA-seq (Table S2.7). Therefore, gene expression and protein purification experiments were carried out for these validated candidates.

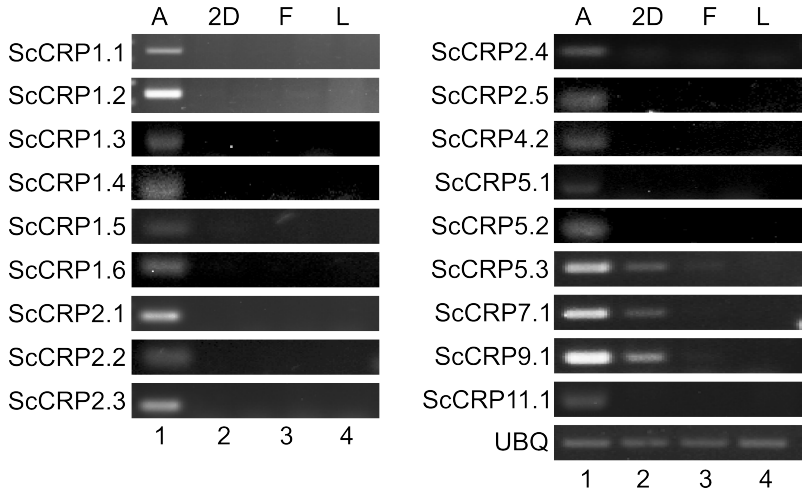


Figure 4.1 RT-PCR validation of 18 candidate genes. Lane 1, wild-type ovules at anthesis stage (A); lane 2, immature ovules at 2DBA stage (2D); lane 3, *frk1* mutant ovules at anthesis stage (F); lane 4, leaf (L). Ubiquitin (UBQ) was used as an internal control.

4.3.3 Candidate Protein Purification

In order to obtain sufficient amount of candidate proteins, a bacterial expression protocol was followed. Except for ScCRP 4.2, which was not expressed using either pGWB8, pET28b or pGEX4T2 vectors in either *E. coli* Rosetta-gami 2 or BL21 strains, all ScCRPs were cloned into pET28b vector, expressed in Rosetta-gami 2 strain and purified by affinity chromatography. Among these, ScCRP1.2, ScCRP1.6 and ScCRP5.1 proteins went through a second separation step of gel filtration for impurity removal (possible multimers arising from intermolecular disulfide bond formation). Here the expression and purification of ScCRP5.1 is reported as an example.

ScCRP5.1 expression was induced by 1mM of IPTG and was expressed in the supernatant in *E. coli*. After purification on the column, fractions were collected and loaded onto a 15% polyacrylamide gel. In Figure 4.2A, five tested fractions showed the desirable size of ScCRP5.1 at 15 kDa. The expression of *ScCRP5.1* was later confirmed by LC-MS/MS sequencing of the cut-band from this gel. In order to remove the impurities from ScCRP5.1, gel filtration chromatography was performed. As shown in Figure 4.2B, larger size impurities eluted first, leaving the fractions after fraction 3 with a single band for ScCRP5.1. These latter fractions were pooled, concentrated and dialyzed into 50 mM Tris·HCl (pH 8.0) for the bead assay.

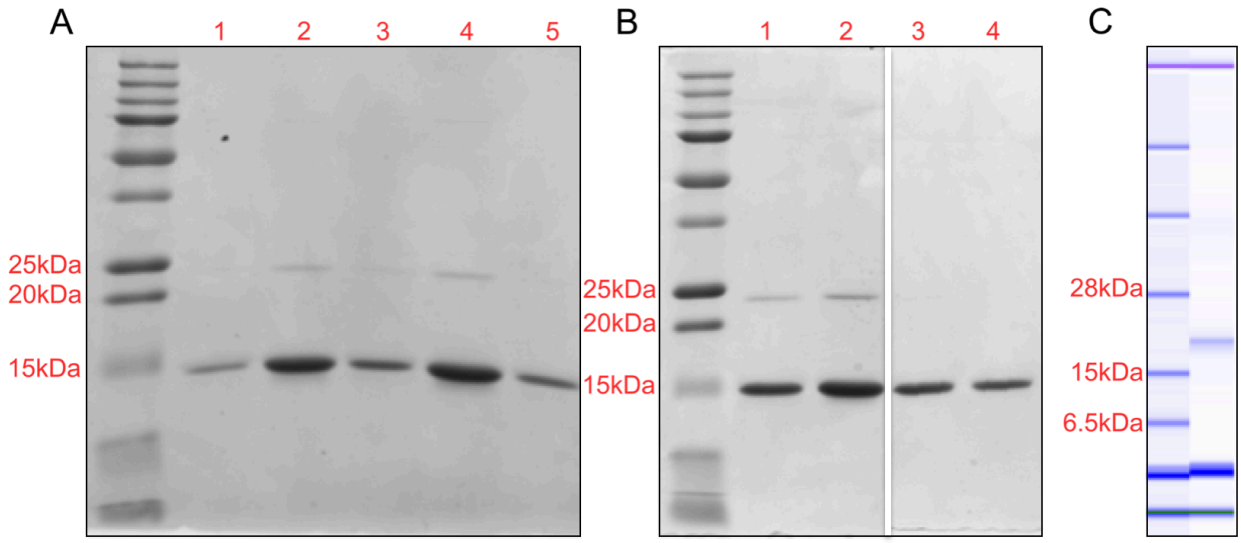


Figure 4.2 Protein purification of ScCRP5.1 by FPLC. (A) After affinity chromatography on a nickel column, polyhistidine-tagged ScCRP5.1 were collected and detected in fractions 1-5 on a 15% SDS-PAGE gel. (B) Gel filtration chromatography using the Superdex 200 column separated ScCRP5.1 at 15 kDa from impurities at 25 kDa after fraction 3. Multiple fractions were collected during gel filtration, while only four representative fractions (1-4) were shown in Figure 4.2B. (C) A simulated gel was constructed by the 2100 Bioanalyzer, showing a major band at ~15 kDa for ScCRP5.1, for the purpose of calculating protein concentration and purity of ScCRP5.1.

Before testing guidance competency of ScCRP5.1, it was loaded onto a microfluidic chip performed on the 2100 Bioanalyzer (Agilent) to evaluate the concentration and purity of this recombinant protein. A simulated gel is shown in Figure 4.2C. The purity of ScCRP5.1 was reported to be 78% and the concentration of this ScCRP5.1 major band was 150 ng/µl. Considering the different size of each candidate protein, the weight concentration was converted to molar concentration. Following these procedures, the final concentration of each candidate used for the PT guidance functional assay, the bead assay, was calculated and summarized in Table 4.2. The newly found 10 candidates are also listed in Table 4.2 in gray color.

Table 4.2 Summary of candidate protein expression and purification

Candidate ID	CRP subgroup	Mature length(aa)	Cys. count	Selection criteria				Protein expression and purification		Attraction competency
				1	2	3	4	Expressed in	Concentration	
ScCRP1.1	EC1-like	205	7	•			•	IB	3 μM	0/30
ScCRP1.2	EC1-like	115	8	•			•	IB	8 μM	2/30
ScCRP1.3	EC1-like	118	7	•			•	IB	67 μM	0/30
ScCRP1.4	EC1-like	103	6	•			•	IB	7 μM	2/30
ScCRP1.5	EC1-like	124	6	•			•	IB	800 μM	0/30
ScCRP1.6	EC1-like	104	6	•			•	IB	915 μM	0/30
ScCRP2.1	Novel ScCRP (4-cysteine motif)	60	6	•			•	IB	23 μM	4/30
ScCRP2.2	Novel ScCRP (4-cysteine motif)	59	6	•			•	IB	110 μM	3/30
ScCRP2.3	Novel ScCRP (4-cysteine motif)	56	6	•			•	IB	160 μM	0/30
ScCRP2.4	Novel ScCRP (4-cysteine motif)	NA	7	•			•	IB	14 μM	4/30
ScCRP2.5	Novel ScCRP (4-cysteine motif)	49	6	•			•	IB	23 μM	0/30
ScCRP4.2	ECA1 gametogenesis related-like	135	7	•			•	not expressed	NA	NA
ScCRP5.1	DEFL	60	6	•			•	Supernatant	15 μM	3/50
ScCRP5.2	DEFL	55	6	•			•	Supernatant	16 μM	5/50
ScCRP5.3	DEFL	NA	7	•			•	Supernatant	19 μM	2/50
ScCRP7.1	S-protein homologue precursor-like	NA	7	•			•	IB	52 μM	0/30
ScCRP9.1	DEFL (TfLURE CRP motif)	NA	10	•			•	Supernatant	12 μM	2/50
ScCRP11.1	DEFL	NA	8	•			•	IB	96 μM	5/67
ScCRPx.66	Novel ScCRP (6-cysteine motif)	43	6	•			•			
ScCRP3.1	LTP-like	89	9			•	•			
ScCRP4.1	ECA1 gametogenesis related-like	86	7		•	•	•			
ScCRP6.1	Thionin-like	89	12			•	•			
ScCRP15.1	Thionin-like	133	18			•	•			
ScCRP1.7	EC1-like	118	7	•			•			
ScCRP1.8	EC1-like	87	6	•			•			
ScCRP4.3	ECA1 gametogenesis related-like	158	6	•			•			
ScCRP4.4	ECA1 gametogenesis related-like	NA	6	•			•			
ScCRP8.1	Snakin-like	89	13	•			•			

In parallel to the bacterial expression system, a plant-based expression system was also used to produce an active form of candidate proteins for PT attraction functional assays. Agroinfiltration-mediated transient expression in *N. benthamiana* was carried out for ScCRP1.4 and ScCRP5.1 proteins. First, RT-PCR was performed to verify their gene expression in infiltrated leaves. Figure 4.3A showed that *ScCRP1.4* and *ScCRP5.1* were overexpressed in infiltrated leaves. Consistent with the transcriptional level, western blot analysis against the polyhistidine-tag confirmed their protein expression in infiltrated leaves (Figure 4.3B).

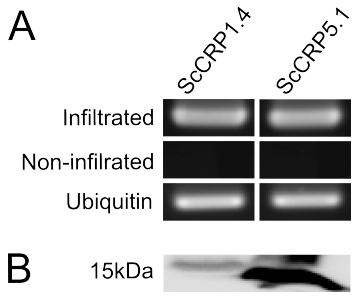


Figure 4.3 Validation of candidate gene expression in *N. benthamiana* leaves. (A) RT-PCR verification of the transcription status of *ScCRP1.4* and *ScCRP5.1* in infiltrated leaves harvested two days after infiltration. (B) Western blot verification of protein expression of ScCRP1.4 and ScCRP5.1. For each lane, 20 μg of total proteins extracted from *N. benthamiana* leaves were loaded on the SDS-PAGE.

Although *ScCRP1.4* and *ScCRP5.1* were expressed in this system, the purification of these two proteins from *N. benthamiana* leaves generated insufficient amount of proteins, which were undetectable on the SDS-PAGE. As a result, the bacterial expression system is used mainly for CRP candidate expression.

4.3.4 Guidance Competency of Candidate Proteins

The guidance competency of candidate proteins was evaluated with the bead assay system, where highly purified proteins are encapsulated into a gelatinous matrix to form microbeads that can be precisely positioned along the path of a PT in order to determine their chemoattracting potential (Okuda, Tsutsui et al. 2009).

Prior to the screening of 17 purified candidates, a positive control for the bead assay was sought. Since wild-type anthesis ovules attract semi *in vivo* grown PTs in *S. chacoense*, ovule exudates were collected as described (Liu, Joly et al. 2015) and tested for their guidance competency. Collection of high concentration of ovule exudates from ~200 ovaries did not induce chemotactic behavior of the PTs in the bead assay. Instead, repulsive response of the PTs to the beads was observed. This may be due to the phenolic compounds that are ubiquitous in plants being collected and placed in the beads in large quantities. As a result, exudates were obtained using lesser ovaries (here only 4 dissected ovaries). After overnight incubation on solid BK medium, exudates were freshly collected, as described previously (Liu, Joly et al. 2015), and used immediately in the bead assay. Figure 4.4 showed that ovule exudates prepared in this method induced chemotactic behavior of 43% of PTs (n = 28), which compares favorably with the percentage observed in the semi *in vivo* assays using mature ovules (59%) in *S. chacoense* (Lafleur, Kapfer et al. 2015). Therefore, mature ovule exudates can be used as a positive control in the candidate screening. The attraction response was observed within 5-10 min after placing the bead. In contrast, BK liquid medium alone in the beads did not attract PTs (0%, n =15).

Seventeen purified candidates were tested in the bead assay, both individually and when all combined. However, PT guidance competency was not observed for current candidates, as summarized in Table 4.2. Examples were given in Figure 4.5.

