Membrane proteins involved in flowering plant gamete interactions

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

vorgelegt von
Philipp Cyprys

aus
Nikolai

im Jahr
2020

Das Promotionsgesuch wurde eingereicht am:

12.6.2020

Die Arbeit wurde angeleitet von:

PD Dr. Stefanie Sprunck

Unterschrift:

Contents

1. Publications

Parts of this work are published in:

Cyprys, P., Lindemeier, M., and Sprunck, S. (2019). Gamete fusion is facilitated by two sperm cell-expressed DUF679 membrane proteins. **Nature Plants** 5 (3): 253–257.

During my PhD I contributed to following other publications:

Kulich, I., Vogler, F., Bleckmann, A., **Cyprys, P.**, Lindemeier, M., Fuchs, I., Krassini, L., Schubert, T., Steinbrenner, J., Beynon, J., Falter-Braun, P., Längst, G., Dresselhaus, T., and Sprunck, S. (2020). ARO scaffolds limit ROP activity by recruiting RENGAPs to polar growth sites. Manuscript in revision

Mergner, J., Frejno, M., List, M., Papacek, M., Chen, X., Chaudhary, A., Samaras, P., Richter, S., Shikata, H., Messerer, M., Lang, D., Altmann, S., **Cyprys, P.**, Zolg, D.P., Mathieson, T., Bantscheff, M., Hazarika, R.R., Schmidt, T., Dawid, C., Dunkel, A., Hofmann, T., Sprunck, S., Falter-Braun, P., Johannes, F., Mayer, K.F.X., Jürgens, G., Wilhelm, M., Baumbach, J., Grill, E., Schneitz, K., Schwechheimer, C., and Kuster, B. (2020). Mass-spectrometry-based draft of the Arabidopsis proteome. **Nature** 579 (7799): 409–414.

Chen, J., Strieder, N., Krohn, N.G., **Cyprys, P.**, Sprunck, S., Engelmann, J.C., and Dresselhaus, T. (2017). Zygotic Genome Activation Occurs Shortly after Fertilization in Maize. **The Plant Cell** 29 (9): 2106–2125.

Resentini, F., **Cyprys, P.**, Steffen, J.G., Alter, S., Morandini, P., Mizzotti, C., Lloyd, A., Drews, G.N., Dresselhaus, T., Colombo, L., Sprunck, S., and Masiero, S. (2017). SUPPRESSOR OF FRIGIDA (SUF4) Supports Gamete Fusion via Regulating Arabidopsis EC1 Gene Expression. **Plant Physiology** 173 (1): 155–166.

Hou, Y., Guo, X., **Cyprys, P.**, Zhang, Y., Bleckmann, A., Le Cai, Huang, Q., Luo, Y., Gu, H., Dresselhaus, T., Dong, J., and Qu, L.-J. (2016). Maternal ENODLs Are Required for Pollen Tube Reception in Arabidopsis. **Current Biology** CB 26 (17): 2343–2350.

2. Summary

Flowering plants have evolved a unique mode of sexual reproduction termed double fertilization, where two female gametes, the egg cell and the central cell, get fertilized by one sperm cell each. The fertilized egg cell gives rise to the zygote, whereas fertilization of the central cell results in formation of the nutrient-storing and embryo-nourishing endosperm. Successful seed formation requires both fertilization events, but molecular mechanisms regulating gamete recognition, attachment and fusion are poorly understood. Prior to this work, only three proteins residing or acting on plant gamete surfaces were known: The sperm membrane protein GAMETE EXPRESSED2 (GEX2) is involved in pre-fusion attachment whereas HAPLESS2 (HAP2) functions as membrane fusogen. Cysteine-rich EGG CELL1 (EC1) peptides are secreted by the egg cell and act as sperm activating molecules by inducing HAP2 re-localization from endomembrane compartments to the sperm surface.

In order to identify novel sperm cell surface proteins with a role in gamete interactions, a protocol for bulk sperm cell isolation from maize was established in this work. Sperm cells were used for RNA-seq transcriptomic profiling and proteomics approaches. The protein composition of membrane-enriched sperm microsomal fractions was analyzed by high-throughput LC-MS/MS. Combined with publicly available *Arabidopsis* sperm transcriptome data, the maize sperm cell proteome and transcriptome data were used to select candidate genes, or gene families, for functional studies in *Arabidopsis*. Translational reporter activity in sperm cells was detected for 5 candidates. Assuming functional redundancies, CRISPR/Cas9 was established and genome edited mutants were screened for fertilization defects.

Loss-of-function mutants of two sperm-expressed DOMAIN OF UNKNOWN FUNCTION 679 Membrane Proteins (DMP8 and DMP9) displayed severe reproductive defects. *dmp8,9* double mutants showed impaired male fertility and up to four unfused *dmp8,9* sperm cells were frequently observed close to the female gametes. Beside fertilization failure, also aborted seeds were observed in *dmp8,9* siliques, caused by single fertilization events. Quantification revealed that *dmp8,9* sperm cells preferentially fertilize the central cell. To address whether gamete adhesion or fusion is impaired in sperm cells lacking functional DMP8 and DMP9, *in situ* cell adhesion assays were performed and showed that *dmp8,9* sperm cells frequently manage to attach to the egg cell, suggesting DMP8 and DMP9 act after gamete attachment but before or during fusion.

DMPs are short four-span transmembrane proteins conserved in *Viridiplantae* (green plants). To gain insight about functionally important DMP9 regions and evolutionary conserved protein function, complementation of the *dmp8,9* mutant was performed.

Deletion constructs expressing truncated DMP9 versions failed to rescue the mutant phenotype. Complementation with most ancient DMP from *Chlamydomonas reinhardtii* (*CrDMP*) was not successful, however a putative *DMP9* ortholog from *Amborella trichopoda* (*AmTrDMP*) managed to partially rescue the fusion-defective *dmp8,9* phenotype, suggesting functional conservation at least in flowering plants.

The biochemical function of DMPs is not yet known, but *Arabidopsis thaliana* DMP1 was suggested to be involved in membrane remodeling. Transient overexpression of DMP9-GFP in tobacco leaf epidermis cells induced DMP1-like membrane remodeling events, which are likely artificial.

Mutant *dmp8,9* sperm cells were furthermore used as a tool to investigate sperm-induced events in the egg cell. Upon sperm cell arrival, the *Arabidopsis* egg cell secretes EC1 proteins to render the sperm cells competent for fusion but the precise moment of triggered EC1 secretion is not yet known. Quantification of EC1-GFP signals in unfertilized egg cells and in those with adjacent, unfused *dmp8,9* sperm cells gave evidence that the secretion of EC1-GFP by the egg cell does not depend on DMP8/9 and is upstream of sperm adhesion. Furthermore, in mammals two plasma membrane-localized TETRASPANINS (TETs) are crucial for gamete interactions and accumulate at the egg-sperm contact site to form fusion competent membrane patches. To investigate whether a similar scenario holds true during flowering plant gamete interactions, the subcellular distribution of egg- and central cell-expressed *Arabidopsis thaliana* TET9-GFP was monitored during interaction with *dmp8,9* sperm cells. A role for TET9 in the formation of fusion-competent sites seems unlikely, as TET9-GFP remained uniformly distributed at the egg- and central cell plasma membrane with attached but unfused *dmp8,9* sperm cells. This suggests that TETs expressed in angiosperm gametes might fulfil different, yet unknown functions.

In summary, two sperm-expressed membrane proteins were discovered that function after gamete attachment to facilitate gamete fusion. It will be an important task for the future to unravel the biochemical-mechanistic function of DMP8 and DMP9, which may involve membrane remodeling, supporting the activity of the fusogen HAP2, or the activation or delivery of HAP2 to the sperm cell surface. The maize sperm transcriptomic and proteomic datasets that have been generated in this work will serve as valuable resources for the identification of more membrane-localized proteins with a role during gamete interaction and fusion.

3. Introduction

Sexual reproduction describes a mode of reproduction in which diploid organisms form haploid gametes via meiosis. The following fusion of gametes of opposite sex leads to the formation of a diploid zygote, and the original number of chromosomes is restored. Sexual reproduction is conserved throughout the eukaryotic tree of life and was likely present even in early pre-metazoans approximately 2 billion years ago, implying a monophyletic origin where sex has evolved once and was already utilized by the last eukaryotic common ancestor (Levin and King, 2013; Speijer, 2016). Sexual reproduction is omnipresent in eukaryotes, but flowering plants have evolved their very own twist on sex, characterized by a unique life cycle and mode of reproduction.

In contrast to the haplontic life cycle of algae, fungi and protists and the diplontic life cycle of animals, land plants undergo a haplodiplontic life cycle where the haploid as well as the diploid phase form individual multicellular generations (Strasburger, 1894): the haploid sporophyte forms spores and the diploid gametophyte forms gametes. In non-vascular land plants such as bryophytes the haploid gametophyte is the dominating generation and forms what is colloquially known as “moss plant”, while the sporophyte is reduced and dependent on the gametophyte (Reski and Cove, 2004).

3.1 Flowering plants reproduce by double fertilization

During the evolution of land plants, the haploid gametophytic generation was subsequently reduced until only present as a few-celled stage in flowering plants (angiosperms), where the diploid sporophyte forms the dominating generation (Figure 1). The sporophyte of angiosperms produces flowers in which specialized cells undergo meiosis resulting in haploid male and female spores which develop to male and female gametophytes, respectively. Ovules and pollen are the angiosperm gametophytes which produce and carry the gametes: egg and central cell (female gametes) and two sperm cells (male gametes).

The two sets of male and female gametes are required for an angiosperm-specific process called double fertilization (Navašin, 1898): both female gametes are fertilized by one sperm cell each. Fusion of the egg cell with one sperm cell gives rise to the diploid zygote, thereby completing the life cycle. The zygote will develop into an embryo within the seed. The endosperm results from fusion of the second sperm cell with the central cell and nourishes the developing embryo (Figure 1). Besides optimizing photosynthetic capacity, water use and downsizing of genomes, the evolvement of double fertilization allows flowering plants to perform internal fertilization independent of water, and is thought to

be one of the major traits that allowed flowering plants to outcompete gymnosperms and ferns, making them the most diverse and successful clade of land plants on earth, nowadays representing over 90 % of all terrestrial vegetation (Boer *et al.*, 2012; Brodribb and Feild, 2010; Simonin and Roddy, 2018; Crepet and Niklas, 2009).

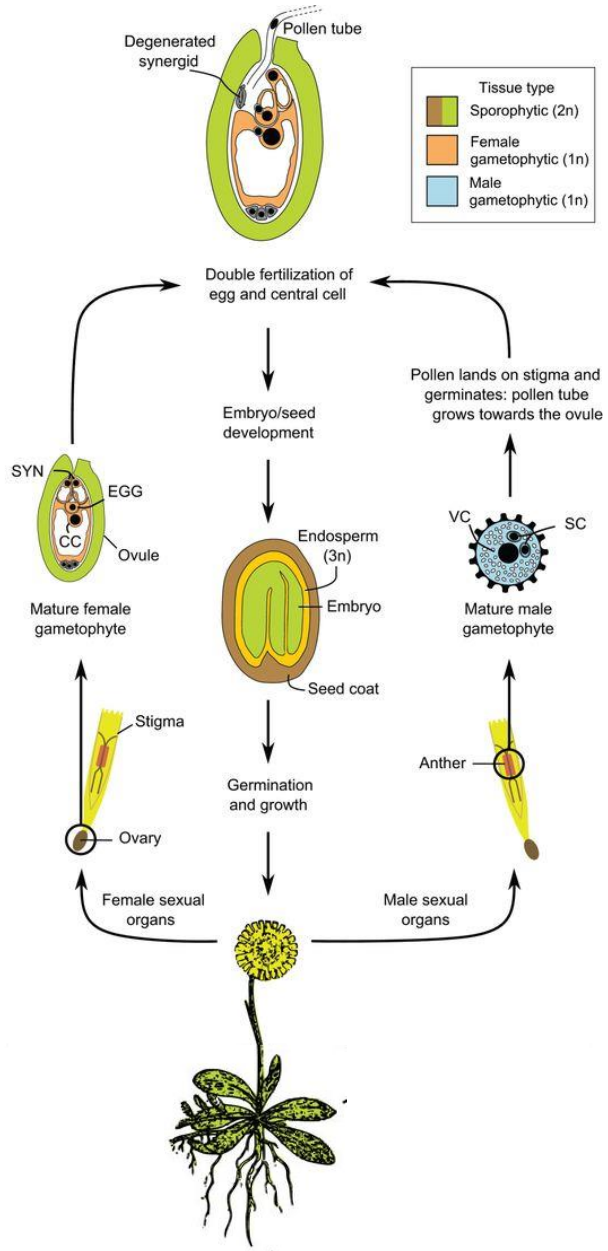


Figure 1 The life cycle of flowering plants. The diploid sporophyte produces flowers that contain male and female sexual organs in which meiosis takes place. Haploid meiosis products develop into gametophytes which carry the gametes. Gametes are are unified during double fertilization giving rise to the next diploid generation. VC = vegetative cell; SC = sperm cells; EGG = Egg cell; CC = Central cell; SYN = Synergids. Modified from Schmidt *et al.*, 2015.

3.2 Female gamete formation in *Arabidopsis thaliana*

In the model angiosperm *Arabidopsis thaliana*, the female gametophyte (embryo sac) develops within the ovule in two distinct developmental steps called megasporogenesis and megagametogenesis (reviewed by Yang *et al.*, 2010; Drews and Koltunow, 2011). During megasporogenesis an ovule primordium is formed and a single subepidermal nucellus cell differentiates into a diploid megaspore mother cell which then undergoes meiosis to form four haploid megaspores. One of the megaspores develops into the functional megaspore while the remaining three undergo programmed cell death (Figure 2). During megasporogenesis, development of sporophytic cell layers called integuments, that will later enclose the embryo sac, is initiated. During megagametogenesis, the functional megaspore undergoes three rounds of mitosis without cytokinesis resulting in a syncytium containing eight nuclei. Subsequent nuclear repositioning, vacuole formation, cellularization and nuclear fusion results in a mature seven-celled female gametophyte (embryo sac) that is almost completely surrounded by the integuments, except for the micropyle, a small opening where the pollen tube will grow in to deliver the two sperm cells for double fertilization (Dresselhaus *et al.*, 2016). The mature embryo sac is a highly polarized structure: one egg cell and two synergid cells are located at the micropylar pole followed by the diploid central cell and three antipodal cells that are positioned at the opposing chalazal pole (Figure 2).

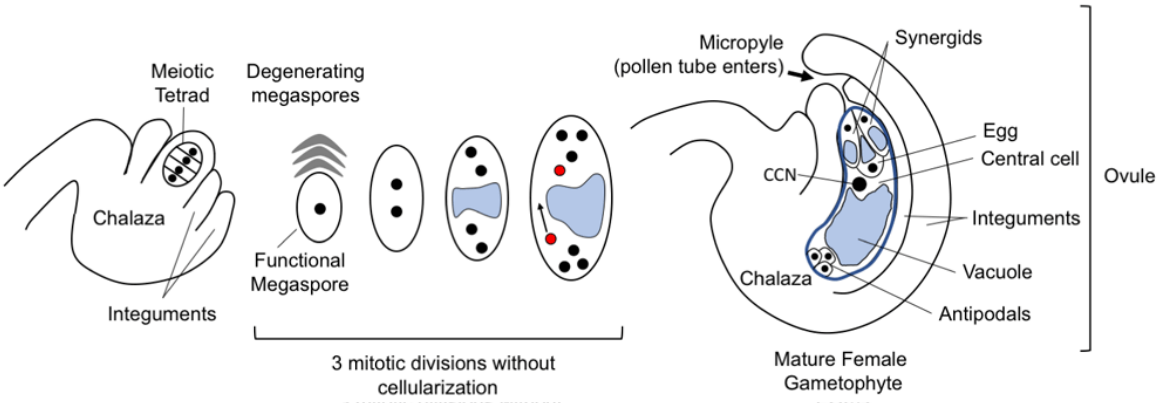


Figure 2 Development of the *Arabidopsis thaliana* female gametophyte. After meiosis, a single haploid cell will enlarge and form the functional megaspore while the remaining products of meiosis degenerate. This process is accompanied by development of surrounding sporophytic cell layers called integuments. The functional megaspore undergoes three mitotic divisions without cytokinesis, accompanied by nuclear positioning and formation of a vacuole (blue). Two polar nuclei (red) migrate to meet each other and fuse to generate the diploid central cell nucleus. The mature female gametophyte consists of seven cells, surrounded by integument layers of sporophytic origin: two synergids, one egg cell, one central cell with large diploid nucleus and three antipodal cells. Modified from Skinner and Sundaresan, 2018.

3.3 Male gamete formation in *Arabidopsis thaliana*

Analogous to the female gametophyte, *Arabidopsis* male gametophyte development consists of two sequential phases termed microsporogenesis and microgametogenesis and takes place within a specialized male reproductive organ called stamen (Figure 3; reviewed by Borg *et al.*, 2009). A diploid pollen mother cell undergoes meiosis, leading to formation and release of four haploid microspores. During microgametogenesis, the microspore nucleus migrates towards the periphery of the microspore, followed by an asymmetric cell division (pollen mitosis I; PMI). The resulting daughter cell consists of a larger vegetative cell and the smaller germ cell, which represents the male germline. Subsequently, the germ cell gets engulfed within the cytoplasm of the vegetative cell, forming a unique cell-within-a-cell structure. A second mitotic division (Pollen mitosis II; PMII) completes male gametophyte development: the germ cell divides to form twin sperm cells required for double fertilization (Figure 3). In plant species that shed tricellular pollen, including *Arabidopsis thaliana* and *Zea mays*, PMII is completed prior to anthesis whereas the majority of angiosperms shed bicellular pollen, where PMII takes place after pollination during growth of the pollen tube.

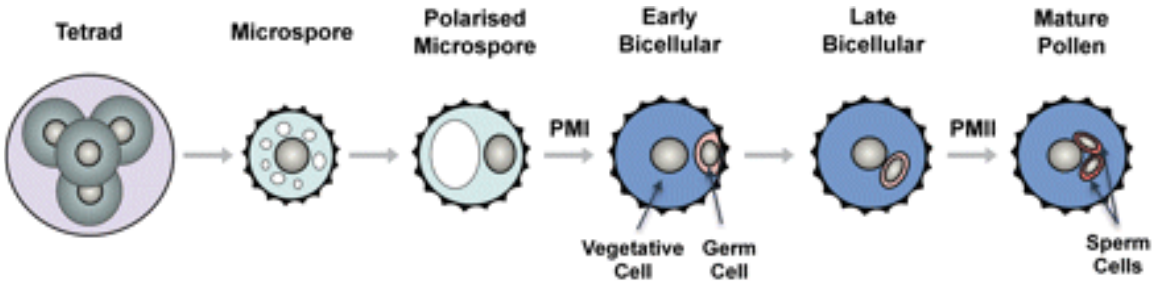


Figure 3 Development of the *Arabidopsis thaliana* male gametophyte. Meiosis of a diploid pollen mother cell leads to the formation of a tetrad followed by release of haploid microspores. Microspores undergo a highly asymmetric cell division called pollen mitosis I (PMI) to produce a bicellular pollen grain consisting of a small generative cell engulfed within the cytoplasm of a large vegetative cell. The generative cell undergoes pollen mitosis II (PMII) to produce two sperm cells. Modified from Borg *et al.*, 2009.

The genetic basis for sperm cell differentiation in *Arabidopsis thaliana* was unraveled by the discovery of DUO POLLEN1 (DUO1), a male germline-specific R2R3 MYB transcription factor: *duo1* pollen produces a single mutant germ cell that is unable to undergo fertilization (Durbarry *et al.*, 2005; Rotman *et al.*, 2005). DUO1 promotes *Arabidopsis* generative cell division and cell cycle progression cooperatively with the transcription factors DAZ1/2 (Borg *et al.*, 2014). DUO1 evolved from an ancient MYB transcription factor in the algal ancestors of land plants and neo-functionalization enabled the recognition of a new *cis*-regulatory element leading to the differentiation of motile sperm (Higo *et al.*,

2018). Modifications in the gene network regulated by DUO1 led to the evolvement of sperm with distinct morphology in land plants, making the emergence of DUO1 the defining event in the evolution of sperm differentiation, accountable for the diverse modes of sexual reproduction in the land plant lineage (Higo *et al.*, 2018). Importantly, DUO1 is a key regulator of sperm cell specification as it directly regulates the expression of sperm-specific genes including *HAPLESS2* (*HAP2*), *GAMETE EXPRESSED 2* (*GEX2*; see 3.6) and the sperm-specific histone variant *HTR10* (Brownfield *et al.*, 2009). Ectopic expression of DUO1 followed by microarray analysis for differentially expressed genes led to the identification of set of genes that is directly transcriptionally regulated by DUO1, many of which code for proteins of unknown biochemical function (Borg *et al.*, 2011).

In some species, e.g. *Nicotiana tabacum*, the two sperm cells have been found to be physically different: these dimorphic sperm species can differ in size and content of organelles such as plastids and mitochondria (Tian *et al.*, 2001). Nuclear sperm dimorphism was observed *in Zeya mays* sperm cells, where one sperm cell contains supernumerary B-chromosomes which are not homologous to the normal set of chromosomes. This sperm cell preferentially fertilizes the egg cell (Shi *et al.*, 1996; Faure *et al.*, 2003). However, the isomorphic sperm cells of *Arabidopsis thaliana* share an equal ability to fertilize the egg cell (Ingouff *et al.*, 2009) and studies with photolabeled sperm provided evidence that the two sperm cells have no preference in fertilizing either of the two female gametes (Hamamura *et al.*, 2011). Fusion of the sperm cells with the two female gametes happens almost simultaneously, with a slight tendency towards sperm-egg plasmogamy (Denninger *et al.*, 2014; Hamamura *et al.*, 2014).

