Identification and Characterization of Membrane-associated Proteins of the *Arabidopsis* Female Gametophyte

DISSERTATION ZUR ERLANGUNG DES
DOKTORGRADES DER NATURWISSENSCHAFT (DR. RER. NAT.)
DER FAKULTÄT FÜR BIOLOGIE UND VORKLINISCHE MEDIZIN
DER UNIVERSITÄT REGENSBURG

vorgelegt von
Thomas Hackenberg

aus
Bad Reichenhall

im Jahr
2017

Das Promotionsgesuch wurde eingereicht am:

Die Arbeit wurde angeleitet von:

PD Dr. Stefanie Sprunck

Unterschrift:

Contents

Contents

Contents

Summary

In the angiosperm life cycle, successful double fertilization marks the transition of the gametophytic to the sporophytic phase. The formation of the male and the female gametophytes and the subsequent interactions of the gametophytes and the male and the female gametes require cell-cell communication which heavily relies on molecules acting on the cell surface.

In this work, a transcriptomic dataset of manually microdissected *Arabidopsis* female gametophytic cells (egg cells, central cells, and synergid cells) was bioinformatically processed to enable for a selection of differentially upregulated candidate genes encoding for cell surface-associated, or extracellular localized proteins. After exhaustive *in silico* analyses the transcriptome data were validated *in vitro* by quantitative PCR, promoter:reporter studies, and whole mount *in situ* hybridization, confirming the array-based data in the majority of cases.

A subset of candidate genes was selected and functionally analyzed by reverse genetics, relying on T-DNA insertion lines, artificial microRNA, and CRISPR/Cas9-mediated genome editing. Moreover, subcellular localization and expression of candidate gene-reporter fusion proteins were determined in stable transgenic *Arabidopsis* lines.

The expression of the secreted SPOROZOITE SURFACE PROTEIN-RELATED (SSPR) started early during megagametogenesis and remained in all cell types of the female gametophyte. CRISPR/Cas9-mediated genome editing suggests SSPR to be functional during progression of the female gametophyte and awaits further investigations.

The role of the membrane-localized family of domain of unknown function 962 (DUF962) proteins in plants was successfully linked to phytosphingosine signal recycling on the molecular level, as the egg cell-expressed member of this family (DUF962_1) was verified to be a true ortholog of *Saccharomyces cerevisiae* MPO1: a key player in the catabolism of 2-hydroxylated C_{16} fatty acids to odd-numbered fatty acids. However, the role of DUF962_1 in the egg cell remains elusive as upon amiRNA-mediated downregulation of *DUF962_1* in the egg cell no effect on this cell was observed but a high frequency of early developmental arrests in the coenocytic female gametophyte.

Expression studies of Tetraspanin GFP fusion proteins in the *Arabidopsis* gametophytes revealed TET8 expression in the mature egg cell, while TET9 was detected in stage FG4 and in the egg cells, the central cells, and weakly in the antipodal cells. Hand pollination with a sperm

cell marker line series suggests that in contrast to some mammalian tetraspanin proteins, TET8 and TET9 are unlikely to contribute to the events of male and female gamete interactions.

In a combined approach of T-DNA insertion lines and CRISPR/Cas9-mediated genome editing a quadruple knock out mutant for the egg cell-expressed Tetraspanins *TET8* and *TET9* and the sperm cell-expressed *TET11* and *TET12* was established, and phenotypically characterized. This revealed a possibly essential function of TET9 in the female gametophyte prior to embryo sac cellularization. Moreover, in a split ubiquitin-based interactor screen for TET9 conducted on an previously established EC-like cell line revealed MEMBRANE STEROID BINDING PROTEIN2 to be a possible membrane-based interactor of TET9.

1 – Introduction

1-1 The Life Cycle of Angiosperms

The life cycle of plants is characterized by an alternating gametophytic, and sporophytic phase. In flowering plants (angiosperms), the visible plant represents the sporophyte that dominates over the gametophyte, which is differentiated from the sporophyte by meiosis and is specialized in sexual reproduction. While the male gametophyte consists of two cell types, one vegetative cell (the pollen grain) in which two sperm cells are enclosed, the female gametophyte consists of four distinct cell types: one egg cell, one central cell, two synergid cells, and three antipodal cells. Successful double fertilization marks the end of the gametophytic phase. During double fertilization one sperm cell fuses with the egg cell forming a zygote and the second sperm cell fuses with the central cell forming the endosperm, respectively. While the zygote develops into the embryo which will give rise to the successive generation, the endosperm assumes a supplementary role in providing nutrients for the embryo but perishes later (McCormick 1993; Yadegari *et al.,* 2004).

1-2 Development of the Polygonum-type Female Gametophyte in *Arabidopsis*

The *Arabidopsis* female gametophyte development can be divided into megasporogenesis and megagametogenesis, which result in the *Polygonum*-type gametophytic pattern that predominates in angiosperms (Figure 1-1) (Drews *et al.,* 2011). Development of the female gametophyte takes place in the ovule primordium, which emerges as protrusion from the inner ovary wall of the placenta (Robinson-Beers *et al.,* 1992). Within the distal end of the developing ovule, a single subepidermal nucellus cell gives rise to the archesporial cell that elongates, polarizes, and finally differentiates into the megaspore mother cell (MMC) (Grossniklaus *et al.,* 1998). In rice, the assignment of germline fate was shown to be restricted by the nucellus-expressed Leucine-rich repeat receptor-like kinase (LRR-RLK) MULTIPLE SPOROCYTES1 (MSP1) (Nonomura *et al.,* 2003), and its coexpressed ligand TAPETUM DETERMINANT-LIKE1A (Zhao *et al.,* 2008). The MSP1 ortholog in *Arabidopsis* is EXCESS MICROSPOROCYTES1/EXTRA SPOROGENOUS CELLS/ (EMS1/EXS), another LRR-RLK which is required for cytokinesis during microsporogenesis (Zhao 2002). Moreover, in order to achieve this the interaction of EMS1/EXS with TAPETUM DETERMINANT1 (TPD1) is required (Jia *et al.,* 2008). TPD1 is a small cysteine-

rich peptide that is secreted from the microsporocyte (PMC) (Grelon *et al.,* 2016). Upon EMS1/EXS-TPD1 binding, periclinal cell divisions of the parietal cells are initiated, which subsequently leads to tapetum cell fate determination (Grelon *et al.,* 2016). Furthermore, the tapetum cells now suppress PMC cell fate assumption (Grelon *et al.,* 2016). The MMC and PMC undergo heterochromatin decondensation to allow for a permissive transcriptional state, in a possibly conserved scenario leading to somatic-to-reproductive cell fate transition (She *et al.,* 2013; She *et al.,* 2014). During the process of megasporogenesis the outer and inner integument layers are initiated from the epidermal cell layer and start to enclose the nucellus (Schneitz *et al.,* 1995).

Megasporogenesis concludes when, accompanied by callose deposition, the MMC undergoes meiosis and gives rise to four haploid megaspores, of which three execute programmed cell death (Rodkiewicz 1970; Webb *et al.,* 1990). The survival of the remaining megaspore was demonstrated to rely on the plasma membrane-localized ARABINOGALACTAN PROTEIN 18 (AGP18) (Zhang *et al.,* 2011; Demesa-Arevalo *et al.,* 2013). Moreover, survival of the megaspore was linked to local callose degradation, which was retained around dying megaspores in *Tillandsia* (Papini *et al.,* 2011).

During the following process of megagametogenesis the functional megaspore undergoes three rounds of karyokineses. Furthermore, the rapidly developing embryo sac is increasingly enclosed by the inner and outer integument layers in an asymmetric manner, thereby establishing a micropylar and chalazal pole (Schneitz *et al.,* 1995) (Figure 1-1 A).

In ovule stage 3-I, the one-nucleate female gametophyte (stage FG1) passes through one round of karyokinesis into stage FG2 where the resulting nuclei are arranged along a longitudinal axis from the chalazal to microyplar pole (Schneitz *et al.,* 1995). Now, the outer integuments enclose the nucellus and inside the developing embryo sac a central vacuole forms from several small vacuoles which marks ovule stage 3-III (FG3) (Schneitz *et al.,* 1995; Christensen *et al.,* 1997).

Stage FG4 is achieved after another round of karyokinesis resulting in four nuclei and the inner integument now enveloping the nucellus as well (Schneitz *et al.,* 1995). During stage FG4 progression, the newly formed nuclei of the chalazal pole rearrange in a longitudinal orientation while the micropylar nuclei remain in orthogonal orientation to the chalazal micropylar axis (Webb *et al.,* 1994; Schneitz *et al.,* 1995; Christensen *et al.,* 1997).

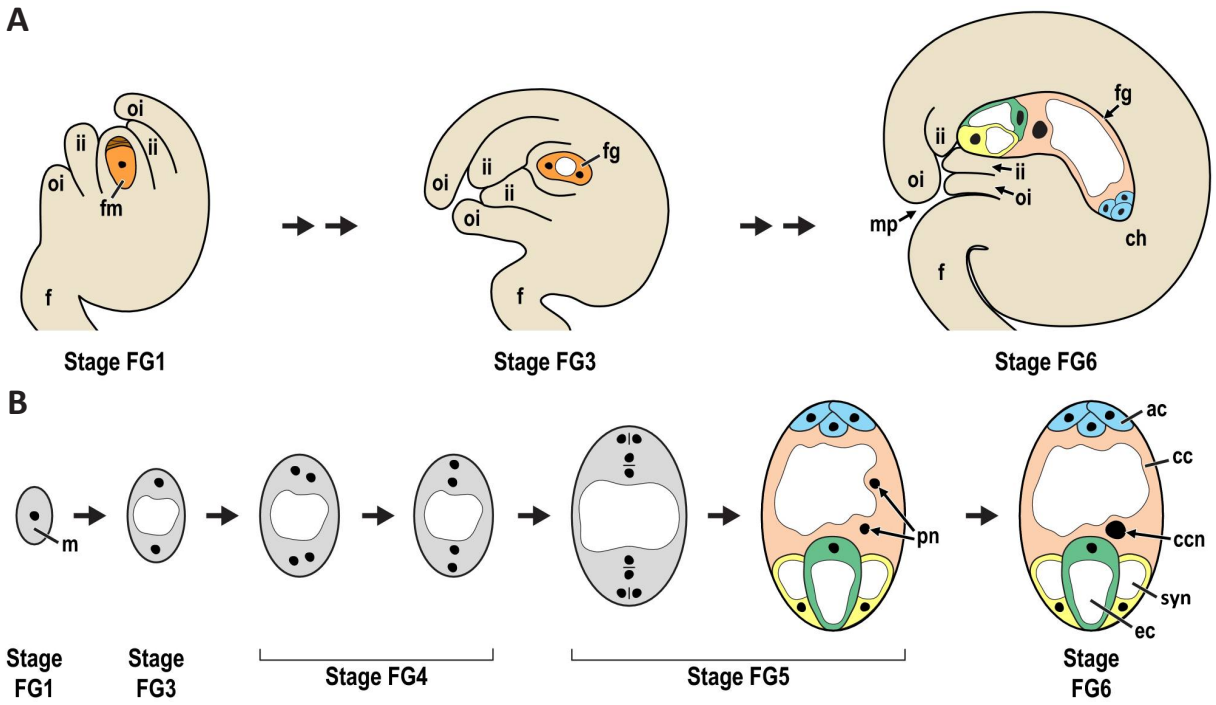


Figure 1-1 Female gametophyte development of the *Polygonum*-type.
(**A**) Schematic overview of the developing female gametophyte from ovule stage 3-I (FG1), with a functional megaspore and short integuments, via ovule stage 3-III (FG3) with a two-celled embryo sac, completely enclosed by the integuments, to the fully cellularized mature ovule stage 3-VI (FG6) with two synergid cells (yellow), and the egg cell (green) located closer to the micropylar pole, the central cell (orange), and at the chalazal pole the three antipodal cells (blue).
(**B**) Schematic overview of megagametogenesis, starting from the functional megaspore in FG1 stage via two consecutive karyokineses to FG4 stage. FG5 stage is reached by mitotic division of all FG4 stage nuclei including cellularization. Finally FG6, the mature stage, is reached after fusion of the polar nuclei to form the homodiploid central cell. Additionally, the seven-celled mature embryo sac contains one haploid egg cell, two haploid synergid cells, and three haploid antipodal cells. ac = antipodal cell, cc = central cell, ccn = central cell nucleus, ch = chalazal pole, ec = egg cell, f = funiculus, fg = female gametophyte, fm = functional megaspore, ii = inner integuments, mp = micropylar pole, oi = outer integuments, pn = polar nuclei, syn = synergid cell, (Schneitz *et al.,* 1995; Drews and Koltunov 2011 with minor modifications).

Now, the coenocytial embryo sac undergoes a last round of karyokinesis and shortly afterwards the most centrally localized polar nuclei migrate towards each other, with the chalazal polar nucleus migrating further than the micropylar nucleus (Christensen *et al.,* 1997). Phragmoblast-mediated cell wall formation between sister and non-sister nuclei completes cellularization and leads to the seven-celled eight-nuclate embryo sac in ovule stage 3-V (FG5) (Webb *et al.,* 1994; Schneitz *et al.,* 1995). Stage FG5 concludes when the polar nuclei fuse with each other to give rise to the homodiploid central cell nucleus (Schneitz *et al.,* 1995). The resulting mature embryo sac contains seven cells – the egg cell and the two synergid cells, forming the egg apparatus, the central cell, and the three chalazal pole-localized antipodal cells (Schneitz *et al.,* 1995) (Figure 1-1 B).

1-3 Cell-fate determination within the mature embryo sac

For embryo sac development, a variety of indispensable genes with association to processes involving chromatin remodeling, mitosis, nuclear migration, intracellular protein sorting, cell expansion, proteolysis, and cellularization were identified (Christensen *et al.,* 1998; Siddiqi *et al.,* 2000; Park *et al.,* 2004; Pagnussat *et al.,* 2005; Blanvillain *et al.,* 2008). These genes are regularly essential for both, male and female gametophyte development (Boavida *et al.,* 2009). However, in contrast to megasporogenesis, so far no cell surface, or extracellular localized proteins were determined crucial for megagametogenesis, with the exception of maize EGG APPARATUS-LIKE1 (EAL1), an egg cell secreted peptide which represses central cell identity in antipodal cells (Krohn *et al.,* 2012).

Furthermore, cell fate decisions at the chalazal pole were shown to depend on the Histidine kinase CKI1-mediated signaling and, connected to CKI1, MYB119 and MYB64 expression (Rabiger *et al.,* 2013; Yuan *et al.,* 2016). The determination of egg cell identity relies on the RWP-RK DOMAIN CONTAINING1 (RKD1) and RWP-RK DOMAIN CONTAINING2 (RKD2) transcription factors (Koszegi *et al.,* 2011). Moreover, the type I MADS-box transcription factors AGAMOUS-LIKE80 (AGL80) and AGAMOUS-LIKE61 (AGL61) function together and are essential for central cell identity assumtion (Portereiko *et al.,* 2006; Bemer *et al.,* 2008; Steffen *et al.,* 2008). Furthermore, the RNA splicing-associated protein LACHESIS is known to restrict egg cell fate assumption in accessory cells like the synergid cells (Gross-Hardt *et al.,* 2007). LACHESIS is regulated by CLOTHO/GAMETOPHYTIC FACTOR1, also a component of the spliceosome, that furthermore regulates restriction of central cell fate (Moll *et al.,* 2008).

However, apart from MSP1 (Nonomura *et al.,* 2003), TDL1 (Zhao *et al.,* 2008), EMS1/EXS (Zhao 2002), TPD1 (Jia *et al.,* 2008), AGP18 (Zhang *et al.,* 2011), and EAL1 (Krohn *et al.,* 2012), almost nothing is known about membrane-associated or secreted proteins involved in signaling during female gametophyte development

1-4 Double Fertilization in Angiosperms

Although apomixis can prove a beneficial short term adaption under specific enviromental conditions like habitat fragmentation or loss of pollinators (Jacquemyn *et al.,* 2012), and was reported for numerous species (Bicknell *et al.,* 2004; Barcaccia *et al.,* 2013), in the long term, sexual reproduction is evolutionary favored in diploid organisms (Bai 2015), as recessive

mutations will be masked less in haploid generations and consequently harmful alleles will be purged (Hojsgaard *et al.,* 2015). Flowering plant sexual reproduction via double fertilization is achieved in a multi-step process which involves the successful germination of the pollen on the stigma, the growth and guidance of the pollen tube through the style, followed by ingrowth into the micropylar region of the ovule, sperm cell delivery by pollen tube burst and receptive synergid degeneration (Dresselhaus *et al.,* 2013). The last step of double fertilization involves sperm cell adhesion to the female gametes, gamete activation and finally plasmogamy followed by karyogamy (Dresselhaus *et al.,* 2013).

1-4-1 Gametophytic Interactions during Double Fertilization

The short distance attraction of the pollen tube to grow through the micropyle towards the female gametophyte is mediated by synergid cell-secreted LURE proteins (Takeuchi *et al.,* 2012). Recently LUREs were found to interact with the pollen-specific and tube tip-localized receptor-like kinase PRK6, which subsequently activates the molecular switch ROP1, a Rho GTPase, and directs the pollen tube growth direction via Rho GTPase signaling in a LURE concentration-dependent manner (Takeuchi *et al.,* 2016). Furthermore, LURE sensing is achieved by the cooperation of PRK6 with PRK1 and PRK3 (Takeuchi *et al.,* 2016). Also, the pollen tube-expressed cell surface-localized leucine-rich repeat receptor-like kinase heteromer MDIS1-MIK confers LURE1 sensing (Wang *et al.,* 2016). The LURE-sensing competency of the pollen tube was achieved by the arabinogalactan polysaccharide AMOR (Mizukami *et al.,* 2016), which also enhances pollen germination (Jiao *et al.,* 2017). In addition, several AGPs were found pistil-expressed along the growth route of the pollen tube towards the ovule and considered to be mediators of female-male crosstalk (Pereira *et al.,* 2016). The pollen tube, carrying two sperm cells tethered to the vegative nucleus, enters the ovule at the micropylar pole and grows close to, and occasionally around, one receptive synergid cell (Denninger *et al.,* 2014) (Figure 1-2 A). The cell surface recognition of the synergid cell and the pollen tube is mediated by FERONIA (FER), a member of the *Catharanthus roseus* subfamily of LRR-RLKs (Huck *et al.,* 2003), (Escobar-Restrepo *et al.,* 2007). In the ovule, FER is specifically expressed in the synergid cells and accumulates asymmetrically in the synergid membrane at the filiform apparatus (Escobar-Restrepo *et al.,* 2007). In the sporophyte, however, FER is ubiquitously expressed and in roots

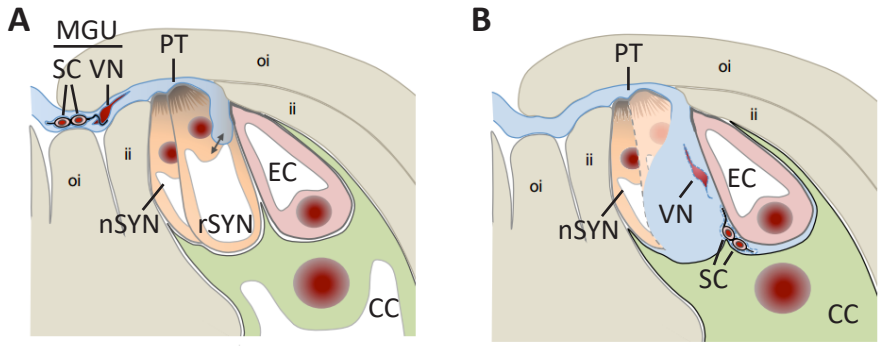


Figure 1-2 Pollen tube arrival, attraction, and perception in Angiosperms.
(**A**) Schematic overview of the micropylar pole of the ovule and the entering pollen tube, containing the male germ unit consisting of two sperm cells tethered to the vegetative nucleus, in contact with the receptive synergid cell. (**B**) Schematic representation of the burst pollen tube and sperm cell discharge towards the fusion site between egg cell, and central cell. EC = egg cell, CC = central cell, ii = inner integument, MGU = male germ unit, nSYN = non receptive synergid cell, oi = outer integument, PT = pollen tube, rSYN = receptive synergid cell, SC = sperm cell, VN = vegetative nucleus (Dresselhaus *et al.,* 2016 with minor additions).

it was shown that FER interacts with the extracellular peptide RAPID ALKANIZATION FACTOR1 (RALF1) (Haruta *et al.,* 2014). Upon RALF1-binding, FER functions as scaffold for formation of a complex composed of multiple LRR-RLKs (Stegmann *et al.,* 2017). In the ovule, the pollen tube perception depends on the interaction of FER with LORELEI, a GPI-anchored, synergid cell-expressed protein (Capron *et al.,* 2008; Li *et al.,* 2015), and on EARLY NODULIN FACTOR14, another GPI-anchored protein with an arabinogalactan glycomodule that was also verified as FER interactor and found essential for pollen tube reception (Hou *et al.,* 2016).

During the phase of physical synergid cell – pollen tube interaction of approximately 30 to 50 minutes, high oscillations in cytosolic Ca^{2+} levels within the receptive synergid cell and the pollen tube were observed, leading, after the strongest increase in cytosolic Ca^{2+} levels, to pollen tube burst and receptive synergid degeneration (Denninger *et al.,* 2014). Reactive oxygen species (ROS), a key component of many signaling pathways including defense response, were shown to be required for pollen tube burst in FER-mediated cytoplasmatic Ca^{2+}-dependent processes, finally delivering the sperm cells to the fusion site (Duan *et al.,* 2014; Ngo *et al.,* 2014; Figure 1-2 B).

1-4-2 Direct Cell-Cell Interactions during Double Fertilization

Once the sperm cells are released from the pollen tube, the egg cell and the central cell were reported to spike in cytosolic calcium levels in a well-timed manner (Denninger *et al.,* 2014; Hamamura *et al.,* 2014). After sperm cell discharge into the fusion site of discontinued cell walls

an average time span of 7.4 minutes passes (Hamamura *et al.,* 2011). During this time span the two connected sperm cells reorient and rearrange towards the respective female gametes until one sperm cell is attached to central cell and egg cell, respectively (Huang *et al.,* 2015). Then the sperm cells separate and plasmogamy occurs with the female gametes, accompanied by another spike in cytosolic Ca^{2+} levels of the egg cell (Hamamura *et al.,* 2011; Denninger *et al.,* 2014; Hamamura *et al.,* 2014). After successful gamete fusion, the endosperm starts to undergo rapid karyokinesis. In between the first and second nuclear division the persistent synergid fuses with the developing endosperm, and the synergid nucleus degenerates in an ethylene-dependent manner to prevent from polytubey (Völz *et al.,* 2013; Maruyama *et al.,* 2015). Elimination of the persistent synergid cell was shown to depend on yet another Arabinogalactan protein (AGP4), termed JAGGER, in a still unknown mechanism (Pereira *et al.,* 2016). Meanwhile, the freshly formed zygote elongates along the future apical-basal axis, establishes polarity, and asymmetrically divides (ten Hove *et al.,* 2015). This asymmetrical division gives rise to a large basal cell which will, with one exception, form the suspensor after a series of transverse divisions, and a small apical cell which will form the embryo (ten Hove *et al.,* 2015).

1-5 Key Molecular Players during Gamete Interactions and Fusion

The final stages of double fertilization are still incompletely understood on the molecular level. GAMETE EXPRESSED 2 (GEX2) a single pass transmembrane protein with extracellular filamin domains forming an immunoglobulin (Ig)-like fold was found essential for attachment of the gametes to each other (Figure 1-3 A; Mori *et al.,* 2014). EGG CELL1.1 (EC1.1), an egg cell-specifically expressed small cysteine-rich protein under transcriptional regulation of SUPPRESSOR OF FRIGIDA (Resentini *et al.,* 2017), needs to be secreted in order to initiate a preferential redistribution of GENERATIVE CELL SPECIFIC1/HAPLESS 2 (GCS1/HAP2) from the endomembrane system of the sperm cells to the plasma membrane, and subsequently enable for gamete fusion (Sprunck *et al.,* 2012). GCS1/HAP2 is a single transmembrane protein required for membrane fusion and fertility not only in angiosperms but also in algae like *Chlamydomonas*, and in parasites like *Plasmodium* (Johnson *et al.,* 2004; von Besser *et al.,* 2006; Mori *et al.,* 2006; Wong *et al.,* 2010) (Figure 1-3 A).

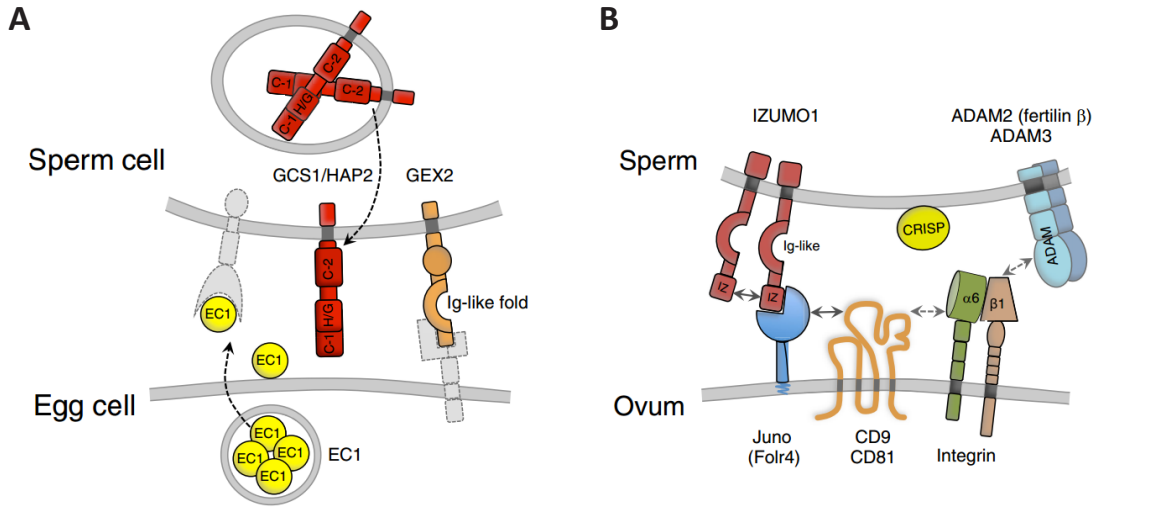


Figure 1-3 Key players of gamete interaction and fusion in flowering plants and in mammals. (**A**) In *Arabidopsis*, upon sperm cell arrival at the fusion site the small cysteine-riche peptide EGG CELL1 (EC1.1) is secreted from the egg cell and initiates preferential relocalization of GENERATIVE CELL SPECIFIC1/HAPLESS2 (GCS1/HAP2) to the sperm cell plasma membrane, which subsequently acts as fusogen and mediates merging of the membranes. GAMETE EXPRESSED2 (GEX2) is also exposed on the sperm cell plasma membrane and mediates attachment to the female gametes. (**B**) Adhesion and recognition during mammalian fertilization is mediated by IZUMO1 on the sperm and JUNO on the egg plasma membrane. Upon IZUMO1-JUNO binding, IZUMO1 undergoes a conformational shift, and dimerizes which abolishes JUNO binding capacity. Furthermore, Tetraspanin proteins CD81, and CD9 were determined essential for fertility, and CD9 mediates assembly of fusion competent sites. Also ADAM proteins and integrins are important mediators of attachment (Dresselhaus *et al.,* 2016).

Furthermore, it contains a large extracellular designated GCS1/HAP2 domain which was shown to trimerize once inserted into the target lipid bilayer in a similar fashion as viral class II fusogens, or the *C.elegans* fusogen EFF-1, and thereby mediates merging of the membranes (Fedry *et al.,* 2017). Hence, it was postulated that these species share a common ancestral gene, as *GCS1/HAP2* assumes the role of an ancient fusogen (Wong *et al.,* 2010; Fedry *et al.,* 2017).

Interestingly, *GCS1/HAP2*, is absent from mammals. During the final stage of mammalian fertilization, the sperm attaches to the egg plasma membrane by the interaction of IZUMO1 (Inoue *et al.,* 2005), which is only exposed to the sperm plasma membrane upon sperm activation (Satouh *et al.,* 2012), and the GPI-anchored folate receptor homolog JUNO (Spiegelstein *et al.,* 2000; Bianchi *et al.,* 2014) (Figure 1-3 B). Notably, JUNO is also expressed and essential for a subpopulation of T-cells in the lyme (Yamaguchi *et al.,* 2007). Upon established IZUMO1-JUNO binding, IZUMO1 undergoes a conformational shift, dimerizes and no longer binds JUNO (Inoue *et al.,* 2015). The resolved IZUMO1 crystal structure (Ohto *et al.,* 2016) revealed structural similarities to domains found in *Plasmodium* proteins named SPECT (sporozoite microneme protein essential for cell traversal) and TRAP1 (thrombospondin repeat anonymous protein

1), two proteins expressed by the invasive sporozoite stage of *Plasmodium berghei* parasites (Nishimura *et al.,* 2016). This shared domain, an extensible beta ribbon domain, can undergo, as shown for TRAP1, a conformational shift and thereby, in combination with additional protein-binding domains, facilitates the gliding motility of TRAP1 (Song *et al.,* 2012).

Until now the exact molecular mechanism of the final sperm egg fusion is unknown. In addition to the JUNO/IZUMO1-mediated recognition and adhesion of mammalian gametes, ADAM proteins, and, partly by interaction with ADAMs, the integrins alpha6 beta1, and alpha9 beta1 are known mediators of sperm egg adhesion (Georgadaki *et al.,* 2016). Furthermore, alpha6 beta1 integrin interacts with the Tetraspanin protein CD9 (Ziyyat *et al.,* 2006) (Figure 1-3 B), which concomitantly accumulates in the contact zone of egg and sperm prior to fusion (Chalbi *et al.,* 2014). Thus, CD9 is considered to laterally organize the gamete fusion machinery on the egg membrane by recruiting necessary cis-partners, as it is known to generate fusion competent sites on the egg membrane (Jegou *et al.,* 2011).

1-6 Tetraspanins, Membrane Microdomains and Sphingolipids

Tetraspanins are a large evolutionary highly conserved family of four pass transmembrane proteins, present in plants and mammals, with their respective N-, and C-termini facing the cytosol, a small occasionally glycosylated extracellular domain, and a large extracellular domain (Wang *et al.,* 2012; Charrin *et al.,* 2014). The large extracellular domain contains several conserve cysteine residues and mediates protein interactions and oligomerization (Schmidt *et al.,* 2016), which was also strongly dependent on palmitoylation of TET9 at the cytosolic side of the transmembrane domains (Figure 1-4; Seigneuret *et al.,* 2001; Berditchevski *et al.,* 2002). Moreover, Tetraspanin proteins are known to indirectly interact with the cytoskeleton via their cytosolic C-termini (Sala-Valdes *et al.,* 2006), and facilitate assembly of supramolecular plasma membrane-based complexes, mostly with integrins (Rubinstein *et al.,* 1994; Yanez-Mo *et al.,* 2009). In mice, Tetraspanin CD81, and CD9 were determined essential for fertilization (Miyado *et al.,* 2000; Rubinstein *et al.,* 2006), and shown to promote muscle cell fusion (Charrin *et al.,* 2013), although they are also known to prevent the fusion of mononuclear phagocytes (Takeda *et al.,* 2003). Additionally, CD9 is known to generate fusion competent sites on the egg membrane (Jegou *et al.,* 2011), in accordance with the concept of Tetraspanin-enriched

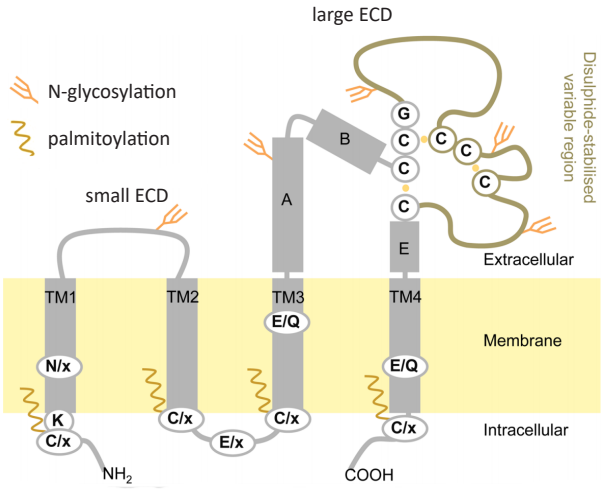


Figure 1-4 The schematic Tetraspanin structure.
A representative Tetraspanin structure with conserved amino acid residues displayed in circles, conserved transmembrane helices 1 - 4, and three conserved helices in the large extracellular domain. Furthermore, palmitoylation, and glycosylation sites, as well as disulfide briges are indicated. ECD = extracellular domain, TM = transmembrane domain (Skaar *et al.,* 2015 with minor additions).

microdomains (Hemler 2003), which pose platforms for mediating processes such as adhesion or exocytosis (Bailey *et al.,* 2011; Perez-Hernandez *et al.,* 2013).

Initially, these microdomains were identified due to their detergent insolubility similar to those of the lipid rafts (Hemler 2003). Lipids rafts constitute hypothetical organizing centers of the plasma membrane with lipid-based sorting mechanisms that in-, and exclude certain proteins (Hancock 2006). Meanwhile, increasing amounts of eukaryotic, and prokaryotic lipid raft-associated proteomic datasets, derived of the analysis of detergent-resistant membrane fractions, are becoming available, revealing lipoproteins, and in eukaryotic datasets tetraspanins (Bae *et al.,* 2004; Dubois *et al.,* 2015; Shah *et al.,* 2015; Toledo *et al.,* 2015).

Notably, these detergent-resistant membrane fractions are predominantly assembled of cholesterol and sphingolipids (Hancock 2006). Sphingolipids arise from the irreversible condensation of palmitoyl-CoA with serine performed by the endoplasmic reticulum-localized serine palmitoyltransferase, thereby forming, after a successive reduction step, Sphinganine (Hanada 2003). This long chain base (LCB) can be N-acylated, then hydroxylated and glycosylated to gain complex glycosphingolipids like phytoceramide (Figure 1-4). Glycosphingolipids are major constituents of the plasma membrane or serve as component of the GPI-anchor (Breslow *et al.,* 2010; Gault *et al.,* 2010). Alternatively, the LCB can be hydroxylated, desaturated, phosphorylated and thereby give rise to a variety of bioactive molecules serving as second messengers or secreted ligands for cell-surface receptors (Breslow *et al.,* 2010; Gault *et al.,*

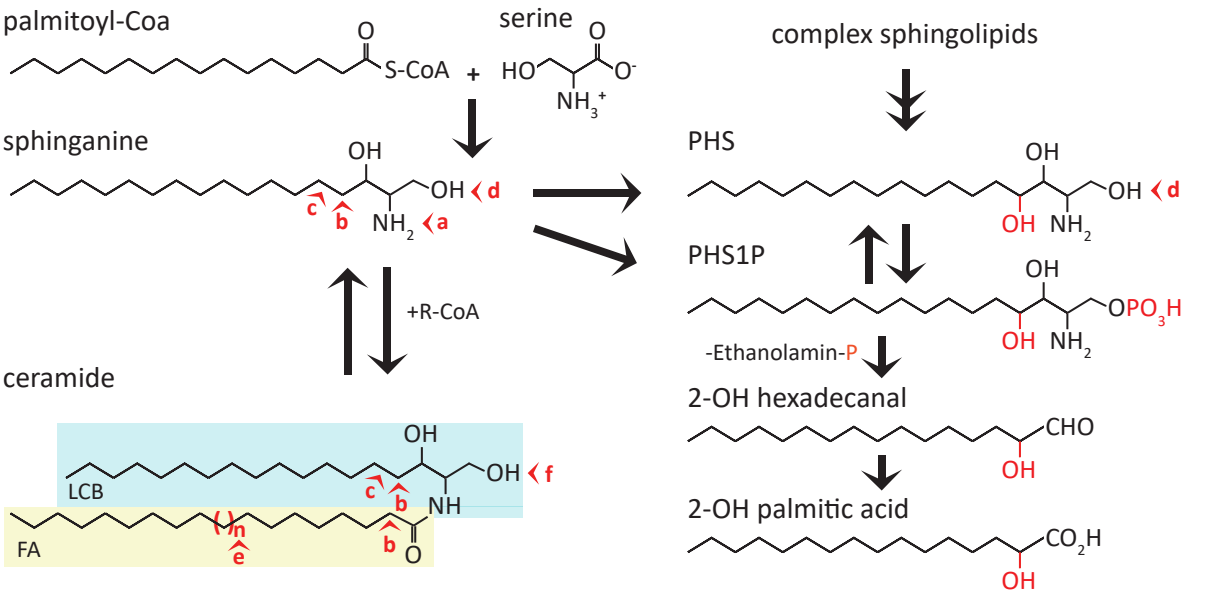


Figure 1-5 Sphingolipid metabolism.
The first and irreversible step in sphingolipid biosynthesis is the condensation of palmitoyl-CoA and serine to 3-ketosphinganine, which, after subsequent reduction, results in the long chain base (LCB) (sphinganine). Subsequently, various modifications of sphinganine are possible: N-acylation (a) yields ceramide; hydroxlation (b) yields phytosphingosine (PHS), the major plant and yeast LCB; desaturation (c) yields sphingosine, the major mammalian LCB; phosphorylation (d) yields (phyto)sphingosine-1-phosphate, an important signaling molecule. Furthermore ceramides vary in fatty acid (FA) chain length (e) (yellow box). The head group position (f) of the LCB part of ceramides (blue box) can be phosphorylated, or derivatized with inositol or mannose moieties, phosphorylcholine, or complex glycans yielding highly diverse glycosphingolipids. Complex sphingolipids are catabolized by deacylation to LCB and FA-aldehyde moieties, here exemplary to PHS. PHS is phosphorylated, then irreversible degraded to phosphorylethanolamine and 2-hydroxyhexadecanal, which can be oxidized further to 2-hydroxyhexadecanal. 2-OH = 2-hydroxy, FA = fatty acid, LCB = long chain base, PHS = phytosphingosine, PHS1P = phytosphingosine-1-phosphate, (with modifications from Breslow and Weissman 2010, and Kondo *et al.,* 2014).

2010). Interestingly, sphingolipids are also actively metabolized in the nucleus, where, for example, sphingosine-1-phosphate was shown to inhibit histone deacetylases and thereby suppress transcription (Hait *et al.,* 2009).

The entire sphingolipid breakdown is mediated via phosphorylation of the LCB, like phytosphingosine (PHS) to phytosphingosine-1-phosphate (PHS1P) and successive irreversible degradation by the mostly endoplasmic reticulum-localized phospholyase to ethanolamine and the respective acyl aldehydes like 2-hydroxyhexadecanal (Tsegaye *et al.,* 2007; Aguilera-Romero *et al.,* 2014) (Figure 1-4). Disturbed sphingolipid homeostasis was reported to have pleiotropic negative effects (Nishikawa *et al.,* 2008) with toxic outcome (Han *et al.,* 2010). Especially accumulation of free LCBs and their phosphorylated counterparts due to necrotrophic funghi-derived fumonisin B1-induced inhibition of the ceramide synthase was shown to be lethal (Wright *et al.,* 2003), as these molecules have a severe impact on cellular survival and pro-death signaling (Pata *et al.,* 2010).

1-7 Aims of this Work

The importance of cell surface-exposed and extracellular-localized proteins during gametophyte and gamete interactions has been emphasized. However, the knowledge about the proteins acting on the surface of the female gametophytic cells and on the sperm cells is scarce and much remains to be investigated.

The aim of this work was to examine the composition of cell surface proteins on female gametophytic cells and on sperm cells, and to select candidate proteins for functional characterization with respect to female gametophyte development and double fertilization.

To achieve this, own and publicly available microarray-based transcriptome data from *Arabidopsis* female gametophytic cells (egg cells, central cells, synergid cells) and sperm cells were subjected to comprehensive bioinformatics studies. The data were processed and analyzed together with microarray data from diverse sporophytic tissues to extract those genes encoding membrane proteins as well as membrane-associated and secreted proteins which are differentially upregulated (enriched) in either one or a combination of female gametophytic cells. This was achieved by comparing (contrasting) the female gametophytic cells and sperm cells with the sporophytic tissues and with each other.

However, the presence of a transcript does not necessarily imply the presence of the translated protein. Therefore, an *Arabidopsis* cell line with an egg cell-like transcriptional profile was used to generate full proteome and membrane proteome data, and to compare them with those of a root-derived callus, to find out whether this approach may serve as a tool to overcome the quantitative restrictions limiting the use of egg cells for proteomics.

Candidate genes, derived from transcriptome studies, were selected to verify the tissue- and/or cell type-specific expression pattern by quantitative PCR, whole mount *in situ* hybridization, and microscopic analyses of transgenic *Arabidopsis* plants expressing either promoter:reporter or genomic:reporter constructs.

Based on these data, a subset of candidate proteins was selected to study their subcellular localization in the developing female gametophyte and during double fertilization. The function of three candidate protein families was investigated by reverse genetics, including the CRISPR/Cas9 approach, and protein-protein interaction studies.

2 – Results

2-1 Acquisition, Processing, and Evaluation of the Transcriptomic Data of *Arabidopsis* Gametophytic Cells

2-1-1 Previous Work: Isolation of *Arabidopsis* Female Gametophytic Cells by Micromanipulation, and ATH1 GeneChip Hybridization

In a previous effort by Lucija Soljic (Soljic 2012), the three major cell types of the *Arabidopsis* female gametophyte were isolated by manual microdissection (MM) as described in Englhart *et al.,* 2017 (in press), and subjected to single cell transcriptome analysis using the GeneChip® Arabidopsis ATH1 Genome Array. Hybridization was performed in four batches as provided by the "Kompetenzzentrum für Fluoreszente Bioanalytik" (Regensburg). Four CEL datasets were obtained, containing four replicates for the egg cell (EC), and three replicates each for the central cell (CC), and the synergid cells (SYN). Soljic excluded the second (E1) egg cell replicate from further studies. The remaining CEL data were processed, including sperm cell (SC)-derived data (Borges *et al.,* 2008), as described in Soljic 2012, and resulted in 10065 EC, 11641 CC, 8728 SYN, and 5829 sperm cells expressed genes called „present" in three of three replicates (Soljic 2012).

2-1-2 Identification of Differentially Expressed Genes encoding Membrane-associated and Secreted Proteins

To acquire a more in-depth transcriptome analysis a combined approach was realized in collaboration with Dr. Maxim Messerer and Dr. Daniel Lang (Helmholtz Center Munich). Here, the CEL data of all ten replicates of female gametophytic cells isolated by MM (Solijc 2012), and published laser-assisted micro dissected (LAM) female gametophytic cells (Wuest *et al.,* 2010), sperm cells (Borges *et al.,* 2008), globular embryo (Spencer *et al.,* 2007; Slane *et al.,* 2014), endosperm of germinating seeds (Penfield *et al.,* 2006; Dekkers *et al.,* 2013), leaves (Usadel *et al.,* 2008; Tang *et al.,* 2012), roots (Stepanova *et al.,* 2007; Lin *et al.,* 2011), and seedlings (Zheng *et al.,* 2009; He *et al.,* 2016) were included in a comprehensive analysis. Also the same analysis was performed while excluding the LAM-derived data of female gameotphytic cells (Wuest *et al.,* 2010). Expression data from the developing endosperm four days post

fertilization (Day *et al.,* 2007; Day *et al.,* 2008) were non includable into the analysis due to platform compatibility issues between [ATH1-121501] Affymetrix Arabidopsis ATH1 Genome Array and HortResearch_Arabidopsis_27K_Operon long oligo set_V1.0. For both analyses (MM+LAM and MM) all respective 78 and 69 CEL files (Supplemental Table 1) were subjected to frozen Robust-Multi-array Analysis (fRMA) from the Bioconductor package frma (McCall *et al.,* 2010) for normalization, background subtraction, and batch-effect correction, which revealed good clustering of the respective replicates in multi dimensional scale (MDS) analysis, and identified the root- and the SC-derived replicates to be most distant from the remaining samples (Figure 2-1 A, B).

Intensity value cutoff was defined as mean value of the negative controls, times two the standard deviation. The individual overlap of probe sets with signal intensities above the cutoff value (5.599934), excluding internal controls, between the replicates of MM-isolated female gametophytic cells is depicted in Figure 2-1 C – F. The maximum overlap ranged from 60% (EC), via 68% (SYN), and 72% (SC), to 78% (CC). These probe sets corresponded to 11509 genes in the EC, 13186 genes in the CC, 12044 genes in the SYN, and 14056 genes in the SC, above the predefine cutoff value in all replicates. In total, the MM+LAM and MM-only analyses identified 19059 genes (19684 probe sets) and 19259 genes (19907 probe sets) above cutoff, respectively.

Arabidopsis thaliana gene descriptions and gene ontology (GO) terms were downloaded from TAIR v10. Prediction of transmembrane domains, and subcellular protein localizations were obtained from Dr. Rainer Schwacke directly as referenced in ARAMEMNON v8.1 (Schwacke *et al.,* 2003), and SUBAcon (Tanz *et al.,* 2013; Hooper *et al.,* 2014), respectively. To address transcriptomic differences between gametophytic cells (EC, CC, SYN, SC) and the sporophytic tissues (endosperm, embryo, leaf, root, seedling) differential gene expression was calculated using the GLM functionality of the limma package (Ritchie *et al.,* 2015), also including a corrective term for possible batch effects (Supplemental Table 5-1).

Differentially expressed genes (DEGs) were determined by contrasting of the female gametophytic cells and male gametes against sporophytic tissues, and by contrasting the gametophytic cells and gametes against each other. Furthermore, DEG annotation was extended by TAIR10-Subcellular_Predictions.xlsx, and TAIR10_functional_descriptions.

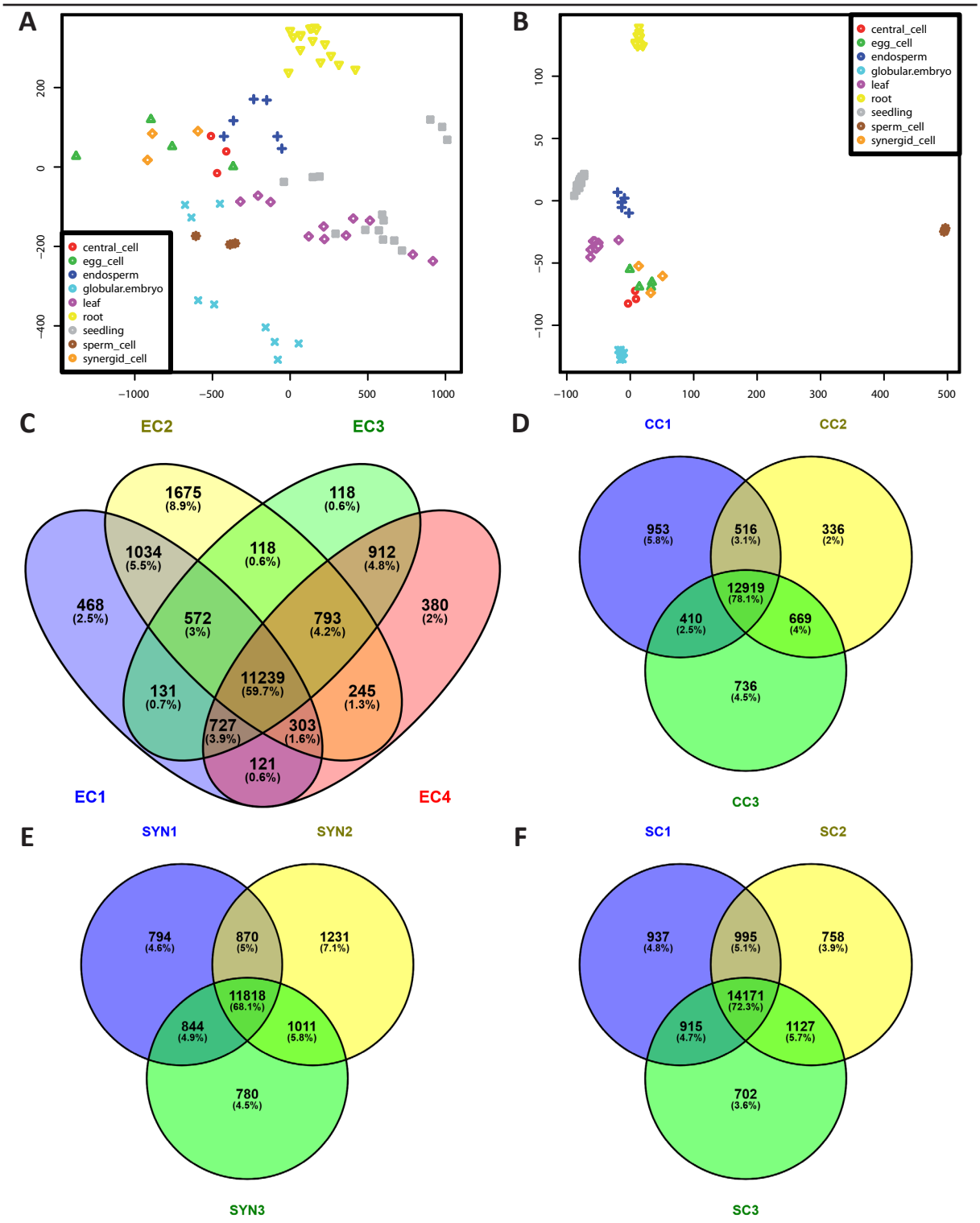


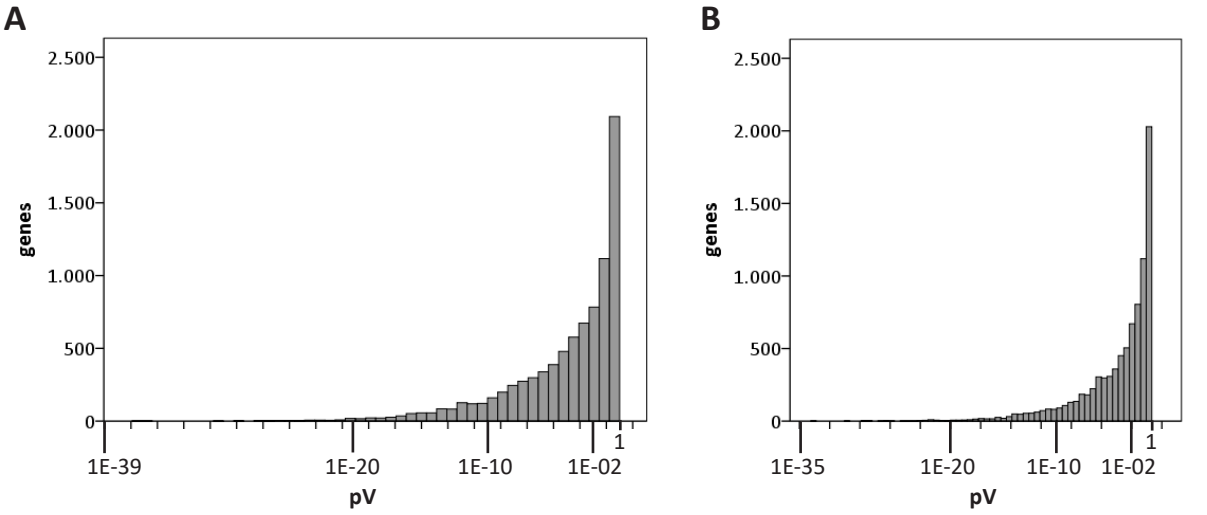
Figure 2-1 Transcriptomic data of female gametophytic cells.
69 ATH1-121501 CEL files of manual microdissected (MM) female gametophytic cells (egg cells, central cells, synergid cells), male gametes (sperm cells), and sporophytic tissues (endosperm, embryo, leaf, root, seedling) were subjected to frozen robust Multi-Array analysis. Multidimensional scale plot depiction before (**A**) and after (**B**) normalization, background subtraction, and batch effect correction revealed good replicate clustering. The MDS plots (**A, B**) were contributed by Dr. Maxim Messerer.
The detected probe sets of MM female gametophytic cells and sperm cells after background subtraction and their distribution onto the individual replicates of the egg cell (**C**), the central cell (**D**), the synergid cells (**E**), and the sperm cells (**F**). EC = egg cell, CC = central cell, SC = sperm cell, SYN = synergid cell.

txt obtained from TAIR v10. Gene functional descriptions, cellular compartments (CO), and localization predictions were assigned to the respective AGIs. Also, data derived of probe sets corresponding to multiple AGIs were removed.

As the aim of this study was to address cell-to-cell signaling and cell-surface interactions, all mitochondrial and chloroplastidial-derived gene identifiers were removed along all genes identifiers corresponding to putatively neither membrane-associated nor extracellular localized proteins. After calculating sporophytic and gametophytic contrasts in total 8506 (MM+LAM) and 8608 (MM-only) DEGs encoding putative memrane-associated, or secreted proteins remained.

2-1-3 Statistical and Gene Ontology-based Evaluation of the obtained Transcriptomic Data

To determine which of the two conducted analyses (MM+LAM, and MM) will be used for in-



C

pV	MM+LAM	MM
DEGs	8506	8608
Mean	0.146	0.172
SE(Mean)	0.003	0.003
Median	0.003	0.011
SD	0.259	0.272

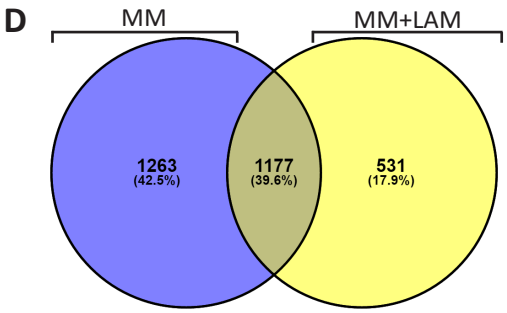


Figure 2-2 Statistical analysis of the LAM+MM and MM analyses.
Probability value frequencies of 8506 DEGs of the MM+LAM (**A**) and 8608 DEGs of the MM-only analysis (**B**), and the respective statistical summary depicting mean and median values (**C**) showed a lower median value for the LAM+MM analysis. (**D**) This diagram contains all DEGs of the EC, the CC, and the SYN with an FC of 1.5 or higher, from either of the reference contrasts (sporophytic and gametophytic), excluding SC-derived DEGs, and shows the 2440 DEGs identified in the dataset derived of the MM analysis (blue subset), and the 1708 DEGs identified in the dataset derived of the LAM+MM analysis (yellow subset). The non intersection subsets were used for GO term analyses (Table 2-1).

detail bioinformatic characterization, both were compared by descriptive statistics and GO term overrepresentation.

First of all, the obtained adjusted probability value (pV) frequencies of the "EC versus sporophytic tissue" contrast of the MM+LAM and MM analyses were compared. By calculating the median, a higher value (0.011) for the four EC replicate number analysis (MM) was obtained than for the seven EC replicate number analysis (MM+LAM) which was at 0.003 (Figure 2-2 A - C). Based on these findings, the MM+LAM analysis should be preferred over the MM analysis.

To limit the chance of including non-DEGs into further studies, the maximum valid pV was set to 0.001, consequently DEGs with higher pVs were excluded. To assess the differences in DEG composure of the MM+LAM and MM analyses in a GO-dependent, biological processes-related context and in line with the working hypothesis that relatively strong expressed genes will participate in cell-to-cell communications and gamete interactions, a log2 fold change (FC) cut off, to represent upregulated genes, was specified to 1.5 or higher. Upregulated DEGs from sporophytic, and gametophytic contrasts were pooled and SC-only derived DEGs were temporarily excluded to estimate the overlap and non intersection subsets of identified DEGs in LAM+MM and MM analyses. 2440 DEGs originated from the MM analysis and 1708 DEGs originated from the MM+LAM analysis, which overlapped in 1177 DEGs (40%) (Figure 2-2 D). The individual subsets of the DEGs (1263 MM-only, 1177 intersection, and 531 MM+LAM-only)

Table 2-1 Overrepresented biological processes of the DEGs encoding membrane-associated or secreted proteins that were identified only in the MM-LAM analysis or only in the MM analysis but not in both.

All 531 DEGs of the MM+LAM-only non intersection subset (**A**), and all 1263 DEGs of the MM-only non intersection subset (**B**) were subjected to PANTHER-based statistical overrepresentation test. Fold overrepresentation was calculated relative to the *Arabidopsis* full GO term reference database (*Arabidopsis* Ref). Results were sorted by ascending pV, the pV threshold was set to ≤ 0.05. Only overrepresented processes are shown. pV = probability value.

A

GO biological process complete	Arabidopsis Ref(27352)	MM+LAM -only (531)	expected	fold overrep.	pV
photosynthesis (GO:0015979)	159	16	3.09	5.18	3.41E-04
establishment of localization (GO:0051234)	1776	66	34.48	1.91	9.40E-04
transport (GO:0006810)	1758	65	34.13	1.90	1.38E-03
localization (GO:0051179)	1846	66	35.84	1.84	3.50E-03
single-organism transport (GO:0044765)	791	35	15.36	2.28	1.78E-02

B

GO biological process complete	Arabidopsis Ref(27352)	MM-only (1263)	expected	fold overrep.	pV
establishment of localization (GO:0051234)	1776	215	82.27	2.61	3.82E-34
transport (GO:0006810)	1758	213	81.43	2.62	7.87E-34
localization (GO:0051179)	1846	219	85.51	2.56	1.31E-33
vesicle-mediated transport (GO:0016192)	284	56	13.16	4.26	1.75E-15
organic substance transport (GO:0071702)	943	111	43.68	2.54	3.94E-15

were used for PANTHER-based GO term biological process statistical overrepresentation tests, with a predefined pV cutoff of 0.05 and revealed differences in DEG composure (Table 2-1). Notably, “photosynthesis” (GO:0015979) was found most overrepresented by factor 5.18 in the non intersection subset corresponding to the MM-LAM-only DEGs (Table 2-1 A). The remaining overrepresented biological processes like “establishment of localization” (GO:0051234), and “transport” (GO:0006810) were identified in both non intersection subsets, while “vesicle-mediated transport” (GO:0016192) was restricted to the MM-only non intersection subset (Table 2-1 B). Despite the pre-defined pV, and FC cutoffs, the MM+LAM analysis would likely include false positive DEGs derived of photosynthetic tissue and exclude potentially valuable DEGs and was discarded consequently. Therefore, an in-depth *in silico* transcriptome analysis was restricted to the data from SCs and female gametophytic cells, acquired by MM.

2-1-4 *In silico* Analysis of the Gametophytic Transcriptome Data

The gametophytic transcriptome data were either compared to sporophytic tissues, or the respective other gametophytic cells. The resulting DEGs, with a FC ≥ 0.05 or FC ≤ -0.05, and a pV ≤ 0.001, encoding for membrane-associated or secreted proteins, obtained by these two comparisons, were characterized by descriptive statistics and statistical GO term enrichment analysis.

2-1-4-1 *In silico* Analysis of the Gametophytic DEGs obtained from the Sporophytic Contrast

To assess differences and similarities between the gametophytic cells on the level of differential gene expression, basic descriptive statistics were applied to the 3306, 3482, 3364, and 5126 DEGs of the EC, the CC, the SYN, and the SC, respectively after contrasting to sporophytic tissues. The FC median values for all four gametophytic cell types were determined negative between -0.939 (EC), and -1.114 (CC). Similar findings applied to the FC mean values, which were slightly negative in all four cases ranging from -0.125 (SYN) to -0.368 (SC). The skewness, which defines the asymmetry of the distribution around its mean value (Brys *et al.,* 2004), was positive for all four cell types. Median and skewness clearly demonstrated a higher frequency of genes on the left side of the distribution, which resembled downregulation. Finally, kurtosis, which determines the tailedness (Westfall 2014) was found negative for the EC (-0.397), and

the SC (-0.308) indicating fewer than expected, by comparison to the normal distribution, strongly regulated genes. Kurtosis was even for the CC (-0.032), and positive for the SYN (0.256), indicating more strongly regulated genes in the SYN. Also, minimum and maximum values for differential gene expression were SYN-derived. These statistic values are all summed up in Figure 2-3 A, and gene expression frequency histograms for all gametophytic cells are depicted in Figure 2-3 B – E. Judging from gene expression frequencies, the synergid cell was found slightly more different from the gametes than the gametes from each other.

A

	DEGs	Mean	Median	Skewness	SE(Skewness)	Kurtosis	SE(Kurtosis)	Min.	Max.
EC	3321	-0.216	-0.939	0.445	0.042	-0.397	0.085	-5.363	8.316
CC	3497	-0.182	-1.114	0.556	0.041	-0.032	0,083	-6.216	9.430
SYN	3372	-0.125	-1.025	0.455	0.042	0.256	0.084	-7.881	10.778
SC	5126	-0.368	-1.100	0.426	0.034	-0.308	0.068	-7.079	8.921

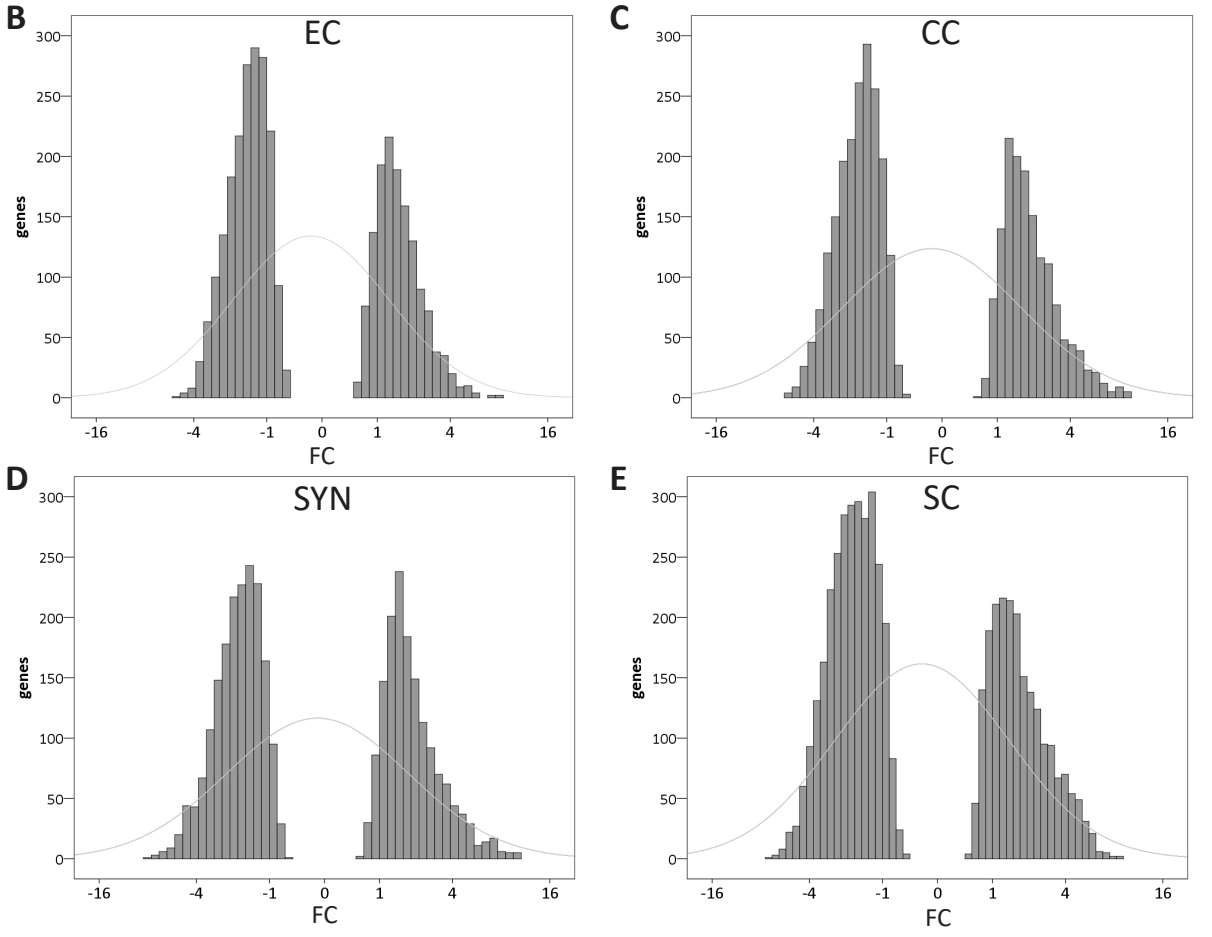


Figure 2-3 Descriptive statistics applied to the DEGs of the female gametophytic cells and male gametes identified by sporophytic contrasting.
Basis descriptive statistics were applied to the DEGs, with pV ≤ 0.001 of the EC, the CC, the SYN, and the SC after contrasting to the sporophyte and summed up (**A**). Log2 fold change frequencies of the DEGs for the EC (**B**), the CC (**C**), the SYN (**D**), and the SC (**E**) depicted as histograms. Grey lines indicate the normal distribution. For all four cell types less DEGs were upregulated than downregulated, which was determined by positive skewness and negative median values. Furthermore, only the SYN showed positive kurtosis thereby demonstrating more than expected strongly regulated genes and a difference from the gametes which had negative kurtosis (EC, SC) and almost even kurtosis (CC). EC = egg cell, CC = central cell, SC = sperm cell, SE = standard error, SYN = synergid cell.

By utilizing GO term enrichment analysis, the possible biological and molecular differences between the gametophytic cells, encoded by the identified DEGs, were addressed. Again the same DEGs, previously utilized for the descriptive statistic analyses, were transformed into GeneEntrezIDs, affiliated with their respective expression values as determined by sporophytic contrasting and subjected to PANTHER-based statistical GO term enrichment analysis. The maximum pV for statistically valid enrichment or depletion of processes was set to 0.05. Similarities in enrichment of biological processes, depicted as PANTHER-slim unified terms, were identified between all female gametophytic cells, especially regarding transport-associated processes like “intracellular protein transport” (GO:0006886), and “protein transport” (GO:0015031) (Table 2-2 A). Also, “RNA metabolic process” (GO:0016070) was found enriched in all female gametophytic cells, while “protein glycosylation” (GO:0006486), and “exocytosis” (GO:0006887) was exclusively found enriched in the SYN. Neither of these processes were found enriched in the SC, here only “DNA metabolic process” (GO:0006259), “DNA repair” (GO:0006281), and “DNA recombination” (GO:0006310) were identified as enriched (Table 2-2 A). GO term-depleted biological processes were related to “carbohydrate transport” (GO:0008643) in the EC and the CC, “defense response to bacterium” (GO:0042742) in the SYN, and a number of metabolic processes like “protein metabolic process” (GO:0019538), “rRNA metabolic process” (GO:0016072), or “translation” (GO:0006412) and “protein folding” (GO:0006457) in the SC (Table 2-2 B). These findings demonstrated, on transcriptomic level, the distance between the SCs, and the cells of the female gametophyte (FG), which had already been observed earlier by MDS plot analysis and moreover indicated slight differences between the SYN, and the female gametes.

The identified 7168 DEGs with an FC ≥ 0.05 or FC ≤ -0.05, (pV ≤ 0.001), after sporophytic contrasting, were distributed relatively evenly onto the individual intersection subsets of gametophytic cells types and male gametes, with the exception of the SC non intersection subset which contained 1874 DEGs (26%), and the EC, CC, SYN, SC intersection subset with 1102 DEGs (15%; Figure 2-4 A). When applying an FC cutoff for upregulation of 1.5 or higher to all identified DEGs, 2438 DEGs remained. Here, the SC non intersection subset still was largest with 798 DEGs (33%), while the EC, CC, SYN, SC intersection had diminished to 120 DEGs (5%), indicating that gametophytic cells shared many downregulated, or slightly upregulated genes

Table 2-2 Enriched and depleted biological processes in gametophytic cells after contrasting to sporophytic tissues.
All identified DEGs with a pV ≤ 0.001 and their affiliated expression values per respective gametophytic cell type were subjected to PANTHER-based GO term enrichment analysis, depicted as enriched (**A**), and depleted (**B**) GO-slim biological processes and sorted by ascending pV (cutoff pV ≤ 0.05). Enriched biological processes found in all female gametophytic cells are depicted in bold letters.