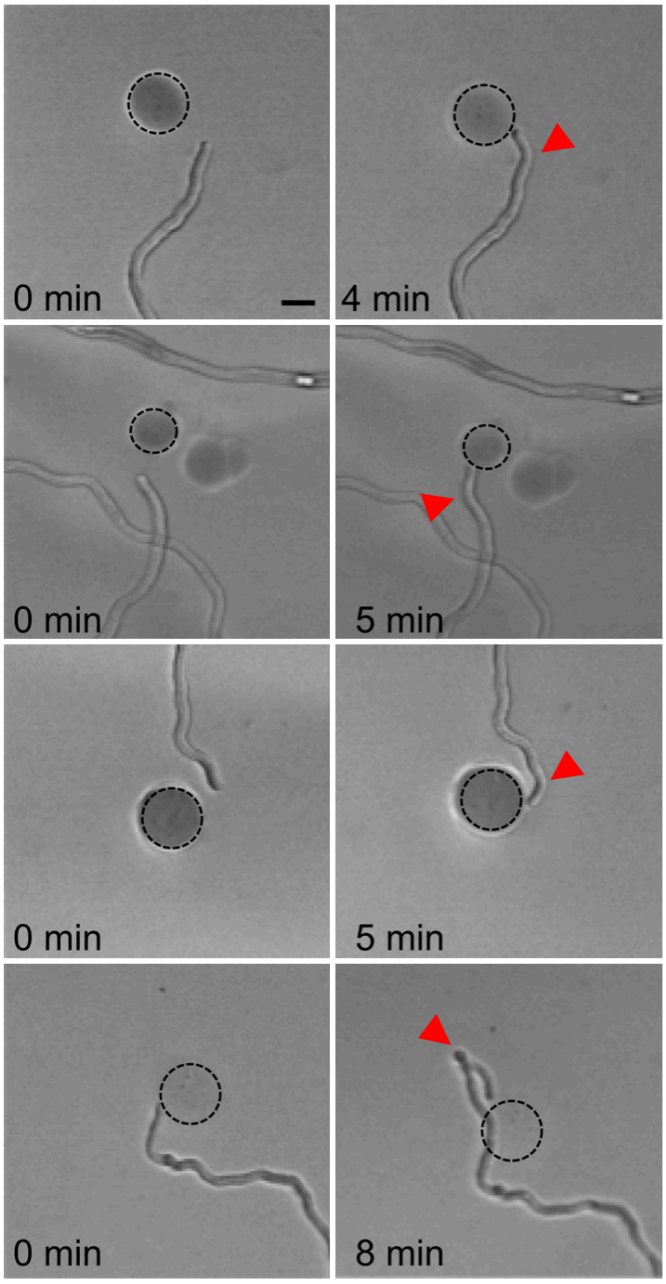


Figure 4.2 Chemotropism of semi *in vivo* grown PTs toward ovule exudates collected from flowers at anthesis stage in *S. chacoense*. Four examples were shown. Pictures were taken immediately after placing the beads and marked as time zero on the left panel. The right panel shows PT behavior 4-8 min after placing the beads. Dotted lines define the position of the beads. Arrowhead indicates the position where PTs turn toward ovule exudates. Scale bar = 20 μm.

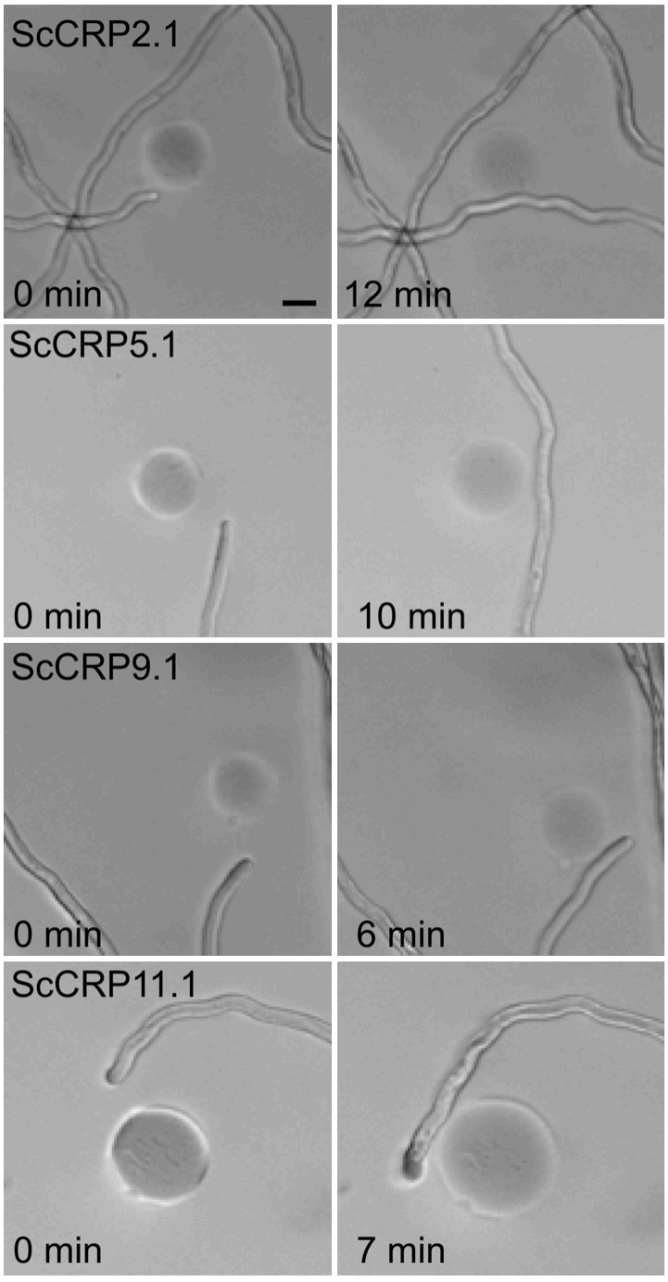


Figure 4.3 Purified candidates did not induce chemotactic behavior of semi *in vivo* grown PTs in the bead assay. The guidance competency of purified candidates were tested both individually and when all combined. Examples were given for candidates ScCRP2.1, ScCRP5.1, ScCRP9.1 and ScCRP11.1. PTs continued growing after passing by the beads. Pictures were taken immediately after placing the beads and marked as time zero on the left panel. The right panel shows PT behavior 6-12 min after placing the beads. Scale bar = 20 μm.

4.4 Discussion

4.4.1 Expression of Candidates In A Biologically Active Conformation For Functional Assay

At the beginning of this project, two methods were tried in order to express candidates in an active form. First, a bacterial expression system using the expression vector pBAD/gIII (Invitrogen) was chosen to prevent protein aggregation into inclusion bodies by targeting recombinant proteins to the periplasmic space of *E. coli*. Nevertheless, all candidate genes were still expressed in the inclusion bodies (data not shown). Alternatively, *Agrobacterium*-mediated transient expression of ScCRP1.4 and ScCRP5.1 proteins in *N. benthamiana* leaves (Figure 4.3) did not generate sufficient amount of soluble proteins for the functional assays.

Finally, all genes were cloned into the pET28b vector and expressed in *E. coli* strain Rosetta-gami 2, which allows for enhanced disulfide bond formation for CRP candidates. As shown in Table 4.2, among expressed candidates, only four genes were expressed in the soluble fraction of *E. coli*. For the rest, refolding was required and conducted according to the LURE-type CRP refolding protocol (Okuda, Tsutsui et al. 2009, Takeuchi and Higashiyama 2012).

Currently, no guidance activity has been detected for the 17 purified CRPs. This project is still ongoing, aiming to express the other 11 candidate CRPs listed in Table 4.1. However, one reason that hindered us from isolating the chemoattractants may lie in the limited activity of recombinant proteins, since the refolding protocol adapted from the LURE-type CRPs research may not be efficient to properly refold CRPs identified in *S. chacoense*, especially when certain cysteine motifs were found to be solanaceous-specific (Figure 2.4).

To remedy this, our laboratory plans to work on producing biologically active proteins in transgenic *S. chacoense* roots transformed with *Agrobacterium rhizogenes*, following a protocol described previously (Rivoal and Hanson 1994). In this system, candidate genes will be constructed in the binary vector pGA643, followed by transformation into *A. rhizogenes* strain A4. *Solanum chacoense* petioles will be cut into 4 cm segments. Each segment will be inserted in MS solid medium (Murashige and Skoog 1962) at one end and infected with agrobacteria at the other end. After 10 d, roots will develop and be excised from the infection sites. RT-PCR

validated transgenic roots that carry candidate genes will be decontaminated using 100μg ml^{-1} carbenicillin. Sterile roots will be treated with 2,4-D to obtain friable calli. The fine callus can be used directly to perform guidance test in the semi *in vivo* assay (Lafleur, Kapfer et al. 2015). Guidance competency can be determined if semi *in vivo* PTs grow toward the callus expressing secreted candidate proteins. Alternatively, the callus can be cultured in liquid medium, where secreted candidate proteins can be collected, lyophilized and tested in the bead assay. In either way, protein purification and refolding processes can be avoided and therefore a higher amount of recombinant proteins can be expected as compared to the *N. benthamiana* expression system.

In parallel to this plant expression system, the yeast expression system using the pPICα vector in *Pichia pastoris* strain (X-33, SMD1168H, or KM71H) will also be used to express candidate genes. The pPICα vector allows for the expression of recombinant proteins as fusion proteins to an N-terminal peptide encoding the *Saccharomyces cerevisiae* α-factor secretion signal (Thermal Fisher Scientific). This will result in the sufficient secretion of CRP candidates into the culture medium. In addition, *P. pastoris* strongly prefers respiratory growth, which creates a great advantage to its culturing in high cell-density environment for high-yield protein production, as compared to fermentative yeast strains. At last, this strain is capable of many post-translational modifications performed by higher eukaryotic cells, such as folding, disulfide bond formation, and both N- and O-glycosylation.

As shown in Figure 2.5 and Table 2.1, glycoprotein fractions purified through ConA column from anthesis ovule extracts enhanced PT attraction by 18%, as compared to total ovule extracts in *S. chacoense*, suggesting that the chemoattractants may be bearing N- or O-glycosylation modification. The use of *Pichia* yeast expression system will largely facilitate the expression of potentially glycosylated CRP candidates.

4.4.2 Optimization of PT Growth Medium For the Bead Assay

In *Torenia,* semi *in vivo* grown PTs become competent to react to micropylar chemoattractant LURE2 only when they travel a sufficient amount of time (7 h) after emerging on the growth medium. PTs that only freshly exit the bottom end of the style (1 h) do not respond to LURE2, which suggests that the additional PT elongation on the medium after exiting the style is important for the acquisition of LURE-binding activity (Okuda, Suzuki et al. 2013).

In *S. chacoense*, PT growth is supported by the BK medium (Brewbaker and Kwack 1963), which contains 0.1 mg/ml H_3BO_3, 0.3 mg/ml $Ca(NO_3)_2$, 0.1 mg/ml KNO_3 and 0.2 mg/ml $MgSO_4$, as supplied by 5% of sucrose. The pH of the medium is adjusted to 5.5.

In the bead assay, with limited pollination to obtain individual tubes, semi *in vivo* PTs exit the style and can usually continue to grow on the BK solid medium for ~ 5 h before growth arrest or tube burst. To avoid using freshly exited PTs, PTs that elongated ~ 3 h after emerging from the style were used in the bead assay. It is possible that, at this time point, PTs are not competent enough to react to recombinant peptides, even though ovule exudates were used with success to guide PTs under the same condition, which is likely due to the signaling molecules involved in competency control of PTs that reside in the exudates.

Following this rationale, the maximum 5 h of PT elongation on the medium after exiting the style may become a bottleneck for the functional assay. Therefore, optimization on different ion concentrations, sucrose concentration of the BK medium may be tried to support longer PT growth and to increase the chance of isolating the attractants in *S. chacoense.*

4.5 Supplementary data

Table S4.1 Primer sequences used in RT-PCR of candidate proteins.