3.4 Sperm cell delivery – the progamic phase

During the evolution of land plants, also sexual reproduction adapted to life in a nonaquatic environment. The sperm cells of flowering plants have lost their motility and are delivered to the female gametes by a pollen tube, a process called siphonogamy. After a compatible pollen grain has landed, adhered and hydrated on the stigma of the flower, the vegetative cell germinates a pollen tube, a rapidly tip-growing cell that navigates through the pistil and is guided towards the micropyle of the ovule, where it bursts to deliver the two sperm cells to the site of double fertilization (Figure 4A; Dresselhaus and Franklin-Tong, 2013). The precise delivery of exactly two sperm cells into each ovule requires extensive communication between pollen tube and female reproductive tissues as the pollen tube interacts with at least seven different pistil cell types prior to completing its journey (Palanivelu and Tsukamoto, 2012). Involved signaling components are as

diverse as amino acids, glycans, phytohormones, Ca^{2+}- and peptide-signaling (Dresselhaus and Franklin-Tong, 2013; Vogler *et al.*, 2014; Mizukami *et al.*, 2016).

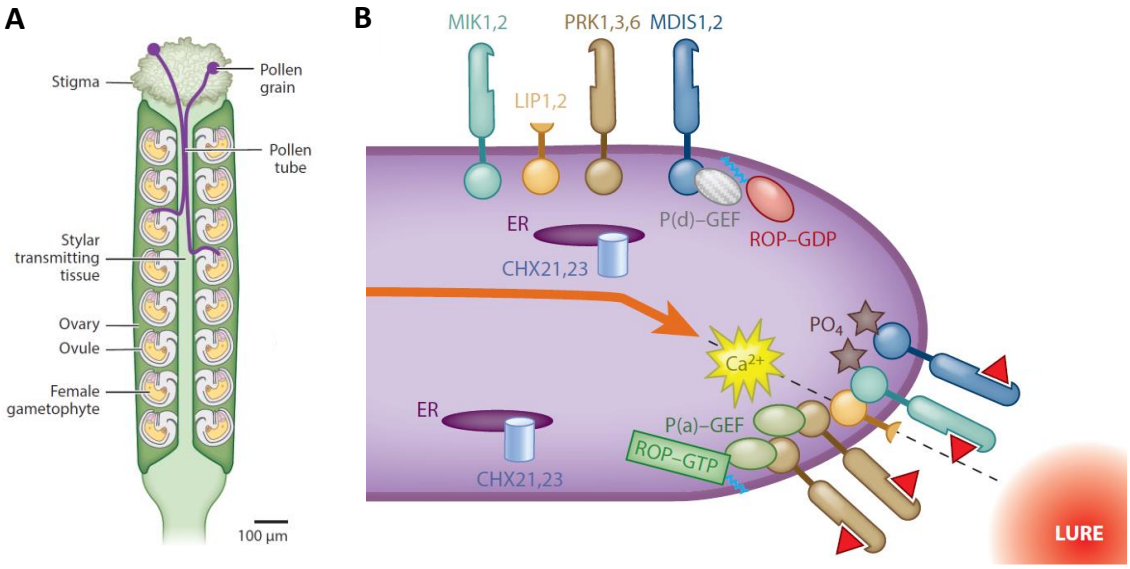


Figure 4 The pollen tube journey towards the synergids of the ovule in *Arabidopsis thaliana*. A, Pollen grains land on the stigma and germinate a pollen tube that enters the transmitting tissue of the style before exiting onto the interior surfaces of the ovary. The pollen tube then enters an ovule and contacts the female gametophyte. **B,** Synergid-secreted LURE proteins guide pollen tubes into the micropyle via receptor complexes. Pollen tube tip-localized RLK heteromers PRK1-3-6 and MDIS1/2-MIK1/2 bind LUREs (red triangles) followed by autophosphorylation and activation of ROP (ROP-GTP) and relocalization of a tip-focused Ca^{2+} maximum. This ultimately leads to reorientation of pollen tube tip growth (orange arrow) towards the LURE gradient (dashed line). Modified from Johnson *et al.*, 2019.

Attraction of the pollen tube towards the micropyle of the ovule is achieved by the synergids, professional secretory cells of the embryo sac (Higashiyama *et al.*, 2001). Located at the micropylar pole of the embryo sac, synergids secrete small, cysteine-rich LURE peptides that act as pollen tube attractants *in vivo* and are sufficient to attract pollen tubes *in vitro* (Okuda *et al.*, 2009). LUREs diverged fast during angiosperm evolution and act in a species-preferential manner by stimulating the emergence of conspecific pollen tubes from the transmitting tract, thereby promoting interspecific isolation (Takeuchi and Higashiyama, 2012; Zhong *et al.*, 2019a). Besides LUREs, other highly polymorphic Cysteine-Rich Peptides (CRPs) such as *Zea mays* EA1 and the non-species-specific XIUQIUs from *Arabidopsis* have been found to mediate pollen tube attraction, making CRPs the major class of pollen tube chemoattractants (Márton *et al.*, 2005; Zhong *et al.*, 2019a). LUREs are perceived by pollen tube tip-localized receptors belonging to the class of Receptor-Like Kinases (RLKs), the largest class of plant surface receptors (Figure 4B). The RLK heteromers PRK6-PRK1-PRK3 and MDIS1/2-MIK1/2 act in LURE sensing and induce a signaling cascade involving phosphorylation, Ca^{2+}- and Rho-GTPase signaling, ultimately

leading to a reorientation of pollen tube tip growth towards the LURE signal (Takeuchi and Higashiyama, 2016; Wang *et al.*, 2016; Malho and Trewavas, 1996).

Once the pollen tube has been guided to grow through the micropyle, its tip comes into contact with the one of the two synergid cells (receptive synergid) of the embryo sac, ceases growth and bursts (Palanivelu and Preuss, 2006). A complex set of molecular interactions is involved in the maintenance of pollen tube integrity during growth, pollen tube reception by the synergid and pollen tube burst leading to the explosive discharge of the two sperm cells into the female gametophyte. Among others, these include interactions between pollen tube and synergid-expressed RLKs and their ligands belonging to the large RAPID ALKALINIZATION FACTOR (RALF) family class of CRPs. GPI-anchored membrane proteins act cooperatively with RLKs and serve as chaperones and co-receptors (reviewed by Johnson *et al.*, 2019).

Pollen tube burst and discharge of the sperm cells into the female gametophyte is associated with degeneration of the receptive synergid (Leydon *et al.*, 2015). In case of failed fertilization by one or both of the sperm cells delivered by the first pollen tube, flowering plants have evolved a remarkable fertilization recovery system: the second, persisting synergid cell can attract an additional pollen tube for the delivery of another sperm pair (Kasahara *et al.*, 2012; Beale *et al.*, 2012). Furthermore, successful double fertilization triggers fusion of the persistent synergid with the much larger endosperm, thereby diluting synergid-expressed LURE pollen tube attractants to prevent attraction of supernumerary pollen tubes and polyspermy (Maruyama *et al.*, 2015).

3.5 Flowering plant sperm cell biology

Angiosperm sperm cells do not only differ from animal and primitive plant sperm cells by the absence of flagella, but also show several unique features such as the absence of a secondary cell wall, a small cytoplasmic volume and condensed chromatin (Southworth and Cresti, 1997).

Due to the engulfment of the generative cell by the vegetative cell during microgametogenesis (see 3.3), the sperm pair is surrounded by an endomembrane of the vegetative cell (Figure 5). A membrane projection extending from one sperm cell tethers the sperm pair to the outer envelope of the vegetative nucleus and forms a cytoplasmic connection between one sperm cell and the nucleus of the vegetative cell, forming a functional assemblage termed male germ unit (MGU; Dumas *et al.*, 1985; McCue *et al.*, 2011).

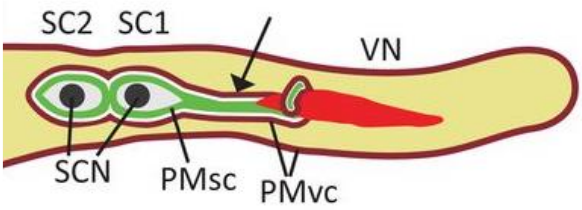


Figure 5 Morphology of the *Arabidopsis thaliana* male germ unit (MGU). The MGU consists of two interconnected sperm cells (SC1, SC2) surrounded by an endomembrane derived from the vegetative cell (PMvc; brown). The plasma membrane (PMsc; green) of one sperm cell forms a projection (arrow) that is tethered to the nucleus of the vegetative cell (VN; red) forming a stable association. The MGU is delivered to the site of double fertilization by the tip-growing pollen tube (yellow) and is transportet with the VN preceding the two sperm cells (PT growth direction: rightwards). Modified from Sprunck *et al.*, 2014.

The resulting stable association between the sperm pair and the vegetative nucleus is maintained until pollen tube discharge. The MGU enters the pollen tube only after rapid tip growth is initiated and is transported with the vegetative nucleus taking the lead (Vogler *et al.*, 2015). MGU integrity is essential for successful fertilization as shown by fertility problems in mutants where MGU movement or integrity is altered (Zhou and Meier, 2014).

After the sperm pair has been discharged into the embryo sac, the vegetative membrane needs to be removed to allow the gamete surfaces to come into contact (Sprunck, 2020). Although the underlying mechanism is unclear, it is conceivable that this membrane disintegrates due to the strong mechanical forces encountered by the MGU during explosive pollen tube discharge. This hypothesis is supported by the observation that the vegetative membrane ruptures easily when sperm cells are released from *in vitro* grown *Arabidopsis* pollen tubes (Sprunck *et al.*, 2014).

After becoming liberated from the vegetative membrane, local modification and removement of putative sperm cell wall material would be required to allow gamete plasma membranes to come into contact (Sprunck, 2020). Although presence of cell wall material surrounding *Arabidopsis thaliana* sperm cells has not yet been investigated, electron micrography of *Papaver rhoeas* generative cells and *Plumbago zeylanica* sperm cells revealed a thin and amorphous layer of cell wall material (McCue *et al.*, 2011; Russell and Cass, 1981). The role and fate of sperm cell wall material during flowering plant gamete interactions remains to be investigated. However, studies on *Chlamydomonas reinhardtii* showed that local degradation of the cell wall by a metalloprotease is a prerequisite for gamete fusion (Pan and Snell, 2000).

After pollen mitosis II the sperm pair forms a stable connection, which is formed during cytokinesis and remains just prior to fertilization (Charzynska and Lewandowska, 1990; Hamamura *et al.*, 2011). Electron micrography of the *Plumbago zeylanica* MGU followed by 3D reconstruction showed that this area is characterized by invaginations, enlarging the surface contact between the two sperm cells, similar to *Arabidopsis thaliana*, where interlocking cell borders at the sperm-sperm interface were observed (Russell, 1984;

McCue *et al.*, 2011). Notably, some sperm-expressed membrane proteins such as Tetraspanins TET11/12 accumulate at this membrane-rich region, which was proposed to represent a MGU-specific microdomain with unknown biological function (Boavida *et al.*, 2013).

3.6 Proteins involved in direct gamete interactions

Both gamete fusion events during double fertilization need to take place in a highly coordinated manner to ensure reproductive success and prevent polyspermy, requiring proteins localized to or acting on the cell surface of the opposing sex gametes. Although fertilization mechanisms differ profoundly between animals and flowering plants (e.g two fertilization events in flowering plants, sperm motility in animals), the discovery of such proteins (Figure 6) has shown that it is a sequential process, characterized by conserved cellular and molecular concepts (Dresselhaus *et al.*, 2016; Sprunck, 2020).

Live imaging of double fertilization in *Arabidopsis thaliana* revealed that the explosive release of the pollen tube content pushes the egg cell-central cell interface open and that the sperm pair is positioned in close proximity to the female gametes (Hamamura *et al.*, 2011). Two gamete fusion events take place between four and ten minutes after sperm cell discharge (Hamamura *et al.*, 2011). During this narrow timeframe the sperm pair needs to separate, attach to one female gamete and the fusion machinery needs to be locally assembled and activated. Tight adhesion between male and female gametes is required to bring opposing plasma membranes into close proximity of about 10 nm, which is considered the critical distance for a fusogenic protein (fusogen) to accomplish merging of the lipid bilayers (Hernández and Podbilewicz, 2017).

3.6.1 Gamete attachment

In flowering plants (Figure 6A) pre-fusion attachment is achieved by GEX2, a sperm-specific single pass transmembrane protein containing a large extracellular domain containing Filamin repeats and a short cytoplasmatic C-terminus (Mori *et al.*, 2014). GEX2-GFP localizes to the sperm plasma membrane and *gex2* mutants fail to attach to the female gametes, leading to severe fertility defects due to fertilization failure and single fertilization events. In case only one of the *gex2* two sperm cells fertilizes a female gamete this is preferentially the central cell (Mori *et al.*, 2014). A putative GEX2 interaction partner on female gametes remains to be discovered and the remaining ~25% successful double fertilization events by homozygous *gex2* sperm suggest that pre-fusion attachment in angiosperms involves more factors. Notably, the GEX2 ectodomain contains Filamin

domain repeats that are predicted to form a Ig-like fold similar to mammalian attachment factor IZUMO1 and the gamete attachment factor FUS1 from *Chlaydomonas reinhardtii* (Inoue *et al.*, 2005; Misamore *et al.*, 2003), suggesting gamete attachment factors from distant phyla may have adopted a similar tertiary structure, despite showing low primary sequence similarity (Mori *et al.*, 2014).

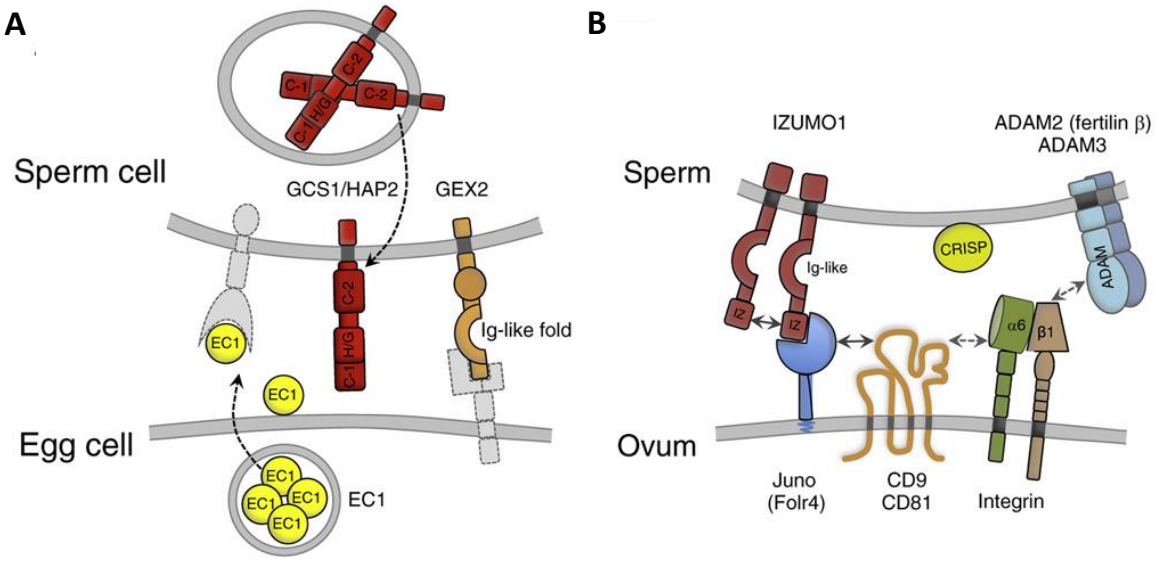


Figure 6 Molecular players for direct gamete interactions in *Arabidopsis thaliana* (A) and mammals (B). A, In *Arabidopsis thaliana,* the egg cell secretes cysteine-rich EC1 proteins upon sperm arrival. EC1 activates sperm cells by an unknown mechanism, leading to relocalization of the fusogen HAP2 from sperm endomembranes to the plasma membrane. Pre-fusion attachment is facilitated by sperm-expressed GEX2, a single pass transmembrane protein containing extracellular filamin repeats that are predicted to form an Immunoglobulin (Ig)-like fold. **B,** The ectodomain of the sperm-expressed single pass transmembrane protein IZUMO forms an Ig-like fold and binds to the oocyte-expressed GPI-anchor protein JUNO via its IZUMO-Domain (IZ). Tetraspanin CD9 accumulates at the fusion site upon IZUMO1-Juno interaction and interacts laterally with other membrane proteins on the egg (e.g. integrins). CD9 might have overlapping roles with Tetraspanin CD81. Other factors include egg-expressed integrin α6β1 dimers, sperm-expressed ADAMs (A Disintegrin and Metalloprotease) and sperm-binding Cysteine-rich secretory proteins (CRISP). Modified from Dresselhaus *et al.*, 2016.

The molecular mechanism by which GEX2 promotes tight adhesion between male and female gametes is not known, however the discovery of the sperm-localized mammalian single-pass transmembrane protein IZUMO1 (Figure 6B), also exhibiting an extracellular Immunoglobulin-like fold and its female binding partner JUNO, a GPI-anchor protein expressed on the egg cell plasma membrane, suggests that protein-protein interactions by gamete surface proteins are crucial for tight adhesion (Inoue *et al.*, 2005; Satouh *et al.*, 2012; Bianchi *et al.*, 2014). Fertilization-relevant GPI-anchored proteins were also found in *Arabidopsis thaliana*, such as LORELEI (LRE) and EARLY NODULIN-LIKE14 (EN14) which represent synergid-expressed determinants of pollen tube reception (Tsukamoto *et al.*, 2010; Hou *et al.*, 2016), or the pollen tube-expressed LORELEI-like-GPI-anchored proteins LLG2/LLG3, which function in maintenance of cellular integrity during pollen tube growth

(Mecchia *et al.*, 2017; Ge *et al.*, 2019). This highlights the role of GPI-anchored proteins during the progamic phase of angiosperm sexual reproduction, however the presence of such proteins on the gamete surfaces and putative roles in direct gamete interactions remain to be discovered.

While IZUMO1-JUNO is the only male-female protein interaction pair essential for tight gamete adhesion identified in any known living organism to date, other non-essential mammalian attachment factors such as sperm-expressed ADAM (A Disintegrin and A Metalloprotease) proteins, interacting with egg-expressed Integrins as well as Cysteine-rich Secretory Proteins (CRISPs), seem to facilitate attachment of the sperm to the oocyte (Almeida *et al.*, 1995; Eto *et al.*, 2002; Desiderio *et al.*, 2010; Cohen *et al.*, 2008). Other fertilization-relevant proteins in mammals include the egg-expressed Tetraspanins (TETs) CD9 and CD81 (Kaji *et al.*, 2000; Rubinstein *et al.*, 2006). CD9 accumulates at the at the egg-sperm contact site prior to fusion to generate fusion competent sites (Chalbi *et al.*, 2014; Jégou *et al.*, 2011). Concomitant with TETs functioning in the organization of multimolecular membrane complexes, it is thought that egg-expressed TET-enriched microdomains organize and cluster gamete interaction proteins on the egg surface by recruiting *cis*-interaction partners such as integrins (Hemler, 2005; Ziyyat *et al.*, 2006). Notably AtTET9, one of the 17 *Arabidopsis thaliana* Tetraspanins localizes to the plasma membranes of egg and central cell (Boavida *et al.*, 2013). However, a functional role of AtTET9 during gamete interactions in angiosperms remains to be discovered, and it is not known whether angiosperm TETs behave similar to mammalian CD9 and CD81 by accumulating at the contact site of male and female gametes.

3.6.2 Gamete fusion

After gamete plasma membranes have been brought into close apposition, high kinetic energy barriers need to be overcome to promote the merging of two lipid bilayers, a process that is facilitated by dedicated proteins called fusogens (Hernández and Podbilewicz, 2017).

While pre-fusion attachment during mammal gamete interaction is well understood on a molecular level, the mammalian fusogen mediating lipid bilayer merging of the gamete plasma membranes remains to be discovered. In *Arabidopsis thaliana*, a sperm cell fusogen candidate *HAPLESS2* (*HAP2*) was initially discovered in a genetic screen for haploid-disrupting (*hapless*) mutant alleles that fail to transmit to the progeny through the male parent (Johnson *et al.*, 2004). An independently discovered *hap2* allele, termed *generative cell–specific 1* (*gcs1*), was identified using generative cells isolated from *Lilium*

longiflorum pollen (Mori *et al.*, 2006). It was shown that *HAP2/GCS1* (henceforth *HAP2*) codes for a single pass transmembrane protein that localizes to the sperm cell plasma membrane where it acts as crucial fertilization factor, as *hap2* mutant sperm cells fail to fuse with the female gametes (Mori *et al.*, 2006; Besser *et al.*, 2006). HAP2 is a highly conserved protein with orthologs present in genomes of all major eukaryotic taxa except fungi, suggesting that HAP2 was already present in an ancestor common to all eukaryotes (Wong and Johnson, 2010). Studies in the green algae *Chlamydomonas reinhardtii*, the malaria parasite *Plasmodium berghei* and *Arabidopsis thaliana* showed that *hap2* mutant male gametes successfully adhere to the female gametes, but fail to fuse (Liu *et al.*, 2008; Mori *et al.*, 2014).

Only recently, crystallographic studies on the *Chlamydomonas reinhardtii* HAP2 ectodomain revealed astonishing structural similarities between HAP2 and viral class II fusion proteins, that are utilized by enveloped viruses to enter their host cells (Fédry *et al.*, 2017). Class II fusion proteins are characterized by three extracellular domains, a single transmembrane domain and a smaller, variable intracellular domain. Membrane fusion is facilitated by trimerization which is concomitant with insertion of a previously hidden hydrophobic protein segment, the "fusion loop", into the target membrane, followed by drastic conformational changes (Figure 7). The trimers fold back on themselves thereby bringing the two lipid bilayers into intimate contact followed by mixing of the outer leaflets (hemifusion) and formation of a fusion pore (reviewed by Kielian and Rey, 2006).