A

cell type	Panther GO-slim Biological process_enriched	genes	pV
EC	**RNA metabolic process** (GO:0016070)	63	4.5E-06
EC	**intracellular protein transport** (GO:0006886)	166	9.8E-06
EC	nucleobase-containing compound metabolic process (GO:0006139)	199	1.1E-05
EC	**protein transport** (GO:0015031)	171	2.1E-05
EC	vesicle-mediated transport (GO:0016192)	133	3.9E-05
EC	phosphate-containing compound metabolic process (GO:0006796)	189	1.9E-02
CC	**RNA metabolic process** (GO:0016070)	42	1.6E-04
CC	nucleobase-containing compound metabolic process (GO:0006139)	168	9.3E-04
CC	primary metabolic process (GO:0044238)	692	4.1E-03
CC	**intracellular protein transport** (GO:0006886)	139	9.9E-03
CC	**protein transport** (GO:0015031)	142	1.1E-02
CC	metabolic process (GO:0008152)	890	3.9E-02
SYN	**protein transport** (GO:0015031)	187	4.4E-10
SYN	**intracellular protein transport** (GO:0006886)	181	5.9E-10
SYN	vesicle-mediated transport (GO:0016192)	141	5.9E-10
SYN	**RNA metabolic process** (GO:0016070)	48	1.2E-03
SYN	transport (GO:0006810)	402	3.4E-03
SYN	localization (GO:0051179)	408	4.5E-03
SYN	nucleobase-containing compound metabolic process (GO:0006139)	164	9.0E-03
SYN	protein glycosylation (GO:0006486)	59	1.6E-02
SYN	exocytosis (GO:0006887)	30	3.8E-02
SC	DNA metabolic process (GO:0006259)	43	6.6E-04
SC	DNA repair (GO:0006281)	20	4.6E-03
SC	DNA recombination (GO:0006310)	9	1.3E-02

B

cell type	Panther GO-slim Biological process_depleted	genes	pV
EC	carbohydrate transport (GO:0008643)	45	1.6E-02
EC	ion transport (GO:0006811)	97	4.1E-02
CC	homeostatic process (GO:0042592)	67	3.7E-03
CC	carbohydrate transport (GO:0008643)	57	2.6E-02
SYN	defense response to bacterium (GO:0042742)	49	1.7E-03
SYN	secondary metabolic process (GO:0019748)	43	1.1E-02
SYN	homeostatic process (GO:0042592)	56	4.1E-02
SC	translation (GO:0006412)	152	1.7E-14
SC	protein metabolic process (GO:0019538)	556	1.2E-05
SC	cellular component biogenesis (GO:0044085)	120	1.3E-05
SC	cellular amino acid metabolic process (GO:0006520)	76	1.5E-04
SC	rRNA metabolic process (GO:0016072)	22	1.6E-04
SC	biosynthetic process (GO:0009058)	301	1.3E-03
SC	primary metabolic process (GO:0044238)	1122	6.7E-03
SC	protein folding (GO:0006457)	41	9.9E-03
SC	cellular amino acid catabolic process (GO:0009063)	12	1.0E-02
SC	glycolysis (GO:0006096)	16	4.1E-02

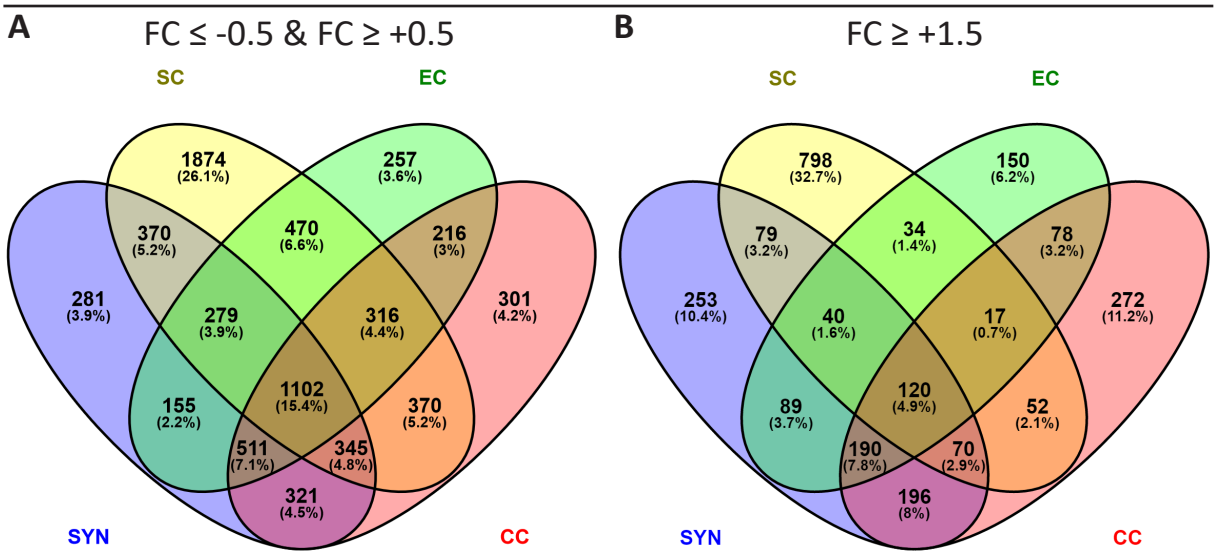


Figure 2-4 Distribution of the DEGs, identified by sporophytic contrasting, among the gametophytic cell types.
Venn diagramm depiction of DEGs with a pV ≤ 0.001 after sporophytic contrasting and their distribution onto the individual gametophytic cell types. All 7168 DEGs (**A**), and only upregulated DEGs with an FC ≥ 1.5 (2438) (**B**). While the SC shared least overlap with the female gametophytic cells 26% (**A**), and 33% (**B**), the remaining subsets contained comparable amounts of DEGs, with the only exception of the EC, CC, SYN, SC intersection subset which was second largest with 1102 DEGs (15%) in (**A**), but reduced by 10% after applying the 1.5 FC cutoff for upregulation. By comparing female gametophytic cells and male gametes to the sporophyte demonstrated that these gamete/gametophytic cell types shared few upregulated genes. EC = egg cell, CC = central cell, FC = log2 fold change, SC = sperm cell, SYN = synergid cell.

but few relatively strongly differentially upregulated genes (Figure 2-4 B). Among these 2438 upregulated DEGs, 485 DEGs encoded for predicted extra cellular localized proteins, of which 264 DEGs were found expressed by the cells of the female gametophyte but not the SCs.

2-1-4-2 *In silico* Analysis of the Gametophytic DEGs obtained from the Gametophytic Contrast

To distinguish the gametophytic cells from each other in a more detailed, transcriptomic way, every gametophytic cell was compared to the respective other cells. These gametophytic contrasts resulted in 2685, 2572, 2210, and 4378 identified DEGs among the EC, CC, SYN, and SC, respectively and these DEG frequencies were subjected to descriptive statistical analysis. Here, the median values for the respective gametophytic cells were determined to positive values for EC (0.691), CC (0.956), and SYN (1.043), while SC was negative (-1.026). Skewness (Brys *et al.,* 2004), was slightly negative for EC (-0.124), even for CC, and SC, and slightly positive for SYN (0.182; Figure 2-5 A – E). Median and skewness pointed out that the EC was, among gametophytic cells, the one with most upregulated genes, while the opposite applied to the SC. Kurtosis was around even for all cell types, only the SYN with a kurtosis value of

0.104 showed the expression-wise strongest regulated genes. Analogous to the sporophytic contrasts described earlier, again PANTHER-based GO term statistical enrichment analysis was performed to assess direct GO term-related differences between the gametophytic cells. While the EC-derived DEGs only showed enrichment in “cellular component biogenesis” (GO:0044085), defined as for example biosynthesis of constituent macromolecules, assembly, and arrangement of constituent parts of a cellular component (AmiGO 2 v2.4.26), and the CC in “generation of precursor metabolites and energy” (GO:0006091), “fatty acid metabolic process” (GO:0006631), the SYN interestingly was found enriched in most processes, e.g. “protein-“ and “primary metabolic processes” (GO:0019538, GO:0044238), “vesicle-mediated

A

	DEGs	Mean	Median	Skewness	SE(Skewness)	Kurtosis	SE(Kurtosis)	Min.	Max.
EC	2694	0.009	0.641	-0.124	0.047	-0.619	0.094	-5.163	6.748
CC	2582	0.321	0.956	0.073	0.048	-0.698	0.096	-4.840	6.118
SYN	2213	0.423	1.043	0.182	0.052	-0.072	0.104	-4.708	7.629
SC	4404	-0.312	-1.026	0.064	0.037	-0.780	0.074	-7.671	7.544

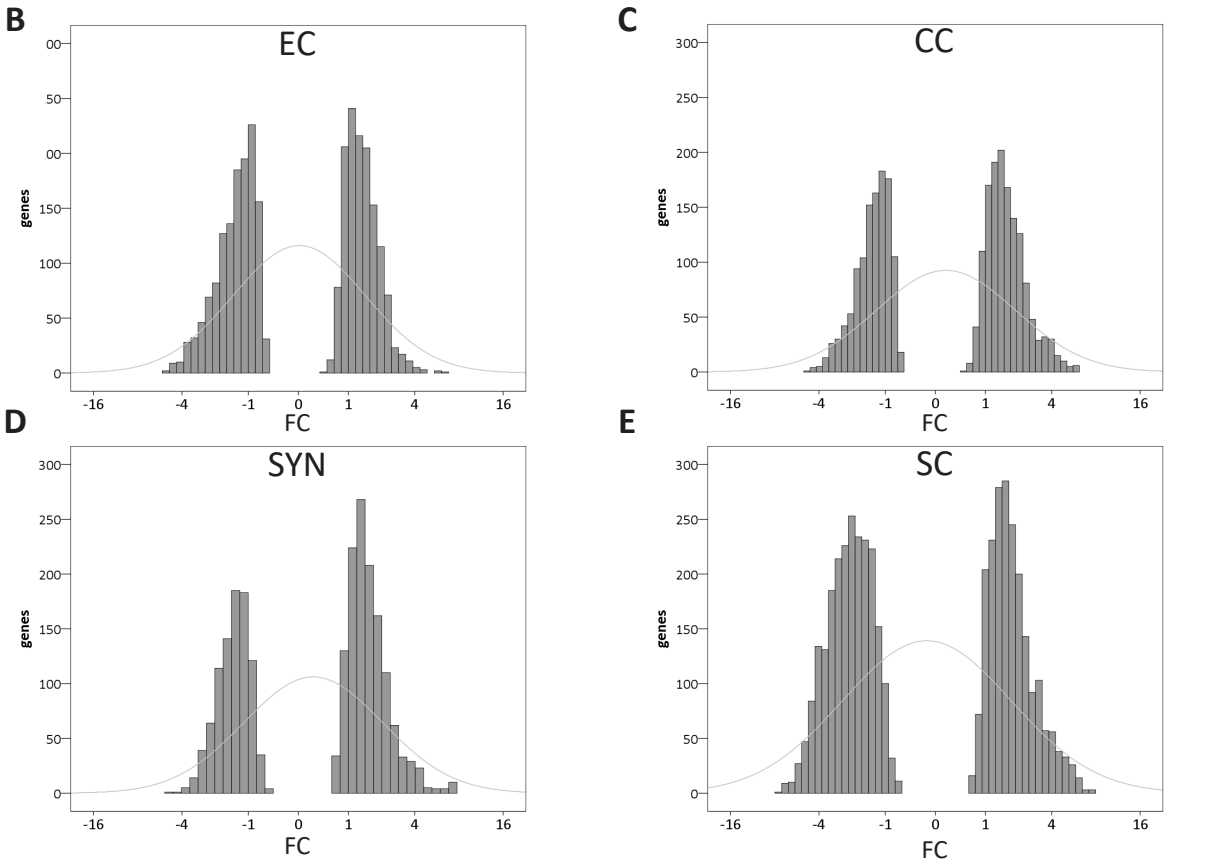


Figure 2-5 Descriptive statistics applied to the DEGs of female gametophytic cells obtained from gametophytic contrasting.

Basis descriptive statistics were applied to DEGs, with a pV ≤ 0.001, of the EC, the CC, the SYN, and SC after contrasting to gametophytic cells and summed up (**A**). Frequencies of the DEGs for the EC (**B**), the CC (**C**), the SYN (**D**), and the SC (**E**) depicted as histograms. Grey lines indicate the normal distribution. For all female gametophytic cells more DEGs were found upregulated than downregulated, determined by the positive median values. Only the median value calculated from SC DEG frequencies was negative. EC = egg cell, CC = central cell, FC = log2 fold change, SC = sperm cell, SE =standard error, SYN = synergid cell.

transport", and "protein transport" (GO:0016192, GO:0015031). Again "DNA recombination" (GO:0006310) was found enriched in the SC (Table 2-3 A). No depleted GO terms were identified in EC, only "mitosis" (GO:0007067) in the CC, "secondary metabolic process" (GO:0019748) in the SYN, and a vast number of depleted processes in SC, like "protein metabolic process" (GO:0019538), and "primary metabolic process" (GO:0044238), or "cellular component biogenesis" (GO:0044085; Table 2-3 B). In summary, GO term statistical enrichment analysis of gamete/gametophytic DEGs showed that the EC was not depleted in any biological processes, the SYN was found enriched in most biological processes, and SCs were identified as the cell type with the most depleted biological processes.

When examining the distribution of DEGs, after gametophytic contrasting, into the individual

Table 2-3 Enriched and depleted biological processes in gametophytic cells after contrasting to gametophytic cells.
All identified DEGs with a pV ≤ 0.001 and their affiliated expression values per gametophytic cell type were subjected to PANTHER-based GO term statistical enrichment analysis, depicted as enriched (**A**), or depleted (**B**) GO-slim biological process, sorted by ascending pV (cutoff pV ≤ 0.05). EC = egg cell, CC = central cell, SC = sperm cell, SYN = synergid cell.

A

cell type	Panther GO-slim Biological process_enriched	genes	pV
EC	cellular component biogenesis (GO:0044085)	73	1.4E-02
CC	generation of precursor metabolites and energy (GO:0006091)	44	8.0E-03
CC	translation (GO:0006412)	96	1.6E-02
CC	fatty acid metabolic process (GO:0006631)	18	1.9E-02
SYN	translation (GO:0006412)	92	1.8E-06
SYN	protein metabolic process (GO:0019538)	260	2.8E-05
SYN	cellular component biogenesis (GO:0044085)	61	6.9E-03
SYN	vesicle-mediated transport (GO:0016192)	118	2.3E-02
SYN	protein transport (GO:0015031)	155	3.0E-02
SYN	primary metabolic process (GO:0044238)	529	3.3E-02
SYN	intracellular protein transport (GO:0006886)	150	3.8E-02
SC	biological regulation (GO:0065007)	240	4.7E-03
SC	DNA recombination (GO:0006310)	7	4.4E-02

B

cell type	Panther GO-slim Biological process_depleted	genes	pV
EC	none		
CC	mitosis (GO:0007067)	17	4.3E-02
SYN	secondary metabolic process (GO:0019748)	23	4.8E-04
SC	translation (GO:0006412)	136	0.0E+00
SC	protein metabolic process (GO:0019538)	491	1.0E-13
SC	primary metabolic process (GO:0044238)	970	7.8E-10
SC	cellular component biogenesis (GO:0044085)	118	8.8E-10
SC	biosynthetic process (GO:0009058)	265	1.3E-07
SC	metabolic process (GO:0008152)	1211	1.0E-05
SC	protein folding (GO:0006457)	34	1.2E-05
SC	rRNA metabolic process (GO:0016072)	22	1.3E-04
SC	cellular amino acid metabolic process (GO:0006520)	53	1.5E-04
SC	generation of precursor metabolites and energy (GO:0006091)	70	1.2E-03
SC	glycolysis (GO:0006096)	11	1.1E-02
SC	cellular amino acid catabolic process (GO:0009063)	12	1.7E-02
SC	carbohydrate metabolic process (GO:0005975)	188	2.2E-02

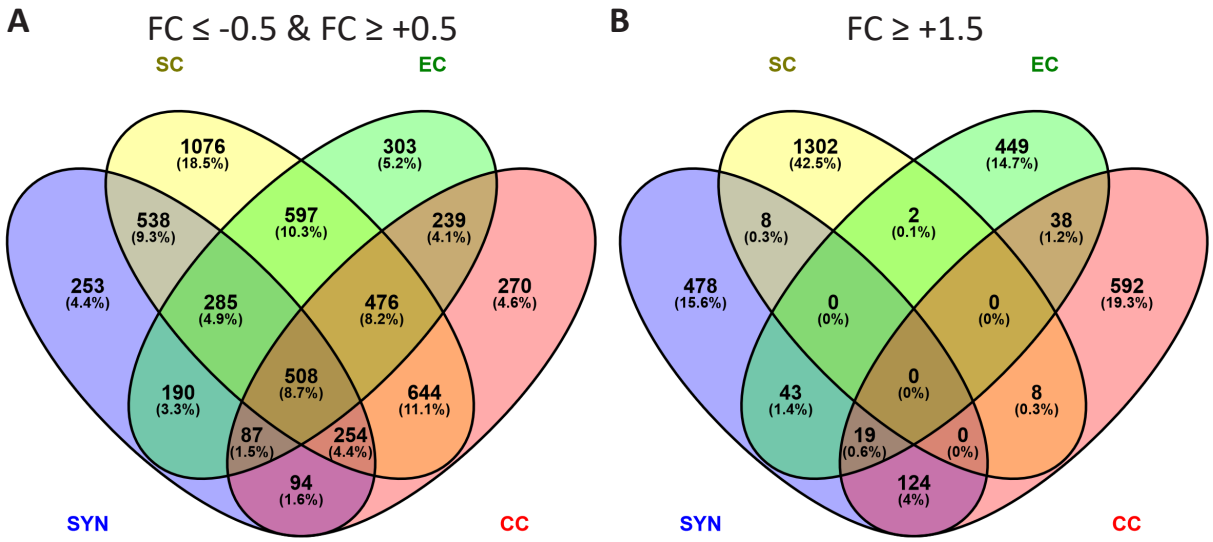


Figure 2-6 Distribution of the DEGs, identfied by gametophytic contrasting among the gametophytic cells.
Distribution of all 6117 DEGs (**A**), and FC ≥ 1.5 upregulated DEGs (3063) (**B**), with pV ≤ 0.001, onto the female gametophytic cells and the male gametes after gametophytic contrasting. The DEGs (**A**) were relatively even distributed onto all subsets, with the exception of the non intersection subset of the SC containing 1076 DEGs (19%). The largest intersection subset was the CC and SC subset with 644 DEGs (11%) (**A**). After applying the cutoff of FC ≥ 1.5 upregulation mostly only DEGs in the indivdual non intersection subsets of the cell types remained, indicating within the gametophytes, cell type-specific differentially upregulated genes (**B**). EC = egg cell, FC = log2 fold change, CC = central cell, SC = sperm cell, SYN = synergid cell.

gametophytic cell types the most overlap was found between SC and CC with 644 DEGs (11%), and again the non intersection subset of SCs was found largest with 1076 DEGs (19%) (Figure 2-6 A). By restricting the identified DEGs of the gametophytic contrasts to the ones upregulated with a FC of 1.5 or higher, 3063 DEGs remained. Most of theses DEGs were cell-type specifically upregulated in only one cell type with relative equal amounts of DEGs in the famel gametophytic cells (15% - 19%; Figure 2-6 B). However, this analysis identified 38 upregulated DEGs in the EC and the CC (Supplemental Table 5-2), while the male and the female gametes shared no upregulated DEGs.

2-1-4-3 The DEGs identified by either Sporophytic or Gametophytic Contrasts show little Overlap

The initial bioinformatic analysis of the transcriptome data aimed at the identification of DEGs in the cells of the female gametophyte and male gametes. This was achieved by either contrasting the transcriptome data to sporophytic tissue, or to the respective other gametophytic cells. In the EC, 3306 DEGs were identified after sporophytic and 2685 DEGs after gametophytic contrasting (Figure 2-7 A). When applying the FC cutoff for upregulation of 1.5, only 718 and

551 DEGs of the EC remained, respectively (Figure 2-7 A). By comparing the FC 1.5 upregulated DEGs identified in both contrasts for all gametophytic cells, the overlap for the EC was at 20% (209 DEGs), and was at 34% for the CC (453 DEGs), 31% for the SYN (403 DEGs), and largest for the SC with 51% (849 DEGs; Figure 2-7 B – E). This comparably large overlap of the DEGs, identified in both comparisons of the SC transcriptome data, indicates the distance of the transcriptomic profile of the SCs to the female gametophytic cells and sporophytic tissues.

In summary, the contrasting of the gametes and the SYN to each other, identified genes that were differentially regulated within the gametophytes and revealed differences in enriched biological processes between the SYN and the gametes based on the transcriptomic data.

A

Contrasts	DEGs FC ≤ -0.5 FC ≥ 0.5 pV ≤ 0.001	DEGs FC ≥ 1.5 pV ≤ 0.001
EC vs sporophyte	3306	718
CC vs sporophyte	3482	995
SYN vs sporophyte	3364	1037
SC vs sporophyte	5126	1210
EC vs CC, SYN, SC	2685	551
CC vs EC, SYN, SC	2572	781
SYN vs EC, CC, SC	2210	672
SC vs EC, CC, SYN	4378	1320

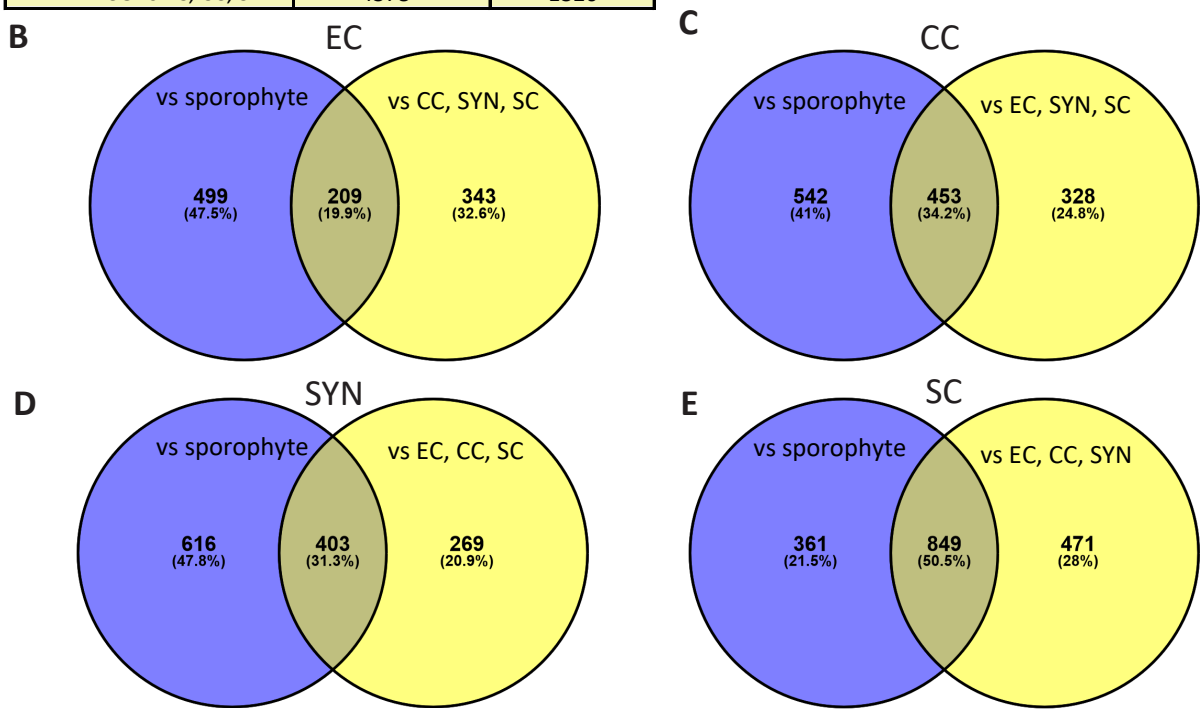


Figure 2-7 DEGs determined by sporophytic and gametophytic contrasts and the overlap between the DEGs identified in the respective contrasts.
All DEGs, and FC ≥ 1.5 upregulated DEGs identified in the female gametophytic cells and the sperm cells after sporophytic contrasting (blue), and gametophytic contrasting (yellow) (**A**). Overlap between FC ≥ 1.5 upregulated DEGs after sporophytic and gametophytic contrasting for EC (**B**), CC (**C**), SYN (**D**), and SC (**E**). EC = egg cell, CC = central cell, FC = log2 fold change, SC = sperm cell, SYN = synergid cell.

Moreover, the DEG composure of the SCs was found different from all female gametophytic cells after gametophytic, as well as sporophytic contrasting.

Relatively even, but potentially gametophyte-specific, expressed genes, that were expressed by multiple cells of the female gametopyte and the SCs, were mostly eliminated by the gametophytic contrasting, while the sporophytic contrasting highlighted gametophyte-enriched genes. In order to investigate cell-to-cell communication between the cells of the FG and SCs and furthermore to investigate putative gamete attachment-mediating proteins, preferentially EC- and/or CC-expressed genes were selected from both data contrasting approaches.

2-2 The RKD2-induced EC-like Cell Line as Tool for EC Membrane-Proteomics

2-2-1 Characterization of the EC-like Cell Line

The female gametophytic cells such as the EC are deeply embedded in the sporophytic tissue and therefore the isolation of sufficient amounts of ECs to conduct protein biochemistry was not feasible. Since transcription-derived data cannot give accurate predictions for the corresponding protein localization and abundance (Evans *et al.,* 2003), an additional approach was pursued.

The ectopic expression of the *RKD2* transcription factor had been shown to assign putative EC-fate to non differentiating cells on the transcriptional level (Koszegi *et al.,* 2011). Therefore,

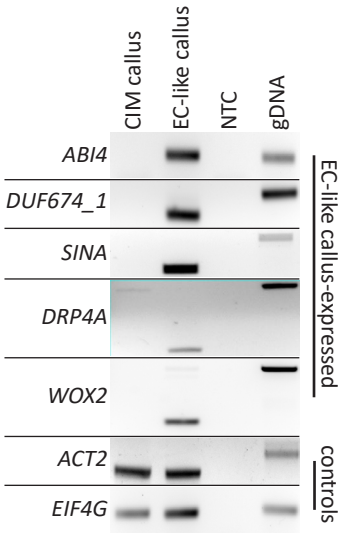


Figure 2-8 EC-like cell line identity verification by RT-PCR.
Transcript presence of five, in EC-like callus compared to CIM-callus, upregulated genes was verified by RT-PCR, and thereby EC-like callus cell line identity. *ACT2* and *EIF4G* served as controls for template abundance. CIM = callus inducing medium derived cell line, gDNA = genomic DNA, NTC = non template control.

proteomic data of this cell line could help in circumventing the given EC protein quantity limitations. Establishing, and in detail analysis of this egg cell-like (EC-like) callus line was part of a different study (Urban 2016). A control cell line was obtained from *Arabidopsis* root segments by callus inducing medium, generated and maintained by Monika Kammerer. The EC-like callus line was maintained by Ingrid Fuchs. Previous microarray-based transcriptomic analyses identified 351 upregulated genes, when comparing an RKD2-induced EC-like callus cell line with a phytohormone-induced callus cell line (Koszegi *et al.,* 2011). Five of these upregulated genes were selected, and the presence of transcript in our EC-like cell line was confirmed by RT-PCR for *ABI4*, *DUF674_1* (=AT5G01150), *DRP4A*, *SINA*, and *WOX2*, thereby verifying the EC-like cell identity (Figure 2-8).

2-2-2 Acquisition of Proteomic Data of the EC-like Cell Line

The potential of this EC-like callus line to serve as tool for biochemical studies was elicited. Two proteomic approaches were pursued utilizing microsomal fractionation (MF), and ammonium persulfate (AMS) and trichloroacetic acid(TCA)-mediated protein precipitations. The resulting proteins extracts were analyzed by mass spectrometry (MS) in collaboration with Dr. Julia Mergner. All mitochondrial-derived or chloroplastidal-derived proteins were removed from the obtained MS-derived data from subsequent analyses.

Since membrane and membrane-associated proteins were major focus of this work, MFs of the EC-like callus, and callus-derived from callus inducing medium (CIM) were generated in four and three replicates, respectively and analyzed in collaboration with Dr. Julia Mergner by mass spectrometry (MS). The proteomic data obtained from Dr. Mergner was manually processed

Table 2-4 Overview of detected proteins in EC-like callus samples.
Proteins detected in microsomal fractions of EC-like callus samples (**A**), and proteins detected in the full proteomic approach of EC-like callus samples (**B**). log2 fold change was calculated from the ratio of LFQ signal intensities of EC-like callus to CIM callus samples. Putative membrane-association or extracellular localization of proteins were assigned according to Aramemnon, TAIR and SUBAcon. FC = log2 fold change, FP = full proteomic approach, MF = microsomal fractionation approach.

A	**EC-like callus (MF)**	**detected**	**FC > 0**	**FC > 1.5**	**EC-like callus-only**
	all proteins	2959	2005	1143	939
	membrane / secreted proteins	1316	817	434	355
B	**EC-like callus (FP)**	**detected**	**FC > 0**	**FC > 1.5**	
	all proteins	7583	4180	759	
	membrane / secreted proteins	2059	1048	212	

by removing all peptides that matched multiple proteins, and by removing all proteins that were only present in one replicate. The ratios of obtained MS-derived signal intensities (LFQ values) between EC-like callus and CIM callus samples were calculated and log2 transformed. Consequently, 2959 proteins in the EC-like callus cell line samples, and 2412 proteins in CIM callus cell line samples were identified. In total, 3351 proteins were identfied in samples of both callus cell lines, of which 939 proteins were absent in CIM callus samples and thereby presumably EC-like specific (28%; Figure 2-9 A). Of these 939 EC-like-specific proteins only 355 were predicted to be membrane-associated or extracellular localized (Table 2-4 A).

Furthermore, to conduct a comprehensive full proteome (FP) approach, proteins from EC-like callus and CIM-callus cell lines protein extracts were precipitated by TCA, and AMS in three technical replicates each. The following MS-based analysis was again conducted in collaboration with Dr. Julia Mergner. Furthermore, log2 transformed fold changes of the ratios of EC-like callus to CIM callus derived LFQ intensities were calculated, and tested for significant differential expression by two-tailed students t-test (interval of confidence = 0.05).

The full proteomic analysis had identified 7583 proteins (Table 2-4 B), of which 2059 proteins were predicted membrane-associated or secreted proteins. By calculating the overlap of the FP data and the MF-based MS data, 4634 proteins were newly identified (58%; Figure 2-9 B). SUBAcon, Aramemonon and TAIR-based protein localization data and transmembrane domain predictions were assigned to the proteins identified in both approaches, then the overlap between at least FC 1.5 upregulated EC-like callus expressed membrane-associated or secreted proteins was determined (Figure 2-9 C). This revealed 571 EC-like callus upregulated

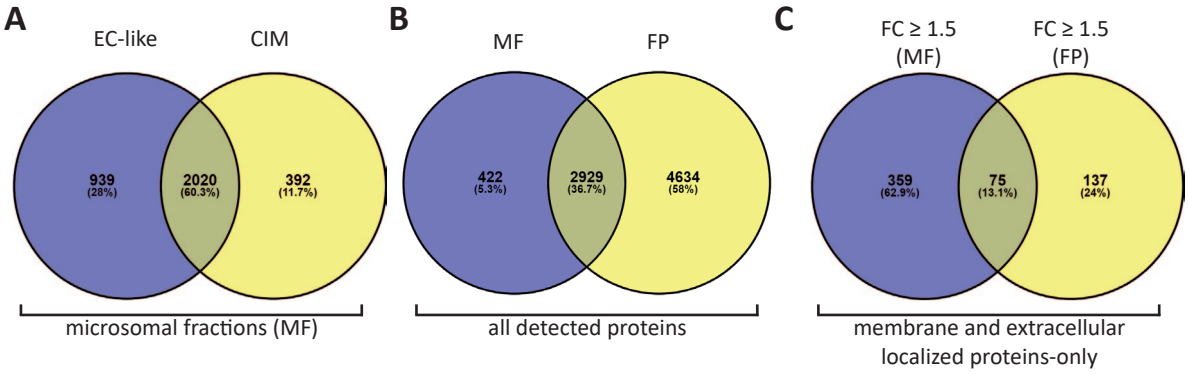


Figure 2-9 Acquired proteomic datasets of EC-like and CIM callus cell lines.
Overlap of 3351 detected proteins in microsomal fractions (MF) of EC-like and CIM callus in more than one replicate (**A**). All 7985 proteins identified by MF-based, and the full proteomic (FP) approach in EC-like and CIM-callus cell lines and their overlap (**B**). Overlap between the FC ≥ 1.5 upregulated EC-like callus membrane-associated and secreted proteins of the MF and FP approaches (**C**). CIM = callus inducing medium-derived cell line, FP = full proteomic approach, MF = microsomal fractions.

putative membrane-associated or extracellular localized proteins in the combined approaches, with only 13% overlap and the vast majority of proteins (63%) had been obtained by the MF approach (Figure 2-9 C), thereby suggesting successful enrichment of membrane proteins in the MF approach.

2-2-3 *In silico* Evaluation of Proteomic Data of the EC-like Cell Line

To assess the quality of the data which had been obtained by MF of the EC-like callus and the CIM-callus samples, and subsequent MS analysis, all 2959 proteins identified in the MF-derived EC-like callus samples were subjected to PANTHER-based statistical overrepresentation test, with aim to identify overrepresented cellular components. This analysis revealed indeed statistically valid overrepresentation of many microsomal compartment-derived cellular components like “nuclear envelope” (GO:0005635), “cytoplasmic membrane-bound vesicle”

Table 2-5 Overrepresented cellular components in the EC-like proteomic data of microsomal fractions.
All 2959 identified proteins in EC-like data of microsomal fractions were subjected to PANTHER-based statistical overrepresentation test. Cellular component overrepresentation was calculated of 2950 proteins (with assigned GO-terms) in relation to the *Arabidopsis* full GO term reference database with 27352 proteins, depicted as GO-slim CO terms, with a maximum pV of 0.05, and sorted by fold overrepresentation in descending order. Only overrepresented processes are shown.

PANTHER GO-Slim Cellular Component	Arabidopsis REF(27352)	EC-like (2950)	expected	fold. overrep.	pV
nuclear envelope (GO:0005635)	58	36	6.26	5.75	1.2E-14
tubulin complex (GO:0045298)	17	10	1.83	5.45	1.1E-03
cytoplasmic membrane-bounded vesicle (GO:0016023)	31	18	3.34	5.38	8.3E-07
nuclear outer membrane-ER membrane network (GO:0042175)	71	34	7.66	4.44	9.8E-11
endosome (GO:0005768)	49	23	5.28	4.35	4.8E-07
mitochondrial inner membrane (GO:0005743)	41	19	4.42	4.3	1.1E-05
nucleolus (GO:0005730)	108	44	11.65	3.78	1.4E-11
vesicle coat (GO:0030120)	57	23	6.15	3.74	6.9E-06
endoplasmic reticulum (GO:0005783)	148	59	15.96	3.7	3.8E-15
ribosome (GO:0005840)	266	105	28.69	3.66	8.8E-27
ribonucleoprotein complex (GO:0030529)	538	201	58.03	3.46	2.3E-48
SNARE complex (GO:0031201)	65	24	7.01	3.42	1.8E-05
microtubule (GO:0005874)	77	26	8.3	3.13	3.1E-05
mitochondrion (GO:0005739)	139	46	14.99	3.07	4.0E-09
Golgi apparatus (GO:0005794)	171	52	18.44	2.82	4.8E-09
cytosol (GO:0005829)	546	156	58.89	2.65	6.9E-25
macromolecular complex (GO:0032991)	1574	431	169.76	2.54	5.4E-67
integral to membrane (GO:0016021)	393	98	42.39	2.31	5.9E-12
protein complex (GO:0043234)	1110	271	119.72	2.26	1.2E-32
cytoplasm (GO:0005737)	2440	556	263.16	2.11	3.5E-61
organelle (GO:0043226)	2646	586	285.38	2.05	9.4E-61

(GO:0016023), or “vesicle coat” (GO:0030120; Table 2-5). Furthermore, to determine putative overrepresented biological processes in the EC-like callus, the 939 proteins that had been identified in the EC-like callus MF samples-only were subjected to PANTHER-based statistical overrepresentation test, which revealed a variety of overrepresented biological processes such as “mitotic cell cycle” (GO:0000278), “seed development” (GO:0048316), “embryo development” (GO:0009790), but also “reproductive structure development” (GO:0048608), and “single organism reproductive process” (GO:0044702; Table 2-6). This indicated an EC-like as well as embryo/seed-like membrane protein composure of this RKD2-induced cell line.

Table 2-6 Overrepresented biological processes in EC-like-only proteomic data of microsomal fractions.
All 939 proteins identified only in the EC-like data of microsomal fractions were subjected to PANTHER-based statistical overrepresentation test. Biological process fold overrepresentation was only calculated relative to the *Arabidopsis* full GO term reference database with 27352 proteins. Biological processes were sorted by fold overrepresentation in descending order with a statistical cutoff of pV ≤ 0.05. Only overrepresented processes are shown, and processes of particular interest are depicted in bold letters.

GO biological process complete	Arabidopsis REF (27352)	EC-like-only (939)	expected	fold overrep.	pV
mitotic cell cycle (GO:0000278)	159	18	5.46	3.3	3.5E-02
organelle fission (GO:0048285)	176	19	6.04	3.14	3.9E-02
seed development (GO:0048316)	511	50	17.54	2.85	2.4E-07
fruit development (GO:0010154)	544	51	18.68	2.73	6.6E-07
embryo development ending in seed dormancy (GO:0009793)	408	37	14.01	2.64	4.0E-04
embryo development (GO:0009790)	423	37	14.52	2.55	9.5E-04
intracellular transport (GO:0046907)	373	32	12.81	2.5	8.3E-03
establishment of localization in cell (GO:0051649)	387	33	13.29	2.48	6.7E-03
establishment of protein localization (GO:0045184)	488	41	16.75	2.45	6.4E-04
protein transport (GO:0015031)	482	39	16.55	2.36	3.1E-03
protein localization (GO:0008104)	523	42	17.95	2.34	1.5E-03
organelle organization (GO:0006996)	955	76	32.79	2.32	6.4E-08
cellular localization (GO:0051641)	474	37	16.27	2.27	1.3E-02
macromolecule localization (GO:0033036)	738	56	25.34	2.21	1.3E-04
reproductive system development (GO:0061458)	947	64	32.51	1.97	8.8E-04
reproductive structure development (GO:0048608)	947	64	32.51	1.97	8.8E-04
post-embryonic development (GO:0009791)	1112	74	38.18	1.94	1.9E-04
cellular component organization (GO:0016043)	1790	116	61.45	1.89	1.4E-07
developmental process involved in reproduction (GO:0003006)	1101	70	37.8	1.85	2.2E-03
cellular component organization or biogenesis (GO:0071840)	1991	126	68.35	1.84	8.7E-08
organic substance transport (GO:0071702)	943	59	32.37	1.82	2.5E-02
reproductive process (GO:0022414)	1267	77	43.5	1.77	3.3E-03
single organism reproductive process (GO:0044702)	1138	69	39.07	1.77	1.3E-02
reproduction (GO:0000003)	1273	77	43.7	1.76	3.9E-03
system development (GO:0048731)	1363	81	46.79	1.73	4.1E-03
localization (GO:0051179)	1846	109	63.37	1.72	8.0E-05
establishment of localization (GO:0051234)	1776	104	60.97	1.71	2.5E-04
transport (GO:0006810)	1758	101	60.35	1.67	9.3E-04
single-multicellular organism process (GO:0044707)	2017	110	69.24	1.59	3.2E-03
multicellular organism development (GO:0007275)	1967	107	67.53	1.58	5.0E-03
anatomical structure development (GO:0048856)	2191	115	75.22	1.53	1.0E-02
multicellular organismal process (GO:0032501)	2133	111	73.23	1.52	2.2E-02
developmental process (GO:0032502)	2347	119	80.57	1.48	3.3E-02
cellular process (GO:0009987)	9004	371	309.11	1.2	2.9E-02
Unclassified (UNCLASSIFIED)	7776	291	266.95	1.09	0.0E+00

Finally, to investigate whether the EC-like callus had the potential to serve as a tool in EC membrane proteomics, the overlap between the obtained EC transcriptomic data and the EC-like proteomic data was determined. Here, the putative membrane-associated, and secreted proteins encoded by all EC-DEGs (FC > 0 versus sporophyte), and the EC-like callus-expressed putative membrane-associated or secreted proteins (FC > 0 EC-like callus / CIM callus) of the MF and FP approaches were included into a three way Venn diagram. This revealed an overlap between EC transcriptome and EC-like proteome of 628 corresponding AGIs with assigned putative membrane-associated or secreted proteins (18%; Figure 2-10).

By utilizing the 320 AGIs of the EC-expressed gene / EC-like callus protein intersections and subsequent PANTHER-based statistical overrepresentation analysis, a number of statistically valid overrepresented biological processes were discovered. Again, the most overrepresented biological process was “phagocytosis” (GO:0006909), but apparently also processes that had been found enriched in the female gametophytic cells like “exocytosis” (GO:0006887), and “protein glycosylation” (GO:0006486) were identified (Table 2-7 A).

On the molecular level, especially “SNAP-receptor activity” (GO:0005484) was found overrepresented (Table 2-7 B).

In conclusion, the overlap between putative membrane, and extracellular localized proteins of the EC and the EC-like callus, as determined here, indicated similarities in enriched, and

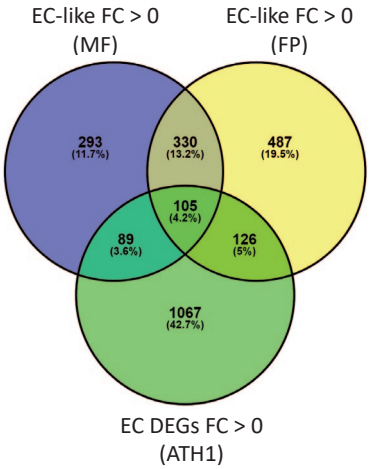


Figure 2-10 Overlap between EC transcriptome and EC-like callus proteome, restricted to putative membrane, or extracellular localized proteins after an FC > 0 cutoff.
This Venn diagramm depicts AGIs corresponding to putative membrane localized, or extracellular proteins only. It combines proteins identified in the MF (817) and the FP (1048) precipitations of the EC-like callus cell line with calculated FC > 0 over CIM callus samples, with 1387 DEGs of the EC (FC > 0 versus sporophyte). The overlap of the EC transcriptome and the EC-like callus proteome in regard to membrane-associated and secreted proteins was determined to 320 proteins (13%). ATH1 = Affymetrix ATH1 GeneChip transcriptome data, FP = full proteome approach, MF = microsomal fractionation, FC = log2 fold change.

Table 2-7 Overrepresented biological processes in EC as well as EC-like cell lines, associated with putative membrane or extracellular localized proteins.
All 320 AGIs found in the intersection between EC expressed and EC-like callus expressed / upregulated proteins with putative membrane, or extracellular localization were subjected to PANTHER-based statistical overrepresentation testing. Obtained results were depicted as GO-slim biological process (**A**) and GO-slim molecular fuction (**B**) sorted by descending fold overrepresentation (calculated by comparison to expected GO terms from *Arabidopsis* whole GO term reference dataset), pV cutoff ≤ 0.05. *ATPase activity coupled to transmembrane movement of compounds.

PANTHER GO-Slim Biological Process	Arabidopsis REF(27352)	EC-like ∩ EC (320)	exp.	fold overrep.	pV
phagocytosis (GO:0006909)	14	4	0.16	24.42	4.6E-03
exocytosis (GO:0006887)	124	9	1.45	6.2	3.5E-03
vesicle-mediated transport (GO:0016192)	516	31	6.04	5.14	3.8E-11
protein localization (GO:0008104)	216	11	2.53	4.35	1.1E-02
protein transport (GO:0015031)	1215	30	14.21	2.11	2.1E-02
localization (GO:0051179)	2044	50	23.91	2.09	1.2E-04
transport (GO:0006810)	1963	48	22.97	2.09	2.1E-04

PANTHER GO-Slim Molecular Function	Arabidopsis REF(27352)	EC-like ∩ EC (320)	exp.	fold overrep.	pV
SNAP receptor activity (GO:0005484)	62	6	0.73	8.27	1.5E-02
ATPase activity* (GO:0042626)	140	9	1.64	5.49	7.1E-03
pyrophosphatase activity (GO:0016462)	478	16	5.59	2.86	2.9E-02
protein binding (GO:0005515)	1286	30	15.05	1.99	4.3E-02
hydrolase activity (GO:0016787)	2326	49	27.21	1.8	6.7E-03

overrepresented biological processes and molecular functions, linked to vesicle transport, and vesicle docking. Furthermore, the EC-like callus proteomic data had revealed, aside early seed-associated processes, also overrepresentation in reproductive processes and thereby qualifies, with certain reservations, for future EC-related investigations.

2-3 Selection of Candidate Genes for Data Validation, and in-depth Functional Studies

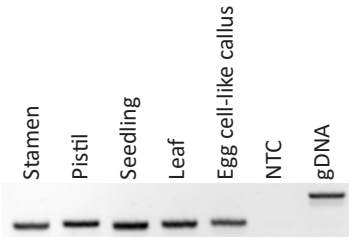
On the molecular level, cell-to-cell signaling in gametophytic cells during gamete recognition and fusion is still incompletely understood. In order to identify new players, preferentially FC 1.5 or higher upregulated DEGs, identified by contrasting to sporophytic tissues, encoding predicted membrane-associated, or extracellular localized proteins were chosen for further analysis and assembled into a list of candidate proteins. This list was composed of proteins highly likely to be involved in signaling, or in mediating adhesion. Furthermore, secreted proteins and unknown proteins were included as well. All genes, including annotations and associated expression values were assembled into a comprehensive heatmap.

Consequently, 57 prime candidate genes and their closest sequence-related family members were recruited, in total 309 genes. To validate the data derived from transcriptomics and bioinformatic analyses, a variety of methods were applied. From the list of prime candidate genes, 20 genes were analyzed by real-time quantitative reverse transcription PCR (qPCR), to determine gene expression strength and pattern in *Arabidopsis.* Furthermore, most promising genes (20) were selected for transcript localization in the female gametophyte, and 10 candidate genes for promoter:reporter studies in stable transgenic *Arabidopsis* plant lines.

2-3-1 Relative Quantification of Gene Expression of selected Candidate Genes

To assess candidate gene expression strength and pattern in different *Arabidopsis* organs and in the EC-like cell line, qPCR was utilized as it is a well established approach to validate microarray-derived expression data on smaller scale (Git *et al.,* 2010). Therefore, pistils of floral stage 12c and stamen of early floral stage 13 were harvested from multiple *Arabidopsis* (ecotype Col-0) plants to serve as exemplary generative tissue samples, while 14 day old sterile grown seedlings, and rosette leaves of bolting plants (both ecotype Col-0) represented vegetative tissue samples. Also, the EC-like cell line (Chapter 2-2) was used to isolate RNA from and subsequently prepare the respective first-strand cDNA templates for qPCR. Genomic DNA-free template was verified for all samples by amplification of *GAPC* using an intron-flanking primer pair and separation by agarose gel electrophoresis (Figure 2-11 A).

A



B

	All	- *5-FCL*	- *FBX2*	- *UBI10*	- *PP2AA3*	- *GAPC*
F	3.200	2.583	2.124	3.196	2.202	3.072
F(crit)	2.866	3.056	3.056	3.056	3.056	3.056
pV	0.035	0.080	0.128	0.044	0.118	0.049

C

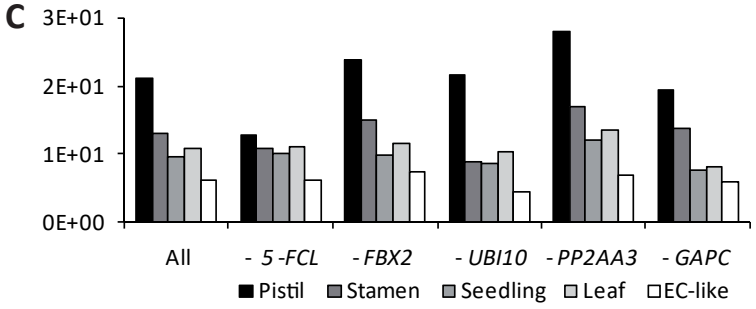
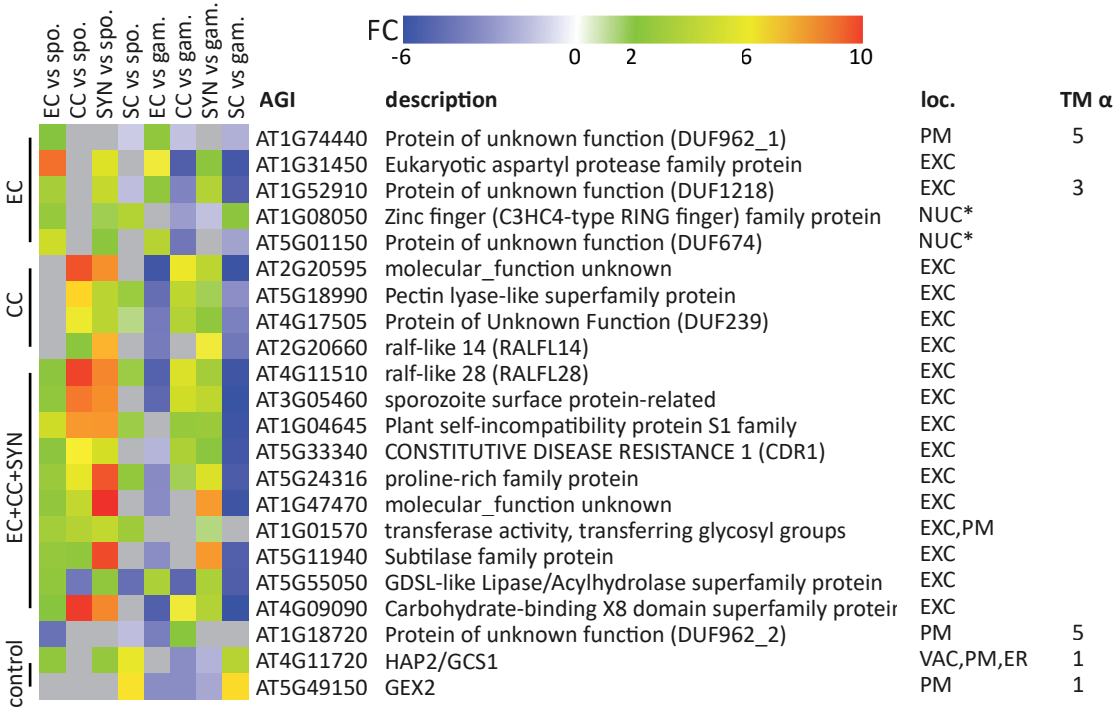


Figure 2-11 Template and reference gene validation for qPCR.
Verification of gDNA-free cDNA template for qPCR by amplification of *GAPC* (**A**). A series of one way ANOVA tests, to determine variance per tissue type in dependence of all tested genes, identified the combination of references genes excluding *5-FCL* (= **-** *5-FCL*) as the gene introducing most variance with F < F(crit) and the most acceptable probability value (**B**). Bar charts displaying the variance per tissue, calculated for all five reference genes genes, and for the individual combinations of only four reference genes (**C**). gDNA = genomic DNA, NTC = non template control, pV = probability value.

Since gene expression strengths of pistil, stamen, leaf, seedling and the EC-like cell line should be compared, five potential reference genes were picked: *FBX2* and *5-FCl* were proposed by the RefGenes tool integrated in Genevestigator (Hruz *et al.,* 2011), *UBI10*, *GAPC*, and *PP2AA3* were chosen as previously identified as valid reference genes (Kudo *et al.,* 2016). Primer pair-dependent PCR amplification efficiencies (E) were determined by a three step dilution of the cDNA template and subsequent calculation via linear regression analysis of obtained quantification cycle (Cq) values (Supplemental Table 5-3). To estimate the reference gene quality for the used samples, which reflected in introduced variance, every reference gene was tested against the remaining four by application of one way ANOVA testing. Equal variance, as determined by F < F(crit), was met by excluding either *5-FCL*, or *FBX2*, or *PP2AA3*. The best pV was obtained for excluding *5-FCL* (pV = 0.08; Figure 2-11 B, C), therefore the geometric mean of *FBX2-*, *UBI10-*, *GAPC-*, and *PP2AA3*-derived Cq values was used for the calculation of normalized relative quantities (NRQ) of the tested candidate genes in the subsequent qPCR experiments.



AGI	description	loc.	TM α
AT1G74440	Protein of unknown function (DUF962_1)	PM	5
AT1G31450	Eukaryotic aspartyl protease family protein	EXC	
AT1G52910	Protein of unknown function (DUF1218)	EXC	3
AT1G08050	Zinc finger (C3HC4-type RING finger) family protein	NUC*	
AT5G01150	Protein of unknown function (DUF674)	NUC*	
AT2G20595	molecular_function unknown	EXC	
AT5G18990	Pectin lyase-like superfamily protein	EXC	
AT4G17505	Protein of Unknown Function (DUF239)	EXC	
AT2G20660	ralf-like 14 (RALFL14)	EXC	
AT4G11510	ralf-like 28 (RALFL28)	EXC	
AT3G05460	sporozoite surface protein-related	EXC	
AT1G04645	Plant self-incompatibility protein S1 family	EXC	
AT5G33340	CONSTITUTIVE DISEASE RESISTANCE 1 (CDR1)	EXC	
AT5G24316	proline-rich family protein	EXC	
AT1G47470	molecular_function unknown	EXC	
AT1G01570	transferase activity, transferring glycosyl groups	EXC,PM	
AT5G11940	Subtilase family protein	EXC	
AT5G55050	GDSL-like Lipase/Acylhydrolase superfamily protein	EXC	
AT4G09090	Carbohydrate-binding X8 domain superfamily protein	EXC	
AT1G18720	Protein of unknown function (DUF962_2)	PM	5
AT4G11720	HAP2/GCS1	VAC,PM,ER	1
AT5G49150	GEX2	PM	1

Figure 2-12 Selected candidate genes and controls for qPCR.
In total, 20 candidate genes were selected: five with preferential EC expression, four with preferential CC expression, and ten with mixed gametophytic expression, also *GCS1/HAP2, GEX2*, and *EC1.1* (not featured by ATH1 GeneChip) were selected as controls for the respective male and female gametophytic tissues. Here, transcriptomic expression data was depicted as heatmap with localization data (SUBAcon), and predicted transmembrane domains (ARAMEMNON v8.1). ER = endoplasmic reticulum, EXC = extracellular, FC = log2 fold change, gam. = gametophyte, loc. = localization, NUC = nucleus, PM = plasma membrane, TM α = alpha helices, VAC = vacuole, spo. = sporophyte,* = previously annotated as membrane localized, meanwhile outdated.

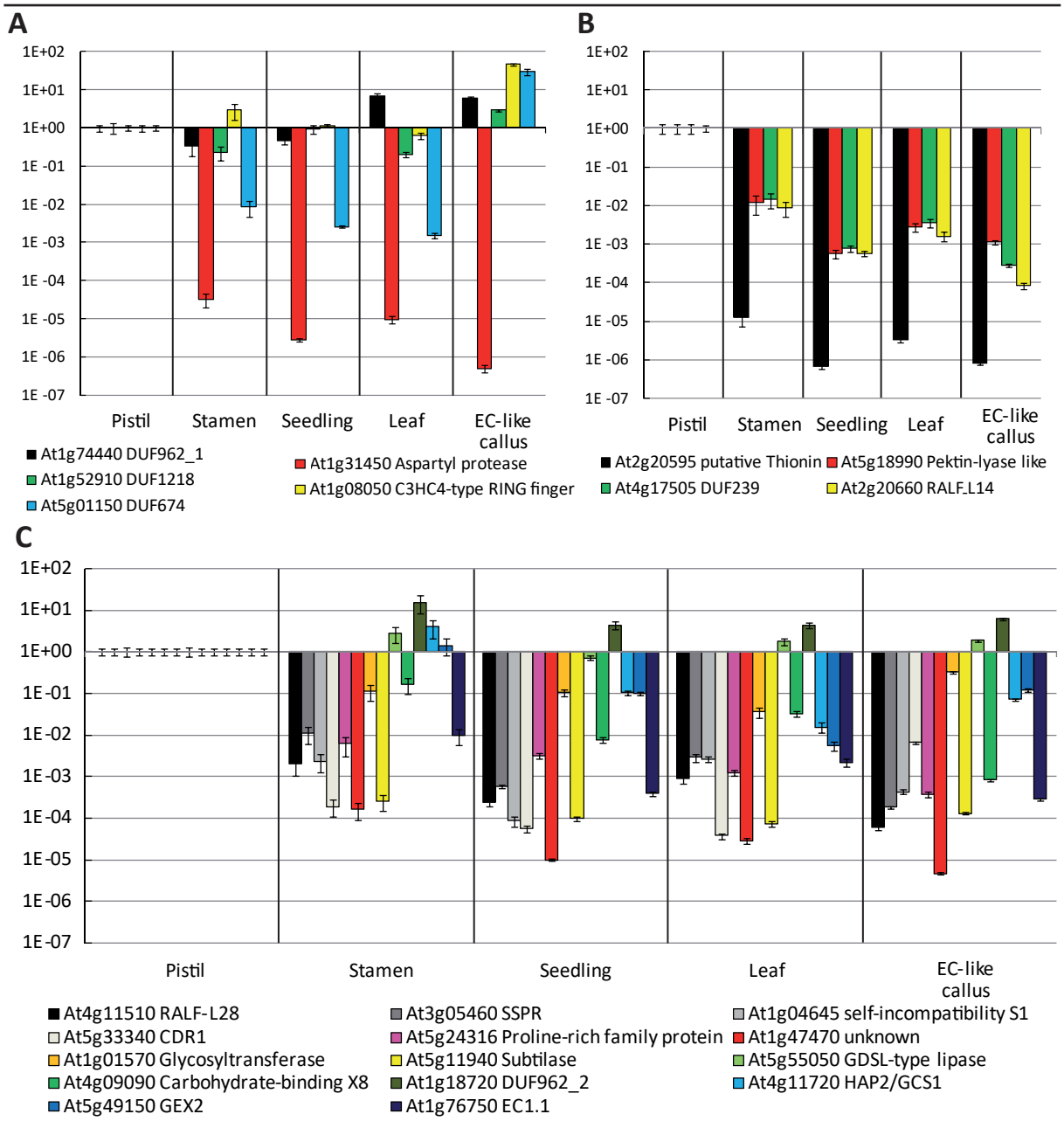


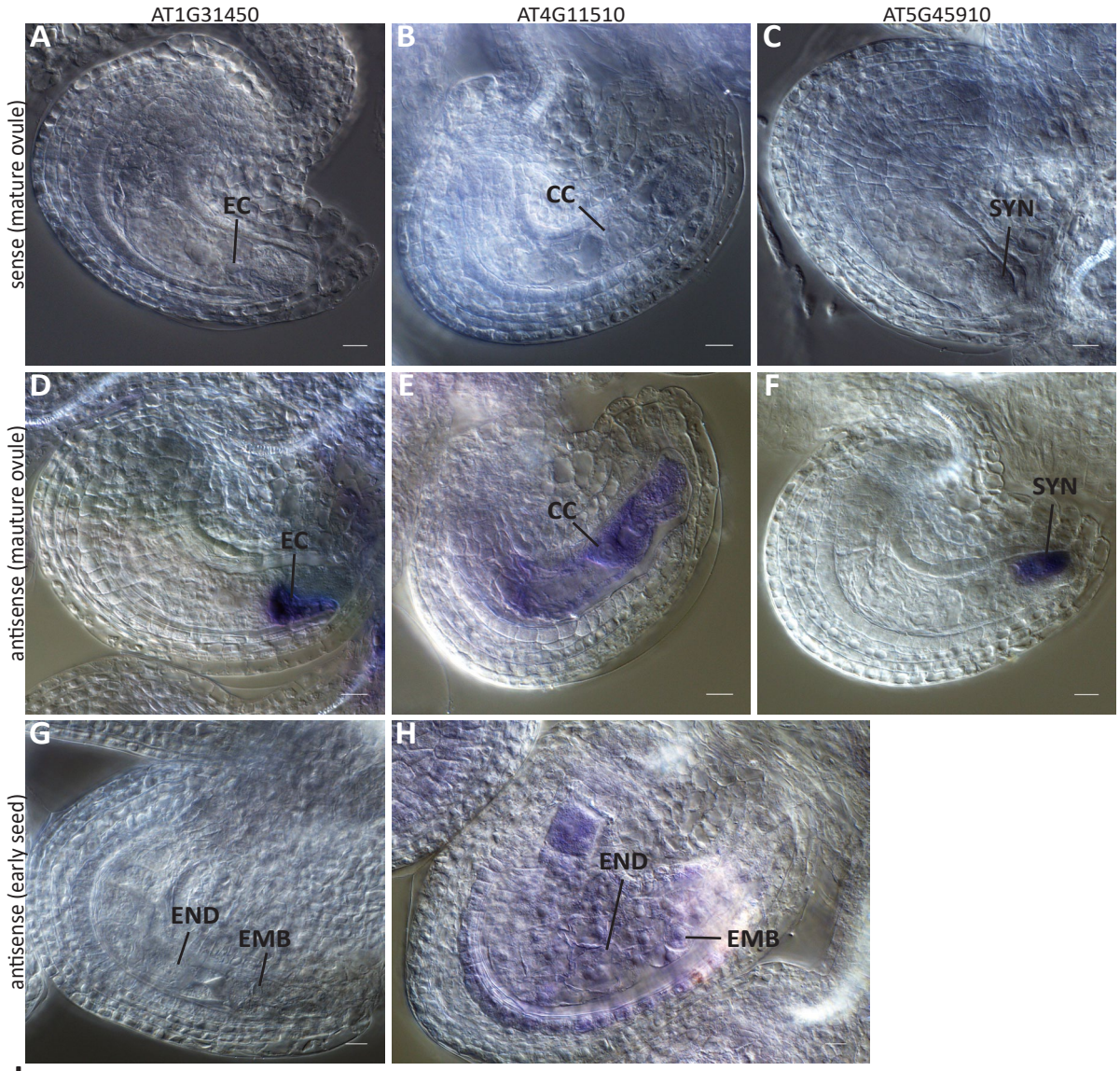
Figure 2-13 Normalized relative quantities of candidate genes per tissue as determined by qPCR.
Normalized relative quantities (relative to the pistil) of selected candidate genes. Candidate genes were either preferential EC-expressed (**A**), CC-expressed (**B**), or expressed by more than one cell of the female gametophyte (**C**)., *GCS1/HAP2*, and *EC1.1* were control genes (**C**). Normalized relative quantities are depicted on a log10 scale. Bars indicate normalized relative standard errors.

Three control genes *GCS1/HAP2* (Johnson *et al.,* 2004, Mori *et al.,* 2006), *GEX2* (Mori *et al.,* 2014), and *EC1.1* (Sprunck *et al.,* 2012) were included as gamete-expressed marker genes for male and female generative tissues (Figure 2-12). The 20 genes of interest for gene expression analysis by qPCR were distributed into three subgroups: Preferentially EC not CC-expressed (i), CC not EC-expressed (ii), and expressed by the female gametophytic cells (iii, Figure 2-12). Three out of five EC-expressed genes were, compared to the stamen, the seedling, and the leaf, found to be enriched in the pistil tissue (AT1G31450, AT5G01150, AT1G52910). Also, with the

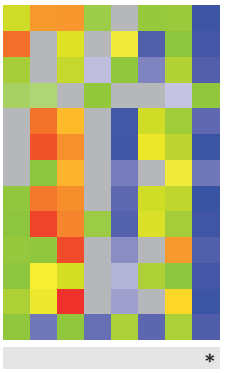
exception of a very strong pistil-expressed putative aspartyl protease (AT1G31450), all genes were found expressed by, and even upregulated in the EC-like callus as well (Figure 2-13 A). The four CC-expressed genes matched the expectations, given by the transcriptomic analysis of gametophytic cells, very well, with all genes being pistil-specific (Figure 2-13 B). Nine out of 11 candidates, which are expressed by more than one female gametophytic cell, were strongest expressed in pistil (Figure 2-13 C). Like expected the SC-expressed genes *GCS1/HAP2* and *GEX2* were determined strongest expressed in stamen, and *EC1.1* was found strongest expressed in pistil (Figure 2-13 C). In summary, qPCR results verified transcripts of 14 candidate genes as pistil-specific (FC > 10^2) and one more candidate gene transcript as pistil-enriched (> FC 10).

2-3-2 *In situ* Detection of Candidate Gene Transcripts in the *Arabidopsis* Ovule

Since the pistil tissue is composed of a variety of cell types in which female gametophytic cells are underrepresented, cellular resolution is required to verify the cell type-specific localization of transcripts. Here, a new adaptation of RNA *in situ* hybridization (Hejatko *et al.,* 2006) was used to validate the acquired transcriptome data of single isolated gametophytic cells. To perform gene-specific whole mount *in situ* hybridizations (WISH), the cDNA sequences of the 20 candidate genes and the cDNAs of the closest gene family members, were obtained from TAIR v10 and aligned by CLC Main Workbench v6. The least conserved section of sequence spanning approximately 200 to 500 bp, or in a few exceptions the full length coding sequence, was amplified by PCR and subsequently cloned into pCR2.1 blunt vector (Supplemental Table 5-4). Furthermore, the already available *EC1.1* CDS in the pCR2.1 blunt vector, provided by Dr. Stefanie Sprunck, was included to serve as positive control for the entire WISH procedure. Consequently, DIG-labeled sense and antisense probes were generated by *in vitro* transcription. The hybridization was performed with *Arabidopsis* (Col-0) pistils of the floral stages 12b/c to 15. Detection of hybridized probes was achieved using anti-DIG-AP antibody and colorimetric staining using BCIP and NBT performed in microtiter plates and stopped after 20 minutes to 4 hours. In total, 21 genes were analyzed by WISH (*EC1.1* included as positive control), of which 14 were successfully detected (Figure 2-14). Among the tested genes, 12 had been determined pistil-specific expressed by qPCR (Chapter 2-3-1). Five probes, of which two genes had been determined pistil-enriched (AT1G74440, AT4G09090), failed to provide results.



I



AGI	observed	descriptions
AT1G04645	SYN	Plant self-incompatibility protein S1 family
AT1G31450	EC	Eukaryotic aspartyl protease family protein
AT1G52910	ovule, ES	Protein of unknown function (DUF1218)
AT1G73200	ovule, ES	Putative integral membrane protein conserved region (DUF2404)
AT2G06090	CC	Plant self-incompatibility protein S1 family
AT2G20595	CC	molecular_function unknown
AT2G20660	SYN	ralf-like 14 (RALFL14)
AT3G05460	SYN, CC, END	sporozoite surface protein-related
AT4G11510	ES, END, EMB	ralf-like 28 (RALFL28)
AT5G11940	SYN	Subtilase family protein
AT5G33340	CC, END	CONSTITUTIVE DISEASE RESISTANCE 1 (CDR1)
AT5G45910	SYN	GDSL-like Lipase/Acylhydrolase superfamily protein
AT5G55050	ovule, ES	GDSL-like Lipase/Acylhydrolase superfamily protein
AT1G76750	EC	EC1.1

Figure 2-14 Successful transcript localizations by WISH in *Arabidopsis* ovules.
Pistils of floral stage 12c - 15 were hybridized with DIG-labeled sense (**A - C**) or antisense probes (**D - H**). AT1G31450 (**A, D, G**) transcript was found in mature EC (**D**) and was absent in the embryo (**G**). AT4G11510 (**B, E, H**) mRNA was detected in the mature embryo sac and was also present within the developing endosperm, and embryo. AT5G45910 (**C, F**) transcripts localized to the SYN (**C**). Heatmap depicting fold change-derived of the transcriptome analysis combined with the WISH-based observations, and functional descriptions (**I**). EC = egg cell, EMB = embryo, END = endosperm, ER = endoplasmic reticulum, ES = embryo sac, EXC = extra cellular, CC = central cell, FC = log2 fold change, SYN = synergid cell. Scale is 10µm. * EC1.1 is not featured by the ATH1 GeneChip.

Three transcripts were localized to the ovule and embryo sac but no cell type in particular, 11 transcripts were specifically detected in the female gametophytic cells and not in the surrounding maternal tissues (Figure 2-14 A – I).

When comparing the results of the single cell transcriptome analysis with the WISH-based detection of candidate gene transcripts, in seven out of 11 cases, deviations in localization were identified. For example AT1G31450 was found EC-specific in WISH experiments but did not show the expected expression in the SYN (Figure 2-14 A, D, G, I). Also, AT5G45910 was found SYN-specific instead of being expressed in all three cell types of the FG (Figure 2-14 C, F, I). These findings suggested either an initially too low defined threshold during bioinformatic processing or possible cross-contaminations within the transcriptomic data of isolated gametophytic cells. Post-fertilization WISH studies were not conducted on genes that had been found SYN-specific (AT1G04645, AT2G20660, AT5G11940, AT5G45910). Three out of six gamete-expressed genes, like AT4G11540, were also detected in early seed developmental stages (Figure 2-14 B, E, H). The combination of qPCR (Chapter 2-3-1) and WISH determined AT1G31450 truly EC-specific, AT2G20595 CC-specific, AT2G20660, AT1G04645, and AT5G11940 SYN-specific. While AT3G05460 was found exclusively expressed in the SYN, CC, and developing endosperm, AT5G33340 gene expression was only detected in the CC and the developing endosperm.

2-3-3 Verification of the Promoter Activities of Selected Genes in *Arabidopsis* Ovules

Besides direct detection of transcripts via WISH, stable plant lines expressing promoter:reporter fusions offer an alternative for verification of cell type-specific promoter acitivity if poor probe performance or probe cross-hybridizations due to highly identical transcripts cause problems. Here, putative promoter sequences of ten candidate genes were cloned upstream of the nuclear localized reporter NLS-GFP or mCherry-H2B (Table 2-8).

DUF674_1 promoter activity was confirmed to be EC-specific (Figure 2-15 A, B). Two DUF674_1 sequence-related genes were not featured by the ATH1 GeneChip, but investigated for promoter activity nevertheless. While *DUF674_2* promoter activity could not be detected, *DUF674_3* promoter activity was found antipodal cell-specific (Figure 2-15 C, D). Furthermore,

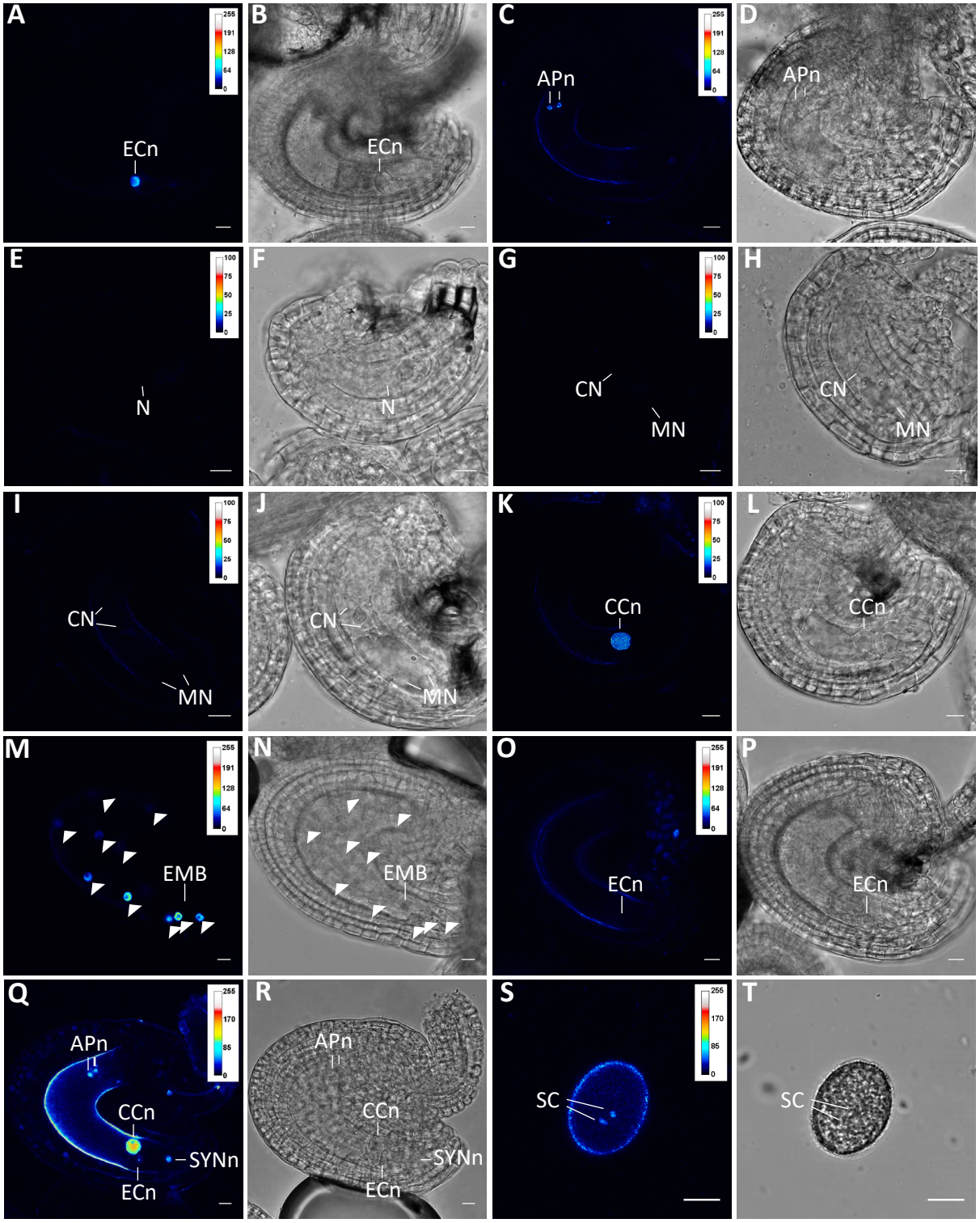


Figure 2-15 Verification of cell type-specific expression by promoter:reporter studies in *Arabidopsis* ovules.
For ten candidate genes stable transgenic *Arabidopsis* lines expressing an endogenous promoter:NLS-GFP or promoter:mCherry-H2B, construct were generated. *DUF674_1p*:NLS-GFP was found to be EC-specifically expressed (**A, B**), and *DUF674_3p*:mCherry-H2B was determined antipodal cell-specific (**C, D**). *DUF962_2p*:mCherry-H2B was not detected in FG1 (**E, F**) FG3 (**G, H**), FG4 (**I, J**), but in the CC (**K, L**), and the endosperm (**M, N**). *DUF962_1p*:mCherry-H2B could not be detected in the EC (**O, P**). *DUF962_3p*:mCherry-H2B was found expressed in all cells of the female gametophyte (**Q, R**) and the SCs (**S, T**). Triangles indicate endosperm nuclei. APn = antipodal cell nucleus, CCn = central cell nucleus, CN = chalazal nucleus, ECn = egg cell nucleus, EMB = embryo, MN = micropylar nucleus, N = functional megaspore, SC = sperm cell, SYNn = synergid cell nucleus. Scale is 10µm.