Primer name	Primer sequence
ScCRP1.1_F	CCCGGATCCGGTCCAAGGTTTCGGGCC
ScCRP1.1_R	CCCAAGCTTTTAATCTCCTTTATTTGCAAATGAAAA
ScCRP1.2_F	CCCGGATCCGCAATTACTCGAAGGAGGAAGA
ScCRP1.2_R	CCCAAGCTTTTAATTTTTCGTACTATCTTCATTCTTCAC
ScCRP1.3_F	CCCGGATCCGAGCGTCTCACAGTTACTCGG
ScCRP1.3_R	CCCAAGCTTTTAGTTCTTAATTTTGCGTACAATCTT
ScCRP1.4_F	CCCGGATCCGACCAACATGTCCAAAATTAAACC
ScCRP1.4_R	CCCAAGCTTCTACAGCGATTCAGGTGAAGG
ScCRP1.5_F	CCCGGATCCGACAATCACAAATTTAAGCCATGTT
ScCRP1.5_R	CCCAAGCTTTCACTCATTTATTATATCCAAATTAGG
ScCRP1.6_F	CCCGGATCCGCGACCTTTAATCACATTTAGTAGTGC
ScCRP1.6_R	CCCAAGCTTTTATGGAGCCAAATTAGGAGTTG
ScCRP2.1_F	CCCGGATCCGTATCCTCCAAGAATAGAAAGTATTTGT
ScCRP2.1_R	CCCAAGCTTTCACCCACTAATATAACCCCCA
ScCRP2.2_F	CCCGGATCCGTATCCTCCACGAATAGAAAGTGTTT
ScCRP2.2_R	CCCAAGCTTTCACCCACTAATATAACCCCCA
ScCRP2.3_F	CCCGGATCCGTATCCTCCATCAATAGAAGAAACTTGT
ScCRP2.3_R	CCCAAGCTTTTATGAGCCATGAGTCTCATCATATT
ScCRP2.4_F	CCCGGATCCGAATTGTAAAAAGTACCCTCCATCAA
ScCRP2.4_R	CCCAAGCTTTTACAAATTATCGGACCTTTTATATAGGTC
ScCRP2.5_F	CCCGGATCCGTGTACCCGATCTCCTTCATCA
ScCRP2.5_R	CCCAAGCTTTTATTCAAGATAACAATTAAATTCAAGAAG
ScCRP4.2_F	CCCGGATCCATGGCAAGACTCAATATCTTCTCAT
ScCRP4.2_R	CCCAAGCTTTTATGGTAGTATTGGAGGTGAAAAGG
ScCRP5.1_F	CCCGGATCCGAATAAATTTGTAGGTGGATTAGCATG
ScCRP5.1_R	CCCAAGCTTTCAGGATTTAGACTCTGTCGGTG
ScCRP5.2_F	CCCGGATCCGTATCCAAATGAATGCTTAAAAGATGA
ScCRP5.2_R	CCCAAGCTTTTAAGCATGAGATAAATTTGGATCATT
ScCRP5.3_F	CCCGGATCCGTTTGGAGATTGTTATCATGATGG
ScCRP5.3_R	CCCAAGCTTTTATAAATTTAGATTATCATCTGTATATGTAATACCAA
ScCRP7.1_F	CCCGGATCCGTATAATTGTAATATTCCTGATTTTCCAGTATG
ScCRP7.1_R	CCCAAGCTTTTATTGGGAATACATTTTATCACCTTTAAC
ScCRP9.1_F	CCCGGATCCGATGAGAATTGTACCAATGATATATGGA
ScCRP9.1_R	CCCAAGCTTTTACTTGCAACAATCTGTGCCAT
ScCRP11.1_F	CCCGGATCCGAGGAATTTGGAGAGGAAATG
ScCRP11.1_R	CCCAAGCTTTTAAACAAGAATACATCTATAAATACC

4.6 Acknowledgments

I thank Minako Kaneda from laboratory Geitmann for providing the pET28b vector. I thank Jonathan Soulard and Mathieu Beauchemin for their critical review and detailed corrections for this chapter. This work was supported by FRQNT (Fonds de Recherche du Québec – Nature et technologies).

5. Conclusion and Perspectives

The aim of this project was to study ES-expressed genes that play an important role in pollen–pistil interactions, more specifically, in PT guidance processes. To this end, ovular signals were studied at both transcriptome and proteome levels. Comparative analyses were conducted on PT attracting ovules (anthesis) vs. PT nonattracting ovules (2DBA and *frk1*) to identify the chemoattractants.

Chapter II provides an overview of ovule transcriptome profiling in *S. chacoense*. First, we *de novo* assembled the ovule transcriptome for *S. chacoense* and its two close relatives *S. gandarillasii* and *S. tarijense.* This part highlights the design of a data curation protocol that has the strength to detect and repair ~15000 frameshifted and chimeric contigs, accounting for 15% of total assembled sequences. This protocol allowed us to conduct transcriptome analyses on solid ground and can be readily adapted to the curation of transcriptome in other non-model species. Comparative analyses revealed 818 genes as ES-expressed candidates (anthesis vs. *frk1*) and 1355 genes involved in late stages of ovule maturation (anthesis vs. 2DBA). A subset of 284 genes was found present in anthesis ovules and was low/absent in *frk1* and 2DBA ovules. These genes are likely to assume an ES-localization and especially associated with late stages of ES maturation during the FG6 to FG7 transition. Meanwhile, since neither 2DBA nor *frk1* ovules attract PTs in the semi *in vivo* assay, whereas anthesis ovules are fully functional in attracting PTs, these 284 genes constitute a candidate gene pool for ovular PT guidance in *S. chacoense*.

Second, this study demonstrated that the glycoprotein fraction of ovule extracts enhanced PT guidance competency by 18%, as compared to total ovule extracts, suggestive of a positive regulation between the glycosylation status of ovular signals and PT guidance. Such relationship has only been observed in the pre-ovular guidance stage, where heavily glycosylated TTS proteins in the style attract semi *in vivo* grown PTs in tobacco and *N. alata* (Cheung, Wang et al. 1995, Wu, Wong et al. 2000). During ovular guidance stage, identified micropylar chemoattractants lack this post-translational modification (Marton, Cordts et al. 2005, Okuda, Tsutsui et al. 2009, Kanaoka, Kawano et al. 2011, Takeuchi and Higashiyama 2012). Nevertheless, glycosylation can still be adopted in other guidance steps, i.e. funicular and long-distance guidance, or in the same micropylar guidance step, which would be unique to *Solanum*

species. As a next step, it will be interesting to characterize the protein composition in the glycoprotein fraction of ovule extracts and to identify the active component that enhances PT guidance in *S. chacoense*. It is possible that this protein belongs to the CRP superfamily, since various CRPs are shown to play a role in pollen–pistil interactions, especially the micropylar guidance chemoattractants in *Arabidopsis* or *Torenia* are small secreted CRPs. Or it could also be a high molecular weight glycoprotein, like the stylar TTS (PT growth stimulant and attractant) (Cheung, Wang et al. 1995, Wu, Wong et al. 2000) and the stylar 120K glycoproteins (capable of binding S-RNase) (Cruz-Garcia, Nathan Hancock et al. 2005). To have a general idea of size of this protein, the glycoprotein faction of total ovule extracts in *S. chacoense* can be separated into low molecular weight (< 30 kDa) and high molecular weight fractions (> 100 kDa), using Amicon ultra centrifugal filters. Fractions containing low, medium and high molecular weight glycoproteins will be applied to the two-choice semi *in vivo* guidance assay individually, to narrow down the active component in enhancing PT guidance. Then mass spectrometry sequencing will be used to study and identify this key molecule.

So far, with 33852 CDS-containing contigs assembled for *S. chacoense* ovule transcriptome, Chapter II mainly focused on 818 ES-expressed genes, which accounts for 2.4% of the total transcriptome, leaving the vast majority of data open to future analyses. Two studies have been planned to carry on the ovule transcriptomic study in *S. chacoense*. According to differential gene expression analyses, 679 genes showed at least a 2-fold increased expression in *frk1* ovules that entirely lack an ES, compared to anthesis ovules (Table S2.4). These genes are expressed in the sporophytic tissues of the ovule, whose expression is repressed by the presence of an ES. This special gene regulation suggests a gametophyte to sporophyte crosstalk, substantiating these genes in FG development and possibly in pollen–pistil interactions. Now that the micropylar guidance cues have been isolated from the synergid cells of *Torenia* and *Arabidopsis* and from the egg apparatus (egg cell and synergid cells) of maize, other signaling molecules involved in funicular guidance, long-distance guidance and the competency control of PT guidance are largely unexplored. Ovular sporophytic tissues may possibly possess these important key factors, which can be isolated through studying the 679 *frk1*-upregulated genes in *S. chacoense*.

Another projected study is to analyze the transcription factors found in *S. chacoense*, focusing on the ones enriched in the reproductive tissues (ovules) relative to the sporophytic tissues (leaf),

using current differential gene expression analyses. During FG development, transcription factors play a pivotal regulatory role during female gamete differentiation. For example, in *Arabidopsis*, two paralogous *R2R3-MYB* transcription factors, namely *FOUR LIPS* and *MYB88*, affect ovule development and fertility by regulating the cell cycle entry into megasporogenesis (Makkena, Lee et al. 2012). During late stages of megagametogenesis, *MYB98* is required for synergid cell differentiation (Kasahara, Portereiko et al. 2005) and regulates a subset of genes that are required during filiform apparatus formation and PT guidance (Punwani, Rabiger et al. 2007, Punwani, Rabiger et al. 2008). In *Arabidopsis*, 1428 transcription factors were identified, and interestingly, 17 were confirmed for their specific expression in the FG using promoter fusion analysis (Wang, Zhang et al. 2010).

From our preliminary results, Figure S2.2B reported that 624 genes were associated with "transcription factor activity" (GO: 0001071), which becomes the entry point for dissecting the regulatory networks in *S. chacoense* ovules. Of these, 124 genes showed ovule-enriched expression (≥ 5-fold more abundantly expressed in anthesis ovules than in leaf, p-value ≤ 0.05). These genes encoded MYB, MADS-box, HD-Zip and ARFs-type transcription factors. By comparing 2DBA and *frk1* ovule transcriptome to anthesis one, a series of transcription factors that are specifically associated with late stages of ES maturation can be identified and characterized in future studies.

Chapter III describes for the first time the secretome of the plant ovule during its last maturation stages, a period where it becomes competent to attract PTs. In this study, a novel tissue-free collection method was designed to gather ovule exudates that substantially reduced the cytosolic contamination to 0.79%, whereas the conventional collection methods typically show a cytosolic contamination of 1 to 3%. Secretome comparisons using ovules vs. other vegetative tissues (such as roots) revealed that among 305 identified ovule-secreted proteins (OSPs), 58% are ovule-specific, indicating common factors are secreted into the extracellular matrix, regardless of the source tissues. Such factors include pathogenesis-related (PR) proteins, chitinase and β-1,3-glucanase that are generally associated with plant defense against pathogen. Among the ovule-specific secreted proteins (oOSPs), proteins potentially involved in pollen–pistil interactions were identified, including a S-RNase (self-incompatibility), GABA-T (PT growth and funicular

guidance), PELPIII (rejection of interspecific PTs), and four CRPs that are potential key signaling molecules in various pollen–pistil interactions.

Comparative secretomic analysis revealed that 128 OSPs were upregulated in anthesis ovules relative to 2DBA ovules, and thus were defined as candidate proteins involved in ovule development and PT guidance. Of these 128 OSPs, differential gene expression analyses indicated that the majority (83%) were not regulated at the RNA level, which suggested that even though the RNAs were similarly produced at the 2DBA stage, their protein secretion was strictly regulated during ES development and the secretion status was critical to successful PT guidance. Hence, this research presents a secretomic view toward pollen–pistil interactions and PT guidance study. It demonstrated ovule secretome study as a robust complement to transcriptome studies.

To date, genes that are remotely activated from the ovule upon pollination events have been identified in *Solanum* species (van Eldik, Ruiter et al. 1997). Considering the satisfying results obtained from ovule secretome study, another project has been planned to study the pollination-induced effect by collecting ovule exudates from different ovule conditions, in which the style has been pollinated for 0, 24, 39, 42 and 48 h, where fertilization in *S. chacoense* takes place at ~ 40 h after pollination. Through mass spectrometry sequencing and comparative secretomic analysis, secreted proteins that are induced upon pollination and fertilization events can be isolated.

Chapter II and III together laid a solid foundation for studying PT guidance in solanaceous species. In fact, the subject of PT chemotropism has been studied for more than a century. In 1893, Molisch observed *in vitro* PT chemotropism toward acids and secretions of the gynoecium, especially those of the stigma (Molisch 1893). In 1894, Miyoshi discovered that PT directional growth was controlled by sugars (Miyoshi 1984). Later on in 1962, external calcium ions were suggested as a potential PT attractant in *Antirrhinum majus* using the *in vitro* assays (Mascarenhas and Machlis 1962). Only until recent years, the semi *in vivo* assay system was developed (Higashiyama, Kuroiwa et al. 1998). Compared to the *in vitro* pollen germination and growth on the medium, the semi *in vivo* assay allows PTs to travel through the pistil and thus renders these PTs competent to react to the attractants (Qin, Leydon et al. 2009). Using this

system, it was demonstrated that *Arabidopsis* ovules emanate heat-liable, diffusible, developmentally regulated and species-specific PT attractants, pointing to a proteinaceous nature of the attractant (Palanivelu and Preuss 2006).

So far, the chemoattractants in *Arabidopsis* and *Torenia* have been identified, which belong to the CRP superfamily. Through Chapter II and III, over 400 ovule-expressed genes were identified as potentially involved in PT guidance processes in *S. chacoense*. Chapter IV focused on the evaluation of CRPs as chemoattractants in PT guidance. Of the 28 CRP candidates that were significantly under-presented in PT nonattracting ovules (2DBA and *frk1*) vs. attracting ovules (anthesis) at the mRNA level and/or protein secretion level, 17 CRPs were expressed and purified in satisfying quantities. As compared to anthesis ovule exudates that induced PT chemotropism in the bead assay, purified candidates, mixed or individually, did not induce PT chemotropism.