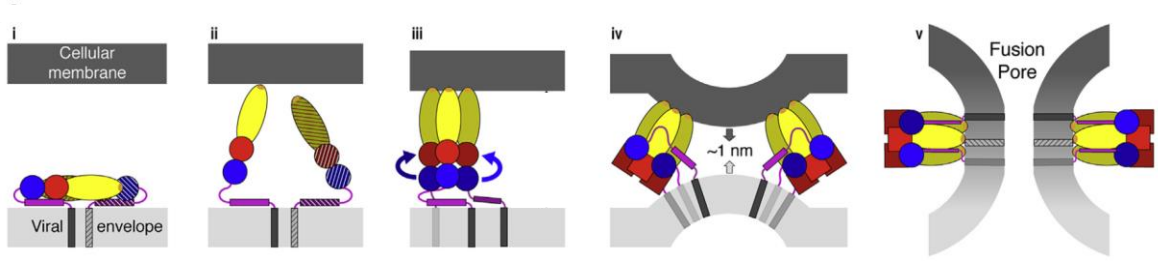


Figure 7 The molecular mechanism of membrane fusion catalyzed by viral class II fusion proteins. The three extracellular domains of class II fusion proteins are colored red, yellow and blue**. i,** Preceding fusion, the fusion loop (orange) is buried with in the pre-fusion homodimer. **ii,** The acidic environment of the endosome triggers dimer dissociation and exposure of the fusion loops. **iii,** Insertion of the fusion loop into the target membrane is concomitant with homotrimerization, and structural rearrangement of domain III (blue arrows), forming an intermediate state. **iv,** Domain III relocation redirects the stems towards the fusion loops thereby pulling the opposing membranes against each other (grey arrows). **v,** Upon contact of the two membranes, fusion pore formation is initiated via formation of a hemifusion intermediate. Modified from Baquero *et al.*, 2019.

The structure of the *Arabidopsis thaliana* HAP2 ectodomain was recently solved and requirement of an apical amphipathic fusion helix for membrane insertion *in vitro* and gamete fusion *in vivo* was demonstrated (Fedry *et al.*, 2018). The same study also revealed

strong diversification of the HAP2 membrane insertion motifs between species, possibly to accommodate to the highly diverse membrane environments that HAP2 encountered during eukaryotic evolution (Fedry *et al.*, 2018). So far, HAP2 has not been identified in mammals, however ectopically expressed *Arabidopsis thaliana* HAP2 can promote mammalian cell-cell fusion in cultured cells when present in both fusing cells (Valansi *et al.*, 2017).

3.6.3 Gamete activation

During viral infection of a host cell, molecular rearrangements that result in the exposure of the fusion loop of class II fusion proteins are triggered by the low pH of the endosome (Helenius *et al.*, 1980). Similarly, activity of sperm-expressed HAP2 needs to be regulated, probably to prevent polyspermy, or the inadvertent fusion of the two sperm cells (Sprunck, 2020). Notably, in *Arabidopsis thaliana* HAP2-YFP predominantly localizes to sperm endomembrane compartments just prior to fertilization and gets redistributed to the sperm surface by a remarkable mechanism: Upon sperm cell arrival, the egg cell secretes small cysteine-rich EGG CELL1 (EC1) proteins previously stored in intracellular vesicular structures (Figure 6B). EC1 proteins are suggested to act as sperm-activating molecules, since stimulation of released sperm cells with peptides spanning the conserved EC1 signature motifs induces the relocalization of HAP2-YFP from endomembrane compartments to the plasma membrane. Thus, *Arabidopsis thaliana* sperm cells are activated before fusion to render them competent for merging with the female gametes (Sprunck *et al.*, 2012). Consequently, EC1-deficient mutants show severe sperm fusion defects, as well as very delayed gamete fusion events (Sprunck *et al.*, 2012; Rademacher and Sprunck, 2013). EC1 proteins seem to have evolved specifically in the angiosperm lineage and *EC1*-like transcripts were also detected in the egg cells of *Amborella trichopoda*, the sister species to all other angiosperms, indicating that EC1-mediated sperm activation is an ancient and conserved mechanism during double fertilization (Sprunck *et al.*, 2014; Flores-Tornero *et al.*, 2019). Neither the trigger nor the exact timepoint of EC1 secretion are known. Live cell imaging of cytosolic Ca^{2+} dynamics in the egg cell during sperm cell delivery and gamete fusion showed that the egg cell responds to pollen tube discharge with a short Ca^{2+}-spike and this Ca^{2+} signal might trigger secretion of EC1 by the egg cell (Denninger *et al.*, 2014; Hamamura *et al.*, 2014).

3.7 Aims of this work

Both gamete fusion events are required for flowering plant seed formation. As seeds form the basis for most of our food and livestock feed, gaining molecular understanding of the processes involved in seed formation is crucial. This includes the knowledge on how gametes interact with each other on the molecular level and how the two sperm cells precisely fuse with one female gamete each. However only three proteins localized to or acting on the surfaces of angiosperm gametes are known to date.

This work aimed to identify novel sperm-expressed membrane proteins with a role during gamete interactions. To achieve this, a sperm cell isolation procedure from maize pollen was established and maize sperm cells were subjected to transcriptomic and proteomic profiling: Sperm cell total proteins and membrane protein-enriched sperm microsomal fractions were measured by LC-MS/MS and sperm cell transcriptomic profiles were generated by RNAseq. Screening these datasets together with *Arabidopsis* sperm cell transcriptome data aimed to identify promising new plasma membrane-localized candidates.

Arabidopsis thaliana was chosen as model organism for detailed functional analysis of selected proteins. This involved the validation of sperm-specific expression and membrane localization of individual candidates in stable translational reporter lines, as well as loss-of-function studies. Assuming functional redundancies between gene family members, this work also aimed to establish multiplex genome editing by CRISPR/Cas9 to generate higher order mutants to screen for fertilization defects. Promising mutants were subjected to in-depth analysis with the aim to exclude developmental, progamic or zygotic effects and to identify mutants with a truly sperm-related effect. *In situ* sperm attachment assays were established to address whether gamete adhesion or fusion is impaired in mutants with fusion-defective sperm.

4. Results

Maize was chosen as a model plant for generating sperm cell transcriptomic and proteomic data because of its agronomic importance and the possibility of harvesting pollen on a massive scale for subsequent sperm cell isolation. With a diameter of around 10 µm (Cyprys, 2015), maize sperm cells are larger than *Arabidopsis* sperm cells (1-2 µm; Borges *et al.*, 2008), allowing better morphological analysis, cell counting and assessment of viability. Finally, maize pollen can be disrupted by osmotic shock thereby gently releasing the sperm cells into the surrounding solution, without the necessity to apply mechanical force.

4.1 Isolation of maize sperm cells

Zea mays is a tricellular pollen species and mature maize pollen grains are composed of three cells (Figure 8A). Two interconnected sperm cells are completely enclosed within the vegetative cell cytoplasm, surrounded by an endogenous membrane of the vegetative cell. A membrane projection tethers one sperm cell to the nucleus of the vegetative cell, forming a physical connection termed male germ unit. The whole structure is encapsulated by a cell wall with a germination pore, a cavity through which the pollen tube will emerge.

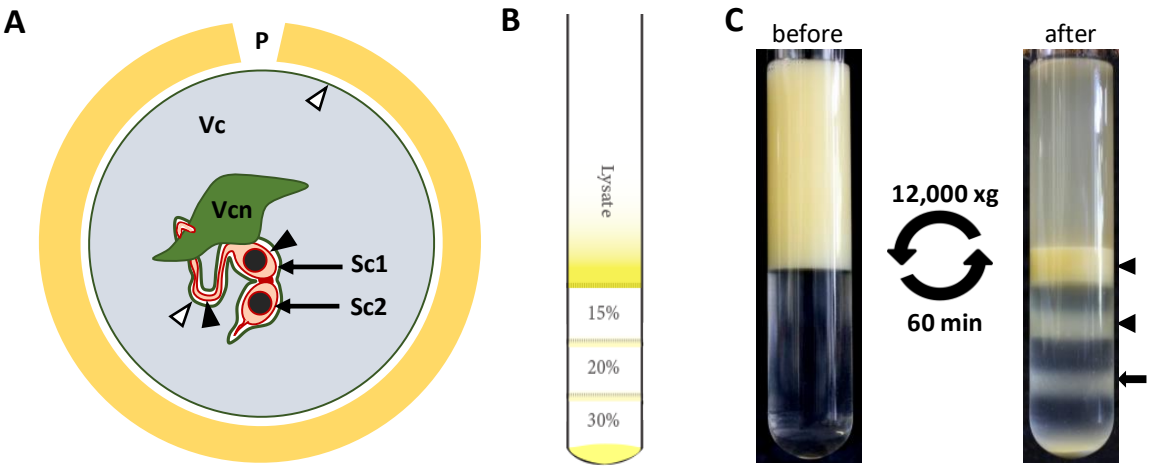


Figure 8 Isolation of maize sperm cells. A, Schematic drawing of a mature maize pollen grain showing two interconnected sperm cells (Sc), tethered to the pollen vegetative nucleus (Vcn). Sperm cells are enclosed in a common membrane of the pollen vegetative cell (green; hollow arrowhead) within its cytoplasm (Vc). The sperm plasma membrane (red; arrowhead) of one sperm cell (Sc1) forms a membrane projection which is physically associated with the nucleus of the vegetative cell (Vcn). The pollen cell wall (yellow) consisting of exine, intine and a cavity termed germination pore (P) is simplified in this drawing. **B,** Schematic drawing of the Percoll gradient used for sperm cell enrichment. After osmotic treatment, the filtrated pollen lysate containing sperm cell protoplasts is layered on top of a discontinuous gradient consisting of 3 phases (30% (v/v) Percoll; 20% (v/v) Percoll; 15% (v/v) Percoll in 0.55 M Mannitol). **C,** After centrifugation three distinct interphases are observed: while the two uppermost interphases mostly accumulate starch and cellular debris (arrowheads), the 20%/30% interphase (arrow) is enriched in sperm cells.

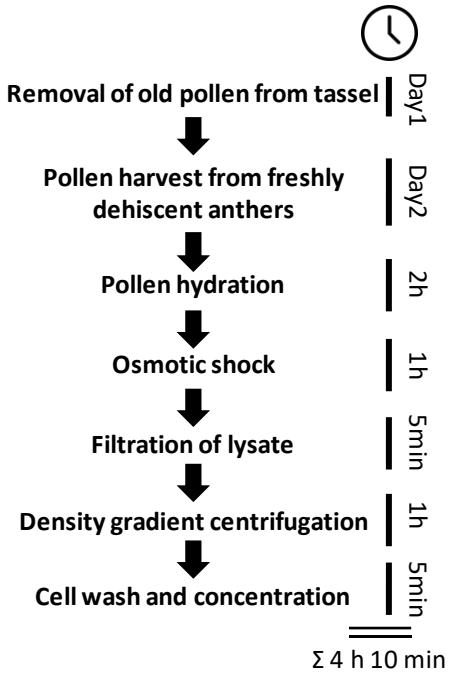


Figure 9 Timeline of the maize sperm cell isolation procedure. With this protocol sperm cell isolation can be performed in as little as 4 h 10 min using standard molecular biology lab equipment.

A protocol described by Dupuis *et al.* (1987) combining pollen rupture by osmotic shock and sperm enrichment on a discontinuous Percoll density gradient was improved at the beginning of this work. The sperm cell isolation procedure started with fresh pollen, after old pollen grains and anthers were removed the evening before by shaking tassels inside a paper bag. After harvesting fresh pollen, it was hydrated by incubating the pollen grains in a vessel under high humidity, followed by pollen grain rupture using a 0.55 M Mannitol solution as osmotic agent, to release the sperm cells into the solution. The lysate was filtered to remove empty and unruptured pollen grains and layered on top of a discontinuous Percoll gradient (Figure 8B), followed by centrifugation. After centrifugation, three interphases became visible as faint yellowish lines (Figure 8C). While the upper two interphases mostly consisted of starch granules and cellular debris, the lower (20/30 %) interphase was enriched in sperm cells. Sperm cell-containing fractions were carefully aspirated and subjected to washing, concentration and cell counting. For the calculation of sperm cell yield it was assumed that 1 mg of maize pollen consists of approximately 2,500 pollen grains (Dupuis *et al.*, 1987) and each pollen grain contains two sperm cells, resulting in a theoretical maximum yield (100 %) of 5 million sperm cells per gram of pollen. Sperm cell isolations performed in the course of this work yielded between 0.5 and 2.5 million sperm cells per gram of pollen (≙ 10 % - 50% yield). It became apparent that the sperm cell isolation procedure was less efficient when pollen harvested on hot summer days was used as starting material.

Figure 9 summarizes the timeline of the protocol and pinpoints the speed of the procedure, which takes only around 4 hours from pollen harvest to isolated sperm cells and can be performed with standard molecular biology laboratory equipment.

4.2 Sperm cell morphology

Isolated sperm cells were subjected to DIC microscopy for morphological analysis. The sperm cells appeared as round protoplasts with a diameter of approximately 10 µm and lacked a visible cell wall (Figure 10A). The sperm cells showed a tendency to cluster together, and aggregates containing numerous sperm cells were frequently observed. This indicates a sticky sperm cell surface and a general tendency to attach to other cells. Despite being highly enriched in sperm cells, the isolated fractions showed some contamination with starch granules, which are highly abundant in maize pollen, and became visible as ellipsoid spots of varying size.

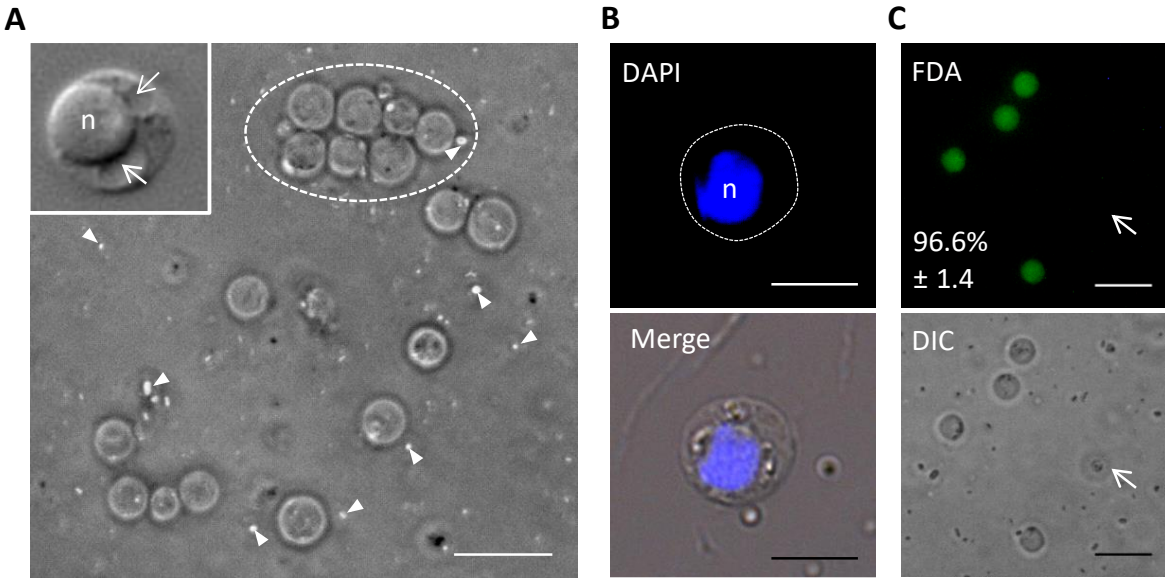


Figure 10 Morphological characteristics of isolated maize sperm cells. A, DIC microscopy reveals that maize sperm cells appear as spherical protoplasts with a diameter of approximately 10 µm and tend to agglutinate in the 0.55 M Mannitol solution (dashed circle). The isolation procedure yields fractions that are highly enriched in sperm cells, however some contaminating starch granules are visible (arrowheads). A close-up image (inset in **A**) of a single sperm cell reveals a large nucleus (n) and two vacuoles (arrows). **B,** DAPI staining highlights the large maize sperm cell nucleus (n). **C,** Assessment of cell viability of freshly isolated sperm cells by FDA staining. 96.6% ± 1.4% of the sperm cells show a strong fluorochromatic reaction and were classified as viable. A non-viable, FDA-negative sperm cell is marked (arrow). Scale bars in (**A**) and (**C**) = 20 µm; scale bar in (**B**) = 10 µm.

When individual sperm cells were inspected (Figure 10B), they revealed a large nucleus, comprising a majority of the cell volume, and several smaller vacuoles (Inset in Figure 10A). Sperm cell viability was evaluated directly after isolation using Fluorescein Diacetate (FDA). Due to the gentle and rapid isolation protocol, freshly isolated sperm cells showed high viability rates of 96.6 % ± 1.4 % as indicated by a strong fluorochromatic reaction (Figure 10C).

4.3 Proteomic analysis of maize sperm cells

Having developed a robust isolation protocol, maize sperm cells were employed in different experiments: To shed light on the sperm cell membrane protein composition, sperm cell microsomal protein fractions were subjected to mass spectrometric analysis ("Membrane proteome"). Furthermore, protein extracts from whole sperm cells lysates were measured in a similar way ("Total proteome"). Isolated sperm cells were also used for RNA-seq experiments to investigate the maize sperm transcriptome (published in Chen *et al.*, 2017).

For proteomics, isolated sperm cells were lysed by shock freezing in liquid nitrogen and a clearing step (10,000 xg) was employed to remove cellular debris and starch granules resulting in a cleared total protein lysate. Membrane protein-enriched microsomal fractions were prepared from cleared protein lysates by ultracentrifugation (100,000 xg). Microsomal fractions became visible as translucent yellowish pellets and were washed and resuspended in resuspension buffer by vigorous pipetting.

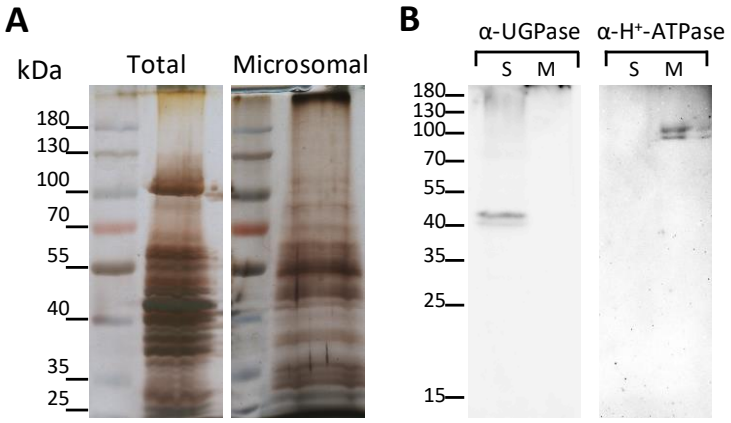


Figure 11 Total and microsomal proteins of maize sperm cells. **A,** Protein profiles of total sperm protein extract and resolubilized sperm microsomal pellet. Proteins were separated by SDS-PAGE, followed by silver staining **B,** Western blot of separated sperm soluble (S) and microsomal (M) fractions after ultracentrifugation. Antibodies directed against a cytoplasmic UDP-glucose pyrophosphorylase (UGPase; ~45 kDa) and a plasma membrane H^+-ATPase (~100 kDa) served as a marker for soluble and membrane-bound proteins, respectively.

Electrophoretic separation of microsomal and total protein fractions by SDS-PAGE followed by silver staining revealed that they can be distinguished by their distinct protein profiles (Figure 11A). The quality of membrane-enriched fractions was assessed by electrophoretic separation of supernatant and microsomal pellets, followed by western blotting and incubation with antibodies directed against a cytoplasmic (α-UGPase) and plasma membrane-bound (α-H^+-ATPase) marker protein (Figure 11B). UGPase signal

(expected MW: 51.6 kDa) appeared exclusively in the supernatant whereas H^+-ATPase signal (expected MW: 95 kDa) was detectable in the microsomal fractions only, suggesting that microsomal fractionation was successful.

Total sperm cell protein (SC) and sperm cell microsomal protein (SCM) samples were subjected to proteomic profiling by LC-MS/MS. Measurements of three biological replicates each, data analysis and visualization was performed by Dr. Julia Mergner at the Institute of Proteomics and Bioanalytics, Technical University Munich.

SC samples yielded a union of 4,799 protein groups, identified in at least one of three biological replicates. A union of 1,774 protein groups were identified in the SCM samples. The majority of protein groups identified were successfully mapped to an individual gene locus, resulting in 67.8 % (SC) and 59.9 % (SCM) of unambiguous protein groups (Figure 12A). Analysis of the overlap across the replicates revealed 3,326 SC protein groups (≙ 69.8 %) and 1,181 SCM protein groups (≙ 68 %) that were detected in all three biological replicates of the respective experiment, thereby representing the core SC and SCM proteome (Figure 12B). These core SC and SCM protein groups showed a high overlap, with 1,050 protein groups identified for both experiments which accounts for 88.9 % of all identified core SCM protein groups. 131 protein groups were exclusively identified in the SCM experiments, contrasted by 2,276 exclusive core SC protein groups (Figure 12C). To identify protein groups enriched in the SCM experiments, a label-free quantification (LFQ) approach was chosen due to the high overlap between SC and SCM protein groups and because normalization yielded comparable abundance ranges for the individual SC and SCM samples (Figure 12D). Based on this normalization, protein groups enriched in the SCM samples were calculated considering the Log_2 fold change of LFQ intensities. Significantly enriched protein groups were determined with a student´s t-test with a multiple testing-adjusted p-value of 0.05 set as cut off (Figure 12E). 148 protein groups were found to be significantly enriched in the SCM dataset.