Table 2-8 Generated stable *Arabidopsis* promoter:reporter lines and detected promoter activities in *Arabidopsis* ovules.
Estimated promoter sequences of the respective genes were amplified and cloned upstream of a nuclear localized fluorescent reporter protein (NLS-GFP_{3x} / mCherry-H2B). For only five out of ten selected genes, promoter activities were detected. AP = antipodal cell, EC = egg cell, END = endosperm, CC = central cell, FG = female gametophyte developmental stage, SC = sperm cell, SYN = synergid cell.

AGI	name	included bp upstream of the start codon	oligo pair	expression (cell type)
AT1G61550	S-locus lectin RLK	825	104/105	no signal
AT1G61550	S-locus lectin RLK	870	242/285	no signal
AT1G74440	DUF962_1	1987	106/107	no signal
AT1G18720	DUF962_2	868	270/271	CC, END
AT3G09085	DUF962_3	761	264/265	CC, EC, SYN, AP, SC, END
AT5G01150	DUF674_1	551	108/109	EC
AT5G01130	DUF674_2	496	238/239	no signal
AT5G01140	DUF674_3	382	240/241	AP
AT3G05460	SSPR	985	550/571	FG2-FG5, CC, (EC), SYN, AP, END
AT1G08050	putative E3 Ligase	661	061/062	no signal
AT5G67550	unknown protein	1044	075/076	no signal

DUF962_2p:mCherry-H2B expression was monitored throughout the individual stages of female gametophyte development were it was found absent of stage FG1, FG3, and FG4 (Figure 2-15 E – J). However, in mature ovules, DUF962_2 promoter expression was specifically detected in the CC, and persisted during endosperm development but was found absent in the embryo (Figure 2-15 K - N). *DUF962_1* had been determined EC-expressed by the transcriptomic profiling of isolated female gametophytic cells but DUF962_1p:mCherry-H2B could not be detected, though (Figure 2-15 O, P). Moreover, *DUF962_3p:mCherry-H2B* expression was detected in all cells of the mature FG, and additionally in SCs (Figure 2-15 Q – T). The data of the *SSPR* promoter activities will be presented in Chapter 2-5. In total, the findings of the promoter:reporter studies in *Arabidopsis* ovules did not qualify for ongoing efforts since for only 50% of the selected genes promoter activities could be detected (Table 2-8).

2-4 Functional Studies of the Selected Genes: Overview

Based upon previous findings and the initial transcriptome studies, a subset of 14 genes were selected and the subcellular localization of these candidate genes was investigated using gene:reporter fusions, expressed by transgenic *Arabidopsis* lines. For ten candidate genes the corresponding fusion proteins were detected within gametophytic cells (Table 2-9). To gain insights on putative protein functions in the female gametophyte, T-DNA mediated knock outs were pursued for 14 genes. Out of 16 investigated T-DNA insertion lines only five lines were found depleted in transcript of the respective gene (Table 2-10). However, none of the T-DNA lines analyzed in this study showed a reduced seed set. Furthermore, the CRISPR/Cas9 system was applied to effectively mutagenize SSPR (Chapter 2-5), *TE7*, *TET9*, and *TET12* (Chapter 2-7).

Table 2-9 Selected candidate genes for promoter:gene:reporter studies in *Arabidopsis*.
Selected candidate genes, or gene family members, name, used promoter length in base pairs (bp) upstream of the start codon, the oligo pairs used for amplification and cloning of promoters, and/or of the CDS/gene sequences are shown. Vector backbones for plant expression and expression patterns are indicated. EC = egg cell, END = endosperm, CC = central cell, COT = cotelydons, PT = pollen tube, SYN = synergid cell, ZY = zygote. The reporter line for TETRASPANIN9 was provided by Dr. Stefanie Sprunck.

AGI	name	promoter size [bp]	oligo pair	Vector backbone	sequence	expression (cell/tissue type)
AT1G61550	S-locus lectin RLK	870	104/193	pGWB550	genomic:GFP	no signal
AT1G74440	DUF962_1	1987	266/267 + 116/117	pB7FWG2	promoter:CDS:GFP	no signal
AT1G74440	DUF962_1	1987	106/117	pGWB550	genomic:GFP	EC, (ZY), leaf, root
AT1G18720	DUF962_2	868	268/269 + 272/273	pB7FWG2	promoter:CDS:GFP	CC, END, leaf, COT
AT1G18720	DUF962_2	868	270/271 + 272/299	pGWB550	genomic:GFP	CC, END, leaf, COT
AT3G09085	DUF962_3	761	264/318	pGWB550	genomic:GFP	no signal
AT5G01150	DUF674_1	551	214/215 + 114/115	pB7FWG2	promoter:CDS:GFP	no signal
AT3G05460	SSPR	985	572/573	pCambia2300	genomic:GFP	SYN, CC, END
AT1G08050	putative E3 Ligase	661	458/459	pGWB550	genomic:GFP	EC
AT5G67550	unknown protein	1044	460/461	pGWB550	genomic:GFP	no signal
AT4G30430	TETRASPANIN9	1571	-	pB7FWG2	promoter:CDS:GFP	EC, CC, root
AT2G23810	TETRASPANIN8	934	164/165	pGWB550	genomic:GFP	EC, leaf
AT4G28050	TETRASPANIN7	250	166/167	pGWB550	genomic:GFP	SYN
AT5G27830	putative folate receptor-like	367	401/400	pGWB550	genomic:GFP	PT, SYN
AT4G17505	DUF239	655	632/633	pCambia2300	genomic:GFP	CC
AT4G09090	beta 1,3 endoglucosidase	701	303/302	pGWB550	genomic:GFP	CC

Table 2-10 Analysis of T-DNA insertion lines of selected candidate genes.
T-DNA insertion positions were mapped by sequencing and, in homozygous individuals, the transcript abundance of the respective genes was investigated by RT-PCR. The seed set of homozygous individuals was compared to the wild type. The GABI_278_H08 insertion was determined absent of the UTR of *TET9* as previously annotated. ho = homozygous, k.o. = knock out, NA = not available, ND = not determined, WT = wild type.

AGI	name	T-DNA insertion	position [bp]	genotype	comment	seed set
AT1G61550	S-locus lectin RLK	SAIL_63_G01	NA	WT		
AT1G74440	DUF962_1	GABI_179_G08	-45 (UTR)	ho	no k.o.	WT-like
AT1G74440	DUF962_1	SALK_059087	543 (Exon)	ho	no k.o.	WT-like
AT1G74440	DUF962_1	SALK_118340	1242 (UTR)	ho	no k.o.	WT-like
AT5G01150	DUF674_1	SALK_025213	400 (Exon)	ho	k.o.	WT-like
AT5G01130	DUF674_2	SALK_122879	NA (Exon)	ho	k.o.	WT-like
AT5G01140	DUF674_3	SALK_055941	953 (Exon)	ho	k.o.	WT-like
AT3G05460	SSPR	SALK_009445	-262	ho	no k.o.	WT-like
AT1G08050	putative E3 Ligase	SAIL_1244_B08	1529 (Exon)	ho	no k.o.	WT-like
AT5G67550	unknown protein	SALK_052167	370 (Exon)	ho	no k.o.	WT-like
AT2G23810	TETRASPANIN8	SALK_136039	402 (Exon)	ho	k.o.	WT-like
AT3G30430	TETRASPANIN9	GABI_278_H08	-			
AT1G18520	TETRASPANIN11	SALK_109259	454 (Exon)	ho	k.o.	WT-like
AT4G09090	beta 1,3 endoglucosidase	SAIL_317_C07	203(Exon)	ho	no k.o.	WT-like
AT1G16180	putative serine incorporator	SALK_084244	260 (Exon)	ho	no k.o.	WT-like
AT5G16900	putative LRR-type RLK	SALK_113523	ND (Exon)	ho	no k.o.	WT-like

In the following, three protein families will be described in more detail and the efforts that were made to shed light on their function in female gametes.

2-5 DEGs Encoding Putative Embryo Sac-secreted Proteins

Cell-to-cell mobile signals are an essential component of intercellular communication. The initial transcriptomic analysis of female gametophytic cells and male gametes had identified 264 DEGs upregulated (FC ≥ 1.5) over sporophytic tissues encoding for putative extracellular localized proteins. A subset of three candidate gene families was considered for further characterization (Figure 2-16).

Selected was a DUF239-containing protein of unknown function, of which seven more family members were found upregulated in female gametophytic cells and two more in SCs (Figure 2-16 A asterisks). The AT4G17505-encoded DUF239 was registered as PF03080 PFAM v31, which was predominantly found in eukaryota with 71 recorded species and 1817 sequences, while only 15 species with 18 recorded sequences corresponded to the bacteria superkingdom. The majority of eukaryotic PF03080-containing sequences belonged to plant orders like poales (505), or brassicales (358) whereas the kingdom of fungi contained 48 sequences in 16 species. Furthermore, a protein with a Carbohydrate-binding X8 domain, according to PFAM v31 categorized as PF07983, was selected (AT4G09090). The two by far most dominant protein architectures of X8 domain-encoding genes are in 1637 sequences the combination of the X8-domain with an additional glycosyl hydrolase family 17 domain (PF00332), or, like AT4G09090,

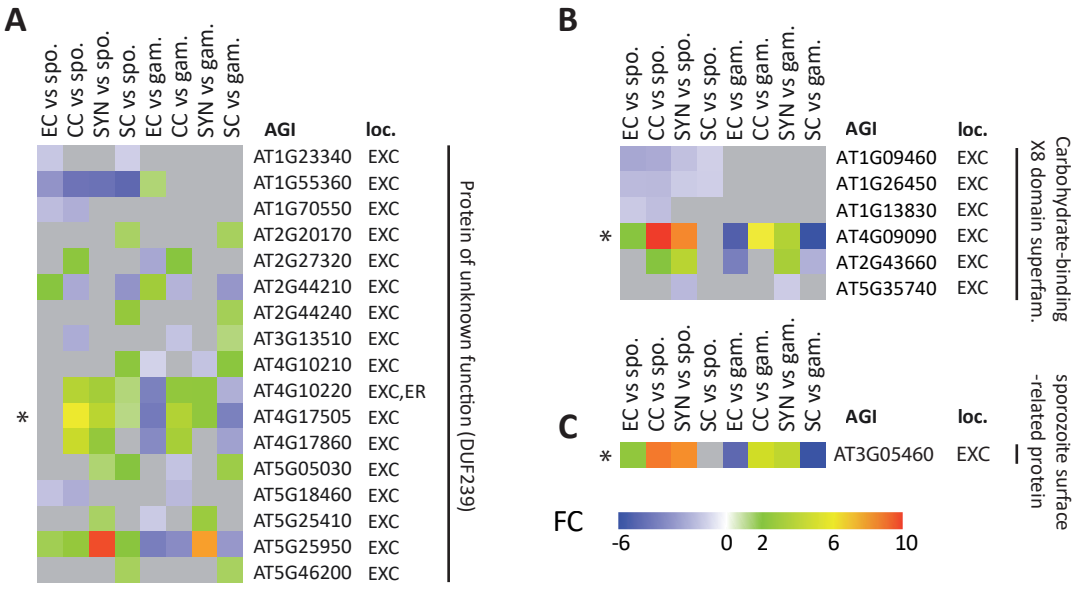


Figure 2-16 Expression of five gene families encoding putative embryo sac secreted proteins. Calculated fold change of DEG-families encoding putative extracellular-localized proteins according to SUBA4. Functional descriptions are summed up for unknown proteins, like DUF239-containing proteins (**A**), Carbohydrate-bindin X8 domain containing proteins (**B**), and the sporozoite surface-related protein (**C**). Asterisks indicate candidate genes. EXC = extracellular, ER = endoplasmic reticulum, FC = log2 fold change, gam. = gametophyte, loc. = localization, spo. = sporophyte.

1362 sequences without any additional functional domains. In *Arabidopsis,* 49 X8-domain containing proteins were known, of which two were found upregulated in female gametophytic cells (Figure 2-16 B asterisks). The so far uncharacterized, *SPOROZOITE SURFACE PROTEIN-RELATED* (*SSPR*), which is a single copy gene with the strongest expression in the CC (Figure 2-16 C), was selected for functional studies (Chapter 2-5-2).

2-5-1 DUF239, and X8 domain-containing Proteins are Secreted by the CC

To address the subcellular localization of candidate proteins encoded by AT4G17505, and AT4G09090 in the cells of the FG, transgenic reporter plant lines expressing translational GFP fusions were generated. The respective genes were cloned into pCambia2300 and pGWB550, including their putative promoter sequences of 655bp and 701bp, respectively. AT4G17505g:GFP was not detected in developing ovules of stage 3-III and stage stage 3-IV (not shown), but found in the mature FG within the CC, and between the CC-surrounding cell types like the inner integuments (Figure 2-17 A, B). After fertilization, no GFP signals were detected (not shown).

Analysis of AT4G09090g:GFP in mature ovules revealed that this X8 domain-containing protein localized extracellularly in the intercellular space surrounding the embryo sac (Figure 2-17 C,

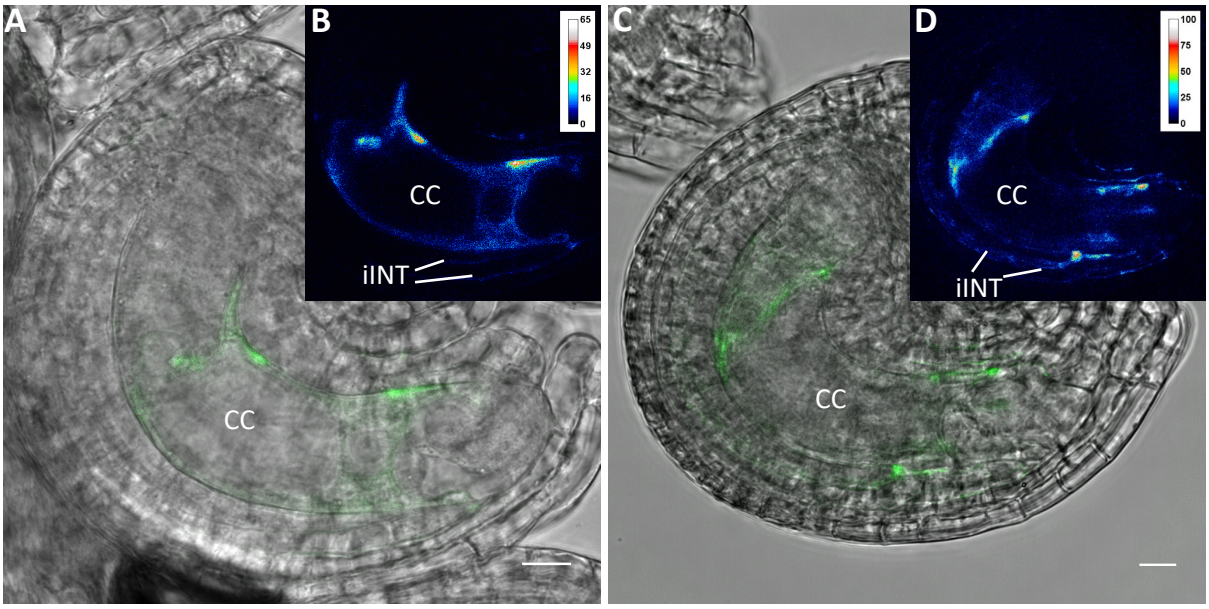


Figure 2-17 Two central cell-secreted proteins in *Arabidopsis* ovules.
AT4G17505g:GFP, encoding for a DUF239-containing protein was detected in the CC, and outside of the mature embryo sac reaching into intercellular space of the inner integument layers (**A, B**). *AT4G09090g:GFP*, encoding a Carbohydrate-binding X8 domain protein, was also detected in the intercellular space surrounding the mature embryo sac (**C, D**). CC = central cell, iINT = inner integument. Scale is 10µm.

D). However, six obtained homozygous plants of the T-DNA insertion line SAIL_317_C07, a T-DNA insertion, which was mapped to the first exon 203 bp downstream of the start codon of AT4G09090, was seed set-wise not significantly different from WT and thereby not pursued any further.

2-5-2 SSPR is an Ovule-specific Secreted Protein

AT3G05460, termed DD27 (Steffen *et al.,* 2007), or *SPOROZOITE SURFACE PROTEIN-RELATED* (*SSPR*), had, despite being an ovule-specific single copy gene in *Arabidopsis*, not been functionally characterized so far. Expression of *SSPR* was determined by promoter:reporter and gene:reporter plant lines utilizing a putative promoter sequence of 985bp upstream of the respective start codon, and by WISH.

CLSM analysis of floral stage 12a/b to stage 14 pistils of SSPRp:mCherry-H2b plant lines revealed stage FG2 to be the start of detectable *SSPR* promoter activity, which grew stronger in mature ovules were all cells of the FG expressed the SSPRp:mCherry-H2b reporter protein (Figure 2-18 A – D). The SSPR promoter activity persisted post fertilization into stage 4-V ovules (Figure 2-18 E, F). WISH of stage 3-VI and 4-V ovules revealed the highest *SSPR* mRNA abundance in

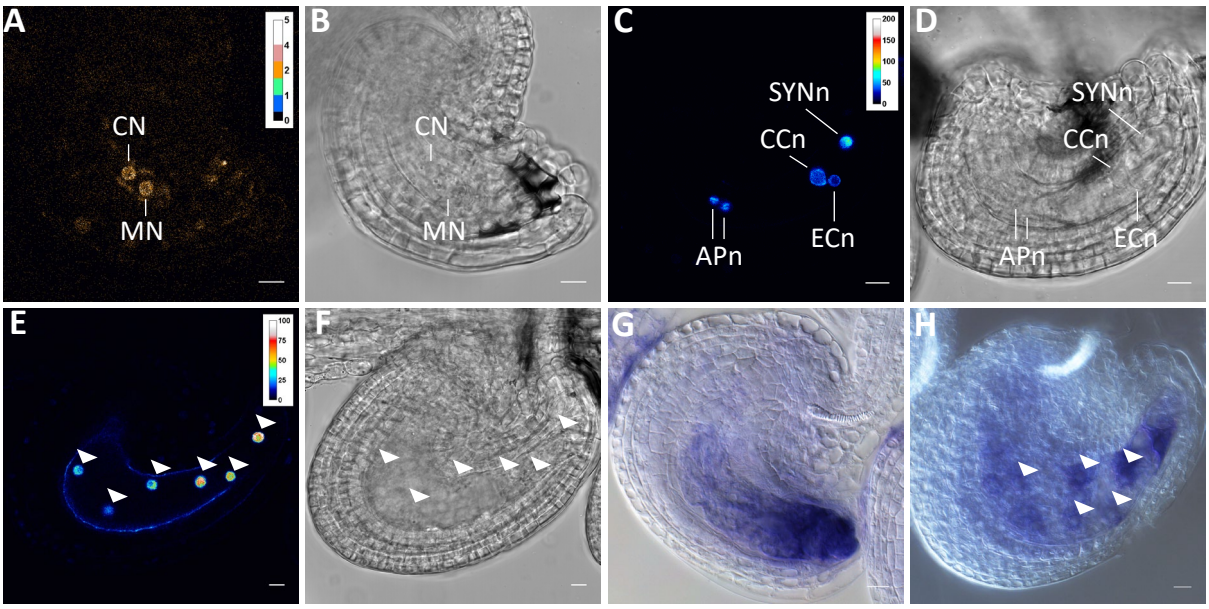


Figure 2-18 *SSPR* promoter activities and detected transripts in *Arabidopsis* ovules.
First *SSPRp:mCherry-H2B* expression was detected in FG2 ovules (**A, B**). In the mature female gametophyte the promoter activity appreared strongest (**C, D**). After fertilization, in stage 4-V ovules strong SSPR promoter activity was detectable in the syncytical endosperm (**E, F**).
Most *SSPR* transcript was detected at the micropylar pole of the mature ovule (**G**), while in stage 4-V ovules most transcript was localized close to the endosperm nuclei. Triangles indicate endosperm nuclei. APn = antipodal cell nucleus, CCn = central cell nucleus, CN = chalazal nucleus, ECn = egg cell nucleus, MN = micropylar nucleus, SYNn = synergid cell nucleus. Scale is 10µm.

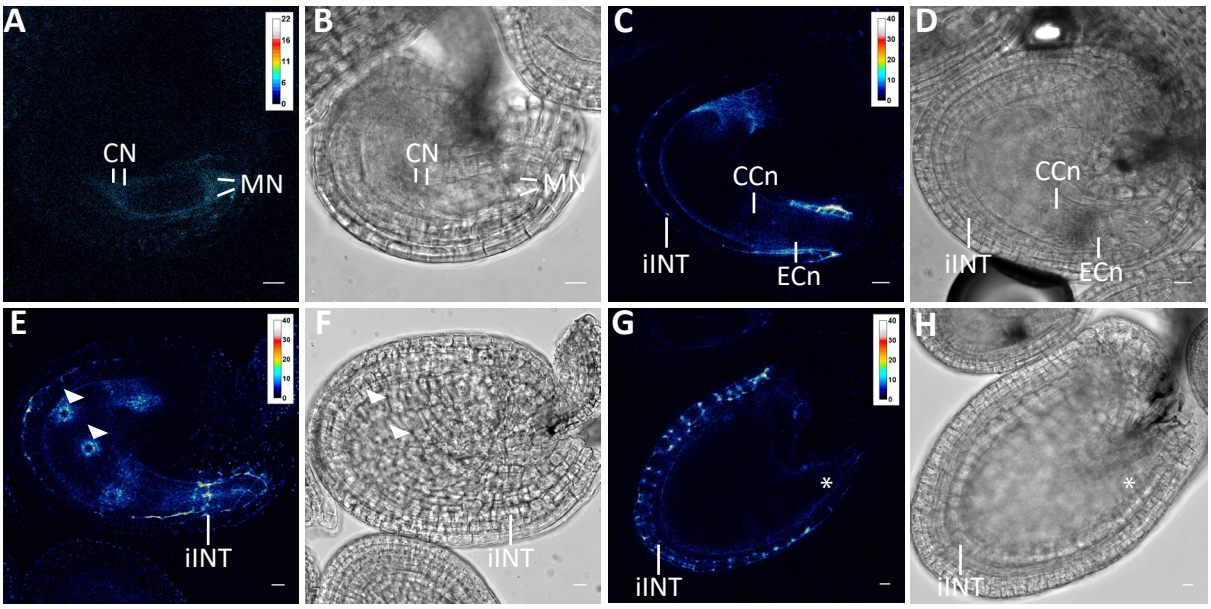


Figure 2-19 Detection of SSPRg:GFP in *Arabidopsis* ovules.
SSPRg:GFP was not detected before stage FG4, where it was only detectable in the endomembrane system of the developing embryo sac (**A, B**). In mature ovules, the fluorescence-derived signal was stronger and preferentially localized to the intercellular space surrounding the embryo sac, including the inner integuments (**C, D**). In stage 4-V ovules, the fusion protein was still detectable in the endomembrane system of the syncytical endosperm, and between the cells of inner integuments (**E, F**), while in stage 4-VI ovules, SSPRg:GFP appeared to be restricted to the inner integuments and absent in the endosperm. The triangles indicate the endosperm nuclei. The asterisks indicates the embryo. ECn = egg cell nucleus, CCn = central cell nucleus, CN = chalazal nucleus, iINT = inner integuments, MN = micropylar nucleus. Scale is 10µm.

the egg apparatus of mature ovules and persisting transcript levels post fertilization (Figure 2-18 G, H). By CLSM-based analysis of floral stage 12b to 15 pistils of *SSPRg:GFP*-expressing plant lines, the GFP-derived fluorescence was non detectable in stage 3-II ovules (not shown) but appeared within the developing embryo sac during late stage FG4 (Figure 2-19 A, B). In mature ovules, the most SSPRg:GFP fusion protein was detected in the micropylar region of the ovule in the intercellular space of the cells surrounding the embryo sac (Figure 2-19 C, D). Post fertilization, in stage 4-V ovules, SSPRg:GFP was observed in the syncytial endosperm as could be seen by most ER-derived fluorescent signals surrounding the endosperm nuclei. Furthermore, SSPRg:GFP could be detected between the layers of the inner integuments (Figure 2-19 E, F). In later seed developmental stages approximately stage 4-VI with a clearly visible early embryo, SSPRg:GFP was absent in the embryo and endosperm. However, the GFP-derived fluorescence appeared between the cells of the inner integuments and between the inner integument layers in a slightly patchy manner (Figure 2-19 G, H).

2-5-3 The *SSPR* Knock Out arrested during FG Development

To elucide a possible function of SSPR during ovule and seed development, a reverse genetics approach was pursued utilizing two available T-DNA insertion lines (SALK_009445, SAIL_869_E01). By PCR-based genotyping all plants of the SAIL_869_E01 T-DNA line were determined WT. However, among the obtained SALK_009445 lines, two homozygous plants were identified. The SALK_009445 insertion was mapped by sequencing and determined to be 262bp upstream of the start codon of *SSPR* (Figure 2-20 A). To address *SSPR* transcript abundance in these plant lines, mRNA of floral stage 12b/c pistils of two homozygous SALK_009445 plants, one genotypic WT SAIL_869_E01 plant, and a true WT was isolated and used for cDNA synthesis. The RT-PCR with gene-specific amplification of *SSPR*, and three template control genes (*GAPC*, *ACT2*, and *UBI10*) determined the homozygous SALK_009445 insertion line not to be depleted in *SSPR* transcript (Figure 2-20 B). Since RT-PCRs indicated one homozygous SALK_009445 plant line to have less *SSPR* transcript, WISH was performed with *SSPR* gene-specific antisense and sense probes of floral stage 12b/c pistils of this SALK_009445 plant, but this clearly demonstrated

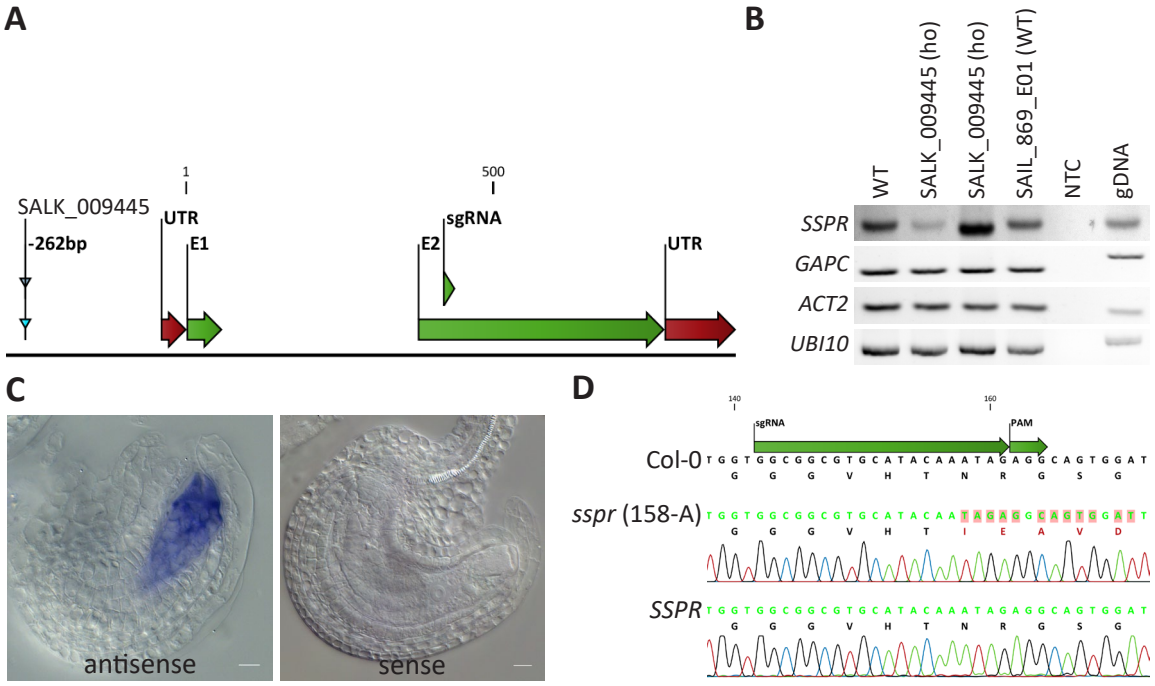


Figure 2-20 Analysis of *SSPR* T-DNA and established CRISPR/Cas9 lines.
The genomic locus of *SSPR*, with the verified position of the T-DNA insertion SALK_009445, and the sgRNA binding site for CRISPR/Cas9 (**A**). RT-PCR analysis of two homozygous SALK_009445 plants, a segregated WT of the T-DNA line SAIL_869_E01, and the WT, with *ACT2, GAPC,* and *UBI10* serving as template controls (**B**). WISH using *SSPR* gene-specific antisense and sense probes identified significant mRNA levels of *SSPR* in SALK_009445 ovules of floral stage 12b/c pistils. CRISPR/Cas9-mediated gene editing occured in one out of 17 T1 plant lines, generating one homozygous *sspr* mutant (158-A) with a frame shift caused by a deletion after position 158 (in CDS) as can seen by this sequencing-derived chromatogram. E = exon, gDNA = genomic DNA, NTC = non template control, UTR = untranslated region. Scale is 10µm.

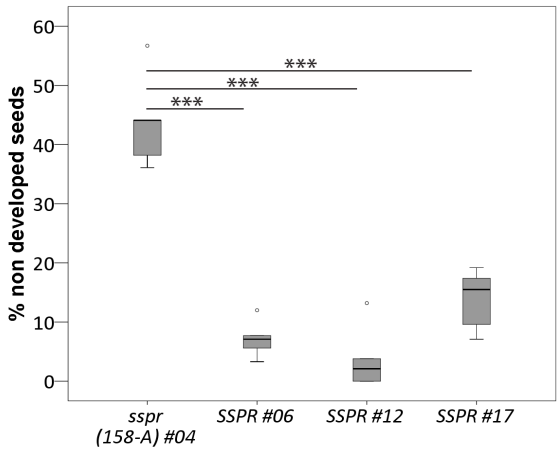


Figure 2-21 Seed sets of *SSPR* CRISPR/Cas9 plants.
Relative non developed seeds of four independent T1 plant lines, obtained from transformation of WT with a *SSPR*-targeting CRISPR/Cas9 construct. The *sspr* (158-A) plant #04 was found homozygous and seed set-wise statistically signifcant different from all three control plant lines, as determined by two-tailed students t-test.

abundant *SSPR* transcript (Figure 2-20 C). Since the T-DNA lines were proven unsuitable for any functional studies regarding *SSPR*. Instead, a cloned CRISPR/Cas9 construct targeting *SSPR* was used to generate stable transgenic *Arabidopsis* lines. In total, 17 transgenic generation 1 (T1) plant lines were analyzed and one plant was found mutated in the target locus. In the CDS of *SSPR*, a homozygous mutation was detected after nucleotide position 158 where a single adenosine was found deleted which led to a shift in the translational frame (Figure 2-20 D). Consequently, the 152 amino acid long protein was altered in the CRISPR/Cas9 line *sspr* (158-A) #04 from the 38th position on and truncated after amino acid position 58.

The obtained *sspr* mutant plant line #04, and three control plant lines (#06, #12, #17) with unaltered *SSPR* gDNA locus were analyzed regarding their seed set. By two-tailed students t-test the rate of undeveloped seeds of the mutant plant was found statistically significant different from all three control plant lines (Figure 2-21). The median values of undeveloped seeds were determined to be 44% for the homozygous *sspr* mutant plant line #04, and 7%, 2% and 16% for the three control lines #06, #12, and #17, respectively (Figure 2-21).

Clearing of floral stage 12c pistils and analysis of the FG development by DIC microscopy of the homozygous mutant plant line and control lines #06, and #17, revealed, in *sspr* mutant pistils, three predominant developmental stages of the FG, stage FG1, FG4 and FG6 (Figure 2-22 A - C). The ovules in floral stage 12c pistils of the homozygous *sspr* mutant line appeared, judging from the developmental stage of the sporophytic tissues, like stage 3-VI ovules accumulated

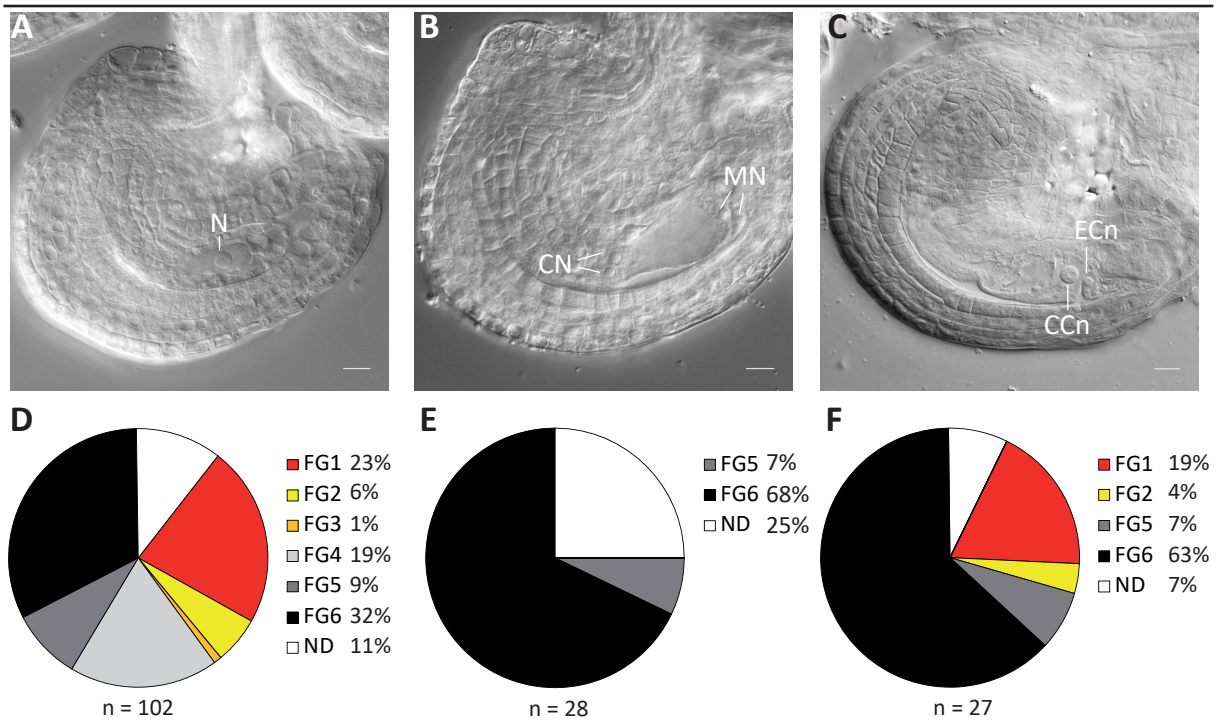


Figure 2-22 Female gametophyte developmental stages in floral stage 12c of CRISPR/Cas9 plant lines.
Clearing of floral stage 12c pistils revealed three major female gametophyte development stages in the *sspr* mutant CRISPR/Cas9 line (158-A) #04 with stage FG1 (**A**), FG4 (**B**), FG6 (**C**) while the respective ovules appeared morphologically mature. In the mutant plant line #4 29% of the female gametophytes arrested in stage FG1 and FG2, and 19% in stage FG3 and FG4 (**D**). In the control plant line #06 most female gametophytes had reached mature stage FG6 (**E**), while in control plant line #17 23% of the female gametophytes were delayed in stage FG1 and FG2 (**F**). CCn = central cell nucleus, CN = chalazal nucleus, ECn = egg cell nucleus, N = functional megaspore, n = numer of ovules, ND = not determined. Scale is 10µm.

arrested FGs in stage FG1 and FG2 (29%), and 20% in stage FG3 and FG4 (Figure 2-22 D). In the control plant line #06 the vast majority (68%) of analyzed FGs had reached maturity (FG6) (Figure 2-22 E). Ovule clearing of the control plant line #17 on the other hand also revealed 23% of the FG arrested in stage FG1 and FG2 while 63% had reached maturity (Figure 2-22 F). In summary, silique and FG analysis demonstrated that the homozygous CRISPR/Cas9 *sspr* mutant plant line #04 exhibited a significant amount of non developed seeds (44%) and a corresponding amount of developmentally arrested FGs in stage FG1 to FG4 (49%).

2-6 DEGs encoding Membrane-localized Proteins

Almost nothing is known about membrane-localized proteins of the female gametes. The transcriptomic approach identified 20 differentially regulated cysteine-rich RLKs, 11 differentially regulated S-domain-type RLKs (Supplemental Figure 5-1), and 104 differentially regulated LRR-RLKs (Supplemental Figure 5-2). After rather unsuccessful pilot experiments with

a small subset of selected RLKs, the focus of this work shifted towards the characterization of a yet unknown protein family with a domain of unkown function (DUF) 962 and the tetraspanin protein family. Both families pose integral membrane proteins with multiple transmembrane domains and are expressed in the female gametes.

2-6-1 The DUF962-containing Gene Family: Phylogeny

The *DUF962_1* gene, AT1G74440, was determined as EC-expressed and upregulated over the sporophyte. In order to determine closely *DUF962_1*-related genes, a database-facilitated search for putative family members and the following reconstruction of phylogenic relationships between the individual genes was conducted. To elucidate the unknown DUF962_1 protein function and a possible role in the EC, the expression was verified on protein level and a reverse

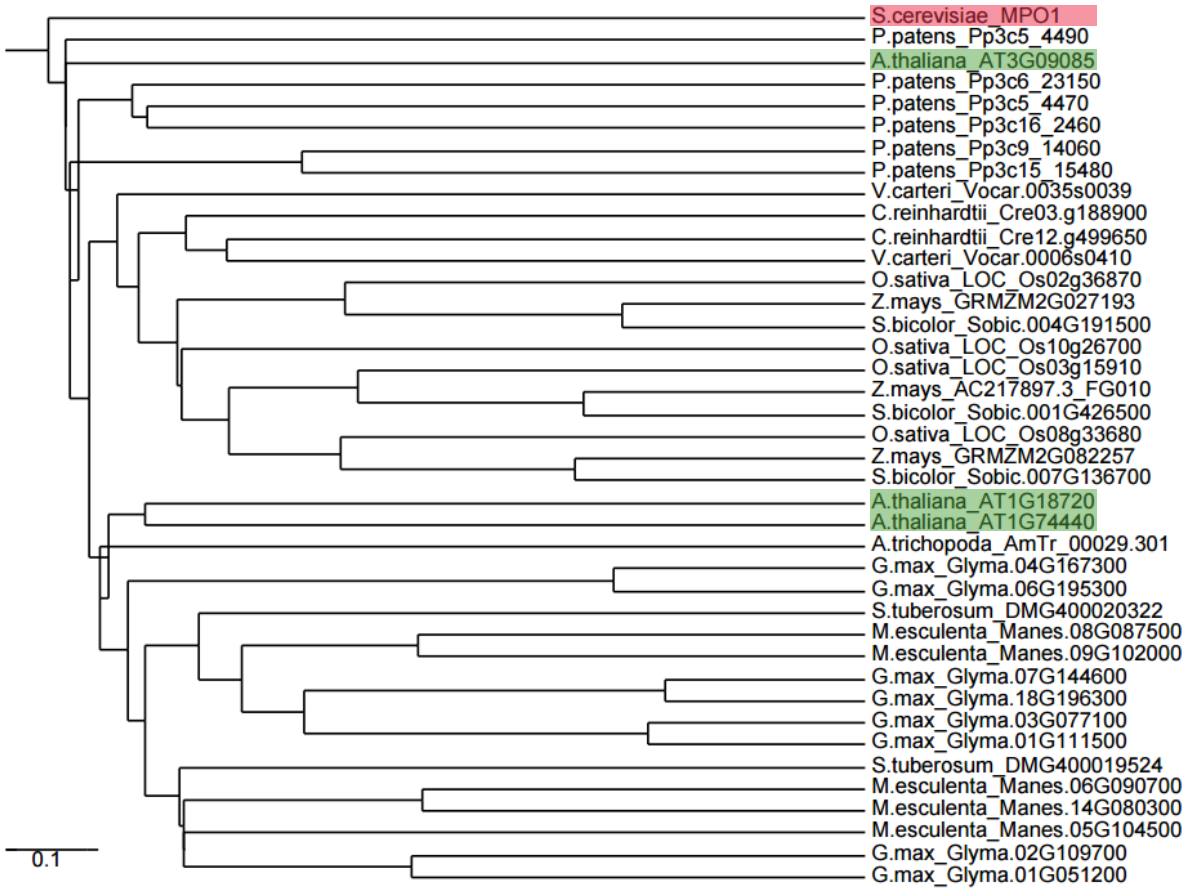


Figure 2-23 Phylogenic relationships of *DUF962* CDS of selected species.
Phylogenic reconstruction of *DUF962*-containing CDS of selected species (*Arabidopsis* CDS are boxed in green). AT1G74440 (*DUF962_1*) and AT1G18720 (*DUF962_2*) are closely related to each other while a more distant relative, AT3G09085 (*DUF962_3*) was closer related to yeast *MPO1* (red box), or *Physcomitrella*. *A. thaliana = Arabidopsis thaliana, A. trichopoda = Aristolochia trichopodia, C. reinhardtii = Clamoydomonas reinhardtii, G. max = Glycine max, M. esculenta = Manes esculenta, O. sativa = Oryza sativa, P. patens = Physcomitrella patens, S. bicolor = Sorghum bicolor, S. cervisiae = Saccharomyces cereviae, S. tuberosum = Solanum tuberosum, V. carteri = Volvox carteri, Z. mays = Zea* mays.

genetic approach utilizing artificial micro RNA was employed.

Based on TAIR v10, *DUF962_1* is closely sequence related to AT1G18720 (*DUF962_2*), since both genes encode the domain of unknown function (DUF) 962. DUF962 is categorized as PFAM (PF) 06127, which is also shared by a third gene in *Arabidopsis*, AT3G09085 (*DUF962_3*). PF06127 is present in a large variety of organisms spanning from unicellular organisms, like bacteria and yeast, via lower plants such as mosses, to higher plants. According to Pfam database v30.0, 383 eukaryotic and 526 bacterial species, containing a total of 1660 sequences with PF06127 have been identified. The *DUF962* coding sequences (CDS) of a small selection of species including *Arabidopsis*, Soybean, Apple, Maize, Rice, *Sorghum*, *Physcomitrella*, *Chlamydomonas*, *Volvox*, and yeast were used for reconstruction of phylogenic relationships. The genetic distance between *Arabidopsis DUF962_1* (AT1G74440) and *DUF962_2* (AT1G18720) was found smallest, and they were closer related to the respective *DUF962* CDS of *Glycine* and *Solanum*, than to *Arabidopsis DUF962_3* which was closer related to the DUF962 CDS of *Saccharomyces* and *Physcomitrella* (Figure 2-23).

2-6-2 DUF962 Proteins of *Arabidopsis* and Yeast are Conserved

To evaluate the sequence conservation of DUF962 proteins, the *Arabidopsis* and the yeast DUF962 CDS were translated into protein, aligned by MUSCLE and the FASTA output was imported into CLC Main Workbench v6 for visualization. The 209 amino residue long protein alignment revealed that the four DUF962 proteins shared a conserved region at the N-terminal end of the protein and another conserved region in the central section of the protein sequence (Figure 2-24).

Arabidopsis DUF962 proteins shared prediction-based common features like the absence of an N-terminal secretion signal peptide, multiple transmembrane domains and 80% of the

Table 2-11 Pair wise comparisons of *Arabidopsis* DUF962 proteins.
Summary of the annotated domains of *Arabidopsis* DUF962 proteins, including predicted transmembrane domain, DUF962 domain size, protein size, isoelectric point, and the amino acid identities determined by pairwise comparison to each other. aa = amino acid., IEP = isoelectric point, SP = secretion signal peptide, TMD = transmembrane domain.

	SP (pred.)	TMD (pred.)	length [aa]	DUF962 [aa]	size [kDa]	IEP (calc.)	% aa identity to At1g74440	% aa identity to At1g18720
At1g74440 (DUF962_1)	no	4-5	208	7 - 171	23.4	9.16	100	74.76
At1g18720 (DUF962_2)	no	4-5	206	7 - 171	23.1	9.28	74.76	100
At3g09085 (DUF962_3)	no	1-3	112	3 - 96	13.2	9.19	23.42	27.93

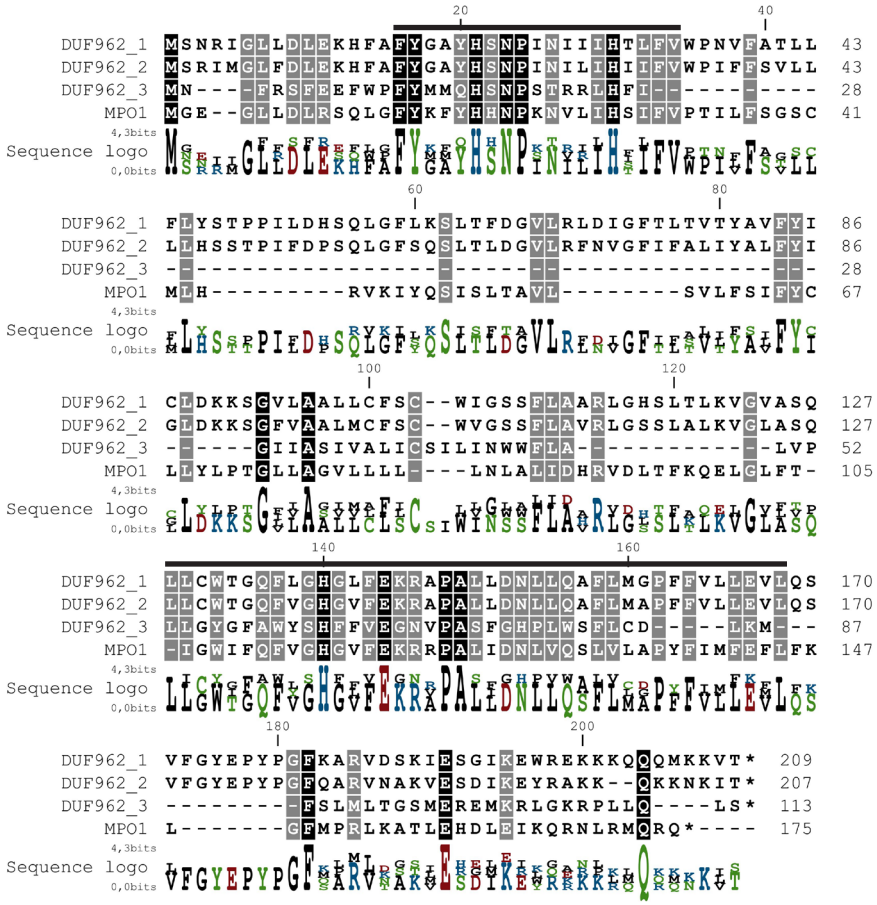


Figure 2-24 *Arabidopsis* and *Saccharmoyes* DUF962 protein alignment.
An alignment of the three *Arabidopsis* DUF962 proteins and the one *Saccharomyces* DUF962 protein (MPO1) demonstrated a high sequence conservation in two parts of the alignment by black (100% match) and grey (> 50% match) boxed residues and the black bar. The colors of the consensus sequence are defined by polarity: black = neutral, nonpolar; blue = basic, polar; green = neutral, polar; red = acidic, polar.

sequence being taken up by the DUF with no additional predicted major functional domains (Table 2-11). Additionally, pair-wise comparison of these proteins revealed that DUF962_1, and DUF962_2 shared 75% amino acid identity and were almost double the size of DUF962_3 which shared less than 28% of amino acid identity with either DUF962_1, and DUF962_2 (Table 2-11).

2-6-3 *DUF962* Genes are Expressed in the Gametophyte and the Sporophyte

Transcriptome analysis of gametophytic cells, after sporophytic contrasting had identified *DUF962_1* to be upregulated in the EC (FC = 2.05), while *DUF962_2*, and *DUF962_3* were not found differentially expressed. However, DUF962_2 was upregulated in the CC (FC 2.28) when compared to the gametophytic cells and SCs (Supplemental Figure 5-3). Examinations of *DUF962_1*, and *DUF962_2* gene expression in *Arabidopsis* pistils, stamen, seedling, leaf, and the EC-like callus by qPCR had determined, relative to the pistil, higher *DUF962_1* transcript

abundance in the leaf, and the EC-like cell line, while the stamen and seedling appeared depleted in *DUF962_1* transcript. The amount of *DUF962_2* transcript was found highest in the stamen, and lowest in pistil tissue. Despite these findings, the localization of *DUF962* gene transcripts by WISH had failed. Therefore, investigations on the cellular localization of DUF962 gene products were continued by detection of DUF962 GFP fusion proteins in transgenic *Arabidopsis* lines.

2-6-4 The DUF962 Gene Family is expressed in the Female Gametes and the Sporophyte

In order to verify the DUF962 expression in the gametes and determine the respective subcellular localization of fusion proteins, a set of constructs was cloned and used to generate stable transgenic *Arabidopsis* reporter lines. The genomic sequences of *DUF962_1, DUF962_2*,

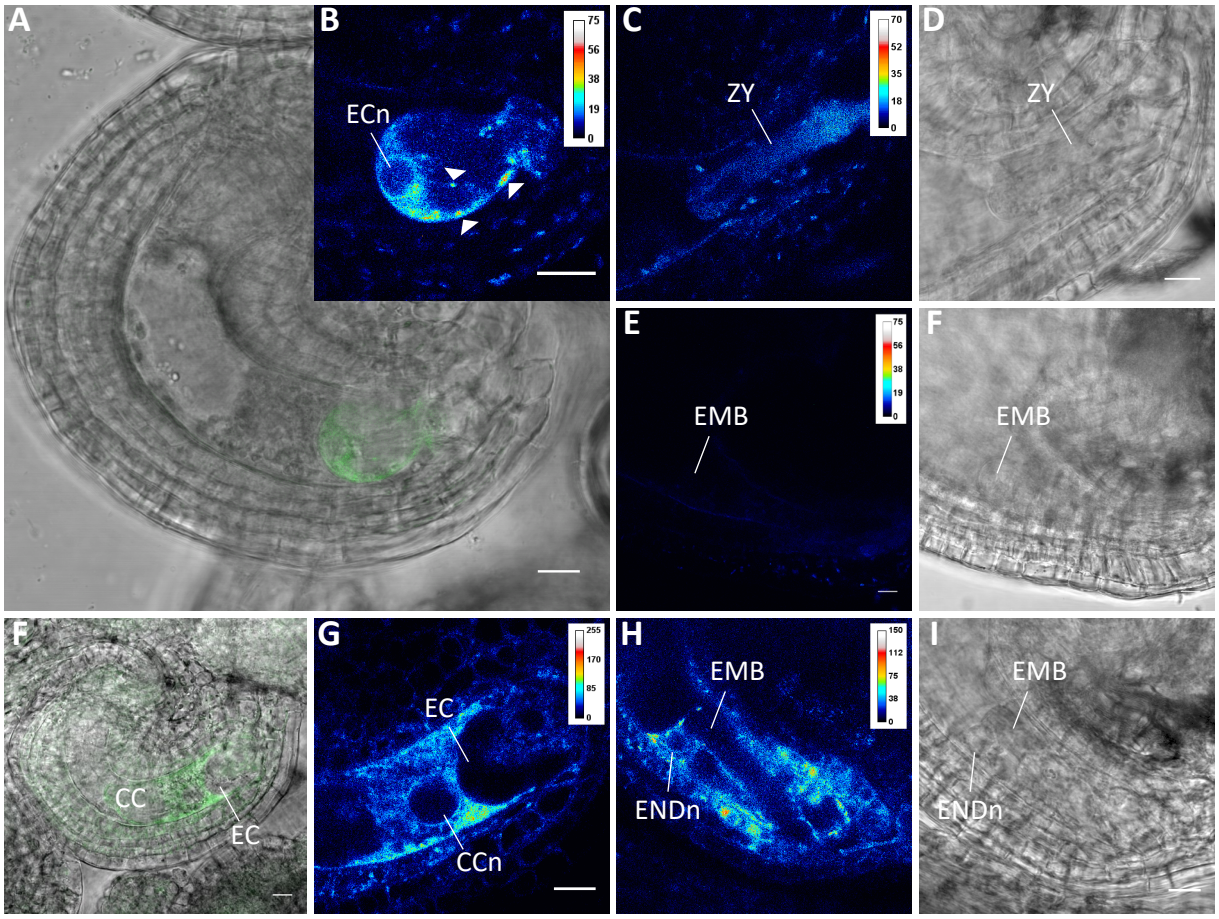


Figure 2-25 Localization of DUF962_1g:GFP and DUF962_2g:GFP in *Arabidopsis* floral tissues. DUF962_1g:GFP was localized to the endomembrane system of the EC (**A, B**), was weakly detected in the zygote (**C, D**) and absent in the early embryo (**E, F**). DUF962_2g:GFP was detected in the endomembrane system of the CC, and inner integuments (**F, G**) and also in the developing endosperm but absent in the embryo (**H, I**). ECn = egg cell nucleus, EMB = embryo, ENDn = endosperm nucleus, CCn = central cell nucleus, ZY = zygote. Triangle indicate accumulated GFP-derived signal. Scale is 10µm.

and *DUF962_3* were obtained from TAIR v10, and 1987, 1990, and 868 bp upstream of the start codon were defined as putative promoter elements, respectively. DUF962 genomic sequences were cloned into vectors for plant expression with C-terminally fused GFP. 12 independent T1 *DUF962_1g:GFP*-expressing reporter lines were investigated by CLSM. During ovule development DUF962_1g:GFP was not detectable in the stages FG1 to FG4 (Supplemental Figure 5-4 A – F), but in mature *Arabidopsis* ECs, DUF962_1g:GFP was found expressed and localized to the endomembrane system, as can be seen by the fluorescent signal surrounding the nucleus and additional dot-like structures within the cell and at the cell periphery (Figure 2-25 A, B triangle). Post fertilization, DUF962_1g:GFP-derived signal was hardly above the detection limit in the zygote (Figure 2-25 C, D), and completely absent in early embryonic stages (Figure 2-25 E, F). Interestingly, in *DUF962_1p:DUF962_1CDS-GFP* plant lines no GFP could be detected (not shown).

DUF962_2g:GFP was clearly absent in the EC, but localized to the endomembrane system of the CC, and the inner integument layers (Figure 2-25 F, G). Moreover, DUF962_2g:GFP-derived fluorescent signals were observed in the developing endosperm and the inner integument layers, while the fluorescent signal was absent in the embryo (Figure 2-26 H, I). qPCR experiments of WT stamen tissue had suggested strong DUF962_2g:GFP-expression in

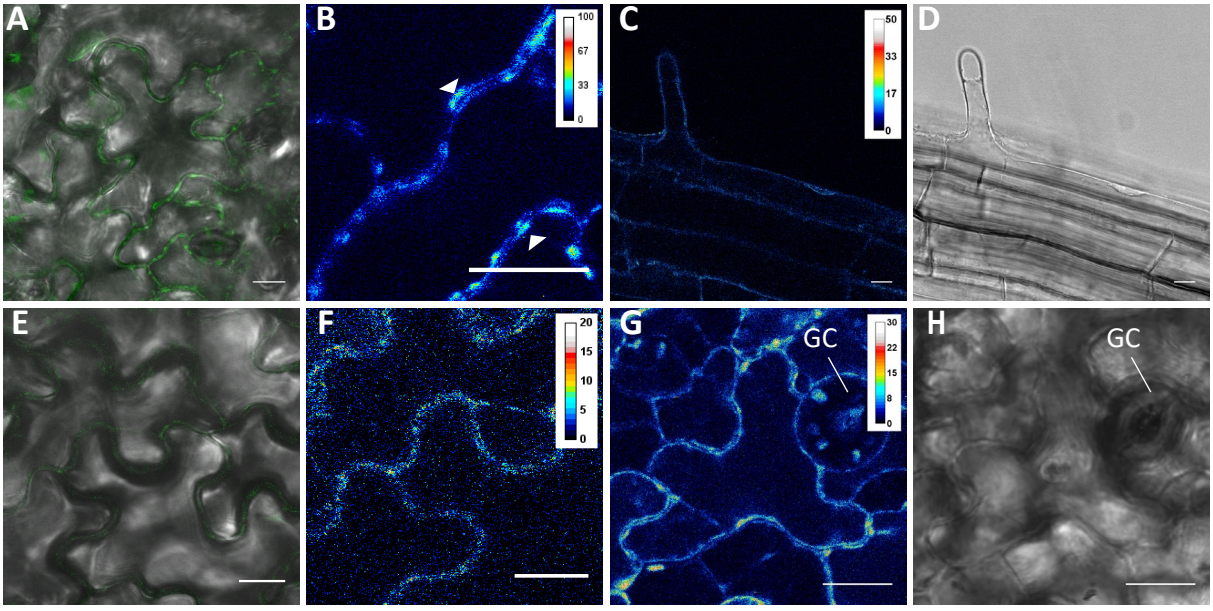


Figure 2-26 Localization of DUF962_1g:GFP, and DUF962_2g:GFP in *Arabidopsis* vegetative tissues.
DUF962_1g:GFP was detected in leaves of bolting plants (**A, B**) and weakly in roots (**C, D**). DUF962_2g:GFP was weakly detected in leaves of bolting plants (**E, F**) and stronger DUF962_2g:GFP expression was observed in cotyledons of two week old seedlings including guard cells (**G, H**). Triangle indicate GFP-derived signal accumulations. GC = guard cell. Scale is 10µm.

stamen but this could not be observed (not shown). In vegetative tissues, DUF962_1g:GFP was detected in leaves of bolting plants, with a similar subcellular distribution like in the EC cytoplasm. It was not detected in the cotyledons of two week old seedlings (not shown), but very weakly in roots (Figure 2-26 A – D). DUF962_2g:GFP was weakly detected in leaves of bolting plants, and also in cotyledons of two week old seedlings including the guard cells (Figure 2-26 E - H), and not in roots (not shown). DUF962_3g:GFP however, could not be detected at all although promoter:reporter lines had determined DUF962_3 promoter activity in all female gametophytic cells, SCs, and leaves.

2-6-5 The *DUF962_1* genomic locus and available T-DNA insertion lines

To verify the annotation of the untranslated regions (UTR) of *DUF962_1*, the corresponding cDNA sequence was obtained from TAIR v10, and to avoid the analysis of truncated sequences in further studies the 5' and 3' UTRs were determined by rapid amplification of cDNA ends (RACE). By this, the annotated UTRs were corrected to a length of 58 nucleotides to the 5' end, and 103 nucleotides to the 3' end. Furthermore, *DUF962_1* harbors an uncharacterized long noncoding RNA (lncRNA) in antisense orientation within the first intron (Figure 2-27 A). The positions of three available T-DNA insertions (GABI_179_G08, SALK_059087, and SALK_1198340), were verified by sequencing and added to the sequence annotation (Figure 2-28 A). These three T-DNA insertion lines were propagated into homozygous generations and mRNA was extracted from floral tissue of each plant. Subsequently, RT-PCR on the respective cDNA templates, including the WT control, showed, when compared to the WT, no significantly altered transcript levels for *DUF962_1* in any of the T-DNA insertion lines, thereby rendering them useless for functional studies (Figure 2-27 B).

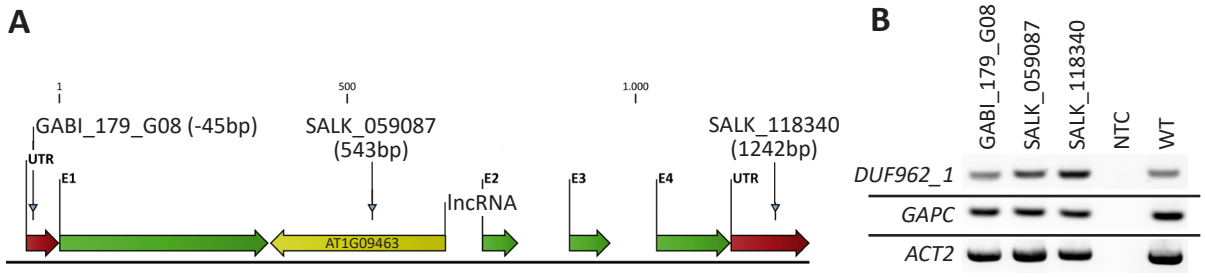


Figure 2-27 DUF962_1 genomic locus and available T-DNA insertion lines.
DUF962_1 genomic locus with RACE-verified UTRs, exon-intron structure, and three available verified T-DNA insertion loci (**A**). RT-PCR of T-DNA insertion lines (GABI_179_G08, SALK_059087, and SALK_1198340), showed no significantly altered *DUF962_1* transcript levels. E = exon, lnc = long noncoding RNA, NTC = non template control, UTR = untranslated region, WT = wildtype.

2-6-6 Knock Down of *DUF962* Genes by amiRNA

As the expression of *DUF962_1* was, in gametophytic cells, EC-specific and neither overlapping with *DUF962_2* nor with *DUF962_3*, which had not been detected on protein level, a single gene-specific knock out in the EC could provide insights into a possible biological function of *DUF962_1* in the EC. Due to the lack of suitable T-DNA insertion lines, an artificial micro RNA (amiRNA) targeting *DUF962_1* and *DUF962_2* (amiR^{DUF962}) was designed and subsequently cloned under the transcriptional control of a set of promoters.

The *EC1.1* promoter (Ingouff *et al.,* 2009) was used for EC-specific expression of the amiR^{DUF962}, as by CLSM analysis of *EC1.1p:GFP*-expressing plant lines GFP-fluorescence could only be observed in the mature EC (Supplemental Figure 5-5 A - F). The CC-specific expression of the amiR^{DUF962} was achieved by the *DD65* promoter (Steffen *et al.,* 2007). Furthermore, for sporophyte-specific amiRNA^{DUF962} expression the *CaMV 35S* promoter (Koziel *et al.,* 1984) sequence was utilized, which was determined not to be active in the ovule and the FG (Supplemental Figure 5-5 G, H). Additionally, the *KNUCKLES* promoter (*KNUp*), and the promoter sequences of AT1G21670, which both had been reported active during the development of the FG (Supplemental Figure 5-2 I – P; Tucker *et al.,* 2012) were utilized to express the amiRNA^{DUF962}. Consequently, these

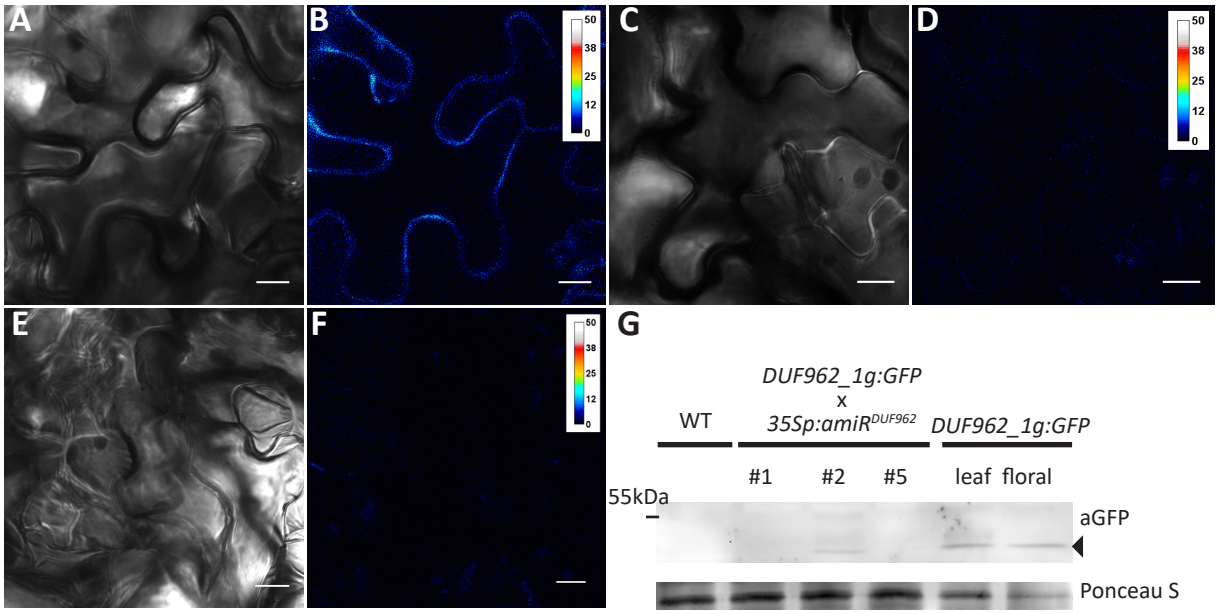


Figure 2-28 Verification of artificial micro RNA-mediated knock down of *DUF962_1*.
CLSM analysis of the leaves of *DUF962_1g:GFP*-expressing plant lines (**A, B**), and of *DUF962_1g:GFP x 35Sp:amiR*^{DUF962}-expressing plant lines (**C, D**), and the WT (**E, F**) did not detect GFP-derived signal in DUF962_1g:GFP x 35Sp:amiR plant lines (**C, D**), thereby indicating a true *DUF962_1* knock down. Western blotting of leaf extract verified the absence of fusion protein in F1 progeny of 35Sp:amiR^{DUF962} and *DUF962_1g:GFP* plants #1, #5, while #2 showed residual signal. DUF962_1g:GFP leaf, and floral tissue served as positive control, leaves of the WT as negative control. Scale is 10µm.

constructs and an empty Destination vector control (EC1.1p:Gateway) were used for floral dip-mediated plant transformation.

To verify the functionality of the amiR^{DUF962}, one *DUF962_1g:GFP* reporter line was crossed with one *35Sp:amiR^{DUF962}* line and the resulting offspring was analyzed by CLSM to determine DUF962_1g:GFP protein levels in leaves. The presence of both T-DNAs within the genome of five F1 plant lines were verified by resistance-based selection, and PCR-based genotyping. By CLSM analysis of leaves of bolting plants of the WT, *DUF962_1g:GFP*, and *DUF962_1g:GFP* x *35Sp:amiR^{DUF962}* plant lines, GFP-derived signal was found absent in leaves of five *DUF962_1g:GFP* x *35Sp:amiR^{DUF962}* - expressing plant lines, but could be detected in the *DUF962_1g:GFP* parental line (Figure 2-28 A – F). To provide biochemical evidence, protein extracts from leaf tissue of the WT, and of *DUF962_1g:GFP* x *35Sp:amiR^{DUF962}* plant #1, #2, #5, and the positive control (*DUF962_1g:GFP*) were prepared. Western blots of the protein extracts were probed with anti-GFP antibody and subsequent chemiluminescence detection showed no detectable DUF962_1:GFP-derived signal in *DUF962_1g:GFP* x *35Sp:amiRNA^{DUF962}* plant #1, #5, while plant #2 retained residual signal (Figure 2-28 G). These findings clearly demonstrated an amiRNA^{DUF962}-mediated knock down of DUF962_1:GFP on the protein level.

2-6-7 The Seed Sets of *Arabidopsis DUF962* amiRNA-Knock Down Lines

To address a possible effect of the *DUF962_1* knock down on the number of non developed seeds. Five siliques each, of ten WT plants, three *ec1* quintuple mutant plants (Sprunck *et al.,* 2012), 26 *EC1.1p:amiR^{DUF962}*-expressing plant lines, 11 *DD65p:amiR^{DUF962}*-expressing plant lines, 12 *35Sp:amiR^{DUF962}*-expressing plant lines, eight *AT1G21670p:amiR^{DUF962}*-expressing plant lines, seven *KNUp:amiR^{DUF962}* -expressing plant lines, and eight *EC1.1p:Gateway* control plant lines were harvested and cleared over night. The estimated amount of non developed seeds per silique was calculated relative to the amount of estimated maximum seeds per silique, then averaged. The resulting seed sets per genotype were subjected to a two-tailed students t-test, referencing the WT.

Notably, the *EC1.1p:amiRNA* plant lines were found seed set-wise statistically significant different from the WT. However, these amiR^{DUF96}-expressing lines showed a larger variation in seed set when compared to the *ec1* quintuple mutant (Figure 2-29 A). By contrast, the

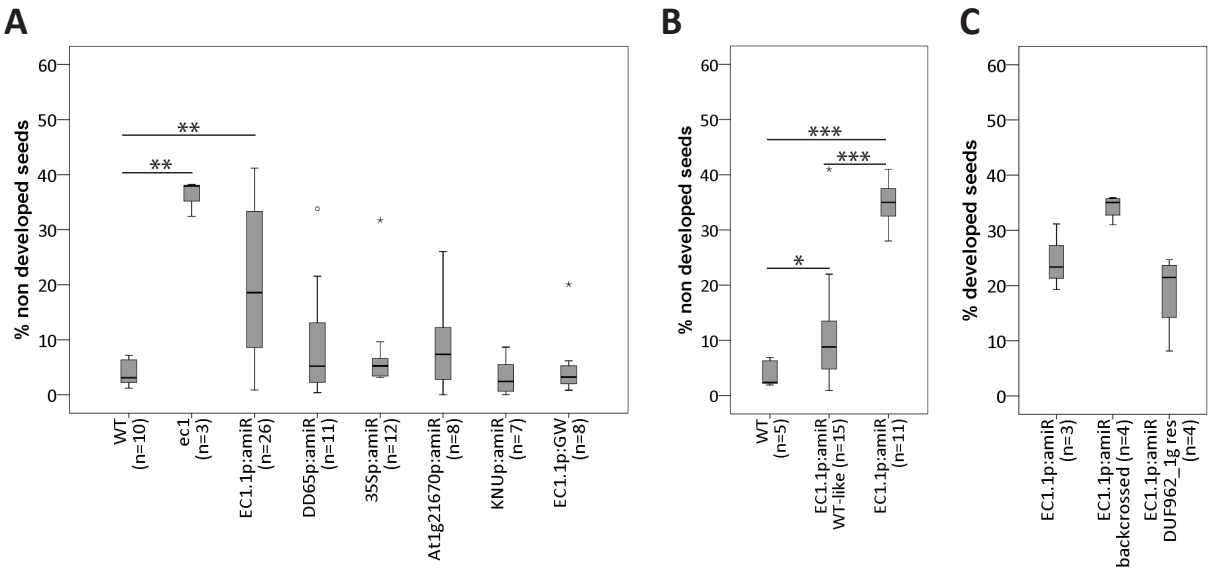


Figure 2-29 Relative non developed seeds in siliques of *Arabidopsis DUF962_1/2* knock down mutant plant lines, and controls.
Two-tailed students t-tests were performed to determine statistical significance of seed set data. Relative non developed seeds in % per silique and genotype.
(**A**) Seed sets of 10 WT plants and 3 *ec1* quintuple mutant plants, and n-independent plant lines expressing *EC1.1p:amiRNA^{DUF962}, DD65p:amiRNA^{DUF962}, 35Sp:amiRNA^{DUF962}, AT1G21670p:amiRNA^{DUF962}, KNUp:amiRNA^{DUF962}*, and *EC1.1p:Gateway* (**A**). The *ec1* quintuple mutant, and *EC1.1p:amiRNA^{DUF962}* expressing plant lines were found seed set-wise statistically significant different from the WT.
(**B**) *EC1.1p:amiRNA* expressing plant lines were grouped into statistically significant WT-like phenotype plant lines, and statistically significant mutant-like phenotype plant lines.
(**C**) One mutant-like phenotype plant line was backcrossed into the WT twice. The mutant genotype of the respective offspring was verified for four plants but these did not show a statistically significant different seed set when compared to offspring (three plants) of the parental line. Transformation of a mutant-like phenotype plant line with an amiRNA-resistant *DUF962_1g^{res}* gene and analysis of four verified transgenic mutant offspring plant lines did seed set-wise not statistically significant deviate from the respective offspring of the parental plant line.

seed sets of the *DD65p:amiR^{DUF962}*, *35Sp:amiR^{DUF962}*, *AT1G21670p:amiR^{DUF962}*, *KNUp:amiR^{DUF962}*, and *EC1.1p:Gateway* mutant genotypes were not significantly affected (Figure 2-29 A). The *EC1.1p:amiR^{DUF962}* plant lines exhibited a high amount of seed set-dependent variation between individual plant lines, ranging from WT-like levels to almost 60% of non developed seeds (Figure 2-29 A). By subjecting the seed sets of the individual *EC1.1p:amiR^{DUF962}* plant lines to two-tailed students t-test, two subpopulations were identified: 15 plant lines had a statistically significant WT-like seed set, whereas 11 plant lines had a statistically significant different seed set compared to the WT, and WT-like mutant plant lines (Figure 2-30 B). These 11 plant lines were therefore termed “mutant-like phenotype plant lines” and showed a median value of 35% non developed seeds (Figure 2-29 B).