To carry on the last part of this project, recently identified CRPs candidates obtained through ovule transcriptome curation and secretome study will continue to be expressed in the bacterial expression system and tested for their guidance competency. At the same time, a few candidates can be expressed in the *Agrobacterium rhizogene*-mediated plant expression system to achieve active protein expression in transgenic calli, which can be co-incubated directly with semi *in vivo* grown PTs for guidance test, so as to avoid protein purification and refolding processes. Or alternatively, candidates can be expressed in *Pichia pastoris* yeast expression system, which has the advantage of generating secreted proteins with high-yield and is capable of many post-translational modifications, such as protein refolding, disulfide bond formation, N- and O-glycosylation performed by higher eukaryotic cells. Since the chemoattractants in *S. chacoense* are potentially glycosylated proteins as demonstrated by the semi *in vivo* assays (Figure 2.5 and Table 2.1), the yeast expression system will largely contribute to the production of active CRP candidates with potential glycosylation modification.

With the advantages of a 34X sequencing depth of the Illumina data obtained from *S. chacoense* ovules, with an optimized hybrid 454/Illumina assembly that identified more candidate CRPs recently, and with the possibility to perform guidance test using transgenic calli directly without protein purification and refolding steps, the chemoattractants in *S. chacoense* will eventually

come to light. At that cheerful moment, investigation on the sub-cellular localization of these attractants can be conducted by fusing the promoter fragment with GUS reporter gene to study their sub-cellular localizations at the RNA level, and by fusing the protein coding sequence with GUS reporter gene to study protein localization pattern after secretion.

Second, the RNAi knockdown and overexpression lines can be generated to characterize the attractants, in terms of seed set and fruit size, ES integrity and positioning of female gametes, possible male defects, and the guidance defects of mutant ovules *in vivo* and in semi *in vivo* assays. Most interestingly, these examinations will clarify to which guidance step these attractants are specifically associated.

Then orthologous chemoattractants in other solanaceous species can be identified, as facilitated by the available genome sequences and EST databases from several *Solanum* species, as well as by the *S. gandarillasii* and *S. tarijense* ovule transcriptome assembled in this study. The structure/function relationship of the chemoattractants can be studied using site-directed mutagenesis of substituted amino acids detected by sequence alignment. At last, *Solanum* species crosses that have shown breeding barriers in ovular guidance step can be genetically engineered to produce PT-attracting female signals in order to achieve a new seed, which promises the gene flow of reproductive isolated species, and to a broader view, the production of more seeds as food source in the world.

6. Reference

Agrawal, G. K., N. S. Jwa, M. H. Lebrun, D. Job and R. Rakwal (2010). "Plant secretome: unlocking secrets of the secreted proteins." Proteomics **10**(4): 799-827.

Alandete-Saez, M., M. Ron and S. McCormick (2008). "GEX3, expressed in the male gametophyte and in the egg cell of Arabidopsis thaliana, is essential for micropylar pollen tube guidance and plays a role during early embryogenesis." Mol Plant **1**(4): 586-598.

Altschul, S. F., W. Gish, W. Miller, E. W. Myers and D. J. Lipman (1990). "Basic local alignment search tool." J Mol Biol **215**(3): 403-410.

Amien, S., I. Kliwer, M. L. Marton, T. Debener, D. Geiger, D. Becker and T. Dresselhaus (2010). "Defensin-like ZmES4 mediates pollen tube burst in maize via opening of the potassium channel KZM1." PLoS Biol **8**(6): e1000388.

Armenta-Medina, A., W. Huanca-Mamani, N. Sanchez-Leon, I. Rodriguez-Arevalo and J. P. Vielle-Calzada (2013). "Functional analysis of sporophytic transcripts repressed by the female gametophyte in the ovule of Arabidopsis thaliana." PLoS One **8**(10): e76977.

Babiychuk, E., S. Kushnir, E. Belles-Boix, M. Van Montagu and D. Inze (1995). "Arabidopsis thaliana NADPH oxidoreductase homologs confer tolerance of yeasts toward the thiol-oxidizing drug diamide." J Biol Chem **270**(44): 26224-26231.

Badri, D. V., C. De-la-Peña, Z. Lei, D. K. Manter, J. M. Chaparro, R. L. Guimaraes, L. W. Sumner and J. M. Vivanco (2012). "Root secreted metabolites and proteins are involved in the early events of plant-plant recognition prior to competition." PLoS One **7**(10): e46640.

Baumberger, N. and D. C. Baulcombe (2005). "Arabidopsis ARGONAUTE1 Is an RNA Slicer That Selectively Recruits microRNAs and Short Interfering RNAs." PNAS **102**(33): 11928-11933.

Bemer, M., M. Wolters-Arts, U. Grossniklaus and G. C. Angenent (2008). "The MADS domain protein DIANA acts together with AGAMOUS-LIKE80 to specify the central cell in Arabidopsis ovules." Plant Cell **20**(8): 2088-2101.

Bendtsen, J. D., L. J. Jensen, N. Blom, G. Von Heijne and S. Brunak (2004). "Feature-based prediction of non-classical and leaderless protein secretion." Protein Eng Des Sel **17**(4): 349-356.

Bianchi, E., B. Doe, D. Goulding and G. J. Wright (2014). "Juno is the egg Izumo receptor and is essential for mammalian fertilization." Nature **508**(7497): 483-487.

Blanc, G., A. Barakat, R. Guyot, R. Cooke and M. Delseny (2000). "Extensive duplication and reshuffling in the Arabidopsis genome." Plant Cell **12**(7): 1093-1101.

Bleckmann, A., S. Alter and T. Dresselhaus (2014). "The beginning of a seed: regulatory mechanisms of double fertilization." Front Plant Sci **5**: 452.

Boavida, L. C., J. D. Becker and J. A. Feijó (2005). "The making of gametes in higher plants." Int J Dev Biol **49**(5-6): 595-614.

Boggs, N. A., K. G. Dwyer, P. Shah, A. A. McCulloch, J. Bechsgaard, M. H. Schierup, M. E. Nasrallah and J. B. Nasrallah (2009). "Expression of distinct self-incompatibility specificities in Arabidopsis thaliana." Genetics **182**(4): 1313-1321.

Bolger, A., F. Scossa, M. E. Bolger, C. Lanz, F. Maumus, T. Tohge, H. Quesneville, S. Alseekh, I. Sorensen, G. Lichtenstein, E. A. Fich, M. Conte, H. Keller, K. Schneeberger, R. Schwacke, I. Ofner, J. Vrebalov, Y. Xu, S. Osorio, S. A. Aflitos, E. Schijlen, J. M. Jimenez-Gomez, M. Ryngajllo, S. Kimura, R. Kumar, D. Koenig, L. R. Headland, J. N. Maloof, N. Sinha, R. C. van Ham, R. K. Lankhorst, L. Mao, A. Vogel, B. Arsova, R. Panstruga, Z. Fei, J. K. Rose, D. Zamir, F. Carrari, J. J. Giovannoni, D. Weigel, B. Usadel and A. R. Fernie (2014). "The genome of the stress-tolerant wild tomato species Solanum pennellii." Nat Genet **46**(9): 1034-1038.

Bosch, M., J. S. Knudsen, J. Derksen and C. Mariani (2001). "Class III pistil-specific extensin-like proteins from tobacco have characteristics of arabinogalactan proteins." Plant Physiol **125**(4): 2180-2188.

Bradford, M. M. (1976). "A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding." Anal Biochem **72**: 248-254.

Brewbaker, J. L. and B. H. Kwack (1963). "The Essential Role of Calcium Ion in Pollen Germination and Pollen Tube Growth." Am J Bot **50**(9): 859-865.

Carreno, F. R., C. N. Goni, L. M. Castro and E. S. Ferro (2005). "14-3-3 epsilon modulates the stimulated secretion of endopeptidase 24.15." J Neurochem **93**(1): 10-25.

Chae, K., B. J. Gonong, S. C. Kim, C. A. Kieslich, D. Morikis, S. Balasubramanian and E. M. Lord (2010). "A multifaceted study of stigma/style cysteine-rich adhesin (SCA)-like Arabidopsis lipid transfer proteins (LTPs) suggests diversified roles for these LTPs in plant growth and reproduction." J Exp Bot **61**(15): 4277-4290.

Chae, K. and E. M. Lord (2011). "Pollen tube growth and guidance: roles of small, secreted proteins." Ann Bot **108**(4): 627-636.

Chantha, S. C., B. S. Emerald and D. P. Matton (2006). "Characterization of the plant Notchless homolog, a WD repeat protein involved in seed development." Plant Mol Biol **62**(6): 897-912.

Chen, Y. H., J. P. Du, J. S. Liu, Y. F. Zhu and M. X. Zou (2010). "Functional Analysis of DUF784 Genes in Micropylar Pollen Tube Guidance." Chinese J Biochem and Mol Biol **26**(10): 903-910.

Chen, Y. H., H. J. Li, D. Q. Shi, L. Yuan, J. Liu, R. Sreenivasan, R. Baskar, U. Grossniklaus and W. C. Yang (2007). "The central cell plays a critical role in pollen tube guidance in Arabidopsis." Plant Cell **19**(11): 3563-3577.

Cheng, F. Y., K. Blackburn, Y. M. Lin, M. B. Goshe and J. D. Williamson (2009). "Absolute protein quantification by LC/MS(E) for global analysis of salicylic acid-induced plant protein secretion responses." J Proteome Res **8**(1): 82-93.

Chettoor, A. M. and M. M. Evans (2015). "Correlation between a loss of auxin signaling and a loss of proliferation in maize antipodal cells." Front Plant Sci **6**: 187.

Cheung, A. Y., H. Wang and H. M. Wu (1995). "A floral transmitting tissue-specific glycoprotein attracts pollen tubes and stimulates their growth." Cell **82**(3): 383-393.

Chevalier, D., M. Batoux, L. Fulton, K. Pfister, R. K. Yadav, M. Schellenberg and K. Schneitz (2005). "STRUBBELIG defines a receptor kinase-mediated signaling pathway regulating organ development in Arabidopsis." PNAS **102**(25): 9074-9079.

Chevalier, E., A. Loubert-Hudon and D. P. Matton (2013). "ScRALF3, a secreted RALF-like peptide involved in cell-cell communication between the sporophyte and the female gametophyte in a solanaceous species." Plant J **73**(6): 1019-1033.

Chevalier, E., A. Loubert-Hudon, E. L. Zimmerman and D. P. Matton (2011). "Cell-cell communication and signalling pathways within the ovule: from its inception to fertilization." New Phytol **192**(1): 13-28.

Cho, W. K., X. Y. Chen, H. Chu, Y. Rim, S. Kim, S. T. Kim, S. W. Kim, Z. Y. Park and J. Y. Kim (2009). "Proteomic analysis of the secretome of rice calli." Physiol Plant **135**(4): 331-341.

Chung, H. J., P. C. Sehnke and R. J. Ferl (1999). "The 14-3-3 proteins: cellular regulators of plant metabolism." Trends Plant Sci **4**(9): 367-371.

Coimbra, S., J. Almeida, V. Junqueira, M. L. Costa and L. G. Pereira (2007). "Arabinogalactan proteins as molecular markers in Arabidopsis thaliana sexual reproduction." J Exp Bot **58**(15-16): 4027-4035.

Coimbra, S. and R. Salema (1999). "Ultrastructure of the Developing and Fertilized Embryo Sac of Amaranthus hypochondriacus L." Ann Bot **84**(6): 781-789.

Conesa, A. and S. Gotz (2008). "Blast2GO: A comprehensive suite for functional analysis in plant genomics." Int J Plant Genomics **2008**: 619832.

Conesa, A., S. Gotz, J. M. Garcia-Gomez, J. Terol, M. Talon and M. Robles (2005). "Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research." Bioinformatics **21**(18): 3674-3676.

Cornish, E. C., J. M. Pettitt, I. Bonig and A. E. Clarke (1987). "Developmentally controlled expression of a gene associated with self-incompatibility in Nicotiana alata." Nature **326**(6108): 99-102.

Cruz-Garcia, F., C. Nathan Hancock, D. Kim and B. McClure (2005). "Stylar glycoproteins bind to S-RNase in vitro." Plant J **42**(3): 295-304.

Daigle, C. and D. P. Matton (2015). "Genome-wide analysis of MAPKKKs shows expansion and evolution of a new MEKK class involved in solanaceous species sexual reproduction." BMC Genomics **16**: 1037.

de Graaf, B. H., B. A. Knuiman, J. Derksen and C. Mariani (2003). "Characterization and localization of the transmitting tissue-specific PELPIII proteins of Nicotiana tabacum." J Exp Bot **54**(380): 55-63.

DeLille, J. M., P. C. Sehnke and R. J. Ferl (2001). "The arabidopsis 14-3-3 family of signaling regulators." Plant Physiol **126**(1): 35-38.

Denninger, P., A. Bleckmann, A. Lausser, F. Vogler, T. Ott, D. W. Ehrhardt, W. B. Frommer, S. Sprunck, T. Dresselhaus and G. Grossmann (2014). "Male-female communication triggers calcium signatures during fertilization in Arabidopsis." Nat Commun **5**: 4645.

Domingos, P., A. M. Prado, A. Wong, C. Gehring and J. A. Feijó (2015). "Nitric Oxide: A Multitasked Signaling Gas in Plants." Mol Plant **8**(4): 506-520.