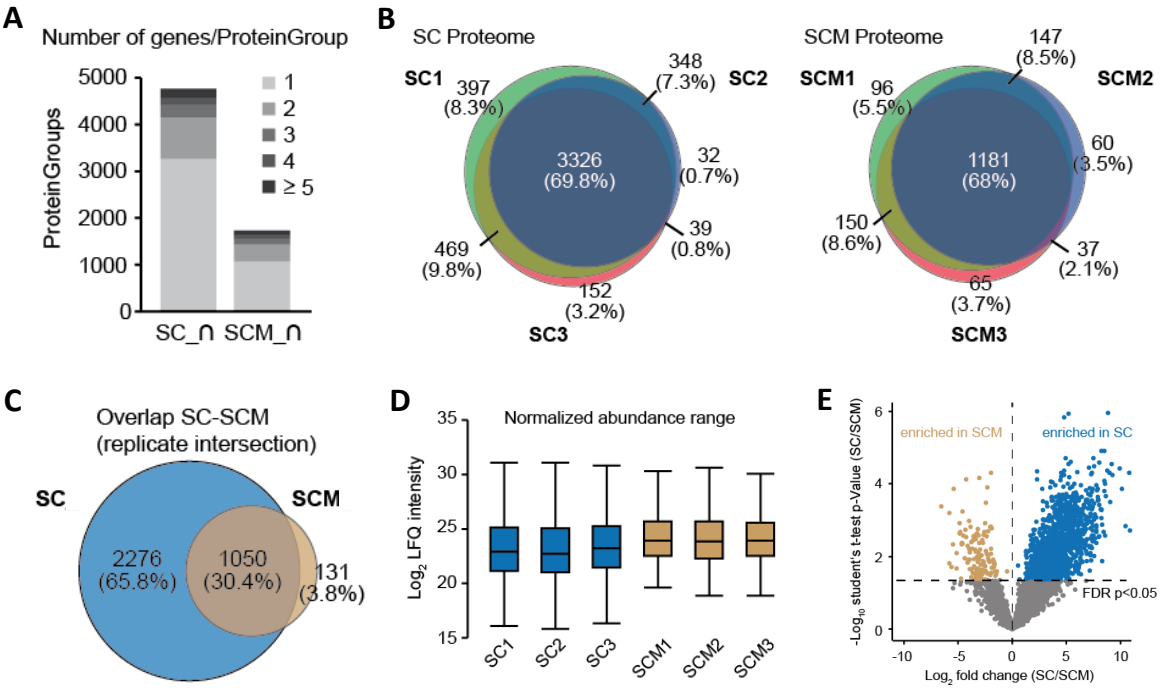


Figure 12 *In silico* analysis of sperm cell total (SC) and sperm cell membrane (SCM) proteomic datasets. A, Number of gene loci assigned per protein group. 67.8 % (SC) and 59.9 % (SCM) of all protein groups are defined by a single gene. **B,** Intersection of the three biological replicates measured for the sperm cell total proteome dataset (SC1-SC3) and the sperm cell microsomal fractions (SCM1-SCM3). A core of 3,326 protein groups (69.8 %) was identified in all three SC replicates, whereas 1,191 protein groups (68%) were detected in all three SCM replicates. **C,** Overlap between core SC and core SCM protein groups. SC and SCM show an high overlap with only 131 protein groups that were identified in the SCM dataset exclusively. **D,** Normalized label-free quantification intensity abundance ranges for SC1-3 and SCM1-3. Due to the high overlap between SC and SCM, a quantitative approach using LFQ was pursued. **E,** Enrichment analysis of SCM protein groups. The Log_2 fold change of LFQ intensities was calculated based on the normalized intensities shown in (**D**). Proteins significantly enriched in SCM fractions were determined by student´s t-test. A multiple testing-adjusted p-value of 0.05 was set as cut off. 148 protein groups were found to be significantly enriched in the SCM dataset.

The SC and SCM datasets were screened for sperm cell membrane proteins already described in the literature (Table 1). Indeed, putative maize orthologs of sperm-expressed GEX2, GEX3, HAP2 and a sperm cell-specific Aquaporin TONOPLAST INTRINSIC PROTEIN 5;1 (TIP5;1) were found. While ZmGEX2, ZmGEX3 and ZmTIP5;1 were detected in both datasets, it came as a surprise that ZmHAP2 was found only in the sperm cell total proteome. Furthermore, the sperm-specific maize Phospholipase MATRILINEAL (MTL) was also detected in both datasets.

Table 1 Previously described sperm cell membrane proteins identified in the sperm cell proteome datasets. Gene identifier (AGPv4), gene names and references are shown. The number of replicates, in which the proteins were identified in the SC and SCM experiments are indicated.

AGPv4	Name	Replicates SC	Replicates SCM	Reference
Zm00001d005781	GEX2	1 of 3	1 of 3	Engel *et al.*, 2003
Zm00001d044419	GEX3	3 of 3	3 of 3	Engel *et al.*, 2003
Zm00001d026178	TIP5;1	3 of 3	3 of 3	Wudick *et al.*, 2014
Zm00001d006177	HAP2	3 of 3	not detected	Besser *et al.*, 2006
Zm00001d029412	MTL	3 of 3	1 of 3	Kelliher *et al.*, 2017

1,774 protein groups were found in the SCM dataset in at least 1 of 3 biological replicates (Figure 12A). These consisted of 3,322 individual proteins. Due to the poor annotation of the maize genome, more than half of SCM proteins (1,736 of 3,322 proteins ≙ 52.3 %) lacked functional annotation ("Uncharacterized protein"). To gain insight on subcellular localization of SCM proteins, the prediction algorithm WoLF PSORT was applied on the SCM dataset. Furthermore, data on SCM proteins were retrieved from the Aramemnon plant protein database, and the AramTMCon consensus algorithm was employed to shed light on membrane proteins containing transmembrane α-helices. These analyses were performed by Dr. Maxim Messerer at the Institute for Plant Genome and Systems Biology, Helmholtz Centre Munich.

Prediction of subcellular localization with WoLF PSORT was successful for 2,824 proteins and their distribution is shown in Figure 13. The majority of the proteins identified in SCM were predicted to localize either to plastids (33.1 %), the cytosol (28.5 %) or the nucleus (18.4 %), making these compartments accountable for 80 % of all proteins identified in the SCM dataset. A surprisingly low number of proteins was predicted to localize to the plasma membrane (148 proteins ≙ 5.2 %) or endomembrane compartments (ER: 33 proteins ≙ 1.2 %; Golgi: 8 proteins ≙ 0.3 %). Prediction of transmembrane α-helices revealed that at least one transmembrane helix is predicted for 603 of the 3,322 SCM proteins (≙ 18.2 %). Furthermore, 21 SCM proteins were annotated in the Aramemnon database as membrane-bound by a lipid modification (e.g. GPI-anchor, Palmitoylation). In the same database, 83 SCM proteins were annotated as membrane-associated (e.g. as part of a membrane-localized multiprotein complex). In total, a surprisingly low number of 707 SCM proteins (≙ 21.3 %) were predicted as membrane-localized or membrane-associated. This suggests that despite the quality control (Figure 11) indicated otherwise, membrane protein enrichment by ultracentrifugation was not successful. This is also reflected by the high number of predicted cytosolic contaminants (28.5 % of all SCM proteins) in the SCM dataset.

In a quantitative approach, 148 protein groups were found to be significantly enriched in the SCM dataset (Figure 12E). Among those were the 131 protein groups which were exclusively found in the SCM dataset (Figure 12C). The 148 SCM-enriched protein groups consisted of 196 individual proteins. Those were considered as high priority candidates for functional studies, however prediction of subcellular localization (Figure 13) revealed that the majority of SCM-enriched proteins are predicted to localize to the nucleus (≙ 45.1 %), the cytosol (≙ 23.6) or to plastids (≙ 18.5 %). Only five proteins (≙ 0.26 %) were predicted to localize the plasma membrane. 7 proteins were predicted to localize to endomembrane compartments (ER: 4 proteins ≙ 0.21 %; Golgi: 3 proteins ≙ 0.15 %). Analysis in the Aramemnon Database revealed that only 25 of the 196 SCM-enriched proteins (≙ 6.1 %) contain at least one transmembrane α-helix. Notably, none of the sperm cell-expressed membrane proteins described in the literature (Table 1) were found among the 196 SCM-enriched proteins.

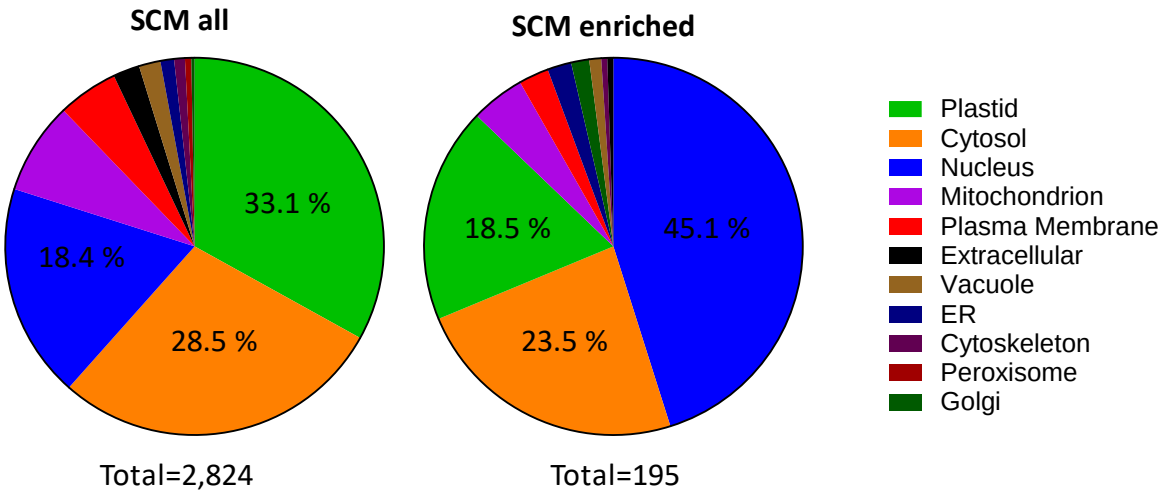


Figure 13 Predicted subcellular localization for all SCM protein (SCM all) and for proteins significantly enriched in the SCM dataset (SCM enriched). Subcellular localization was predicted with WoLF PSORT and is shown as parts of a whole. Prediction was successful for 2,824 SCM proteins. These were predicted to localize to following compartments: 934 plastidal proteins (≙ 33.1 %), 804 cytosolic proteins (≙ 28.5 %), 519 nuclear proteins (≙ 18.4 %), 223 mitochondrial proteins (≙ 7.9 %), 146 plasma membrane-localized proteins (≙ 5.2 %), 64 extracellular proteins (≙ 2.3 %), 52 vacuolar proteins (≙ 1.8 %), 33 ER-localized proteins (≙ 1.2 %), 26 cytoskeleton proteins (≙ 0.9 %), 15 peroxisomal proteins (≙ 0.5 %) and 8 Golgi-localized proteins (≙ 0.3 %). Among the 195 significantly enriched SCM proteins, 88 nuclear proteins (≙ 45.1 %), 46 cytosolic proteins (≙ 23.5 %), 36 plastidal proteins (≙ 18.5 %), 9 mitochondrial proteins (≙ 4.6 %), 5 plasma membrane-localized proteins (≙ 2.6 %), 4 ER-localized proteins (≙ 2.1 %), 3 Golgi-localized proteins (≙ 1.5 %), 2 vacuolar proteins (≙ 1 %), 1 cytoskeleton protein (≙ 0.5 %) and 1 extracellular protein (≙ 0.5 %) were predicted. Prediction with WoLF PSORT was performed by Dr. Maxim Messerer.

4.4 Candidate protein identification and selection

Although enrichment analysis did not yield satisfying results, several published sperm membrane proteins were identified in the SC and SCM datasets respectively (Table 1), suggesting that the datasets as a whole can be considered as suitable for candidate identification. The whole dataset was therefore screened for promising proteins or protein families with transmembrane domains and for membrane-associated proteins. Selected candidate proteins (Table 2) for functional studies included two Receptor-like Kinases (RLKs) and proteins containing following conserved domains: LUNG SEVEN TRANSMEMBRANE RECEPTOR (LUSTR), PLACENTA-SPECIFIC GENE 8 (PLAC8), BAX INHIBITOR 1-RELATED (BI-1), SOLUBLE N-ETHYLMALEIMIDE-SENSITIVE FACTOR ATTACHMENT PROTEIN 25 (SNAP25) and TB2/DP1/HVA22-RELATED (HVA22).

Table 2 Candidate proteins identified in the proteomic dataset. Maize gene identifiers (AGPv4), their presence in the maize sperm proteomic datasets (SC = total proteome; SCM = membrane proteome) and putative *Arabidopsis thaliana* (*A.th*) orthologs are indicated.

AGPv4	Maize Description	Maize SC Proteome	Putative *A.th* ortholog	
			AGI	Name
Zm00001d023459	HVA22-like protein	SC, SCM	At2g36020	*HVA22J*
Zm00001d034431	Bax inhibitor-1 family protein	SC, SCM	At1g03070	*LFG4*
Zm00001d003079	Protein GPR107	SC, SCM	At3g09570	*LUSTR*
Zm00001d048551	Cell number regulator 6 (PLAC8)	SC, SCM	At1g68610	*PCR11*
Zm00001d050074	Putative leucine-rich repeat receptor-like protein kinase family protein	SC, SCM	At1g73080	*PEPR1*
Zm00001d035992	Putative receptor-like protein kinase family protein	SC, SCM	At5g59650	*RLK1*
Zm00001d015414	Putative receptor-like protein kinase family protein	SC, SCM	At4g20790	*RLK2*
Zm00001d016686	SNAP25 homologous protein SNAP33	SC, SCM	At1g13890	*SNAP30*

Functional studies on putative orthologs were performed in *Arabidopsis thaliana* as the *Arabidopsis* generation time is shorter*,* generation of transgenic plants works with much higher efficiency and a plethora of marker lines and protocols to dissect the process of double fertilization on a cellular level are available.

Identification of true orthologs between maize and *Arabidopsis thaliana* was facilitated by publicly deposited *Arabidopsis* transcriptome data in the Genevestigator software: the expression pattern of *Arabidopsis thaliana* gene family members containing the domains indicated above were screened and genes showing strong and/or specific expression in sperm cells were chosen as candidates (Table 2).

In a second approach, *in silico* co-expression analysis in *Arabidopsis thaliana* was performed with sperm-specific *GEX2* in the Genevestigator software, using the “anatomy” dataset containing microarray data from 105 *Arabidopsis thaliana* tissues including sperm

cells. Gene products were subsequently analyzed for presence of transmembrane alpha helices using the Aramemnon plant membrane protein database and subcellular localization was predicted using the SUBA4 online tool in order to identify transmembrane, membrane-anchored or secreted proteins (Supplemental table 2). Among the Top 20 *GEX2* co-expressed genes were genes encoding published sperm cell-expressed membrane proteins such as HAP2 (Mori *et al.*, 2006), GEX1 (Alandete-Saez *et al.*, 2011) and TET11 (Boavida *et al.*, 2013). From this list the genes encoding CYSTEINE-RICH RECEPTOR-LIKE PROTEIN KINASE 43 (CRK43), PLANT CADMIUM RESISTANCE 11 (PCR11) and PROTEIN OF UNKNOWN FUNCTION DUF679 (DAU2/DMP9) were chosen as candidates for localization and functional studies. Notably, a putative PCR11 ortholog was also identified in the SCM dataset (Zm00001d048551).

A comprehensive summary of candidate genes selected for functional studies in *Arabidopsis thaliana* can be found in Table 3.

Table 3 Candidate genes selected for expression and loss-of-function studies in *Arabidopsis*. Gene identifier (AGI), name, promoter length in base pairs (bp) upstream of the start codon and oligo pairs used to clone translational reporter constructs are indicated. Known DUO1 target genes (Borg *et al.* 2011) are indicated. Subcellular localization as predicted by the SUBAcon algorithm (Hooper *et al.* 2014) is shown (PM = plasma membrane). The presence for N-terminal signal peptides was evaluated with SignalP 4.0 (Petersen *et al.* 2011). The number of transmembrane alpha helices predicted by the AramTMcon algorithm (Schwacke *et al.* 2003) is indicated. Protein subcellular localization as observed in stable *Arabidopsis thaliana* translational reporter lines and expression in male and female reproductive tissues is shown (syn = synergid; pt = pollen tube; sc = sperm cell). Candidates labeled with ([a]) were selected from maize sperm proteome data. Candidates labeled with ([b]) were selected from *GEX2* co-expression analysis in *Arabidopsis thaliana*. *DMP8* and *DMP10* were selected assuming functional redundancies with *DMP9*.

AGI	Name	Promoter length	Cloning primers	DUO1 Target?	SUBAcon prediction	Signal peptide	TM alpha helices	Subcellular localization	Expression male	Expression female
At4g28670	*CRK43*[b]	452 bp	#207/#209	no	PM	Yes	1	not analyzed	not detected	not detected
At1g09157	*DMP8*	374 bp	#337/#338	no	PM	No	4	PM	sc	not detected
At5g39650	*DMP9*[b]	1,526 bp	#223/#224	yes	PM	No	4	PM	sc	not detected
At5g27370	*DMP10*	410 bp	#339/#340	no	PM	No	4	not analyzed	not detected	not detected
At2g36020	*HVA22J*[a]	2,095 bp	#290/#291	no	secreted	No	2	endomembrane	pt; not in sc	not analyzed
At1g03070	*LFG4*[a]	639 bp	#228/#229	no	vacuole	No	7	PM/endomembrane	sc	not analyzed
At3g09570	*LUSTR*[a]	1,067 bp	#142/#124	no	PM	Yes	7	endomembrane	pt; not in sc	syn
At1g68610	*PCR11*[ab]	5,814 bp	#251/#252	yes	PM	No	1	PM/endomembrane	sc	not analyzed
At1g73080	*PEPR1*[a]	797 bp	#125/#127	no	PM	Yes	1	PM	not detected	Integuments
At5g59650	*RLK1*[a]	444 bp	#270/#271	no	PM	Yes	1	not analyzed	not detected	not analyzed
At4g20790	*RLK2*[a]	3,177 bp	#288/#289	no	PM	Yes	1	PM	sc	not analyzed
At1g13890	*SNAP30*[a]	368 bp	#286/#287	no	nucleus	No	0	not analyzed	pt; not in sc	not analyzed

4.5 Expression and loss-of-function studies

Twelve *Arabidopsis thaliana* candidate genes were selected to investigate their expression in male and female reproductive tissues and to verify their subcellular localization (see Table 3). Based on these results, candidates were selected for loss-of-function studies by generating CRISPR/Cas9 genome-edited mutants.

4.5.1 Translational reporters

Stable transgenic lines expressing the respective candidate as translational C-terminal GFP fusion under control of the endogenous promoter were generated and candidate expression and subcellular localization was analyzed in male (pollen, *in vitro* germinated pollen tubes) and female (ovules) reproductive tissues via CLSM (Figure 14). The results for *DMP8, DMP9,* and *DMP10* will be presented separately in chapter 4.6.

No signal was detected in reporter lines expressing *DMP10g:GFP, CRK43g:GFP* or *RLK1g:GFP* in any of the analyzed tissues (not shown). *PEPR1g:GFP* was absent in pollen and pollen tubes (not shown) but a plasma membrane-localized signal was detected in unfertilized ovules, where signals were observed in the integuments surrounding the embryo sac and became stronger towards the micropylar pole of the ovule (Supplemental figure 1). Ovules expressing *LUSTRg:GFP* showed signals within the synergid cells, accumulating as dotted structures at the filiform apparatus (Supplemental figure 1).

HVA22J-GFP, LUSTR-GFP and SNAP30-GFP signals were detected in pollen grains and in germinated pollen tubes. However, the GFP signal was excluded from sperm cells and their outline became visible as dark background (Figure 14, hollow arrowheads), suggesting that these three reporter constructs are expressed in the vegetative cell of the pollen grain but not in sperm cells.

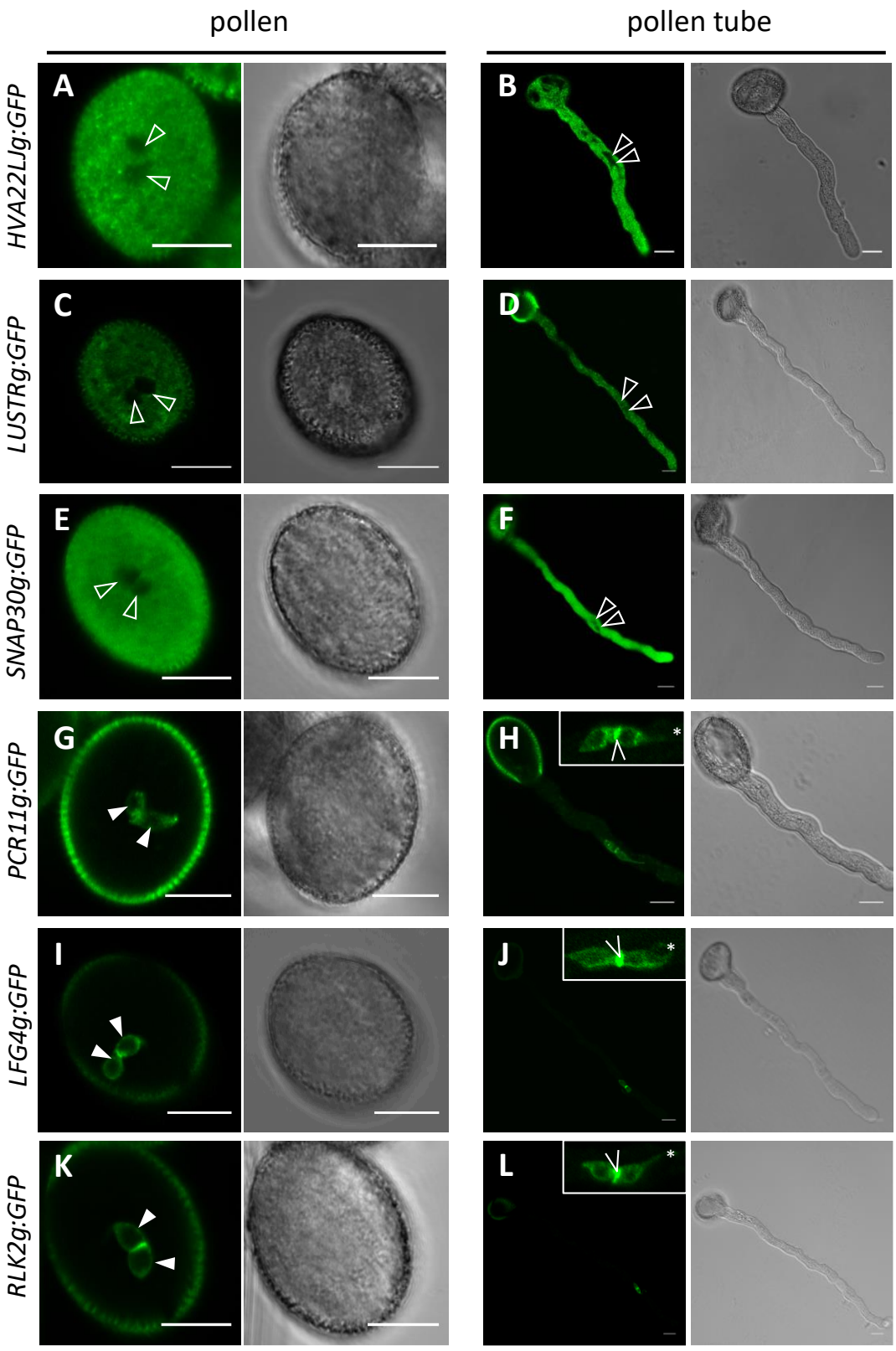


Figure 14 Localization of candidate GFP-fusion proteins in *Arabidopsis* pollen and pollen tubes. HVA22J-GFP (**A, B**), LUSTR-GFP (**C, D**) and SNAP30-GFP (**E, F**) were detected in the vegetative cell of pollen grains and tubes, while being excluded from sperm cells (hollow arrowheads). PRC11-GFP (**G, H**), LFG4-GFP (**I, J**), and RLK2-GFP (**K, L**) showed signal in sperm cells (arrows). Insets show signals at the sperm plasma membrane, including the membranous extension associated with the vegetative cell nucleus (asterisks). Strong signal at the connection between the sperm pair (open arrowheads) became apparent in pollen tubes. Scale bars = 10 µm.