The mutant-like phenotype plant line #19 was selected for two approaches with the goal to rule out an integration effect due to the insertion of multiple T-DNAs, and to reverse the mutant-like phenotype to a WT-like phenotype. Pollen of the mutant-like phenotype plant line

#19 was used for pollination of the WT. Subsequently obtained mutant-like F1 plants were verified by antibiotic-based selection, and PCR-based genotyping, then backcrossed into the WT background once more. Presence of the *EC1.1p:amiR*^{DUF962} T-DNA was again verified in the F2 generation by antibiotic-based selection and PCR-based genotyping. Siliques of four plants were cleared and the respective relative seed sets were calculated and found seed set-wise not statistically significant different from offspring of the parent mutant-like phenotype plant line #19 (Figure 2-29 C).

Moreover, the mutant-like phenotype plant line #19 was transformed with a putative complementation construct DUF962_1g^{res}, in order to revert resulting offspring to a WT-like phenotype. It is known that a total of two mismatches in the amiRNA binding site at the positions 1, or between positions 15, and 21 do not significantly interfere with amiRNA performance (Li *et al.,* 2013), therefore silent mutations were introduced in *DUF962_1g* at the amiRNA binding positions 4, 11, 13, and 19, to generate a mutated DUF962_1res gene to which the *amiR*^{DUF962} would potentially not hybridize anymore. The presence of both *EC1.1p:amiR*^{DUF962} and *DUF962_1g*^{res} T-DNAs was verified by antibiotic-based selection, and PCR-based genotyping in four plant lines but these were seed set-wise statistically not significant different from offspring of the parental mutant-like phenotype plant line #19 (Figure 2-29 C).

In summary, expression of the amiR^{DUF962} by the *EC1.1* promoter led to statistically significant less developed seeds per silique in 11 out of 26 T1 plant lines. This seed set-dependent phenotype was not reverted by introducing the putative amiR^{DUF962} resistant construct *DUF962_1g*^{res} into the mutant-like phenotype plant line #19, nor was it reverted to a WT-like phenotype in transgenic F2 progeny of the twice backcrossed (into the WT) plant line #19.

2-6-8 The FG of *EC1.1p:amiR*^{DUF962}-expressing Plant Lines Arrest in Early Developmental Stages

To elucidate the reasons for the reduced seed set in *EC1.1p:amiR*^{DUF962}-expressing plant lines the development of the FG gametophytes in the mutant-like plant lines #16, #19, #22, #26, #28, #32, and the WT were investigated by Feulgen staining of floral stage 12a – 12c pistils. Ovules of mutant-like phenotype plant lines in stage 3-I looked undistinguishable from those of the WT (Figure 2-30 A, B). However, the formation and enlarging of the central vacuole that

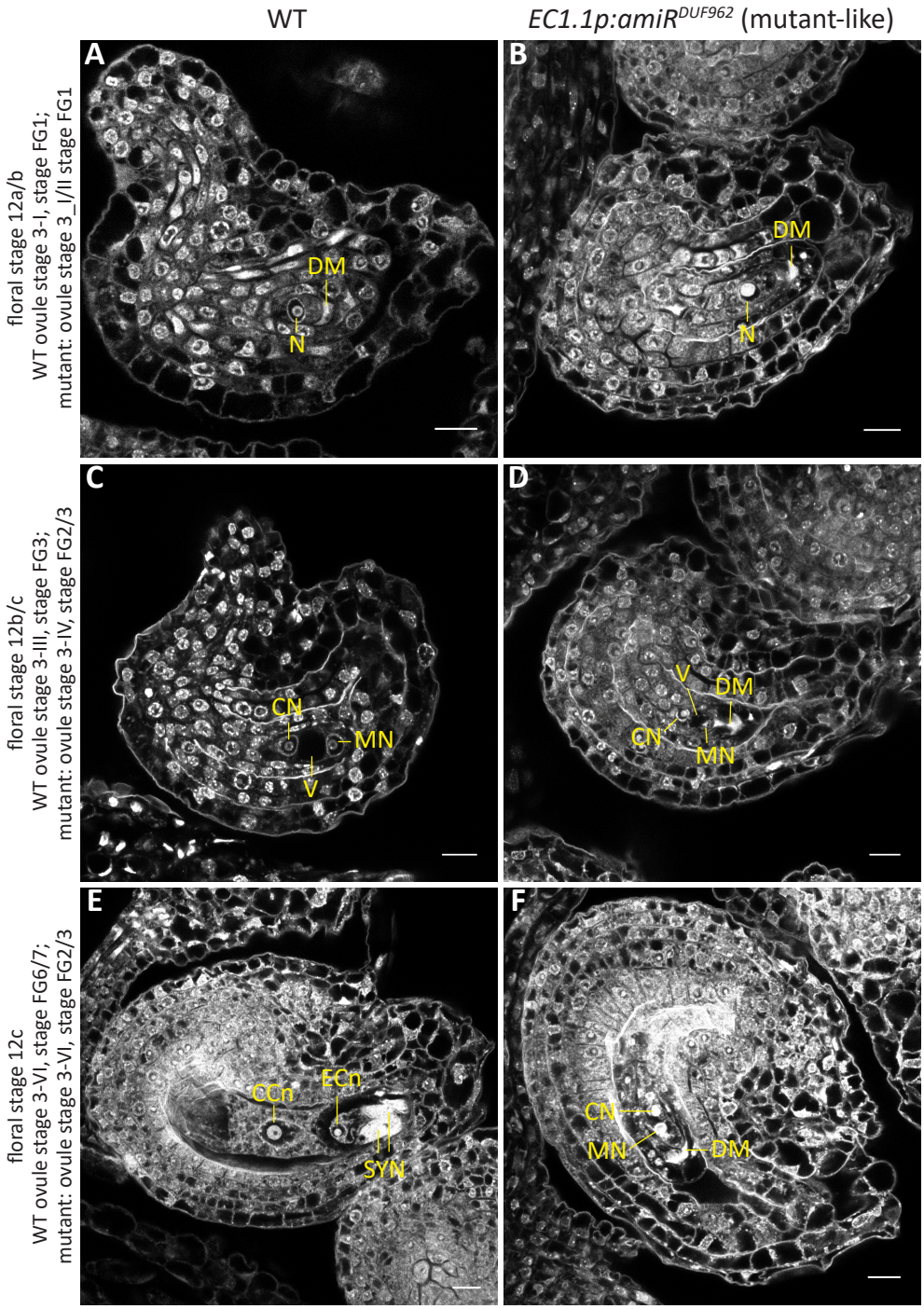


Figure 2-30 Female gametophyte development in *EC1.1p:amiR*^{DUF962} knock down lines and the WT.

Optical sections of Feulgen-stained pistils of floral stage 12a/b (**A, B**), 12b/c (**C, D**), 12c (**E, F**) of the WT (**A, C, E**) and the *EC1.1p:amiR*^{DUF962} mutant-like phenotype plant lines (**B, D, F**). A developmental phenotype was identified during the FG2 to FG3 transition where the central vacuole, that separates the chalazal and the micropylar nucleus, did not fully form (**C, D**). The mutant embryo sacs remained, when compared to the WT, small, halted karyokinesis and did not reach maturity (**E, F**). CCn = central cell nucleus, CN = chalazal nucleus, DM = degenerated megaspores, ECn = egg cell nucleus, MN = micropylar nucleus, N = functional megaspore, V = vacuole. Scale is 10µm.

separates the chalazal and micropylar nucleus and characterizes the FG2 to FG3 transition did not fully occur in the mutant-like plant lines (Figure 2-30 C, D). Moreover, the chalazal, and micropylar nucleus did not undergo further karyokinesis eventually leading to FG6, but instead

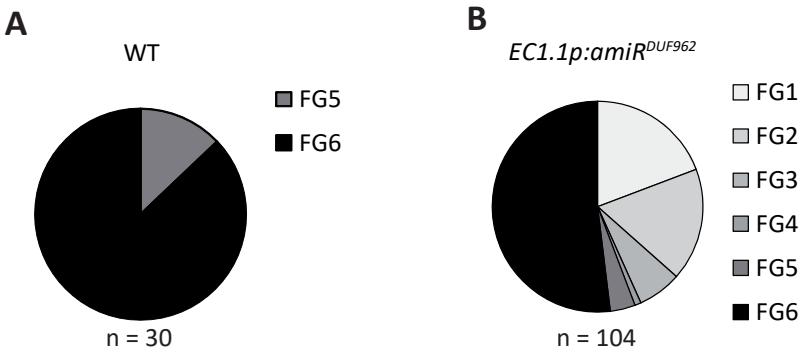


Figure 2-31 Distribution of FG developmental stages in floral stage 12c pistils of *EC1.1p:amiR*^{DUF962} mutant-like phenotype plant lines, and WT.
Ovules of floral stage 12c pistils of heterozygous *EC1.1p:amiR*^{DUF962} mutant-like phenotype plant lines #16, #19, and WT were analyzed by Feulgen staining. Resulting distribution of observed female gametophyte developmental stages for 30 WT (**A**), and 104 *EC1.1p:amiR*^{DUF962} (**B**) ovules showed 43% developmentally arrested ovules in the mutant plant lines.

the developing embryo sac remained small and arrested in stage FG2/3, while the surrounding sporophytic tissue developed normally, which led to an obvious difference between the WT and the mutant ovules (Figure 2-30 E, F). The relative amounts of observed FG developmental stages in the two heterozygous *EC1.1p:amiRNA*^{DUF962} mutant-like phenotype plant lines #16, #19, and the WT were determined by quantifying the phenotypes of Feulgen-stained floral stage 12c pistils. Of 30 WT ovules four (13%) were determined to be in stage FG5, whereas the remaining ovules had reached maturity (FG6), while in 104 analyzed ovules of the mutant-like plant lines, 45 (43%) of the ovules remained in stage FG1 to FG3 (Figure 2-31 A, B).
This frequent arrest in FG development observed in 42% of the *EC1.1p:amiRNA*^{DUF962} -expressing plant lines will be discussed in detail later (Chapter 3-7). So far DUF962_1 could not be linked to a molecular function, but this was achieved by a switch in the model system.

2-6-9 The *S. cerevisiae* Δ*mpo1* Mutant is Complemented by *Arabidopsis* DUF962_1

DUF962 proteins are harbored by a variety of organisms, among them *Saccharomyces cerevisiae*, which contains only one DUF962-encoding gene. This gene DUF962-encoding gene was determined non-essential for cell viability in yeast, but a full knock out was found to have a slightly reduced fitness in minimal medium (Breslow *et al.,* 2008; Qian *et al.,* 2012). Furthermore, during the logarithmic growth phase, this DUF962 mutant was shown to accumulate a breakdown product of phytosphingosine (PHS) in the growth medium. Hence, the yeast DUF962 protein was determined to be required to catalyze the conversion

of 2-hydroxypalmitoyl-CoA to pentadecanoic acid, and termed Metabolism of PHS to Odd-numbered fatty acids 1 (MPO1) (Kondo *et al.,* 2014). As DUF962 proteins are conserved, a functional complementation of the yeast mutant Δ*mpo1* with *Arabidopsis DUF962_1*, and *DUF962_3* was pursued.

In order to prepare for the functional complementation of the Δ*mpo1* mutant, the CDS of *DUF962_1*, *DUF962_3*, and *MPO1* were cloned under the transcriptional control of the PMA1 promoter, and subsequently used for transformation of Δ*mpo1* cells.

As radiolabelled PHS was not available for purchase, in a pilot experiment, the Δ*mpo1* mutant and the corresponding parental WT strain (BY4741) were fed 37MBq [9,10] ^{3}H palmitic acid (PA) directly, and after four hours of logarithmic growth and metabolic conversion of PA into PHS and other PA/PHS-derived compounds, the lipids were extracted from the cell-free growth medium and the cell pellets. Lipid extracts with a specific activity of 1.85MBq each, were resolved by normal-phase TLC, and subsequently exposed to X-ray films. The resulting pattern revealed an additional band in the growth medium-derived sample of the Δ*mpo1* mutant (Figure 2-32 asterisks), which resembled previous findings (Kondo *et al.,* 2014).

Thereby, the Δ*mpo1* mutant complementation assay was conducted with 7.4MBq of [9,10]^{3}H PA which were fed to the established logarithmically growing yeast clones with the genotypes Δ*mpo1,* Δ*mpo1 PMA1p:DUF962_1*, Δ*mpo1 PMA1p:DUF962_3*, Δ*mpo1 PMA1p:MPO1* (as positive control), and the WT. Subsequently, after four hours, the lipids were extracted from

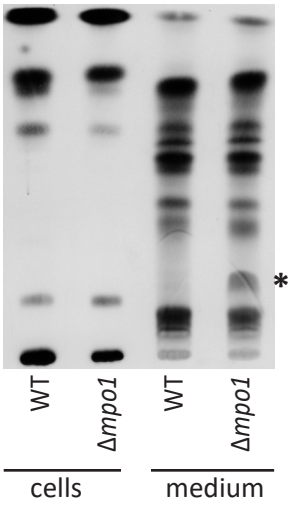


Figure 2-32 Lipid extracts of ^{3}H PA-fed *Δmpo1* mutant and WT.
Logarthmic growing yeast WT and *Δmpo1* cells were fed 37MBq of [9,10]^{3}H palmitic acid, and the lipids were extracted from the cell pellets and the cell-free growth medium. The lipid samples with a specific activity of 185kBq each, were resolved by normal-phase TLC with Hexane:Dieethylether:Acetic Acid 30:70:1. After X-ray film exposure and development, a *Δmpo1* mutant-specific band was identified in the medium-derived sample (asterisks).

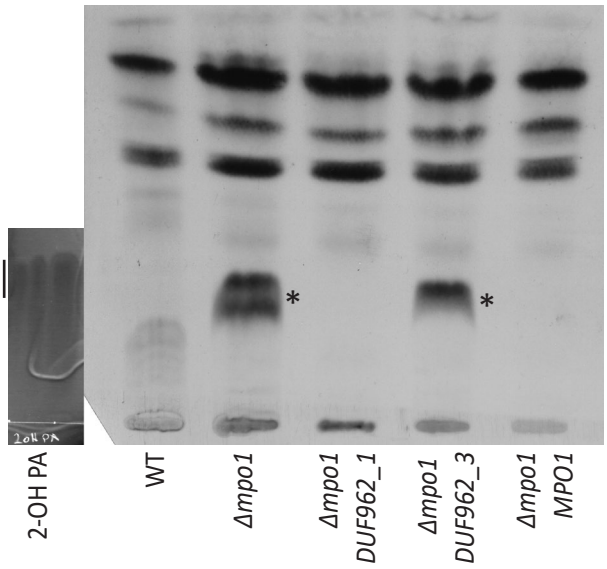


Figure 2-33 Lipid extracts of ^{3}H PA-fed yeast mutants, and WT adressing PHS-catabolism related phenotypic complementation.
Lipids were extracted from the cell-free growth medium, and 31kBq per sample were spotted and resolved by normal-phase TLC with Hexane:Dieethylether:Acetic Acid 30:70:1, along with a 2-hydroxy palmitic acid standard. The lane corresponding to the standard was cut from the plate and stained with Rhodamine6G. The remaining plate was coated with 7% (w/v) PPO and exposed to a X-ray film. The realigned plates revealed an accumulation of signal (asterisks) at the migration height of 2-hydroxypalmitic acid (black bar) in Δ*mpo1,* and Δ*mpo1 PMA1p:DUF962_3* cells. 2-OH PA = -hydroxypalmitic acid.

the cell-free growth medium, and the cell pellets. After resolving with normal-phase TLC, no differences in the patterns of 85.1kBq cell pellet derived-lipid samples of the individual genotypes were observed (not shown).

Lipid extracts of the cell-free growth medium with a corresponding radioactivity of 31kBq, as determined by liquid scintillation counting, each were resolved by normal-phase TLC, along with a non radioactive 2-hydroxypalmitic acid (2-OH PA) standard. The lane corresponding to the 2-OH PA standard was cut off and stained with Rhodamine6G, while the remaining plate was exposed to a X-ray film, and developed. Afterwards, the plate fragments were realigned and on the corresponding migration height of 2-OH PA a band each for the genotypes Δ*mpo1*, and Δ*mpo1 PMA1p:DUF962_3* was identified (Figure 2-33 asterisks). The retardation factors, for the obious bands were determined to the relative similar values of 3.2 for 2-OH PA, and to 3.5 for Δ*mpo1*, and 3.4 for Δ*mpo1 PMA1p:DUF962_3*.

Notably, the 2-OH PA corresponding band of Δ*mpo1 PMA1p:DUF962_3* appeared weaker than the one the Δ*mpo1* genotype sample which appeared as double band (Figure 2-34 asterisks). These findings demonstrated a PA/PHS-catabolism dependent phenotypic rescue of the Δ*mpo1* mutant by *PMA1p:DUF962_1*, *PMA1p:MPO1*, and possibly partly by *PMA1p:DUF962_3.*

2-6-10 The Role of Tetraspanins in *Arabidopsis* Gametophytic Cells

Mammalian tetraspanins are involved in a variety of signaling and development-related processes, and were found essential for mammalian fertilization (Evans 2012; Charrin *et al.,* 2014). Similar functions are hypothesized for *Arabidopsis* tetraspanin proteins but remain to be shown. To address the elusive biological function of gametophytic-expressed *TET7*, *TET8*, *TET9*, *TET11*, and *TET12* in the context of membrane-interactions during *Arabidopsis* gametophytic development and gamete interactions the expression of TET7, TET8, and TET9 GFP fusion proteins was investigated in *Arabidopsis* vegetative and generative tissues. Furthermore, a simultaneous knock out of multiple tetraspanins (*TET7, TET8, TET9, TET11,* and *TET12*) was achieved and the resulting phenotype was investigated. In addition, the split ubiquitin system was exploited for a TET9 interactor screen.

2-6-11 Two Subclades of the Tetraspanin family are Preferentially Expressed in Gametophytic Cells

By the initial transcriptomic analysis of gametophytic cells 17 out of 33 *Arabidopsis* tetraspanins, were found differentially expressed in *Arabidopsis* gametophytic cells. By contrasting to

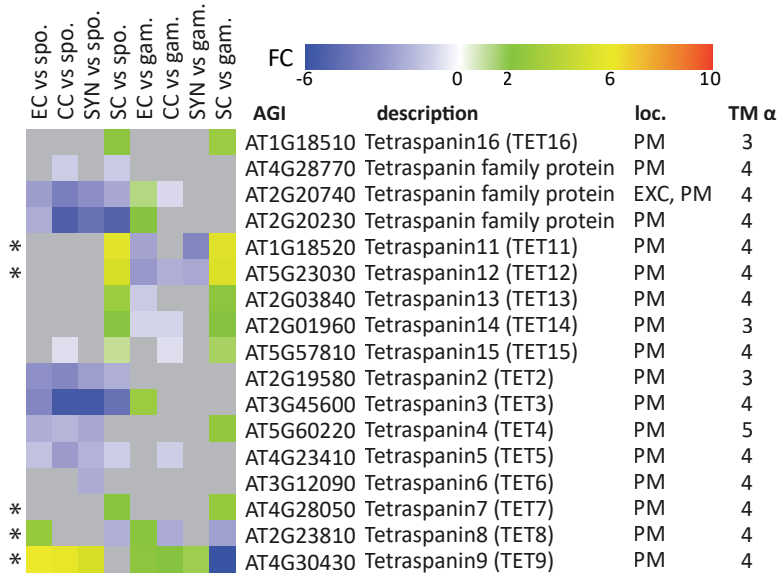


Figure 2-34 Differentially expressed Tetraspanins according to the transcriptomic data of female gametophytic cells and male gametes.
The analysis of the transcriptomic data of female gametophytic cells, and SCs identified 17 differentially expressed Tetraspanin genes. The heatmap depiction shows *TET11* and *TET12,* which were determined highly SC-specific, and *TET7, TET8*, and *TET9* that were found enriched in female gametophytic cells. Asterisks indicates candidate genes, protein localization as determined by SUBAcon and predicted transmembrane domains by Aramemnon are displayed. EXC = extracellular, gam. = gametophyte, loc. = localization, PM = plasma membrane, spo. = sporophyte, TM α = transmembrane domains by alpha helices.

sporophytic tissue, *TET7* was found upregulated in the SCs (FC = 1.87), *TET8* upregulated in the EC (FC = 1.59), *TET9* upregulated in the EC (FC = 6.0), the CC (FC = 5.78), and the SYN (FC = 5.24), and *TET11* and *TET12* were both upregulated in the SCs (each FC = 5.65; Figure 2-35). The analyses were thus restricted to preferentially gametophytic-expressed *TET7*, *TET8*, *TET9*, *TET11*, and *TET12* (Figure 2-34 asterisks). The tetraspanin genes of interest all encode proteins with four predicted transmembrane domains and SUBAcon-based localization to the plasma membrane. The *Arabidopsis* tetraspanin expression pattern had already been in the focus of a previous investigation (Boavida *et al.,* 2013), but those results deviated slightly from the findings of this work.

For expression studies and subcellular localization the genomic sequences of *TET7* and *TET8,* including 250bp and 934bp putative promoter sequences, respectively, were cloned into pGWB550 and used for the generation of stable *Arabidopsis* lines. A homozygous reporter line for *TET9:GFP*, (in pB7FWG2.0), with an endogenous 1571 bp putative promoter fragment, was already established and provided by Dr. Stefanie Sprunck. These tetraspanin reporter lines were used for CLSM-based detection of the respective tetraspanin GFP fusion proteins in floral stage 12b – 12c pistils, pollen, and germinated pollen tubes (PT) of floral stage 13

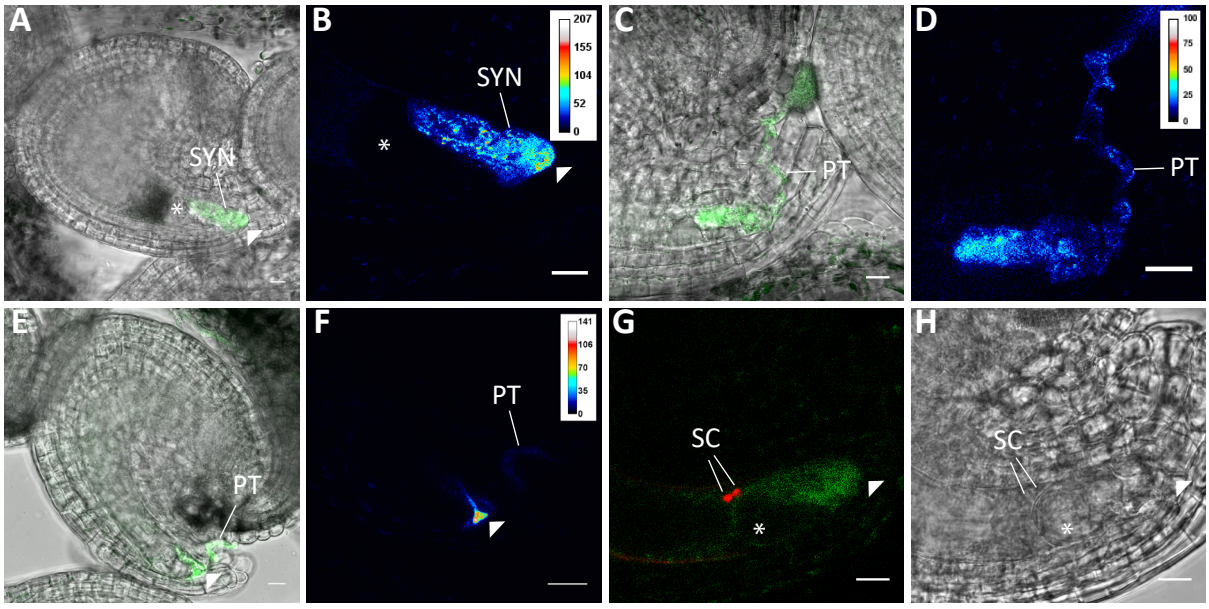


Figure 2-35 Localization of TET7g:GFP in mature ovules and during double fertilization.
The micropylar pole of the ovule is indicated by the triangle and the position of the EC by the asterisks. In mature ovules, TET7g:GFP localized to the SYN and accumulated at the filiform apparatus (**A, B**). It was also localized in pollen tubes, here a *TET7g:GFP* pollinated WT plant (**C, D**). Upon pollen tube arrival of self-pollinated *TET7g:GFP* plants, the strongest detected fluorescence was observed in the optical section of the micropylar pole of the ovule (**E, F**). After pollen tube burst, the TET7:GFP-derived signal grew weaker very quickly and was hardly detectable prior to gamete fusion. PT = pollen tube, SC = sperm cell. SYN = synergid cell. Scale is 10µm.

stamen, in leaves of bolting plants, in cotyledons of 14 day old seedlings, and in roots. In mature ovules, *TET7g:GFP* was found expressed in the SYN with an accumulation towards the filiform apparatus (Figure 2-35 A, B triangle). Pollen (not shown) and the germinated PT also expressed TET7g:GFP, as observed in *TET7g:GFP* pollinated WT pistils. However, like in the SYN the subcellular localization resembled not exclusively plasma membrane-localization but rather a localization in the endomembrane system (Figure 2-35 C, D). The optical section focusing on the micropylar entry point of the PT into the ovule reached the highest GFP-derived signal intensities of self pollinated *TET7g:GFP* plants (Figure 2-35 E, F triangle). When pollinating TET7g:GFP plant lines with a red SC marker line (Ingouff *et al.,* 2007), the microscopy-based analysis after PT burst, receptive SYN degeneration, and SC delivery to the fusion site revealed that the GFP-derived fluorescence grew more diffuse and weaker very quickly (Figure 2-35 G, H).

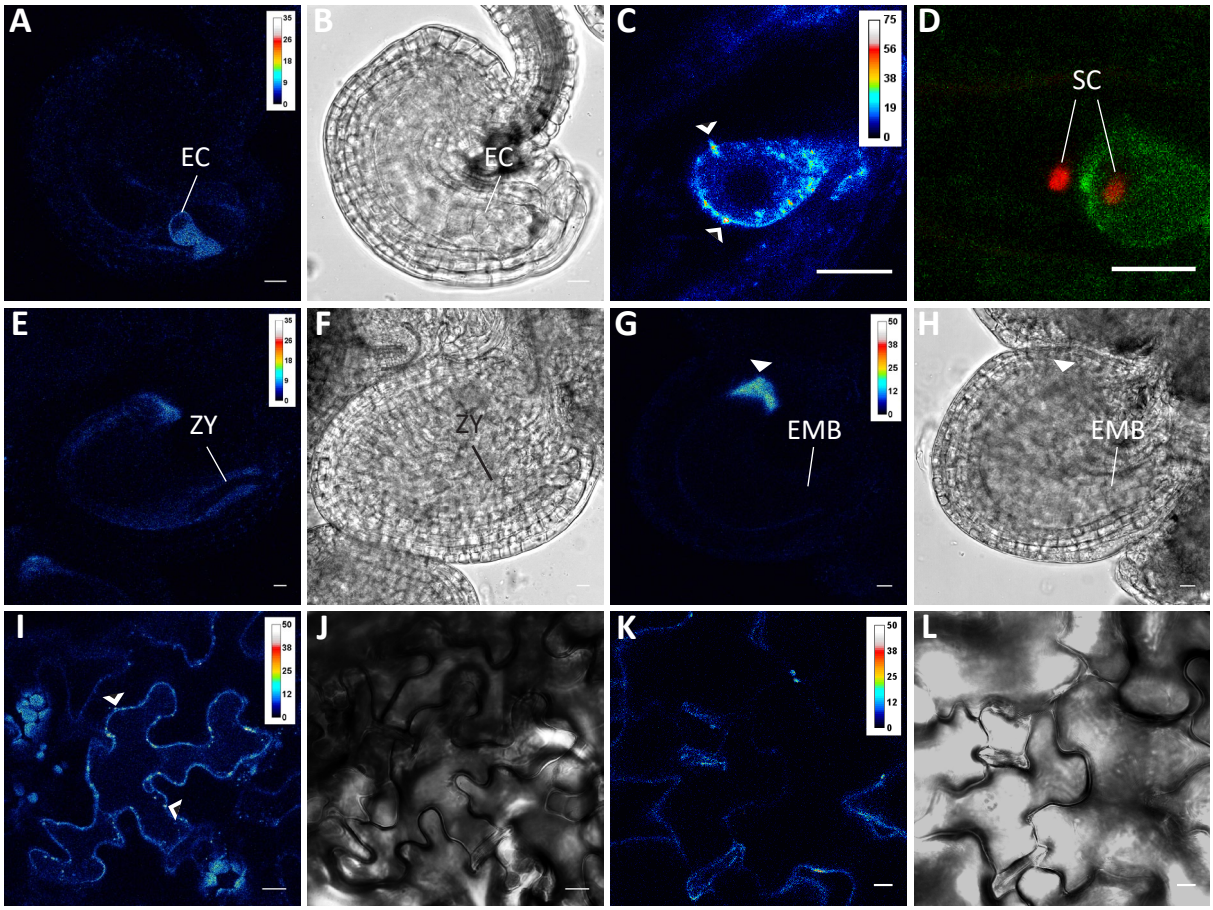


Figure 2-36 Localization of TET8g:GFP in *Arabidopsis* TET8g:GFP reporter lines.
In the ovule, TET8g:GFP was localized to the mature EC, where it displayed patchy plasma membrane and endomembrane localization (**A - C** arrowhead). Upon gamete fusion no obvious redistribution of TET8g:GFP was observed (**D**). TET8g:GFP was already absent in the zygote (**E, F**), but reappeared at the chalazal pole (triangle) once the seed development reached an early embryonic stage (**G, H**). TET8g:GFP was also detected in the leaves in a patchy manner (**I, J** arrowhead), and weakly in the cotyledons (**K, L**). EC = egg cell, EMB = embryo, SC = sperm cell, ZY = zygote. Scale is 10µm.

Analogous investigations were performed on *TET8g:GFP* expressing plant lines. In the ovule, TET8g:GFP localized to the mature EC of unpollinated and pollinated pistils in mostly the endomembrane system with a patchy accumulation of GFP-derived signals near the border of the plasma membrane (Figure 2-36 A – C, arrowheads). Upon gamete fusion no obvious change in subcellular localization of TET8g:GFP was observed (Figure 2-37 D). Post fertilization, TET8g:GFP was absent in the zygote (Figure 2-36 E, F) but reappeared during ovule stage 4-VI at the chalazal pole of the ovule (Figure 2-36 G, H triangle). In the leaves as well as the cotyledons TET8g:GFP was detected (Figure 2-36 I – L), however not in roots (not shown).

TET9p:TET9-GFP plants were investigated as well, here TET9:GFP was not visible in ovule stage 3-II, but appeared above detection limit in stage FG4 (Figure 2-37 A – D). The strongest GFP-derived signal was detected in mature female gametophyte with signals in the EC, and the

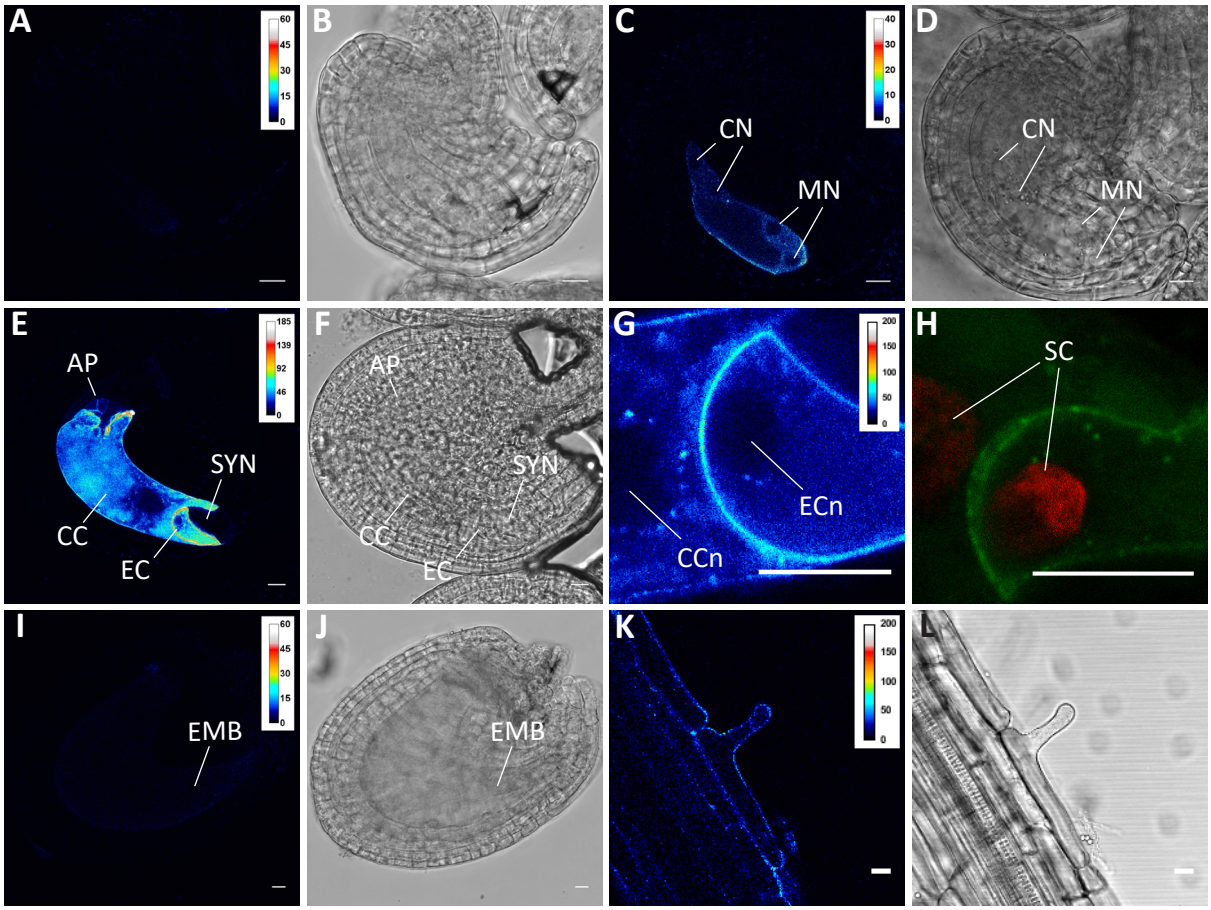


Figure 2-37 Localization of TET9-GFP in *Arabidop*sis TET9-GFP gene:reporter lines.
TET9-GFP was determined absent in ovule stage 3-II (**A, B**), appeared in stage FG4 (**C, D**), and was expressed strongest in the mature EC and CC and weakly in the antipodal cells (E, F). Upon gamete fusion no obious change in the TET9:GFP subcellular localization was observed: unpollinated (**G**), shortly after gamete fusion (**H**). Post fertilization TET9:GFP was absent of the ovule (I, J). TET9g:GFP was detected in roots (**K, L**). AP = antipodal cells, EC = egg cell nucleus, EMB = embryo, CCn = central cell nucleus, CN = chalazal nucleus, MN = micropylar nucleus, SC = sperm cell, SYN = synergid cell. Scale is 10µm.

CC and weakly in the antipodal cells (Figure 2-37 E, F). Pollination studies of TET9-GFP pistils with a red SCs marker line (Ingouff *et al.,* 2007) could not provide evidence of an involvement of TET9-GFP in the process of double fertilization, based upon localization studies. Post fertilization, the fluorescent signal was absent in the embryo and the endosperm (Figure 2-38 I, J). Also in leaves and cotyledons no GFP-derived signal could be detected (not shown). However, roots did express TET9-GFP (Figure 2-37 K, L). In summary, in the male and female gametophytes TET7g:GFP localized to the SYN and the PT, TET8g:GFP localized to the EC, and TET9-GFP localized to the developing embryo sac, and predominantely to the EC and to the CC.

2-6-12 A *tet8 tet11* Double Knock has a WT-like Seed Set

To possibly assign a biological function to *TET7*, *TET8*, *TET9*, *TET11*, and *TET12* in female and male gametophytes a reverse genetic approach was pursued. Homozygous T-DNA insertion lines of *tet8* (SALK_136039) and *tet11* (SALK_109259) were obtained, and insertion loci were verified and mapped by sequencing to 402 and 453bp downstream of the start codon, respectively (Supplemental Figure 5-6). Homozygous *tet8* and *tet11* plants were crossed, and subsequently propagated into the F2 generation, which was selected for double homozygous individuals by PCR-based genotyping. By RT-PCR with *TET8*, *TET11*, *ACT2*, and *GAPC* of cDNAs from the WT, and the *tet8 tet11* pistil, as well as stamen, a full knock out for *TET8*, and *TET11* was verified in the *tet8 tet11* double homozygous line #8 (Figure 2-38 A). However, the relative

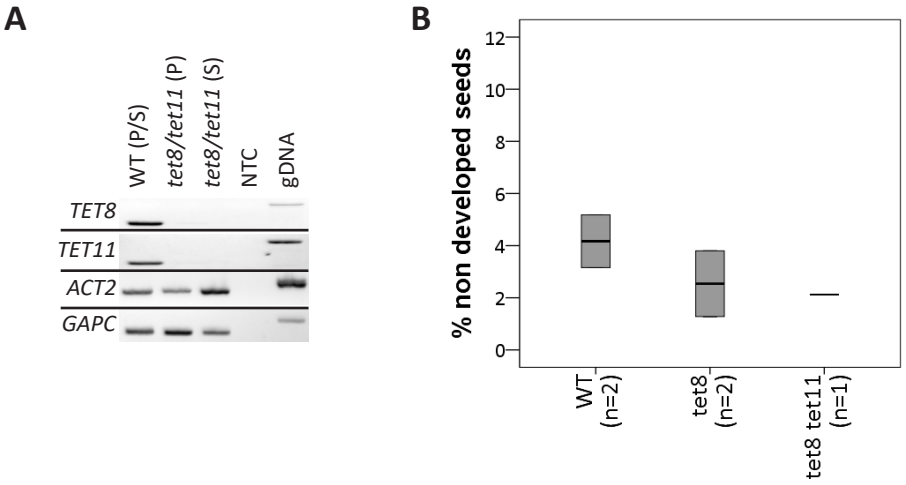


Figure 2-38 The double homozygous knock out line for *tet8* and *tet11*.
The double homozygous *tet8 tet11* line #08 was determined a *tet8 tet11* double knock out by RT-PCR of pistil, and stamen-derived templates, including *ACT2*, and *GAPC* as template controls (**A**). The relative non developed seeds of the homozygous *tet8* and the double homozygous knock out *tet8 tet11* were all determined WT-like. gDNA = genomic DNA, n = number of plants, NTC = non template control, P = pistil, S = stamen.

rate of non developed seeds per five siliques of two *tet8* plants, the *tet8 tet11* line #8, and two WT plants did not show obvious differences in seed set (Figure 2-38 B).

2-6-13 Simultaneous Tetraspanin Knock Outs by the CRISPR/Cas9 System

As the *tet8 tet11* double knock exhibited a WT-like phenotype, and neither for *TET9* nor *TET12* suitable T-DNA insertion lines were available, the CRISPR/Cas9 system was employed to simultaneously knock out a combination of gametophytic-expressed tetraspanins. Since *TET9* and *TET12* were also expressed by the gametes and to circumvent complementation effects caused by functional redundancies, the double homozygous *tet8 tet11* mutant was transformed with a CRISPR/Cas9 construct for simultaneous editing of *TET9* and *TET12* in *tet8 tet11* mutant background. Also, WT was transformed with a CRISPR/Cas9 construct for simultaneous editing of *TET7* and *TET12.* The target loci of 30 obtained T1 plant lines for putative *tet9* and *tet12* knock outs in the *tet8 tet11* mutant background, and of six T1 plant lines with putative *tet7* and *tet12* knock outs in WT background were sequenced.

Three homozygous *tet7* plant lines, and one homozygous *tet12* mutant were identified in genetic WT background (Table 2-12). Furthermore, two homozygous *tet12* mutations, and one heterozygous *tet12* mutation, as well as one homozygous *tet12* heterozygous *tet9* mutation were identified in the *tet8 tet11* genetic background (Table 2-12). The respective mutated DNA sequences for *TET7*, *TET9*, and *TET12* were translated into proteins to investigate the genome editing effects on protein level in regard to the translational frame (Table 2-12). For *TET7*, one mutant allele was recovered with a deletion of one amino acid after position 26, while the

Table 2-12 Obtained tetraspanin mutant plant lines by CRISPR/Cas9-mediated genome editing.
Obtained plant lines in the WT or the *tet8 tet11* double homozygous mutant background. The identified positions of insertions or deletions relative from start codon, and the type of mutation on DNA level are indicated in brackets. The resulting effect onto the translational frame downstream of the mutated locus is also depicted. For line 1 and line 6 each, two mutant alleles were obtained.

plant	genetic background	CRISPR genotype	mutation	effect on translation
1	Col-0	*tet7* (79+T); (79+C)	insertion	**frameshift**
2	Col-0	*tet7* (78-3nt)	deletion	in frame
5	Col-0	*tet7* (78-5nt)	deletion	**frameshift**
6	Col-0	*tet12* (321+A); (321+T)	insertion	**frameshift**
13	*tet8 tet11*	*tet12* (319-42nt)	deletion	in frame
18	*tet8 tet11*	*tet12* (319+T)	insertion	**frameshift**
25	*tet8 tet11*	*tet12* (311-8nt) *tet9* $^{+/-}$ (98-A)	deletion, deletion	**frameshift, frameshift**
30	*tet8 tet11*	*tet12* $^{+/-}$ (319-3nt)	deletion	in frame

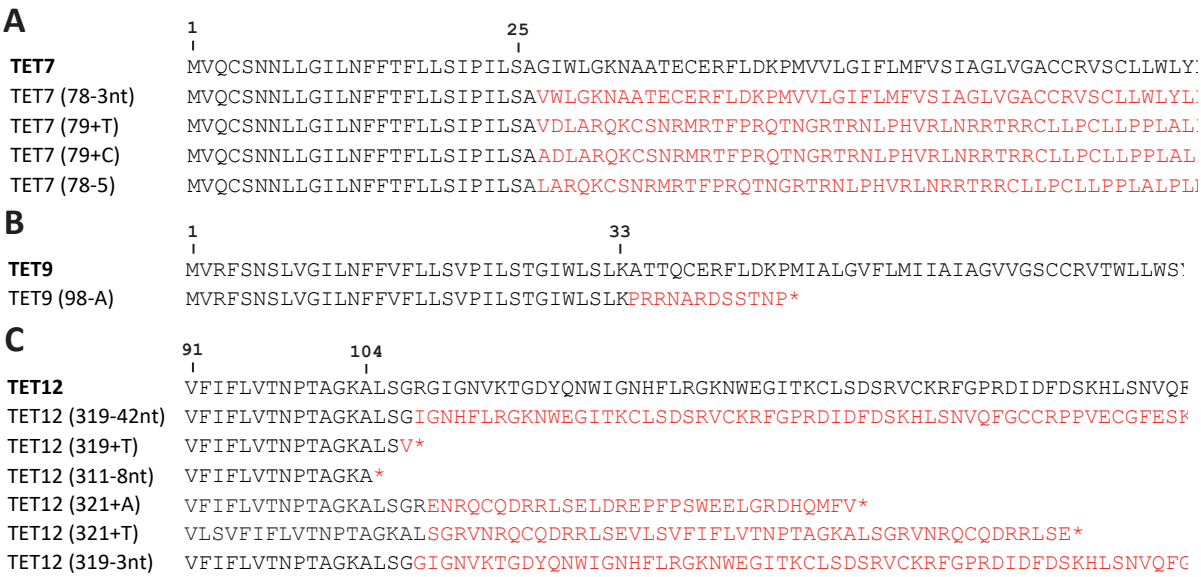


Figure 2-39 The TET7, TET8, TET9 protein sequences after CRISPR/Cas9-mediated genome editing of *TET7, TET8*, and *TET9*.
Alignments of WT protein sequences (depicted in bold letters), and the respective mutated protein sequences for TET7 (**A**), TET9 (**B**), and TET12 (**C**). Amino acid residues not matching the corresponding WT amino acid residue are depicted in red. Numers indicate amino acid positions relative from the translational start.

other three alleles lead to a frame shift, and severely altered protein sequences after position 26 (Figure 2-39 A). For *TET9*, one heterozygous mutant was obtained with a severely altered protein after position 33 and a premature stop codon after amino acid residue 45 (Figure 2-39 B). Six mutant alleles were obtained for *TET12*, resulting in two in frame deletions with 14, or one amino acid deleted after position 107. In total, four mutant alleles with translational frame shifts in TET12, leading to severely altered and truncated proteins after the amino acid positions 104, 106, and 108, respectively had been generated (Figure 2-39 C).

From this set of mutant plant lines a tetraspanin quadruple mutant was identified. The *tet8 tet11* T-DNA double knock out plant also contained a homozygous mutation in *tet12* resulting in a truncated TET12 protein after position 104, and a heterozygous, severe mutation in *TET9.* Thereby this *tet8 tet11 tet12 tet9*$^{+/-}$ quadruple mutant was of special interest for subsequent analysis.

2-6-14 Seed Sets of CRISPR/Cas9-mediated Tetraspanin Knock Out Lines

Phenotypic analyses were conducted on developing siliques of the tetraspanin CRISPR/Cas9 mutant plant lines. The relative rate of non developed seeds for homozygous *tet7* single mutations in the WT background showed almost full seed set, with a median value of 2% (*tet7* #01), or a high variation from 0% to 47% of non developed seeds, but a median value of 2%

(*tet7* #02) (Figure 2-41 A). Also, a control plant line without any occurred CRISP/Cas9-mediated genome editing event in the target locus (*TET12 TET7* #04) showed also a high variation from 4% to 41% of non developed seeds with the corresponding median value of 12% (Figure 2-40 A). Lastly, a homozygous *tet12* single mutation plant line (tet12 #06), had 2% to 32% non developed seeds, with a median value of 6% (Figure 2-40 A).

Analogous investigations were made for CRISPR/Cas9 plant lines in the *tet8 tet11* mutant background. For six control plant lines (#06, #07, #09, #10, #11, #20) without CRISPR/Cas9-mediated mutation of the target loci the median values for non developed seeds ranged from 1% to 23%, while the plant lines #13, and #18, which were found mutated in *tet12* loci in the *tet8 tet11* mutant background, displayed median values of 44% and 41% of non developed seeds (Figure 2-40 B). Notably, the homozygous *tet8 tet11 tet12* and heterozygous *tet9* quadruple mutant plant line #25 also had a reduced seed set with a median value of 44% non developed seeds (Figure 2-40 B).

In summary, when compared to the respective control plant lines the *tet7*, and *tet12* single knock outs seemed to have no effect on the seed set. The observed high variations in seed set were likely caused by ongoing pest control problems and thereby heavily stressed plants. However, *tet8 tet11 tet12* plant lines showed significantly reduced seed sets, although the observed rate of 44%, and 41% of non developed seeds did not correlate with the triple homozygous mutant genotype.

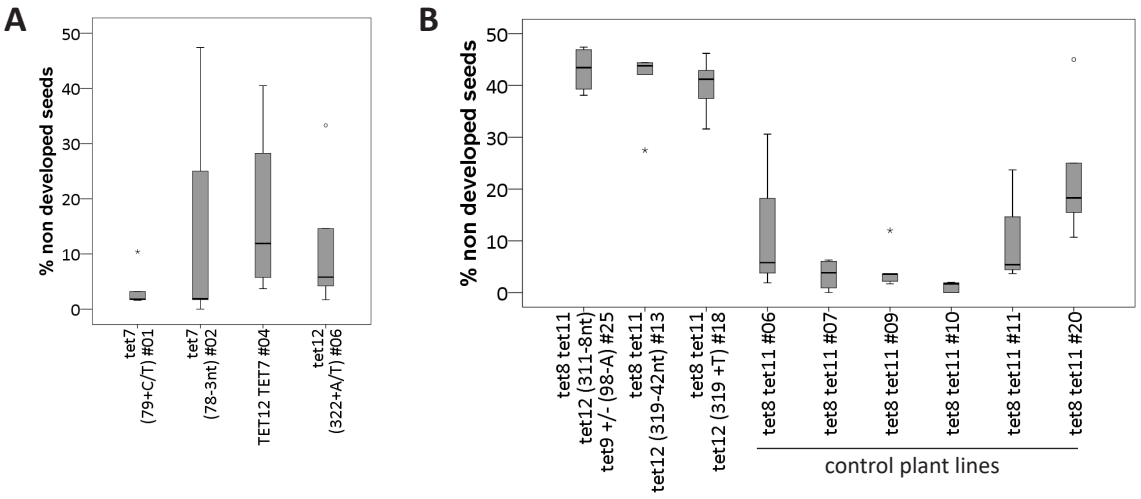


Figure 2-40 Seed sets of tetraspanin mutant CRISPR/Cas9 plant lines.
Relative rates of non developed seeds of the *tet7* mutant plant lines #01, #02, and the *tet12* mutant plant line #06 in the WT genetic background, and one mutant control plant line without mutated target loci #04 (**A**). Relative rates of non developed seeds of obtained CRISPR/Cas9 mutant plants in the *tet8 tet11* genetic background, with mutated genome loci for *tet12* and *tet9,* and non mutated loci.

2-6-15 Female Gametophytes of *tet8 tet11 tet12* Mutant Lines Arrest in FG1

The two obtained homozygous *tet8 tet11 tet12* triple mutant lines (#13, #18) exhibited, in contrast to the acquired *tet12* single mutant line #06, a severely reduced seed set. To elucidate the cause for this phenotypic abnormality the triple mutant lines #13 and #18 were analyzed by Feulgen staining of floral stage 12b/c pistils. This analysis revealed developmentally delayed FGs in stage FG1 and FG2 in both mutant plant lines #13, #18, aside developmentally unremarkable mature ovules (Figure 2-41 A – D).

In total, in floral stage 12b/c pistils of the homozygous triple mutant plant lines #13, and #18, 41% and 50% of the FGs remained in stage FG1 or FG2 while 54% and 43% had reached the mature developmental stage, respectively (Figure 2-42 A, B). Both plant lines had been determined homozygous for all three mutations but seed set data and the observed FG developmental stages indicated a rather heterozygous gametophytic effect caused by the insertion of the CRISPR/Cas9-containing T-DNA than a *tet8 tet11 tet12* mutation-dependent

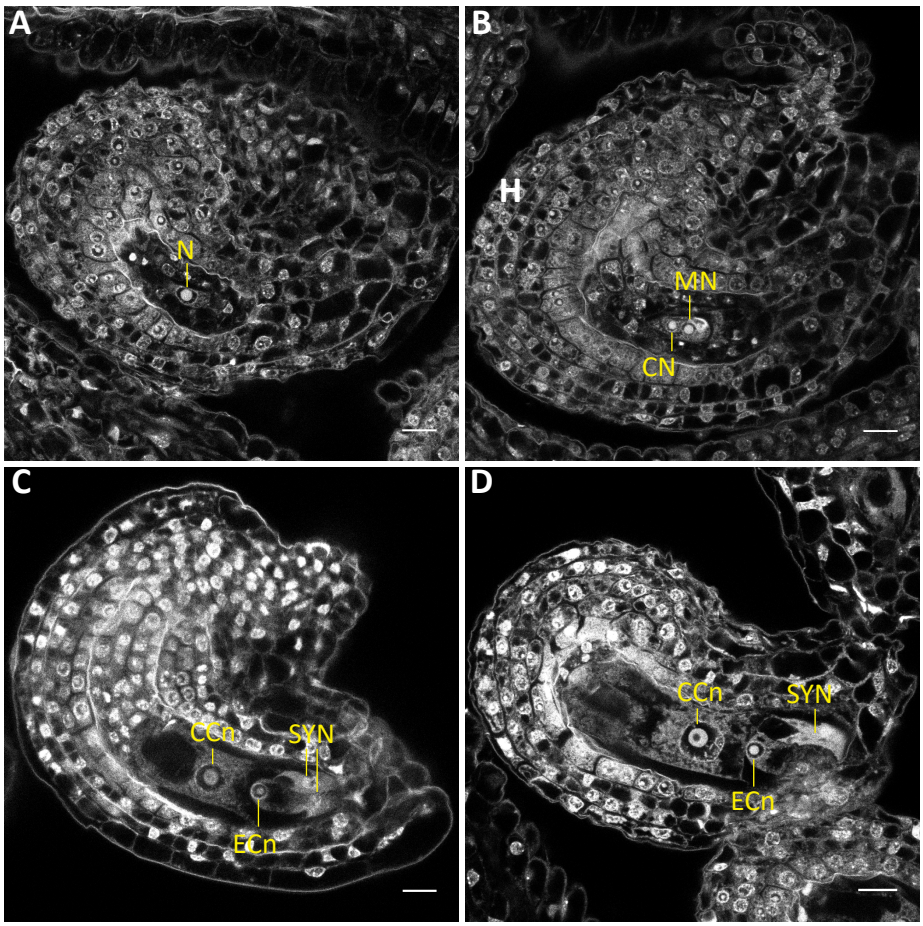


Figure 2-41 Feulgen staining of floral stage 12b/c pistils of *tet8 tet11 tet12* plant lines.
Arrested ovules in stage FG1 (**A**) and stage FG2 --(**B**), and mature ovules in FG6/7 (**C, D**) of floral stage 12b/c pistils of the *tet8 tet11 tet12* lines #13 (**A, C**), and #18 (**B, D**), respectively. CCn = central cell nucleus, CN = chalazal nucleus, ECn = egg cell nucleus, MN = micropylar nucleus, N = functional megaspore, SYN = synergid cell. Scale is 10µm.

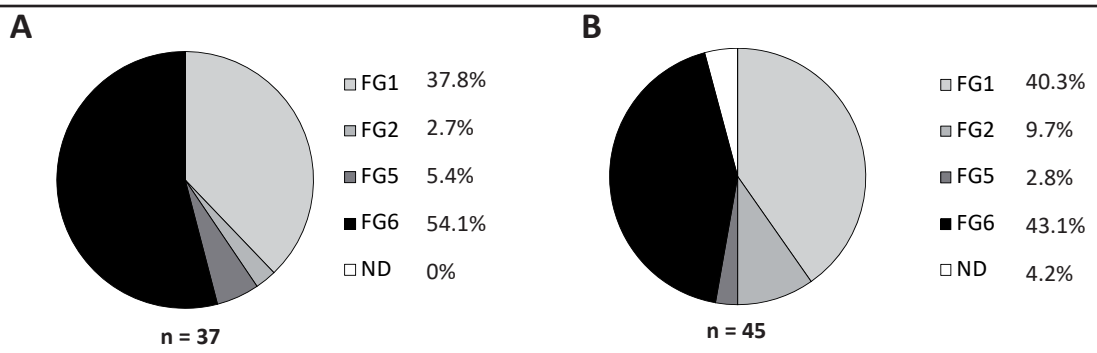


Figure 2-42 Observed female gametophyte developmental stages in floral stage 12b/c pistils of *tet8 tet11 tet12* plant lines #13 and #18.
Observed female gametophyte developmental stages in floral stage 12b/c pistils of homozygous *tet8 tet11 tet12* plant lines #13 (**A**), and #18 (**B**), both showed accumulation of arrested female gametophytes in stage FG1. n = number of ovules, ND = not determined.

cause for this phenotype.

2-6-16 Female Gametophytes of the *tet8 tet11 tet12 tet9*$^{+/-}$ Mutant Line #25 Arrest in FG4

To determine whether the reduced seed set in *tet8 tet11 tet12 tet9*$^{+/-}$ quadruple mutant plant line #25 was caused by a defective FG development, floral stage 12b/c and 14 pistils of this quadruple mutant, and the WT were subjected to ovule clearing and DIC microscopy, as well as Feulgen staining. In the early stage FG4, the chalazal and micropylar nuclei of the developing embryo sac appeared smaller in the mutant plant line #25 than in the WT (Figure 2-43 A, B). In rare cases only one micropylar nucleus was observed in mutant ovules (three-celled embryo sac; Figure 2-43 A, B). During late FG4 stage, when the chalazal nuclei rearrange prior to cellularization, the nuclei in the quadruple mutant still had not enlarged and were hard to identify (Figure 2-43 C, D).

By investigating ovules and developing seeds of floral stage 14 pistils via clearing and DIC microscopy, only three major developmental stages were identified: (i) WT-sized ovules containing embryo sacs without visible nuclei; (ii) fully cellularized FG6/FG7 stage ovules with visible EC, CC, and SYNs; (iii) early seeds with visible syncytical endosperm and developing embryo (Figure 2-44 A – C). To gain a better subcellular resolution, Feulgen staining was performed of floral stage 12b/c pistils of the *tet8 tet11 tet12 tet9*$^{+/-}$ quadruple mutant plant line #25, and the WT. In ovules of the quadruple mutant plant line #25, in rare cases only one micropylar nucleus was observed, and the nuclei of the developing embryo sac appeared smaller and disorganized, when compared to WT stage FG4 where gametophytic nuclei

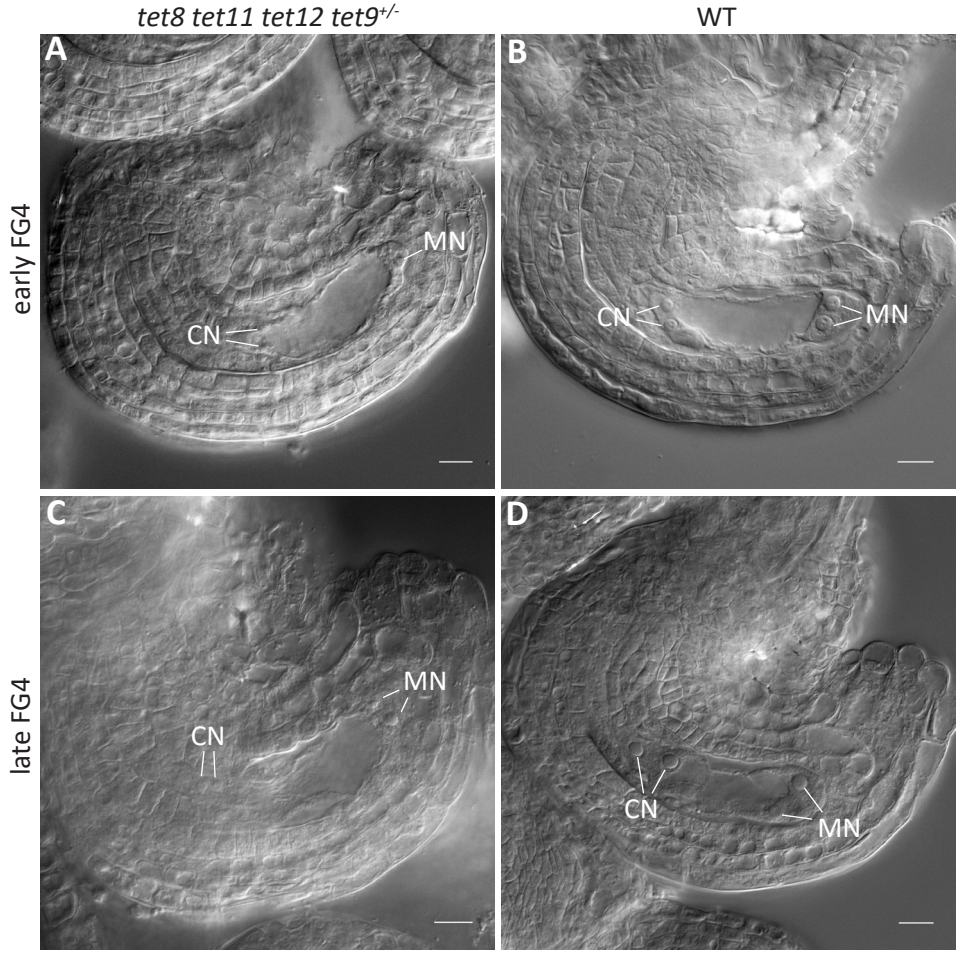


Figure 2-43 Clearing of ovules of the *tet8 tet11 tet12 tet9*$^{+/-}$ mutant line and the WT.
Cleared ovules of *tet8 tet11 tet12 tet9* +/- plant line #25 (**A, C**), and the WT (**B, D**) in early (**A, B**), and late (**C, D**) FG4 developmental phase revealed smaller nuclei in the female gametophyte of the mutant plant line. CN = chalazal nuclei, MN = micropylar nuclei. Scale is 10µm.

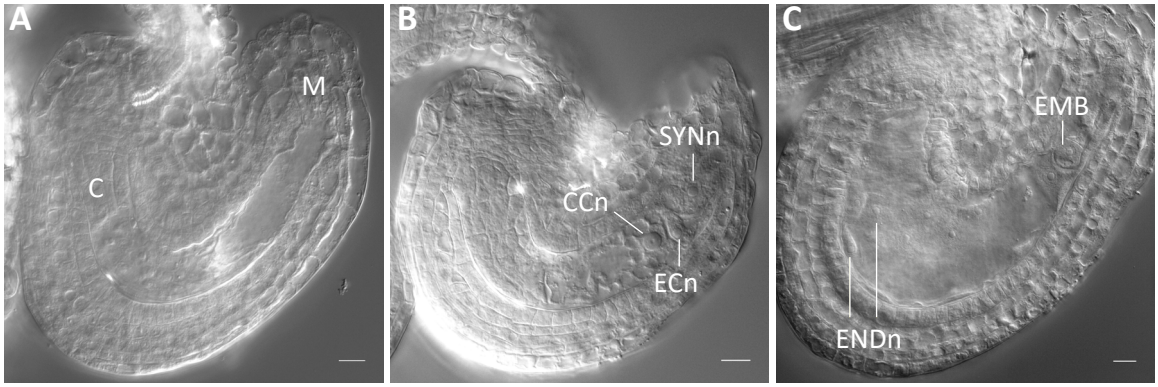


Figure 2-44 Ovules and early seeds in floral stage 14 pistils of the *tet8 tet11 tet12 tet9* +/- mutant line.
Cleared ovules of *tet8 tet11 tet12 tet9*$^{+/-}$ plant line #25 of floral stage 14 pistils showed either empty embryo sacs (**A**), fully developed mature female gametophytes (**B**), or early seeds (**C**). C = chalazal pole, CCn = central cell nucleus, ECn = egg cell nucleus, EMB = embryo, ENDn = endosperm nucleus, M = micropylar pole, SYNn = synergid cell nucleus. Scale is 10µm.

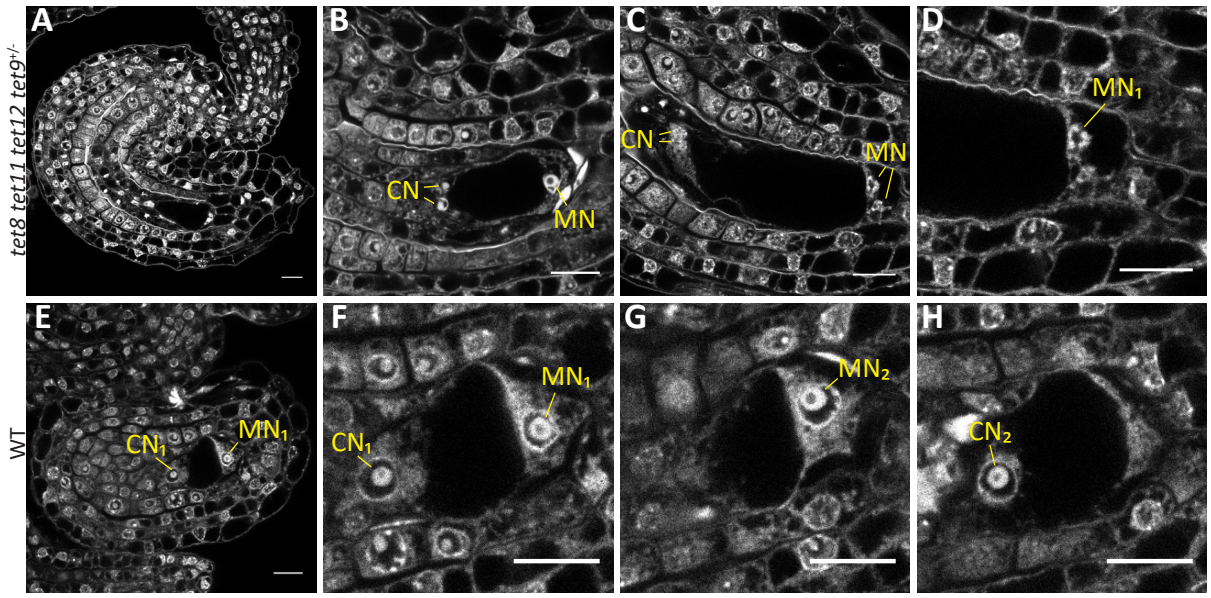


Figure 2-45 Feulgen staining of ovules of the *tet8 tet11 tet12 tet9*$^{+/-}$ plant line and the WT.
Optical sections through Feulgen-stained ovules of *tet8 tet11 tet12 tet9*$^{+/-}$ plant line #25 (**A - D**) and the WT (**E - H**). In comparison to the WT (**E - H**), in the mutant line, disrupted mutant nuclei (**C, D**) and rarely only one micropylar nucleus (**B**) were detected. (**D**) is a magnification of (**C**). CN = chalazal nuclei, MN = micropylar nuclei. Scale is 10µm.

maintain a more circular form (Figure 2-45 A – H).

The clearing-based (Figure 2-46 A), and the Feulgen staining-based (Figure 2-46 B) analyses of floral stage 12 b/c pistils of the quadruple mutant plant line #25 had determined at least 44% and 34% of the FG arrested in stage FG3 to FG4, respectively. Only 41% and 42% of the FGs had reached maturity (Figure 2-46 A, B). When compared to the WT pistils of floral stage 12b/c an obvious difference came to light as here only 2.5% of the ovules remained in stage FG4, and 81% had reached the mature ovule stage 3-VI (Figure 2-46 C). Clearing of floral stage 14 pistils of the quadruple mutant plant line revealed 30% mature FGs, 28% developing seeds, and 35% of the FGs still in developmental stage FG3/4, without or barely visible nuclei within the immature embryo sac (Figure 2-46 D). In summary, phenotypic analysis of developing ovules of the homozygous *tet8 tet11* plant line with a CRISPR/Cas9-mediated homozygous *tet12* mutation and a heterozygous *tet9* mutation had shown disintegrated, smaller nuclei in non cellularized embryo sacs in 35% to 42% of the FG, which resembled earlier findings from the relative rates of non developed seeds (Chapter 2-6-14).

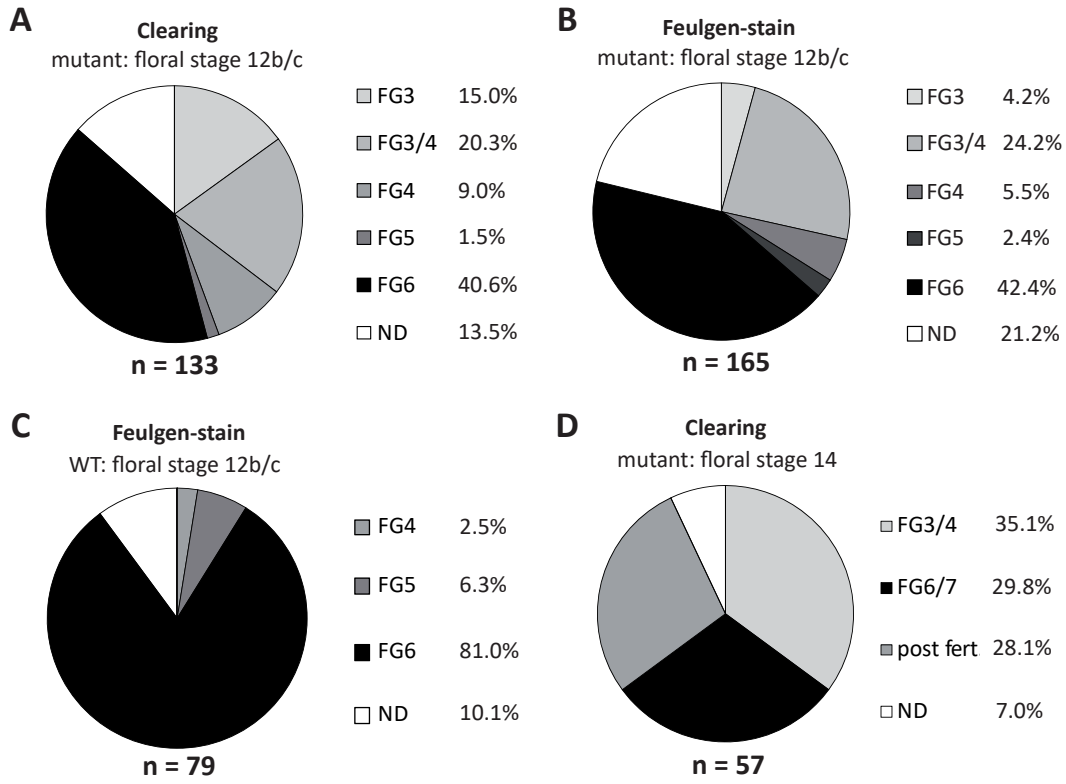


Figure 2-46 Observed female gametophyte developmental stages in ovules of the mutant line with the *tet8 tet11 tet12 tet9* +/- genotype and the WT.
Determined female gametophyte developmental stages in *tet8 tet11 tet12 tet9* quadruple mutant plant line #25 (**A, B, D**), and WT (**C**) in floral stage 12c (**A, B, C**), and floral stage 14 (**D**) pistils. Showed approximately one third of the developing embryo sacs in the *tet8 tet11 tet12 tet9*$^{+/-}$ quadruple mutant arrested or dead in stage FG3 to FG4. n = number of ovules, ND = not determined.