Domínguez, E., J. A. Mercado, M. A. Quesada and A. Heredia (1999). "Pollen sporopollenin: degradation and structural elucidation." Sex Plant Reprod **12**(3): 171-178.

Dong, J., S. T. Kim and E. M. Lord (2005). "Plantacyanin plays a role in reproduction in Arabidopsis." Plant Physiol **138**(2): 778-789.

Dong, Y. H., A. Kvarnheden, J.-L. Yao, P. W. Sutherland, R. G. Atkinson, B. A. Morris and R. C. Gardner (1998). "Identification of pollination-induced genes from the ovary of apple (Malus domestica)." Sex Plant Reprod **11**(5): 277-283.

Dorion, S., Parveen, J. Jeukens, D. P. Matton and J. Rivoal (2005). "Cloning and characterization of a cytosolic isoform of triosephosphate isomerase developmentally regulated in potato leaves." Plant Science **168**(1): 183-194.

Dresselhaus, T. and N. Franklin-Tong (2013). "Male-female crosstalk during pollen germination, tube growth and guidance, and double fertilization." Mol Plant **6**(4): 1018-1036.

Drews, G. N. and A. M. Koltunow (2011). "The female gametophyte." Arabidopsis Book **9**: e0155.

Drews, G. N., D. Wang, J. G. Steffen, K. S. Schumaker and R. Yadegari (2011). "Identification of genes expressed in the angiosperm female gametophyte." J Exp Bot **62**(5): 1593-1599.

Duan, Q., D. Kita, E. A. Johnson, M. Aggarwal, L. Gates, H. M. Wu and A. Y. Cheung (2014). "Reactive oxygen species mediate pollen tube rupture to release sperm for fertilization in Arabidopsis." Nat Commun **5**: 3129.

Eberle, C. A., N. O. Anderson, B. M. Clasen, A. D. Hegeman and A. G. Smith (2013). "PELPIII: the class III pistil-specific extensin-like Nicotiana tabacum proteins are essential for interspecific incompatibility." Plant J **74**(5): 805-814.

Eisenbach, M. and L. C. Giojalas (2006). "Sperm guidance in mammals - an unpaved road to the egg." Nat Rev Mol Cell Biol **7**(4): 276-285.

Emanuelsson, O., H. Nielsen, S. Brunak and G. von Heijne (2000). "Predicting subcellular localization of proteins based on their N-terminal amino acid sequence." J Mol Biol **300**(4): 1005-1016.

Escobar-Restrepo, J. M., N. Huck, S. Kessler, V. Gagliardini, J. Gheyselinck, W. C. Yang and U. Grossniklaus (2007). "The FERONIA receptor-like kinase mediates male-female interactions during pollen tube reception." Science **317**(5838): 656-660.

Feijo, J. A. (2010). "The mathematics of sexual attraction." J Biol **9**(3): 18.

Fernandez-Pozo, N., N. Menda, J. D. Edwards, S. Saha, I. Y. Tecle, S. R. Strickler, A. Bombarely, T. Fisher-York, A. Pujar, H. Foerster, A. Yan and L. A. Mueller (2015). "The Sol Genomics Network (SGN)--from genotype to phenotype to breeding." Nucleic Acids Res **43**(Database issue): D1036-1041.

Floerl, S., A. Majcherczyk, M. Possienke, K. Feussner, H. Tappe, C. Gatz, I. Feussner, U. Kues and A. Polle (2012). "Verticillium longisporum infection affects the leaf apoplastic proteome, metabolome, and cell wall properties in Arabidopsis thaliana." PLoS One **7**(2): e31435.

Foote, H. C., J. P. Ride, V. E. Franklin-Tong, E. A. Walker, M. J. Lawrence and F. C. Franklin (1994). "Cloning and expression of a distinctive class of self-incompatibility (S) gene from Papaver rhoeas L." PNAS **91**(6): 2265-2269.

Frank, J., H. Kaulfurst-Soboll, S. Rips, H. Koiwa and A. von Schaewen (2008). "Comparative analyses of Arabidopsis complex glycan1 mutants and genetic interaction with staurosporin and temperature sensitive3a." Plant Physiol **148**(3): 1354-1367.

Germain, H., E. Chevalier and D. P. Matton (2006). "Plant bioactive peptides: an expanding class of signaling molecules." Botany **84**(1): 1-19.

Germain, H., E. Chevalier, S. Caron and D. P. Matton (2005). "Characterization of five RALF-like genes from Solanum chacoense provides support for a developmental role in plants." Planta **220**(3): 447-454.

Germain, H., M. Gray-Mitsumune, J. Houde, R. Benhamman, T. Sawasaki, Y. Endo and D. P. Matton (2013). "The Solanum chacoense ovary receptor kinase 11 (ScORK11) undergoes tissue-dependent transcriptional, translational and post-translational regulation." Plant Physiol Biochem **70**: 261-268.

Germain, H., M. Gray-Mitsumune, E. Lafleur and D. P. Matton (2008). "ScORK17, a transmembrane receptor-like kinase predominantly expressed in ovules is involved in seed development." Planta **228**(5): 851-862.

Goldberg, E. E., J. R. Kohn, R. Lande, K. A. Robertson, S. A. Smith and B. Igic (2010). "Species selection maintains self-incompatibility." Science **330**(6003): 493-495.

Goldman, M. H., M. Pezzotti, J. Seurinck and C. Mariani (1992). "Developmental expression of tobacco pistil-specific genes encoding novel extensin-like proteins." Plant Cell **4**(9): 1041-1051.

Goring, D. R. and S. J. Rothstein (1992). "The S-locus receptor kinase gene in a self-incompatible Brassica napus line encodes a functional serine/threonine kinase." Plant Cell **4**(10): 1273-1281.

Goulet, C., C. Goulet, M. C. Goulet and D. Michaud (2010). "2-DE proteome maps for the leaf apoplast of Nicotiana benthamiana." Proteomics **10**(13): 2536-2544.

Grabherr, M. G., B. J. Haas, M. Yassour, J. Z. Levin, D. A. Thompson, I. Amit, X. Adiconis, L. Fan, R. Raychowdhury, Q. Zeng, Z. Chen, E. Mauceli, N. Hacohen, A. Gnirke, N. Rhind, F. di Palma, B. W. Birren, C. Nusbaum, K. Lindblad-Toh, N. Friedman and A. Regev (2011). "Full-length transcriptome assembly from RNA-Seq data without a reference genome." Nat Biotechnol **29**(7): 644-652.

Gray-Mitsumune, M. and D. P. Matton (2006). "The Egg apparatus 1 gene from maize is a member of a large gene family found in both monocots and dicots." Planta **223**(3): 618-625.

Gray-Mitsumune, M., M. O'Brien, C. Bertrand, F. Tebbji, A. Nantel and D. P. Matton (2006). "Loss of ovule identity induced by overexpression of the fertilization-related kinase 2 (ScFRK2), a MAPKKK from Solanum chacoense." J Exp Bot **57**(15): 4171-4187.

Guan, Y., J. Lu, J. Xu, B. McClure and S. Zhang (2014). "Two Mitogen-Activated Protein Kinases, MPK3 and MPK6, Are Required for Funicular Guidance of Pollen Tubes in Arabidopsis." Plant Physiol **165**(2): 528-533.

Guo, F. Q., M. Okamoto and N. M. Crawford (2003). "Identification of a plant nitric oxide synthase gene involved in hormonal signaling." Science **302**(5642): 100-103.

Gupta, R. and R. Deswal (2012). "Low temperature stress modulated secretome analysis and purification of antifreeze protein from Hippophae rhamnoides, a Himalayan wonder plant." J Proteome Res **11**(5): 2684-2696.

Gupta, S., V. Wardhan, S. Verma, S. Gayali, U. Rajamani, A. Datta, S. Chakraborty and N. Chakraborty (2011). "Characterization of the secretome of chickpea suspension culture reveals pathway abundance and the expected and unexpected secreted proteins." J Proteome Res **10**(11): 5006-5015.

Hafidh, S., J. Fíla and D. Honys (2016). "Male gametophyte development and function in angiosperms: a general concept." Plant Reprod doi: 10.1007/s00497-015-0272-4.

Hafidh, S., D. Potesil, J. Fila, J. Fecikova, V. Capkova, Z. Zdrahal and D. Honys (2014). "In search of ligands and receptors of the pollen tube: the missing link in pollen tube perception." Biochem Soc Trans **42**(2): 388-394.

Hamamura, Y., S. Nagahara and T. Higashiyama (2012). "Double fertilization on the move." Curr Opin Plant Biol **15**(1): 70-77.

Hamamura, Y., C. Saito, C. Awai, D. Kurihara, A. Miyawaki, T. Nakagawa, M. M. Kanaoka, N. Sasaki, A. Nakano, F. Berger and T. Higashiyama (2011). "Live-cell imaging reveals the dynamics of two sperm cells during double fertilization in Arabidopsis thaliana." Curr Biol **21**(6): 497-502.

Haweker, H., S. Rips, H. Koiwa, S. Salomon, Y. Saijo, D. Chinchilla, S. Robatzek and A. von Schaewen (2010). "Pattern recognition receptors require N-glycosylation to mediate plant immunity." J Biol Chem **285**(7): 4629-4636.

Hawkes, J. G. (1962). "Introgression in certain wild potato species." Euphytica **11**(1): 26-35.

Higashiyama, T. (2002). "The synergid cell: attractor and acceptor of the pollen tube for double fertilization." J Plant Res **115**(1118): 149-160.

Higashiyama, T. (2010). "Peptide signaling in pollen–pistil interactions." Plant Cell Physiol **51**(2): 177-189.

Higashiyama, T., R. Inatsugi, S. Sakamoto, N. Sasaki, T. Mori, H. Kuroiwa, T. Nakada, H. Nozaki, T. Kuroiwa and A. Nakano (2006). "Species preferentiality of the pollen tube attractant derived from the synergid cell of Torenia fournieri." Plant Physiol **142**(2): 481-491.

Higashiyama, T., H. Kuroiwa, S. Kawano and T. Kuroiwa (1998). "Guidance in vitro of the pollen tube to the naked embryo sac of torenia fournieri." Plant Cell **10**(12): 2019-2032.

Higashiyama, T., H. Kuroiwa and T. Kuroiwa (2003). "Pollen-tube guidance: beacons from the female gametophyte." Curr Opin Plant Biol **6**(1): 36-41.

Higashiyama, T. and H. Takeuchi (2015). "The Mechanism and Key Molecules Involved in Pollen Tube Guidance." Annu Rev Plant Biol **66**: 393-413.

Higashiyama, T., S. Yabe, N. Sasaki, Y. Nishimura, S. Miyagishima, H. Kuroiwa and T. Kuroiwa (2001). "Pollen tube attraction by the synergid cell." Science **293**(5534): 1480-1483.

Hijmans, R. J., D. M. Spooner, A. R. Salas, L. Guarino and J. de la Cruz (2002). Atlas of Wild Potatoes. Rome, International Plant Genetic Resources Institute.

Hirsch, C. D., J. P. Hamilton, K. L. Childs, J. Cepela, E. Crisovan, B. Vaillancourt, C. N. Hirsch, M. Habermann, B. Neal and C. R. Buell (2014). "Spud DB: A Resource for Mining Sequences, Genotypes, and Phenotypes to Accelerate Potato Breeding." Plant Gen **7**: 1-12.

Holt, W. V. and A. Fazeli (2010). "The oviduct as a complex mediator of mammalian sperm function and selection." Mol Reprod Dev **77**(11): 934-943.

Horade, M., M. M. Kanaoka, M. Kuzuya, T. Higashiyama and N. Kaji (2013). "A microfluidic device for quantitative analysis of chemoattraction in plants." RSC Advances **3**(44): 22301-22307.

Hornett, E. A. and C. W. Wheat (2012). "Quantitative RNA-Seq analysis in non-model species: assessing transcriptome assemblies as a scaffold and the utility of evolutionary divergent genomic reference species." BMC Genomics **13**: 361.

Huang, B.-Q. and S. Russell (1992). Female Germ Unit: Organization, Isolation, and Function, International Rev Cytology **140**: 233-293.

Hulskamp, M., K. Schneitz and R. E. Pruitt (1995). "Genetic Evidence for a Long-Range Activity That Directs Pollen Tube Guidance in Arabidopsis." Plant Cell **7**(1): 57-64.

Inoue, N., M. Ikawa, A. Isotani and M. Okabe (2005). "The immunoglobulin superfamily protein Izumo is required for sperm to fuse with eggs." Nature **434**(7030): 234-238.

Inoue, N., M. Ikawa and M. Okabe (2008). "Putative sperm fusion protein IZUMO and the role of N-glycosylation." Biochem Biophys Res Commun **377**(3): 910-914.

Johnston, A. J., P. Meier, J. Gheyselinck, S. E. Wuest, M. Federer, E. Schlagenhauf, J. D. Becker and U. Grossniklaus (2007). "Genetic subtraction profiling identifies genes essential for Arabidopsis reproduction and reveals interaction between the female gametophyte and the maternal sporophyte." Genome Biol **8**(10): R204.