Five candidates (*PCR11*, *LFG4, RLK2, DMP8* and *DMP9*) showed a sperm cell-specific translational reporter activity (Figure 14, Figure 21). In the pollen grain, the GFP signal appeared at the surface of the two sperm cells. In *PCR11g:GFP*-expressing lines, the fusion protein was detected in the endomembranes but also decorated the sperm periphery (Figure 14G,H). Fluorescence was diffuse and granular, resembling the subcellular localization observed for translational HAP2 reporters (Besser *et al.*, 2006; Mori *et al.*, 2006; Mori *et al.*, 2014). RLK2-GFP signal was observed exclusively at the sperm plasma membrane as expected for a Receptor-like Kinase (Figure 14K,L). Despite being predicted as a vacuolar protein (Table 3), subcellular localization of *LFG4g:GFP* was similar to that of RLK2-GFP, suggesting plasma membrane localization (Figure 14 I,J).

Imaging of sperm cells in *in vitro* germinated pollen tubes can be advantageous as the male germ unit is transported in a stretched-out conformation. CLSM analysis of the five sperm expressed candidates in pollen tubes confirmed a subcellular localization of GFP fusion proteins at the periphery of the two sperm cells. Signal was also observed at the plasma membrane projection that tethers one sperm cell to the vegetative nucleus (Figure 14, asterisks). Within the pollen tube, it became apparent that all sperm-expressed candidates accumulated GFP signal in a membrane subdomain at the interface contact between the sperm pair (Figure 14, open arrowheads). Imaging in pollen tubes also revealed that all sperm-expressed candidates were absent from the pollen tube plasma membrane, however *LFG4g:GFP* expressing lines showed weak GFP signal within the pollen tube cytoplasm.

Despite containing 2 transmembrane α-helices, SUBAcon predicted HVA22J as secreted protein (Table 3). Accordingly, *HVA22Jg:GFP* signal was detected in the cytoplasm of pollen and pollen tubes (Figure 14A,B). While being predicted as a plasma membrane-localized protein, LUSTR-GFP signal was detected within pollen and pollen tubes but was absent from the pollen tube plasma membrane (Figure 14C,D). Signal in *SNAP30g:GFP*-expressing lines was distributed uniformly within the cytoplasm of the pollen grain and pollen tube, although SUBAcon predicted localization to the nucleus (Table 3; Figure 14E,F). The uniform intracellular distribution of GFP signal for these candidates suggested cleavage of the reporter during posttranslational processing, physiological degradation or protein turnover. To investigate this scenario, protein extracts from *HVA22Jg:GFP*, *LUSTRg:GFP* and *SNAP30g:GFP* expressing lines were prepared from mixed inflorescence tissues followed by electrophoretic separation, western blotting and probing with an α-GFP antibody (Figure 15).

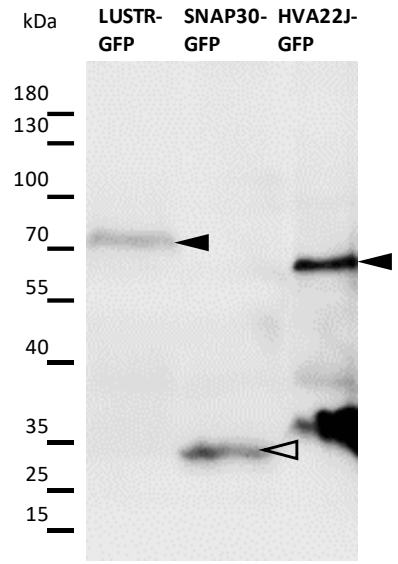


Figure 15 Western blot of protein extracts from selected candidate marker lines. Inflorescence protein extracts from LUSTR-GFP (78.7 kDa) and HVA22J-GFP (58.2 kDa)-expressing plant lines showed signal at the expected molecular weight for the fusion proteins (arrowheads). SNAP30-GFP (57.8 kDa) was not detected, but a signal around the molecular weight of free GFP became visible (hollow arrowhead).

Western blotting confirmed that the signals observed in *HVA22Jg:GFP* and *LUSTRg:GFP* expressing lines can be attributed to full length fusion proteins as apparent signals matched with the calculated molecular weight (HVA22J-GFP: 58.2 kDa; LUSTR-GFP: 78.7 kDa). However, a cleaved product was additionally observed for *HVA22Jg:GFP,* as indicated by a strong signal around 35 kDa. Protein extracts from a *SNAP30g:GFP*-expressing line lacked the expected signal (SNAP30-GFP: 57.8 kDa). Nevertheless, a prominent signal below 35 kDa became visible, corresponding the size of free GFP (GFP: 28.7 kDa). This suggests that the *SNAP30* promoter is active in pollen, however the cytosolic signal observed in *SNAP30g-GFP* expressing lines does not reflect this candidate´s subcellular localization.

In summary, the translational reporter studies in *Arabidopsis* reproductive tissues showed that expression of five candidates was sperm cell-specific. The predicted plasma membrane localization was confirmed for four candidates. LFG4 was predicted to localize to the vacuole but the fusion protein was observed at the sperm cell periphery. Furthermore, all sperm-expressed candidates showed signal at the plasma membrane projection connecting one sperm cell to the vegetative nucleus. Sperm-specific expression and surface localization made those five genes prime candidates for subsequent functional studies.

4.5.2 Loss-of-function studies

To gain insights on putative protein functions in sperm cells, a reverse genetics approach was chosen and stable gene-edited *Arabidopsis thaliana* lines were generated in the Columbia-0 (*Col-0*) background and in the sperm cell nuclear marker line *HTR10p:HTR10-mRFP* using the CRISPR/Cas9 system. Genome editing events in the CRISPR/Cas9 mutants generated in this thesis are summarized in Supplemental table 1. Single guide RNAs (sgRNAs) were designed to target the single exon of *PCR11*, the first exon of *RLK2* or the first exon of *LFG4* (Figure 16A). Transgenic progeny was screened for genome editing events by PCR-amplifying genomic regions surrounding the sgRNA binding sites, followed

by sequencing and assembly of the obtained chromatograms to the wild type reference genomic sequence (Figure 16B).

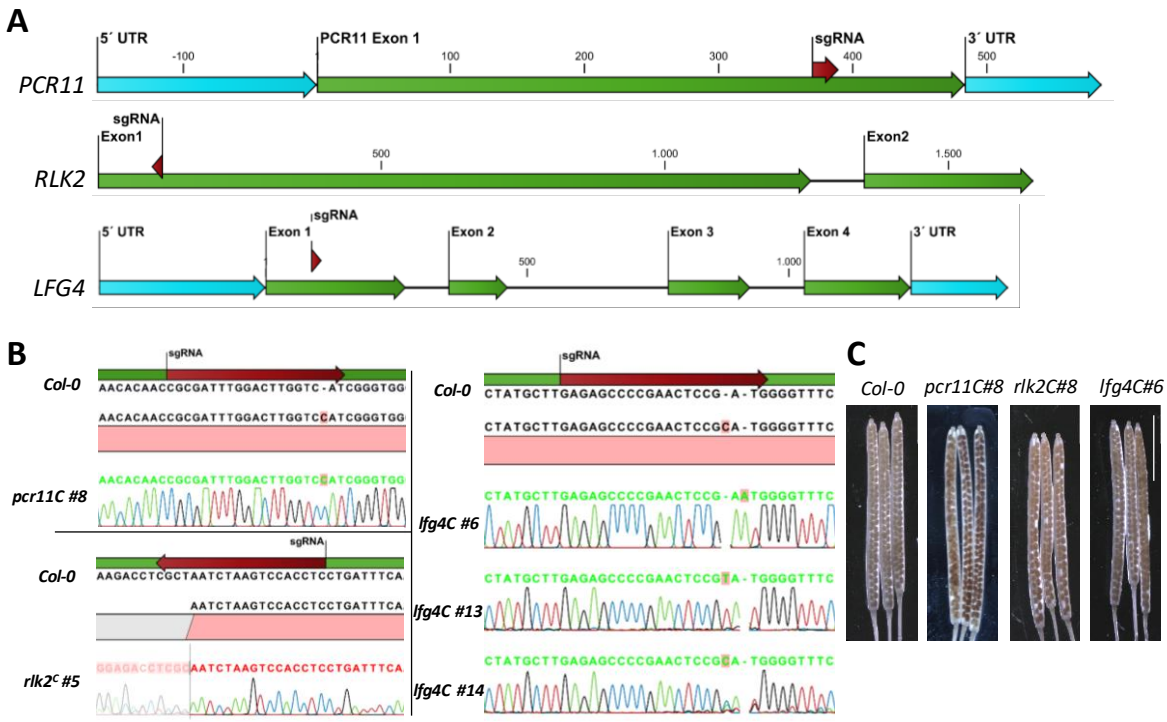


Figure 16 Generation of CRISPR/Cas9 genome edited mutants for *PCR11, RLK2* and *LFG4*. A, Genomic structures of *PLANT CADMIUM RESISTANCE 11* (*PCR11*), *RECEPTOR-LIKE KINASE 2* (*RLK2*) and *LIFEGUARD 4* (*LFG4*). sgRNA binding sites are indicated. Numbers are base pairs relative to the start codon. Note that no UTRs were annotated for *RLK2*. **B,** Genome editing events were identified by PCR amplifying regions surrounding the sgRNA binding site followed by sequencing and assembling to the reference genomic sequence (*Col-0*). Single base insertions for PCR11C and LFG4C are highlighted in red. **C,** Siliques from genome edited mutants were cleared and analyzed for unfertilized ovules and aborted seeds but full seed set comparable to the wild type (*Col-0*) was observed for all analyzed T1 lines. Scale bar in (**C**) = 5 mm.

For *PCR11* 20 independent T1 lines were analyzed and only one line was found mutated in the target locus (*pcr11C#8*) with a homozygous +C insertion at position 387 after the start codon. The resulting shift in the translational frame leads to formation of a premature stop codon and a truncated protein of 157 amino acids. Of 12 independent T1 lines screened for *RLK2*, only one genome edited mutant was identified (*rlk2C*#5). The exact nature of this trans-heterozygous mutation was impossible to decipher from the superimposed chromatogram. However, both *RLK2* alleles are edited and a heterozygous mutation can be excluded. *LFG4* was successfully mutagenized and screening of 15 independent T1 lines resulted in three mutants (*lfg4C#6, #13, #14*), carrying frameshift-inducing single base insertions with premature stop codons.

To assess fertilization defects in the different genome edited lines, seed set was chosen as a first read out. Almost mature siliques were cleared and screened for unfertilized ovules or aborted seeds. However, none of the genome edited lines showed reduced seed set

(Figure 16C) and the candidates were not analyzed any further. The phenotype of *DMP8/9/10* loss-of-function mutants will be presented in the following chapter.

4.6 Discovery of sperm-specific DMPs

DUF679 DOMAIN MEMBRANE PROTEIN 9 (*DMP9,* Kasaras and Kunze, 2010) was identified as co-expressed with *GEX2* in an *in-silico* screen using publicly available *Arabidopsis thaliana* microarray data (see Supplemental table 2). *DMP9* is also known as *DUO1 ACTIVATED UNKNOWN* (*DAU2*) and a direct target gene of the male germline-specific MYB transcription factor DUO1 (see chapter 3.3), a core transcriptional regulator of the male germline-specific differentiation program in *Arabidopsis* (Borg *et al.*, 2011).

4.6.1 The conserved DUF679: domain architecture

Domain of unknown function (DUF) 679 domain-containing membrane proteins (DMPs) are small integral membrane proteins with 3-5 predicted transmembrane alpha helices. They are characterized by a conserved protein domain of unknown function annotated as DUF679 (Pfam PF05078). DMP protein sequences from selected species were retrieved from UniProt and the HMMTOP server was used to predict transmembrane alpha helices. Representative protein structures of DUF679 proteins from selected species are shown in Figure 17.

The conserved DUF679 core comprises the transmembrane helices. Four transmembrane α-helices are predicted for most DMPs, and the N- and C-termini of AtDMP9 are predicted to face the cytosol. Notably, the N- and C-terminal parts of DMPs from different species are quite diverse and differ considerably in amino acid sequence and length. Distinguished examples are DMPs from the lycophyte *Selaginella moellendorffii* which are considerably shorter, and the single *DMP* copy found in the genomes of *Volvox carteri* and *Chlamydomonas reinhardtii*, characterized by an extended C-terminus of > 100 amino acids (Figure 17). While four transmembrane α-helices are predicted for most DMPs by HMMTOP, Analysis of DMP protein sequences from lycophytes and algae revealed DMPs with three (*Selaginella moellendorffii, Volvox carteri*) or five (*Chlamydomonas reinhardtii)* predicted transmembrane helices (Figure 17). An odd number of transmembrane helices suggests that the termini in those DMPs face opposing sides of the membrane.

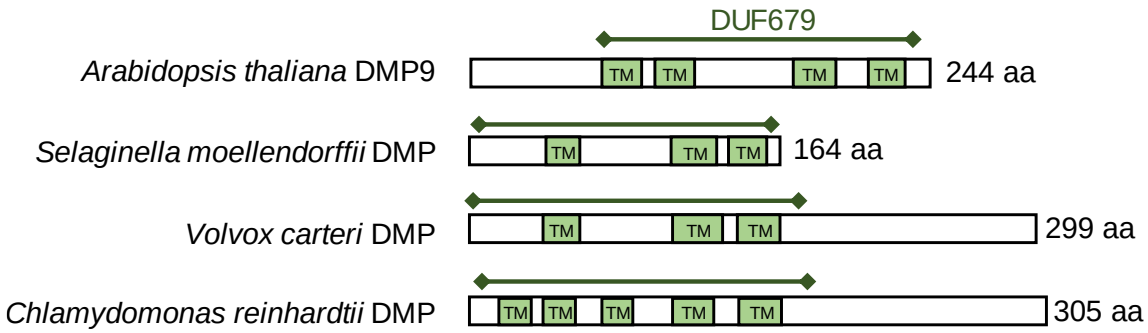


Figure 17 Protein structures of representative DMPs from selected species. Proteins are drawn to scale. Transmembrane α-helices as predicted by HMMTOP are indicated as green boxes. The conserved DUF679 domain as predicted by Pfam (PF05078) is indicated above the schemes. *Arabidopsis thaliana* DMP9 (AT5G39650) is a short protein (244 aa) characterized by 4 transmembrane alpha helices and an extended amino-terminus, whereas *Selaginella moellendorffii* DMP (SELMODRAFT_68639) is considerably shorter (164 aa) and only 3 transmembrane domains are predicted. HMMTOP predicts 3 transmembrane helices for *Volvox carteri* DMP (VOLCADRAFT_95012) and 5 transmembrane helices for *Chlamydomonas reinhardtii* DMP (Cre07.g312850). Both algal DMPs are characterized by an extended C-terminus. aa = amino acids, TM = transmembrane α-helix.

4.6.2 Phylogeny

Analysis of species distribution in the Pfam database revealed that the genomes of archaea, prokaryotes and animals do not contain *DMP* coding genes. However, DMP orthologs can be found in all clades of the *Viridiplantae* ranging from the unicellular algae *Chlamydomonas reinhardtii* to angiosperms. DUF679-containing proteins were also identified in the mushroom-forming fungi *Sphaerolobus stellatus* and *Exidia glandulosa* but these were considerably longer (439 to 462 amino acids) and differed from the green plant clade DMPs. To gain a deeper insight about evolutionary relationships among DMPs, a phylogenic tree was constructed with full length DUF679-containing protein sequences from selected species retrieved from InterPro (Figure 18). The tree covered the whole evolutionary history of DUF679 ranging from unicellular and multicellular algae (*Chlamydomonas reinhardtii*, *Volvox carteri*) and bryophytes (*Physcomitrella patens*) to gymnosperms (*Araucaria cunninghamii*) basal angiosperms (*Amborella trichopoda*) monocots (*Zea mays*) and dicots (*Solanum, Arabidopsis thaliana)*. The longer DUF679-containing proteins from fungi served as outgroup. Phylogenic reconstruction revealed that *Arabidopsis thaliana* DMP8 and DMP9 form a subclade with DMP10 (Figure 18).

To obtain insights about species distribution, the number of DUF679 containing sequences for selected species was retrieved from the Pfam database (Table 4). While the genomes of green algae and mosses and gymnosperms contain 1-2 DMPs, the number increases in angiosperms and varies from 8 (*Ananas comosus*) to 19 (*Oryza sativa subsp. indica*) DMPs in monocots and 4 (*Genlisea aurea*) to 26 (*Brassica napus*) DMPs in dicots.

Table 4 Species distribution of DMPs. The number of DMPs was retrieved from Uniprot and shows that the DUF679 expanded specifically in the flowering plant lineage.

Species	Number of DMPs	Species	Number of DMPs
Chlamydomonas reinhardtii	1	*Zea mays*	14
Physcomitrella patens	2	*Oryza sativa subsp. indica*	19
Araucaria cunninghammii	1	*Solanum lycopersicum*	7
Amborella trichopoda	5	*Arabidopsis thaliana*	10

This demonstrates that despite being an ancient protein family, angiosperms expanded their DUF679 protein repertoire. The almost exclusive and ubiquitous occurrence of DUF679 in the *Viridiplantae* indicates a functional role for DMPs in plant-specific processes.

4.6.3 DMPs in *Arabidopsis thaliana*

The *Arabidopsis thaliana* genome encodes 10 DMP proteins (AtDMPs). The multiple sequence alignment constructed with AtDMP protein sequences (Figure 19) reveals that the conserved DUF679 comprises the majority of the protein potions including transmembrane helices and inter-transmembrane helix loops. However, the N-termini of AtDMPs differ considerably in length (12 - 69 amino acids) and in their amino acid sequence. The AramTmCon consensus algorithm predicts 4 transmembrane spanning alpha helices for *Arabidopsis thaliana* DMPs resulting in a protein topology resembling Tetraspanins, with amino- and carboxy-termini facing the cytosol and a cytoplasmic loop 2. DMP8 and DMP9 have the longest N-terminus (69 amino acids) and differ from other *Arabidopsis thaliana* DMPs by a slightly elongated loop 2. It should be noted that the extended loop 2 of DMP8 and DMP9 comprises additional 10 amino acids (“IAEKEIPVTD”) showing 100 % conservation between DMP8 and DMP9. While DMP10 also shows a slightly elongated loop 2, this addition is absent from all other AtDMPs, thereby presenting an exclusive feature of DMP8 and DMP9.

The SUBAcon algorithm was used to predict the subcellular localization for all 10 *Arabidopsis thaliana* DMPs. The majority of AtDMPs was predicted to localize either to the plasma membrane exclusively (DMP4, 5, 8-10), or to the plasma membrane and the vacuole (DMP1, 2). DMP3 and DMP7 were predicted to localize to plastids and DMP6 exclusively to the vacuole. Analysis of AtDMP protein sequences using SignalP 4.0 signal peptide prediction server indicates that AtDMP N-termini do not contain a signal peptide for secretion.