3 – Discussion

3-1 Female Gametophytic Cells in the Focus: What We Know and What We Don't Know

The final stages of double fertilization are mediated by a number of proteins, located at the cell surface or the intercellular space. Integral membrane proteins like FERONIA (FER) and POLLEN RECEPTOR-LIKE KINASE6 (PRK6) and MALE DISCOVERER1 (MDIS1) and MDIS1-INTERACTING RECEPTOR LIKE KINASE1 (MIK1) and MIK2, GENERATIVE CELL SPECIFIC1/HAPLESS2 (GCS1/HAP2), or GAMETE EXPRESSED2 (GEX2) are essential for pollen tube (PT) perception, gamete attachment and gamete fusion, respectively (Johnson *et al.,* 2004; Mori *et al.,* 2006, Escobar-Restrepo *et al.,* 2007; Mori *et al.,* 2014; Takeuchi *et al.,* 2016; Wang *et al.,* 2016). In addition, small cysteine-rich secreted peptides (CRPs) are indispensible for intergamete-signaling as only upon triggered secretion of EGG CELL1.1 sperm cell attachment to, or fusion with, the female gametes can be executed (Sprunck *et al.,* 2012). Moreover, to mediate PT guidance synergid-secreted EGG APPARATUS1 and LURE1, are required in maize and *Arabidopsis*, respectively (Marton *et al.,* 2005, Okuda *et al.,* 2009, Takeuchi *et al.,* 2012). Furthermore, extracellular localized GPI-anchored proteins, like LORELEI, JAGGER, and members of the EARLY NODULIN-LIKE family, mediate reception of the PT (Capron *et al.,* 2008; Hou *et al.,* 2016; Pereira *et al.,* 2016). Also, protein N-glycosylation via endoplasmic reticulum-localized TURAN and EVAN is required for the PT reception pathway (Lindner *et al.,* 2015). Even though the picture of double fertilization in angiosperms is getting pieced together, there are many things we still don't know.

It was proposed there is signaling between the female, and between the female and male gametes required to prevent from single fertilization events and ensure double fertilization (Huang *et al.,* 2015). Cell ablation and mutant screens have shown that the synergid cell can assume egg cell identity (Gross-Hardt *et al.,* 2007; Pagnussat *et al.,* 2007; Moll *et al.,* 2008, Kirioukhova *et al.,* 2011; Kong *et al.,* 2015), which most likely involves mobile signals between the egg cell and the synergid cells.

Plant mobile macromolecules (Guan *et al.,* 2017) like small mobile RNAs (Ibarra *et al.,* 2012; Dunoyer *et al.,* 2013), or secreted peptides are interesting candidates to mediate such events. Moreover, the crosstalk between the maternal sporophyte and the embedded female

gametophyte (FG) is still elusive (Figueiredo *et al.,* 2016). Finally, the respective GEX2 receptor on the surface of the female gametes, or the receptor for the EC1 peptides remain yet to be discovered.

3-2 Bridging the Gap: Transcriptomics

The Affymetrix ATH1 121501 GeneChip serves as popular tool to acquire expression data but features only about 24000 genes and thereby no exhaustive analysis of the *Arabidopsis* transcriptome. In the past, FG defective mutants, like *determinant infertile1,* have served as tools for subtractive transcriptomic profiling of mutant ovules and wild type ovules (Johnston *et al.,* 2007; Jones-Rhoades *et al.,* 2007; Steffen *et al.,* 2007) in order to identify FG-specific genes. The combination of laser-assisted microdissection (LAM) with array hybridization and later RNA seq, enabled to acquire the first *Arabidopsis* transcriptome data of female gametophytic cells and central cells, respectively (Wuest *et al.,* 2010; Schmid *et al.,* 2012).

However, the latest efforts made have not addressed the secreted and cell surface-localized proteins of the FG, especially gametes. In a previous work of transcriptional profiling of manually microdissected (MM) live female gametophytic cells (Soljic 2012) 10065, 11641, and 8728 genes had been determined present in egg cell (EC), central cell (CC), and synergid cells (SYN), respectively. In this study, the transcriptomic data of Soljic 2012 was processed with another 59 hybridization data of sporophytic tissues. MDS plot analysis revealed good clustering of the probe sets and to the probe sets present in all replicates the respective AGIs were assigned: 11509, 12186, and 12044 genes were determined present in the EC, the CC and the SYN, respectively.

In comparison, the LAM of female gametophytic cells and the ATH1 GeneChip hybridizations had revealed 7171 EC-expressed genes, 7287 CC-expressed genes, and 5628 SYN-expressed genes (Wuest *et al.,* 2010). In addition to those genes, another 10320 CC-expressed genes were detected by LAM of CCs and subsequent RNA seq (Schmid *et al.,* 2012).

The main focus of this study presented here was to identify, in comparison to sporophytic tissues, differentially expressed genes (DEGs) encoding membrane-associated and extracellular localized proteins of the FG in regard to cell-to-cell communication and attachment.

Initially, the transcriptomic datasets of LAM-isoalted, and MM female gametophytic cells were

processed in a combined approach to obtain more robust data. However, GO term analysis of the resulting DEGs revealed presumably sporophytic-derived transcripts within the LAM-derived transcriptomic dataset and therefore this approach was discarded. Consequently, the analysis was performed on MM-isolated female gametophytic cells and transcriptomic data of *Arabidopsis* sperm cells, isolated by fluorescence-activated cell sorting (FACS) (Borges *et al.,* 2008). This combined analysis revealed 19259 transcripts above cut off within the EC, the CC, the SYN, and the SCs. Of these 19259 genes, 14056 genes were sperm cell-derived (59% of ATH1 GeneChip represented genes) which was an unexpected number. In comparison, the initial analysis by Borges *et al.,* 2008 identified 5829 sperm cell-expressed genes (24% of represented genes). Array hybridizations using the Affymetrix Rice Genome Array had identified 10732 genes (22% of represented genes) expressed in rice SCs (Russell *et al.,* 2012). As this poses a surprising difference, the threshold values of the respective cutoff values need to be adjusted and validation experiments on transcript level should be performed by using isolated *Arabidopsis* sperm cells.

To determine differential expression, the expression values of the female gametophytic cells and the male gametes were contrasted with each other and with a selection of sporophytic tissues. Mitochondrial and chloroplastidial-derived gene identifiers were removed from the analysis finally resulting in 7168 DEGs (compared to the sporophytic tissues), encoding for membrane-associated or secreted proteins of which 2438 DEGs were determined enriched with a FC ≥ 1.5.

3-3 What's the Hold-up? Data Validation!

Cell-type specific expression of selected DEGs within the FG was verified by promoter:reporter-expressing plant lines and *in situ* hybridizations. Generating stable *Arabidopsis* lines to express promoter:reporter constructs was proven unsatisfactory as for certain genes detectable GFP-based fluorescence was absent. Gene:reporter fusions of these genes showed the expected signals though. These results may be caused by the fact that an intact exon-intron structure can affect transcription (Rose *et al.,* 2008; Karve *et al.,* 2011) and an intact intron-exont structure can be more important than proximal promoter fragments in determining the site of transcription initiation (Gallegos *et al.,* 2017).

However, if the promoter sequence proves functional and drives the reporter transcription, this rigid system cannot account for cell to cell mobility of specific transcripts (Thieme *et al.,* 2015). Moreover, it introduces a noticeable delay due to reporter protein stability. Additionally, T-DNA position effects can lead to false positive results (Thompson *et al.,* 2001; Schubert *et al.,* 2004). On the other hand, promoter-driven expression of fluorescent reporter proteins offer powerful tools to monitor biological processes *in vivo* (Xiao *et al.,* 2010; Denninger *et al.,* 2014) and should not be neglected all together.

For validation of the transcriptome data with cellular resolution, *in situ* hybridization was found superior, as it enabled for a higher throughput than promoter:reporter lines, investigations were conducted in the genetic WT, and transcripts were localized directly. Interestingly, in some cases the localization of the transcripts was shared by less cells of the FG than expected according to the bioinformatic analysis. This applied only to the cell types with low-expressed genes. In order to visualize the transcripts of these genes, prolonged staining procedures may be required in some cases. Weak transcripts that cannot be visualized all together might, however also indicate too low cutoff values during bioinformatic processing. Otherwise, the cells of the FG are known to acquire cell fate in dependence of each other (Tekleyohans *et al.,* 2017), thereby immature collected cells may express residual transcripts that are, in the fully mature gametophyte, restricted to only one cell type. Lastly, prolonged enzyme treatment during the isolation procedure can cause plasmogamy between isolated ECs, CCs, and SYNs (Englhart *et al.,* 2017), cause stress-induced gene expression or possibly disturb cell fate repression as the cells get separated from each other within the isolation buffer.

Additionally, brightfield or DIC microscopy proved disadvantageous as the CC partly engulfs the cells of the egg apparatus and thereby in unfortunate oriented ovules does not allow for distinguishing of the diffuse staining between the individual cell types. Here, employing fluorescent *in situ* probes allows for acquiring optical sections by confocal microscopy (Levsky *et al.,* 2003). Moreover, high affinity LNA probes can efficiently hybridize short sequences in case of alternatively spliced exons (Darnell *et al.,* 2010), or untranslated regions to prevent from cross-hybridization.

3-4 Gametophytic Cells Show Distinct Expression Profiles of Genes Encoding Cell Surface Proteins

MDS plot analysis of all detected transcripts above cutoff determined the female gametophytic cells close to each other and much closer to sporophytic tissues like the leaf than to the SCs, which clustered far away from all other tissues and cell types. This resembles findings of principal component analysis of rice gametes (Russel *et al.,* 2012). Although the female gametophytic cells had been clustering together, differences were observed when the DEGs encoding for gametophyte-enriched putative membrane-associated, or extracellular localized proteins (excluding transcripts with plastidial-derived gene identifier) were subjected to GO term analysis in regard to enriched biological processes.

The SYN serves as a specialized “transfer cell”, that acquires the function of a glandular cell to produce and secrete substances to direct the growth of the PT (Van Went 1970; Okuda *et al.,* 2009; Dresselhaus *et al.,* 2013). This secretory function of the SYN is reflected by the finding that the biological processes encoded by the SYN-derived transcripts are protein transport, glycosylation, and exocytosis for peptide secretion. This in turn is supported by the SYNs extensive microtubule and actin network that is oriented towards the filiform apparatus (Webb *et al.,* 1994), suggesting polar vesicle transport to the micropylar pole of the SYN.

In *Arabidopsis*, PT attraction is not exclusively mediated by the SYN but also by the CC, which contributes by supplying a CBP1-mediated transcriptional landscape required for pollen tube attraction (Li *et al.,* 2015). Furthermore, the enriched biological processes found in the CC-derived transcriptome data, when compared to the other female gametophytic cells and male gametes, suggest that the CC expresses transcripts that encode proteins required for the generation of precursor metabolites and energy. This could serve as preparation for nourishing the embryo and the rapid growth of the endosperm post fertilization. Once fertilized, the CC undergoes numerous nuclear divisions in quick succession(Brown *et al.,* 1999) – a process that was also observed in a fertilization-independent manner in a number of mutants (Chaudhury *et al.,* 1997) (Figueiredo *et al.,* 2015). Peculiarly, that, in gametophytic comparison, the CC was depleted in mitosis-related biological processes though.

Lastly, after sporophytic contrasting, the EC shared enriched biological processes with the CC and the SYN, like protein transport related processes, but was in direct gametophytic

comparison not found depleted in any biological processes.

3-5 I'm Gonna Send Them to Outta Space: Secreted Proteins of the Female Gametophyte

Signal transduction between non attached cells relies on mobile signals, which range from, e.g phytohormones, to sugar moieties like AMOR (Mizukami *et al.,* 2016), and secreted small proteins or peptides like EA1 or the CLE family (Marton *et al.,* 2005; Fiume *et al.,* 2011). The dataset of female gametophytic cells features numerous embryo sac-secreted, or cell surface localized proteins like the RAPID ALKANIZATION FACTOR-LIKE (RAFL) peptides (Murphy *et al.,* 2014). Own transcriptome data revealed that RALFL28 is enriched in the FG cells, with strong expression in the CC and the SYN, while *RALFL14* is preferentially expressed in the SYN and *RALFL18* is enriched in the CC and the SYN (Figure 3-1).

Furthermore, WISH analyses for *RALFL28* validated the expression of *RALFL28* in the FG cells. In *Arabidopsis*, there are 39 RALFL genes, *RALFL14* and *RALFL18* belong to clade IV, and *RALFL28* to clade I (Cao *et al.,* 2012). RALFL-proteins are linked to, e.g inhibition of cell elongation (Pearce *et al.,* 2001; do Canto *et al.,* 2014), or, in *Solanum*, to cell polarity establishment of the embryo sac (Chevalier *et al.,* 2013). It is possible that, unlike to clade I to III, clade IV proteins are not true RALFs, just RALF-related (Campbell *et al.,* 2017). Even though clade IV is overrepresented in *Brassicaceae* (Campbell *et al.,* 2017), only *RALFL8* has been analyzed and linked to stress response in *Arabidopsis* roots (Atkinson *et al.,* 2013). Since it was shown that FERONIA (FER)

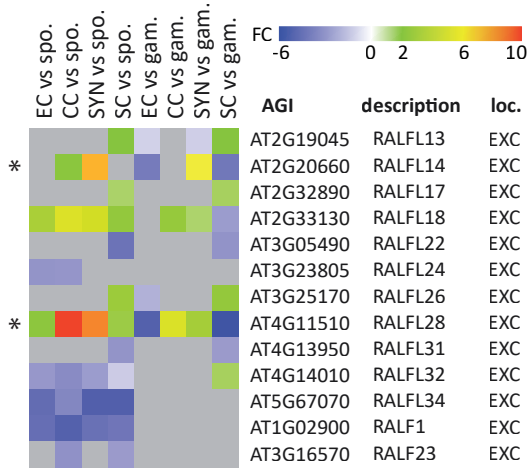


Figure 3-1 Differentially expressed members of the rapid alkanization factor-like gene family. Heatmap depiction of DEGs of the rapid alkanization factor (like). Asterisks indicates candidate genes. EXC = extracellular, FC = log2 fold changegam. = gametophyte, spo. = sporophyte.

interacts with RALF1 (Haruta *et al.,* 2014), and possibly other RALF peptides (Stegmann *et al.,* 2017), and FER overlaps in expression with the SYN-expressed RALFL14, it would be interesting to test a putative interaction of RALFL14 with FER.

Another transcript was found EC-specific and absent post fertilization by WISH: AT1G31450, encoding a predicted extracellular localized putative aspartyl protease. AT1G31450 and the closely sequence-related to *CONSTITUTIVE DISEASE RESISTANCE1* (*CDR1*) were by far (factor 10^4) strongest expressed in pistil tissue as determined by qPCR. Additional WISH experiments identified *CDR1* transcript in the CC and at the chalazal pole of the developing endosperm but not the EC (Supplemental Figure 5-7). Previous studies had shown CDR1 to be an apoplastic protease, with yet unknown endogenous targets for proteolytic cleavage (Xia *et al.,* 2004). However, sporophytic overexpression of *CDR1* had led to accumulation of salicylic acid and activation of defense genes accompanied with strictly restricted oxidative bursts to few cells in leaves and following cell death (Xia *et al.,* 2004).

It is known that programmed cell death is required during multiple stages of *Arabidopsis* seed formation like SYN and PT degeneration, the cell death of the antipodal cells and the nucellus cells post ferilizaton (Dominguez *et al.,* 2014; Xu *et al.,* 2016). Furthermore, CDR1s mode of action was located upstream of salicylic acid-mediated pathogen defense signaling by the proposed processing of a peptide elicitor (Xia *et al.,* 2004). Members of this strongly ovule expressed protein family could therefore contribute to a constantly heightened state of alertness of the ovule and developing seed regarding pathogen attack, or even facilitate cell death of certain cell types pre or post fertilization.

Among the DEGs, encoding secreted proteins, AT4G09090, a member of the family of Carbohydrate-binding X8 domain proteins was investigated. qPCR had determined AT4G09090 strongly pistil expressed and gene:reporter line analysis verified the respective fusion protein to be secreted from the CC and localized in close proximity to and around the embryo sac. Another X8 domain containing protein, OLE-e10 from *Olea europea* (olive), had been shown to bind beta 1,3 glucan (callose) (Barral *et al.,* 2005). In plants, callose is very often found in plasmodesmata (PD), which are important regulators of symplastic cell-to-cell trafficking of micro- and macromolecules (De Storme *et al.,* 2014). Cell to cell signaling via PD can be restricted by callose deposition in the PDs, which can also restrict pathogen movement

(Li *et al.,* 2012). Also, clogging of the PD with callose can enable the symplatically isolated cells to develop less affected from the surrounding cells (De Storme *et al.,* 2014). However, plasmodesmata-localizing proteins are known to localize in specles (Vaddepalli *et al.,* 2014), which was not the case for AT4G09090-GFP. Even though the X8 domain does not have an enzymatic function in a similar PD-localizing but GPI-anchored protein (Simpson *et al.,* 2009), callose binding of AT4G09090 might be sterically hindered by the C-terminal GFP and thereby AT4G09090:GFP may be mislocalizing. Moreover, AT4G09090 also does not have a predicted GPI-anchor sequence and it might serve yet a different function around the embryo sac with lesser requirement for a PD-restricted localization. Unfortunately, reverse genetics by AT4G09090 T-DNA plant lines was hampered by functional redundancy and overlapping gene expression patterns.

3-6 SSPR, Ready for Development or Defense?

SPOROZOITE SURFACE PROTEIN-RELATED (SSPR) is an uncharacterized single copy gene. The expression of SSPR-GFP fusion protein under control its endogenous promoter revealed that SSPR is heavily secreted from the embryo sac and during endosperm development. Homology modeling using SWISS-MODEL revealed structural similarities to the 2yip.3A-L beta sheet structure of *Sarcocystis muris* microneme protein 2 (SML-2; Figure 3-2 A, B). Since SML-2 is a galactose/galactosamine-specific lectin (Muller *et al.,* 2011), lectin properties of SSPR are likely. *Sarcocystis* is an intracellular parasite in mammals, that gains entry into cells by secretory organelles (micronemes) which are filled with proteins mediating cell-surface attachment to the target cells. Similar mechanisms are utilized by *Plasmodium*, which harbors the sporozoite surface protein precursor 2 (SSP2), a thrombospondin-related anonymous protein (TRAP) (Rogers *et al.,* 1992).

To determine where and if *Arabidopsis* SSPR contains, apart from an N-terminal secretion signal, any functional domain a KEGG-based motif search of the SSPR sequences was conducted. Consequently, one putative salp15 motif was identified in the SSPR sequence, spanning from the amino acid positions 40 to 70 (Figure 3-2 C). Salp15 is a protein which is only found in *Iodes scapularis* (tick) saliva (Anguita *et al.,* 2002), and known to bind a non linear region of the D1 domain of the human T-cell receptor CD4 (Juncadella *et al.,* 2008). The variable D1

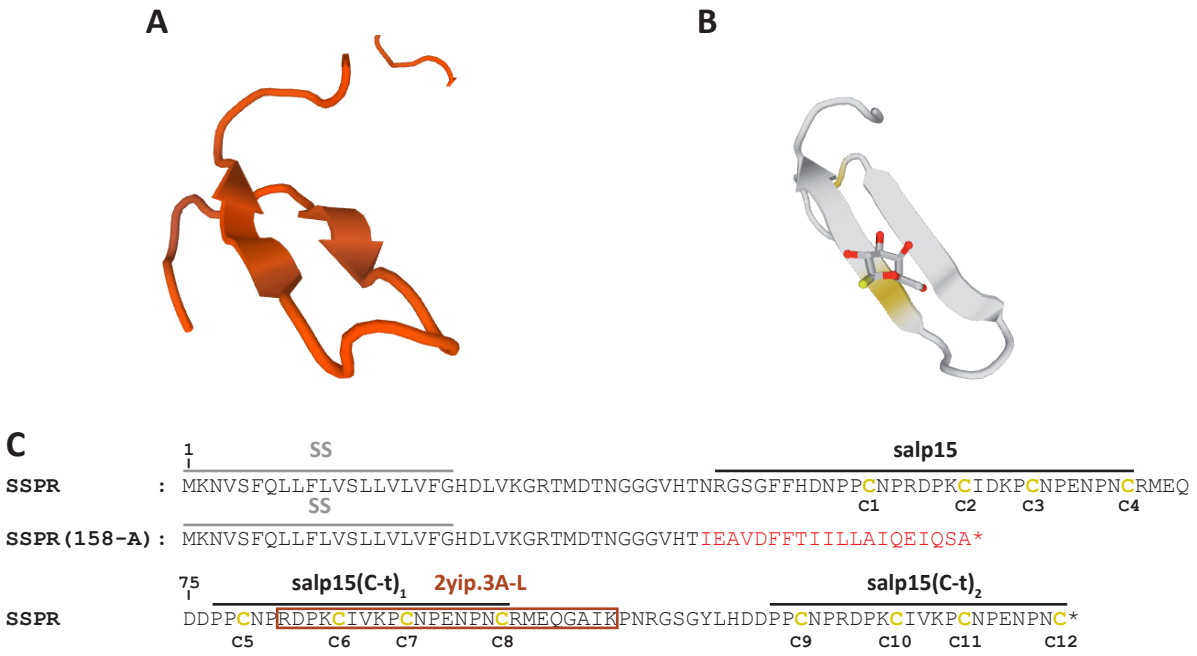


Figure 3-2 *Arabidopsis* SSPR protein sequences, functional and structural domains.
Homology modelling of SSPR (**A**) revealed a structural homology to the 1-Thio-beta-galactose binding domain (2yip.3A-L) of *Sarcocystis muris* protein 2 (SML-2) (**B**). The *Arabidopsis* SSPR WT sequence, with predicted functional and structural domains, and the CRISPR/Cas9-mutated SSPR sequence. Altered SSPR amino acid residues, caused by a translational frame shift, are depicted in red letters. The predicted secretion signal, the salp15 motif, and two truncated salp15 motifs (C-t) are indicated. The 2yip.3A-L corresponding region of SML-2 is boxed in brown and partially overlaps with the salp15 motif sequence. C = conserved cysteine position, SS = secretion signal.

domain of CD4 consists of Ig-like domains (Maddon *et al.,* 1985; Ryu *et al.,* 1990). Upon salp15-CD4 binding the CD4 receptor heterodimerization is prevented, the Src tyrosine kinase p65lck-dependent signaling cannot be initiated, and consequently CD4(+) T-cell activation via calcium mobilization is repressed (Anguita *et al.,* 2002). The SSPR sequence features two additional truncated salp15 motif sequences that are comprised of the 22 most C-terminal amino acids of salp15 (Figure 3-2 C). The binding of salp15 to CD4 was shown to heavily rely on the 20 most C-terminal localized amino acids of salp15, which contain four conserved cysteine residues (Juncadella *et al.,* 2008), that are also shared by all three SSPR salp15 motifs.

Furthermore, within the amino acid sequence of SSPR, the sequence of salp15 motif partly overlaps with the sequence of the galactose-binding beta sheet structure of 2yip.3A-L of SML-2. Glycosylation of the extracellular domains of the human T-cell receptor CD4 is restricted to the D3 and D4 domains of CD4, though (Maddon *et al.,* 1985; Ryu *et al.,* 1990). Thereby, the binding of salp15 to the D1 domain of CD4 is most likely achieved differently.

Notably, the HIV1 envelope glycoprotein 120 (gp120) also binds the D1 domain of the CD4 receptor, which then triggers a conformational change in gp120 and initiates virus-host

membrane fusion (Ryu *et al.,* 1990; Moore *et al.,* 1992). In a competitive assay, salp15 reduced the binding of gp120 to CD4 and it was shown that the 20 most C-terminal amino acids of salp15 bound to a conserved region of gp120 (Juncadella *et al.,* 2008). This region corresponds to a β-sheet structure in gp120 that faces towards CD4 in the CD4-gp120 bound state and thereby salp15 putatively hampers virus-host fusion via binding to CD4 and gp120 (Juncadella *et al.,* 2008).

Since SSPR combines multiple salp15 domains and their Ig-like domain binding properties, with the structural homology to the galactose binding domain of SML-2, it became an attractive idea to study candidate to study the function of SSPR in the *Arabidopsis* FG.

The homozygous *sspr* mutant (158-A) line #04 generated by CRISPR/Cas9-mediated gene editing encoded for a severely truncated version of SSPR lacking the salp15 motifs (Figure 3-2 C). The residual protein would thereby be non-functional. Furthermore, the seed set of this sspr mutant line #04 was reduced by 43%, with the corresponding FGs arrested in stage FG1 and FG2 (29%), as well as in stage FG3 and FG4 (20%). The SSPR expression data in the ovule correlates with these early stages of FG development. However, an essential SSPR function is not very likely as a higher penetrance of the observed phenotype would be expected in a homozygous mutant of this single copy gene. Nevertheless, SSPR may still have a defense-related function in the FG by inhibiting viral cell fusion or it may yet suppress signaling cascades, like tick salp15 does, by blocking the extracellular domain of a receptor within the ovule until SSPR expression is abolished in later seed developmental stages. To address the function of SSPR in the developing FG, it will be essential to generate and analyze more mutant CRISPR/Cas9 lines for SSPR. Lastly, if SSPR works as signaling molecule within the ovule and seed, and retains its lectin properties, TRICEPS (Frei *et al.,* 2012) is a promising method that offers the potential to identify the corresponding glycosylated cell-surface receptor.

3-7 I Sense a Disturbance in the Apoplast: Gamete-expressed RLKs

Membrane-associated RLKs are prime candidates to mediate the signaling events during gamete interactions and fusion, a process which is preceeded by Ca^{2+}-dependent signaling (Denninger *et al.,* 2014; Hamamura *et al.,* 2014). However, the underlying cis-, and more importantly trans-interactions between the individual gametes remain largely elusive. The transcriptome dataset

of female gametophytic cells features a large number of differentially expressed receptor (like) kinases (RLKs) in the EC, and the CC. From analysis of the transcriptome data of the 44 member family of cysteine-rich RLKs, four members were found strongly upregulated in the female gametes: CRK10, CRK23, CRK24, and CRK30, while CRK26 and CRK27 were found upregulated in the CC and the SYN, and the EC and the SYN, respectively (Supplemental Figure 5-1 A).

In qPCR experiments of seven to 21 day old seedlings the simulation of oxidative stress caused immediate upregulation of CRK10, CRK23, CRK27 within one hour, and within eight hours the upregulation of CRK26 (Wrzaczek *et al.,* 2010). Furthermore, light stress caused the upregulation of CRK30, and CRK24 expression was increased by the bacterial elicitor flg22 (Wrzaczek *et al.,* 2010). These CRKs contain a DUF26 with a conserved cysteine-motif (C-8X-2X-C), similar to the motif in GRIM REAPER which is known to be involved in ROS-associated cell death (Wrzaczek *et al.,* 2009; Wrzaczek *et al.,* 2010). Consequently, this DUF26 CRK family was proposed to act as reactive oxygen species (ROS) receptors of the apoplast, as cysteines are sensitive to redox modifications (Wrzaczek *et al.,* 2010).

ROS and Ca^{2+} are the fastest interconnected ways for signal transduction in plant cells (Baxter *et al.,* 2014; Gilroy *et al.,* 2014), and important mediators of PCD, for example of the pollen tube (Duan *et al.,* 2014). Moreover, extracellular ROS are also required for cell expansion in root hairs, the growing PT, and the leaves while the individual ROS moieties assume antagonistic functions (Singh *et al.,* 2016). In addition, ROS were shown to be critical for embryo sac development and increased ROS levels in the embryo sac lead to arrested embryogenesis (Martin *et al.,* 2013). Hence, functional studies on these members of uncharacterized CRKs could provide additional insights on the requirement for apoplastic ROS-sensing during embryo sac development or on the level of gamete interactions.

In addition, four members of the family of S-locus Lectin RLKs of the SD1 clade (Shiu *et al.,* 2003) were determined upregulated in the female gametophytic cells by the transcriptomic approach (Supplemental Figure 5-1 B). When browsing publicly available data via Genevestigator (Hruz *et al.,* 2008), AT1G61490 and AT1G61500 showed microbe-associated molecular pattern (MAMP)-responsiveness in roots and leaves, respectively. AT1G61440 was upregulated upon ABA treatment of suspension culture cells (Okamoto *et al.,* 2012). AT1G61390, which was only found differentially upregulated in the CC, showed a weak response on transcriptomic

level to MAMPs, brassinolide and boric acid treatment of seedlings (Goda 2004, Tintor *et al.,* 2013). Moreover, these four Lectin S-domain RLKs were, on sequence level, highly similar to *LORE*, another member of the SD1 clade, which is involved in MAMP-triggered immunity (Ranf *et al.,* 2015). Thereby, these RLKs are most likely associated to defense-responses and not necessarily prime candidates for the events of gamete recognition and interaction.

The largest group of RLKs in *Arabidopsis* are the leucine-rich repeat receptor like kinases (LRR-RLKs), of 228 members forming 14 distinct clades (Shiu *et al.,* 2003). In the microarray-based expression data 104 LRR-RLK genes were determined differentially expressed (Supplemental Figure 5-2). 20 LRR-RLKs were found upregulated and of those eight LRR-RLKs in female gametes only. These were, for example, TMK3 (type IX) which is involved in auxin-dependent growth regulation (Dai *et al.,* 2013), and BSR880 (type-III), a Brassinosteroid kinase-interacting receptor (Xu *et al.,* 2014). The remaining six LRR-RLKs remain to be functionally characterized and showed a plant-wide relative ubiquitous expression pattern. In a comprehensive *Arabidopsis*-wide screen of the expression pattern of all 228 LRR-RLKs only three had been found flower-specific (Wu *et al.,* 2016): The POLLEN RECEPTOR LIKE KINASE 6, which is involved in LURE sensing and PT guidance (Takeuchi *et al.,* 2016), STRUBELLIG-RECEPTOR FAMILY 5, that does not have an assigned function yet (Eyuboglu *et al.,* 2007), and the EF-TU RECEPTOR which mediates PAMP-triggered defense (Zipfel *et al.,* 2006).

In conclusion, the combination of the comprehensive study of Wu et al., 2016 (Wu *et al.,* 2016) and this work did not identify a female gamete-specific LRR-RLK. However, a cell type-specific expression of the RLK may not be necessary as, for example, the LLR-RLK FER, which is broadly expressed, mediates a variety of different processes such as PT perception, immune signaling, root hair growth promotion, or mechanical signal transduction, depending on the respective tissue, developmental stage, coreceptor, and ligand (Huck *et al.,* 2003; Escobar-Restrepo *et al.,* 2007; Duan *et al.,* 2010; Shih *et al.,* 2014; Li *et al.,* 2015; Mao *et al.,* 2015; Stegmann *et al.,* 2017).

3-8 Like a Stick to the Head: Reverse genetics of *DUF962*

Acquiring gene-specific knock out plant lines from the large collections of T-DNA insertion plant lines (Alonso *et al.,* 2003; Sessions *et al.,* 2002; Rosso *et al.,* 2003) has long been a primary approach in *Arabidopsis* reverse genetics. If T-DNA insertion lines do not provide the desired knock outs, or the insertion positions were incorrectly mapped in the first place, silencing RNA (siRNA) via long double stranded RNA, or artificial micro RNA (amiRNA) can provide alternative solutions to acquire a gene-specific knock down (Ossowski *et al.,* 2008).

The *Arabidopsis* DUF962-containing genes expression patterns were determined in stable transgenic *Arabidopsis* plant lines expressing *DUF962_1g:GFP*, *DUF962_2g:GFP*, and *DUF962_3:GFP*, respectively. While DUF962_3:GFP could not be detected at all. DUF962_1g:GFP was localized in the endomembrane system of the mature ECs, the leaves and the roots, while DUF962_2g:GFP was detected in the the mature CC, the developing endosperm, the leaves, and the cotyledons. Three obtained homozygous T-DNA insertion lines for *DUF962_1* had WT-like *DUF962_1* transcript levels. Moreover, in the beginning of this study for *DUF962_2* no T-DNA insertion lines were available. To conduct a functional study on DUF962 genes nonetheless, an amiRNA construct, targeting *DUF962_1* and *DUF962_2* (amiR^{DUF962}) was generated. Expression of amiR^{DUF962} in transgenic *Arabidopsis* lines under the transcriptional control of a sporophyte-specific promoter, a CC-specific promoter, or expression during FG development, starting from stage FG1, did not affect the seed set of the respective plant lines. However, *EC1.1p:amiR*^{DUF962} expressing knock down lines exhibited in 42% of the generated plant lines an early female gametophytic arrest.

In the past, several studies explicitly addressed identification of gametophytic mutants in *Arabidopsis,* or described genes found essential for the FG development (Christensen *et al.,* 1998; Christensen *et al.,* 2002; Acosta-Garcia *et al.,* 2004; Pagnussat *et al.,* 2005). Most of the observed phenotypic abnormalities within the FGdevelopment were characterized by an early arrest during stage FG1 and FG2. But notably, only 2% of the mutant plant lines showed a gametophytic defect (Pagnussat *et al.,* 2005). Therefore 42% of the plant lines expressing the *EC1.1p:amiR*^{DUF962} construct exhibiting an arrest in early FG developmental stages was rather unexpected and required further investigations, as the *EC1.1* promoter (Ingouff *et al.,* 2009) was expected to drive an EC-specific expression of amiR^{DUF962}.

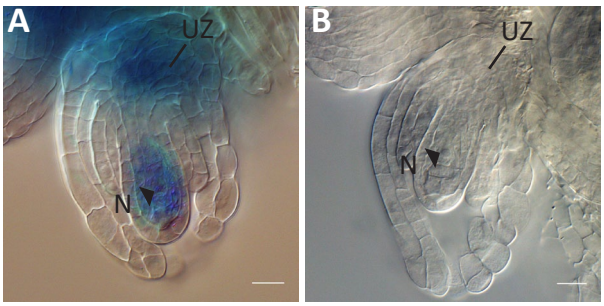


Figure 3-3 Observed GUS signals in young ovules of *EC1.1p:GUS* plant lines and the WT after over night staining.
Pistils of floral stage 12a of eight individual heterozygous *EC1.1p:GUS* plant lines (**A**) and WT (**B**) were examined by GUS-staining. Two out of eight *EC1.1p:GUS*-expressing plant lines showed staining in the functional megaspore in ovule stage 3-I, and seven out of eight plant lines showed staining in the region of the vasculature unloading zone (**A**), while the respective WT control was unstained (**B**). N = functional megaspore, UZ = vasculature unloading zone. Scale is 10µm.

To re-investigate the *EC1.1* promoter activity during FG development, extended GUS staining of eight individual T1 plant lines expressing *EC1.1p:GUS* was performed. Notably, in stage 3-I ovules, GUS-derived signals was observed from the functional megaspore in two out of eight lines, and in seven out of eight cases GUS staining in the region of the vasculature unloading zone, while the WT remained unstained (Figure 3-3). After overnight-staining of mature ovules these eight T1 *EC1.1p:GUS*-expressing plant lines showed staining of the EC in all eight plant lines (Supplemental Figure 5-8). However, stage 3-II/III ovules of *EC1.1p:GUS* reporter lines did not show any detectable staining in the corresponding FGs (not shown). Furthermore, no detectable expression of DUF962 proteins was observed in early FGs, nor could an introduced amiRNA resistant gene *DUF962_1g-amiR^{res}* revert this FG arrest. Frequent ovule abortions in the two-nucelate FG stage, like the arrests observed in the *EC1.1p:amiR^{DUF962}* plant lines, have been reported in salt-stressed plants as a cause of water depletion (Sun *et al.,* 2004). However, the respective control plant lines for the *EC1.1p:amiR^{DUF962}* did not show any comparable effects. Therefore, the observed arrest during the early FG development indicates a T-DNA position insertion effect in the generated amiRNA lines, even though the high frequency of this phenotype would suggest an amiR^{DUF962}-mediated effect, as the amiR^{DUF962} had been verified functional in leaves.

3-9 DUF962_1 is a True Ortholog of Yeast MPO1

Studies in yeast have verified *Arabidopsis* DUF962_1 to be a true ortholog of *Saccharomyces* MPO1, as the accumulation of 2-hydroxy palmitic acid (2-OH PA) in the growth medium was abolished in DUF962_1 expressing *Δmpo1* cells. Consequently, DUF962_1 should be renamed to AtMPO1.1. As the amino acid sequence identity between DUF962_1 and DUF962_2 was determined to 75%, an identical protein function is very likely, thereby DUF962_2 qualifies for AtMPO1.2. More controversial is DUF962_3 which, albeit closer related to yeast MPO1, could not fully revert the *Δmpo1* mutant specific accumulation of 2-OH PA in the YPD medium. Moreover, expression of DUF962_3:GFP could not be detected in *Arabidopsis* and thereby DUF962_3 might be a pseudogene.

Yeast growing in glucose-supplied medium is under glucose-induced transcriptional repression (Kresnowati *et al.,* 2006). Studies in yeast showed determined MPO1 to be transcriptional upregulated upon the glucose depletion-induced diauxic shift (Brauer *et al.,* 2005), and even more upregulated in response to stationary phase (Gasch *et al.,* 2000). The diauxic shift marks the transition from log-phase into stationary phase and is characterized by increased stress due to nutrient exhaustion (Werner-Washburne *et al.,* 1993). Furthermore, in a competitive growth experiment conducted in synthetic exudates medium (minimal medium), the *Δmpo1* had shown only 86% of WT fitness (Qian *et al.,* 2012). Thereby, the obvious explanation is that the transcriptional regulation of MPO1 enables for execution of an alternative C_{16} fatty acid oxidation pathway for energy generation. On the molecular level, this is further strengthened by the fact that the N-terminal region of MPO1 has weak sequence similarity to an NAD^+ binding domain (Kondo *et al.,* 2014), which shares conserved amino acids with *Arabidopsis* MPO1 proteins.

As MPO1 was found endoplasmic reticulum localized (Huh *et al.,* 2003; Kondo *et al.,* 2014), and AtMPO1.1:GFP was also shown to localize to the endomembrane system, a non-plastidial mode of action of MPO1 is likely. Notably, Faa1p, a key enzyme of the fatty acid oxidation pathway which usually localizes to the peroxisomes was shown to additionally localize to other cellular compartments like the plasma membrane and to mediate fatty acid transport (Black *et al.,* 2007; Houten *et al.,* 2010; Fan *et al.,* 2014). However, as MPO1 is proposed to act as an alpha oxidase, the endoplasmic reticulum-based 2-OH PA breakdown results in

either endoplasmic, or cytosolic-localized pentadecanoic acid (C_{15}), CO_2 and NADH, and would thereby not directly contribute to respiratory-based ATP generation, but require a NAD^+/NADH shuttle which consequently results in less ATP-yield (Shen *et al.,* 2006).

In opposite to even-numbered fatty acids, the beta oxidation of odd-numbered fatty acids such as pentadecanoic acid yields as final products succinyl-CoA and propionyl-CoA. Propionyl-CoA has to be regenerated under ATP consumption before it can be channeled back into the carbon metabolism (McCarthy *et al.,* 2001), while more importantly, succinyl-CoA can readily be utilized in the citric acid cycle (Krebs 1940).

Intermediates of the citric acid cycle serve as a variety of precursors for molecular components like amino acids, or nucleotides. The formation of glutamate or subsequently of glutamine, both utmost important carriers for nitrogen, is achieved by amination of alpha ketoglutarate (Forde *et al.,* 2007) – thereby, within the citric acid cycle, potentially depleting the direct precursor of succinyl-CoA. Consequently, alpha oxidation of fatty acids to odd-numbered moieties and subsequent beta oxidation poses an anaplerotic pathway to replenish the citric acid cycle with succinyl-CoA when need arises during stress or aging (Pfeuffer *et al.,* 2016), and especially for increased levels of protein biosynthesis (Owen *et al.,* 2002). However, plants and microorganisms are capable of generating succinate during the glyoxylate cycle (Eastmond *et al.,* 2001). To conclude, the generation of odd-numbered fatty acids by alpha oxidation, primarily for production of energy and metabolic intermediates, appears, in plants and microorganisms, less preferable than a combination of the beta oxidation pathway and the glyoxylate cycle.

3-10 The Fat and the Furious: Sphingolipid Signaling and MPO1

In yeast, it was shown that odd-numbered fatty acids were generated from phytosphingosine (PHS; Kondo *et al.,* 2014), which poses, in yeast and plants, a substantial amount of the plasma membrane (Pata *et al.,* 2010). Sphingolipids are synthesized from serine and palmitoyl-CoA by the endoplasmic reticulum-localized SERINE PALMITOYLTRANSFERASE (Hanada 2003). The long-chain bases of sphingolipids are known to be important second messengers in their phosphorylated and dephosphorylated form with reported antagonistic function in ROS-mediated cell death, respectively (Shi *et al.,* 2007; Pata *et al.,* 2010). Furthermore, signaling

functions vary as chain desaturation and hydroxylation at position four (C4) was reported to alter bioactivity of phytosphingosine-1-phosphate (PHS1P) in plants and confer binding to the heterotrimeric G protein GPA1 in order to control stomatal aperture, downstream of abscisic acid signaling (Coursol *et al.,* 2005). In rice, a knock down of the stigma-expressed Dihydrosphingosine C4 Hydroxylase, which mediates the hydroxylation of dihydrosphingosine at position C4 to phytosphingosine, and dihydroceramide to phytoceramide, was found sterile (Imamura *et al.,* 2007).

In plants and yeast, the degradation of phosphorylated sphingolipids is mediated by the LONG CHAIN BASE PHOSPHATE LYASE (LCB-P), an enzyme that localizes to the endoplasmic reticulum but retains the active center on the cytosolic side (Tsegaye *et al.,* 2007). LCB-P irreversibly cleaves sphingosine-phosphate between the C2 to C3 position and thereby yields ethanolamine phosphate and a long-chain aldehyde, usually (2-Hydroxy)hexadecanal (Tsegaye *et al.,* 2007). Thereby, the LCB-P contributes to (phyto)sphingosine / (phyto)sphingosine-1-phosphate homeostasis upon heightened or deregulated sphingosine production (Tsegaye *et al.,* 2007). Downstream of PHS/PHS1P-mediated signaling, MPO1 catalyzes the formation of pentadecanoic acid and thereby clears out 2-hydroxy palmitoyl-CoA (Kondo *et al.,* 2014), probably to maintain a steady-state of enzymatic degradation. Without MPO1, this steady-state cannot be maintained in a sufficient manner, thereby the accumulating palmitoyl-CoA is possibly hydrolyzed to 2-OH PA and released from the cell in a timely fashion (Kondo *et al.,* 2014). Pentadecanoic acid, once generated, can partially be salvaged in form of glycerophospholipids (Kondo *et al.,* 2014) or utilized for generation of metabolic intermediates and energy. Pentadecanoic acid may furthermore serve a yet unknown purpose, as two species of (C_{17}) odd-numbered fatty acids were isolated from the culture medium of a yeast-like fungus and assigned with antibiotic activity (Rezanka *et al.,* 2015).

AtMPO1 proteins are expressed in the leaves including the stomata, were PHS1P-dependent signaling was reported to regulate stomatal aperture (Coursol *et al.,* 2005). As MPO1 proteins clear phytosphingosine catabolic products, an altering plant-available water supply could give indications if PHS-dependent stomata closure correlates with MPO1 expression levels.

In the fertilized EC, AtMPO1.1 expression was abolished very quickly. This raises the question if the zygote shuts down PHS-dependent signaling pathways. Furthermore, transcriptional

profiling of seedling roots had shown a strong upregulation of *AtMPO1.1*, but not *AtMPO1.2,* upon auxin treatment (Lewis *et al.,* 2013). Auxin, however, was reported to play an important role during embryo development (Long *et al.,* 2006; Pagnussat *et al.,* 2009).

Moreover, also post fertilization a different regulation of AtMPO1.1 and AtMPO1.2 expression was observed. While the EC shuts down AtMPO1.1 expression after fertilization, the developing endosperm continuously expresses AtMPO1.2. The endosperm assumes a supplementary role during seed maturation and development, and provides nutrients to the embryo (Lafon-Placette *et al.,* 2014), which results in the accumulation of vast amounts of oils in embryo and endosperm (Li *et al.,* 2006). Here, a role of AtMPO1.2 as enzyme in the salvage pathway of phytosphingosine appears feasible. So far, the content of odd-numbered fatty acids within the developing embryo and endosperm of *Arabidopsis* has not been in the focus of investigation (Penfield *et al.,* 2004; Li *et al.,* 2006; Graham 2008), probably due to their low abundance and little understood function.

3-11 The Egg Cell-like Callus, a Tool Worth Exploiting!

As most cellular processes are mediated by proteins, an acquired proteome can provide a functional-oriented picture of the respective cell type. Over the last decade, huge technical advances have greatly reduced the minimum requirements to perform –omics studies and also enabled for single cell proteomics, e.g. of human oocytes (Virant-Klun *et al.,* 2016), or to the level of phosphoproteomics (Su *et al.,* 2017). These high-tech methods will support the characterization of cells on different levels. Nevertheless, the identification of unknown protein-protein interactions cannot rely on single cells yet.

In this work, two methods of protein precipitation and microsomal fractionation (MF) were applied in combination with mass spectrometry to acquire the EC-like callus full proteome and membrane proteome, respectively. This resulted in significant differences between the identified putative membrane-associated or secreted proteins of the respective approaches, indicated by little overlap in EC-like specific and upregulated proteins derived of the MF and the AMS/TCA-mediated protein precipitation approches. As the MF included three ultracentrifugation steps more diverse resolubilization procedures of the precipitated protein pellets, using various detergents, are likely to improve the coverage of identified proteins

(Kalipatnapu *et al.,* 2005). As empoyed here, the combination of both approaches with higher sample replicate numbers may even increase the saturation of detected proteins. Also, since previous transcriptional profiling of another EC-like callus by array-hybridization had identified 351 strongly upregulated genes (Koszegi *et al.,* 2011), of which 226 corresponding proteins (64%) were identified in our EC-like cell line proteome.

While RKD1 is expressed by the cells of the egg apparatus, RKD2 is specifically expressed by the EC and is known to assign EC-like identity (Koszegi *et al.,* 2011). The analysis of the proteomic data revealed proteins assigned to embryo and seed development-related biological processes. Although EC-expressed genes, that are not expressed in the embryo, like *AT1G31450* (this work), or *EC1.1* (Sprunck *et al.,* 2012) were not found expressed by the EC-like cell line, an 18% overlap of EC-expressed genes encoding membrane-associated and secreted proteins with the membrane-associated and secreted proteins detected in the EC-like callus proteome was determined, though. Another possible explanation would be an underlying heterogenous population of cells within the EC-like callus composed of cells with EC-like identity, and of cells with embryo or seed-like identity.

Moreover, it had been shown that constitutively overexpression of transgenes causes transcript level-dependent silencing (Schubert *et al.,* 2004). This could apply as in the EC-like callus the EC-identity fate assumption relies on the *CaMV 35S* promoter-driven expression of *RKD2*. Furthermore, RKD4, an regulator of embryonic development (Waki *et al.,* 2011), was not represented by the EC-like callus proteome either. Additionally, proteins that were EC-expressed, and absent of the embryo like AT1G08050, or TET8 and TET9 had been identified in the EC-like callus proteome. Thereby, although exhibiting a mixed embryo and EC fate identity, this EC-like cell line offers potential for EC-related protein interaction studies.

3-12 Putative Protein Interaction Partners of TET9

Knowledge of protein or genetic interactions is growing rapidly as until September 2016, more than 1 million genetic and protein interactions were deposited in the Biogrid database, and while almost half of those corresponded to human-related research, only 41918 protein interactions referred to *Arabidopsis* (Chatr-Aryamontri *et al.,* 2017). For *Arabidopsis* the two major interactome databases are the membrane-interaction database (MIND) (Jones *et al.,*

2014), and the functional association networks STRING (Szklarczyk *et al.,* 2017). However, both did not provide leads for the examined candidate proteins, like TET9.

Integral membrane proteins are known to require intact lipid bilayers for proper folding and function (Lee 2004; Jamshad *et al.,* 2011). Therefore, protein interaction studies in live membranes, such as the mbSUS (Obrdlik *et al.,* 2004), should be preferred, as was achieved for MIND which in turn is limited by the available clones (currently 3233) for testing. Creative Biolabs provided a service in which a prey library from EC-like callus cDNA was prepared and screened with TET9 as bait.

Despite the numerous biological processes, in which Tetraspanins are involved in (Charrin *et al.,* 2014), the mbSUS screen identified only nine putatively interacting proteins, of which four were verified in the correct reading frame (Figure 3-4).

Among them was MEMBRANE STEROID BINDING PROTEIN 2 (MSBP2), a membrane-associated protein encoded by a cell-to-cell mobile mRNA that was detected plant-wide (Winter *et al.,* 2007; Thieme *et al.,* 2015). MSBP2 has UniProt-assigned binding properties for 1,2-cyclopentanoperhydrophenanthrene (The UniProt 2017), which occurs for example in brassinosteroids or cholesterol. Cholesterol is aside from sphingolipids a component of lipid rafts (Lozano *et al.,* 2016), that are described as hypothetical major ordered regions of the plasma membrane (Simons *et al.,* 1997). Tetraspanin-enriched microdomains also

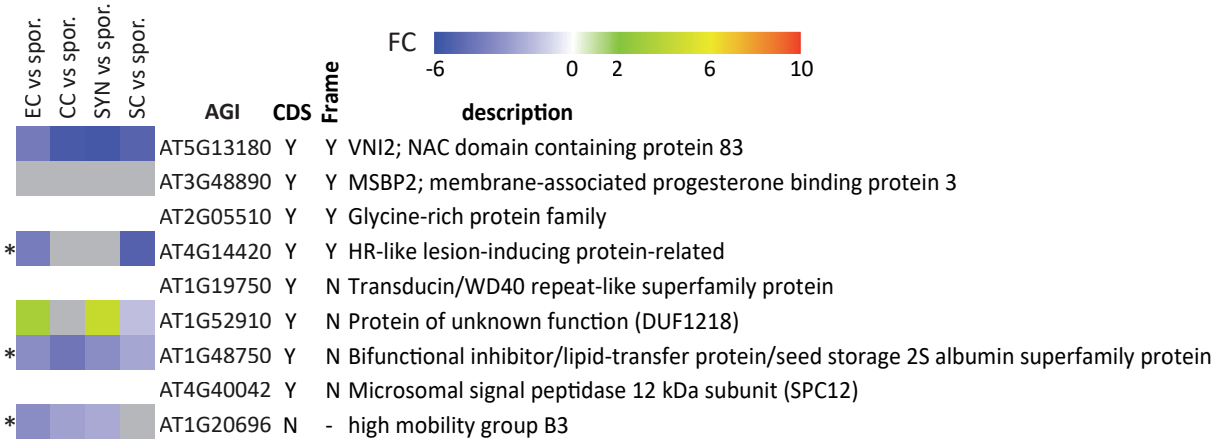


Figure 3-4 Putative TET9 interactors as determined by the mbSUS screen.
Heatmap depiction of transcriptomic data of female gametophytic cells and male gametes after sporophytic contrasting in combination with putative interactors of TET9 (bait) after screen with an EC-like cell line cDNA (prey) library utilizing the split ubiquitin system. Reporter-active clones were sequenced in order to map the sequence to the *Arabidopsis* genome and reading frames. Asterisks indicate proteins identifed as upregulated in the EC-like cell line over the CIM callus in proteomic studies. Genes not featured by the ATH1 GeneChip were left blank in the heatmap. CDS = coding sequence, FC = log2 fold change, gam. = gametophyte.

form functional units on the plasma membrane (Berditchevski *et al.,* 2002; Hemler 2003). Although tetraspanins not necessarily accumulate in lipid rafts (Le Naour *et al.,* 2006), lipid rafts and Tetraspanin-enriched microdomains may yet interact (Charrin *et al.,* 2009). Since the tetraspanin CD81's activity was shown to be modulated upon cholesterol binding recently (Zimmerman *et al.,* 2016), it is possible that TET9 colocalizes with cholesterol within the plasma membrane as well. Thereby, a plasma membrane-based physical interaction of TET9 and MSBP2 appears feasible.

VNI2, a NAC083 containing protein with protein and DNA binding properties, was localized in the nucleus of *35Sp:VNI2-GFP* expressing protoplasts (Yang *et al.,* 2011), was also identified as putative TET9 interactor. VNI2 represses xylem vessel formation by interaction with VASCULATURE-RELATED NAC-DOMAIN7 (Yamaguchi *et al.,* 2010), delays aging and promotes salt stress tolerance as the protein apparently gains increased transcription promoting activity due to higher salt concentrations (Seo *et al.,* 2011; Yang *et al.,* 2011). However, VNI2 does not feature a transmembrane domain, which is in accordance with the SUBA4-based localization for VNI2 in the nucleus (Hooper *et al.,* 2017). Thereby, colocalization of VNI2 with preferentially plasma membrane-localized TET9 is unlikely. The acylation of proteins can lead to attachment of these proteins to the plasma membrane, though (Resh 2004). An conducted acylation motif search on the VNI2 protein sequence (lipid.biocuockoo.org) revealed a S-Farnesylation motif at the amino acid position C248.

In conclusion, the putative interactions of MSBP2 and VNI2 with TET9 should be verified in a second approach such as co-immunoprecipitation but could provide first evidence of a molecular function of TET9.

3-13 The Role of Tetraspanin Proteins in Development

Investigations of *TET7g:GFP*-expressing plant lines showed patchy accumulations of TET7 within the SYN and the growing PT. Furthermore, tetraspanin proteins are known to interact with the cytoskeleton (Detchokul *et al.,* 2014), bind lipids (Zimmerman *et al.,* 2016), or interact with phospholipid binding proteins (He *et al.,* 2011), and are important for exocytosis (Bari *et al.,* 2011). Thereby, judging from the observed localization pattern of TET7:GFP, TET7 is a likely candidate for mediating exocytosis in the secretory very active cells like the SYN and the PT.

The investigated quadruple mutant with a *tet8 tet11 tet12 tet9*$^{+/-}$ genotype showed abnormal amounts of arrested female gametophytes with disintegrating nuclei in stage FG4. This developmental stage coincides with first detectable TET9-GFP expression and is therefore likely a phenotype caused by the absence of functional TET9. Nuclear degradation in stage FG4 has been reported, for example, in the *Arabidopsis* mutants *eda5* and *eda7* (Pagnussat *et al.,* 2005). EDA5 corresponds to a mutation in the putative extracellular localized exostosin-type glycosyltransferase, while in EDA7, a protein of unknown function containing WD40 repeats is affected (Pagnussat *et al.,* 2005). WD40 domain proteins pose a huge family with protein-protein and protein-nucleic acid binding capacity, and no assigned enzymatic function (Xu *et al.,* 2011) and thereby a scaffold function similar to those of the tetraspanin proteins.

In general, nuclear disintegration is observed in plant cells undergoing programmed cell death (PCD) (Reape *et al.,* 2008), which is accompanied by fragmentation of the DNA (Mittler *et al.,* 1997). It is known that the chromatin of cells undergoing PCD condenses at the inner nuclear membrane and degrades (Papini *et al.,* 1999), which resembles Feulgen staining-based observations of stage FG4 nuclei in the tetraspanin quadruple mutant plant line #25. Until comparable stage FG4 developmental arrests, like in the tetraspanin quadruple mutant line #25, are observed in additional independent *tet8 tet9 tet11 tet12* mutant plant lines a T-DNA insertion position effect can, per se, not be ruled out.

3-14 Conclusions and Future Issues

This work provides evidence for the good quality of the transcriptomic data obtained from the manually microdissected live female gametophytic cells. However, none of the candidate proteins investigated in this work revealed a function during the interaction of male and female gametes but exhibited potential functions during female gametophyte development. Meanwhile single cell –omics have become increasingly powerful and allow for parallel sequencing of the methylome and the transcriptome (Angermueller *et al.,* 2016), and more importantly, for identification and quantification of the peptides of a single cell (Budnik *et al.,* 2017). Thereby, this could facilitate the exhaustive characterization of the gamete transcriptomes and the composure of the female gamete cell surface in more detail.

This is especially interesting as the protoplast-like female gametophytic cells were shown to eventually execute plasmogamy with each other in the enzymatic solution used to islolate them from ovules (Englhart *et al.,* 2017). Moreover, the synergid cell has reportedly assumed egg cell identity in mutants (Pagnussat *et al.,* 2007), (Kong *et al.,* 2015), and fuses with the central cell post fertilization (Maruyama *et al.,* 2015). Hence, it can be expected that an unknown fusogenic factor is expressed on the cell surface of the egg cell, central cell and the synergid cell. It will thus be important to continue the functional characterization of more genes identified in this work.

This study and a previous work (Boavida *et al.,* 2013) revealed the male and the female gametophytes express a large number of tetraspanins proteins, of which *TET11* to *TET16* correspond to the male gametophyte, and *TET8* to *TET9* to the female gametophyte, while TET7 is expressed by both gametophytes. Studies of mutant plant lines did not assign an essential function to TET8, TET11 or TET12 in a combined knock out of *tet8*, *tet11*, and *tet12*. To overcome the functional redundancy within the gene family (Boavida *et al.,* 2013), (Wang *et al.,* 2012), and to determine the extent of complementary tetraspanin function regarding *TET11* to *TET16* as all are expressed in the sperm cells, mutant lines of much higher magnitude will be required.

This can be achieved by a combination of the already established *tet11* knock out line and the CRISPR/Cas9 technology that is readily replacing the toolbox of conventional reverse genetics (Noman *et al.,* 2016), and as employed in this work allows for simultaneous and precise

mutagenesis of multiple target loci. Additionally, as *TET8* and *TET9* are both expressed by the egg cell, double knock out mutants for *TET8* and *TET9* could address functional redundancy of TET8 and TET9 in the egg cell and thereby reveal a biological function of TET8 and/or TET9. If reproducible phenotypes are generated by knocking out single tetraspanin genes, like TET9, or by knocking out combinations of tetraspanin genes, it will be essential to revert the phenotype to a WT-like situation by reintroducing the respective tetraspanin genes, thereby effectively neglecting putative secondary effects caused by CRISPR/Cas9 mutagenized off-targets or T-DNA position insertion effects. Phenotypic rescue would also allow to address the evolutionary conservation of tetraspanin proteins by utilizing tetraspanin genes-derived from other phylogenic clades or species. Functional complementation could be achieved by selecting for Cas9-segregated progeny of mutant plant lines or by reintroducing proto spacer motif-mutated tetraspanin sequences.

By utlizing live imaging of the generated gene:reporter plant lines it would be interesting to see if temporal resolution of the events of gametophytic and gamete interactions could provide new insights on possible tetraspanin dynamics during these processes.

Yeast MPO1 and *Arabidopsis* AtMPO1.1 are 2-OH PA catabolic enzymes. Under nutrient starvation conditions yeast Δ*mpo1* mutants displayed a decreased fitness (Qian *et al.,* 2012). *Arabidopsis mpo1.1 mpo1.2* double knock out plant lines will soon be generated by simultaneous expression of already cloned sgRNAs targeting AtMPO1.1 and AtMPO1.2 in combination with Cas9-mediated DNA cleavage. This could provide a seed or leaf lipid content-dependent phenotype. Moreover, morphological or fitness-related phenotypes could be observed under starvation conditions, similar to those observed in yeast mutants.

4 – Experimental Procedures

All material and reagents were used at *pro analysi* (pa) grade and purchased from Sigma-Aldrich Co. LLC. if not stated otherwise. Furthermore, if not stated otherwise dissolved in Millipak-grade water (Merck KGaA), autoclaved 20min at 121°C, 2 MPa, or filter sterilized (0.2µm).
For handling RNA, only water of Biopak-grade (Merck KGaA), and RNAse-free plastic ware was used. Glass ware was heated to 180°C over night or DEPC-treated, followed by autoclavation as described above. Nucleic acid concentrations and cell turbidities, at OD_{600} were measured at the NanoDrop2000 (Thermo Fisher Scientific). All the wild-type plants were progeny of the ecotype *Col-0*, T-DNA insertion lines were obtained from the Arabidopsis Biological Resource Center, or the Nottingham Arabidopsis Stock Centre.

4-1 Computer-based Methods

For various *in silico* analyses, simulations, and predictions following tools were used.

4-1-1 Databases, Online Tools, and Software

“Adobe Creative suite” v4 adobe.com

“Aramemnon” v8.1 aramemnon.uni-koeln.de (Schwacke *et al.,* 2003)

“Array Express“ ebi.ac.uk/arrayexpress/ (Kolesnikov *et al.,* 2015)

“Biomart” v.9 biomart.org (Smedley *et al.,* 2015)

“Breaking Cas” bioinfogp.cnb.csic.es/tools/breakingcas (Oliveros *et al.,* 2016)

“CLC Main Workbench” v6.x clcbio.com

“DAVID” v6.8 david.ncifcrf.gov/conversion.jsp (Huang da *et al.,* 2008)

“eFP Browser” bar.utoronto.ca/efp/cgi-bin/efpWeb.cgi (Winter *et al.,* 2007)

“FiJi” imagej.net/Fiji v1.50g, with CoLoc, and Royal Lookuptable plugins (Schindelin *et al.,* 2012)

“GENEVESTIGATOR®” genevestigator.com (Hruz *et al.,* 2008)

“Gene Expression Ombnibus” ncbi.nlm.nih.gov/geo/ (Edgar *et al.,* 2002; Barrett *et al.,* 2013)

“IBM SPSS Statistics” v23 ibm.com

“KEGG” v82.0 genome.jp/keg/ (Kanehisa *et al.,* 2000; Kanehisa *et al.,* 2016; Kanehisa *et al.,* 2017)

„Microsoft Excel“ v2007 microsoft.com

„MIND" associomics.dpb.carnegiescience.edu/Associomics/Home.html (Jones *et al.,* 2014)

"Morpheus" software.broadinstitute.org/morpheus/#

"MUSCLE" ebi.ac.uk/Tools/msa/muscle/ (Edgar 2004)

"PANTHER" v11 pantherdb.org/ (Mi *et al.,* 2013; Mi *et al.,* 2017)

"Pfam database" v30.0 pfam.xfam.org/ (Finn *et al.,* 2016)

"Phylodendron" iubio.bio.indiana.edu/webapps/Phylodendron/ by D.G. Gilbert

"Phytozome" v12 phytozome.net (Goodstein *et al.,* 2012)

„Primer3web" primer3.ut.ee/ (Koressaar *et al.,* 2007; Untergasser *et al.,* 2012)

"Primer-BLAST" ncbi.nlm.nih.gov/tools/primer-blast/ (Ye *et al.,* 2012)

"Saccharomyces Genome Database" yeastgenome.org (Cherry *et al.,* 1997; Cherry *et al.,* 2012; Engel *et al.,* 2013)

„SignalP" v4.1 cbs.dtu.dk/services/SignalP/ (Petersen *et al.,* 2011)

„SUBAcon" suba3.plantenergy.uwa.edu.au/ (Tanz *et al.,* 2013; Hooper *et al.,* 2014)

"SWISS-MODEL" swissmodel.expasy.org/interactive (Arnold *et al.,* 2006; Kopp *et al.,* 2006 Bordoli *et al.,* 2009; Guex *et al.,* 2009; Kiefer *et al.,* 2009; Biasini *et al.,* 2014; Bienert *et al.,* 2017)

"TAIR" v10 arabidopsis.org (Huala *et al.,* 2001; Rhee *et al.,* 2003; Harris *et al.,* 2004; Lamesch *et al.,* 2010; Lamesch *et al.,* 2012; Berardini *et al.,* 2015)

"T-DNA Express" signal.salk.edu/cgi-bin/tdnaexpress

"UniProt" uniprot.org (The UniProt 2017)

"Venny" v2.1 bioinfogp.cnb.csic.es/tools/venny/ by Oliverso, J.C. (2007-2015)

Phosphorylation site prediction: musite.net (Gao *et al.,* 2010); phosphosite.org (Li *et al.,* 2002), (Bateman *et al.,* 2002; Obenauer *et al.,* 2003)

"WMD3" wmd3.weigelworld.org/cgi-bin/webapp.cgi (Schwab *et al.,* 2005; Ossowski *et al.,* 2008)

4-1-2 Transcriptome Analysis (by Dr. Maxim Messerer)

Inference of differentially expressed genes (DEGs) was performed based on a total of 69 Affymetrix Arabidopsis ATH1 Genome Arrays (ATH1-121501) (Supplemental Table 5-1). In addition to our data, we obtained further datasets from NCBIs Gene Expression Omnibus

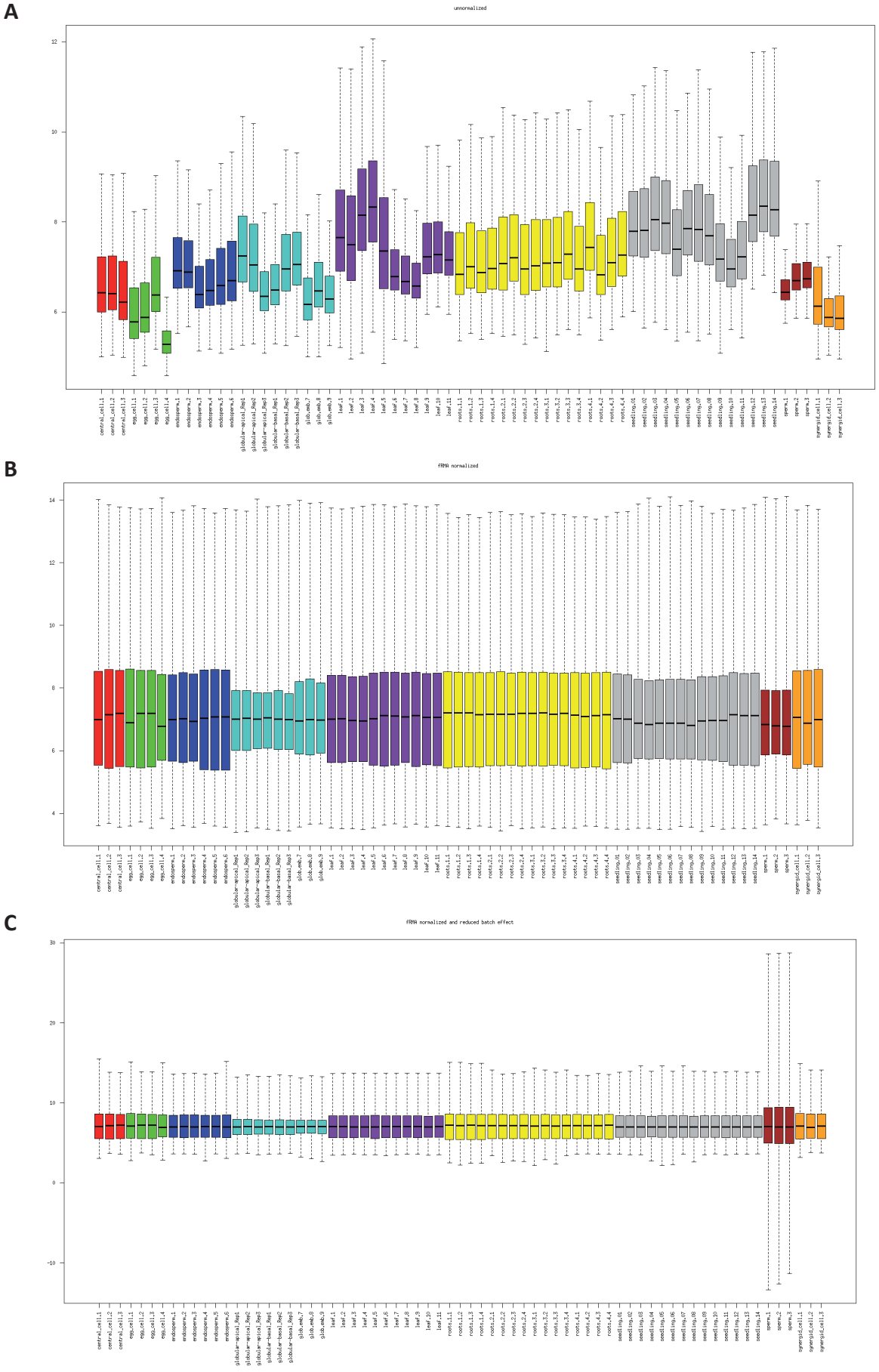


Figure 4-1 Expression values of all microarrays utilized for transcriptome data processing. Microarray-derived expression values before normalization (**A**), after normalization (**B**), after batch effect correction (**C**). These boxplots were provided by Dr. Maxim Messerer.

(GEO) and ArrayExpress. Normalisation of arrays was performed using the frozen Robust-Multi-array Analysis (fRMA) algorithm from the Bioconductor package frma (McCall *et al.,* 2010). As no frozen parameter vectors were available for Affymetrix ATH1 microarray chips from the Bioconductor repository, we utilized the package frmaTools to generate vectors using a training dataset comprised of 246 CEL-files obtained from GEO (82 batches, batch size = 3) (McCall *et al.,* 2011). A cutoff value based on the internal negative controls of the Affymetrix GeneChip was defined. Probes below this cutoff were discarded from the analysis as non-expressed. Quality control was perfomed using box- and MDS plots comparing unnormalized and normalized datasets (Figure 4-1). The DEGs were identified using the GLM functionality of limma package (Ritchie *et al.,* 2015). A corrective term for possible batch effects (column „RNA", Supplemental Table 5-1) was included in the model GLM matrix prior to querying the individual contrasts. The *A. thaliana* gene IDs and GO terms were downloaded from BioMart (Aken *et al.,* 2016) and the gene descriptions from the TAIR homepage (gene_aliases_20130831.txt). The prediction of transmembrane domains was obtained from ARAMEMNON, and the subcellular protein localisation predictions from SUBAcon.

4-1-3 Manual Transcriptome Data Annotation and Processing

TAIR10-Subcellular_Predictions.xlsx, and TAIR10_functional_descriptions.txt was downloaded from TAIR v10, gene functional description, cellular compartment, and localization prediction were assigned to the respective AGIs. All mitochondrial and chloroplastidial-derived gene identifiers were removed as well as all probe sets binding multiple gene transcripts. Furthermore, all genes putatively not membrane-associated or extracellular where excluded from further analysis. All expression values with an adjusted pValue higher than 0.001 were removed from subsequent analysis.

4-1-4 GO-term Analysis

All AGIs were converted to EntrezGeneIDs by DAVID conversion tool, assigned to the respective expression values and uploaded to PANTHER Classification System. Statistical expression test for *A. thaliana* and default settings was performed and results were exported. GO term statistical overrepresentation test was performed with a list of AGIs-only but analogous to GO

term enrichment.