Joly, V. and D. P. Matton (2015). "KAPPA, a simple algorithm for discovery and clustering of proteins defined by a key amino acid pattern: a case study of the cysteine-rich proteins." Bioinformatics **31**(11): 1716-1723.

Jones-Rhoades, M. W., J. O. Borevitz and D. Preuss (2007). "Genome-wide expression profiling of the Arabidopsis female gametophyte identifies families of small, secreted proteins." PLoS Genet **3**(10): 1848-1861.

Jung, Y. H., S. H. Jeong, S. H. Kim, R. Singh, J. E. Lee, Y. S. Cho, G. K. Agrawal, R. Rakwal and N. S. Jwa (2008). "Systematic secretome analyses of rice leaf and seed callus suspension-cultured cells: workflow development and establishment of high-density two-dimensional gel reference maps." J Proteome Res **7**(12): 5187-5210.

Kachroo, A., M. E. Nasrallah and J. B. Nasrallah (2002). "Self-incompatibility in the Brassicaceae: receptor-ligand signaling and cell-to-cell communication." Plant Cell **14** Suppl: S227-238.

Kagi, C., N. Baumann, N. Nielsen, Y. D. Stierhof and R. Gross-Hardt (2010). "The gametic central cell of Arabidopsis determines the lifespan of adjacent accessory cells." PNAS **107**(51): 22350-22355.

Kanaoka, M. M., N. Kawano, Y. Matsubara, D. Susaki, S. Okuda, N. Sasaki and T. Higashiyama (2011). "Identification and characterization of TcCRP1, a pollen tube attractant from Torenia concolor." Ann Bot **108**(4): 739-747.

Kane, N. A., Z. Agharbaoui, A. O. Diallo, H. Adam, Y. Tominaga, F. Ouellet and F. Sarhan (2007). "TaVRT2 represses transcription of the wheat vernalization gene TaVRN1." Plant J **51**(4): 670-680.

Kasahara, R. D., M. F. Portereiko, L. Sandaklie-Nikolova, D. S. Rabiger and G. N. Drews (2005). "MYB98 is required for pollen tube guidance and synergid cell differentiation in Arabidopsis." Plant Cell **17**(11): 2981-2992.

Katoh, K. and D. M. Standley (2013). "MAFFT multiple sequence alignment software version 7: improvements in performance and usability." Mol Biol Evol **30**(4): 772-780.

Kaufmann, H., F. Salamini and R. Thompson (1991). "Sequence variability and gene structure at the self-incompatibility locus of Solanum tuberosum." Mol and General Genet MGG **226**(3): 457-466.

Kim, S., J. C. Mollet, J. Dong, K. Zhang, S. Y. Park and E. M. Lord (2003). "Chemocyanin, a small basic protein from the lily stigma, induces pollen tube chemotropism." PNAS **100**(26): 16125-16130.

Kim, S., M. Park, S. I. Yeom, Y. M. Kim, J. M. Lee, H. A. Lee, E. Seo, J. Choi, K. Cheong, K. T. Kim, K. Jung, G. W. Lee, S. K. Oh, C. Bae, S. B. Kim, H. Y. Lee, S. Y. Kim, M. S. Kim, B. C. Kang, Y. D. Jo, H. B. Yang, H. J. Jeong, W. H. Kang, J. K. Kwon, C. Shin, J. Y. Lim, J. H. Park, J. H. Huh, J. S. Kim, B. D. Kim, O. Cohen, I. Paran, M. C. Suh, S. B. Lee, Y. K. Kim, Y. Shin, S. J. Noh, J. Park, Y. S. Seo, S. Y. Kwon, H. A. Kim, J. M. Park, H. J. Kim, S. B. Choi, P. W. Bosland, G. Reeves, S. H. Jo, B. W. Lee, H. T. Cho, H. S. Choi, M. S. Lee, Y. Yu, Y. Do Choi, B. S. Park, A. van Deynze, H. Ashrafi, T. Hill, W. T. Kim, H. S. Pai, H. K. Ahn, I. Yeam, J. J. Giovannoni, J. K. Rose, I. Sorensen, S. J. Lee, R. W. Kim, I. Y. Choi, B. S. Choi, J. S. Lim, Y. H. Lee and D. Choi (2014). "Genome sequence of the hot pepper provides insights into the evolution of pungency in Capsicum species." Nat Genet **46**(3): 270-278.

Kim, S. G., S. T. Kim, Y. Wang, S. K. Kim, C. H. Lee, K. K. Kim, J. K. Kim, S. Y. Lee and K. Y. Kang (2010). "Overexpression of rice isoflavone reductase-like gene (OsIRL) confers tolerance to reactive oxygen species." Physiol Plant **138**(1): 1-9.

Knapp, S., L. Bohs, M. Nee and D. M. Spooner (2004). "Solanaceae–A model for linking genomics with biodiversity." Comp Funct Genomics **5**(3): 285-291.

Koltunow, A. M. (1993). "Apomixis: Embryo Sacs and Embryos Formed without Meiosis or Fertilization in Ovules." Plant Cell **5**(10): 1425-1437.

Komori, R., Y. Amano, M. Ogawa-Ohnishi and Y. Matsubayashi (2009). "Identification of tyrosylprotein sulfotransferase in Arabidopsis." PNAS **106**(35): 15067-15072.

Konozy, E. H., H. Rogniaux, M. Causse and M. Faurobert (2013). "Proteomic analysis of tomato (Solanum lycopersicum) secretome." J Plant Res **126**(2): 251-266.

Kubo, T., M. Fujita, H. Takahashi, M. Nakazono, N. Tsutsumi and N. Kurata (2013). "Transcriptome analysis of developing ovules in rice isolated by laser microdissection." Plant Cell Physiol **54**(5): 750-765.

Lafleur, E., C. Kapfer, V. Joly, Y. Liu, F. Tebbji, C. Daigle, M. Gray-Mitsumune, M. Cappadocia, A. Nantel and D. P. Matton (2015). "The FRK1 mitogen-activated protein kinase kinase kinase (MAPKKK) from Solanum chacoense is involved in embryo sac and pollen development." J Exp Bot **66**(7): 1833-1843.

Lagace, M., S. C. Chantha, G. Major and D. P. Matton (2003). "Fertilization induces strong accumulation of a histone deacetylase (HD2) and of other chromatin-remodeling proteins in restricted areas of the ovules." Plant Mol Biol **53**(6): 759-769.

Lander, E. S. and M. S. Waterman (1988). "Genomic mapping by fingerprinting random clones: a mathematical analysis." Genomics **2**(3): 231-239.

Langmead, B., C. Trapnell, M. Pop and S. L. Salzberg (2009). "Ultrafast and memory-efficient alignment of short DNA sequences to the human genome." Genome Biol **10**(3): R25.

Lantin, S., M. O'Brien and D. P. Matton (1999). "Pollination, wounding and jasmonate treatments induce the expression of a developmentally regulated pistil dioxygenase at a distance, in the ovary, in the wild potato Solanum chacoense Bitt." Plant Mol Biol **41**(3): 371-386.

Lausser, A., I. Kliwer, K. O. Srilunchang and T. Dresselhaus (2010). "Sporophytic control of pollen tube growth and guidance in maize." J Exp Bot **61**(3): 673-682.

Leshem, Y., C. Johnson and V. Sundaresan (2013). "Pollen tube entry into the synergid cell of Arabidopsis is observed at a site distinct from the filiform apparatus." Plant Reprod **26**(2): 93-99.

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

Lindner, H., S. A. Kessler, L. M. Muller, H. Shimosato-Asano, A. Boisson-Dernier and U. Grossniklaus (2015). "TURAN and EVAN Mediate Pollen Tube Reception in Arabidopsis Synergids through Protein Glycosylation." PLoS Biol **13**(4): e1002139.

Ling, Y., C. Zhang, T. Chen, H. Hao, P. Liu, R. A. Bressan, P. M. Hasegawa, J. B. Jin and J. Lin (2012). "Mutation in SUMO E3 ligase, SIZ1, disrupts the mature female gametophyte in Arabidopsis." PLoS One **7**(1): e29470.

Lituiev, D. S., N. G. Krohn, B. Muller, D. Jackson, B. Hellriegel, T. Dresselhaus and U. Grossniklaus (2013). "Theoretical and experimental evidence indicates that there is no detectable auxin gradient in the angiosperm female gametophyte." Development **140**(22): 4544-4553.

Liu, Y., V. Joly, S. Dorion, J. Rivoal and D. P. Matton (2015). "The plant ovule secretome: a different view toward pollen–pistil interactions." J Proteome Res **14**(11): 4763-4775.

Liu, Y., Z. Yan, N. Chen, X. Di, J. Huang and G. Guo (2010). "Development and function of central cell in angiosperm female gametophyte." Genesis **48**(8): 466-478.

Long, J. A. (2006). "Plant development: new models and approaches bring progress." Development **133**(23): 4609-4612.

Lu, Y., S. Chanroj, L. Zulkifli, M. A. Johnson, N. Uozumi, A. Cheung and H. Sze (2011). "Pollen tubes lacking a pair of K+ transporters fail to target ovules in Arabidopsis." Plant Cell **23**(1): 81-93.

Lum, G. K., J. Meinken, J. Orr, S. Frazier and X. J. Min (2014). "PlantSecKB: the Plant Secretome and Subcellular Proteome KnowledgeBase." Computational Mol Biol **4**(1): 1-17.

Lush, W. M., T. Spurck and R. Joosten (2000). "Pollen Tube Guidance by the Pistil of a Solanaceous Plant." Ann Bot **85**(suppl 1): 39-47.

Luu, D. T., X. Qin, D. Morse and M. Cappadocia (2000). "S-RNase uptake by compatible pollen tubes in gametophytic self-incompatibility." Nature **407**(6804): 649-651.

Makkena S., E. Lee, F. D. Sack, R. S. Lamb (2012). "The R2R3 MYB transcription factors FOUR LIPS and MYB88 regulate female reproductive development." J Exp Bot **63**: 5545-5558.

Margulies, M., M. Egholm, W. E. Altman, S. Attiya, J. S. Bader, L. A. Bemben, J. Berka, M. S. Braverman, Y. J. Chen, Z. Chen, S. B. Dewell, L. Du, J. M. Fierro, X. V. Gomes, B. C. Godwin, W. He, S. Helgesen, C. H. Ho, G. P. Irzyk, S. C. Jando, M. L. Alenquer, T. P. Jarvie, K. B. Jirage, J. B. Kim, J. R. Knight, J. R. Lanza, J. H. Leamon, S. M. Lefkowitz, M. Lei, J. Li, K. L. Lohman, H. Lu, V. B. Makhijani, K. E. McDade, M. P. McKenna, E. W. Myers, E. Nickerson, J. R. Nobile, R. Plant, B. P. Puc, M. T. Ronan, G. T. Roth, G. J. Sarkis, J. F. Simons, J. W. Simpson, M. Srinivasan, K. R. Tartaro, A. Tomasz, K. A. Vogt, G. A. Volkmer, S. H. Wang, Y. Wang, M. P. Weiner, P. Yu, R. F. Begley and J. M. Rothberg (2005). "Genome sequencing in microfabricated high-density picolitre reactors." Nature **437**(7057): 376-380.

Marshall, E., L. M. Costa and J. Gutierrez-Marcos (2011). "Cysteine-rich peptides (CRPs) mediate diverse aspects of cell-cell communication in plant reproduction and development." J Exp Bot **62**(5): 1677-1686.

Martin, M. V., D. F. Fiol, V. Sundaresan, E. J. Zabaleta and G. C. Pagnussat (2013). "oiwa, a female gametophytic mutant impaired in a mitochondrial manganese-superoxide dismutase, reveals crucial roles for reactive oxygen species during embryo sac development and fertilization in Arabidopsis." Plant Cell **25**(5): 1573-1591.

Marton, M. L., S. Cordts, J. Broadhvest and T. Dresselhaus (2005). "Micropylar pollen tube guidance by egg apparatus 1 of maize." Science **307**(5709): 573-576.

Marton, M. L., A. Fastner, S. Uebler and T. Dresselhaus (2012). "Overcoming hybridization barriers by the secretion of the maize pollen tube attractant ZmEA1 from Arabidopsis ovules." Curr Biol **22**(13): 1194-1198.

Mascarenhas, J. P., L. Machlis (1962). "Chemotropic response of Antirrhinum majus pollen to calcium. " Nature **196**: 292-293.

Matsubayashi, Y. and Y. Sakagami (1996). "Phytosulfokine, sulfated peptides that induce the proliferation of single mesophyll cells of Asparagus officinalis L." PNAS **93**(15): 7623-7627.

Matton, D. P., C. Bertrand, G. Laublin and M. Cappadocia (1998). Molecular aspects of self-incompatibility in tuber-bearing Solanum species. Comprehensive Potato Biotechnology S. M. Paul Khurana, C. R. and M. D. Upadhya. New Delhi Malhotra Publishing House: 97-113.