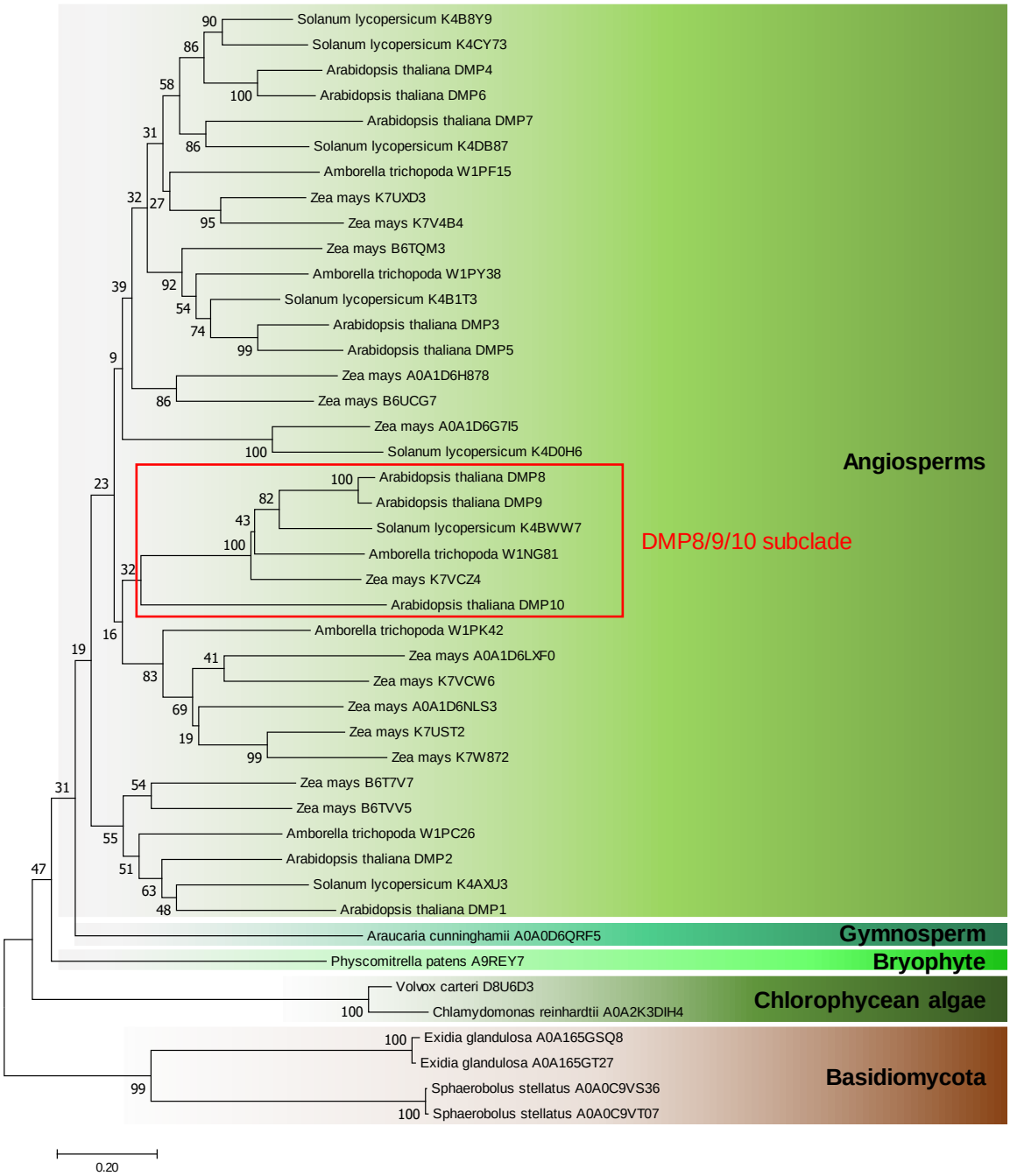


Figure 18 Phylogenetic relationship of selected DUF679-containing proteins. The Neighbor joining phylogenetic tree with 1000 bootstrap replications was constructed using selected full-length DUF679-containing protein sequences from green algae, bryophytes (here: *Physcomitrella*), gymnosperms (here: hoop pine), basal angiosperms (here: *Amborella trichopoda*), and monocot and dicot species (maize, tomato, *Arabidopsis*). Sequences were retrieved from InterPro (PF05078), truncated sequences were not included. The four longer protein sequences (439 to 462 amino acids) from mushroom-forming fungi (subphylum *Agaricomycotina*) with annotated DUF679 served as an outgroup. Bootstrap values are shown next to the branches. Evolutionary distances were computed using the Poisson correction method and are in the units of the number of amino acid substitutions per site. All positions containing gaps and missing data were eliminated. Branch labels indicate species and InterPro accession number. Nomenclature of *Arabidopsis* DMPs according to Kasaras and Kunze, 2010. The DMP8/9/10 subclade is highlighted in red. Modified from Cyprys *et al.*, 2019.

Figure 19 Alignment of the ten *Arabidopsis* DMPs. Four predicted transmembrane α-helices (AramTmCon consensus prediction) are boxed in red. The DUF679 region is labeled by a black line. Note that DMP8 (243 aa) and DMP9 (244 aa) have longer N-termini compared to all other DMP proteins and that DMP8, DMP9 and DMP10 possess a longer loop 2. Amino acid colors indicate residue properties (red: negative charge; blue: positive charge; green: polar; black: hydrophobic). Amino- and carboxy-termini are predicted to be cytosolic (blue dashed lines), suggesting a cytosolic loop 2. TMD = transmembrane domain. Modified from Cyprys *et al.*, 2019.

Phylogenetic reconstruction of evolutionary relationships among AtDMPs revealed that they form 4 clades (Figure 20A): DMP1/2, DMP3/5, DMP4/6/7 and DMP8/9/10. DMP9 is closely related to DMP8 and both share 90.2 % identity (220/244 amino acids identical) suggesting they resulted from gene duplication in the *Arabidopsis thaliana* genome. DMP8 and DMP9 form a subclade with DMP10. Previous promoter-GUS studies across various tissues and developmental stages reported *DMP9* promoter activity in pollen, whereas *DMP8* and *DMP10* promoters were not active in any of the tested plant organs (Kasaras and Kunze, 2010). Screening of publicly available microarray data in Genevestigator revealed that only *DMP1*, *DMP2*, *DMP4* and *DMP9* were present on the *Arabidopsis* genome ATH1 array and that *DMP9* showed strong expression in stamen, pollen and sperm cells (Supplemental figure 2).

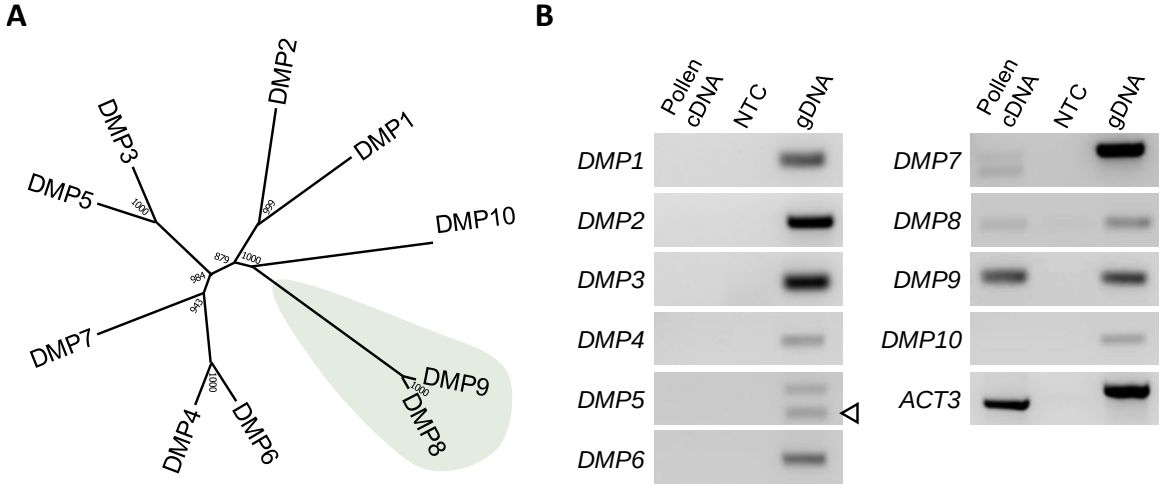


Figure 20 Phylogeny and expression of *Arabidopsis thaliana DMPs* in pollen. **A,** Unrooted UPGMA tree based on amino acid similarities among *Arabidopsis* DMP proteins shows 4 distinct clades. Bootstrap values (1.000 trials) are shown at branch points. Sperm-specific DMP8 and DMP9 are highlighted in green. **B,** Transcript presence of *DMP7, DMP8* and *DMP9* in pollen was demonstrated by RT-PCR. Genomic DNA (gDNA) and an intron-flanking actin primer (*ACT3*) were used as controls. The hollow arrowhead indicates the expected amplicon size for *DMP5*. Modified from Cyprys *et al.*, 2019.

Previous semi-quantitative RT-PCR expression analysis of *AtDMPs* in vegetative tissues and flowers showed that *DMP7* is ubiquitously expressed in all analyzed tissues while *DMP8-10* expression is restricted to flowers (Kasaras and Kunze, 2010). *Arabidopsis* flowers are a complex tissue containing male and female gametophytes but also several sporophytic cell types. To further dissect expression of all 10 *AtDMPs* in male reproductive tissues, *Arabidopsis* pollen was isolated in bulk using a modified vacuum cleaner (see 6.3.13), and pollen-derived cDNA was employed in RT–PCR experiments (Figure 20B). Only *DMP7*, *DMP8* and *DMP9* were found to be transcribed in pollen. Two amplicons were observed for *DMP7*, the only *Arabidopsis thaliana DMP* containing two short introns. The absence

of genomic DNA contamination in the Actin control suggests the presence of spliced and unspliced *DMP7* transcript in pollen.

While promoter:GUS studies suggested *DMP9* promoter activity specifically in pollen (Kasaras and Kunze, 2010), this method lacks spatial resolution due to the small size of *Arabidopsis thaliana* sperm cells and diffusion of the enzymatic product between cells and tissues after histochemical staining. Therefore, stable transgenic *Arabidopsis* lines expressing *DMP7, DMP8, DMP9* and *DMP10* as C-terminal GFP fusion proteins under control of their endogenous promoters were generated, and expression and subcellular localization was investigated in pollen and *in vitro* germinated pollen tubes using CLSM (Figure 21).

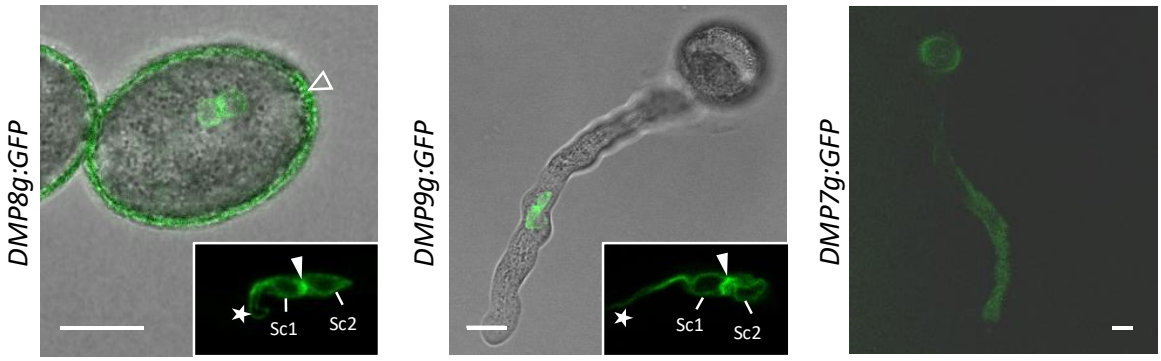


Figure 21 Subcellular localization of DMP-GFP fusion proteins in *Arabidopsis* pollen and pollen tubes. DMP8-GFP and DMP9-GFP fusion proteins are detectable in sperm cells. Close up images of sperm cells in *in vitro* germinated pollen tubes (insets) show that DMP8-GFP and DMP9-GFP localize to the sperm plasma membrane. Signal is also visible at the membranous extension associated with the vegetative nucleus (asterisks). Strong signal at the connection between the sperm pair (arrowheads) became apparent in pollen tubes. The pollen grain exine shows unspecific autofluorescence (hollow arrowhead) that is also present in wild-type pollen. Weak DMP7-GFP signal can be observed in the cytoplasm of *in vitro* germinated pollen tube but was not detected in sperm cells. Scale bars = 10 µm. Modified from Cyprys *et al.*, 2019.

DMP8-GFP and DMP9-GFP signal was detected specifically in the sperm cell plasma membrane including the membranous extension associated with the vegetative nucleus. Signal accumulation was observed at the connection between the two sperm cells. DMP7-GFP signal was absent from pollen (not shown) but weak signal in the pollen tube cytoplasm became visible in *in vitro* germinated pollen tubes. However, DMP7-GFP signal was excluded from sperm cells. DMP10-GFP was not detectable in pollen, pollen tubes or sperm cells (not shown).

4.6.4 *dmp* loss of function mutants display a reduced seed set

Strong expression in sperm cells and protein localization to the sperm plasma membrane made *DMP8* and *DMP9* prime candidates for revere genetics studies to analyze putative functions of these proteins during double fertilization in *Arabidopsis thaliana*.

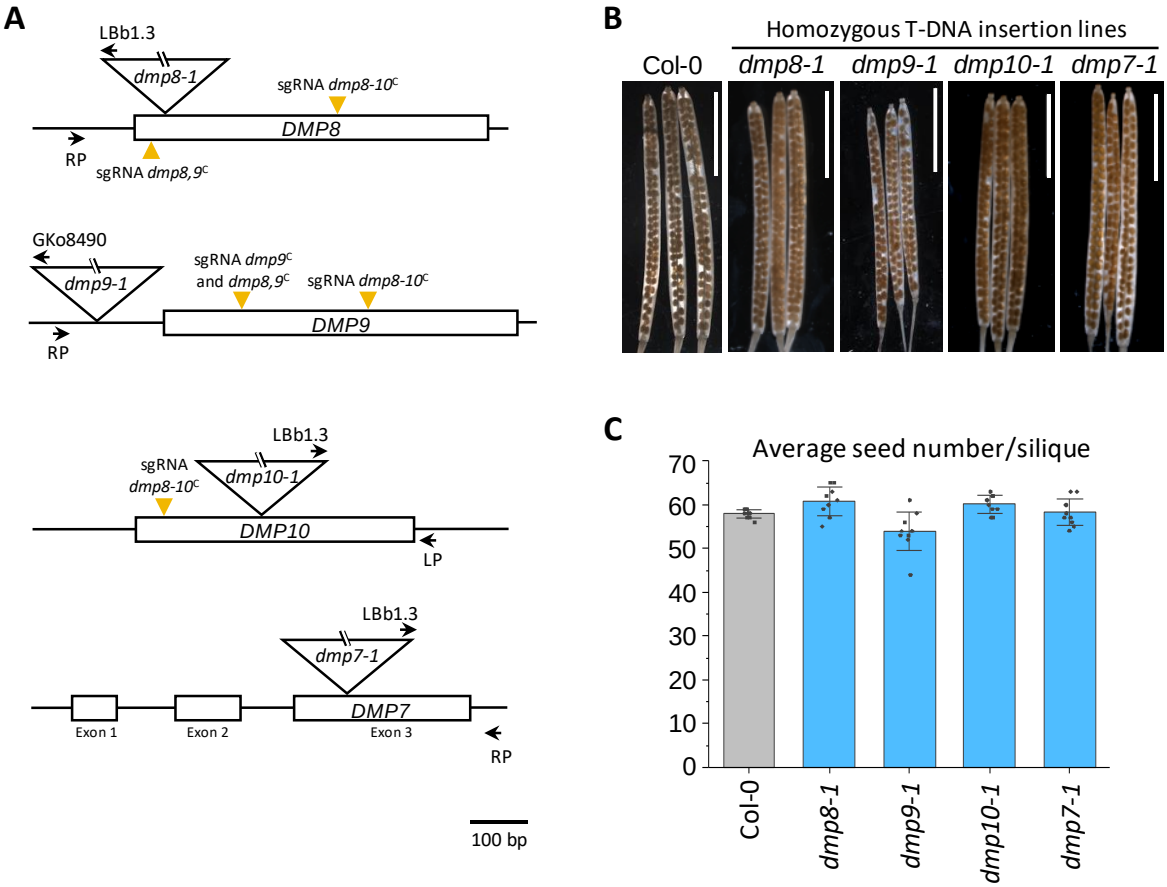


Figure 22 Seed set in siliques from homozygous *dmp* T-DNA insertion lines. A, Genomic structures of *DMPs* and T-DNA positions in the analyzed *dmp8-1* (SALK_131115C), *dmp9-1* (GABI_475G02) and *dmp10-1* (SALK_059670) mutant lines. Additionally, a homozygous T-DNA insertion line for pollen-expressed *DMP7* was investigated (*dmp7-1*; SALK_039261). T-DNA integration loci were validated by sequencing to be in the single exon of *DMP8* (63 bp downstream of start codon), in the promoter region of *DMP9* (24 bp upstream of 5´ UTR), in the single exon of *DMP10* (272 bp downstream of start codon) and in exon 3 of *DMP7* (476bp downstream of start codon), respectively. Position of primers used for genotyping are indicated. Orange triangles point at sgRNA binding sites that have been used to generate the CRISPR/Cas9 genome-edited mutants for *DMP8*, *DMP9* and *DMP10* (shown in Figure 23). **B,** Representative images of cleared siliques harvested from homozygous T-DNA lines in comparison to wild type siliques (Col-0). **C,** Quantification of seed numbers per silique. Data are the means ± s.d. of 10 siliques analyzed per genotype. Scale bars in (**B**) = 5 mm. Modified from Cyprys *et al.*, 2019.

Homozygous T-DNA insertion lines for *dmp7* (SALK_039261), *dmp8* (SALK_131115C), *dmp9* (GABI_475G02) and *dmp10* (SALK_059670) were obtained, and insertion loci were mapped by sequencing to be in the single exon of *DMP8* (63 bp downstream of start codon), in the promoter region of *DMP9* (24 bp upstream of 5´ UTR), in the single exon of *DMP10* (272 bp downstream of start codon), and in exon 3 of *DMP7* (476 bp downstream

of start codon), respectively (Figure 22A). Reproductive defects in *dmp* T-DNA insertion lines were determined by counting the number of developing seeds per silique. T-DNA insertion lines showed no reproductive phenotype and the average seed set was comparable to that of the wild-type (Figure 22B,C).

Assuming functional redundancy among *Arabidopsis thaliana DMPs*, CRISPR/Cas9 was employed to simultaneously knock out multiple *DMPs*, generating genome edited single, double and triple mutants. Mutant alleles were generated for *DMP9* (*dmp9C*) *DMP8* and *DMP9* (*dmp8,9C*), and *DMP8* to *DMP10* (*dmp8–10C*). Furthermore, a *dmp8,9* double mutant was generated by CRISPR/Cas9 genome editing of *DMP9* in the homozygous *dmp8-1* T-DNA background (*dmp9C,8-1*). CRISPR/Cas9 mutants were generated either in wild type (*Col-0*) or in the sperm cell nuclear marker (*HTR10p:HTR10-mRFP*) background. An overview of genome editing events in CRISPR/Cas9 mutants generated in this work can be found in Supplemental table 1.

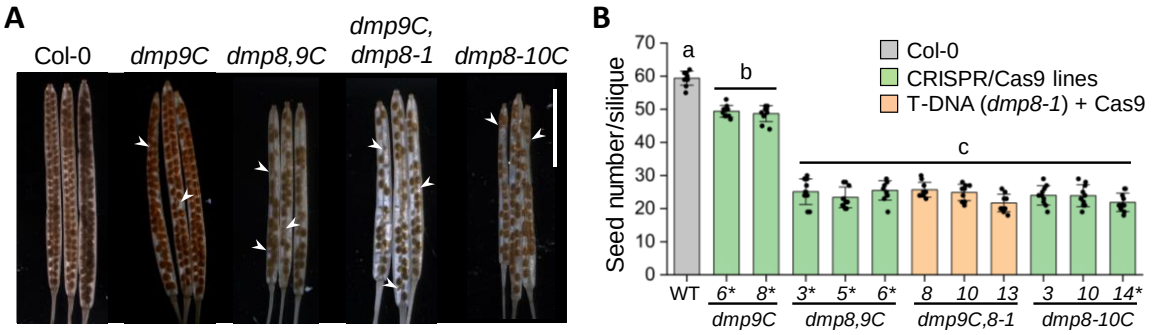


Figure 23 Quantification of seed set in *dmp* mutant siliques. A, Representative images of cleared siliques from the wild-type (Col-0) and individual CRISPR/Cas9 lines generated for *DMP9* (*dmp9C*), *DMP8* and *DMP9* (*dmp8,9C*), *DMP9* in the *dmp8-1* T-DNA insertional mutant (*dmp9C,8-1*) and *DMP8*, *DMP9* and *DMP10* (*dmp8–10C*). The arrowheads point to small, flat seeds. **B,** Quantification of seed number per silique from *Col-0* and independent CRISPR/Cas9 lines. Data are the means ± s.d. of n = 10 siliques per genotype. Significant differences (P < 0.0001 by two-sided Student's t-test) are indicated by different letters. The asterisks indicate CRISPR/Cas9 lines generated in the sperm nuclear marker line *HTR10p:HTR10-mRFP*. WT = wild type. Scale bar = 5 mm. Modified from Cyprys *et al.*, 2019.

A slight but significant reduction of seed set to 74 % occurred in the CRISPR/Cas9 lines *dmp9C_6* and *dmp9C_8*, with 49.4 ± 2.1 and 48.7 ± 2.3 seeds per silique, respectively (Figure 23B). The average percentage of developing seeds declined to 41 % in siliques from the six independent *dmp8 dmp9* double-mutant lines (*dmp8,9C* and *dmp9C,8-1*, respectively), with mean values ranging from 21.7 (±2.7) to 25.7 (±2.3) seeds per silique. No further reduction in seed set was obtained for the triple mutant *dmp8–10C* lines, suggesting functional redundancies between DMP8 and DMP9 but not DMP10. Additionally, small and rather flat seeds became apparent in mutant siliques and were included in the quantification (Figure 23A, arrowheads).

To exclude T-DNA integration effects and off-target effects obtained during CRISPR/Cas9 genome editing on seed set, rescue experiments were performed. *dmp* mutant plant lines devoid of the sgRNA- and Cas9-containing T-DNA were not available at the timepoint of analysis, therefore complementation with Cas9-insensitive constructs was pursued.

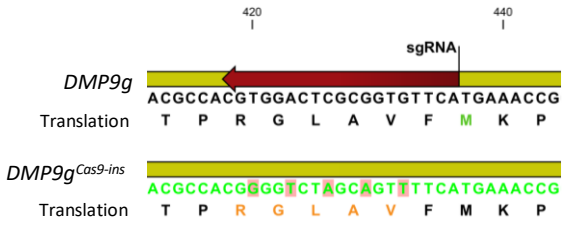


Figure 24 Generation of a Cas9-insensitive *DMP9* genomic fragment. Introduction of synonymous base substitutions at the sgRNA binding site is thought to reduce sgRNA binding efficiency without affecting the translated protein sequence.

Substituting one of the two Gs in the protospacer adjacent motif "NGG" is the most convenient way in generating such a construct. However, substituting one of those two bases would have resulted in an amino acid exchange in the translated protein product. Therefore, Cas9-insensitive complementation constructs were generated by introducing synonymous base substitutions at the *DMP9* sgRNA binding site via site directed mutagenesis (Figure 24).