4-1-5 Heatmap Generation

Annotated expression values were uploaded as Excel spreasheet, the data was saved to *.svg and imported into Adobe Illustrator.

4-1-6 Reconstruction of Phylogenic Relationships

DNA coding sequences of selected organisms were downloaded from available databases (Biomart v.9, Saccharomyces Genome Database) aligned with MUSCLE, standard settings and the resulting Phylip output was vizualized by Phylodendron, set to interleaved output.

4-1-7 Designing CRISPR/Cas9 sgRNA

Target gene sequences were uploaded to Breaking Cas (Oliveros *et al.,* 2016), prediction was achieved with *Streptococcus pyrogenes* 20nmer and NGG PAM site settings, suggested targets sites without putative off-targets were chosen for construct generation.

4-1-8 Manual Proteome Data Processing

Mass spectrometric data-derived from microsomal fractions of EC-like callus and CIM callus samples were manually processed as follows: Non unique matching peptides were removed, splice variants were ignored, and proteins only identified in one sample replika were removed as well. Also, mitochondrial and chloroplastidal-derived gene identifiers were excluded from further analyses. Moreover, the ratios of protein intensities (LFQ values) for EC-like callus and CIM callus samples were calculated and log2 transformed. The full proteomic analysis was performed by Dr. Julia Mergner. The log2 fold change values of the samples derived of TCA and AMS-mediated protein precipitations were arithmetically averaged for final presentation.

4-1-9 Plotting and Statistics

Statistical analysis and plotting of results was done with SPSS v23 or MS Excel 2007. Venn diagrams were drawn with Venny v2.1.

4-1-10 Oligonucleotide Design

Oligonucleotides for amplification purposes were manually designed and then subjected to Primer3web, and Primer-Blast quality control. The primer pair THO-719/720 was provided by Philipp Cyprys. THO-121/2 and THO-530/1 were provided by Dr. Stefanie Sprunck.

4-2 Nucleic Acid-based Methods

4-2-1 Isolation of Plant DNA

To obtain gDNA from plant tissue protocols with minor modifications of either (Edwards *et al.,* 1991) or for higher purity (Richards *et al.,* 2001) were performed. The DNA was dissolved in 10mM Tris-HCl pH8.0, 1mM EDTA.

4-2-2 Isolation of Plasmid DNA

Plasmid DNA was isolated from *E.coli* by spin column purification kits available from Addgene LGC Standards, MACHEREY-NAGEL GmbH & Co. KG, and Thermo Fisher Scientific Inc. according to manufacturers protocol.

4-2-3 Isolation of DNA from Yeast

Plasmid and gDNA from *Saccharomyces cerevisiae* was isolated from yeast according to (Hoffman 2001).

4-2-4 Isolation of mRNA

For gene expression studies, mRNA was isolated using DynaBeads® Oligo$(dT)_{25}$ (Thermo Fisher Scientific Inc.) according to the manufacturers protocol.

4-2-5 Degrading of DNA

DNA was degraded by DNAseI (Thermo Fisher Scientific Inc.) in Mn^{2+} buffer according to manufacturers recommendations, 15min at 37°C.

4-2-6 Synthesis of First-Strand cDNA

If not stated otherwise cDNA was synthesized from isolated DNA-free RNA by RevertAid H

Minus Reverse Transcriptase (Thermo Fisher Scientific Inc.) according to manufacturers protocol.

4-2-7 Rapid amplification of cDNA ends of AT1G74440

RACE (Frohman *et al.,* 1988) was performed with 5'/3' RACE Kit 2nd generation (Roche Diagnostics Deutschland GmbH), with kit supplied enzymes, and the RevertAid H Minus Reverse Transcriptase (Thermo Fisher Scientific. Inc.) according to manufacturers recommendations. The mRNA serving as template was isolated from floral stage 12c pistils.

5' UTR determination: mRNA was poly A-tailed, then transcribed, in combination with a $d(T)_{20}$ primer into first-strand cDNA. Next, the cDNA was column purified and amplified by PCR using the nested gene-specific primer THO-189 and an anchored $d(T)_{16}$-N primer. Then a second round of PCR-based amplification was performed using THO-355 and the anchored $d(T)_{16}$-N primer.

3' UTR determination: mRNA was transcribed into first-strand cDNA in combination with a $d(T)_{20}$ primer, then a gene-specific product was obtained by PCR with THO-353 and the anchored the anchored $d(T)_{16}$-N primer.

Finally, resulting DNAs were cloned into pCR2.1 blunt vector (Thermo Fisher Scientific Inc.). Vectors were sequenced and the readout was mapped to the genomic sequence of AT1G74440. URTs were determined by sequences obtained from 5/8 (5'UTR) and 4/5 (3' UTR) sequenced and positive clones.

4-2-8 Taq Polymerase-based PCR

PCR (Mullis *et al.,* 1986) was done using Taq Polymerase (Thermo Fisher Scientific Inc.) for non-cloning applications including Colony PCR, genotyping, and RT-PCR. Reaction mix was standardized for 20µL and set up according to manufacturers recommendations.

4-2-9 Colony PCR

Colony PCR was performed with Taq polymerase. gDNA or CDS in pENTR dTOPO or pCR2.1 blunt vector backbones (Invitrogen) was amplified with THO-124/125, and if necessary with gene-specific primers. Promoter:reporter and promoter:gateway constructs were detected

by promoter-specific oligos and either GW-reverse (THO-43), NLS-GFP-reverse (THO-64), or mCherry-H2b-reverse (THO-259). Gene:reporter constructs were detected by gene-specific oligos and/or pGWB550 GFP-reverse (THO-317).

4-2-10 Reverse-transcription PCR

RT-PCR was performed on cDNA extracted from EC-like callus and CIM callus samples with following oligonucleotides for *ABI4* (THO491/2), *DUF674_1* (THO-086/7), *DRP4A* (THO-499/494), *SINA* (THO-495/496), *WOX2* (THO-487/8), *EIF4G* (THO530/531). *ACT2* (THO-121/122) was used as control. Moreover, RT-PCR was also performed on T-DNA insertion line samples with oligonucleotides for *DUF962_1* (see qPCR), *DUF674_1* (THO-086/7), *DUF674_2* (THO-378/9), *DUF674_3* (THO-311/312); *SSPR* (see qPCR); *AT1G08050* (see qPCR); *AT5G67550* (THO-088/9), *AT4G09090* (see qPCR), *AT1G16180* (THO-504/5); *AT5G161900* (THO-411/2), *TET8* (THO-225/583), *TET11* (584/5). For control genes *ACT2* (see qPCR), *GAPC* (see qPCR), and *UBI10* (see qPCR) were selected.

4-2-11 Genotyping of Transgenic Plants, and Determination of T-DNA Insertion Positions

See Table 4-1 for the T-DNA insertion lines (Sessions *et al.,* 2002;Rosso *et al.,* 2003; Alonso *et al.,* 2003), and the oligos used for PCR-based detection of the T-DNA borders, the affected genes and the resistance marker genes.

4-2-12 Phusion® High-Fidelity DNA Polymerase-based PCR

Phusion® High-Fidelity DNA Polymerase (Thermo Fisher Scientific) was used for amplification and further use for cloning applications, as recommended by the manufacturer.

4-2-13 Overlap extension PCR

To generate DUF962^{res} overlap extension PCR (Higuchi *et al.,* 1988) was performed with the oligos THO-581/2 and the gene-specific oligos THO-106/513 followed by Gateway cloning into pGWB550.

Table 4-1 Insertion lines and primers used for genotyping of transgenic plants.
List of oligos used for genotyping of T-DNA insertion lines and generated transgenic plants. LP = forward primer, RP = reverse primer, Lb = left border.

oligo	target / orientation
THO-0001	Lb Gkat
THO-0002	GABI_278_H08 LP1
THO-0003	GABI_278_H08 RP1
THO-0004	SALK_LBb1.3
THO-0005	SALK_136039 LP
THO-0006	SALK_136039 RP
THO-0010	SALK_52167 LP
THO-0011	SALK_52167 RP
THO-0014	SALK_084244 LP
THO-0015	SALK_084244 RP
THO-0016	SAIL_63_G01 LP
THO-0017	SAIL_63_G01 RP
THO-0018	SAIL_1244_B08 LP
THO-0019	SAIL_1244_B08 RP
THO-0045	SALK_109259 LP
THO-0046	SALK_109259 RP
THO-0094	HYGROMYCIN B fw
THO-0095	HYGROMYCIN B rev
THO-0126	pDD65 3' fw
THO-0127	amiRNA rev
THO-0128	SALK_118340 LP
THO-0129	SALK_118340 RP
THO-0130	SALK_113523 LP
THO-0131	SALK_113523 RP
THO-0024	pEC1 3'fw
THO-0219	SALK_025213 LP
THO-0220	SALK_025213 RP
THO-0221	SALK_122879 LP
THO-0222	SALK_122879 RP
THO-0223	SALK_055941 LP
THO-0224	SALK_055941 RP
THO-0229	KANAMYCIN A rev
THO-0230	KANAMYCIN A fw
THO-0231	BASTA fw
THO-0232	BASTA rev
THO-0259	mCherry 5' rev
THO-0280	SAIL_317_C07 LP
THO-0281	SAIL_317_C07 RP
THO-0576	SAIL_869_E01 LP
THO-0577	SAIL_869_E01 RP
THO-0577	SALK_009445 RP
THO-0578	SALK_009445 LP

4-2-14 qPCR

Total RNA was extracted of the selected tissues by the TRIZOL® method (Connolly *et al.,* 2006). 1µg of RNA was subjected to DNAseI (Thermo Fisher Scientific Inc.) treatment for 20min at 37°C using 2U enzyme, followed by heat inactivation and transcription into first-strand cDNA by 150U Superscript SIII® reverse transcriptase (Thermo Fisher Scientific Inc.) at 50°C for one hour.

The qPCR (Gibson *et al.,* 1996) reaction was set up with KAPA SYBR® FAST qPCR Master Mix (2x) (Kapa Biosystems) in half reactions (10µL), using G003-SF stripes (Kisker Biotech GmbH & Co KG) and run in Mastercycler epgradient S realplex² and realplex v2.2 (eppendorf AG). Cycling was done in a two-step program, with 98°C 2s, (98°C for 5s, 60°C for 10s)$_{40x}$ including a final product melting curve quality control. Furthermore, sample replicas with obviously unspecific amplified products (past cycle 35) were excluded. Primer pair-dependent PCR amplification efficiency (E) was determined by three-step dilution of the cDNA template. Normalized relative quantities and the standard error of normalized relative quantities were calculated according to (Hellemans *et al.,* 2007) while excluding the standard error derived of the oligo efficiency determination step. Reference genes *FBX2*, *5-FCl* were proposed by the RefGenes tool integrated in Genevestigator (Hruz *et al.,* 2011). *UBI10*, *GAPC*, *PP2AA3* were chosen as previously described (Kudo *et al.,* 2016).

4-2-15 Agarose Gel Electrophoresis

DNA of size larger than 1.5kbp was separated using Tris-acetic acid-EDTA-buffered gel system (Wagner 1964), while smaller DNA strands were separated by Tetraborate-boric acid-buffered gel system (Brody *et al.,* 2004). RNA was separated in 7% formaldehyde MOPS-buffered gel system (Balmain *et al.,* 1982). Nucleic acids were stained by 2x 10^{-5} % (v/v) Ethidium bromide (Sigma-Aldrich Co. LLC.) final concentration, excited on BiometraTi5 (Biometra GmbH), illuminated by UV-B 8W G5 (SankyoDenki) and detected by Allied Vision Technology camera.

4-3 Cloning Procedures

Before further proceeding to subsequent cloning steps all clones were verified by sequencing, after Colony PCR and a plasmid restriction digest. Sequence analysis was accomplished by CLC Main Workbench.

4-3-1 Restriction-Ligation-based Cloning

Restriction-ligation cloning was performed with enzymes purchased from New England BioLabs GmbH, according to manufacturers recommendations. This also applied to green gate cloning (Lampropoulos *et al.,* 2013). AT1G21670p was amplified by THO-519/520, KNUp was amplified by THO-521/2, and cloned into pH2GW7 (Karimi *et al.,* 2002), Moreover, SSPR and AT4G17505 were cloned into pCambia2300 (Table 2-9), AtMPO1 was cloned into pB7FWG2.0 (Karimi *et al.,* 2002) (see Table 2-9 for additional oligo nucleotide identifiers used for promoter cloning).

4-3-2 Gateway Cloning®

All enzyme mixes for Gateway cloning® were purchased from Thermo Scientific Inc., for all setups only half reactions were used with incubation times > 1h, otherwise according to manufacturers recommendations. Entry clones were generated by either using pENTR/D-TOPO® Kit or a combination of Gateway® BP Clonase™ II Enzyme Mix and pDONR2.07. Expression vectors were generated by Gateway™ LR clonase™ II Enzyme Mix. Used vector backbones pGWB550 (Nakagawa *et al.,* 2007) pB7FWG2.0 and pH7GW2.0 (Karimi *et al.,* 2002) and pGWB80.

Yeast complementation constructs DUF962_1 CDS and DUF962_3 CDS were cloned with stop codon using oligos THO-116/513 and THO-298/515 into pDR196-GW (Komarova *et al.,* 2012); MPO1 was amplified by THO-500/3 and cloned into pDR196-GW (Komarova *et al.,* 2012).

4-3-3 amiRNA Construction

Artifical micro RNA construction amiR^{DUF962} was performed after WMD3 instructions and the used oligos were: THO-118 - THO-121, THO100/1. The amiR^{DUF962} sequence (5'-CTGCGCCTTTAGACAATCTAA-3') was cloned into pENTR dTOPO, then under transcriptional control of EC1.1p (Ingouff *et al.,* 2009) in pGWB80, furthermore under transcriptional control

of the DD65p (Steffen et al., 2007), and the 35Sp in pB2GW7.0 (Karimi et al., 2002).

4-3-4 CRISPR/Cas9 Construct Preparation

Guidance RNAs sequence (Table 4-2) were incorporated into the oligos with respective BSA I overhangs for green gate cloning and subsequent CRISPR/Cas9-mediated genome editing (Ran *et al.,* 2013). Constructs were cloned as described (Wang et al., 2015).

For single construct generation, the oligos for SSPR sgRNA (THO-684/685) were phosphorylated by 10u PNK (NEB), 1x T4 Ligase Buffer (NEB), for 1h at 37°C, then heat the PNK was heat inactivated for 20min at 75°C. Afterwards the phosphorylated oligos were complementary annealed by heat denaturation followed by renaturation, and cycle ligation with 20u BsaI-HF (NEB), 200u T4 Ligase (NEB), 1x Cutsmart (NEB), 1µl (10mM) Ribo-ATP (NEB) at (2min 37°C, 2min 16°C)$_{20x}$ 15min 37°C, 10min 85°C in vector pHEE401E (Wang *et al.,* 2015).

Double sgRNA constructs for TET12/9, TET12/7 were generated by amplification of TET7 reverse (THO-688/9), TET9 forward (THO-690/1) and TET9 reverse (THO-696/7), TET12 forward (THO-686/7) from pHEE2E, subsequently cloned into CRISPR/Cas9 vectors pHEE401E (Wang *et al.,* 2015) by cycle ligation (see above).

Table 4-2 small guidance RNAs for CRISPR/Cas9-mediated genome editing of *TET7, TET9*, *TET12,* and *SSPR*.
The designed sgRNAs for the respective target loci, G/C content, and the binding site relativ to the start codon.

target	AGI	sgRNA (20mer)	G/C content [%]	binding site [bp]	location
TET7	AT4G28050	CCGGGATCTGGCTCGGCAAC	70	77-96	Exon1
TET9	AT4G30430	GGCATTGCGTCGTGGCTTTA	55	96-115	Exon1
TET12	AT5G23030	GGAAAAGCGTTATCCGGTAG	50	304-323	Exon1
SSPR	AT3G05460	GGCGGCGTGCATACAAATAG	55	142-161	Exon1

4-4 Organism-based Methods

4-4-1 Cultivation and Strains of Bacteria

Escherichia coli strain DH5α (F^- *endA1 glnV44 thi-1 recA1 relA1 gyrA96 deoR nupG purB20* ϕ80d*lacZ*ΔM15 Δ(*lacZYA-argF*)U169, hsdR17($r_K^- m_K^+$), λ^-), and DB3.1 (F- gyrA462 endA1 glnV44 Δ(sr1-recA) mcrB mrr hsdS20(r_B^-, m_B^-) ara14 galK2 lacY1 proA2 rpsL20(Sm^r) xyl5 Δleu mtl1) were used in this work, DB3.1 only for propagation of plasmids holding the ccdB operon. Bacteria were grown at 37°C, shaking with 200rpm, or if on plate grown upside down, using LB medium. Selection via antibiotics was done using either Carbenicillin 100µg/ml, Kanamycin 50µg/ml, Spectinomycin 100µg/ml.

Agrobacterium tumefaciens was used for plant transformation, with either strain C58C1 or GV3101::pMP90 for pCambia- or pGreen-derived backbones, respectively and always grown at 28°C shaking at 200rpm or upside down if on plate. Selection via antibiotics was done using either Gentamycin 10µg/ml, Tetracyclin 10µg/ml, Rifampicin 12µg/ml, Carbenicillin 100µg/ml, Kanamycin A 100µg/ml, or Spectinomycine 200µg/ml.

4-4-2 Transformation and Generation of Competent Cells

Calcium chloride-dependent chemically competent *E. coli* or *A. tumefaciens* were generated after (Dagert *et al.,* 1979) at growth at 18°C for two days or 28°C for one day, respectively. Transformation was done by heat shock at 42°C or 37°C, respectively.

4-4-3 Cultivation of Yeast

Δmpo1, clone ID 4378 (genotype BY4741 MatA haploid his3Δ1 leu2Δ0 met15Δ0 ura3Δ0), and the respective parental strain BY4741 of *Saccharomyces cerevisiae* were obtained from GE Healthcare Dharmacon Inc., grown at 30°C and cultivated if not stated otherwise in YPD medium (2% (w/v) glucose, 2% (w/v) bacto peptone, 1% (w/v) yeast extract, solidified with 1.5% (w/v) agar if required), shaking at 200rpm or grown upside down if on plate. Selection was done using the respective dropout medium, *Δmpo1* was additionally selected with G418 200µg/ml, bacterial contamination was suppressed by 100µg/ml Carbenicillin.

4-4-4 Yeast Transformation

Cells were transformed after (Ito *et al.,* 1983), resulting colonies were tested by colony PCR, then plasmid and DNA was extracted, and used for retransformation of competent *E. coli*, of which in turn plasmid of sufficient purity could be obtained and analyzed for correct verification of positive yeast clones.

4-4-5 Sterile Plant Growth Conditions

Arabidopsis seeds were surface sterilized by 70% (v/v) EtOH (tech.) 0.5% (v/v) TWEEN20, dispersed on sterile Whatman filter discs with 100% EtOH (tech.), sown on 2.2g/l Murashige & Skoog medium including MES buffer and vitamins (DuchefaDirect), solidified with 0.9% (w/v) phytoagar (DuchefaDirect) in sterile plastic ware. Antibiotics were supplemented at the concentrations: 75µg/ml (Kanamycin A), 30µg/ml (Hygromycin B).

4-4-6 Callus Line Generation and Propagation

The RKD2-induced EC-like callus line was generated by Marc Urban (Urban 2016) and maintained by Ingrid Fuchs. Root segment-derived control callus induced by callus inducing medium (Goren *et al.,* 1979; Ikeuchi *et al.,* 2013) was generated by Monika Kammerer. All tissue was shock frozen in liquid nitrogen and kept at -80°C until further use.

4-4-7 Plant Growth Conditions on Soil

Soil was supplemented with fertilizer, fungicide and pesticide, plants were grown for 6 weeks in a 8h light / 16h dark photoperiod then transferred to 16 h / 8 dark photoperiod at approximately 8000 Lux.

4-4-8 Generation of Stable Transgenic Lines of *Arabidopsis*

If not stated otherwise, several pots of *Arabidopsis* ecotype *Col-0* were transformed with *A. tumefaciens* using the floral dip method (Zhang *et al.,* 2006) using the infiltration solution: 2.2g/L (w/v) Murashige & Skoog medium including MES buffer and Vitamins (DuchefaDirect), 2% (w/v) sucrose and 0.001% (v/v) Silwet L77 (Lehle Seeds). Ripe seeds were sown on soil for BASTA selection (0.5% (v/v) BASTA, 0.5% (v/v) TWEEN 20, sprayed on seedlings three to four

times or alternatively sterile on plate.

4-4-9 *In vitro* Germination of Pollen Tubes

In vitro germination of pollen tubes was performed with minor adjustments after (Rodriguez-Enriquez *et al.,* 2013).

4-4-10 Hand Pollinations

Floral stage 12c pistils were hand pollinated with pollen of floral stage 13 stamen at least six hours prior to investigation.

4-5 Lipid Assays

In the following fatty acids and sphingosines are combined under the term lipids.

4-5-1 Radioactivity Measurements

Radioactivity of tritium-labeled compounds was measured by liquid scintillation counting using Rotiszint® eco plus LSC-Universal cocktail (Carl Roth GmbH + Co. KG), and LS 6000SC (Beckman Coulter GmbH), with a determined device counting efficiency of 40%.

4-5-2 Lipid Handling

For lipid handling direct exposure to light was avoided when possible. Lipids were stored water-free at -20°C and dissolved in $CHCl_3$/MeOH (1:3). Non radioactive PHS was obtained from Enzo Life Sciences Inc., and 2-OH PA was obtained from Sigma-Aldrich Co. LLC., all with > 98% purity. All purchased lipids were dissolved in EtOH.

4-5-3 Conversion of [9,10] ^{3}H PA in ^{3}H PHS and Breakdown Products

[9,10] ^{3}H PA (ART0129) was purchased from HARTMANN ANALYTIC GmbH with 1.11-2.22 TBq/mmol, and fed to logarithmic growing yeast of an OD_{600} of approximately 0.3, along with 5µmol/mL L-serine and incubated for 4h more hours.

4-5-4 Lipid extraction from Cell Pellets

This procedure was modified after (Browse *et al.,* 1986) and (Lester *et al.,* 2013). Yeast cells were harvested at approximately 3000 xg, washed with cold 5% TCA, twice with cold water, then shock frozen in liquid nitrogen for storage at -20°C. Cells were resupended in 1mL 1M HCl in MeOH, 0.6% (v/v) Dimethoxypropane, incubated for 20min at 80°C in a pyrex glass extraction tube with a teflon lined cap. Subsequently, 1.2mL 0.9% NaCl and 0.6mL $CHCl_3$ were added, vortexed, put on ice and the phases were separated for 10min at approx. 3000 xg. The lower phase was removed and dried in the eppendorf Concentrator plus at 45°C (V-AL).

4-5-5 Lipid Extraction from the Growth Medium

This procedure was modified after (Bligh *et al.,* 1959), see schematic (Figure 4-2). To 80mL of cell-free growth medium 200mL MeOH, then 100mL $CHCl_3$ was added to obtain a monophasic solution which was, if needed, stored over night at 4°C. Then successively 20mL 1.8M NaCl, 15mL H_2O, and 35mL $CHCl_3$ were added, mixed and transferred into a separatory funnel. After phase separation, the lower phase of about 180mL was transferred into a round bottom flask. This volume was evaporated with an rotary evaporator, stirring in water bath at 42°C – 45°C,

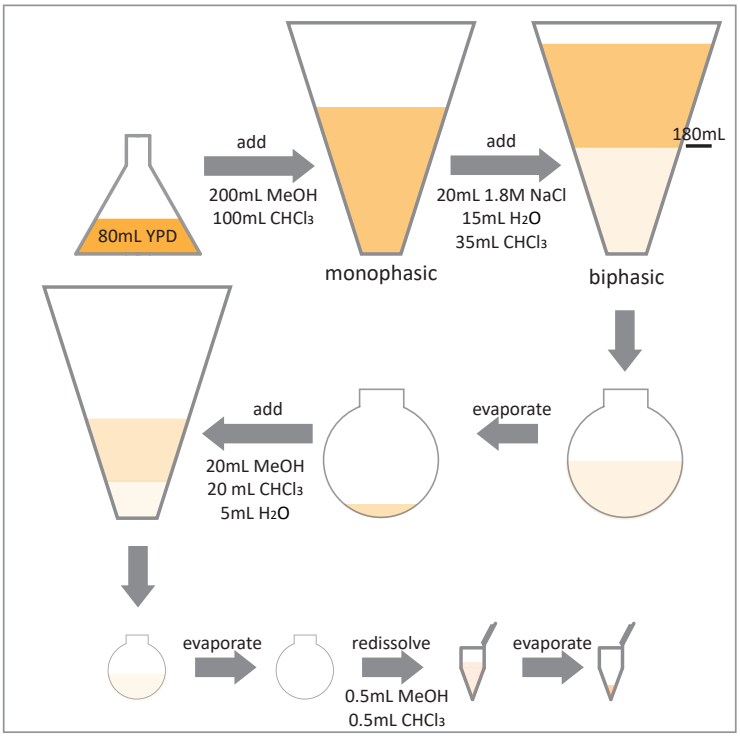


Figure 4-2 Lipid extraction from YPD medium.
Schematic of the workflow established for successful 2-hydroxy palmitic acid extraction from relative large volumes of cell growth medium. $CHCl_3$ = chloroform, MeOH = methanol.

and water jet-derived low vaccum. The remaining low-volume watery phase was supplemented with 20mL MeOH, 20mL $CHCl_3$ and transferred into the separatory funnel, where 5mL H_2O was added and mixed. Phases were allowed to separate. The lower phase was transferred into a 100mL round bottom flask, attached to the rotary evaporator and evaporated under same conditions as before. The dry flask was cooled down and then lipids were eluted with 1mL MeOH/$CHCl_3$ (1:1), which was evaporated in the eppendorf Concentrator plus at 45°C (V-AL), finally lipid extracts were resuspended in 200µL MeOH/$CHCl_3$ (1:3).

4-5-6 Lipid Separation by Thin-layer Chromatography and Detection

Thin-layer chromatography (Stahl 1956) was done by spotting lipid extracts on TLC Silica gel 60 plates (Merck KGaA), 1.5cm from the bottom and at least 1cm from the sides and resolved by 120mL of mobile phase in a gas-saturated glass tank. For one-step TLC Hexane:Dieethylether:AcOH 30:70:1 mobile phase was used and the mobile phase was migrated to at least two thirds of the plate height.

Detection of tritium-labeled compounds was done by coating plates with 7% (w/v) 2,5-diphenyl oxazole in diethylether (Randerath 1970), then exposed to Super HR-HA 30 X-ray films (FujiFilm Corporation) at – 80°C. 2-OH PA was detected by coating the silica plate using $3x10^{-5}$ % (w/v) Rhodamine6G in 95% EtOH (Kishimoto *et al.,* 2001) and scanning with Typhoon9500, plus control software v1.0 (GE Healthcare) excitation at 532nm, detection with 575LP long pass filter.

4-5-7 Retardation Factor Calculation

The retardation factors were calculated by the ratio of the distance of the mobile phase migration to the middle of the respective migrating zone (Sherma 2008).

4-6 Microscopy-based Methods

All microscopy was performed with 1.5mm coverslips.

4-6-1 Flower, Ovule and Female Gametophyte Developmental Stage Classification

Determination of ovule developmental stages according to (Schneitz *et al.,* 1995) in floral stage pistils defined by (Smyth *et al.,* 1990) and female gametophyte developmental stage according to (Christensen *et al.,* 1997).

4-6-2 CLSM

Confocal laser scanning microscopy was achieved with TCS SP8 (Leica Microsystems GmbH). GFP, and Schiff-reagent excitations with 488nm, 20% output power, detection with Hybrid detectors and AOBS set to 495nm – 545nm and 570nm – 630nm, respectively. mCherry, excitation was done with 561nm, detection with Hybrid detectors and AOBS set to 570nm – 630nm. All images were taken with Plan Apochromat 40x oil objective.

4-6-3 DIC Microscopy

DIC microscopy was done using Imager.M2 with Plan APO-CH 40/1.4 DIC VIS-R objective, Axiocam Color 105, and ZEN 2 (blue edition) software (Carl Zeiss Microscopy GmbH).

4-6-4 Silique Clearing

Mature but unopened siliques were incubated over night in EtOH/AcOH (3:1), transferred to 70% EtOH, then imaged at SteREO Discovery, PlanApo S 0.63x, AxioCam MRc5 (Carl Zeiss Microsystems GmbH).

4-6-5 Clearing of Ovules

Clearing of sliced pistils, ovules and early seeds and imaging with DIC was performed after (Hejatko *et al.,* 2006).

4-6-6 Feulgen Staining

Feulgen staining with Schiff-reagent was performed after (Barrell *et al.,* 2005).

4-6-7 GUS Staining

Histochemical Beta-glucuronidase (GUS) staining (Jefferson *et al.,* 1987) was performed on pisils in 25mM PO_4 buffer pH7.0, 1.25mM $K_3Fe(CN)_6$, 1.25mM $K_4Fe(CN)_6$, 0.25% (v/v) Triton X-100 and 1mg/ml X-GLUC (Thermo Fisher Scientific Inc.). Staining was performed from 30min to over night at 37°C in microtiter plates.

4-7 Protein-based Methods

4-7-1 Extraction of Proteins from Plant Tissue

Approximately 100mg frozen tissue (fresh weight) was ground (cold) in a rotatory mill using metal beads, then solubilized in 2mL 0.5M HEPES pH7.5, 150mM NaCl, 10% (w/v) sucrose, 3mM DTT, 1mM EDTA, 1mM PMSF, 1mM Benzamidine, 0.5% (v/v) NP40. Cellular debris was precipitated at 16000 xg and 35000 xg at 4°C for 10 and 15min, respectively. Resulting crude extract was used for IP using GFP-Trap®_M (Chromotek). The GFP-Trap® beads were incubated with the protein extract for 1h at 4°C, washed, one time with W1 0.5M Hepes pH7.5, 150mM NaCl, 3mM DTT, 0.5% (v/v) NP40, two times with W2 0.5M Hepes pH7.5, 450mM NaCl, 3mM DTT, 0.5% (v/v) NP40, and finally one time with W1. Elution of GFP tagged proteins from the beads was achieved by 0.2M glycine pH2.5, followed by magnetical separation and neutralized with 1M Tris HCl pH10.3. Eluted protein was dissolved in SDS loading buffer (50mM Tris-HCl pH6.8, 2% (w/v) SDS, 0.1% bromphenolblue, 20% (v/v) glycerol, 100mM DTT) heated to 42°C for 10min and subjected to SDS page (Laemmli 1970).

4-7-2 Microsomal fractionation of Callus Cell extracts for Mass Spectrometric Analysis

This method was performed with minor modifications after (Uebler 2016). Approximately 200mg frozen callus samples were pulverized in mortar and pestle, and ressolubilized in 2mL of microsomal fractionation buffer, cell debris was precipitated 16k xg, 4°C, supernatant was precipitated at 100000 xg, the pellet washed twice, resolubilized and precipitated two more times. Resulting pellets were solubilized and stored at -20°C. Subsequently 10µg per sample were subjected to mass spectrometry-based analysis.

4-7-3 Protein Precipitations of Callus Cells for Full Proteomic Analysis by Mass Spectrometry

Approximately 200mg of frozen callus sample was pulverized in a cold mortar and pestle. For TCA-mediated protein precipitation (Debro *et al.,* 1956), the sample was resolubilized (for buffer see 4-7-1) then major cellular debris was precipitated 10min at 16000 xg at 4°C. The clear supernatant was supplemented with 1/10 Vol 100% (v/v) TCA, mixed, incubated for 1h at 4°C, spun down at 16000 xg for 30min and subsequently washed with acetone. The pellet was air dryed. The membrane protein precipitation with AMS was performed after (Hurkman *et al.,* 1986). Subsequently obtained proteins were subjected to mass spectrometry-based analysis (10µg per sample).

4-7-4 Mass Spectrometric Analysis (by Dr. Julia Mergner)

For tryptic digestion samples were reduced and alkylated with 10 mM DTT and 55 mM chloroacetamide, respectively. Tryptic in-gel digestion was performed according to standard procedures. Nanoflow liquid chromatography tandem mass spectrometry was performed with a Q-Exactive HF (Thermo Fisher Scientific) mass spectrometer coupled to an EksigentnanoLC-Ultra 1D+ (AB Sciex) or Dionex Ultimate 3000 (Thermo Fisher Scientific). Online peptide fractionation was performed with in-house packed C18 columns using either a 60 min or 110 min linear gradient from 4 to 32 % of solvent B (0.1% formic acid and 5 % DMSO in acetonitrile; solvent A: 0.1 % formic acid and 5 % DMSO in water) at a flow rate of 300 nL/min. The setup consisted of a trap column with 75 um x 2 cm, packed with 5 um particels of ReprosilPur ODS-3 (Dr. Maisch GmbH) and a 75 µm x 40 cm analytical column packed with 3 µm particels of C18 Reprosil Gold 120 (Dr. Maisch GmbH). The Q-Exactive HF was operated in data dependent mode automatically switching between MS (60,000 resolution, 3 x 10^6 AGC target, 50 ms maximum injection time, 360-1,300 m/z scan range, profile mode) and MS2 (15,000 resolution, 1 x 10^5 AGC target, 50 ms maximum injection time, top 20, 1.7 m/z isolation width, 20 s dynamic exclusion, centroid mode). Intensity-based label-free quantification was performed using MaxQuant (version 1.5.3.8) (Cox *et al.,* 2008). Spectra were searched against the TAIR10 Arabidopsis database and the internal contaminants database from MaxQuant(TAIR10_pep_20101214_ updated.fasta, download October 2015). The resulting list of identified proteins was further

filtered for hits from the reverse decoy database, contaminating proteins, proteins only identified by a modified peptide and non-quantifiable proteins. For statistical analysis missing values between the replicate measurements were replaced by random numbers drawn from the normal distribution of the measured data (width 0.3, down shift 1.8 applied separately for each column) using Perseus (version 1.5.5.3) (Tyanova *et al.,* 2016). A two-tailed students t-test (S0 = 0.1, p-vaue = 0.05) was performed to identify significant different fold changes of the $\log_2$ transformed LFQ intensities between EC-like callus and CIM callus samples (Cox *et al.,* 2014).

4-7-5 Measurement of Protein Concentrations

Protein concentrations were determined by Bradford assay (Bradford 1976) using the Protein Assay Dye Reagent (Bio-Rad Laboratories Inc.) according to manufacturers recommendations and measured at the Nanodrop2000 (Thermo Fisher Scientific Inc.). BSA was used to generate a reference curve.

4-7-6 SDS PAGE

SDS PAGE (Laemmli 1970) was performed after heating the samples in SDS loading buffer (50mM Tris-HCl pH6.8, 2% (w/v) SDS, 0.1% bromphenolblue, 20% (v/v) glycerol, 100mM DTT) 10min at 42°C or 95°C for membrane, or soluble proteins, respectively. Electrophoresis was performed in Mini-PROTREAN® (Bio-Rad) for 20min at 10V per gel then to full resolution at 30V per gel.

4-7-7 Staining of SDS gels

After electrophoresis SDS gels were stained over night in CoomassieG250 (Candiano *et al.,* 2004) and destained with water.

4-7-8 Western Blot

After SDS PAGE, proteins were transferred (Burnette 1981) to Amersham Protran 0.45 nitrocellulose (GE Healthcare), in Mini Trans-Blot® Cell (Bio-Rad Laboratries Inc.), 365mA, 1h, stained with Ponceau S (Bannur *et al.,* 1999), blocked with 5% skimmed milk powder or BSA

in TBST 90min at RT or at 4°C over night, probed with 1:2000 anti-GFP antibody (monoclonal mouse, Roche), 4°C over night, and 1:1000 goat anti-mouse-HRP (Sigma) 90min RT, in TBST. Chemiluminescence signals were detected on a Fluorchem FC2 ECL imager using the Pierce ECL Western blotting Substrate.

4-8 WISH

If not stated otherwise all reagents were purchased from Roche Diagnostics Deutschland GmbH. 5µg of the circular pCR2.1 vector containing the respective target sequences were linearized by either EcoRV-HF or BamHI-HF (New England Biolabs GmbH) over night, then spin column purified by QIAquick PCR purification kit (Qiagen GmbH), eluted with 35µl water, and *in vitro* transcribed with Sp6 or T7 DNA-dependent RNA polymerase, respectively (Hejatko *et al.,* 2006). Probe sizes and purity were checked by denaturing agarose gel. The entire hybridization procedure was performed using InsituPro VSi with software v3.1 (Intavis AG) according to given protocol (Hejatko *et al.,* 2006), starting after permeabilization to circumvent damaging the instrument. All prior steps were performed in glass dishes. Detection was done using 1:2000 anti-DIG-AP antibody and colorimetric staining using BCIP and NBT performed in microtiter plates and stopped after 20 minutes to 4 hours.

4-9 The TET9 Protein Interactor Screen using the mbSUS Assay (service by Creative Biolabs)

The identification of putative TET9 interaction partners by the split ubiquitin system (mbSUS) (Obrdlik *et al.,* 2004) was provide as service by Creative™ Biolabs (Project CBLJ11171403B). Therefore, total RNA was extracted from the EC-like cell line, subsequently transcribed into flanked cDNA, and cloned into pPR3-N for prey-library construction, which resulted in 1.25E+07 transformants with an average cDNA-derived insert size of more than 1.2kbp. The cDNA sequences were fused to the C-terminus of the N-terminal half of ubiquitin. Furthermore, a Gateway system compatible *TET9* CDS Entry vector was used for bait construction. Here the C-terminal end of ubiquitin, and the artificial transcription factor LexA-VP16 were fused translationally in frame to the N-terminus of TET9. Upon dimerization of the split ubiquitin at the cytosolic side of the plasma membrane, due to prey and bait protein proximity, the LexA-

VP16 transcription factor was cleaved off and consequently the ß-galactosidase reporter, and the auxotrophic markers *HIS3*, and *ADE2* were activated. To negate self-activation LexA-VP16 cleavage and subsequent activation of the reporter genes, Cub-TET9 was cotransformed with the empty library prey vector pPR3-N and the stringency of selective medium was determined sufficient to suppress autoactivation by using SD-Leu-Trp-His plates. In the following, the TET9 bait, and the EC-like callus-derived cDNA prey library were cotransformed, plated on selective medium as determined, positive clones were restreaked on selective medium and finally tested for ß-galactoidase activity. Of resulting 23 verified positive clones plasmid DNA was isolated, then sequenced for identification with BLAST-based search, which resulted in nine putative TET9-interacting proteins.

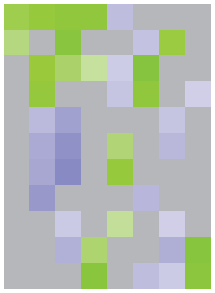
5 – Supplement

5-1 Supplemental Figures

A

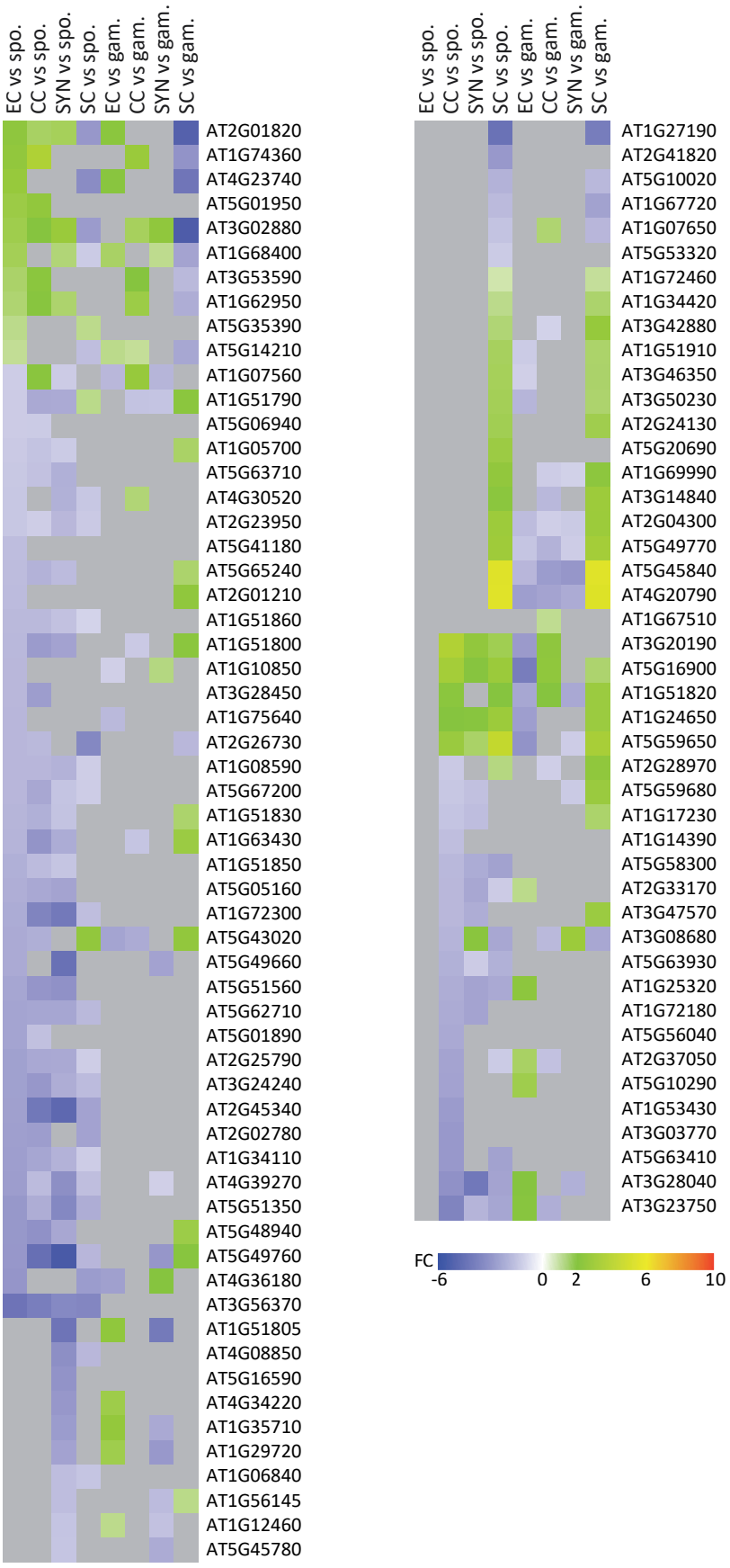
AGI	description	loc.	TM α	TM β
AT4G11460	cysteine-rich RLK (RECEPTOR-like protein kinase) 30 (CRK30)	PM	1	
AT4G23180	cysteine-rich RLK (RECEPTOR-like protein kinase) 10 (CRK10)	PM	1	
AT4G21230	cysteine-rich RLK (RECEPTOR-like protein kinase) 27 (CRK27)	PM	1	
AT4G23310	cysteine-rich RLK (RECEPTOR-like protein kinase) 23 (CRK23)	PM		1
AT1G70520	cysteine-rich RLK (RECEPTOR-like protein kinase) 2 (CRK2)	PM	1	
AT4G23320	cysteine-rich RLK (RECEPTOR-like protein kinase) 24 (CRK24)	PM	1	
AT4G38830	cysteine-rich RLK (RECEPTOR-like protein kinase) 26 (CRK26)	PM	1	
AT4G23260	cysteine-rich RLK (RECEPTOR-like protein kinase) 18 (CRK18)	PM	1	
AT4G23300	cysteine-rich RLK (RECEPTOR-like protein kinase) 22 (CRK22)	PM	1	
AT4G11530	cysteine-rich RLK (RECEPTOR-like protein kinase) 34 (CRK34)	PM	1	
AT4G21400	cysteine-rich RLK (RECEPTOR-like protein kinase) 28 (CRK28)	PM	1	
AT4G23290	cysteine-rich RLK (RECEPTOR-like protein kinase) 21 (CRK21)	PM	1	
AT3G45860	cysteine-rich RLK (RECEPTOR-like protein kinase) 4 (CRK4)	PM	1	
AT4G21410	cysteine-rich RLK (RECEPTOR-like protein kinase) 29 (CRK29)	PM	1	
AT4G23270	cysteine-rich RLK (RECEPTOR-like protein kinase) 19 (CRK19)	PM	1	
AT4G23190	cysteine-rich RLK (RECEPTOR-like protein kinase) 11 (CRK11)	PM	1	
AT4G23210	cysteine-rich RLK (RECEPTOR-like protein kinase) 13 (CRK13)	PM	1	
AT4G11490	cysteine-rich RLK (RECEPTOR-like protein kinase) 33 (CRK33)	PM	1	
AT4G04510	cysteine-rich RLK (RECEPTOR-like protein kinase) 38 (CRK38)	PM	1	
AT5G40380	cysteine-rich RLK (RECEPTOR-like protein kinase) 42 (CRK42)	PM	2	

B

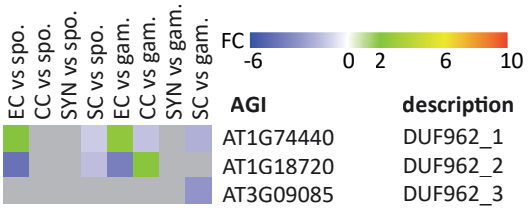


AGI	description	loc.	TM α
AT1G61490	putative S-domain-type receptor protein kinase	PM	1
AT1G61500	putative S-domain-type receptor protein kinase	PM	1
AT1G61440	putative S-domain-type receptor protein kinase	PM	2
AT1G61390	putative S-domain-type receptor protein kinase	PM	2
AT1G11410	putative S-domain-type receptor protein kinase	PM	1
AT1G11330	putative S-domain-type receptor protein kinase	PM	1
AT4G27300	putative S-domain-type receptor protein kinase	PM	1
AT5G35370	putative S-domain-type receptor protein kinase	PM	1
AT4G11900	putative S-domain-type receptor protein kinase	PM	1
AT1G61370	putative S-domain-type receptor protein kinase	PM	1
AT1G11340	putative receptor-like protein kinase (AtRKS1)	PM	2

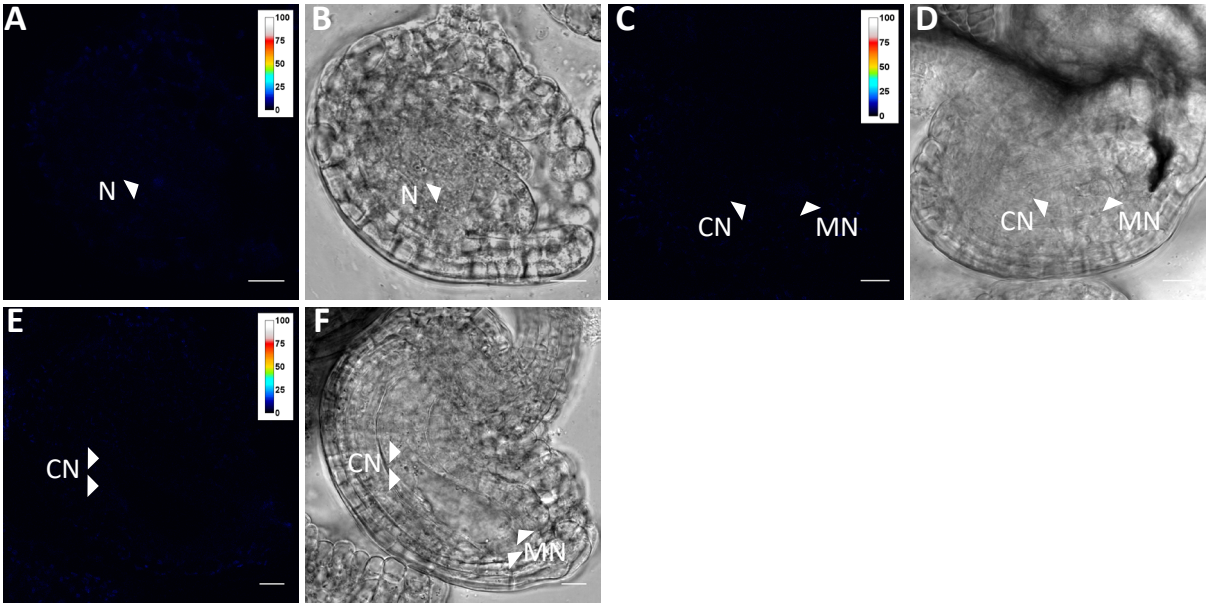
Supplemental Figure 5-1 differentially expressed cysteine-rich and S-domain-type RLKs in *Arabidopsis* female gametophytic cells and male gametes.
The transcriptomic approach identified 20 cysteine-rich RLKs (**A**) and 11 S-domain-type RLKs (**B**) as differentially expressed in the female gamtophytic cells and the male gametes. FC = log2 fold change, gam. = gametophyte, loc. = localization according to SUBAcon, PM = plasma membrane, spo. = sporophyte, TMα = Aramemnon-based prediction for membrane alpha helices, TMβ = Aramemnon-based prediction for membrane beta sheets.



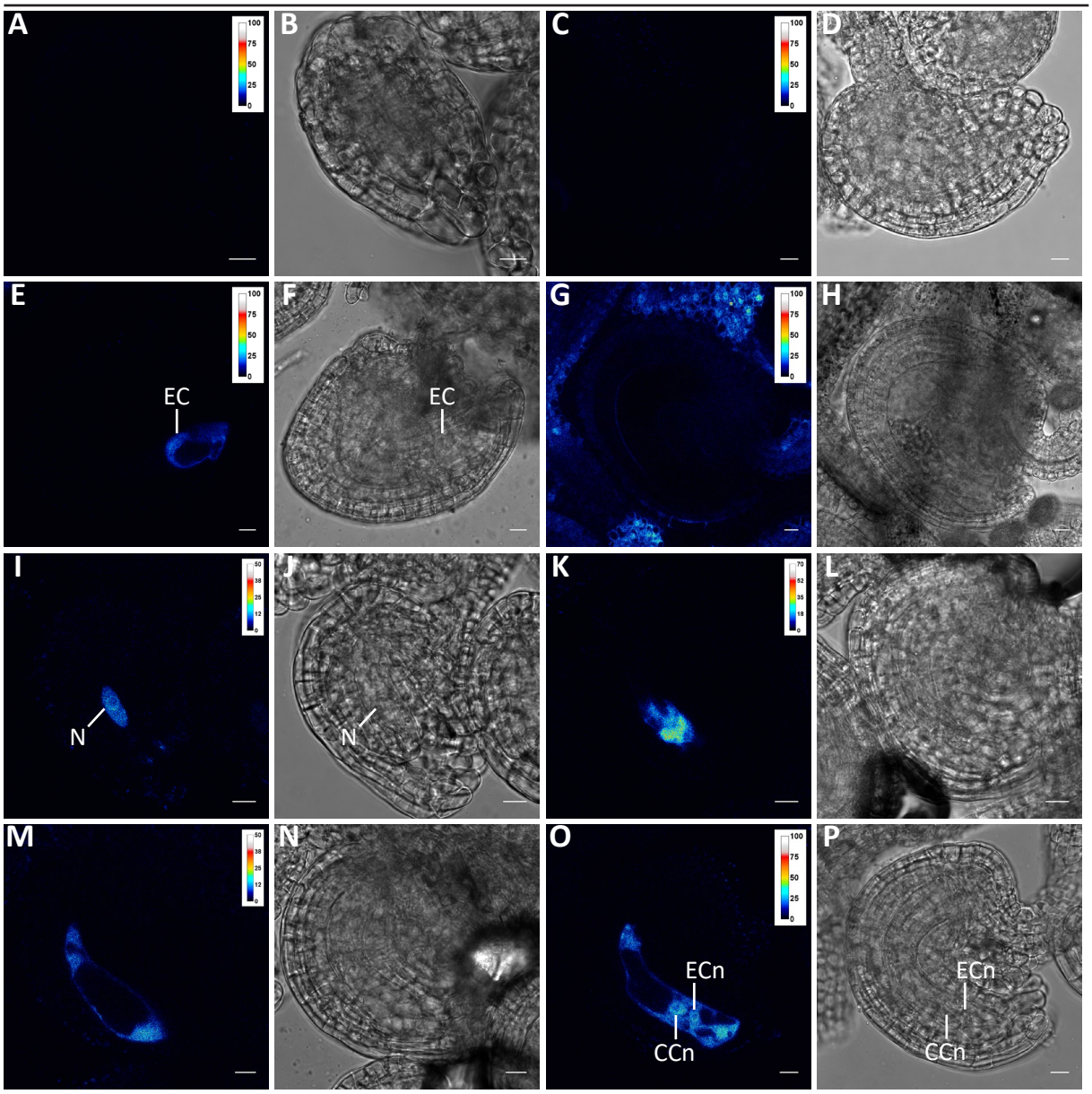
Supplemental Figure 5-2 differentially expressed LRR-RLKs in *Arabidopsis* female gametophytic cells and male gametes.
By the transcriptomic approach 104 differentially expressed LRR-RLKs of the female gametophytic cells and the male gametes were identified. FC = log2 fold change, gam. = gametophyte, spo. = sporophyte



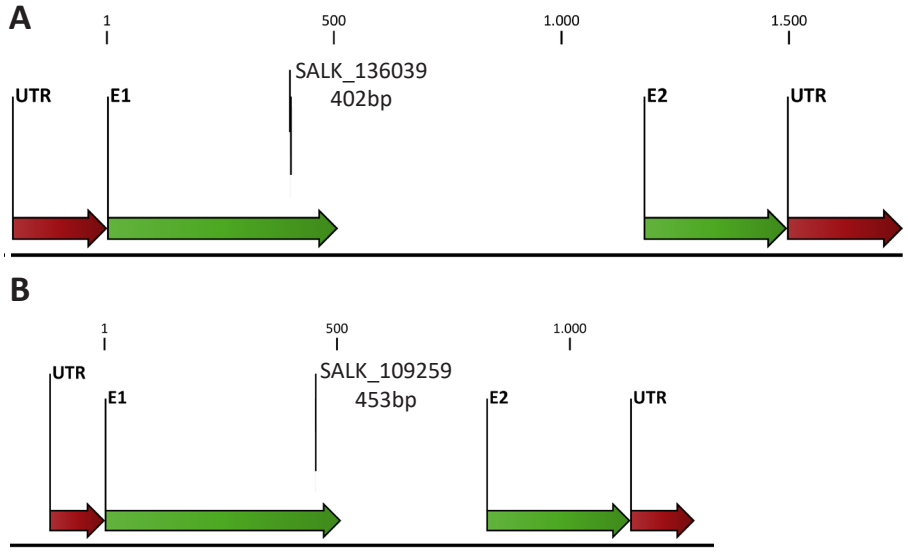
Supplemental Figure 5-3 *DUF962* gene expression in *Arabidopsis* female gametophytic cells and male gametes.
Heatmap depiction of DUF962 gene expression determined by transcriptomic studies of single isolated gametophytic cells. AT1G74440 (*DUF962_1*) was found upregulated in the EC (FC = 2.05 vs sporophyte, FC = 2.49 vs gametophyte), while AT1G18720 (*DUF962_2*) was downregulated in the EC (FC = -3.24 vs sporophyte), and SC (FC = -1.22 vs sporophyte) but found upregulated in the CC when contrasted to the gametophyte (FC = 2.28). AT3G09085 (*DUF962_3*) was found not differentially expressed against the sporophyte. FC = log2 fold change, gam. = gametophyte, spo. = sporophyte.



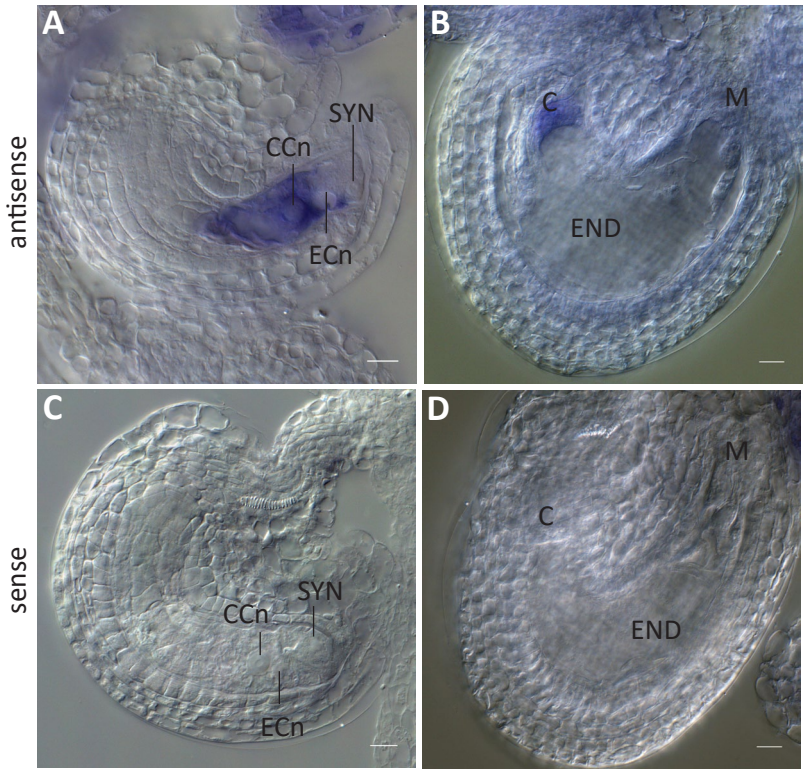
Supplemental Figure 5-4 DUF962_1g:GFP abundance during the female gametophyte development.
Neither in stage FG1 (**A, B**), nor in stage FG3 (**C, D**), or in late stage FG4 (**E, F**) GFP-derived fluorescent signal could be detected by CLSM analysis of *DUF962_1g:GFP*-expression plants. CN = chalazal nucleus, N = functional megaspore, MN = micropylar nucleus. Scale is 10µm.



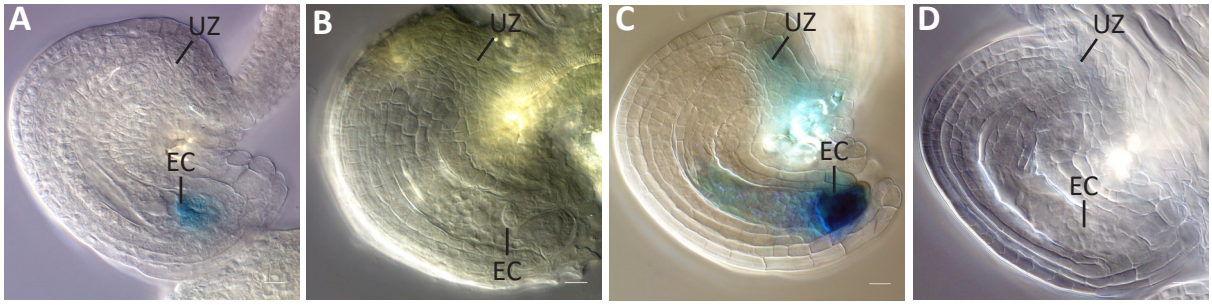
Supplemental Figure 5-5 Detection of GFP expressed by *EC1.1p, 35Sp*, and *AT1G21670p* in developing female gametophytes.
EC1.1p-expressed GFP was neither detected in stage FG1 (**A, B**), nor stage FG2 (**C, D**) but only in mature ECs (**E, F**). GFP was absent in mature ovules but detected in the funiculus of *35Sp:GFP* plant lines (**G, H**). In *AT1G21670p:GFP* plant lines, GFP was detected in the developing female gametophyte of stage FG1 (**I, J**), stage FG2/3 (**K, L**), FG4 (**M, N**), and stage FG6 (**O, P**). ECn = egg cell nucleus, CCn = central cell nucleus, N = functional megaspore. Scale is 10µm.



Supplemental Figure 5-6 *tet8* and *tet11* T-DNA insertion lines.
The genomic loci of *TET8* (**A**), and *TET11* (**B**) as obtained from TAIR v10, including mapped T-DNA insertion postions. E = exon, UTR = untranslated region.



Supplemental Figure 5-7 WISH on mature and early seeds for detection of *CDR1*.
WISH was performed on mature ovules (**A, C**) and early seeds (**B, D**) with DIG-labelled antisense (**A, B**) and sense (**C, D**) probes for *CDR1* transcript detection. ECn = egg cell nucleus, END = endosperm, CCn = central cell nucleus, SYN = synergid cell. Scale is 10µm.



Supplemental Figure 5-8 GUS-staining of *EC1.1p:GUS* plant lines and the WT.
Mature ovules of *EC1.1p:GUS* expressing plant lines (**A, C**) and the WT (**B, D**) were stained 30min (**A, B**) and over night (**C, D**). All eight *EC1.1p:GUS* plant lines showed staining of the EC, extended staining showed additional slight GUS signal from within the embryo sac and the vasculature unloading zone. The WT showed no staining. EC = egg cell, UZ = region of the vasculature unloading zone. Scale is 10µm.

5-2 Supplemental Tables

Supplemental Table 5-1 Analyzed CEL files, their names and respective batches.
For bioinformatic analyses microarray data of (Soljic 2012) and publicly available data obtained from the gene expression omnibus and array express were analyzed comprehensively.

CEL-file	Name	RNA (batch)	Target
A308b_2C.CEL	central_cell.1	Soljic_2012_01	central_cell
A308d_C2.CEL	central_cell.2	Soljic_2012_02	central_cell
A308d_C3.CEL	central_cell.3	Soljic_2012_02	central_cell
CentralCell1.CEL	central_cell.4	Wuest	central_cell
CentralCell2.CEL	central_cell.5	Wuest	central_cell
CentralCell3.CEL	central_cell.6	Wuest	central_cell
A308_eggcells_KFB_labeled.CEL	egg_cell.1	Soljic_2012_01	egg_cell
A308d_E2.CEL	egg_cell.2	Soljic_2012_02	egg_cell
A308d_E3.CEL	egg_cell.3	Soljic_2012_02	egg_cell
A308b_1E.CEL	egg_cell.4	Soljic_2012_01	egg_cell
Egg1.CEL	egg_cell.5	Wuest	egg_cell
Egg2.CEL	egg_cell.6	Wuest	egg_cell
Egg3.CEL	egg_cell.7	Wuest	egg_cell
GSM134299.CEL	endosperm.1	GSE5751	endosperm
GSM134300.CEL	endosperm.2	GSE5751	endosperm
GSM134301.CEL	endosperm.3	GSE5751	endosperm
GSM1010606_101423-005.CEL	endosperm.4	41212	endosperm
GSM1010607_101423-006.CEL	endosperm.5	41212	endosperm
GSM1010608_101423-007.CEL	endosperm.6	41212	endosperm
GSM133776.CEL	globular-apical_Rep1	GSE5730	globular.embryo
GSM133778.CEL	globular-apical_Rep2	GSE5730	globular.embryo
GSM133780.CEL	globular-apical_Rep3	GSE5730	globular.embryo
GSM133777.CEL	globular-basal_Rep1	GSE5730	globular.embryo
GSM133779.CEL	globular-basal_Rep2	GSE5730	globular.embryo
GSM133781.CEL	globular-basal_Rep3	GSE5730	globular.embryo
GSM1468520_nEMB__1.CEL	glob.emb.7	60242	globular.embryo
GSM1468521_nEMB__2.CEL	glob.emb.8	60242	globular.embryo
GSM1468522_nEMB__3.CEL	glob.emb.9	60242	globular.embryo
GSM265846.CEL	leaf.1	GSE10522	leaf
GSM265852.CEL	leaf.2	GSE10522	leaf
GSM265858.CEL	leaf.3	GSE10522	leaf
GSM265864.CEL	leaf.4	GSE10522	leaf
GSM826259.CEL	leaf.5	GSE10522	leaf
GSM738134.CEL	leaf.6	29771	leaf
GSM738135.CEL	leaf.7	29771	leaf
GSM738136.CEL	leaf.8	29771	leaf
PSB_COL_1.CEL	leaf.9	3045	leaf
PSB_COL_2.CEL	leaf.10	3045	leaf
PSB_COL_3.CEL	leaf.11	3045	leaf

Supplemental Table 5-1 (continued)

CEL-file	Name	RNA (batch)	Target
GSM618324.CEL	roots.1.1	25171	root
GSM618326.CEL	roots.1.2	25171	root
GSM618328.CEL	roots.1.3	25171	root
GSM618330.CEL	roots.1.4	25171	root
GSM618332.CEL	roots.2.1	25171	root
GSM618334.CEL	roots.2.2	25171	root
GSM618336.CEL	roots.2.3	25171	root
GSM618338.CEL	roots.2.4	25171	root
GSM618340.CEL	roots.3.1	25171	root
GSM618342.CEL	roots.3.2	25171	root
GSM618344.CEL	roots.3.3	25171	root
GSM618346.CEL	roots.3.4	25171	root
GSM179958.CEL	roots.4.1	7432	root
GSM179959.CEL	roots.4.2	7432	root
GSM179971.CEL	roots.4.3	7432	root
GSM179972.CEL	roots.4.4	7432	root
GSM131217.CEL	seedling.01	GSE5618	seedling
GSM131219.CEL	seedling.02	GSE5618	seedling
GSM131176.CEL	seedling.03	GSE5617	seedling
GSM131184.CEL	seedling.04	GSE5617	seedling
GSM131191.CEL	seedling.05	GSE5617	seedling
GSM131198.CEL	seedling.06	GSE5617	seedling
GSM131206.CEL	seedling.07	GSE5617	seedling
GSM131214.CEL	seedling.08	GSE5617	seedling
GSM131847.CEL	seedling.09	GSE5641	seedling
GSM131848.CEL	seedling.10	GSE5641	seedling
GSM131849.CEL	seedling.11	GSE5641	seedling
GSM439767.CEL	seedling.12	17610	seedling
GSM439768.CEL	seedling.13	17610	seedling
GSM439769.CEL	seedling.14	17610	seedling
At_Sperm_Rep1_ATH1_IGC_FB.CEL	sperm.1	35	sperm_cell
At_Sperm_Rep2_ATH1_IGC_FB.CEL	sperm.2	35	sperm_cell
At_Sperm_Rep3_ATH1_IGC_FB.CEL	sperm.3	35	sperm_cell
A308_c_S1.CEL	synergid_cell.1	Soljic_2012_01	synergid_cell
A308d_S2.CEL	synergid_cell.2	Soljic_2012_02	synergid_cell
A308d_S3.CEL	synergid_cell.3	Soljic_2012_02	synergid_cell
Synergids1.CEL	synergid_cell.4	Wuest	synergid_cell
Synergids2.CEL	synergid_cell.5	Wuest	synergid_cell
Synergids3.CEL	synergid_cell.6	Wuest	synergid_cell

Supplemental Table 5-2 DEGs identified by gametophytic contrasting.
FC ≥ 1.5 pV≤ 0.001 upregulated DEGs shared by multiple cells of the female gametophyte or the sperm cells after gametophytic contrasting. Candidates are boxed.