Matton, D. P., O. Maes, G. Laublin, Q. Xike, C. Bertrand, D. Morse and M. Cappadocia (1997). "Hypervariable Domains of Self-Incompatibility RNases Mediate Allele-Specific Pollen Recognition." Plant Cell **9**(10): 1757-1766.

McClure, B., F. Cruz-Garcia and C. Romero (2011). "Compatibility and incompatibility in S-RNase-based systems." Ann Bot **108**(4): 647-658.

McClure, B. A. and V. Franklin-Tong (2006). "Gametophytic self-incompatibility: understanding the cellular mechanisms involved in "self" pollen tube inhibition." Planta **224**(2): 233-245.

Miller, J. T. and D. M. Spooner (1996). "Introgression of Solanum chacoense (Solanum sect. Petota): Upland Populations Reexamined." Systematic Bot **21**(4): 461-475.

Miyosha, M. (1894). "Über Reizbewegungen der Pollenschläuche. Flora **78**: 76-93.

Molisch, H. (1893). "Zur Physiologie des Pollens mit besonderer Rücksicht auf die chemotropische Bewegungen der Pollenschläuche. " Mathnaturw **102**: 423-448.

Mollet, J. C., S. Y. Park, E. A. Nothnagel and E. M. Lord (2000). "A lily stylar pectin is necessary for pollen tube adhesion to an in vitro stylar matrix." Plant Cell **12**(9): 1737-1750.

Murashige, T. and F. Skoog (1962). "A Revised Medium for Rapid Growth and Bio Assays with Tobacco Tissue Cultures." Physiologia Plantarum **15**(3): 473-497.

Nakagawa, T., T. Kurose, T. Hino, K. Tanaka, M. Kawamukai, Y. Niwa, K. Toyooka, K. Matsuoka, T. Jinbo and T. Kimura (2007). "Development of series of gateway binary vectors, pGWBs, for realizing efficient construction of fusion genes for plant transformation." J Biosci Bioeng **104**(1): 34-41.

Nakai, K. and P. Horton (1999). "PSORT: a program for detecting sorting signals in proteins and predicting their subcellular localization." Trends Biochem Sci **24**(1): 34-36.

O'Brien, M., S. C. Chantha, A. Rahier and D. P. Matton (2005). "Lipid signaling in plants. Cloning and expression analysis of the obtusifoliol 14alpha-demethylase from Solanum chacoense Bitt., a pollination- and fertilization-induced gene with both obtusifoliol and lanosterol demethylase activity." Plant Physiol **139**(2): 734-749.

O'Brien, M., M. Gray-Mitsumune, C. Kapfer, C. Bertrand and D. P. Matton (2007). "The ScFRK2 MAP kinase kinase kinase from Solanum chacoense affects pollen development and viability." Planta **225**(5): 1221-1231.

O'Brien, M., C. Kapfer, G. Major, M. Laurin, C. Bertrand, K. Kondo, Y. Kowyama and D. P. Matton (2002). "Molecular analysis of the stylar-expressed Solanum chacoense small asparagine-rich protein family related to the HT modifier of gametophytic self-incompatibility in Nicotiana." Plant J **32**(6): 985-996.

Ohnishi, T., H. Takanashi, M. Mogi, H. Takahashi, S. Kikuchi, K. Yano, T. Okamoto, M. Fujita, N. Kurata and N. Tsutsumi (2011). "Distinct gene expression profiles in egg and synergid cells of rice as revealed by cell type-specific microarrays." Plant Physiol **155**(2): 881-891.

Okuda, S., T. Suzuki, M. M. Kanaoka, H. Mori, N. Sasaki and T. Higashiyama (2013). "Acquisition of LURE-binding activity at the pollen tube tip of Torenia fournieri." Mol Plant **6**(4): 1074-1090.

Okuda, S., H. Tsutsui, K. Shiina, S. Sprunck, H. Takeuchi, R. Yui, R. D. Kasahara, Y. Hamamura, A. Mizukami, D. Susaki, N. Kawano, T. Sakakibara, S. Namiki, K. Itoh, K. Otsuka, M. Matsuzaki, H. Nozaki, T. Kuroiwa, A. Nakano, M. M. Kanaoka, T. Dresselhaus, N. Sasaki and T. Higashiyama (2009). "Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells." Nature **458**(7236): 357-361.

Olmedo-Monfil, V., N. Duran-Figueroa, M. Arteaga-Vazquez, E. Demesa-Arevalo, D. Autran, D. Grimanelli, R. K. Slotkin, R. A. Martienssen and J. P. Vielle-Calzada (2010). "Control of female gamete formation by a small RNA pathway in Arabidopsis." Nature **464**(7288): 628-632.

Ottenschlager, I., P. Wolff, C. Wolverton, R. P. Bhalerao, G. Sandberg, H. Ishikawa, M. Evans and K. Palme (2003). "Gravity-regulated differential auxin transport from columella to lateral root cap cells." PNAS **100**(5): 2987-2991.

Pagnussat, G. C., M. Alandete-Saez, J. L. Bowman and V. Sundaresan (2009). "Auxin-dependent patterning and gamete specification in the Arabidopsis female gametophyte." Science **324**(5935): 1684-1689.

Palanivelu, R., L. Brass, A. F. Edlund and D. Preuss (2003). "Pollen tube growth and guidance is regulated by POP2, an Arabidopsis gene that controls GABA levels." Cell **114**(1): 47-59.

Palanivelu, R. and D. Preuss (2006). "Distinct short-range ovule signals attract or repel Arabidopsis thaliana pollen tubes in vitro." BMC Plant Biol **6**: 7.

Palanivelu, R. and T. Tsukamoto (2012). "Pathfinding in angiosperm reproduction: pollen tube guidance by pistils ensures successful double fertilization." Wiley Interdiscip Rev Dev Biol **1**(1): 96-113.

Pandey, A., U. Rajamani, J. Verma, P. Subba, N. Chakraborty, A. Datta, S. Chakraborty and N. Chakraborty (2010). "Identification of extracellular matrix proteins of rice (Oryza sativa L.) involved in dehydration-responsive network: a proteomic approach." J Proteome Res **9**(7): 3443-3464.

Park, S. Y., G. Y. Jauh, J. C. Mollet, K. J. Eckard, E. A. Nothnagel, L. L. Walling and E. M. Lord (2000). "A lipid transfer-like protein is necessary for lily pollen tube adhesion to an in vitro stylar matrix." Plant Cell **12**(1): 151-164.

Pereira, A. M., S. Masiero, M. S. Nobre, M. L. Costa, M. T. Solis, P. S. Testillano, S. Sprunck and S. Coimbra (2014). "Differential expression patterns of arabinogalactan proteins in Arabidopsis thaliana reproductive tissues." J Exp Bot **65**(18): 5459-5471.

Petersen, T. N., S. Brunak, G. von Heijne and H. Nielsen (2011). "SignalP 4.0: discriminating signal peptides from transmembrane regions." Nat Methods **8**(10): 785-786.

Petrucco, S., A. Bolchi, C. Foroni, R. Percudani, G. L. Rossi and S. Ottonello (1996). "A maize gene encoding an NADPH binding enzyme highly homologous to isoflavone reductases is activated in response to sulfur starvation." Plant Cell **8**(1): 69-80.

Plattner, S., C. Gruber, J. Stadlmann, S. Widmann, C. W. Gruber, F. Altmann and H. Bohlmann (2015). "Isolation and Characterization of a Thionin Proprotein Processing Enzyme from Barley." J Biol Chem **290**(29):18056-18067.

Pontes, A., Y. Zhang and W. Hu (2013). "Novel functions of GABA signaling in adult neurogenesis." Front Biol doi:10.1007/s11515-013-1270-2.

Portereiko, M. F., A. Lloyd, J. G. Steffen, J. A. Punwani, D. Otsuga and G. N. Drews (2006). "AGL80 is required for central cell and endosperm development in Arabidopsis." Plant Cell **18**(8): 1862-1872.

Potato Genome Consortium (2011). "Genome sequence and analysis of the tuber crop potato." Nature **475**(7355): 189-195.

Prado, A. M., R. Colaco, N. Moreno, A. C. Silva and J. A. Feijó (2008). "Targeting of pollen tubes to ovules is dependent on nitric oxide (NO) signaling." Mol Plant **1**(4): 703-714.

Prado, A. M., D. M. Porterfield and J. A. Feijó (2004). "Nitric oxide is involved in growth regulation and re-orientation of pollen tubes." Development **131**(11): 2707-2714.

Prior, N., S. A. Little, C. Pirone, J. E. Gill, D. Smith, J. Han, D. Hardie, S. J. O'Leary, R. E. Wagner, T. Cross, A. Coulter, C. Borchers, R. W. Olafson and P. von Aderkas (2013). "Application of proteomics to the study of pollination drops." Appl Plant Sci **1**(4).

Punwani, J. A., D. S. Rabiger and G. N. Drews (2007). "MYB98 positively regulates a battery of synergid-expressed genes encoding filiform apparatus localized proteins." Plant Cell **19**(8): 2557-2568.

Punwani, J. A., D. S. Rabiger, A. Lloyd and G. N. Drews (2008). "The MYB98 subcircuit of the synergid gene regulatory network includes genes directly and indirectly regulated by MYB98." Plant J **55**(3): 406-414.

Qin, Y., A. R. Leydon, A. Manziello, R. Pandey, D. Mount, S. Denic, B. Vasic, M. A. Johnson and R. Palanivelu (2009). "Penetration of the stigma and style elicits a novel transcriptome in pollen tubes, pointing to genes critical for growth in a pistil." PLoS Genet **5**(8): e1000621.

Ray, A., J. D. Lang, T. Golden and S. Ray (1996). "SHORT INTEGUMENT (SIN1), a gene required for ovule development in Arabidopsis, also controls flowering time." Development **122**(9): 2631-2638.

Ray, A., K. Robinson-Beers, S. Ray, S. C. Baker, J. D. Lang, D. Preuss, S. B. Milligan and C. S. Gasser (1994). "Arabidopsis floral homeotic gene BELL (BEL1) controls ovule development through negative regulation of AGAMOUS gene (AG)." PNAS **91**(13): 5761-5765.

Regente, M., M. Pinedo, M. Elizalde and L. de la Canal (2012). "Apoplastic exosome-like vesicles: a new way of protein secretion in plants?" Plant Signal Behav **7**(5): 544-546.

Rejon, J. D., F. Delalande, C. Schaeffer-Reiss, C. Carapito, K. Zienkiewicz, J. de Dios Alche, M. Isabel Rodriguez-Garcia, A. Van Dorsselaer and A. J. Castro (2014). "The plant stigma exudate: a biochemically active extracellular environment for pollen germination?" Plant Signal Behav **9**(3): e28274.

Rejon, J. D., F. Delalande, C. Schaeffer-Reiss, C. Carapito, K. Zienkiewicz, J. de Dios Alche, M. I. Rodriguez-Garcia, A. Van Dorsselaer and A. J. Castro (2013). "Proteomics profiling reveals novel proteins and functions of the plant stigma exudate." J Exp Bot **64**(18): 5695-5705.

Renault, H., A. El Amrani, R. Palanivelu, E. P. Updegraff, A. Yu, J. P. Renou, D. Preuss, A. Bouchereau and C. Deleu (2011). "GABA accumulation causes cell elongation defects and a decrease in expression of genes encoding secreted and cell wall-related proteins in Arabidopsis thaliana." Plant Cell Physiol **52**(5): 894-908.

Rholam, M. and C. Fahy (2009). "Processing of peptide and hormone precursors at the dibasic cleavage sites." Cell Mol Life Sci **66**(13): 2075-2091.

Rivoal, J. and A. D. Hanson (1994). "Metabolic Control of Anaerobic Glycolysis (Overexpression of Lactate Dehydrogenase in Transgenic Tomato Roots Supports the Davies-Roberts Hypothesis and Points to a Critical Role for Lactate Secretion." Plant Physiol **106**(3): 1179-1185.

Robinson, M. D., D. J. McCarthy and G. K. Smyth (2010). "edgeR: a Bioconductor package for differential expression analysis of digital gene expression data." Bioinformatics **26**(1): 139-140.

Rodríguez, F. and D. M. Spooner (2009). "Nitrate Reductase Phylogeny of Potato (Solatium sect. Petota) Genomes with Emphasis on the Origins of the Polyploid Species." Systematic Bot **34**(1): 207-219.

Sanchez-Leon, N., M. Arteaga-Vazquez, C. Alvarez-Mejia, J. Mendiola-Soto, N. Duran-Figueroa, D. Rodriguez-Leal, I. Rodriguez-Arevalo, V. Garcia-Campayo, M. Garcia-Aguilar, V. Olmedo-Monfil, M. Arteaga-Sanchez, O. M. de la Vega, K. Nobuta, K. Vemaraju, B. C. Meyers and J. P. Vielle-Calzada (2012). "Transcriptional analysis of the Arabidopsis ovule by massively parallel signature sequencing." J Exp Bot **63**(10): 3829-3842.