Complementation of the *dmp8-10C* mutant with a Cas9-insensitive 2,498 bp genomic *DMP9* fragment including promoter and terminator regions (*DMP9g*$^{Cas9\text{-}ins}$) was successful and restored the seed set in plant lines that were PCR-positive for the complementation construct (Figure 25A,B), concluding that off-target effects during CRISPR/Cas9 genome editing can be excluded. Similarly, site-directed mutagenesis was performed with the *DMP9g-GFP* construct was employed to generate Cas9-insensitive *DMP9g-GFP*$^{Cas9\text{-}ins}$. Rescue experiments with *DMP9g-GFP*$^{Cas9\text{-}ins}$ failed to restore the phenotype (Figure 25C). Screening of *DMP9g-GFP*$^{Cas9\text{-}ins}$ expressing lines for GFP signal by fluorescence microscopy revealed signal at the sperm periphery (Figure 25D), similar to the translational reporter line shown in Figure 21. These results suggest that the C-terminal fusion of GFP renders the DMP9-GFP protein non-functional.

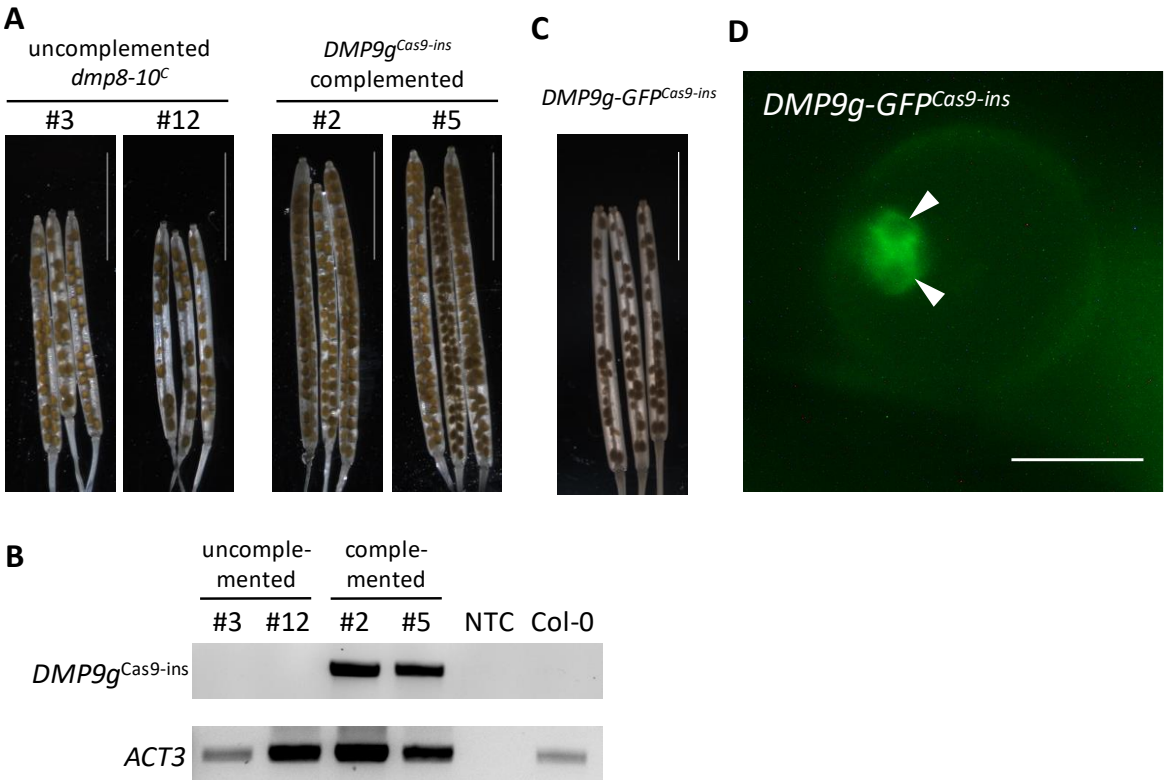


Figure 25 Rescue of the genome-edited *dmp8-10C* mutant with a Cas9-insensitive complementation construct. A, Representative images showing seed set in cleared siliques, harvested from each two *DMP9g*$^{Cas9-ins}$-complemented lines and two uncomplemented *dmp8-10C_3* plants. Seed set is restored in resulting plant lines #2 and #5, which were positive for the Cas9-insensitive genomic construct *DMP9g*$^{Cas9-ins}$. Lines #3 and #12 were negative for the complementation construct and served as controls. **B,** PCR-based genotyping to confirm presence of the complementation construct *DMP9g*$^{Cas9-ins}$ in the independent plant lines shown in (**A**). PCR for *ACTIN3* (*ACT3)* served as DNA quality control. **C,** Complementation with the Cas9-insensitive construct *DMP9g-GFP*$^{Cas9-ins}$ did not restore the phenotype. **D,** Microscopic analysis of *DMP9g-GFP*$^{Cas9-ins}$ pollen grains revealed DMP9-GFP signal at the sperm cell periphery (arrowheads). NTC, non-template control; Col-0, wild type genomic DNA. Scale bars in (**A**) and (**C**) = 5 mm; scale bar in (**D**) = 10 µm. Modified from Cyprys *et al.*, 2019.

4.6.5 *dmp* mutants show impaired male fertility

Although expression and subcellular localization of DMP8/9-GFP indicate a function in sperm cells, reduced fertility can have a multitude of reasons ranging from pollen developmental defects to aberrant pollen tube germination, growth and guidance. Therefore, these processes were dissected in the *dmp8-10C* mutant.

Reciprocal crosses between the *dmp8-10C* mutant and wild type (*Col-0*) were carried out to investigate the role of parental sex on the reduced seed set (Figure 26A). *Col-0* pistils cross-pollinated with *dmp8-10C_3* mutant pollen showed reduced seed set similar to that of self-pollinated *dmp8-10C_3* pistils. When *dmp8-10C_3* pistils were pollinated with *Col-0* pollen almost full seed set was observed. These results demonstrate that male, but not female fertility is affected in *dmp* mutants.

Development of the male gametophyte was assessed by DAPI staining of mature pollen grains. *dmp8-10C_3* mutant pollen was fully developed and two condensed sperm cell nuclei that were strongly stained by DAPI and a diffusely stained vegetative cell nucleus became visible, similar to wild type pollen (Figure 26B). Analysis of female gametophyte development revealed mature ovules containing a fully developed 7-celled embryo sac, as observed in wild type ovules (not shown).

To investigate if *dmp8-10C* pollen is able to germinate and grow *in vitro*, pollen germination assays on solid pollen germination medium were performed. *dmp8-10C* mutant pollen was able to germinate and showed similar growth performance as wild type pollen (Figure 26C). Pollen tube growth and guidance was assayed i*n vivo* by Aniline Blue staining of selfed *dmp8-10C* mutant pistils and control pistils (Figure 26D). Aniline Blue staining of the callose-rich pollen tubes growing through the female reproductive tissues showed that mutant pollen tubes were able to reach the bottom of the ovary. Pollen tube emergence from the transmitting tract and pollen tube guidance towards the micropyle of the ovules was not affected, as *dmp8-10C_3* mutant pollen tubes successfully reached the ovules.

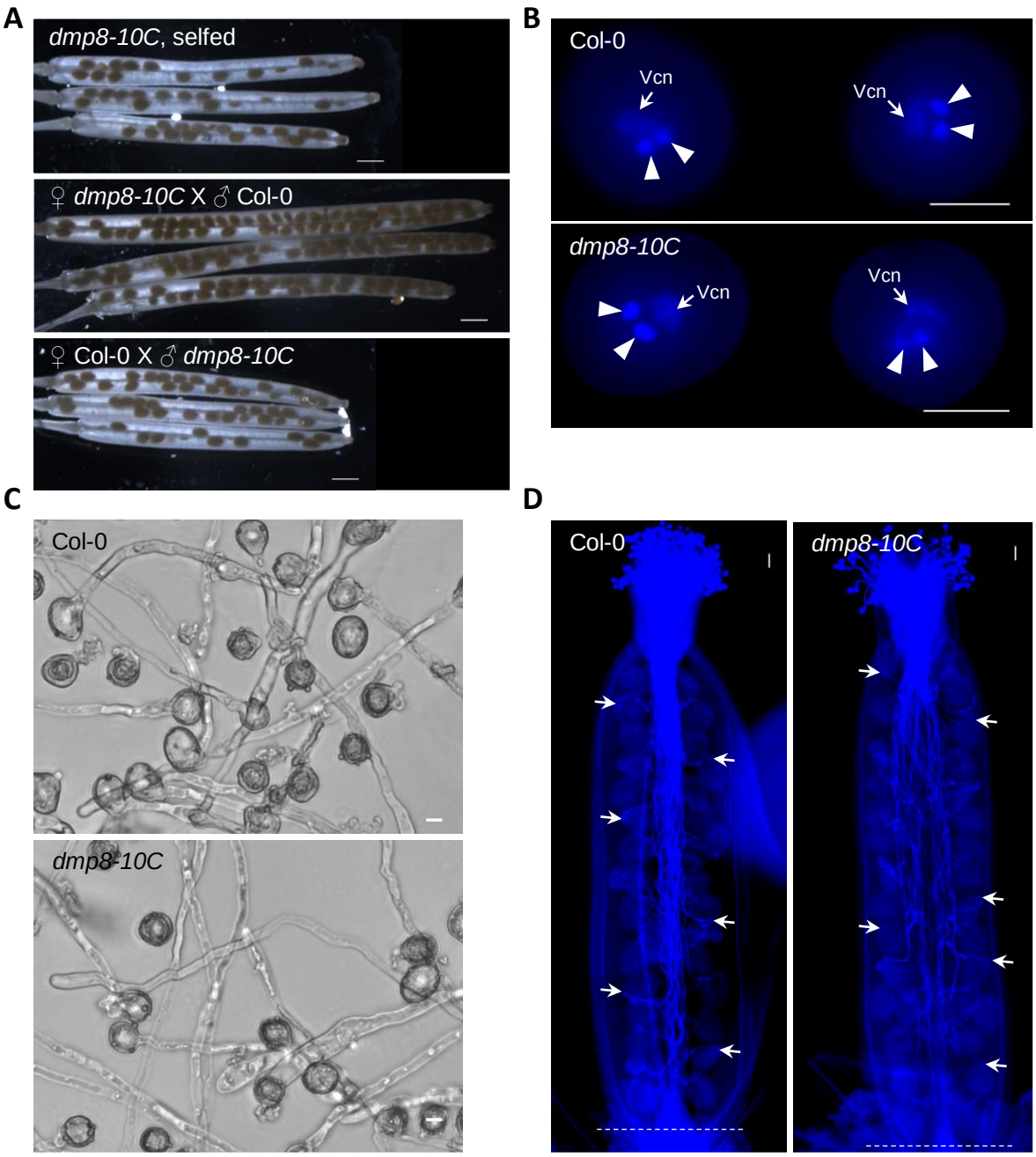


Figure 26 Male fertility is affected in *dmp8-10C* but pollen development, germination and ovule targeting appear normal. A, Reciprocal crosses prove that male fertility is affected in *dmp8-10C.* Wild type (Col-0) pistils, cross-pollinated with *dmp8-10C* mutant pollen, show reduced seed set similar to self-pollinated *dmp8-10C* pistils. **B,** DAPI staining of mature pollen shows that mutant *dmp8-10C* pollen is fully developed, like wild type pollen. Two sperm cell nuclei (arrowheads) and one vegetative cell nucleus (Vcn) are visible **C,** *In vitro* pollen germination demonstrates that mutant *dmp8-10C* pollen shows wild type-like germination and growth. **D,** Aniline blue staining of self-pollinated wild type and *dmp8-10C* pistils at flower stage 14 reveals that growth and guidance of *dmp8-10C* pollen tubes into the stigma, through the style and within the ovary is not affected. Like wild type pollen tubes, *dmp8-10C* pollen tubes reach the bottom of the ovary (dotted line), close to the receptacle. Pollen tube guidance towards the ovules (arrows) appears not to be affected. Scale bars in (**A**) and (**D**) = 1 mm; scale bars in (**B**) and (**C**) = 10 µm. Modified from Cyprys *et al.*, 2019.

In summary, the experiments show that while male fertility is affected in *dmp8-10C_3* mutants, male gametophyte development, pollen tube growth and ovule targeting was comparable to the wild type. Together with the sperm-specific expression and localization of DMP8/9-GFP these results suggest that DMP8 and DMP9 exercise their function redundantly at a step after the two sperm cells have been delivered to the embryo sac by a pollen tube, presumably during gamete activation adhesion or fusion.

4.6.6 *dmp* mutant sperm fail to fuse with the female gametes

Assuming that sperm-expressed *DMP8* and *DMP9* function during a late step in gamete interactions, wild type pistils were pollinated with *dmp8-10C* pollen, and ovules were screened for unfused *dmp8-10C* sperm 30 hours after pollination (HAP), a time point where double fertilization is successfully completed in a wild type situation. The first set of *dmp8-10C* lines generated in this work were in wild type (*Col-0*) background, which made it difficult to observe unfused sperm cells by conventional DIC microscopy. Therefore, ovules were dissected from pistils 30 HAP with *dmp8-10C* mutant pollen and were subjected to Feulgen staining followed by CLSM. Staining of *Arabidopsis* anthers showed that the two sperm cell nuclei and the less condensed vegetative nucleus can be visualized with this method (Figure 27A inset). Wild type pistils pollinated with wild type pollen (30 HAP) and pistils with downregulated *EC1* expression (*ec1-RNAi*) pollinated with wild type pollen (30 HAP) served as control.

CLSM analysis revealed that in a wild type situation double fertilization was successfully completed 30 HAP as indicated by a developing zygote and proliferating endosperm nuclei (Figure 27A). Unfused sperm cells in *EC1*-deficient ovules were previously described (Sprunck *et al.*, 2012; Rademacher and Sprunck, 2013) and, indeed, unfused sperm cells became visible as bright lenticular spots next to the unfertilized egg cell and central cell nucleus (Figure 27B). A similar scenario was observed in wild type pistils pollinated with *dmp8-10C* pollen as unfused sperm cell nuclei were detectable within the embryo sac (Figure 27C).

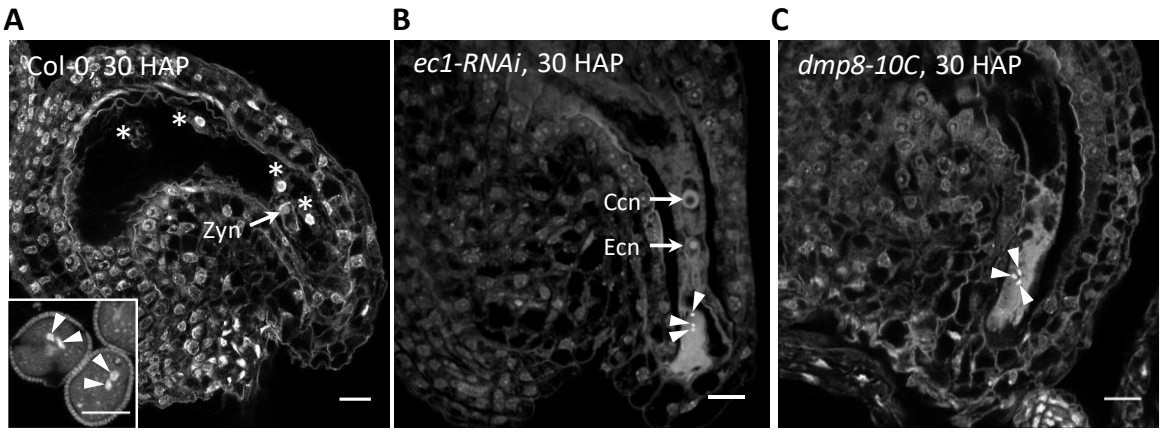


Figure 27 *dmp8-10C* mutant sperm fail to fuse with the female gametes. CLSM single optical sections of Feulgen-stained Col-0 ovules pollinated with either Col-0 (**A**) or *dmp8-10C* (**C**) pollen and *ec1-RNAi* mutant ovules (**B**) 30 HAP. Zygote (Zyn) and endosperm nuclei (asterisks) are visible in *Col-0* ovules. In *dmp8-10C* and *ec1-RNAi* ovules, nuclei of unfused sperm cells (arrowheads) are detectable. Inset in (**A**) shows strong staining of sperm cell nuclei (arrowheads) in Feulgen-stained pollen grains. Ccn = central cell nucleus; Ecn = egg cell nucleus. Scale bars = 10 µm.

These results suggested that *dmp8-10C* sperm cells fail to fuse with the female gametes and remain unfused in the female gametophyte even 30 HAP. For quantitative assessment of unfused sperm cells by fluorescence microscopy, a novel set of genome-edited single, double and triple mutants was generated in the sperm cell nuclear marker line *HTR10p:HTR10-mRFP* (*dmp9C HTR10-mRFP, dmp8,9C HTR10-mRFP* and *dmp8-10C HTR10-mRFP*). The mutants showed reduced seed set similar to the first set of mutants generated in wild type background (Figure 23).

Unfused *dmp8,9C HTR10-mRFP* sperm cell nuclei were readily observed in pollen tube-targeted ovules 42 HAP. Besides two unfused sperm cells, ovules with either one, three or four unfused sperm cells were also detected (Figure 28A). Pollinations performed with control pollen (*HTR10-mRFP)* showed that 98.1 % of the control ovules were successfully fertilized 42 HAP and did not contain unfused sperm cells. By contrast, only 44.4 % of mutant ovules were fertilized. Of the mutant ovules, 55.6 % contained unfused sperm cells. Besides ovules with two unfused *dmp8,9C HTR10-mRFP* sperm cells (31.8 %), ovules with either one/three (18 %) or four (5.8 %) unfused sperm cells were observed. This indicates supernumerary sperm delivery, which has already been described for other fertilization-defective *Arabidopsis* mutants (Kasahara *et al.*, 2012; Beale *et al.*, 2012). In case that the first sperm pair fails to fertilize the female gametes, the persistent synergid cell of the ovule is able to attract a second pollen tube, delivering another pair of sperm cells to ensure reproductive success.

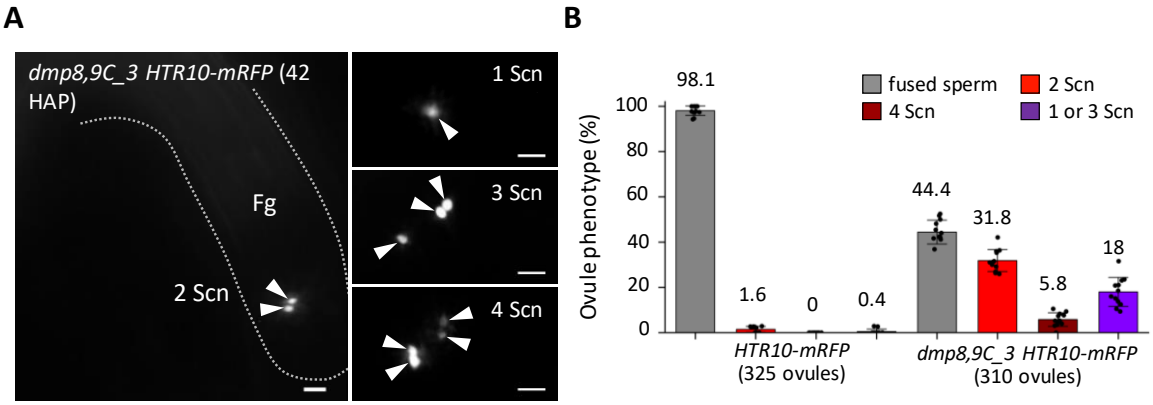


Figure 28 Quantification of unfused sperm cells phenotypes in ovules targeted with *dmp8,9C_3 HTR10-mRFP* pollen tubes. A, Representative images of pollen tube-targeted ovules with unfused *dmp8,9C_3 HTR10-mRFP* sperm nuclei 42 HAP. Either one, two, three or four unfused sperm cell nuclei (Scn) were detected (arrowheads). The dashed line indicates the position of the female gametophyte (Fg). **B,** Quantification of unfused sperm phenotypes in ovules targeted with *dmp8,9C_3 HTR10-mRFP* pollen 42 HAP. Control experiments were performed with *HTR10-mRFP* pollen (wild type for DMP8 and DMP9). Data are the means ± s.d. of n = 12 pistils per genotype. Scale bars = 10 µm. Modified from Cyprys *et al.*, 2019.

Quantification of unfused sperm cells (42 HAP) was also performed with the single mutant *dmp9C_8 HTR10-mRFP.* Although seed set was only slightly reduced in *dmp9C HTR10-mRFP* single mutants (Figure 29A), ovules with two (25.1 %), one or three (10.4 %) or four (2.2 %) unfused sperm cells were frequently observed (Figure 29B). The almost full seed set in the *dmp9C_8 HTR10-mRFP* single mutants suggests, that most of the unfused *dmp9C_8 HTR10-mRFP* sperm cells successfully fertilize the female gametes at a time point later than 42 HAP, thus indicating delayed gamete fusion.

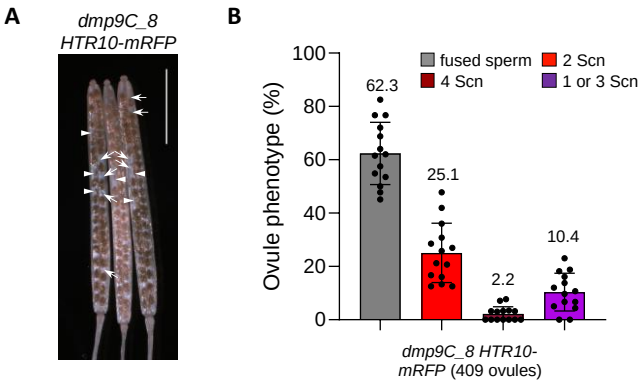


Figure 29 Quantification of unfused sperm cells phenotypes in ovules targeted with *dmp9C_8 HTR10-mRFP* pollen tubes. A, Siliques of homozygous *dmp9C_8 HTR10-mRFP* plants show almost full seed set, with few unfertilized ovules (arrows) and aborted seeds (arrowheads). **B,** Quantification of unfused sperm phenotypes in ovules targeted with *dmp9C_8 HTR10-mRFP* pollen 42 HAP. Data are the means ± s.d. of n = 14 pistils. Scale bar = 5 mm.