Cell type	AGI	description
EC + CC	AT5G10450	a role in mediating oxidative metabolism in stress response
EC + CC	AT4G25450	non-intrinsic ABC protein 8 (NAP8)
EC + CC	AT4G20850	tripeptidyl peptidase ii (TPP2)
EC + CC	AT1G20980	squamosa promoter binding protein-like 14 (SPL14)
EC + CC	AT3G13360	WPP domain interacting protein 3 (WIP3)
EC + CC	AT5G15090	voltage dependent anion channel 3 (VDAC3)
EC + CC	AT3G09690	alpha/beta-Hydrolases superfamily protein
EC + CC	AT5G35630	glutamine synthetase 2 (GS2)
EC + CC	AT1G14820	Sec14p-like phosphatidylinositol transfer family protein
EC + CC	AT2G39450	MTP11
EC + CC	AT3G14920	Peptide-N4-(N-acetyl-beta-glucosaminyl)asparagine amidase A protein
EC + CC	AT5G03350	Legume lectin family protein
EC + CC	AT4G05160	AMP-dependent synthetase and ligase family protein
EC + CC	AT3G24190	Protein kinase superfamily protein
EC + CC	AT3G55090	ABC-2 type transporter family protein
EC + CC	AT5G65020	annexin 2 (ANNAT2)
EC + CC	AT2G21330	fructose-bisphosphate aldolase 1 (FBA1)
EC + CC	AT1G73990	signal peptide peptidase (SPPA)
EC + CC	AT5G17930	MIF4G domain-containing protein / MA3 domain-containing protein
EC + CC	AT5G22640	embryo defective 1211 (emb1211)
EC + CC	AT5G03940	Chloroplast Signal Recognition Particle Subunit
EC + CC	AT1G63900	E3 Ubiquitin ligase family protein
EC + CC	AT5G67470	formin homolog 6 (FH6)
EC + CC	AT5G57200	ENTH/ANTH/VHS superfamily protein
EC + CC	AT1G78400	Pectin lyase-like superfamily protein
EC + CC	AT4G30430	tetraspanin9 (TET9)
EC + CC	AT3G14660	cytochrome P450, family 72, subfamily A, polypeptide 13 (CYP72A13)
EC + CC	AT5G20720	chaperonin 20 (CPN20)
EC + CC	AT3G01280	voltage dependent anion channel 1 (VDAC1)
EC + CC	AT3G22630	20S proteasome beta subunit D1 (PBD1)
EC + CC	AT4G29735	unknown protein
EC + CC	AT3G51800	ATG2
EC + CC	AT3G11270	maternal effect embryo arrest 34 (MEE34)
EC + CC	AT4G19185	nodulin MtN21 /EamA-like transporter family protein
EC + CC	AT3G29185	molecular_function unknown
EC + CC	AT1G50250	FTSH protease 1 (FTSH1)
EC + CC	AT5G58980	Neutral/alkaline non-lysosomal ceramidase
EC + CC	AT2G35060	K+ uptake permease 11 (KUP11)
EC + SYN	AT5G13390	NO EXINE FORMATION 1 (NEF1)
EC + SYN	AT2G45710	Zinc-binding ribosomal protein family protein
EC + SYN	AT2G47730	glutathione S-transferase phi 8 (GSTF8)
EC + SYN	AT1G35510	O-fucosyltransferase family protein
EC + SYN	AT5G10840	Endomembrane protein 70 protein family
EC + SYN	AT1G20200	EMBRYO DEFECTIVE 2719 (EMB2719)
EC + SYN	AT1G15750	TOPLESS (TPL)
EC + SYN	AT3G13920	eukaryotic translation initiation factor 4A1 (EIF4A1)
EC + SYN	AT1G50740	Transmembrane proteins 14C
EC + SYN	AT3G52290	IQ-domain 3 (IQD3)
EC + SYN	AT3G08710	thioredoxin H-type 9 (TH9)
EC + SYN	AT5G24430	Calcium-dependent protein kinase (CDPK) family protein
EC + SYN	AT5G23580	recombinant protein is fully active and induced by Ca2+
EC + SYN	AT1G75760	ER lumen protein retaining receptor family protein
EC + SYN	AT5G64310	arabinogalactan protein 1 (AGP1)
EC + SYN	AT1G76400	Ribophorin I
EC + SYN	AT5G55050	GDSL-like Lipase/Acylhydrolase superfamily protein
EC + SYN	AT3G29810	COBRA-like protein 2 precursor (COBL2)
EC + SYN	AT5G04020	calmodulin binding

Supplemental Table 5-2 (continued)

EC + SYN	AT1G52910	Protein of unknown function (DUF1218)
EC + SYN	AT5G59845	Gibberellin-regulated family protein
EC + SYN	AT5G35380	Protein kinase protein with adenine nucleotide alpha hydrolases-like domain
EC + SYN	AT3G61880	cytochrome p450 78a9 (CYP78A9)
EC + SYN	AT4G35380	SEC7-like guanine nucleotide exchange family protein
EC + SYN	AT1G10540	nucleobase-ascorbate transporter 8 (NAT8)
EC + SYN	AT3G07970	QUARTET 2 (QRT2)
EC + SYN	AT4G18050	P-glycoprotein 9 (PGP9)
EC + SYN	AT2G02990	ribonuclease 1 (RNS1)
EC + SYN	AT3G21620	ERD (early-responsive to dehydration stress) family protein
EC + SYN	AT1G31450	Eukaryotic aspartyl protease family protein
EC + SYN	AT4G02080	secretion-associated RAS super family 2 (SAR2)
EC + SYN	AT5G18500	Protein kinase superfamily protein
EC + SYN	AT3G51550	FERONIA (FER)
EC + SYN	AT1G31440	SH3 domain-containing protein
EC + SYN	AT3G46000	actin depolymerizing factor 2 (ADF2)
EC + SYN	AT5G11740	arabinogalactan protein 15 (AGP15)
EC + SYN	AT3G52930	Aldolase superfamily protein
EC + SYN	AT4G27090	Ribosomal protein L14
EC + SYN	AT5G39740	ribosomal protein L5 B (RPL5B)
EC + SYN	AT1G53850	20S proteasome alpha subunit E1 (PAE1)
EC + SYN	AT1G04750	vesicle-associated membrane protein 721 (VAMP721)
EC + SYN	AT3G14600	Ribosomal protein L18ae/LX family protein
EC + SYN	AT5G58070	temperature-induced lipocalin (TIL)
EC + SC	AT5G11420	molecular_function unknown
EC + SC	AT1G12240	ATBETAFRUCT4
CC + SC	AT1G02810	Plant invertase/pectin methylesterase inhibitor superfamily
CC + SC	AT1G30740	FAD-binding Berberine family protein
CC + SC	AT1G51820	Leucine-rich repeat protein kinase family protein
CC + SC	AT5G46540	P-glycoprotein 7 (PGP7)
CC + SC	AT1G29630	5'-3' exonuclease family protein
CC + SC	AT1G64830	Eukaryotic aspartyl protease family protein
CC + SC	AT5G14870	cyclic nucleotide-gated channel 18 (CNGC18)
CC + SC	AT1G10680	P-glycoprotein 10 (PGP10)
CC + SYN	AT3G12700	Eukaryotic aspartyl protease family protein
CC + SYN	AT5G67080	mitogen-activated protein kinase kinase kinase 19 (MAPKKK19)
CC + SYN	AT1G21870	golgi nucleotide sugar transporter 5 (GONST5)
CC + SYN	AT1G17430	alpha/beta-Hydrolases superfamily protein
CC + SYN	AT3G10460	Plant self-incompatibility protein S1 family
CC + SYN	AT5G51260	HAD superfamily, subfamily IIIB acid phosphatase
CC + SYN	AT3G47040	Glycosyl hydrolase family protein
CC + SYN	AT3G18220	Phosphatidic acid phosphatase (PAP2) family protein
EC + CC + SYN	AT3G14990	Class I glutamine amidotransferase-like superfamily protein
EC + CC + SYN	AT4G14420	HR-like lesion-inducing protein-related
EC + CC + SYN	AT2G40765	unknown protein
EC + CC + SYN	AT2G19970	CAP
EC + CC + SYN	AT2G46140	Late embryogenesis abundant protein
EC + CC + SYN	AT1G78380	glutathione S-transferase TAU 19 (GSTU19)
EC + CC + SYN	AT1G17860	Kunitz family trypsin and protease inhibitor protein
EC + CC + SYN	AT5G16910	cellulose-synthase like D2 (CSLD2)
EC + CC + SYN	AT5G14180	Myzus persicae-induced lipase 1 (MPL1)
EC + CC + SYN	AT4G16500	Cystatin/monellin superfamily protein
EC + CC + SYN	AT4G24920	secE/sec61-gamma protein transport protein
EC + CC + SYN	AT5G58860	cytochrome P450, family 86, subfamily A, polypeptide 1 (CYP86A1)
EC + CC + SYN	AT2G17500	Auxin efflux carrier family protein
EC + CC + SYN	AT3G48930	embryo defective 1080 (EMB1080)
EC + CC + SYN	AT4G03070	AOP1
EC + CC + SYN	AT1G21750	PDI-LIKE 1-1

Supplemental Table 5-2 (continued)

EC + CC + SYN	AT2G47470	UNFERTILIZED EMBRYO SAC 5 (UNE5)
EC + CC + SYN	AT1G14080	fucosyltransferase 6 (FUT6)
EC + CC + SYN	AT3G05310	MIRO-related GTP-ase 3 (MIRO3)
EC + CC + SYN	AT4G38700	Disease resistance-responsive (dirigent-like protein) family protein
EC + CC + SYN	AT4G20460	NAD(P)-binding Rossmann-fold superfamily protein
EC + CC + SYN	AT1G74000	strictosidine synthase 3 (SS3)
EC + CC + SYN	AT2G31390	pfkB-like carbohydrate kinase family protein
EC + CC + SYN	AT5G38330	low-molecular-weight cysteine-rich 80 (LCR80)
EC + CC + SYN	AT5G14030	translocon-associated protein beta (TRAPB) family protein
EC + CC + SYN	AT5G10750	Protein of unknown function (DUF1336)
EC + CC + SYN	AT3G17230	invertase/pectin methylesterase inhibitor family protein
EC + CC + SYN	AT5G54860	Major facilitator superfamily protein
EC + CC + SYN	AT3G53370	S1FA-like DNA-binding protein
EC + CC + SYN	AT4G27500	proton pump interactor 1 (PPI1)
EC + CC + SYN	AT3G48950	Pectin lyase-like superfamily protein
EC + CC + SYN	AT3G10890	Glycosyl hydrolase superfamily protein
EC + CC + SYN	AT1G73610	GDSL-like Lipase/Acylhydrolase superfamily protein
EC + CC + SYN	AT2G23990	early nodulin-like protein 11 (ENODL11)
EC + CC + SYN	AT3G05460	sporozoite surface protein-related
EC + CC + SYN	AT1G12000	Phosphofructokinase family protein
EC + CC + SYN	AT4G16660	heat shock protein 70 (Hsp 70) family protein
EC + CC + SYN	AT4G09090	Carbohydrate-binding X8 domain superfamily protein
EC + CC + SYN	AT1G77510	PDI-like 1-2 (PDIL1-2)
EC + CC + SYN	AT4G18950	Integrin-linked protein kinase family
EC + CC + SYN	AT3G22370	alternative oxidase 1A (AOX1A)
EC + CC + SYN	AT1G60985	SCR-like 6 (SCRL6)
EC + CC + SYN	AT5G33340	CONSTITUTIVE DISEASE RESISTANCE 1 (CDR1)
EC + CC + SYN	AT4G11510	ralf-like 28 (RALFL28)
EC + CC + SYN	AT2G41451	unknown protein
EC + CC + SYN	AT3G17140	Plant invertase/pectin methylesterase inhibitor superfamily protein
EC + CC + SYN	AT3G24510	Defensin-like (DEFL) family protein
EC + CC + SYN	AT3G14850	TRICHOME BIREFRINGENCE-LIKE 41 (TBL41)
EC + CC + SYN	AT1G24400	lysine histidine transporter 2 (LHT2)
EC + CC + SYN	AT4G30590	early nodulin-like protein 12 (ENODL12)
EC + CC + SYN	AT1G71890	SUC5
EC + CC + SYN	AT1G04645	Plant self-incompatibility protein S1 family
EC + CC + SYN	AT5G09370	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
EC + CC + SYN	AT4G07390	Mannose-P-dolichol utilization defect 1 protein
EC + CC + SYN	AT3G29360	UDP-glucose 6-dehydrogenase family protein
EC + CC + SYN	AT4G32470	Cytochrome bd ubiquinol oxidase, 14kDa subunit
EC + CC + SYN	AT2G33150	peroxisomal 3-ketoacyl-CoA thiolase 3 (PKT3)
EC + CC + SYN	AT5G63510	gamma carbonic anhydrase like 1 (GAMMA CAL1)
EC + CC + SYN	AT1G44170	aldehyde dehydrogenase 3H1 (ALDH3H1)
EC + CC + SYN	AT5G61790	calnexin 1 (CNX1)
EC + CC + SYN	AT5G62460	RING/FYVE/PHD zinc finger superfamily protein
EC + CC + SYN	AT4G11600	glutathione peroxidase 6 (GPX6)
EC + CC + SYN	AT5G58710	rotamase CYP 7 (ROC7)
EC + CC + SYN	AT4G24190	SHEPHERD (SHD)
EC + CC + SYN	AT1G27530	Ubiquitin-conjugating enzyme/RWD-like
EC + CC + SYN	AT5G40730	arabinogalactan protein 24 (AGP24)
EC + CC + SYN	AT3G15020	Lactate/malate dehydrogenase family protein
EC + CC + SYN	AT5G27520	peroxisomal adenine nucleotide carrier 2 (PNC2)
EC + CC + SYN	AT1G29310	SecY protein transport family protein
EC + CC + SYN	AT3G51160	MURUS 1 (MUR1)
EC + CC + SYN	AT3G27210	unknown protein
EC + CC + SYN	AT1G74450	molecular_function unknown
EC + CC + SYN	AT4G07960	Cellulose-synthase-like C12 (CSLC12)
EC + CC + SYN	AT5G02490	Heat shock protein 70 (Hsp 70) family protein

Supplemental Table 5-2 (continued)

EC + CC + SYN	AT5G26250	Major facilitator superfamily protein
EC + CC + SYN	AT3G55680	Plant invertase/pectin methylesterase inhibitor superfamily protein
EC + CC + SYN	AT3G11210	SGNH hydrolase-type esterase superfamily protein
EC + CC + SYN	AT1G13680	PLC-like phosphodiesterases superfamily protein
EC + CC + SYN	AT2G20070	molecular_function unknown
EC + CC + SYN	AT4G10220	Protein of Unknown Function (DUF239)
EC + CC + SYN	AT2G41290	strictosidine synthase-like 2 (SSL2)
EC + CC + SYN	AT1G22015	DD46
EC + CC + SYN	AT2G38110	glycerol-3-phosphate acyltransferase 6 (GPAT6)
EC + CC + SYN	AT1G73560	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
EC + CC + SYN	AT1G68290	endonuclease 2 (ENDO 2)
EC + CC + SYN	AT2G23900	Pectin lyase-like superfamily protein
EC + CC + SYN	AT5G56480	Bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily protein
EC + CC + SYN	AT4G20050	QUARTET 3 (QRT3)
EC + CC + SYN	AT2G02515	unknown protein
EC + CC + SYN	AT4G17505	Protein of Unknown Function (DUF239)
EC + CC + SYN	AT3G25160	ER lumen protein retaining receptor family protein
EC + CC + SYN	AT4G32375	Pectin lyase-like superfamily protein
EC + CC + SYN	AT1G74010	Calcium-dependent phosphotriesterase superfamily protein
EC + CC + SYN	AT4G30070	low-molecular-weight cysteine-rich 59 (LCR59)
EC + CC + SYN	AT1G09370	Plant invertase/pectin methylesterase inhibitor superfamily protein
EC + CC + SYN	AT1G26795	Plant self-incompatibility protein S1 family
EC + CC + SYN	AT4G39490	cytochrome P450, family 96, subfamily A, polypeptide 10 (CYP96A10)
EC + CC + SYN	AT5G09730	beta-xylosidase 3 (BXL3)
EC + CC + SYN	AT4G35725	unknown protein
EC + CC + SYN	AT2G06090	Plant self-incompatibility protein S1 family
EC + CC + SYN	AT3G04540	Cysteine-rich protein
EC + CC + SYN	AT1G56620	Plant invertase/pectin methylesterase inhibitor superfamily protein
EC + CC + SYN	AT2G20595	molecular_function unknown
EC + CC + SYN	AT3G03960	TCP-1/cpn60 chaperonin family protein
EC + CC + SYN	AT1G23190	Phosphoglucomutase/phosphomannomutase family protein
EC + CC + SYN	AT4G01870	tolB protein-related
EC + CC + SYN	AT3G62600	ATERDJ3B
EC + CC + SYN	AT3G04120	glyceraldehyde-3-phosphate dehydrogenase C subunit 1 (GAPC1)
EC + CC + SYN	AT3G61110	ribosomal protein S27 (RS27A)
EC + CC + SYN	AT1G56340	calreticulin 1a (CRT1a)
EC + CC + SYN	AT4G29480	Mitochondrial ATP synthase subunit G protein
EC + CC + SYN	AT3G62870	Ribosomal protein L7Ae/L30e/S12e/Gadd45 family protein
EC + CC + SYN	AT5G23740	ribosomal protein S11-beta (RPS11-BETA)
EC + CC + SYN	AT5G20290	Ribosomal protein S8e family protein
EC + CC + SYN	AT2G42210	OEP16-3
EC + CC + SYN	AT1G64190	6-phosphogluconate dehydrogenase family protein
EC + CC + SYN	AT5G52840	NADH-ubiquinone oxidoreductase-related
EC + CC + SYN	AT1G09210	calreticulin 1b (CRT1b)
EC + CC + SYN	AT1G14320	SUPPRESSOR OF ACAULIS 52 (SAC52)
EC + CC + SYN	AT1G48830	Ribosomal protein S7e family protein
EC + CC + SYN	AT2G31490	unknown protein
EC + CC + SYN	AT1G79550	phosphoglycerate kinase (PGK)
EC + CC + SYN	AT4G26910	Dihydrolipoamide succinyltransferase
EC + CC + SYN	AT2G17440	plant intracellular ras group-related LRR 5 (PIRL5)

Supplemental Table 5-3 Determined amplification efficiencies for gene-specific amplification of target genes by qPCR.

By linear regression analysis of the obtained Cq values from a three-step dilution series of the cDNA template, and subsequent qPCR the gene-specific amplification efficiencies were calculated for the designated oligo pairs. E = efficiency.

AGI	oligo pair	E	AGI	oligo pair	E
AT4G21326	676/677	2.04	AT2G20595	544/545	2.04
AT1G52910	602/659	2.00	AT5G33340	678/679	2.11
AT5G55050	598/599	2.07	AT3G05460	652/653	1.92
AT1G74440	188/189	2.13	AT1G08050	086/087	2.04
AT5G01150	190/191	2.04	AT1G04645	425/426	2.01
AT1G31450	538/539	2.08	AT1G01570	660/661	1.90
AT1G18720	313/314	1.84	AT1G76750	720/721	2.11
AT4G11510	540/541	1.91	AT4G11720	680/681	2.00
AT2G20660	542/543	1.74	AT5G49150	682/683	2.04
AT1G47470	648/649	1.84	AT1G13440	526/527	1.89
AT4G17505	662/663	2.07	AT1G13320	646/647	2.07
AT5G24316	650/651	1.95	AT4G05320	528/529	1.91
AT5G18990	654/655	2.03	AT5G21040	588/589	1.75
AT4G09090	516/517	2.04	AT5G13050	586/587	1.77

Supplemental Table 5-4 Generated probes for WISH-based detection of candidate gene transcripts.

To perform WISH on 20 candidate genes, and *EC1.1* as positive control, the respective DIG-labeled probes in sense and antisense orientation were generated by amplification from cDNA or gDNA, subsequent cloning into pCR2.1 blunt vector, linearization and *in vitro* transcription by Sp6, and T7 DNA-dependent RNA polymerase, respectively. The amplicon size plus 98 nucleotides when transcribed with Sp6, or 111 nucleotides when transcribed with T7 DNA-dependent RNA polymerase equals the obtained probe size. Whether the probes produced consistent results is indicated separately. AT1G76750 CDS in the pCR2.1 blunt vector was contributed by Dr. Stefanie Sprunck.

(probe) AGI	oligo pair	amplicon size [bp]	consistent signal observed
AT1G04645	548/549	494	**yes**
AT1G18720	272/273	618	no
AT1G31450	538/539	380	**yes**
(i) AT1G52910	602/603	252	no
(ii) AT1G52910	658/659	528	**yes**
AT1G73200	600/601	306	**yes**
AT1G74440	116/117	624	no
AT1G76750	-	477	**yes**
AT2G06090	534/535	408	**yes**
AT2G20595	545/546	361	**yes**
AT2G20660	542/543	296	**yes**
AT2G23810	590/591	225	no
AT2G43660	466/467	463	no
AT3G05460	546/547	277	**yes**
AT3G09085	287/298	336	no
AT4G09090	462/463	491	no
AT4G11510	540/541	200	**yes**
AT5G11940	536/537	2289	**yes**
AT5G33340	604/605	288	**yes**
AT5G45910	532/533	1240	**yes**
(i) AT5G55050	598/599	224	no
(ii) AT5G55050	656/657	244	**yes**
AT5G67550	592/593	217	no

Table 5-5 Oligo nucleotides used in this work.

oligo	sequence
THO- 1	ATATTGACCATCATACTCATTGC
THO- 2	CTACCGGTTTGTTGCTTATCG
THO- 3	CTAAACCCTATTCCTCCCTCG
THO- 4	ATTTTGCCGATTTCGGAAC
THO- 5	TATCCACCACTCGCGTAAAAG
THO- 6	TGATATGCATCGAAGTTCAAAAC
THO- 10	AAAATCCAGAGAAAACGGCTC
THO- 11	ATTTTGATGGCGAGATCAATG
THO- 12	CTCCTGCTGCTTCAATCATTC
THO- 13	ACTTGTGAAGGCAACAAATGG
THO- 14	TAACCGATGCTGGTAAAATGC
THO- 15	ACCGTCTCCATCCGAAATTAG
THO- 16	TATTATTCAAGGAATTGCGCG
THO- 17	TGTGTCACTGTTACCAGCGAG
THO- 18	CCCGGTTCAGACTTAACCAAC
THO- 19	ATTCATACCTTTGGGTTTGGC
THO- 20	TCCACAATGACTCAAAAACCC
THO- 21	CAAACTTCCTCTCTGACGTCG
THO- 24	ATTACTTGGGCTTTCACTCTACC
THO- 43	CTTTACGATGCCATTGGGATATATCAAC
THO- 45	AATCACTTTTCCTAGCCGTCC
THO- 46	CGTTTCTGAACTCGAAGTTGC
THO- 61	CACCTTTATAGAAAACCCAAATTATAGAGCG
THO- 62	TGTTACACTTTCCGATTTCGATTC
THO- 64	CATGACCTTTCTCTTCTTCTTTGG
THO- 68	TGCCTGTCAACCTCTATGGAC
THO- 69	ACGAGCCTTTACCTCAAGGAG
THO- 70	ATGCGAAATTGATGCAATTTC
THO- 71	AAGGTATGAATTGGAATCCGG
THO- 72	AGGATTGATGCAGTGTGGAAG
THO- 73	AGCATCGTGACCGTTGTAAAC
THO- 75	GGGGACAAGTTTGTACAAAAAAGCAGGCTCTGGACAGAGAGAAGTTCGGAAG
THO- 76	GGGGACCACTTTGTACAAGAAAGCTGGGTTGTTGATGATCGAAAATCATGTTC
THO- 84	CTGCAAATTGCCTTTTCTTATCGA
THO- 86	TCTGCTCAATCCAGCGACAG
THO- 87	GGCTTTCGAGCACCAGTAGC
THO- 94	GATCGAAAAGTTCGACAGCGTCTCC
THO- 95	GGGATCAGCAATCGCGCATA
THO- 100	CTGCAAGGCGATTAAGTTGGGTAAC
THO- 101	GCGGATAACAATTTCACACAGGAAACAG
THO- 104	CACCCAGATTTAGAAAGCATGCAAGG
THO- 106	CACCAGAGATACAAAAATGGGTTACATCATAG
THO- 107	TGGTTTAAAGTTTTCTCTAAACACTTTAG
THO- 108	CACCAGAAACACAAAGTGGAGAGAATCC
THO- 109	TAGCGTAATAGATGAAAGAAAGACG
THO- 114	CACCATGGCTAAAAGTAGTGAGGAGCC
THO- 115	AGTCTCTTCCTTCACTTTCTTTGAAAC
THO- 116	CACCATGAGCAATCGCATC
THO- 117	AGTGACCTTCTTCATCTGCTG
THO- 118	GATTAGATTGTCTAAAGGCGCAGCTCTCTTTTGTATTCCA
THO- 119	AGCTGCGCCTTTAGACAATCTAATCAAAGAGAATCAATGA
THO- 120	AGCTACGCCTTTAGAGAATCTATTCACAGGTCGTGATATG
THO- 121	GAATAGATTCTCTAAAGGCGTAGCTACATATATATTCCTA
THO- 122	GATTTGGCATCACACTTTCTACAATG
THO- 123	GTTCCACCACTGAGCACAATG
THO- 124	TTGTAAAACGACGGCCAG
THO- 125	ACAGCTATGACCATGTAATACGAC
THO- 126	GCAATTGTAAGAAAGCACTTATGC
THO- 127	TTAACGAGTCTAGTTTGAATTTTGG
THO- 128	GCGGAATAACTCATGTGGATG
THO- 129	CTGACTCGCTTTTTCAGGTTG
THO- 130	TGGAGGATCGTCATCGTTATC
THO- 131	AGTGTTTTCCCCAACAGGTTC
THO- 132	TTTTTGCATTTTTGGGTTATCAC
THO- 133	GGTTTAGTTCCCTCCAAATCG

Table 5-5 (continued) Oligo nucleotides used in this work.

oligo	sequence
THO- 134	TATCAGGTTTTGCTCACCAGG
THO- 135	TAGTTGAAGCCGCAGAGAAAG
THO- 136	ATTCATCAGTTTGGATTGCG
THO- 137	TTGCTGACATACCGTCATTTG
THO- 138	GAAGGGCTTGAGGAACAGAAC
THO- 139	ACACAACCCGGAATGTAAGTG
THO- 140	TGCTGATTTCATTCAAGGAGG
THO- 141	TGATTGAGGAAACTCAATGCC
THO- 142	TAAACATACGCGTTACGTCCC
THO- 143	TCTCCATATGGGAAGGAACTG
THO- 144	AAGTCATTTTGTAATGGGGGC
THO- 145	TTGGTGATGAAGATTGTTCCC
THO- 160	GTGTGTTTGTTTCCATGGGAG
THO- 161	AGGTGAGGATGCAAATGAATG
THO- 164	CACCAGCCATTTGATCTTTCAAAGG
THO- 165	AGGCTTATATCCGTAGGTACGG
THO- 166	CACCGATTCATCATGTCAGAATTTTATTTG
THO- 167	CCAACTGCGTTTCCTGTTG
THO- 168	CAGTTTTTGCCTTTGTTGTTAC
THO- 169	ACTGTCATCCCTCTTGTTGTTC
THO- 170	TTTACCTCTTCGCTATGTTCCTC
THO- 171	TTGGACGGCTTACAACAACC
THO- 188	GTCAGTTCCTCGGACATGG
THO- 189	CGTACCCGAAAACAGACTGC
THO- 190	CAAAGTGTACTGTATTAGATGATCTGACC
THO- 191	ATGTCACCAAGGAAGCTCTCAAC
THO- 193	ACGCCCTAGCACCACAGATTG
THO- 214	TTAAGAGCTCAGAAACACAAAGTGGAGAGAATCC
THO- 215	TTAAACTAGTTAGCGTAATAGATGAAAGAAAGACG
THO- 219	GCTCAAAGCAGAGGATGTCAC
THO- 220	TGCGGAAAATTTATGAACGAG
THO- 221	GCTTTTGTTGACGTGCTCTTC
THO- 222	TATTTGTTGCACCTCGAGGTC
THO- 223	CCAAGAACATTAGGGAGGCTC
THO- 224	GCTCAAAGCAGAGGATGTCAC
THO- 225	CAAGGAACATCTTACTGCTCTTCAG
THO- 229	GATGCACTCCGCATACAGC
THO- 230	AAATGGCTAAAATGAGAATATCACC
THO- 231	CACGGTCAACTTCCGTACCGAGC
THO- 232	CATGCCAGTTCCCGTGCTTGAAGC
THO- 238	CACCATTTGTTTTTTGAACATCTACTCTGG
THO- 239	AGCCATATTAGGCCAGATAGACTAAAG
THO- 240	CACCTGAAGTGAACCTGCATTGATG
THO- 241	ACCATTGTCGATCAACCCTTTAAG
THO- 242	CACCAGTGTTCTTGTCTTTTGGGTTTG
THO- 259	GATGGCCATGTTATCCTCCTC
THO- 259	GATGGCCATGTTATCCTCCTC
THO- 264	CACCAAACTAAAATTTCTGAGGAATCTGG
THO- 265	ATCTTCTCTATCAAATTAAAACCCTA
THO- 266	TTAAGAGCTCAGAGATACAAAAATGGGTTACATCATAG
THO- 267	TTAAACTAGTTGGTTTAAAGTTTTCTCTAAACACTTTAG
THO- 268	TTAAGAGCTCTTGAGGCAGAAAATGTGTGG
THO- 269	GGATAAAAAATTCGTTGAGGAAG
THO- 270	CACCTTGAGGCAGAAAATGTGTGG
THO- 271	TTAAACTAGTGGATAAAAAATTCGTTGAGGAAG
THO- 272	CACCATGAGTCGCATAATGGGATTG
THO- 273	ATGTGATCTTATTCTTCTTCTGCTTC
THO- 280	AATTATCAAGGATTGGGGTGG
THO- 281	TTTTTAAAATCTTTGACTTCTTATATGG
THO- 282	GACCTAAAGACTCACGGTTCG
THO- 283	TGGAGCCTCATGGTTAAAGAG
THO- 285	TATTCATCTTCTCTATCAAATTAAAACC
THO- 290	GGGAAAATGAAGATCCTCTGG
THO- 291	CATGCATCTTGTGAACTGTGG
THO- 292	CCTCTGTGTTTTTGTTTTGGC

Table 5-5 (continued) Oligo nucleotides used in this work.

oligo	sequence
THO- 293	GGCATTCTTCAGCATTCTGAG
THO- 294	AAACACAGTGGTTTCTGGGTG
THO- 295	TGCACTAGACCCGTTAGATGC
THO- 296	ACCTAGCTCAGAACGTGCAAG
THO- 297	AGAGTATCATTCCCCAGGTGG
THO- 298	CACCATGAATTTCAGAAGCTTTGAGGAG
THO- 299	TGTGATCTTATTCTTCTTCTGCTTC
THO- 302	AGACACATACACACATTTACCAAAACTG
THO- 303	CACCACATCCTACTCAACTTCCGATGTG
THO- 313	ATTTACTCCAAGCTTTTCTCATGG
THO- 314	TCTGCTTCTTTGCCCTATATTCC
THO- 317	CGTATGTTGCATCACCTTCACC
THO- 318	AGAGAGCTGCAACAATGGC
THO- 352	GCTGACATTTGATGGTGTTTTGC
THO- 353	ACATCTGTTTGGATAAGAAATCTGGTG
THO- 354	CGTACCCGAAAACAGACTGC
THO- 355	CGTTTCTCAAACAAGCCATGTCC
THO- 364	TTTTTTTTTTTTTTTTTTTT(G,A,C)
THO- 365	GTATGGCCAAACGTCTTCG
THO- 400	ATCAGAGCTTGTTCTTCTATTTATAGCTG
THO- 401	CACCGGTTAATCGGACAAATTGAG
THO- 411	AAGCATGAATATACTGTTGGATGAG
THO- 412	TTCGGTCAAGCGATATGTTC
THO- 413	AAAAGTTCTAACATACTATTGGACGATC
THO- 414	AACACCAAAACTGTAGACATCACTC
THO- 415	ACAAACATATTGCTGGACGAAC
THO- 416	GCTGTACACATCGCTCTTCTCAC
THO- 417	AGGGACGTCAAAACTACAAATATATTG
THO- 418	CTATATACATCACTCTTCTCTGTCAACC
THO- 419	GCTCATGGCCAAAATTGC
THO- 420	GGTCAATCACATTTTGGTTTGTG
THO- 421	TTGCCGATTTCGGACTTTC
THO- 422	ACAATCCCAAAGCTATAAACATCAC
THO- 423	AACTGCCAACATATTGCTCGAC
THO- 424	GAAACTGTAGACATCACTCTTCTCACC
THO- 425	CCATACATTGCAAATCCAAGC
THO- 426	TTATTTCCCACTGACAGTCATCAC
THO- 452	TTACATGCGTCAGCTGAGTTG
THO- 453	GTGAGAGCGAAATCAAAGTGG
THO- 454	TTTTAAACTTTTTGATTTTCTACCATC
THO- 455	AACAATGCACCGTCTAACGAG
THO- 456	TGGAGAAAACATTTGATAAAGCC
THO- 457	TCGAAGCAAAAGAACCGTTAG
THO- 458	CACCGAGAGATGGTTAGGGTCTTAGTTATTG
THO- 459	GCTTTGCTCTCCACGACC
THO- 460	CACCGGACAGAGAGAAGTTCGGAAG
THO- 461	TATATTCACTTGTCGTATTTCTGCAG
THO- 462	AAAATTACTTGAATCTCAATTGAATTGTC
THO- 463	ACACACATTTATCAAAACTTGGATCC
THO- 466	CAAAGACATCAAACAAAAACAAATATG
THO- 467	TAAGTAAAAAGTTGAACCATTAATTATTCAAAG
THO- 492	CAGACCGATCCAACATGAACC
THO- 494	AGTGTCTGAAACTCCTTCACCAG
THO- 495	CCAAAAGAGGCTACGAGTTCC
THO- 496	CATAAGAGAATGTCTCGGTGCAG
THO- 497	CCAAGAGCAACCTGAAAAGC
THO- 498	AAGCATTCCGATGTCAATGAG
THO- 499	CGTTGGCAGGAATCAATTTGC
THO- 500	CACCATGGGCGAAGGTTTGCTGG
THO- 501	TTGTCTTTGCATCCGCAAATTTC
THO- 502	TTATTGTCTTTGCATCCGCAAAT
THO- 513	TTAAGTGACCTTCTTCATCTGCTG
THO- 514	TTATGTGATCTTATTCTTCTTCTG
THO- 515	TCAAGAGAGCTGCAACAATGG
THO- 516	AAATATCATCACTTTTGGCACTTC

Table 5-5 (continued) Oligo nucleotides used in this work.

oligo	sequence
THO- 517	CACAATCAACTCCTTCACTACATC
THO- 518	TTTCAAATGTCCTATCATTATCGTC
THO- 518	TTTCAAATGTCCTATCATTATCGTC
THO- 526	TGGGAAAGTGTTGCCATCC
THO- 527	CTTCATTTTGCCTTCAGATTCCTC
THO- 528	ACCCTTGAAGTGGAAAGCTCC
THO- 529	TTCCAGCGAAGATGAGACGC
THO- 532	CACCATGAGAATCAATATGTTATTCATAGTGG
THO- 533	ATTATAAGTAGGCATGGTGAAACG
THO- 534	CACCATGAATAATCTTTTTGTTTTGCTC
THO- 535	CCAACCATATATTTTCACGAAATC
THO- 536	CACCATGATGAGCTCTATAGTCTCATGG
THO- 537	CTGGTCGAAGTACATCAACATTTG
THO- 538	ATACACCGGAGGAGGATACG
THO- 539	AGTCATAGCCAACAAGGAAATCC
THO- 540	GTTCATAAGCAACAACATGAACG
THO- 541	TTATATCATCTTTCGAGGAAGAGC
THO- 542	GAAGCTCTTAATCTTCGCCG
THO- 543	TAAACCGGTAACAATGTGAATGG
THO- 544	TATCAAAAGATAAAAACAGAGTTTGC
THO- 545	CGTGAAGTTAAAAACAGTAGAGATGATC
THO- 546	TAATCCAGAGAATCCAAATTGC
THO- 547	TTTCTTACCGCAATTTGGATTCTC
THO- 548	CATTCATAAATACAAACATCAAAATGG
THO- 549	TATATCTTCATCATCGCAGGAGTAC
THO- 550	CACCTCAAGTAAAAATTGGTTTCTTCAGC
THO- 551	AATTGAATTCATGAGTCGCATAATGGGATTG
THO- 574	CCGGCCCTGCTGGATAACTTACTG
THO- 575	CAGTAAGTTATCCAGCAGGGCCGG
THO- 576	TTTGGAGCCATGTCGTTTATC
THO- 577	TATTTTTCCCCATTCTCGGTC
THO- 578	GGTTTCTTCAGCATGAGCAAG
THO- 581	GAGCTCCAGCACTTCTTGACAACCTACTC
THO- 582	GAGTAGGTTGTCAAGAAGTGCTGGAGCTC
THO- 583	GTCTTGTGGAGAGCAAAGTTTG
THO- 584	GCGTTTTATCACAAGAACTTGTC
THO- 585	GTTGCTCCACGTACCACAGTC
THO- 586	TGCAAAGGAGAAGGGTTGG
THO- 587	TGTTACAGGAATACTTCCATCCTCAAG
THO- 588	GATTTCTTCGAGTACTGTGGAACC
THO- 589	CAATGCTTGTGTGAAGTTCCATATC
THO- 590	GCTTTCAGGAACAACAAGAGG
THO- 591	ATCAAGGTTTAGCCCAAAAGG
THO- 592	CAAAACCTAACAGAATGCCTTGC
THO- 593	CTGCAGAAGAAGTTGGATGAATAG
THO- 598	GTTTTGCTGACGTTACATCTGC
THO- 599	TCATGTAGAGACCAACTGAGTAAGAG
THO- 600	CTAGTAGTTATGTCCATATGCACAAACTC
THO- 601	TCAAAAGCGCTCTGAGTCTTG
THO- 602	CCTCCTAGCTGTGAAGTGATCC
THO- 603	CAAGAAGCAAAATGACTAACATCATC
THO- 604	TACAAGTCTCGGAGGATTTGG
THO- 605	GAGAATTTAACTCTCCCTTGAAAAG
THO- 632	AATTTCTAGAAAATGAATCGAATCACAAAATCAG
THO- 633	AATTGGATCCATAAGGGCAAGGACCAGG
THO- 646	AACGTGGCCAAAATGATGC
THO- 647	CACATTGTCAATAGATTGGAGAGC
THO- 648	GAAAAGCCAAGCTCAAAATGC
THO- 649	CCGGAGCACCAATCTTGC
THO- 650	CTAAGACCTCCCAAGCCATCC
THO- 651	GAAAGGAGGTCGTCGTGTCG
THO- 652	ATGAAGAACGTTTCCTTTCAGC
THO- 653	TCTATTTGTATGCACGCCGC
THO- 654	AAGGCTTACTTGGGCAGAGC
THO- 655	GCTGCTTTTCGTAACCCTGG

Table 5-5 (continued) Oligo nucleotides used in this work.

oligo	sequence
THO- 656	TCACGATAGTGGAGCGCGTAG
THO- 657	GGCAGGGTTGGAAATAATGTC
THO- 658	ATGGCTTCAAAGCTCGTGATC
THO- 659	CTAGTAATGTGGGGTTTGGTAACC
THO- 660	CATACGTGGCGTGTTGTCAG
THO- 661	TCACATTCAGGCTCCCACATG
THO- 662	ACAATCCCATCAGTGAGCTGG
THO- 663	TTAATAAGGGCAAGGACCAGGC
THO- 676	AGCCGACCCATTTGACTTCG
THO- 677	GAAGGCTTTGGACTCGGACAT
THO- 678	TGTACAAGTCTCGGAGGATTTGG
THO- 679	ACCGTTTTGGAAACAGTGTCG
THO- 680	ATGCTACGACTGGTGGGAAG
THO- 681	TGGACTCGATGAAGTTGGTGG
THO- 682	GTTCAACTTCTCAGTTGGAGATCC
THO- 683	TTGGCTGCCATTGGAACAAG
THO- 684	ATTGGCGGCGTGCATACAAATAG
THO- 685	AAACTATTTGTATGCACGCCGCC
THO- 686	ATATATGGTCTCGATTGGAAAAGCGTTATCCGGTAGGTT
THO- 687	TGGAAAAGCGTTATCCGGTAGGTTTTAGAGCTAGAAATAGC
THO- 688	AACCCGGGATCTGGCTCGGCAACAATCTCTTAGTCGACTCTAC
THO- 689	ATTATTGGTCTCGAAACCCGGGATCTGGCTCGGCAACAA
THO- 690	ATATATGGTCTCGATTGGCATTGCGTCGTGGCTTTAGTT
THO- 691	TGGCATTGCGTCGTGGCTTTAGTTTTAGAGCTAGAAATAGC
THO- 696	AACTAAAGCCACGACGCAATGCCAATCTCTTAGTCGACTCTAC
THO- 697	ATTATTGGTCTCGAAACTAAAGCCACGACGCAATGCCAA
THO- 698	GTCCCACATCGCTTAGATAAGAAAACGAAG
THO- 699	AACAGAGGAAGAAGAAATCGATCTGGAA
THO- 719	ATCCTCATGCTCATGGTGGC
THO- 720	CGGAGAGAGGCAGTGTTGAG

6 – References

Acosta-Garcia, G. and J. P. Vielle-Calzada (2004). "A Classical Arabinogalactan Protein Is Essential for the Initiation of Female Gametogenesis in Arabidopsis." Plant Cell **16**(10): 2614-2628.

Aguilera-Romero, A., C. Gehin and H. Riezman (2014). "Sphingolipid Homeostasis in the Web of Metabolic Routes." Biochim Biophys Acta **1841**(5): 647-656.

Aken, B. L., S. Ayling, D. Barrell, L. Clarke, V. Curwen, S. Fairley, J. Fernandez Banet, K. Billis, C. Garcia Giron, T. Hourlier, K. Howe, A. Kahari, F. Kokocinski, F. J. Martin, D. N. Murphy, R. Nag, M. Ruffier, M. Schuster, Y. A. Tang, J. H. Vogel, S. White, A. Zadissa, P. Flicek and S. M. Searle (2016). "The Ensembl Gene Annotation System." Database (Oxford) **2016**.

Alonso, J. M., A. N. Stepanova, T. J. Leisse, C. J. Kim, H. Chen, P. Shinn, D. K. Stevenson, J. Zimmerman, P. Barajas, R. Cheuk, C. Gadrinab, C. Heller, A. Jeske, E. Koesema, C. C. Meyers, H. Parker, L. Prednis, Y. Ansari, N. Choy, H. Deen, M. Geralt, N. Hazari, E. Hom, M. Karnes, C. Mulholland, R. Ndubaku, I. Schmidt, P. Guzman, L. Aguilar-Henonin, M. Schmid, D. Weigel, D. E. Carter, T. Marchand, E. Risseeuw, D. Brogden, A. Zeko, W. L. Crosby, C. C. Berry and J. R. Ecker (2003). "Genome-Wide Insertional Mutagenesis of Arabidopsis Thaliana." Science **301**(5633): 653-657.

Angermueller, C., S. J. Clark, H. J. Lee, I. C. Macaulay, M. J. Teng, T. X. Hu, F. Krueger, S. A. Smallwood, C. P. Ponting, T. Voet, G. Kelsey, O. Stegle and W. Reik (2016). "Parallel Single-Cell Sequencing Links Transcriptional and Epigenetic Heterogeneity." Nat Methods **13**(3): 229-232.

Anguita, J., N. Ramamoorthi, J. W. Hovius, S. Das, V. Thomas, R. Persinski, D. Conze, P. W. Askenase, M. Rincon, F. S. Kantor and E. Fikrig (2002). "Salp15, an Ixodes Scapularis Salivary Protein, Inhibits Cd4(+) T Cell Activation." Immunity **16**(6): 849-859.

Arnold, K., L. Bordoli, J. Kopp and T. Schwede (2006). "The Swiss-Model Workspace: A Web-Based Environment for Protein Structure Homology Modelling." Bioinformatics **22**(2): 195-201.

Atkinson, N. J., C. J. Lilley and P. E. Urwin (2013). "Identification of Genes Involved in the Response of Arabidopsis to Simultaneous Biotic and Abiotic Stresses." Plant Physiology **162**(4): 2028-2041.

Bae, T. J., M. S. Kim, J. W. Kim, B. W. Kim, H. J. Choo, J. W. Lee, K. B. Kim, C. S. Lee, J. H. Kim, S. Y. Chang, C. Y. Kang, S. W. Lee and Y. G. Ko (2004). "Lipid Raft Proteome Reveals Atp Synthase Complex in the Cell Surface." Proteomics **4**(11): 3536-3548.

Bai, S. N. (2015). "The Concept of the Sexual Reproduction Cycle and Its Evolutionary Significance." Front Plant Sci **6**: 11.

Bailey, R. L., J. M. Herbert, K. Khan, V. L. Heath, R. Bicknell and M. G. Tomlinson (2011). "The Emerging Role of Tetraspanin Microdomains on Endothelial Cells." Biochem Soc Trans **39**(6): 1667-1673.

Balmain, A., R. Krumlauf, J. K. Vass and G. D. Birnie (1982). "Cloning and Characterisation of the Abundant Cytoplasmic 7s Rna from Mouse Cells." Nucleic Acids Res **10**(14): 4259-4277.

Bannur, S. V., S. V. Kulgod, S. S. Metkar, S. K. Mahajan and J. K. Sainis (1999). "Protein Determination by Ponceau S Using Digital Color Image Analysis of Protein Spots on Nitrocellulose Membranes." Anal Biochem **267**(2): 382-389.

Barcaccia, G. and E. Albertini (2013). "Apomixis in Plant Reproduction: A Novel Perspective on an Old Dilemma." Plant Reproduction **26**(3): 159-179.

Bari, R., Q. Guo, B. Xia, Y. H. Zhang, E. E. Giesert, S. Levy, J. J. Zheng and X. A. Zhang (2011). "Tetraspanins Regulate the Protrusive Activities of Cell Membrane." Biochem Biophys Res Commun **415**(4): 619-626.

Barral, P., C. Suarez, E. Batanero, C. Alfonso, D. Alche Jde, M. I. Rodriguez-Garcia, M. Villalba, G. Rivas and R. Rodriguez (2005). "An Olive Pollen Protein with Allergenic Activity, Ole E 10, Defines a Novel Family of Carbohydrate-Binding Modules and Is Potentially Implicated in Pollen Germination." Biochem J **390**(Pt 1): 77-84.

Barrell, P. J. and U. Grossniklaus (2005). "Confocal Microscopy of Whole Ovules for Analysis of Reproductive Development: The Elongate1 Mutant Affects Meiosis Ii." Plant J **43**(2): 309-320.

Barrett, T., S. E. Wilhite, P. Ledoux, C. Evangelista, I. F. Kim, M. Tomashevsky, K. A. Marshall, K. H. Phillippy, P. M. Sherman, M. Holko, A. Yefanov, H. Lee, N. Zhang, C. L. Robertson, N. Serova, S. Davis and A. Soboleva (2013). "Ncbi Geo: Archive for Functional Genomics Data Sets--Update." Nucleic Acids Res **41**(Database issue): D991-995.

Bateman, A., E. Birney, L. Cerruti, R. Durbin, L. Etwiller, S. R. Eddy, S. Griffiths-Jones, K. L. Howe, M. Marshall and E. L. Sonnhammer (2002). "The Pfam Protein Families Database." Nucleic Acids Res **30**(1): 276-280.

6 – References

Baxter, A., R. Mittler and N. Suzuki (2014). "Ros as Key Players in Plant Stress Signalling." Journal of Experimental Botany **65**(5): 1229-1240.
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.
Berardini, T. Z., L. Reiser, D. Li, Y. Mezheritsky, R. Muller, E. Strait and E. Huala (2015). "The Arabidopsis Information Resource: Making and Mining the "Gold Standard" Annotated Reference Plant Genome." Genesis **53**(8): 474-485.
Berditchevski, F., E. Odintsova, S. Sawada and E. Gilbert (2002). "Expression of the Palmitoylation-Deficient Cd151 Weakens the Association of Alpha 3 Beta 1 Integrin with the Tetraspanin-Enriched Microdomains and Affects Integrin-Dependent Signaling." J Biol Chem **277**(40): 36991-37000.
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.
Biasini, M., S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, F. Kiefer, T. Gallo Cassarino, M. Bertoni, L. Bordoli and T. Schwede (2014). "Swiss-Model: Modelling Protein Tertiary and Quaternary Structure Using Evolutionary Information." Nucleic Acids Res **42**(Web Server issue): W252-258.
Bicknell, R. A. and A. M. Koltunow (2004). "Understanding Apomixis: Recent Advances and Remaining Conundrums." Plant Cell **16 Suppl**: S228-245.
Bienert, S., A. Waterhouse, T. A. De Beer, G. Tauriello, G. Studer, L. Bordoli and T. Schwede (2017). "The Swiss-Model Repository-New Features and Functionality." Nucleic Acids Res **45**(D1): D313-D319.
Black, P. N. and C. C. Dirusso (2007). "Yeast Acyl-Coa Synthetases at the Crossroads of Fatty Acid Metabolism and Regulation." Biochim Biophys Acta **1771**(3): 286-298.
Blanvillain, R., L. C. Boavida, S. Mccormick and D. W. Ow (2008). "Exportin1 Genes Are Essential for Development and Function of the Gametophytes in Arabidopsis Thaliana." Genetics **180**(3): 1493-1500.
Bligh, E. G. and W. J. Dyer (1959). "A Rapid Method of Total Lipid Extraction and Purification." Can J Biochem Physiol **37**(8): 911-917.
Boavida, L. C., P. Qin, M. Broz, J. D. Becker and S. Mccormick (2013). "Arabidopsis Tetraspanins Are Confined to Discrete Expression Domains and Cell Types in Reproductive Tissues and Form Homo- and Heterodimers When Expressed in Yeast." Plant Physiol **163**(2): 696-712.
Boavida, L. C., B. Shuai, H. J. Yu, G. C. Pagnussat, V. Sundaresan and S. Mccormick (2009). "A Collection of Ds Insertional Mutants Associated with Defects in Male Gametophyte Development and Function in Arabidopsis Thaliana." Genetics **181**(4): 1369-1385.
Bordoli, L., F. Kiefer, K. Arnold, P. Benkert, J. Battey and T. Schwede (2009). "Protein Structure Homology Modeling Using Swiss-Model Workspace." Nat Protoc **4**(1): 1-13.
Borges, F., G. Gomes, R. Gardner, N. Moreno, S. Mccormick, J. A. Feijo and J. D. Becker (2008). "Comparative Transcriptomics of Arabidopsis Sperm Cells." Plant Physiol **148**(2): 1168-1181.
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.
Brauer, M. J., A. J. Saldanha, K. Dolinski and D. Botstein (2005). „Homeostatic Adjustment and Metabolic Remodeling in Glucose-Limited Yeast Cultures." Mol Biol Cell **16**(5): 2503-2517.
Breslow, D. K., D. M. Cameron, S. R. Collins, M. Schuldiner, J. Stewart-Ornstein, H. W. Newman, S. Braun, H. D. Madhani, N. J. Krogan and J. S. Weissman (2008). "A Comprehensive Strategy Enabling High-Resolution Functional Analysis of the Yeast Genome." Nat Methods **5**(8): 711-718.
Breslow, D. K. and J. S. Weissman (2010). "Membranes in Balance: Mechanisms of Sphingolipid Homeostasis." Mol Cell **40**(2): 267-279.
Brody, J. R. and S. E. Kern (2004). "Sodium Boric Acid: A Tris-Free, Cooler Conductive Medium for DNA Electrophoresis." Biotechniques **36**(2): 214-216.
Brown, R. C., B. E. Lemmon, H. Nguyen and O.-A. Olsen (1999). "Development of Endosperm in Arabidopsis Thaliana." Sexual Plant Reproduction **12**(1): 32-42.
Browse, J., P. J. Mccourt and C. R. Somerville (1986). "Fatty Acid Composition of Leaf Lipids Determined after Combined Digestion and Fatty Acid Methyl Ester Formation from Fresh Tissue." Anal Biochem **152**(1): 141-145.
Brys, G., M. Hubert and A. Struyf (2004). "A Robust Measure of Skewness." Journal of Computational and Graphical Statistics **13**(4): 996-1017.
Budnik, B., Levy, E., Slavov, N., (2017). "Mass-spectrometry of single mammalian cells

quantifiesproteome heterogenity during cell differentiation." bioRxiv, preprint
Burnette, W. N. (1981). ""Western Blotting": Electrophoretic Transfer of Proteins from Sodium Dodecyl Sulfate--Polyacrylamide Gels to Unmodified Nitrocellulose and Radiographic Detection with Antibody and Radioiodinated Protein A." Anal Biochem **112**(2): 195-203.
Campbell, L. and S. R. Turner (2017). "A Comprehensive Analysis of Ralf Proteins in Green Plants Suggests There Are Two Distinct Functional Groups." Front Plant Sci **8**: 37.
Candiano, G., M. Bruschi, L. Musante, L. Santucci, G. M. Ghiggeri, B. Carnemolla, P. Orecchia, L. Zardi and P. G. Righetti (2004). "Blue Silver: A Very Sensitive Colloidal Coomassie G-250 Staining for Proteome Analysis." Electrophoresis **25**(9): 1327-1333.
Cao, J. and F. Shi (2012). "Evolution of the Ralf Gene Family in Plants: Gene Duplication and Selection Patterns." Evol Bioinform Online **8**: 271-292.
Capron, A., M. Gourgues, L. S. Neiva, J. E. Faure, F. Berger, G. Pagnussat, A. Krishnan, C. Alvarez-Mejia, J. P. Vielle-Calzada, Y. R. Lee, B. Liu and V. Sundaresan (2008). "Maternal Control of Male-Gamete Delivery in Arabidopsis Involves a Putative Gpi-Anchored Protein Encoded by the Lorelei Gene." Plant Cell **20**(11): 3038-3049.
Chalbi, M., V. Barraud-Lange, B. Ravaux, K. Howan, N. Rodriguez, P. Soule, A. Ndzoudi, C. Boucheix, E. Rubinstein, J. P. Wolf, A. Ziyyat, E. Perez, F. Pincet and C. Gourier (2014). "Binding of Sperm Protein Izumo1 and Its Egg Receptor Juno Drives Cd9 Accumulation in the Intercellular Contact Area Prior to Fusion During Mammalian Fertilization." Development **141**(19): 3732-3739.
Charrin, S., S. Jouannet, C. Boucheix and E. Rubinstein (2014). "Tetraspanins at a Glance." J Cell Sci **127**(Pt 17): 3641-3648.
Charrin, S., M. Latil, S. Soave, A. Polesskaya, F. Chretien, C. Boucheix and E. Rubinstein (2013). "Normal Muscle Regeneration Requires Tight Control of Muscle Cell Fusion by Tetraspanins Cd9 and Cd81." Nat Commun **4**: 1674.
Charrin, S., F. Le Naour, O. Silvie, P. E. Milhiet, C. Boucheix and E. Rubinstein (2009). "Lateral Organization of Membrane Proteins: Tetraspanins Spin Their Web." Biochem J **420**(2): 133-154.
Chatr-Aryamontri, A., R. Oughtred, L. Boucher, J. Rust, C. Chang, N. K. Kolas, L. O'donnell, S. Oster, C. Theesfeld, A. Sellam, C. Stark, B. J. Breitkreutz, K. Dolinski and M. Tyers (2017). "The Biogrid Interaction Database: 2017 Update." Nucleic Acids Res **45**(D1): D369-D379.
Chaudhury, A. M., L. Ming, C. Miller, S. Craig, E. S. Dennis and W. J. Peacock (1997). "Fertilization-Independent Seed Development in Arabidopsis Thaliana." Proc Natl Acad Sci U S A **94**(8): 4223-4228.
Chen, W.-N., J. Zhang, Y. Lin, N. Chen, Z.-H. Zhan, H. S.-H. Chung, Y. Li and Y.-H. Shi (2013). "Particle Swarm Optimization with an Aging Leader and Challengers." IEEE Transactions on Evolutionary Computation **17**(2): 241-258.
Cherry, J. M., C. Ball, S. Weng, G. Juvik, R. Schmidt, C. Adler, B. Dunn, S. Dwight, L. Riles, R. K. Mortimer and D. Botstein (1997). "Genetic and Physical Maps of Saccharomyces Cerevisiae." Nature **387**(6632 Suppl): 67-73.
Cherry, J. M., E. L. Hong, C. Amundsen, R. Balakrishnan, G. Binkley, E. T. Chan, K. R. Christie, M. C. Costanzo, S. S. Dwight, S. R. Engel, D. G. Fisk, J. E. Hirschman, B. C. Hitz, K. Karra, C. J. Krieger, S. R. Miyasato, R. S. Nash, J. Park, M. S. Skrzypek, M. Simison, S. Weng and E. D. Wong (2012). "Saccharomyces Genome Database: The Genomics Resource of Budding Yeast." Nucleic Acids Res **40**(Database issue): D700-705.
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.
Christensen, C. A., S. W. Gorsich, R. H. Brown, L. G. Jones, J. Brown, J. M. Shaw and G. N. Drews (2002). "Mitochondrial Gfa2 Is Required for Synergid Cell Death in Arabidopsis." Plant Cell **14**(9): 2215-2232.
Christensen, C. A., E. J. King, J. R. Jordan and G. N. Drews (1997). "Megagametogenesis in Arabidopsis Wild Type and the Gf Mutant." Sexual Plant Reproduction **10**(1): 49-64.
Christensen, C. A., S. Subramanian and G. N. Drews (1998). "Identification of Gametophytic Mutations Affecting Female Gametophyte Development in Arabidopsis." Dev Biol **202**(1): 136-151.
Clough, S. J. and A. F. Bent (1998). "Floral Dip: A Simplified Method for Agrobacterium-Mediated Transformation of Arabidopsis Thaliana." Plant J **16**(6): 735-743.
Connolly, M. A., P. A. Clausen and J. G. Lazar (2006). "Preparation of Rna from Plant Tissue Using Trizol." CSH Protoc **2006**(1).
Coursol, S., H. Le Stunff, D. V. Lynch, S. Gilroy, S. M. Assmann and S. Spiegel (2005). "Arabidopsis

Sphingosine Kinase and the Effects of Phytosphingosine-1-Phosphate on Stomatal Aperture." Plant Physiol **137**(2): 724-737.
Cox, J., M. Y. Hein, C. A. Luber, I. Paron, N. Nagaraj and M. Mann (2014). "Accurate Proteome-Wide Label-Free Quantification by Delayed Normalization and Maximal Peptide Ratio Extraction, Termed Maxlfq." Mol Cell Proteomics **13**(9): 2513-2526.
Cox, J. and M. Mann (2008). "Maxquant Enables High Peptide Identification Rates, Individualized P.P.B.-Range Mass Accuracies and Proteome-Wide Protein Quantification." Nat Biotechnol **26**(12): 1367-1372.
Dagert, M. and S. D. Ehrlich (1979). "Prolonged Incubation in Calcium Chloride Improves the Competence of Escherichia Coli Cells." Gene **6**(1): 23-28.
Dai, N., W. Wang, S. E. Patterson and A. B. Bleecker (2013). "The Tmk Subfamily of Receptor-Like Kinases in Arabidopsis Display an Essential Role in Growth and a Reduced Sensitivity to Auxin." PLoS One **8**(4): e60990.
Darnell, D. K., S. Stanislaw, S. Kaur and P. B. Antin (2010). "Whole Mount in Situ Hybridization Detection of Mrnas Using Short Lna Containing DNA Oligonucleotide Probes." RNA **16**(3): 632-637.
Day, R. C., R. P. Herridge, B. A. Ambrose and R. C. Macknight (2008). "Transcriptome Analysis of Proliferating Arabidopsis Endosperm Reveals Biological Implications for the Control of Syncytial Division, Cytokinin Signaling, and Gene Expression Regulation." Plant Physiol **148**(4): 1964-1984.
Day, R. C., L. Mcnoe and R. C. Macknight (2007). "Evaluation of Global Rna Amplification and Its Use for High-Throughput Transcript Analysis of Laser-Microdissected Endosperm." Int J Plant Genomics: 61028.
De Storme, N. and D. Geelen (2014). "Callose Homeostasis at Plasmodesmata: Molecular Regulators and Developmental Relevance." Front Plant Sci **5**: 138.
Debro, J. R. and A. Korner (1956). "Solubility of Albumin in Alcohol after Precipitation by Trichloroacetic Acid: A Simplified Procedure for Separation of Albumin." Nature **178**(4541): 1067.
Dekkers, B. J., S. Pearce, R. P. Van Bolderen-Veldkamp, A. Marshall, P. Widera, J. Gilbert, H. G. Drost, G. W. Bassel, K. Muller, J. R. King, A. T. Wood, I. Grosse, M. Quint, N. Krasnogor, G. Leubner-Metzger, M. J. Holdsworth and L. Bentsink (2013). "Transcriptional Dynamics of Two Seed Compartments with Opposing Roles in Arabidopsis Seed Germination." Plant Physiol **163**(1): 205-215.
Demesa-Arevalo, E. and J. P. Vielle-Calzada (2013). "The Classical Arabinogalactan Protein Agp18 Mediates Megaspore Selection in Arabidopsis." Plant Cell **25**(4): 1274-1287.
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.
Detchokul, S., E. D. Williams, M. W. Parker and A. G. Frauman (2014). "Tetraspanins as Regulators of the Tumour Microenvironment: Implications for Metastasis and Therapeutic Strategies." Br J Pharmacol **171**(24): 5462-5490.
Do Canto, A. M., P. H. O. Ceciliato, B. Ribeiro, F. a. O. Morea, A. a. F. Garcia, M. C. Silva-Filho and D. S. Moura (2014). "Biological Activity of Nine Recombinant Atralf Peptides: Implications for Their Perception and Function in Arabidopsis." Plant Physiology and Biochemistry **75**: 45-54.
Dominguez, F. and F. J. Cejudo (2014). "Programmed Cell Death (Pcd): An Essential Process of Cereal Seed Development and Germination." Frontiers in Plant Science **5**.
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.
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.
Duan, Q. H., D. Kita, C. Li, A. Y. Cheung and H. M. Wu (2010). "Feronia Receptor-Like Kinase Regulates Rho Gtpase Signaling of Root Hair Development." Proceedings of the National Academy of Sciences of the United States of America **107**(41): 17821-17826.
Dubois, L., K. K. Ronquist, B. Ek, G. Ronquist and A. Larsson (2015). "Proteomic Profiling of Detergent Resistant Membranes (Lipid Rafts) of Prostasomes." Mol Cell Proteomics **14**(11): 3015-3022.

Dunoyer, P., C. Melnyk, A. Molnar and R. K. Slotkin (2013). "Plant Mobile Small Rnas." Cold Spring Harb Perspect Biol **5**(7).
Eastmond, P. J. and I. A. Graham (2001). "Re-Examining the Role of the Glyoxylate Cycle in Oilseeds." Trends Plant Sci **6**(2): 72-78.
Edgar, R., M. Domrachev and A. E. Lash (2002). "Gene Expression Omnibus: Ncbi Gene Expression and Hybridization Array Data Repository." Nucleic Acids Res **30**(1): 207-210.
Edgar, R. C. (2004). "Muscle: Multiple Sequence Alignment with High Accuracy and High Throughput." Nucleic Acids Res **32**(5): 1792-1797.
Edwards, K., C. Johnstone and C. Thompson (1991). "A Simple and Rapid Method for the Preparation of Plant Genomic DNA for Pcr Analysis." Nucleic Acids Res **19**(6): 1349.
Engel, S. R. and J. M. Cherry (2013). "The New Modern Era of Yeast Genomics: Community Sequencing and the Resulting Annotation of Multiple Saccharomyces Cerevisiae Strains at the Saccharomyces Genome Database." Database (Oxford) **2013**: bat012.
Englhart, M., Soljic, L., and Sprunck, S., (2017). "Manual Isolation of Living Cells from the Arabidopsis thaliana Female Gametophyte by Micromanipulation" *in press* Mol Methods
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.
Evans, E. J., L. Hene, L. M. Sparks, T. Dong, C. Retiere, J. A. Fennelly, R. Manso-Sancho, J. Powell, V. M. Braud, S. L. Rowland-Jones, A. J. Mcmichael and S. J. Davis (2003). "The T Cell Surface--How Well Do We Know It?" Immunity **19**(2): 213-223.
Evans, J. P. (2012). "Sperm-Egg Interaction." Annu Rev Physiol **74**: 477-502.
Eyuboglu, B., K. Pfister, G. Haberer, D. Chevalier, A. Fuchs, K. F. Mayer and K. Schneitz (2007). „Molecular Characterisation of the Strubbelig-Receptor Family of Genes Encoding Putative Leucine-Rich Repeat Receptor-Like Kinases in Arabidopsis Thaliana." BMC Plant Biol **7**: 16.
Fan, J., C. Yan, R. Roston, J. Shanklin and C. Xu (2014). "Arabidopsis Lipins, Pdat1 Acyltransferase, and Sdp1 Triacylglycerol Lipase Synergistically Direct Fatty Acids toward Beta-Oxidation, Thereby Maintaining Membrane Lipid Homeostasis." Plant Cell **26**(10): 4119-4134.
Fedry, J., Y. Liu, G. Pehau-Arnaudet, J. Pei, W. Li, M. A. Tortorici, F. Traincard, A. Meola, G. Bricogne, N. V. Grishin, W. J. Snell, F. A. Rey and T. Krey (2017). "The Ancient Gamete Fusogen Hap2 Is a Eukaryotic Class Ii Fusion Protein." Cell **168**(5): 904-915 e910.
Figueiredo, D. D., R. A. Batista, P. J. Roszak and C. Kohler (2015). "Auxin Production Couples Endosperm Development to Fertilization." Nat Plants **1**: 15184.
Figueiredo, D. D. and C. Kohler (2016). "Bridging the Generation Gap: Communication between Maternal Sporophyte, Female Gametophyte and Fertilization Products." Curr Opin Plant Biol **29**: 16-20.
Finn, R. D., P. Coggill, R. Y. Eberhardt, S. R. Eddy, J. Mistry, A. L. Mitchell, S. C. Potter, M. Punta, M. Qureshi, A. Sangrador-Vegas, G. A. Salazar, J. Tate and A. Bateman (2016). "The Pfam Protein Families Database: Towards a More Sustainable Future." Nucleic Acids Res **44**(D1): D279-285.
Fiume, E., M. Monfared, J. Jun and J. C. Fletcher (2011). "Cle Polypeptide Signaling Gene Expression in Arabidopsis Embryos." Plant Signal Behav **6**(3): 443-444.
Forde, B. G. and P. J. Lea (2007). "Glutamate in Plants: Metabolism, Regulation, and Signalling." J Exp Bot **58**(9): 2339-2358.
Frei, A. P., O. Y. Jeon, S. Kilcher, H. Moest, L. M. Henning, C. Jost, A. Pluckthun, J. Mercer, R. Aebersold, E. M. Carreira and B. Wollscheid (2012). "Direct Identification of Ligand-Receptor Interactions on Living Cells and Tissues." Nat Biotechnol **30**(10): 997-1001.
Frohman, M. A., M. K. Dush and G. R. Martin (1988). "Rapid Production of Full-Length Cdnas from Rare Transcripts: Amplification Using a Single Gene-Specific Oligonucleotide Primer." Proc Natl Acad Sci U S A **85**(23): 8998-9002.
Gall, J. G. and M. L. Pardue (1969). "Formation and Detection of Rna-DNA Hybrid Molecules in Cytological Preparations." Proc Natl Acad Sci U S A **63**(2): 378-383.
Gallegos, J. E. and A. B. Rose (2017). "Intron DNA Sequences Can Be More Important Than the Proximal Promoter in Determining the Site of Transcript Initiation." The Plant Cell: tpc.00020.02017.
Gao, J., J. J. Thelen, A. K. Dunker and D. Xu (2010). "Musite, a Tool for Global Prediction of General and Kinase-Specific Phosphorylation Sites." Mol Cell Proteomics **9**(12): 2586-2600.
Gasch, A. P., P. T. Spellman, C. M. Kao, O. Carmel-Harel, M. B. Eisen, G. Storz, D. Botstein and P. O. Brown (2000). "Genomic Expression Programs in the Response of Yeast Cells to Environmental Changes." Mol Biol Cell **11**(12): 4241-4257.

Gault, C. R., L. M. Obeid and Y. A. Hannun (2010). "An Overview of Sphingolipid Metabolism: From Synthesis to Breakdown." Adv Exp Med Biol **688**: 1-23.
Georgadaki, K., N. Khoury, D. A. Spandidos and V. Zoumpourlis (2016). "The Molecular Basis of Fertilization (Review)." Int J Mol Med **38**(4): 979-986.
Gibson, U. E., C. A. Heid and P. M. Williams (1996). "A Novel Method for Real Time Quantitative Rt-Pcr." Genome Res **6**(10): 995-1001.
Gilroy, S., N. Suzuki, G. Miller, W. G. Choi, M. Toyota, A. R. Devireddy and R. Mittler (2014). "A Tidal Wave of Signals: Calcium and Ros at the Forefront of Rapid Systemic Signaling." Trends Plant Sci **19**(10): 623-630.
Git, A., H. Dvinge, M. Salmon-Divon, M. Osborne, C. Kutter, J. Hadfield, P. Bertone and C. Caldas (2010). "Systematic Comparison of Microarray Profiling, Real-Time Pcr, and Next-Generation Sequencing Technologies for Measuring Differential Microrna Expression." RNA **16**(5): 991-1006.
Goda, H. (2004). "Comprehensive Comparison of Auxin-Regulated and Brassinosteroid-Regulated Genes in Arabidopsis." Plant Physiology **134**(4): 1555-1573.
Goodstein, D. M., S. Shu, R. Howson, R. Neupane, R. D. Hayes, J. Fazo, T. Mitros, W. Dirks, U. Hellsten, N. Putnam and D. S. Rokhsar (2012). "Phytozome: A Comparative Platform for Green Plant Genomics." Nucleic Acids Res **40**(Database issue): D1178-1186.
Goren, R., A. Altman and I. Giladi (1979). "Role of Ethylene in Abscisic Acid-Induced Callus Formation in Citrus Bud Cultures." Plant Physiol **63**(2): 280-282.
Graham, I. A. (2008). "Seed Storage Oil Mobilization." Annu Rev Plant Biol **59**: 115-142.
Grelon, M., J. Huang, T. Zhang, L. Linstroth, Z. Tillman, M. S. Otegui, H. A. Owen and D. Zhao (2016). "Control of Anther Cell Differentiation by the Small Protein Ligand Tpd1 and Its Receptor Ems1 in Arabidopsis." PLOS Genetics **12**(8): e1006147.
Gross-Hardt, R., C. Kagi, N. Baumann, J. M. Moore, R. Baskar, W. B. Gagliano, G. Jurgens and U. Grossniklaus (2007). "Lachesis Restricts Gametic Cell Fate in the Female Gametophyte of Arabidopsis." PLoS Biol **5**(3): e47.
Grossniklaus, U. and K. Schneitz (1998). "The Molecular and Genetic Basis of Ovule and Megagametophyte Development." Semin Cell Dev Biol **9**(2): 227-238.
Guan, D., B. Yan, C. Thieme, J. Hua, H. Zhu, K. R. Boheler, Z. Zhao, F. Kragler, Y. Xia and S. Zhang (2017). "Plamom: A Comprehensive Database Compiles Plant Mobile Macromolecules." Nucleic Acids Res **45**(D1): D1021-D1028.
Guex, N., M. C. Peitsch and T. Schwede (2009). „Automated Comparative Protein Structure Modeling with Swiss-Model and Swiss-Pdbviewer: A Historical Perspective." Electrophoresis **30 Suppl 1**: S162-173.
Hait, N. C., J. Allegood, M. Maceyka, G. M. Strub, K. B. Harikumar, S. K. Singh, C. Luo, R. Marmorstein, T. Kordula, S. Milstien and S. Spiegel (2009). "Regulation of Histone Acetylation in the Nucleus by Sphingosine-1-Phosphate." Science **325**(5945): 1254-1257.
Hamamura, Y., M. Nishimaki, H. Takeuchi, A. Geitmann, D. Kurihara and T. Higashiyama (2014). "Live Imaging of Calcium Spikes During Double Fertilization in Arabidopsis." Nat Commun **5**: 4722.
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.
Han, S., M. A. Lone, R. Schneiter and A. Chang (2010). "Orm1 and Orm2 Are Conserved Endoplasmic Reticulum Membrane Proteins Regulating Lipid Homeostasis and Protein Quality Control." Proc Natl Acad Sci U S A **107**(13): 5851-5856.
Hanada, K. (2003). "Serine Palmitoyltransferase, a Key Enzyme of Sphingolipid Metabolism." Biochim Biophys Acta **1632**(1-3): 16-30.
Hancock, J. F. (2006). "Lipid Rafts: Contentious Only from Simplistic Standpoints." Nat Rev Mol Cell Biol **7**(6): 456-462.
Harris, M. A., J. Clark, A. Ireland, J. Lomax, M. Ashburner, R. Foulger, K. Eilbeck, S. Lewis, B. Marshall, C. Mungall, J. Richter, G. M. Rubin, J. A. Blake, C. Bult, M. Dolan, H. Drabkin, J. T. Eppig, D. P. Hill, L. Ni, M. Ringwald, R. Balakrishnan, J. M. Cherry, K. R. Christie, M. C. Costanzo, S. S. Dwight, S. Engel, D. G. Fisk, J. E. Hirschman, E. L. Hong, R. S. Nash, A. Sethuraman, C. L. Theesfeld, D. Botstein, K. Dolinski, B. Feierbach, T. Berardini, S. Mundodi, S. Y. Rhee, R. Apweiler, D. Barrell, E. Camon, E. Dimmer, V. Lee, R. Chisholm, P. Gaudet, W. Kibbe, R. Kishore, E. M. Schwarz, P. Sternberg, M. Gwinn, L. Hannick, J. Wortman, M. Berriman, V. Wood, N. De La Cruz, P. Tonellato, P. Jaiswal, T. Seigfried, R. White and C. Gene

Ontology (2004). "The Gene Ontology (Go) Database and Informatics Resource." Nucleic Acids Res **32**(Database issue): D258-261.
Haruta, M., G. Sabat, K. Stecker, B. B. Minkoff and M. R. Sussman (2014). "A Peptide Hormone and Its Receptor Protein Kinase Regulate Plant Cell Expansion." Science **343**(6169): 408-411.
He, B., Y. H. Zhang, M. M. Richardson, J. S. Zhang, E. Rubinstein and X. A. Zhang (2011). "Differential Functions of Phospholipid Binding and Palmitoylation of Tumour Suppressor Ewi2/Pgrl." Biochem J **437**(3): 399-411.
He, F., S. Yoo, D. Wang, S. Kumari, M. Gerstein, D. Ware and S. Maslov (2016). "Large-Scale Atlas of Microarray Data Reveals the Distinct Expression Landscape of Different Tissues in Arabidopsis." Plant J **86**(6): 472-480.
Hejatko, J., I. Blilou, P. B. Brewer, J. Friml, B. Scheres and E. Benkova (2006). "In Situ Hybridization Technique for Mrna Detection in Whole Mount Arabidopsis Samples." Nat Protoc **1**(4): 1939-1946.
Hellemans, J., G. Mortier, A. De Paepe, F. Speleman and J. Vandesompele (2007). "Qbase Relative Quantification Framework and Software for Management and Automated Analysis of Real-Time Quantitative Pcr Data." Genome Biol **8**(2): R19.
Hemler, M. E. (2003). "Tetraspanin Proteins Mediate Cellular Penetration, Invasion, and Fusion Events and Define a Novel Type of Membrane Microdomain." Annu Rev Cell Dev Biol **19**: 397-422.
Higuchi, R., B. Krummel and R. K. Saiki (1988). "A General Method of in Vitro Preparation and Specific Mutagenesis of DNA Fragments: Study of Protein and DNA Interactions." Nucleic Acids Res **16**(15): 7351-7367.
Hoffman, C. S. (2001). "Preparation of Yeast DNA." Curr Protoc Mol Biol **Chapter 13**: Unit13 11.
Hojsgaard, D. and E. Horandl (2015). "A Little Bit of Sex Matters for Genome Evolution in Asexual Plants." Front Plant Sci **6**: 82.
Hooper, C. M., I. R. Castleden, S. K. Tanz, N. Aryamanesh and A. H. Millar (2017). "Suba4: The Interactive Data Analysis Centre for Arabidopsis Subcellular Protein Locations." Nucleic Acids Res **45**(D1): D1064-D1074.
Hooper, C. M., S. K. Tanz, I. R. Castleden, M. A. Vacher, I. D. Small and A. H. Millar (2014). "Subacon: A Consensus Algorithm for Unifying the Subcellular Localization Data of the Arabidopsis Proteome." Bioinformatics **30**(23): 3356-3364.
Hou, Y., X. Guo, P. Cyprys, Y. Zhang, A. Bleckmann, L. Cai, Q. Huang, Y. Luo, H. Gu, T. Dresselhaus, J. Dong and L. J. Qu (2016). "Maternal Enodls Are Required for Pollen Tube Reception in Arabidopsis." Curr Biol **26**(17): 2343-2350.
Houten, S. M. and R. J. A. Wanders (2010). "A General Introduction to the Biochemistry of Mitochondrial Fatty Acid B-Oxidation." Journal of Inherited Metabolic Disease **33**(5): 469-477.
Hruz, T., O. Laule, G. Szabo, F. Wessendorp, S. Bleuler, L. Oertle, P. Widmayer, W. Gruissem and P. Zimmermann (2008). "Genevestigator V3: A Reference Expression Database for the Meta-Analysis of Transcriptomes." Adv Bioinformatics **2008**: 420747.
Hruz, T., M. Wyss, M. Docquier, M. W. Pfaffl, S. Masanetz, L. Borghi, P. Verbrugghe, L. Kalaydjieva, S. Bleuler, O. Laule, P. Descombes, W. Gruissem and P. Zimmermann (2011). "Refgenes: Identification of Reliable and Condition Specific Reference Genes for Rt-Qpcr Data Normalization." BMC Genomics **12**: 156.
Huala, E., A. W. Dickerman, M. Garcia-Hernandez, D. Weems, L. Reiser, F. Lafond, D. Hanley, D. Kiphart, M. Zhuang, W. Huang, L. A. Mueller, D. Bhattacharyya, D. Bhaya, B. W. Sobral, W. Beavis, D. W. Meinke, C. D. Town, C. Somerville and S. Y. Rhee (2001). "The Arabidopsis Information Resource (Tair): A Comprehensive Database and Web-Based Information Retrieval, Analysis, and Visualization System for a Model Plant." Nucleic Acids Res **29**(1): 102-105.
Huang Da, W., B. T. Sherman, R. Stephens, M. W. Baseler, H. C. Lane and R. A. Lempicki (2008). "David Gene Id Conversion Tool." Bioinformation **2**(10): 428-430.
Huang, J., Y. Ju, X. Wang, Q. Zhang and Sodmergen (2015). "A One-Step Rectification of Sperm Cell Targeting Ensures the Success of Double Fertilization." J Integr Plant Biol **57**(5): 496-503.
Huck, N., J. M. Moore, M. Federer and U. Grossniklaus (2003). "The Arabidopsis Mutant Feronia Disrupts the Female Gametophytic Control of Pollen Tube Reception." Development **130**(10): 2149-2159.
Huh, W. K., J. V. Falvo, L. C. Gerke, A. S. Carroll, R. W. Howson, J. S. Weissman and E. K. O'shea (2003). "Global Analysis of Protein Localization in Budding Yeast." Nature **425**(6959): 686-691.
Hurkman, W. J. and C. K. Tanaka (1986). "Solubilization of Plant Membrane Proteins for Analysis by Two-Dimensional Gel Electrophoresis." Plant Physiol **81**(3): 802-806.
Ibarra, C. A., X. Feng, V. K. Schoft, T. F. Hsieh, R. Uzawa, J. A. Rodrigues, A. Zemach, N. Chumak,

A. Machlicova, T. Nishimura, D. Rojas, R. L. Fischer, H. Tamaru and D. Zilberman (2012). "Active DNA Demethylation in Plant Companion Cells Reinforces Transposon Methylation in Gametes." Science **337**(6100): 1360-1364.