Sanders, P. M., A. Q. Bui, K. Weterings, K. N. McIntire, Y.-C. Hsu, P. Y. Lee, M. T. Truong, T. P. Beals and R. B. Goldberg (1999). "Anther developmental defects in Arabidopsis thaliana male-sterile mutants." Sex Plant Reprod **11**(6): 297-322.

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

Sato, Y., N. Sugimoto, T. Higashiyama and H. Arata (2015). "Quantification of pollen tube attraction in response to guidance by female gametophyte tissue using artificial microscale pathway." J Biosci Bioeng **120**(6): 697-700.

Sauter, M. (2015). "Phytosulfokine peptide signalling." J Exp Bot **66**(17): 5161-5169.

Shpak, E. D., C. T. Berthiaume, E. J. Hill and K. U. Torii (2004). "Synergistic interaction of three ERECTA-family receptor-like kinases controls Arabidopsis organ growth and flower development by promoting cell proliferation." Development **131**(7): 1491-1501.

Silverstein, K. A., W. A. Moskal, Jr., H. C. Wu, B. A. Underwood, M. A. Graham, C. D. Town and K. A. VandenBosch (2007). "Small cysteine-rich peptides resembling antimicrobial peptides have been under-predicted in plants." Plant J **51**(2): 262-280.

Smyth, D. R., J. L. Bowman and E. M. Meyerowitz (1990). "Early flower development in Arabidopsis." Plant Cell **2**(8): 755-767.

Sommer-Knudsen, J., W. M. Lush, A. Bacic and A. E. Clarke (1998). "Re-evaluation of the role of a transmitting tract-specific glycoprotein on pollen tube growth." Plant J **13**(4): 529-535.

Song, X., L. Yuan and V. Sundaresan (2014). "Antipodal cells persist through fertilization in the female gametophyte of Arabidopsis." Plant Reprod **27**(4): 197-203.

Song, Y., C. Zhang, W. Ge, Y. Zhang, A. L. Burlingame and Y. Guo (2011). "Identification of NaCl stress-responsive apoplastic proteins in rice shoot stems by 2D-DIGE." J Proteomics **74**(7): 1045-1067.

Soulard, J., N. Boivin, D. Morse and M. Cappadocia (2014). "eEF1A is an S-RNase binding factor in self-incompatible Solanum chacoense." PLoS One **9**(2): e90206.

Sprunck, S., T. Hackenberg, M. Englhart and F. Vogler (2014). "Same same but different: sperm-activating EC1 and ECA1 gametogenesis-related family proteins." Biochem Soc Trans **42**(2): 401-407.

Sprunck, S., S. Rademacher, F. Vogler, J. Gheyselinck, U. Grossniklaus and T. Dresselhaus (2012). "Egg cell-secreted EC1 triggers sperm cell activation during double fertilization." Science **338**(6110): 1093-1097.

Steentoft, C., S. Y. Vakhrushev, H. J. Joshi, Y. Kong, M. B. Vester-Christensen, K. T. Schjoldager, K. Lavrsen, S. Dabelsteen, N. B. Pedersen, L. Marcos-Silva, R. Gupta, E. P. Bennett, U. Mandel, S. Brunak, H. H. Wandall, S. B. Levery and H. Clausen (2013). "Precision mapping of the human O-GalNAc glycoproteome through SimpleCell technology." EMBO J **32**(10): 1478-1488.

Steffen, J. G., I. H. Kang, J. Macfarlane and G. N. Drews (2007). "Identification of genes expressed in the Arabidopsis female gametophyte." Plant J **51**(2): 281-292.

Steffen, J. G., I. H. Kang, M. F. Portereiko, A. Lloyd and G. N. Drews (2008). "AGL61 interacts with AGL80 and is required for central cell development in Arabidopsis." Plant Physiol **148**(1): 259-268.

Stein, J. C., B. Howlett, D. C. Boyes, M. E. Nasrallah and J. B. Nasrallah (1991). "Molecular cloning of a putative receptor protein kinase gene encoded at the self-incompatibility locus of Brassica oleracea." PNAS **88**(19): 8816-8820.

Stewman, S. F., M. Jones-Rhoades, P. Bhimalapuram, M. Tchernookov, D. Preuss and A. R. Dinner (2010). "Mechanistic insights from a quantitative analysis of pollen tube guidance." BMC Plant Biol **10**: 32.

Stuhrwohldt, N., R. I. Dahlke, A. Kutschmar, X. Peng, M. X. Sun and M. Sauter (2015). "Phytosulfokine peptide signaling controls pollen tube growth and funicular pollen tube guidance in Arabidopsis thaliana." Physiol Plant **153**(4): 643-653.

Sundaresan, V. and M. Alandete-Saez (2010). "Pattern formation in miniature: the female gametophyte of flowering plants." Development **137**(2): 179-189.

Swanson, R., A. F. Edlund and D. Preuss (2004). "Species specificity in pollen–pistil interactions." Annu Rev Genet **38**: 793-818.

Takayama, S., H. Shiba, M. Iwano, H. Shimosato, F. S. Che, N. Kai, M. Watanabe, G. Suzuki, K. Hinata and A. Isogai (2000). "The pollen determinant of self-incompatibility in Brassica campestris." PNAS **97**(4): 1920-1925.

Takeuchi, H. and T. Higashiyama (2012). "A species-specific cluster of defensin-like genes encodes diffusible pollen tube attractants in Arabidopsis." PLoS Biol **10**(12): e1001449.

Tebbji, F., A. Nantel and D. P. Matton (2010). "Transcription profiling of fertilization and early seed development events in a solanaceous species using a 7.7 K cDNA microarray from Solanum chacoense ovules." BMC Plant Biol **10**: 174.

Terry, M. E. and B. A. Bonner (1980). "An Examination of Centrifugation as a Method of Extracting an Extracellular Solution from Peas, and Its Use for the Study of Indoleacetic Acid-induced Growth." Plant Physiology **66**(2): 321-325.

Thom, I., M. Grote, J. Abraham-Peskir and R. Wiermann (1998). "Electron and X-ray microscopic analyses of reaggregated materials obtained after fractionation of dissolved sporopollenin." Protoplasma **204**(1-2): 13-21.

Tomato Genome Consortium (2012). "The tomato genome sequence provides insights into fleshy fruit evolution." Nature **485**(7400): 635-641.

Tunc-Ozdemir, M., C. Rato, E. Brown, S. Rogers, A. Mooneyham, S. Frietsch, C. T. Myers, L. R. Poulsen, R. Malho and J. F. Harper (2013). "Cyclic nucleotide gated channels 7 and 8 are essential for male reproductive fertility." PLoS One **8**(2): e55277.

Uebler, S., T. Dresselhaus and M. L. Marton (2013). "Species-specific interaction of EA1 with the maize pollen tube apex." Plant Signal Behav **8**(10) doi: 10 4161/psb 25682.

van Eldik, G. J., R. K. Ruiter, P. H. Colla, M. M. van Herpen, J. A. Schrauwen and G. J. Wullems (1997). "Expression of an isoflavone reductase-like gene enhanced by pollen tube growth in pistils of Solanum tuberosum." Plant Mol Biol **33**(5): 923-929.

Voinnet, O., S. Rivas, P. Mestre and D. Baulcombe (2003). "An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus." Plant J **33**(5): 949-956.

Wang D., C. Zhang, D. J. Hearn, I. H. Kang, J. A. Punwani, M. I. Skaggs, G. N. Drews, K. S. Schumaker, R. Yadegari (2010). "Identification of transcription-factor genes expressed in the Arabidopsis female gametophyte." BMC Plant Biol **10**: 110.

Wang, H., L. C. Boavida, M. Ron and S. McCormick (2008). "Truncation of a protein disulfide isomerase, PDIL2-1, delays embryo sac maturation and disrupts pollen tube guidance in Arabidopsis thaliana." Plant Cell **20**(12): 3300-3311.

Wang, H., H. M. Wu and A. Y. Cheung (1993). "Development and Pollination Regulated Accumulation and Glycosylation of a Stylar Transmitting Tissue-Specific Proline-Rich Protein." Plant Cell **5**(11): 1639-1650.

Wang, S. S., F. Wang, S. J. Tan, M. X. Wang, N. Sui and X. S. Zhang (2014). "Transcript profiles of maize embryo sacs and preliminary identification of genes involved in the embryo sac-pollen tube interaction." Front Plant Sci **5**: 702.

Webb, M. C. and E. G. Williams (1988). "The Pollen Tube Pathway in the Pistil of Lycopersicon peruvianum." Ann Bot **61**(4): 415-423.

Wen, F., H. D. VanEtten, G. Tsaprailis and M. C. Hawes (2007). "Extracellular proteins in pea root tip and border cell exudates." Plant Physiol **143**(2): 773-783.

Wheeler, M. J., B. H. de Graaf, N. Hadjiosif, R. M. Perry, N. S. Poulter, K. Osman, S. Vatovec, A. Harper, F. C. Franklin and V. E. Franklin-Tong (2009). "Identification of the pollen self-incompatibility determinant in Papaver rhoeas." Nature **459**(7249): 992-995.

Willemse, M. T. M. and M. A. W. Franssen-Verheijen (1988). Ovular Development and Pollen Tube Growth in the Ovary of Gasteria verrucosa (Mill.) H. Duval as Condition for Fertilization.

Sex Reprod in Higher Plants. M. Cresti, P. Gori and E. Pacini, Springer Berlin Heidelberg: 357-362.

Wu, H. M., H. Wang and A. Y. Cheung (1995). "A pollen tube growth stimulatory glycoprotein is deglycosylated by pollen tubes and displays a glycosylation gradient in the flower." Cell **82**(3): 395-403.

Wu, H. M., E. Wong, J. Ogdahl and A. Y. Cheung (2000). "A pollen tube growth-promoting arabinogalactan protein from nicotiana alata is similar to the tobacco TTS protein." Plant J **22**(2): 165-176.

Wu, K., M. F. Rooney and R. J. Ferl (1997). "The Arabidopsis 14-3-3 multigene family." Plant Physiol **114**(4): 1421-1431.

Wuest, S. E., K. Vijverberg, A. Schmidt, M. Weiss, J. Gheyselinck, M. Lohr, F. Wellmer, J. Rahnenfuhrer, C. von Mering and U. Grossniklaus (2010). "Arabidopsis female gametophyte gene expression map reveals similarities between plant and animal gametes." Curr Biol **20**(6): 506-512.

Yamamoto, Y., M. Nishimura, I. Hara-Nishimura and T. Noguchi (2003). "Behavior of vacuoles during microspore and pollen development in Arabidopsis thaliana." Plant Cell Physiol **44**(11): 1192-1201.

Yang, H., N. Kaur, S. Kiriakopolos and S. McCormick (2006). "EST generation and analyses towards identifying female gametophyte-specific genes in Zea mays L." Planta **224**(5): 1004-1014.

Yang, W.-C., D.-Q. Shi and Y.-H. Chen (2010). "Female Gametophyte Development in Flowering Plants." Ann Rev Plant Biol **61**(1): 89-108.

Yang, Y. W., K. N. Lai, P. Y. Tai and W. H. Li (1999). "Rates of nucleotide substitution in angiosperm mitochondrial DNA sequences and dates of divergence between Brassica and other angiosperm lineages." J Mol Evol **48**(5): 597-604.

Yang, Z. (2007). "PAML 4: phylogenetic analysis by maximum likelihood." Mol Biol Evol **24**(8): 1586-1591.

Yetisen, A. K., L. Jiang, J. R. Cooper, Y. Qin, R. Palanivelu and Y. Zohar (2011). "A microsystem-based assay for studying pollen tube guidance in plant reproduction." J Micromechanics and Microengineering **21**(5).

Yu, G. H., J. Zou, J. Feng, X. B. Peng, J. Y. Wu, Y. L. Wu, R. Palanivelu and M. X. Sun (2014). "Exogenous gamma-aminobutyric acid (GABA) affects pollen tube growth via modulating putative Ca2+-permeable membrane channels and is coupled to negative regulation on glutamate decarboxylase." J Exp Bot **65**(12): 3235-3248.

Yu, H. J., P. Hogan and V. Sundaresan (2005). "Analysis of the female gametophyte transcriptome of Arabidopsis by comparative expression profiling." Plant Physiol **139**(4): 1853-1869.

Zhou, L., S. A. Bokhari, C. J. Dong and J. Y. Liu (2011). "Comparative proteomics analysis of the root apoplasts of rice seedlings in response to hydrogen peroxide." PLoS One **6**(2): e16723.

Zhou, P., K. A. Silverstein, L. Gao, J. D. Walton, S. Nallu, J. Guhlin and N. D. Young (2013). "Detecting small plant peptides using SPADA (Small Peptide Alignment Discovery Application)." BMC Bioinformatics **14**: 335.

Zinkl, G. M., B. I. Zwiebel, D. G. Grier and D. Preuss (1999). "Pollen-stigma adhesion in Arabidopsis: a species-specific interaction mediated by lipophilic molecules in the pollen exine." Development **126**(23): 5431-5440.

Zonia, L. and T. Munnik (2008). "Vesicle trafficking dynamics and visualization of zones of exocytosis and endocytosis in tobacco pollen tubes." J Exp Bot **59**(4): 861-873.