Imamura, T., H. Kusano, Y. Kajigaya, M. Ichikawa and H. Shimada (2007). "A Rice Dihydrosphingosine C4 Hydroxylase (Dsh1) Gene, Which Is Abundantly Expressed in the Stigmas, Vascular Cells and Apical Meristem, May Be Involved in Fertility." Plant Cell Physiol **48**(8): 1108-1120.

Ingouff, M., Y. Hamamura, M. Gourgues, T. Higashiyama and F. Berger (2007). "Distinct Dynamics of Histone3 Variants between the Two Fertilization Products in Plants." Curr Biol **17**(12): 1032-1037.

Ingouff, M., T. Sakata, J. Li, S. Sprunck, T. Dresselhaus and F. Berger (2009). „The Two Male Gametes Share Equal Ability to Fertilize the Egg Cell in Arabidopsis Thaliana." Curr Biol **19**(1): R19-20.

Inoue, N., Y. Hagihara, D. Wright, T. Suzuki and I. Wada (2015). "Oocyte-Triggered Dimerization of Sperm Izumo1 Promotes Sperm-Egg Fusion in Mice." Nat Commun **6**: 8858.

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.

Ito, H., Y. Fukuda, K. Murata and A. Kimura (1983). "Transformation of Intact Yeast Cells Treated with Alkali Cations." J Bacteriol **153**(1): 163-168.

Jacquemyn, H., L. De Meester, E. Jongejans and O. Honnay (2012). "Evolutionary Changes in Plant Reproductive Traits Following Habitat Fragmentation and Their Consequences for Population Fitness." Journal of Ecology **100**(1): 76-87.

Jamshad, M., Y. P. Lin, T. J. Knowles, R. A. Parslow, C. Harris, M. Wheatley, D. R. Poyner, R. M. Bill, O. R. Thomas, M. Overduin and T. R. Dafforn (2011). "Surfactant-Free Purification of Membrane Proteins with Intact Native Membrane Environment." Biochem Soc Trans **39**(3): 813-818.

Jefferson, R. A., T. A. Kavanagh and M. W. Bevan (1987). "Gus Fusions: Beta-Glucuronidase as a Sensitive and Versatile Gene Fusion Marker in Higher Plants." EMBO J **6**(13): 3901-3907.

Jegou, A., A. Ziyyat, V. Barraud-Lange, E. Perez, J. P. Wolf, F. Pincet and C. Gourier (2011). "Cd9 Tetraspanin Generates Fusion Competent Sites on the Egg Membrane for Mammalian Fertilization." Proc Natl Acad Sci U S A **108**(27): 10946-10951.

Jia, G., X. Liu, H. A. Owen and D. Zhao (2008). "Signaling of Cell Fate Determination by the Tpd1 Small Protein and Ems1 Receptor Kinase." Proceedings of the National Academy of Sciences **105**(6): 2220-2225.

Jiao, J., A. G. Mizukami, S. Sankaranarayanan, J. Yamguchi, K. Itami and T. Higashiyawma (2017). "Structure-Activity Relation of Amor Sugar Molecule That Activates Pollen-Tubes for Ovular Guidance." Plant Physiol **173**(1): 354-363.

Johnson, M. A., K. Von Besser, Q. Zhou, E. Smith, G. Aux, D. Patton, J. Z. Levin and D. Preuss (2004). "Arabidopsis Hapless Mutations Define Essential Gametophytic Functions." Genetics **168**(2): 971-982.

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.

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.

Jones, A. M., Y. Xuan, M. Xu, R. S. Wang, C. H. Ho, S. Lalonde, C. H. You, M. I. Sardi, S. A. Parsa, E. Smith-Valle, T. Su, K. A. Frazer, G. Pilot, R. Pratelli, G. Grossmann, B. R. Acharya, H. C. Hu, C. Engineer, F. Villiers, C. Ju, K. Takeda, Z. Su, Q. Dong, S. M. Assmann, J. Chen, J. M. Kwak, J. I. Schroeder, R. Albert, S. Y. Rhee and W. B. Frommer (2014). "Border Control--a Membrane-Linked Interactome of Arabidopsis." Science **344**(6185): 711-716.

Juncadella, I. J., R. Garg, T. C. Bates, E. R. Olivera and J. Anguita (2008). "The Ixodes Scapularis Salivary Protein, Salp15, Prevents the Association of Hiv-1 Gp120 and Cd4." Biochem Biophys Res Commun **367**(1): 41-46.

Kalipatnapu, S. and A. Chattopadhyay (2005). "Membrane Protein Solubilization: Recent Advances and Challenges in Solubilization of Serotonin1a Receptors." IUBMB Life **57**(7): 505-512.

Kanehisa, M., M. Furumichi, M. Tanabe, Y. Sato and K. Morishima (2017). "Kegg: New Perspectives on Genomes, Pathways, Diseases and Drugs." Nucleic Acids Res **45**(D1):

D353-D361.
Kanehisa, M. and S. Goto (2000). "Kegg: Kyoto Encyclopedia of Genes and Genomes." Nucleic Acids Res **28**(1): 27-30.
Kanehisa, M., Y. Sato, M. Kawashima, M. Furumichi and M. Tanabe (2016). "Kegg as a Reference Resource for Gene and Protein Annotation." Nucleic Acids Res **44**(D1): D457-462.
Karimi, M., D. Inze and A. Depicker (2002). "Gateway Vectors for Agrobacterium-Mediated Plant Transformation." Trends Plant Sci **7**(5): 193-195.
Karve, R., W. Liu, S. G. Willet, K. U. Torii and E. D. Shpak (2011). "The Presence of Multiple Introns Is Essential for Erecta Expression in Arabidopsis." RNA **17**(10): 1907-1921.
Kiefer, F., K. Arnold, M. Kunzli, L. Bordoli and T. Schwede (2009). „The Swiss-Model Repository and Associated Resources." Nucleic Acids Res **37**(Database issue): D387-392.
Kirioukhova, O., A. J. Johnston, D. Kleen, C. Kagi, R. Baskar, J. M. Moore, H. Baumlein, R. Gross-Hardt and U. Grossniklaus (2011). "Female Gametophytic Cell Specification and Seed Development Require the Function of the Putative Arabidopsis Incenp Ortholog Wyrd." Development **138**(16): 3409-3420.
Kishimoto, K., R. Urade, T. Ogawa and T. Moriyama (2001). "Nondestructive Quantification of Neutral Lipids by Thin-Layer Chromatography and Laser-Fluorescent Scanning: Suitable Methods for "Lipidome" Analysis." Biochem Biophys Res Commun **281**(3): 657-662.
Kolesnikov, N., E. Hastings, M. Keays, O. Melnichuk, Y. A. Tang, E. Williams, M. Dylag, N. Kurbatova, M. Brandizi, T. Burdett, K. Megy, E. Pilicheva, G. Rustici, A. Tikhonov, H. Parkinson, R. Petryszak, U. Sarkans and A. Brazma (2015). "Arrayexpress Update--Simplifying Data Submissions." Nucleic Acids Res **43**(Database issue): D1113-1116.
Komarova, N. Y., S. Meier, A. Meier, M. S. Grotemeyer and D. Rentsch (2012). „Determinants for Arabidopsis Peptide Transporter Targeting to the Tonoplast or Plasma Membrane." Traffic **13**(8): 1090-1105.
Kondo, N., Y. Ohno, M. Yamagata, T. Obara, N. Seki, T. Kitamura, T. Naganuma and A. Kihara (2014). "Identification of the Phytosphingosine Metabolic Pathway Leading to Odd-Numbered Fatty Acids." Nat Commun **5**: 5338.
Kong, J., S. Lau and G. Jurgens (2015). "Twin Plants from Supernumerary Egg Cells in Arabidopsis." Curr Biol **25**(2): 225-230.
Kopp, J. and T. Schwede (2006). "The Swiss-Model Repository: New Features and Functionalities." Nucleic Acids Res **34**(Database issue): D315-318.
Koressaar, T. and M. Remm (2007). "Enhancements and Modifications of Primer Design Program Primer3." Bioinformatics **23**(10): 1289-1291.
Koszegi, D., A. J. Johnston, T. Rutten, A. Czihal, L. Altschmied, J. Kumlehn, S. E. Wust, O. Kirioukhova, J. Gheyselinck, U. Grossniklaus and H. Baumlein (2011). "Members of the Rkd Transcription Factor Family Induce an Egg Cell-Like Gene Expression Program." Plant J **67**(2): 280-291.
Koziel, M. G., T. L. Adams, M. A. Hazlet, D. Damm, J. Miller, D. Dahlbeck, S. Jayne and B. J. Staskawicz (1984). "A Cauliflower Mosaic Virus Promoter Directs Expression of Kanamycin Resistance in Morphogenic Transformed Plant Cells." J Mol Appl Genet **2**(6): 549-562.
Krebs, H. A. (1940). "The Citric Acid Cycle and the Szent-Gyorgyi Cycle in Pigeon Breast Muscle." Biochem J **34**(5): 775-779.
Kresnowati, M. T., W. A. Van Winden, M. J. Almering, A. Ten Pierick, C. Ras, T. A. Knijnenburg, P. Daran-Lapujade, J. T. Pronk, J. J. Heijnen and J. M. Daran (2006). „When Transcriptome Meets Metabolome: Fast Cellular Responses of Yeast to Sudden Relief of Glucose Limitation." Mol Syst Biol **2**: 49.
Krohn, N. G., A. Lausser, M. Juranic and T. Dresselhaus (2012). "Egg Cell Signaling by the Secreted Peptide Zmeal1 Controls Antipodal Cell Fate." Dev Cell **23**(1): 219-225.
Kudo, T., Y. Sasaki, S. Terashima, N. Matsuda-Imai, T. Takano, M. Saito, M. Kanno, S. Ozaki, K. Suwabe, G. Suzuki, M. Watanabe, M. Matsuoka, S. Takayama and K. Yano (2016). "Identification of Reference Genes for Quantitative Expression Analysis Using Large-Scale Rna-Seq Data of Arabidopsis Thaliana and Model Crop Plants." Genes Genet Syst **91**(2): 111-125.
Laemmli, U. K. (1970). "Cleavage of Structural Proteins During the Assembly of the Head of Bacteriophage T4." Nature **227**(5259): 680-685.
Lafon-Placette, C. and C. Kohler (2014). "Embryo and Endosperm, Partners in Seed Development." Curr Opin Plant Biol **17**: 64-69.
Lamesch, P., T. Z. Berardini, D. Li, D. Swarbreck, C. Wilks, R. Sasidharan, R. Muller, K. Dreher, D. L. Alexander, M. Garcia-Hernandez, A. S. Karthikeyan, C. H. Lee, W. D. Nelson, L. Ploetz, S. Singh, A. Wensel and E. Huala (2012). "The Arabidopsis Information Resource (Tair): Improved

Gene Annotation and New Tools." Nucleic Acids Res **40**(Database issue): D1202-1210.
Lamesch, P., K. Dreher, D. Swarbreck, R. Sasidharan, L. Reiser and E. Huala (2010). "Using the Arabidopsis Information Resource (Tair) to Find Information About Arabidopsis Genes." Curr Protoc Bioinformatics **Chapter 1**: Unit1 11.
Lampropoulos, A., Z. Sutikovic, C. Wenzl, I. Maegele, J. U. Lohmann and J. Forner (2013). "Greengate---a Novel, Versatile, and Efficient Cloning System for Plant Transgenesis." PLoS One **8**(12): e83043.
Le Naour, F., M. Andre, C. Boucheix and E. Rubinstein (2006). "Membrane Microdomains and Proteomics: Lessons from Tetraspanin Microdomains and Comparison with Lipid Rafts." Proteomics **6**(24): 6447-6454.
Lee, A. G. (2004). "How Lipids Affect the Activities of Integral Membrane Proteins." Biochim Biophys Acta **1666**(1-2): 62-87.
Lester, R. L., B. R. Withers, M. A. Schultz and R. C. Dickson (2013). "Iron, Glucose and Intrinsic Factors Alter Sphingolipid Composition as Yeast Cells Enter Stationary Phase." Biochim Biophys Acta **1831**(4): 726-736.
Levsky, J. M. and R. H. Singer (2003). "Fluorescence in Situ Hybridization: Past, Present and Future." J Cell Sci **116**(Pt 14): 2833-2838.
Lewis, D. R., A. L. Olex, S. R. Lundy, W. H. Turkett, J. S. Fetrow and G. K. Muday (2013). "A Kinetic Analysis of the Auxin Transcriptome Reveals Cell Wall Remodeling Proteins That Modulate Lateral Root Development in Arabidopsis." Plant Cell **25**(9): 3329-3346.
Li, C., F. L. Yeh, A. Y. Cheung, Q. Duan, D. Kita, M. C. Liu, J. Maman, E. J. Luu, B. W. Wu, L. Gates, M. Jalal, A. Kwong, H. Carpenter and H. M. Wu (2015). "Glycosylphosphatidylinositol-Anchored Proteins as Chaperones and Co-Receptors for Feronia Receptor Kinase Signaling in Arabidopsis." Elife **4**.
Li, H. J., S. S. Zhu, M. X. Zhang, T. Wang, L. Liang, Y. Xue, D. Q. Shi, J. Liu and W. C. Yang (2015). "Arabidopsis Cbp1 Is a Novel Regulator of Transcription Initiation in Central Cell-Mediated Pollen Tube Guidance." Plant Cell **27**(10): 2880-2893.
Li, J., Y. Ning, W. Hedley, B. Saunders, Y. Chen, N. Tindill, T. Hannay and S. Subramaniam (2002). "The Molecule Pages Database." Nature **420**(6916): 716-717.
Li, J. F., H. S. Chung, Y. Niu, J. Bush, M. Mccormack and J. Sheen (2013). "Comprehensive Protein-Based Artificial Microrna Screens for Effective Gene Silencing in Plants." Plant Cell **25**(5): 1507-1522.
Li, W., Y. Zhao, C. Liu, G. Yao, S. Wu, C. Hou, M. Zhang and D. Wang (2012). "Callose Deposition at Plasmodesmata Is a Critical Factor in Restricting the Cell-to-Cell Movement of Soybean Mosaic Virus." Plant Cell Rep **31**(5): 905-916.
Li, Y., F. Beisson, M. Pollard and J. Ohlrogge (2006). "Oil Content of Arabidopsis Seeds: The Influence of Seed Anatomy, Light and Plant-to-Plant Variation." Phytochemistry **67**(9): 904-915.
Lin, W. D., Y. Y. Liao, T. J. Yang, C. Y. Pan, T. J. Buckhout and W. Schmidt (2011). "Coexpression-Based Clustering of Arabidopsis Root Genes Predicts Functional Modules in Early Phosphate Deficiency Signaling." Plant Physiol **155**(3): 1383-1402.
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.
Long, J. A., C. Ohno, Z. R. Smith and E. M. Meyerowitz (2006). "Topless Regulates Apical Embryonic Fate in Arabidopsis." Science **312**(5779): 1520-1523.
Lozano, M. M., J. S. Hovis, F. R. Moss and S. G. Boxer (2016). "Dynamic Reorganization and Correlation among Lipid Raft Components." Journal of the American Chemical Society **138**(31): 9996-10001.
Maddon, P. J., D. R. Littman, M. Godfrey, D. E. Maddon, L. Chess and R. Axel (1985). "The Isolation and Nucleotide Sequence of a Cdna Encoding the T Cell Surface Protein T4: A New Member of the Immunoglobulin Gene Family." Cell **42**(1): 93-104.
Mao, D. D., F. Yu, J. Li, B. Van De Poel, D. Tan, J. L. Li, Y. Q. Liu, X. S. Li, M. Q. Dong, L. B. Chen, D. P. Li and S. Luan (2015). "Feronia Receptor Kinase Interacts with S-Adenosylmethionine Synthetase and Suppresses S-Adenosylmethionine Production and Ethylene Biosynthesis in Arabidopsis." Plant Cell and Environment **38**(12): 2566-2574.
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.
Maruyama, D., R. Volz, H. Takeuchi, T. Mori, T. Igawa, D. Kurihara, T. Kawashima, M. Ueda, M. Ito, M. Umeda, S. Nishikawa, R. Gross-Hardt and T. Higashiyama (2015). "Rapid Elimination of the Persistent Synergid through a Cell Fusion Mechanism." Cell **161**(4): 907-918.
Mccall, M. N., B. M. Bolstad and R. A. Irizarry (2010). "Frozen Robust Multiarray Analysis (Frma)." Biostatistics **11**(2): 242-253.
Mccall, M. N. and R. A. Irizarry (2011). "Thawing Frozen Robust Multi-Array Analysis (Frma)." BMC Bioinformatics **12**: 369.
Mccarthy, A. A., H. M. Baker, S. C. Shewry, T. F. Kagawa, E. Saafi, M. L. Patchett and E. N. Baker (2001). "Expression, Crystallization and Preliminary Characterization of Methylmalonyl Coenzyme a Epimerase from Propionibacterium Shermanii." Acta Crystallogr D Biol Crystallogr **57**(Pt 5): 706-708.
Mccormick, S. (1993). "Male Gametophyte Development." Plant Cell **5**(10): 1265-1275.
Mi, H., X. Huang, A. Muruganujan, H. Tang, C. Mills, D. Kang and P. D. Thomas (2017). "Panther Version 11: Expanded Annotation Data from Gene Ontology and Reactome Pathways, and Data Analysis Tool Enhancements." Nucleic Acids Res **45**(D1): D183-D189.
Mi, H., A. Muruganujan, J. T. Casagrande and P. D. Thomas (2013). "Large-Scale Gene Function Analysis with the Panther Classification System." Nat Protoc **8**(8): 1551-1566.
Mittler, R. and E. Lam (1997). "Characterization of Nuclease Activities and DNA Fragmentation Induced Upon Hypersensitive Response Cell Death and Mechanical Stress." Plant Mol Biol **34**(2): 209-221.
Miyado, K., E. Mekada and K. Kobayashi (2000). "[a Crucial Role of Tetraspanin, Cd9 in Fertilization]." Tanpakushitsu Kakusan Koso **45**(10): 1728-1734.
Mizukami, A. G., R. Inatsugi, J. Jiao, T. Kotake, K. Kuwata, K. Ootani, S. Okuda, S. Sankaranarayanan, Y. Sato, D. Maruyama, H. Iwai, E. Garenaux, C. Sato, K. Kitajima, Y. Tsumuraya, H. Mori, J. Yamaguchi, K. Itami, N. Sasaki and T. Higashiyama (2016). "The Amor Arabinogalactan Sugar Chain Induces Pollen-Tube Competency to Respond to Ovular Guidance." Curr Biol **26**(8): 1091-1097.
Moise, J. A., S. Han, L. Gudynaite-Savitch, D. A. Johnson and B. L. A. Miki (2005). "Seed Coats: Structure, Development, Composition, and Biotechnology." In Vitro Cellular & Developmental Biology-Plant **41**(5): 620-644.
Moll, C., L. Von Lyncker, S. Zimmermann, C. Kägi, N. Baumann, D. Twell, U. Grossniklaus and R. Groß-Hardt (2008). „Clo/Gfa1andatoare Novel Regulators of Gametic Cell Fate in Plants." The Plant Journal **56**(6): 913-921.
Moore, J. P., Q. J. Sattentau, P. J. Klasse and L. C. Burkly (1992). "A Monoclonal Antibody to Cd4 Domain 2 Blocks Soluble Cd4-Induced Conformational Changes in the Envelope Glycoproteins of Human Immunodeficiency Virus Type 1 (Hiv-1) and Hiv-1 Infection of Cd4+ Cells." J Virol **66**(8): 4784-4793.
Mori, T., T. Igawa, G. Tamiya, S. Y. Miyagishima and F. Berger (2014). "Gamete Attachment Requires Gex2 for Successful Fertilization in Arabidopsis." Curr Biol **24**(2): 170-175.
Mori, T., H. Kuroiwa, T. Higashiyama and T. Kuroiwa (2006). "Generative Cell Specific 1 Is Essential for Angiosperm Fertilization." Nat Cell Biol **8**(1): 64-71.
Muller, J. J., M. S. Weiss and U. Heinemann (2011). "Pan-Modular Structure of Microneme Protein Sml-2 from the Parasite Sarcocystis Muris at 1.95 a Resolution and Its Complex with 1-Thio-Beta-D-Galactose." Acta Crystallogr D Biol Crystallogr **67**(Pt 11): 936-944.
Mullis, K., F. Faloona, S. Scharf, R. Saiki, G. Horn and H. Erlich (1986). "Specific Enzymatic Amplification of DNA in Vitro: The Polymerase Chain Reaction." Cold Spring Harb Symp Quant Biol **51 Pt 1**: 263-273.
Murphy, E. and I. De Smet (2014). "Understanding the Ralf Family: A Tale of Many Species." Trends Plant Sci **19**(10): 664-671.
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.
Ngo, Q. A., H. Vogler, D. S. Lituiev, A. Nestorova and U. Grossniklaus (2014). "A Calcium Dialog Mediated by the Feronia Signal Transduction Pathway Controls Plant Sperm Delivery." Dev Cell **29**(4): 491-500.
Nishikawa, M., K. Hosokawa, M. Ishiguro, H. Minamioka, K. Tamura, I. Hara-Nishimura, Y. Takahashi, K. Shimazaki and H. Imai (2008). "Degradation of Sphingoid Long-Chain Base

1-Phosphates (Lcb-1ps): Functional Characterization and Expression of Atdpl1 Encoding Lcb-1p Lyase Involved in the Dehydration Stress Response in Arabidopsis." Plant Cell Physiol **49**(11): 1758-1763.

Nishimura, K., L. Han, E. Bianchi, G. J. Wright, D. De Sanctis and L. Jovine (2016). "The Structure of Sperm Izumo1 Reveals Unexpected Similarities with Plasmodium Invasion Proteins." Curr Biol **26**(14): R661-662.

Noman, A., M. Aqeel and S. He (2016). "Crispr-Cas9: Tool for Qualitative and Quantitative Plant Genome Editing." Front Plant Sci **7**: 1740.

Nonomura, K., K. Miyoshi, M. Eiguchi, T. Suzuki, A. Miyao, H. Hirochika and N. Kurata (2003). "The Msp1 Gene Is Necessary to Restrict the Number of Cells Entering into Male and Female Sporogenesis and to Initiate Anther Wall Formation in Rice." Plant Cell **15**(8): 1728-1739.

Obenauer, J. C., L. C. Cantley and M. B. Yaffe (2003). "Scansite 2.0: Proteome-Wide Prediction of Cell Signaling Interactions Using Short Sequence Motifs." Nucleic Acids Res **31**(13): 3635-3641.

Obrdlik, P., M. El-Bakkoury, T. Hamacher, C. Cappellaro, C. Vilarino, C. Fleischer, H. Ellerbrok, R. Kamuzinzi, V. Ledent, D. Blaudez, D. Sanders, J. L. Revuelta, E. Boles, B. Andre and W. B. Frommer (2004). "K+ Channel Interactions Detected by a Genetic System Optimized for Systematic Studies of Membrane Protein Interactions." Proc Natl Acad Sci U S A **101**(33): 12242-12247.

Ohto, U., H. Ishida, E. Krayukhina, S. Uchiyama, N. Inoue and T. Shimizu (2016). "Structure of Izumo1-Juno Reveals Sperm-Oocyte Recognition During Mammalian Fertilization." Nature **534**(7608): 566-569.

Okamoto, M., Y. Tsuboi, H. Goda, T. Yoshizumi, Y. Shimada and T. Hirayama (2012). "Multiple Hormone Treatment Revealed Novel Cooperative Relationships between Abscisic Acid and Biotic Stress Hormones in Cultured Cells." Plant Biotechnology **29**(1): 19-34.

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.

Oliveros, J. C., M. Franch, D. Tabas-Madrid, D. San-Leon, L. Montoliu, P. Cubas and F. Pazos (2016). "Breaking-Cas-Interactive Design of Guide Rnas for Crispr-Cas Experiments for Ensembl Genomes." Nucleic Acids Res **44**(W1): W267-271.

Ossowski, S., R. Schwab and D. Weigel (2008). "Gene Silencing in Plants Using Artificial Micrornas and Other Small Rnas." The Plant Journal **53**(4): 674-690.

Owen, O. E., S. C. Kalhan and R. W. Hanson (2002). "The Key Role of Anaplerosis and Cataplerosis for Citric Acid Cycle Function." J Biol Chem **277**(34): 30409-30412.

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.

Pagnussat, G. C., H. J. Yu, Q. A. Ngo, S. Rajani, S. Mayalagu, C. S. Johnson, A. Capron, L. F. Xie, D. Ye and V. Sundaresan (2005). "Genetic and Molecular Identification of Genes Required for Female Gametophyte Development and Function in Arabidopsis." Development **132**(3): 603-614.

Pagnussat, G. C., H. J. Yu and V. Sundaresan (2007). "Cell-Fate Switch of Synergid to Egg Cell in Arabidopsis Eostre Mutant Embryo Sacs Arises from Misexpression of the Bel1-Like Homeodomain Gene Blh1." Plant Cell **19**(11): 3578-3592.

Papini, A., S. Mosti and L. Brighigna (1999). "Programmed-Cell-Death Events During Tapetum Development of Angiosperms." Protoplasma **207**(3-4): 213-221.

Papini, A., S. Mosti, E. Milocani, G. Tani, P. Di Falco and L. Brighigna (2011). "Megasporogenesis and Programmed Cell Death in Tillandsia (Bromeliaceae)." Protoplasma **248**(4): 651-662.

Park, S. K., D. Rahman, S. A. Oh and D. Twell (2004). "Gemini Pollen 2, a Male and Female Gametophytic Cytokinesis Defective Mutation." Sex Plant Reprod **17**(2): 63-70.

Pata, M. O., Y. A. Hannun and C. K. Ng (2010). "Plant Sphingolipids: Decoding the Enigma of the Sphinx." New Phytol **185**(3): 611-630.

Pearce, G., D. S. Moura, J. Stratmann and C. A. Ryan, Jr. (2001). "Ralf, a 5-Kda Ubiquitous Polypeptide in Plants, Arrests Root Growth and Development." Proc Natl Acad Sci U S A **98**(22): 12843-12847.

Penfield, S., Y. Li, A. D. Gilday, S. Graham and I. A. Graham (2006). "Arabidopsis Aba Insensitive4 Regulates Lipid Mobilization in the Embryo and Reveals Repression of Seed Germination by

the Endosperm." Plant Cell **18**(8): 1887-1899.
Penfield, S., E. L. Rylott, A. D. Gilday, S. Graham, T. R. Larson and I. A. Graham (2004). "Reserve Mobilization in the Arabidopsis Endosperm Fuels Hypocotyl Elongation in the Dark, Is Independent of Abscisic Acid, and Requires Phosphoenolpyruvate Carboxykinase1." Plant Cell **16**(10): 2705-2718.
Pereira, A. M., A. L. Lopes and S. Coimbra (2016). "Arabinogalactan Proteins as Interactors Along the Crosstalk between the Pollen Tube and the Female Tissues." Frontiers in Plant Science **7**.
Pereira, A. M., A. L. Lopes and S. Coimbra (2016). "Jagger, an Agp Essential for Persistent Synergid Degeneration and Polytubey Block in Arabidopsis." Plant Signal Behav **11**(8): e1209616.
Pereira, A. M., M. S. Nobre, S. C. Pinto, A. L. Lopes, M. L. Costa, S. Masiero and S. Coimbra (2016). ""Love Is Strong, and You're So Sweet": Jagger Is Essential for Persistent Synergid Degeneration and Polytubey Block in Arabidopsis Thaliana." Mol Plant **9**(4): 601-614.
Perez-Hernandez, D., C. Gutierrez-Vazquez, I. Jorge, S. Lopez-Martin, A. Ursa, F. Sanchez-Madrid, J. Vazquez and M. Yanez-Mo (2013). "The Intracellular Interactome of Tetraspanin-Enriched Microdomains Reveals Their Function as Sorting Machineries toward Exosomes." J Biol Chem **288**(17): 11649-11661.
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.
Pfeuffer, M. and A. Jaudszus (2016). "Pentadecanoic and Heptadecanoic Acids: Multifaceted Odd-Chain Fatty Acids." Advances in Nutrition **7**(4): 730-734.
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.
Qian, W., D. Ma, C. Xiao, Z. Wang and J. Zhang (2012). "The Genomic Landscape and Evolutionary Resolution of Antagonistic Pleiotropy in Yeast." Cell Rep **2**(5): 1399-1410.
Rabiger, D. S. and G. N. Drews (2013). "Myb64 and Myb119 Are Required for Cellularization and Differentiation During Female Gametogenesis in Arabidopsis Thaliana." PLoS Genet **9**(9): e1003783.
Ran, F. A., P. D. Hsu, C. Y. Lin, J. S. Gootenberg, S. Konermann, A. E. Trevino, D. A. Scott, A. Inoue, S. Matoba, Y. Zhang and F. Zhang (2013). „Double Nicking by Rna-Guided Crispr Cas9 for Enhanced Genome Editing Specificity." Cell **154**(6): 1380-1389.
Randerath, K. (1970). "An Evaluation of Film Detection Methods for Weak Beta-Emitters, Particularly Tritium." Anal Biochem **34**: 188-205.
Ranf, S., N. Gisch, M. Schaffer, T. Illig, L. Westphal, Y. A. Knirel, P. M. Sanchez-Carballo, U. Zahringer, R. Huckelhoven, J. Lee and D. Scheel (2015). „A Lectin S-Domain Receptor Kinase Mediates Lipopolysaccharide Sensing in Arabidopsis Thaliana." Nat Immunol **16**(4): 426-433.
Reape, T. J., E. M. Molony and P. F. Mccabe (2008). "Programmed Cell Death in Plants: Distinguishing between Different Modes." J Exp Bot **59**(3): 435-444.
Resentini, F., P. Cyprys, J. G. Steffen, S. Alter, P. Morandini, C. Mizzotti, A. Lloyd, G. N. Drews, T. Dresselhaus, L. Colombo, S. Sprunck and S. Masiero (2017). "Suppressor of Frigida (Suf4) Supports Gamete Fusion Via Regulating Arabidopsis Ec1 Gene Expression." Plant Physiol **173**(1): 155-166.
Resh, M. D. (2004). "Membrane Targeting of Lipid Modified Signal Transduction Proteins." Subcell Biochem **37**: 217-232.
Rezanka, T., I. Kolouchova and K. Sigler (2015). "Precursor Directed Biosynthesis of Odd-Numbered Fatty Acids by Different Yeasts." Folia Microbiol (Praha) **60**(5): 457-464.
Rhee, S. Y., W. Beavis, T. Z. Berardini, G. Chen, D. Dixon, A. Doyle, M. Garcia-Hernandez, E. Huala, G. Lander, M. Montoya, N. Miller, L. A. Mueller, S. Mundodi, L. Reiser, J. Tacklind, D. C. Weems, Y. Wu, I. Xu, D. Yoo, J. Yoon and P. Zhang (2003). "The Arabidopsis Information Resource (Tair): A Model Organism Database Providing a Centralized, Curated Gateway to Arabidopsis Biology, Research Materials and Community." Nucleic Acids Res **31**(1): 224-228.
Richards, E., M. Reichardt and S. Rogers (2001). "Preparation of Genomic DNA from Plant Tissue." Curr Protoc Mol Biol **Chapter 2**: Unit2 3.
Ritchie, M. E., B. Phipson, D. Wu, Y. Hu, C. W. Law, W. Shi and G. K. Smyth (2015). "Limma Powers Differential Expression Analyses for Rna-Sequencing and Microarray Studies." Nucleic Acids Res **43**(7): e47.
Robinson-Beers, K., R. E. Pruitt and C. S. Gasser (1992). "Ovule Development in Wild-Type Arabidopsis and Two Female-Sterile Mutants." Plant Cell **4**(10): 1237-1249.
Rodkiewicz, B. (1970). "Callose in Cell Walls During Megasporogenesis in Angiosperms." Planta

93(1): 39-47.
Rodriguez-Enriquez, M. J., S. Mehdi, H. G. Dickinson and R. T. Grant-Downton (2013). "A Novel Method for Efficient in Vitro Germination and Tube Growth of Arabidopsis Thaliana Pollen." New Phytol **197**(2): 668-679.
Rogers, W. O., M. D. Rogers, R. C. Hedstrom and S. L. Hoffman (1992). "Characterization of the Gene Encoding Sporozoite Surface Protein 2, a Protective Plasmodium Yoelii Sporozoite Antigen." Mol Biochem Parasitol **53**(1-2): 45-51.
Rose, A. B., T. Elfersi, G. Parra and I. Korf (2008). "Promoter-Proximal Introns in Arabidopsis Thaliana Are Enriched in Dispersed Signals That Elevate Gene Expression." Plant Cell **20**(3): 543-551.
Rosso, M. G., Y. Li, N. Strizhov, B. Reiss, K. Dekker and B. Weisshaar (2003). "An Arabidopsis Thaliana T-DNA Mutagenized Population (Gabi-Kat) for Flanking Sequence Tag-Based Reverse Genetics." Plant Mol Biol **53**(1-2): 247-259.
Rubinstein, E., F. Le Naour, M. Billard, M. Prenant and C. Boucheix (1994). "Cd9 Antigen Is an Accessory Subunit of the Vla Integrin Complexes." Eur J Immunol **24**(12): 3005-3013.
Rubinstein, E., A. Ziyyat, M. Prenant, E. Wrobel, J. P. Wolf, S. Levy, F. Le Naour and C. Boucheix (2006). "Reduced Fertility of Female Mice Lacking Cd81." Dev Biol **290**(2): 351-358.
Russell, S. D., X. Gou, C. E. Wong, X. Wang, T. Yuan, X. Wei, P. L. Bhalla and M. B. Singh (2012). "Genomic Profiling of Rice Sperm Cell Transcripts Reveals Conserved and Distinct Elements in the Flowering Plant Male Germ Lineage." New Phytol **195**(3): 560-573.
Ryu, S. E., P. D. Kwong, A. Truneh, T. G. Porter, J. Arthos, M. Rosenberg, X. P. Dai, N. H. Xuong, R. Axel, R. W. Sweet and Et Al. (1990). "Crystal Structure of an Hiv-Binding Recombinant Fragment of Human Cd4." Nature **348**(6300): 419-426.
Sala-Valdes, M., A. Ursa, S. Charrin, E. Rubinstein, M. E. Hemler, F. Sanchez-Madrid and M. Yanez-Mo (2006). "Ewi-2 and Ewi-F Link the Tetraspanin Web to the Actin Cytoskeleton through Their Direct Association with Ezrin-Radixin-Moesin Proteins." J Biol Chem **281**(28): 19665-19675.
Satouh, Y., N. Inoue, M. Ikawa and M. Okabe (2012). "Visualization of the Moment of Mouse Sperm-Egg Fusion and Dynamic Localization of Izumo1." J Cell Sci **125**(Pt 21): 4985-4990.
Schindelin, J., I. Arganda-Carreras, E. Frise, V. Kaynig, M. Longair, T. Pietzsch, S. Preibisch, C. Rueden, S. Saalfeld, B. Schmid, J. Y. Tinevez, D. J. White, V. Hartenstein, K. Eliceiri, P. Tomancak and A. Cardona (2012). "Fiji: An Open-Source Platform for Biological-Image Analysis." Nat Methods **9**(7): 676-682.
Schmid, M. W., A. Schmidt, U. C. Klostermeier, M. Barann, P. Rosenstiel and U. Grossniklaus (2012). „A Powerful Method for Transcriptional Profiling of Specific Cell Types in Eukaryotes: Laser-Assisted Microdissection and Rna Sequencing." PLoS One **7**(1): e29685.
Schmidt, T. H., Y. Homsi and T. Lang (2016). "Oligomerization of the Tetraspanin Cd81 Via the Flexibility of Its Delta-Loop." Biophys J **110**(11): 2463-2474.
Schneitz, K., M. Hulskamp and R. E. Pruitt (1995). "Wild-Type Ovule Development in Arabidopsis Thaliana: A Light Microscope Study of Cleared Whole-Mount Tissue." The Plant Journal **7**(5): 731-749.
Schubert, D., B. Lechtenberg, A. Forsbach, M. Gils, S. Bahadur and R. Schmidt (2004). „Silencing in Arabidopsis T-DNA Transformants: The Predominant Role of a Gene-Specific Rna Sensing Mechanism Versus Position Effects." Plant Cell **16**(10): 2561-2572.
Schwab, R., J. F. Palatnik, M. Riester, C. Schommer, M. Schmid and D. Weigel (2005). „Specific Effects of Microrna on the Plant Transcriptome." Dev Cell **8**(4): 517-527.
Schwacke, R., A. Schneider, E. Van Der Graaff, K. Fischer, E. Catoni, M. Desimone, W. B. Frommer, U. I. Flugge and R. Kunze (2003). „Aramemnon, a Novel Database for Arabidopsis Integral Membrane Proteins." Plant Physiol **131**(1): 16-26.
Seigneuret, M., A. Delaguillaumie, C. Lagaudriere-Gesbert and H. Conjeaud (2001). "Structure of the Tetraspanin Main Extracellular Domain. A Partially Conserved Fold with a Structurally Variable Domain Insertion." J Biol Chem **276**(43): 40055-40064.
Seo, P. J. and C. M. Park (2011). "Signaling Linkage between Environmental Stress Resistance and Leaf Senescence in Arabidopsis." Plant Signal Behav **6**(10): 1564-1566.
Sessions, A., E. Burke, G. Presting, G. Aux, J. Mcelver, D. Patton, B. Dietrich, P. Ho, J. Bacwaden, C. Ko, J. D. Clarke, D. Cotton, D. Bullis, J. Snell, T. Miguel, D. Hutchison, B. Kimmerly, T. Mitzel, F. Katagiri, J. Glazebrook, M. Law and S. A. Goff (2002). "A High-Throughput Arabidopsis Reverse Genetics System." Plant Cell **14**(12): 2985-2994.
Shah, A., D. Chen, A. R. Boda, L. J. Foster, M. J. Davis and M. M. Hill (2015). "Raftprot: Mammalian Lipid Raft Proteome Database." Nucleic Acids Res **43**(Database issue): D335-338.

She, W. and C. Baroux (2014). "Chromatin Dynamics During Plant Sexual Reproduction." Front Plant Sci **5**: 354.
She, W., D. Grimanelli, K. Rutowicz, M. W. Whitehead, M. Puzio, M. Kotlinski, A. Jerzmanowski and C. Baroux (2013). "Chromatin Reprogramming During the Somatic-to-Reproductive Cell Fate Transition in Plants." Development **140**(19): 4008-4019.
Shen, W., Y. Wei, M. Dauk, Y. Tan, D. C. Taylor, G. Selvaraj and J. Zou (2006). "Involvement of a Glycerol-3-Phosphate Dehydrogenase in Modulating the Nadh/Nad+ Ratio Provides Evidence of a Mitochondrial Glycerol-3-Phosphate Shuttle in Arabidopsis." Plant Cell **18**(2): 422-441.
Sherma, J. (2008). "Planar Chromatography." Analytical Chemistry **80**(12): 4253-4267.
Shi, L., J. Bielawski, J. Mu, H. Dong, C. Teng, J. Zhang, X. Yang, N. Tomishige, K. Hanada, Y. A. Hannun and J. Zuo (2007). "Involvement of Sphingoid Bases in Mediating Reactive Oxygen Intermediate Production and Programmed Cell Death in Arabidopsis." Cell Res **17**(12): 1030-1040.
Shih, H. W., N. D. Miller, C. Dai, E. P. Spalding and G. B. Monshausen (2014). "The Receptor-Like Kinase Feronia Is Required for Mechanical Signal Transduction in Arabidopsis Seedlings." Current Biology **24**(16): 1887-1892.
Shiu, S. H. and A. B. Bleecker (2003). "Expansion of the Receptor-Like Kinase/Pelle Gene Family and Receptor-Like Proteins in Arabidopsis." Plant Physiol **132**(2): 530-543.
Siddiqi, I., G. Ganesh, U. Grossniklaus and V. Subbiah (2000). "The Dyad Gene Is Required for Progression through Female Meiosis in Arabidopsis." Development **127**(1): 197-207.
Simons, K. and E. Ikonen (1997). "Functional Rafts in Cell Membranes." Nature **387**(6633): 569-572.
Simpson, C., C. Thomas, K. Findlay, E. Bayer and A. J. Maule (2009). "An Arabidopsis Gpi-Anchor Plasmodesmal Neck Protein with Callose Binding Activity and Potential to Regulate Cell-to-Cell Trafficking." Plant Cell **21**(2): 581-594.
Singh, R., S. Singh, P. Parihar, R. K. Mishra, D. K. Tripathi, V. P. Singh, D. K. Chauhan and S. M. Prasad (2016). "Reactive Oxygen Species (Ros): Beneficial Companions of Plants' Developmental Processes." Frontiers in Plant Science **7**.
Skaar, K., H. J. Korza, M. Tarry, P. Sekyrova and M. Hogbom (2015). "Expression and Subcellular Distribution of Gfp-Tagged Human Tetraspanin Proteins in Saccharomyces Cerevisiae." PLoS One **10**(7): e0134041.
Slane, D., J. Kong, K. W. Berendzen, J. Kilian, A. Henschen, M. Kolb, M. Schmid, K. Harter, U. Mayer, I. De Smet, M. Bayer and G. Jurgens (2014). „Cell Type-Specific Transcriptome Analysis in the Early Arabidopsis Thaliana Embryo." Development **141**(24): 4831-4840.
Smedley, D., S. Haider, S. Durinck, L. Pandini, P. Provero, J. Allen, O. Arnaiz, M. H. Awedh, R. Baldock, G. Barbiera, P. Bardou, T. Beck, A. Blake, M. Bonierbale, A. J. Brookes, G. Bucci, I. Buetti, S. Burge, C. Cabau, J. W. Carlson, C. Chelala, C. Chrysostomou, D. Cittaro, O. Collin, R. Cordova, R. J. Cutts, E. Dassi, A. Di Genova, A. Djari, A. Esposito, H. Estrella, E. Eyras, J. Fernandez-Banet, S. Forbes, R. C. Free, T. Fujisawa, E. Gadaleta, J. M. Garcia-Manteiga, D. Goodstein, K. Gray, J. A. Guerra-Assuncao, B. Haggarty, D. J. Han, B. W. Han, T. Harris, J. Harshbarger, R. K. Hastings, R. D. Hayes, C. Hoede, S. Hu, Z. L. Hu, L. Hutchins, Z. Kan, H. Kawaji, A. Keliet, A. Kerhornou, S. Kim, R. Kinsella, C. Klopp, L. Kong, D. Lawson, D. Lazarevic, J. H. Lee, T. Letellier, C. Y. Li, P. Lio, C. J. Liu, J. Luo, A. Maass, J. Mariette, T. Maurel, S. Merella, A. M. Mohamed, F. Moreews, I. Nabihoudine, N. Ndegwa, C. Noirot, C. Perez-Llamas, M. Primig, A. Quattrone, H. Quesneville, D. Rambaldi, J. Reecy, M. Riba, S. Rosanoff, A. A. Saddiq, E. Salas, O. Sallou, R. Shepherd, R. Simon, L. Sperling, W. Spooner, D. M. Staines, D. Steinbach, K. Stone, E. Stupka, J. W. Teague, A. Z. Dayem Ullah, J. Wang, D. Ware, M. Wong-Erasmus, K. Youens-Clark, A. Zadissa, S. J. Zhang and A. Kasprzyk (2015). "The Biomart Community Portal: An Innovative Alternative to Large, Centralized Data Repositories." Nucleic Acids Res **43**(W1): W589-598.
Smyth, D. R., J. L. Bowman and E. M. Meyerowitz (1990). "Early Flower Development in Arabidopsis." Plant Cell **2**(8): 755-767.
Soljic, L. (2013). Microarray Analysis of Single Isolated Cells of the Female Gametophyte Reveals Potential Regulators of Female Germline Development in Arabidopsis Thaliana. Regensburg, Universität Regensburg.
Song, G., A. C. Koksal, C. Lu and T. A. Springer (2012). „Shape Change in the Receptor for Gliding Motility in Plasmodium Sporozoites." Proc Natl Acad Sci U S A **109**(52): 21420-21425.
Spencer, M. W., S. A. Casson and K. Lindsey (2007). "Transcriptional Profiling of the Arabidopsis Embryo." Plant Physiol **143**(2): 924-940.
Spiegelstein, O., J. D. Eudy and R. H. Finnell (2000). „Identification of Two Putative Novel

Folate Receptor Genes in Humans and Mouse." Gene **258**(1-2): 117-125.
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.
Stahl, E. (1956). "[Thin-Layer Chromatography; Methods, Influencing Factors and an Example of Its Use]." Pharmazie **11**(10): 633-637.
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.
Stegmann, M., J. Monaghan, E. Smakowska-Luzan, H. Rovenich, A. Lehner, N. Holton, Y. Belkhadir and C. Zipfel (2017). „The Receptor Kinase Fer Is a Ralf-Regulated Scaffold Controlling Plant Immune Signaling." Science **355**(6322): 287-+.
Stepanova, A. N., J. Yun, A. V. Likhacheva and J. M. Alonso (2007). "Multilevel Interactions between Ethylene and Auxin in Arabidopsis Roots." Plant Cell **19**(7): 2169-2185.
Su, Y., Q. Shi and W. Wei (2017). "Single Cell Proteomics in Biomedicine: High-Dimensional Data Acquisition, Visualization, and Analysis." Proteomics **17**(3-4).
Sun, K., K. Hunt and B. A. Hauser (2004). "Ovule Abortion in Arabidopsis Triggered by Stress." Plant Physiol **135**(4): 2358-2367.
Szklarczyk, D., J. H. Morris, H. Cook, M. Kuhn, S. Wyder, M. Simonovic, A. Santos, N. T. Doncheva, A. Roth, P. Bork, L. J. Jensen and C. Von Mering (2017). "The String Database in 2017: Quality-Controlled Protein-Protein Association Networks, Made Broadly Accessible." Nucleic Acids Res **45**(D1): D362-D368.
Takeda, Y., I. Tachibana, K. Miyado, M. Kobayashi, T. Miyazaki, T. Funakoshi, H. Kimura, H. Yamane, Y. Saito, H. Goto, T. Yoneda, M. Yoshida, T. Kumagai, T. Osaki, S. Hayashi, I. Kawase and E. Mekada (2003). "Tetraspanins Cd9 and Cd81 Function to Prevent the Fusion of Mononuclear Phagocytes." J Cell Biol **161**(5): 945-956.
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.
Takeuchi, H. and T. Higashiyama (2016). "Tip-Localized Receptors Control Pollen Tube Growth and Lure Sensing in Arabidopsis." Nature **531**(7593): 245-248.
Tang, X., M. H. Lim, J. Pelletier, M. Tang, V. Nguyen, W. A. Keller, E. W. Tsang, A. Wang, S. J. Rothstein, J. J. Harada and Y. Cui (2012). "Synergistic Repression of the Embryonic Programme by Set Domain Group 8 and Embryonic Flower 2 in Arabidopsis Seedlings." J Exp Bot **63**(3): 1391-1404.
Tanz, S. K., I. Castleden, C. M. Hooper, M. Vacher, I. Small and H. A. Millar (2013). "Suba3: A Database for Integrating Experimentation and Prediction to Define the Subcellular Location of Proteins in Arabidopsis." Nucleic Acids Res **41**(Database issue): D1185-1191.
Tekleyohans, D. G., T. Nakel and R. Gross-Hardt (2017). "Patterning the Female Gametophyte of Flowering Plants." Plant Physiol **173**(1): 122-129.
Ten Hove, C. A., K. J. Lu and D. Weijers (2015). "Building a Plant: Cell Fate Specification in the Early Arabidopsis Embryo." Development **142**(3): 420-430.
The Uniprot, C. (2017). "Uniprot: The Universal Protein Knowledgebase." Nucleic Acids Res **45**(D1): D158-D169.
Thieme, C. J., M. Rojas-Triana, E. Stecyk, C. Schudoma, W. Zhang, L. Yang, M. Minambres, D. Walther, W. X. Schulze, J. Paz-Ares, W. R. Scheible and F. Kragler (2015). "Endogenous Arabidopsis Messenger Rnas Transported to Distant Tissues." Nat Plants **1**(4): 15025.
Thompson, A. and M. J. Gasson (2001). "Location Effects of a Reporter Gene on Expression Levels and on Native Protein Synthesis in Lactococcus Lactis and Saccharomyces Cerevisiae." Appl Environ Microbiol **67**(8): 3434-3439.
Tintor, N., A. Ross, K. Kanehara, K. Yamada, L. Fan, B. Kemmerling, T. Nurnberger, K. Tsuda and Y. Saijo (2013). "Layered Pattern Receptor Signaling Via Ethylene and Endogenous Elicitor Peptides During Arabidopsis Immunity to Bacterial Infection." Proc Natl Acad Sci U S A **110**(15): 6211-6216.
Toledo, A., A. Perez, J. L. Coleman and J. L. Benach (2015). "The Lipid Raft Proteome of Borrelia Burgdorferi." Proteomics **15**(21): 3662-3675.
Tsegaye, Y., C. G. Richardson, J. E. Bravo, B. J. Mulcahy, D. V. Lynch, J. E. Markham, J. G. Jaworski, M. Chen, E. B. Cahoon and T. M. Dunn (2007). "Arabidopsis Mutants Lacking Long Chain Base Phosphate Lyase Are Fumonisin-Sensitive and Accumulate Trihydroxy-18:1 Long

Chain Base Phosphate." J Biol Chem **282**(38): 28195-28206.
Tsutsui, H. and T. Higashiyama (2016). "Pkama-Itachi Vectors for Highly Efficient Crispr/Cas9-Mediated Gene Knockout in Arabidopsis Thaliana." Plant Cell Physiol.
Tucker, M. R., T. Okada, Y. Hu, A. Scholefield, J. M. Taylor and A. M. Koltunow (2012). "Somatic Small Rna Pathways Promote the Mitotic Events of Megagametogenesis During Female Reproductive Development in Arabidopsis." Development **139**(8): 1399-1404.
Tyanova, S., T. Temu, P. Sinitcyn, A. Carlson, M. Y. Hein, T. Geiger, M. Mann and J. Cox (2016). "The Perseus Computational Platform for Comprehensive Analysis of (Prote)Omics Data." Nat Methods **13**(9): 731-740.
Uebler, S. (2016). Analysis of Cell-Cell Communication During Reproductive Processes by Ea1-Like Peptides in Maize. Regensburg, Universität Regensburg.
Untergasser, A., I. Cutcutache, T. Koressaar, J. Ye, B. C. Faircloth, M. Remm and S. G. Rozen (2012). „Primer3--New Capabilities and Interfaces." Nucleic Acids Res **40**(15): e115.
Urban, M. (2016). Gametogenesis-Related Small Rnas and Argonaute Proteins in Arabidopsis Thaliana. Regensburg, Universität Regensburg.
Usadel, B., O. E. Blasing, Y. Gibon, F. Poree, M. Hohne, M. Gunter, R. Trethewey, B. Kamlage, H. Poorter and M. Stitt (2008). „Multilevel Genomic Analysis of the Response of Transcripts, Enzyme Activities and Metabolites in Arabidopsis Rosettes to a Progressive Decrease of Temperature in the Non-Freezing Range." Plant Cell Environ **31**(4): 518-547.
Vaddepalli, P., A. Herrmann, L. Fulton, M. Oelschner, S. Hillmer, T. F. Stratil, A. Fastner, U. Z. Hammes, T. Ott, D. G. Robinson and K. Schneitz (2014). "The C2-Domain Protein Quirky and the Receptor-Like Kinase Strubbelig Localize to Plasmodesmata and Mediate Tissue Morphogenesis in Arabidopsis Thaliana." Development **141**(21): 4139-4148.
Van Went, J. L. (1970). "The Ultrastructure of the Synergids of Petunia." Acta Botanica Neerlandica **19**(2): 121-132.
Virant-Klun, I., S. Leicht, C. Hughes and J. Krijgsveld (2016). "Identification of Maturation-Specific Proteins by Single-Cell Proteomics of Human Oocytes." Mol Cell Proteomics **15**(8): 2616-2627.
Völz, R., J. Heydlauff, D. Ripper, L. Von Lyncker and R. Gross-Hardt (2013). "Ethylene Signaling Is Required for Synergid Degeneration and the Establishment of a Pollen Tube Block." Dev Cell **25**(3): 310-316.
Von Besser, K., A. C. Frank, M. A. Johnson and D. Preuss (2006). "Arabidopsis Hap2 (Gcs1) Is a Sperm-Specific Gene Required for Pollen Tube Guidance and Fertilization." Development **133**(23): 4761-4769.
Wagner, M. (1964). "[a Method for the Direct Determination of Desoxyribonuclease Following Agar Gel Electrophoresis]." J Chromatogr **14**: 107-108.
Waki, T., T. Hiki, R. Watanabe, T. Hashimoto and K. Nakajima (2011). "The Arabidopsis Rwp-Rk Protein Rkd4 Triggers Gene Expression and Pattern Formation in Early Embryogenesis." Curr Biol **21**(15): 1277-1281.
Wang, F., K. Vandepoele and M. Van Lijsebettens (2012). "Tetraspanin Genes in Plants." Plant Sci **190**: 9-15.
Wang, Z.-P., H.-L. Xing, L. Dong, H.-Y. Zhang, C.-Y. Han, X.-C. Wang and Q.-J. Chen (2015). "Egg Cell-Specific Promoter-Controlled Crispr/Cas9 Efficiently Generates Homozygous Mutants for Multiple Target Genes in Arabidopsis in a Single Generation." Genome Biology **16**(1).
Wang, T., L. Liang, Y. Xue, P. F. Jia, W. Chen, M. X. Zhang, Y. C. Wang, H. J. Li and W. C. Yang (2016). "A Receptor Heteromer Mediates the Male Perception of Female Attractants in Plants." Nature **531**(7593): 241-244.
Webb, M. C. and B. E. S. Gunning (1990). "Embryo Sac Development in Arabidopsis-Thaliana .1. Megasporogenesis, Including the Microtubular Cytoskeleton." Sexual Plant Reproduction **3**(4): 244-256.
Webb, M. C. and B. E. S. Gunning (1994). "Embryo Sac Development in Arabidopsis-Thaliana .2. The Cytoskeleton During Megagametogenesis." Sexual Plant Reproduction **7**(3): 153-163.
Werner-Washburne, M., E. Braun, G. C. Johnston and R. A. Singer (1993). "Stationary Phase in the Yeast Saccharomyces Cerevisiae." Microbiol Rev **57**(2): 383-401.
Westfall, P. H. (2014). "Kurtosis as Peakedness, 1905 - 2014. R.I.P." Am Stat **68**(3): 191-195.
Winter, D., B. Vinegar, H. Nahal, R. Ammar, G. V. Wilson and N. J. Provart (2007). "An "Electronic Fluorescent Pictograph" Browser for Exploring and Analyzing Large-Scale Biological Data Sets." PLoS One **2**(8): e718.
Wong, J. L. and M. A. Johnson (2010). "Is Hap2-Gcs1 an Ancestral Gamete Fusogen?" Trends Cell Biol **20**(3): 134-141.

Wright, B. S., J. W. Snow, T. C. O'brien and D. V. Lynch (2003). "Synthesis of 4-Hydroxysphinganine and Characterization of Sphinganine Hydroxylase Activity in Corn." Arch Biochem Biophys **415**(2): 184-192.
Wrzaczek, M., M. Brosche, H. Kollist and J. Kangasjarvi (2009). „Arabidopsis Gri Is Involved in the Regulation of Cell Death Induced by Extracellular Ros." Proc Natl Acad Sci U S A **106**(13): 5412-5417.
Wrzaczek, M., M. Brosche, J. Salojarvi, S. Kangasjarvi, N. Idanheimo, S. Mersmann, S. Robatzek, S. Karpinski, B. Karpinska and J. Kangasjarvi (2010). "Transcriptional Regulation of the Crk/Duf26 Group of Receptor-Like Protein Kinases by Ozone and Plant Hormones in Arabidopsis." BMC Plant Biol **10**: 95.
Wu, Y., Q. Xun, Y. Guo, J. Zhang, K. Cheng, T. Shi, K. He, S. Hou, X. Gou and J. Li (2016). "Genome-Wide Expression Pattern Analyses of the Arabidopsis Leucine-Rich Repeat Receptor-Like Kinases." Mol Plant **9**(2): 289-300.
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.
Xia, Y. J., H. Suzuki, J. Borevitz, J. Blount, Z. J. Guo, K. Patel, R. A. Dixon and C. Lamb (2004). "An Extracellular Aspartic Protease Functions in Arabidopsis Disease Resistance Signaling." Embo Journal **23**(4): 980-988.
Xiao, Y. L., J. C. Redman, E. L. Monaghan, J. Zhuang, B. A. Underwood, W. A. Moskal, W. Wang, H. C. Wu and C. D. Town (2010). "High Throughput Generation of Promoter Reporter (Gfp) Transgenic Lines of Low Expressing Genes in Arabidopsis and Analysis of Their Expression Patterns." Plant Methods **6**: 18.
Xu, C. and J. Min (2011). "Structure and Function of Wd40 Domain Proteins." Protein Cell **2**(3): 202-214.
Xu, P., S. L. Xu, Z. J. Li, W. Tang, A. L. Burlingame and Z. Y. Wang (2014). "A Brassinosteroid-Signaling Kinase Interacts with Multiple Receptor-Like Kinases in Arabidopsis." Mol Plant **7**(2): 441-444.
Xu, W., E. Fiume, O. Coen, C. Pechoux, L. Lepiniec and E. Magnani (2016). "Endosperm and Nucellus Develop Antagonistically in Arabidopsis Seeds." Plant Cell **28**(6): 1343-1360.
Yadegari, R. and G. N. Drews (2004). "Female Gametophyte Development." Plant Cell **16 Suppl**: S133-141.
Yamaguchi, M., M. Ohtani, N. Mitsuda, M. Kubo, M. Ohme-Takagi, H. Fukuda and T. Demura (2010). "Vnd-Interacting2, a Nac Domain Transcription Factor, Negatively Regulates Xylem Vessel Formation in Arabidopsis." Plant Cell **22**(4): 1249-1263.
Yamaguchi, T., K. Hirota, K. Nagahama, K. Ohkawa, T. Takahashi, T. Nomura and S. Sakaguchi (2007). "Control of Immune Responses by Antigen-Specific Regulatory T Cells Expressing the Folate Receptor." Immunity **27**(1): 145-159.
Yanez-Mo, M., O. Barreiro, M. Gordon-Alonso, M. Sala-Valdes and F. Sanchez-Madrid (2009). "Tetraspanin-Enriched Microdomains: A Functional Unit in Cell Plasma Membranes." Trends Cell Biol **19**(9): 434-446.
Yang, S. D., P. J. Seo, H. K. Yoon and C. M. Park (2011). "The Arabidopsis Nac Transcription Factor Vni2 Integrates Abscisic Acid Signals into Leaf Senescence Via the Cor/Rd Genes." Plant Cell **23**(6): 2155-2168.
Ye, J., G. Coulouris, I. Zaretskaya, I. Cutcutache, S. Rozen and T. L. Madden (2012). "Primer-Blast: A Tool to Design Target-Specific Primers for Polymerase Chain Reaction." BMC Bioinformatics **13**: 134.
Yuan, L., Z. Liu, X. Song, C. Johnson, X. Yu and V. Sundaresan (2016). "The Cki1 Histidine Kinase Specifies the Female Gametic Precursor of the Endosperm." Dev Cell **37**(1): 34-46.
Zhang, X., R. Henriques, S. S. Lin, Q. W. Niu and N. H. Chua (2006). "Agrobacterium-Mediated Transformation of Arabidopsis Thaliana Using the Floral Dip Method." Nat Protoc **1**(2): 641-646.
Zhang, Y., J. Yang and A. M. Showalter (2011). "Atagp18 Is Localized at the Plasma Membrane and Functions in Plant Growth and Development." Planta **233**(4): 675-683.
Zhao, D. Z. (2002). "The Excess Microsporocytes1 Gene Encodes a Putative Leucine-Rich Repeat Receptor Protein Kinase That Controls Somatic and Reproductive Cell Fates in the Arabidopsis Anther." Genes & Development **16**(15): 2021-2031.
Zhao, X., J. De Palma, R. Oane, R. Gamuyao, M. Luo, A. Chaudhury, P. Herve, Q. Xue and J. Bennett (2008). "Ostdl1a Binds to the Lrr Domain of Rice Receptor Kinase Msp1, and Is

Required to Limit Sporocyte Numbers." Plant J **54**(3): 375-387.
Zheng, Y., N. Ren, H. Wang, A. J. Stromberg and S. E. Perry (2009). "Global Identification of Targets of the Arabidopsis Mads Domain Protein Agamous-Like15." Plant Cell **21**(9): 2563-2577.
Zimmerman, B., B. Kelly, B. J. Mcmillan, T. C. Seegar, R. O. Dror, A. C. Kruse and S. C. Blacklow (2016). "Crystal Structure of a Full-Length Human Tetraspanin Reveals a Cholesterol-Binding Pocket." Cell **167**(4): 1041-1051 e1011.
Zipfel, C., G. Kunze, D. Chinchilla, A. Caniard, J. D. Jones, T. Boller and G. Felix (2006). "Perception of the Bacterial Pamp Ef-Tu by the Receptor Efr Restricts Agrobacterium-Mediated Transformation." Cell **125**(4): 749-760.
Ziyyat, A., E. Rubinstein, F. Monier-Gavelle, V. Barraud, O. Kulski, M. Prenant, C. Boucheix, M. Bomsel and J. P. Wolf (2006). "Cd9 Controls the Formation of Clusters That Contain Tetraspanins and the Integrin Alpha 6 Beta 1, Which Are Involved in Human and Mouse Gamete Fusion." J Cell Sci **119**(Pt 3): 416-424.

7 – Abbreviations

2-OH PA	2-hydroxy palmitic acid
AMS	ammoniapersulfate
BCIP	5-bromo-4-chloro-3-indolyl-phosphate
CC	central cell
CDS	coding sequence
CLSM	scanning laser microscopy
CIM	callus inducing medium
CN	chalazal nucleus
Cq	quantification cycle
CRISPR	clustered regularly interspaced short palindromic repeats
DEG	differentially expressed gene
DIC	differential interference contrast
DIG	digoxigenin
DNA	deoxyribunocleic acid
EC	egg cell
ER	endoplasmic reticulum
EXC	extra cellular
FA	fatty acid
FC	log2 fold change
FG	female gametophyte
FP	full proteome
gDNA	genomic DNA
GFP	green fluorescent protein
GO	gene ontology
GUS	beta glucuronidase
LAM	laser-assisted microdissection
LRR-RLK	leucine-rich repeat receptor-like kinase
MF	microsomal fractionation
MGU	male germ unit
MM	manual microdissection
MN	micropylar nucleus
MS	mass spectrometry
NAD	nicotine amide dinucleotide
NBT	nitrotetrazolim blue chloride
NRQ	normalized relative quantity
NTC	non template control
PA	palmitic acid
PCR	polymerase chain reaction
PHS	phytosphingosine
PHS1P	phytosphingosine-1-phosphate
PM	plasma membrane
PPO	2,5-diphenyl oxazole
PT	pollen tube
pV	probability value
qPCR	reverse-transcriptase quantitative real-time PCR
ROS	reactive oxygen species
RLK	receptor-like kinase
RNA	ribonucleic acid
RT-PCR	reverse-transcriptase PCR
SC	sperm cell
SE	standard error
SYN	synergid cell
TCA	trichloroacetic acid
T-DNA	transfer DNA
VAC	vacuole
VN	vegetative nucleus
WISH	whole mount *in situ* hybridization
WT	wild type

8 – List of Figures

8 – List of Figures

9 – List of Tables

10 – Publications

Parts of this work will be published in

Thomas Hackenberg, Lucija Soljic, Maxim Messerer, Daniel Lang, Klaus F. Mayer, and Stefanie Sprunck: Membrane-associated and secreted proteins of the *Arabidopsis* female gametophyte, identified by transcriptome analyses.
(manuscript in preparation)

Marc Urban, Nicolas Strieder, Andrea Bleckmann, **Thomas Hackenberg**, Maurits Evers, Lucija Šoljić, Thomas Dresselhaus, Julia C Engelmann*, Stefanie Sprunck*:Small RNA pathway components in the female reproductive cell lineage of *Arabidopsis thaliana.*
(manuscript in preparation)
*shared corresponding authorship

11 – Acknowledgements

I want to thank PD Dr. Stefanie Sprunck for supervising me, and offering this diverse project to me in the first place, with all its possibilities to explore and investigate. I am grateful that I was always free to pursue my own approaches, while benefiting from her valuable input and optimism and a big thanks for putting all the effort especially in correcting my thesis.

Also, I want to thank Prof. Dr Thomas Dresselhaus and Prof. Dr. Martin Parniske for mentoring and sharing their thoughts and ideas on the project.

Moreover, I want to thank Prof. Dr. Kay Schneitz for assessing my thesis.

I want to thank Prof. Dr. Thomas Dresselhaus because even before I started my Ph.D. I have always been most welcome in his department.

Many thanks to all my collaborators that contributed to the project. Especially Dr. Maxim Messerer and Dr. Daniel Lang for the invaluable help with the sea of transcriptomic data and Dr. Julia Mergner for the aquisition and processing of the proteomic data.

Big thanks to Andrea for always happily sharing her sheer endless pool of knowledge and expertise. Another big thanks to Uli for being the best Skitriporganizer ever and, apart from that, many thanks for the invaluable scientific discussions and advice - and the non-scientific discussions.

Big thanks to the Lab and Department as everyone was contributing to the relaxed and productive atmosphere, especially Alex, Hans, Brian, Marc, Sanna, Maria, Frank (!), Monika, Ingrid, Angelika, Michael, Annemarie, and Phil – best lab buddy ever.

Finally, thanks to Angelika and Volker for the neverending support in so many ways, without you I would not have gotten that far. I also want to thank my sister Andrea! aaaand many thanks to my friends and Karin for always supporting my cause.

12 – Eidesstattliche Erklärung

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit ohne unzulässige Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus anderen Quellen direkt oder indirekt übernommenen Daten und Konzepte sind unter Angabe des Literaturzitats gekennzeichnet.

Die Arbeit wurde bisher weder im In- noch im Ausland in gleicher oder ähnlicher Form einer anderen Prüfungsbehörde vorgelegt.

Thomas Hackenberg

Regensburg,