4.6.7 *dmp* mutant sperm preferentially fertilize the central cell

The observation of odd numbers of unfused sperm cells in *dmp* mutants suggests that a single sperm has successfully fertilized a female gamete whereas the other sperm did not. This phenomenon is termed single fertilization and has also been observed for the *gex2* mutant impaired in sperm adhesion (Mori *et al.*, 2014). Ovules that have started seed development after single fertilization will abort later on. Indeed, small and rather flat seeds were detected in almost mature *dmp* mutant siliques (Figure 23A arrowheads). Dissection of *dmp8,9C_3 HTR10-mRFP* siliques during maturation revealed that next to unfertilized ovules and viable seeds, these siliques harbor smaller white seeds that will later become brown and shriveled (Figure 30A), indicating single fertilization events.

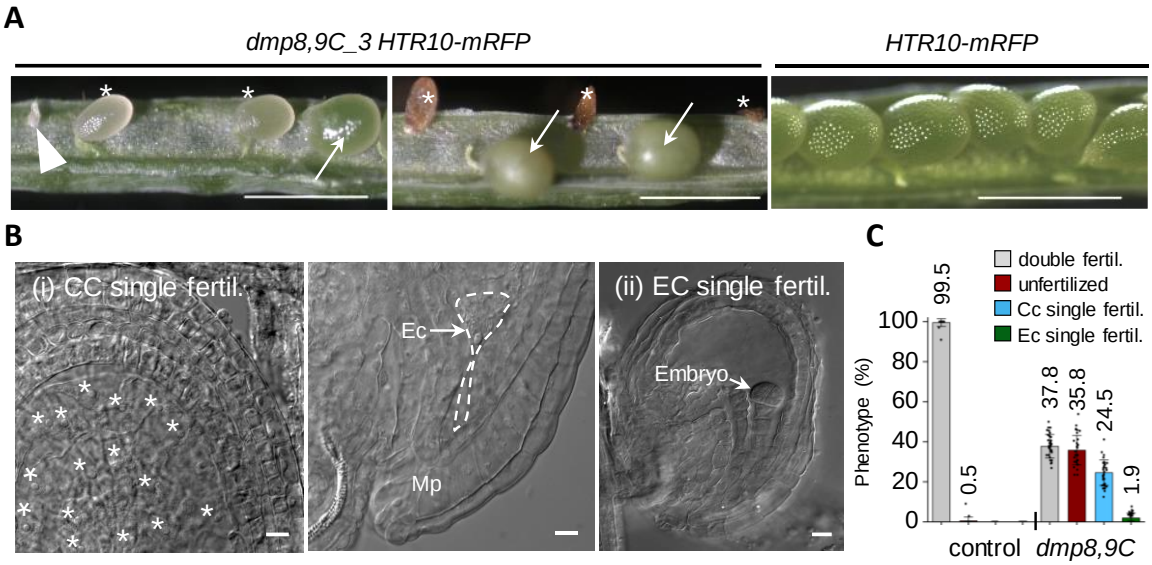


Figure 30 Quantification of single fertilization events by *dmp8,9C* sperm cells. A, Dissected siliques showing an unfertilized ovule (arrowhead), aborted seeds (asterisks) and viable seeds (arrows) in *dmp8,9C_3 HTR10-mRFP*. In young siliques (left panel), aborted seeds remain white and stay smaller than viable ones. Later on, they become brown and shriveled (middle panel). *HTR10-mRFP* siliques were used as control (right panel). **B,** Single fertilization events detected in cleared *dmp8,9C_3 HTR10-mRFP* ovules 5 DAP. (i) Single central cell fertilization with dividing endosperm nuclei (asterisks) and an unfertilized egg cell (Ec) in the same ovule, close to the micropyle (Mp). (ii) An ovule containing an embryo but lacking endosperm. **C,** Quantification of ovule phenotypes in *dmp8,9C_3 HTR10-mRFP* (796 ovules, n = 34 pistils) and the control (*HTR10-mRFP*; 631 ovules, n = 24 pistils) 5 DAP. Data are the means ± s.d. Scale bars in (A) = 1 mm ; scale bars in (B) = 10 µm. Modified from Cyprys *et al.*, 2019.

To investigate and quantify the occurrence of single fertilization events by *dmp8,9C* sperm cells, developing *dmp8,9C_3 HTR10-mRFP* siliques were dissected 5 DAP and ovules were cleared for DIC microscopy. Pistils pollinated with *HTR10-mRFP* pollen served as a control. In the control 99.5 % of ovules were fertilized and displayed both developing endosperm and embryo. On the contrary, 24.5 % fertilized ovules with proliferating endosperm nuclei but lacking a detectable embryo were detected in *dmp8,9C_3 HTR10-mRFP* pistils.

Eventually, a collapsed egg cell was found at their micropyle, suggesting single fertilization of the central cell (Figure 30B(i)). Only in very rare cases (15 out of 796 ovules; 1.9 %), single fertilization of the egg cell was observed, indicated by ovules containing an embryo but no endosperm nuclei (Figure 30B(ii)). These seeds abort much earlier than their single central cell fertilized counterpart. Taking into account all fertilization events (Figure 30C) 62.3 % of the central cells, but only 39.5 % of the egg cells fused with *dmp8,9C_3 HTR10-mRFP* sperm cells, suggesting that *dmp8,9C* mutant sperm preferentially fertilize the central cell and that sperm–egg fusion is more dependent on functional DMP8 and DMP9 than is sperm fusion with the central cell.

Autonomous endosperm development in the absence of central cell fertilization is a phenomenon frequently observed in apomictic plants but has also been described for a set of *Arabidopsis thaliana fis (FERTILIZATION INDEPENDENT SEED DEVELOPMENT*) mutants (Ohad *et al.*, 1996; Chaudhury *et al.*, 1997; Grossniklaus *et al.*, 1998). To exclude that the proliferating endosperm nuclei observed in endosperm-only ovules are due to this phenomenon, *dmp8-10C HTR10-mRFP* pollen was used to pollinate the egg cell nuclear marker line *EC1.1p:NLS3xGFP*. Ovules were dissected from pollinated pistils and analyzed by CLSM.

♀ *EC1.1p:NLS3xGFP* X ♂ *dmp8-10^{C} HTR10-mRFP*

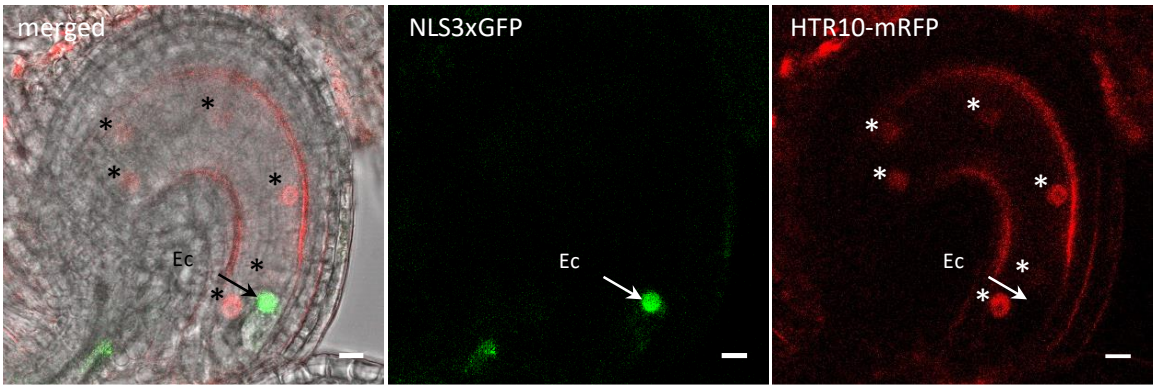


Figure 31 Paternally-derived HTR10-mRFP fluorescence in dividing endosperm nuclei indicate single central cell fertilization with mutant *dmp8-10C HTR10-mRFP* sperm. Emasculated pistils of the egg cell marker line *EC1.1p:NLS-3xGFP* were hand-pollinated with mutant *dmp8-10C_14 HTR10-mRFP* pollen. 30 HAP, approximately 24 % ovules (for quantification see Figure 30C) exhibit dividing endosperm nuclei (asterisks) with residual red fluorescence of paternally-derived *HTR10-mRFP* (middle panel). By contrast, the green fluorescent nucleus of the unfertilized egg cell (Ec) is clearly visible but lacks paternally derived red fluorescence (right panel). Scale bar = 10 µm. Modified from Cyprys *et al.*, 2019.

Figure 31 depicts an ovule with proliferating endosperm nuclei, showing residual red fluorescence of sperm cell-derived HTR10-mRFP. By contrast, the unfertilized egg cell shows a single green fluorescing nucleus but lacks red fluorescence. This result provides

evidence that the phenotype observed in Figure 30B can be directly attributed to single fertilization of the central cell caused by a *dmp8,9C* mutant sperm cell.

4.6.8 *dmp* mutant sperm cells can attach to the female gametes

It is known that molecular events that take place in the late phases of gamete interaction involve 3 phases: sperm activation, gamete adhesion and gamete fusion (see chapter 3.6). In order to further dissect the phase during which DMP8 and DMP9 exercise their biochemical function, *in situ* sperm cell adhesion assays with female gametes within the ovule were performed. Enzymatic removal of surrounding cell wall material allows the female gamete protoplasts to round up, resulting in a certain degree of movement within the embryo sac. This allows to discriminate between attached and non-attached sperm cells.

In order to visualize the egg plasma membrane, a marker line expressing tagRFP fused to the remorin anchor of *Medicago truncatula* symbiotic Remorin 1 (*MtSYMREM1*) under the control of the egg cell-specific *EC1.1* promoter (*EC1.1p:tagRFP-RemA*) was developed (Lindemeier, 2017). Analysis of *EC1.1p:tagRFP*-RemA-expressing ovules by CLSM revealed that the egg cell plasma membrane is specifically labeled in this marker line and the outline of the egg became visible as a single cone-shaped cell at the micropylar pole (Figure 32A). Successful cell wall degradation was indicated by almost spherically shaped egg and central cell protoplasts within the ovule (Figure 32B). For unfused sperm cells, varied positions after protoplasting were expected, depending on whether the sperm cells managed to separate and/or adhere to one of the female gametes (Figure 32C): non-attached sperm cells will slip down towards the micropyle, when the receptive synergid cell degenerates (1). Attached but unfused sperm pairs may either adhere to the protoplast of the egg cell (2) or the central cell (3). In case the sperm pair separates, either of them may attach to one of the female gametes (4). Without protoplasting, unfused *dmp8,9C_3 HTR10-mRFP* sperm cells were detected close to the egg cell even 24 HAP, when wild-type sperm cells had already accomplished double fertilization. (Figure 32D(i)). Notably, the largest proportion of unfused sperm cells adhered to the egg cell plasma membrane after protoplasting (Figure 32D(i)-(iv)). In 82.6 % of the ovules with unfused *dmp8,9C_3 HTR10-mRFP* sperm cells, these remained attached to the egg cell plasma membrane either as a pair (58.8 ± 4.6 %) or as single sperm (23.8 ± 2.5 %). In the remaining ovules with unfused sperm cells (17.4 %), these were neither attached to the egg nor to the central cell (Figure 32G).

The same assay was repeated to evaluate attachment of sperm cells to the central cell. Due to the lack of a central cell plasma membrane marker with sufficiently strong fluorescence, the marker line *DD65p:GFP* (Steffen *et al.*, 2007), labeling the cytoplasm of the central cell, was used for pollinations with *dmp8,9C_3 HTR10-mRFP* pollen. However, *dmp8,9C_3 HTR10-mRFP* were never found to be attached to the central cell (Figure 32E).

Furthermore, gamete attachment assays with heterozygous *ec1-RNAi/+* plants, in which egg cell-specific *EC1* expression is reduced in 50 % of the ovules, and sperm fusion failure rates of 44–47 % are observed (Sprunck *et al.*, 2012), were carried out (Lindemeier, 2017).In these assays, 76 % ± 3.3 % of the unfused *HTR10-mRFP* sperm cells *in ec1-RNAi* ovules were neither attached to the egg cell nor to the central cell but slipped down towards the micropyle after protoplasting (Figure 32F), suggesting that sperm cell adhesion and activation in *Arabidopsis* are interdependent processes. In the remaining ovules with unfused sperm cells, these were either attached to the egg cell (17 % ± 4.9%) or to the central cell (7 % ± 2.9%). Furthermore, non-attached sperm cells in *ec1-RNAi* ovules stayed together as a pair (Figure 32F), suggesting a correlation between EC1-mediated sperm activation and the physical separation of the sperm pair.

In summary, the results from *in situ* sperm cell adhesion assays suggest a model in which the sperm cells are delivered to the female gametophyte as an interconnected pair that initially attaches to the female gametes. Once the sperm pair manages to separate, a decoupled sperm that is not attached to the egg cell can fuse with the central cell, whereas sperm–egg fusion is much more dependent on the action of DMP8 and DMP9. This is in line with the observation that *dmp8,9* mutant sperm cells preferentially fertilize the central cell (Figure 30C).

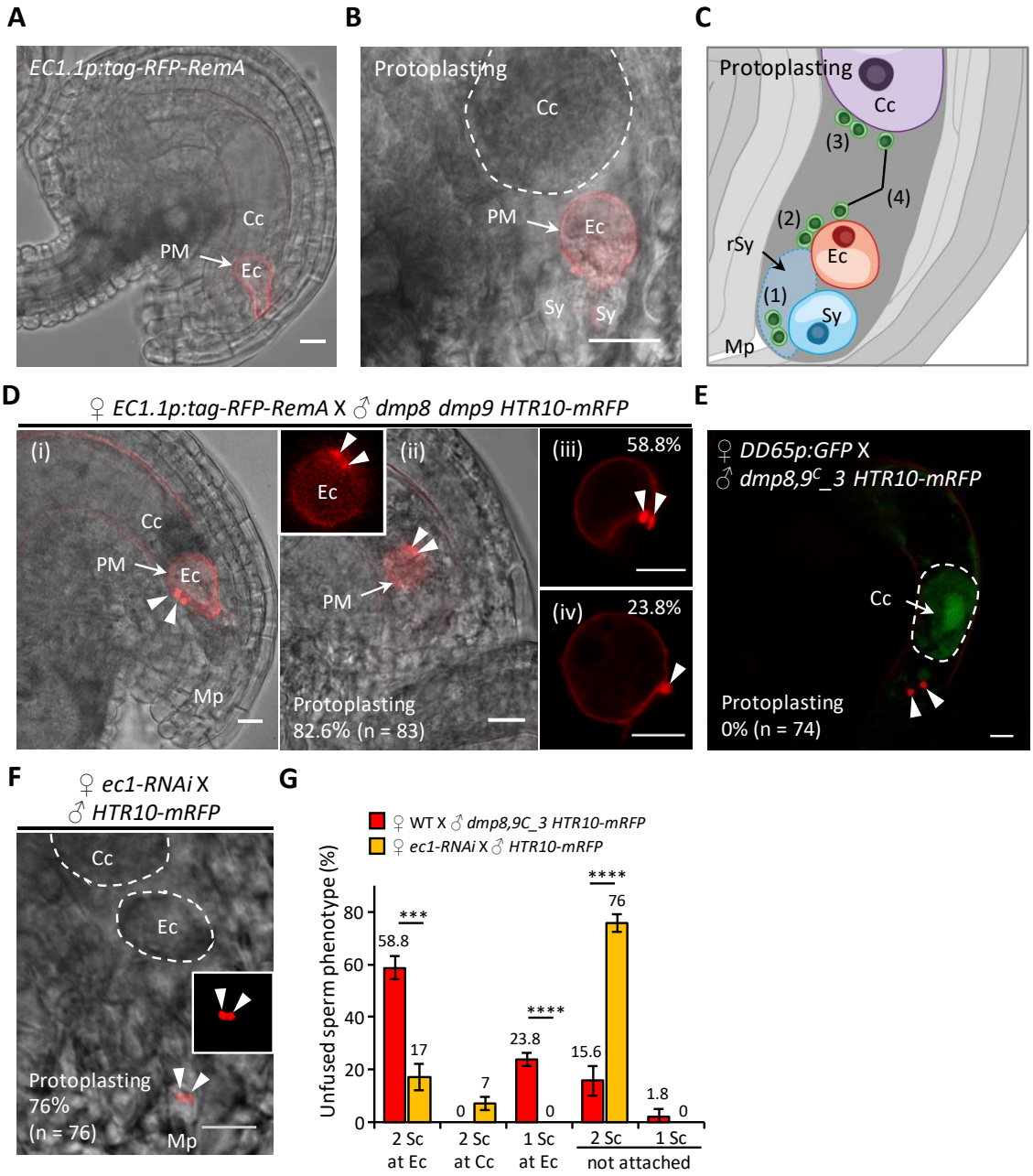


Figure 32 Unfused *dmp8,9C_3 HTR10-mRFP* sperm cells can attach to the female gametes. A, Egg cell plasma membrane (PM) marker line used for protoplasting of the egg cell (Ec), central cell (Cc) and synergid cells (Sy) within the ovule. **B,** Succesful protoplasting was indicated by spherical egg cell and central cell protoplasts. **C,** Potential sperm cell positions after protoplasting: non-attached sperm cells (1); a sperm pair attached to the egg cell (2) or central cell (3); separated sperm pair attached to the female gametes (4). **D,** The egg cell plasma membrane marker 24 HAP with *dmp8 dmp9* mutant pollen (83 ovules; n = 3 independent experiments). (i) Unfused *dmp8–10C_14 HTR10-mRFP* sperm cells (arrowheads) in close proximity to the egg plasma membrane before protoplasting. (ii) Unfused *dmp8–10C_14 HTR10-mRFP* sperm cells attached to the egg cell after protoplasting. Phenotype of unfused *dmp8,9C_3 HTR10-mRFP* sperm cell pairs (iii) and single sperm (iv) attaching to the egg plasma membrane (quantification shown in G). **E,** Protoplasting experiments with central cell marker *DD65p:GFP*. Unfused *dmp8,9C_3 HTR10-mRFP* sperm cells (arrowheads) or single sperm were never found attached to the central cell. **F,** Protoplasting of *ec1-RNAi* ovules 20 HAP with *HTR10-mRFP* pollen. In the majority of cases, the unfused sperm cell pair (arrowheads) locates close to the micropyle. **G,** Quantification of unfused sperm cell phenotypes in protoplasted ovules shown in (**D**) and (**F**). Data are the means ± s.d. of n = 3 (*dmp8,9C_3 HTR10-mRFP*) and n = 4 *(ec1-RNAi*) independent experiments. Significant differences (two-sided Student's t-test) are indicated (****P < 0.0001). Scale bars = 10 µm. Modified from Cyprys *et al.*, 2019. Images shown in (**A**), (**B**), (**C**), and (**F**) were generated by Dr. Maria Lindemeier (Lindemeier, 2017).

4.6.9 All domains of DMP9 are important for function

In order to identify functionally important regions in DMP9, rescue experiments with mutated DMP9 versions were performed. Truncated DMP9 versions were tested for their ability to complement the gamete fusion defective phenotype observed in *dmp8,9C* mutants.

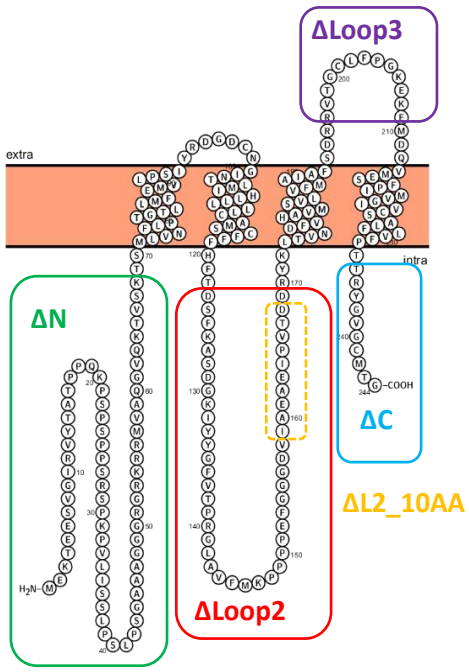


Figure 33 Beads-on-a-string model of DMP9 inserted into the plasma membrane. DMP9 is characterized by 4 predicted transmembrane α-helices, amino- and carboxy-termini facing the cytosol, two smaller extracellular loops (loop 1, loop 3) and a larger intracellular loop (loop 2). Amino acids missing in the truncated DMP9 versions are highlighted with boxes. Amino acids are in one letter code.

Figure 33 depicts a beads-on-a-string model for DMP9 inserted into the plasma membrane, reflecting its Tetraspanin-like topology consisting of cytoplasmic N- and C-termini and a large intracellular loop (loop 2), contrasted by two smaller loops 1 and 3, facing the extracellular space. The genomic Cas9-insensitive construct which successfully complemented the *dmp8,9* mutant (see Figure 25) was used as template for site directed mutagenesis, thereby generating constructs that express truncated forms of DMP9 under control of the endogenous promoter. Five DMP9-deletion constructs were generated in total resulting truncation of DMP9 either at the N-terminus ($DMP9_{\Delta N}$; lacking amino acids 1-68), the intracellular loop 2 ($DMP9_{\Delta Loop2}$; lacking amino acids 123-169), the extracellular loop 3 ($DMP9_{\Delta Loop3}$; lacking amino acids 197-209) or the C-terminus ($DMP9_{\Delta C}$; lacking amino acids 235-244). Considering the extended loop 2 of DMP8 and DMP8 (see Figure 19), a 10 amino acid deletion within loop 2 of DMP9 ($DMP9_{\Delta L2\_10aa}$; lacking amino acids 159-168) was introduced. For $DMP9_{\Delta Loop2}$ and $DMP9_{\Delta Loop3}$ each three amino acids adjacent to the predicted transmembrane helices were left untouched to allow formation of a minimal loop and to avoid altered protein topology caused by two consecutive transmembrane α-helices (Figure 33). The deletion constructs were transformed into the *dmp8,9C_3 HTR10-mRFP* mutant and the ability of the truncated proteins to rescue the fusion defective phenotype was evaluated by quantification of seed numbers in almost mature, cleared siliques (Figure 34).

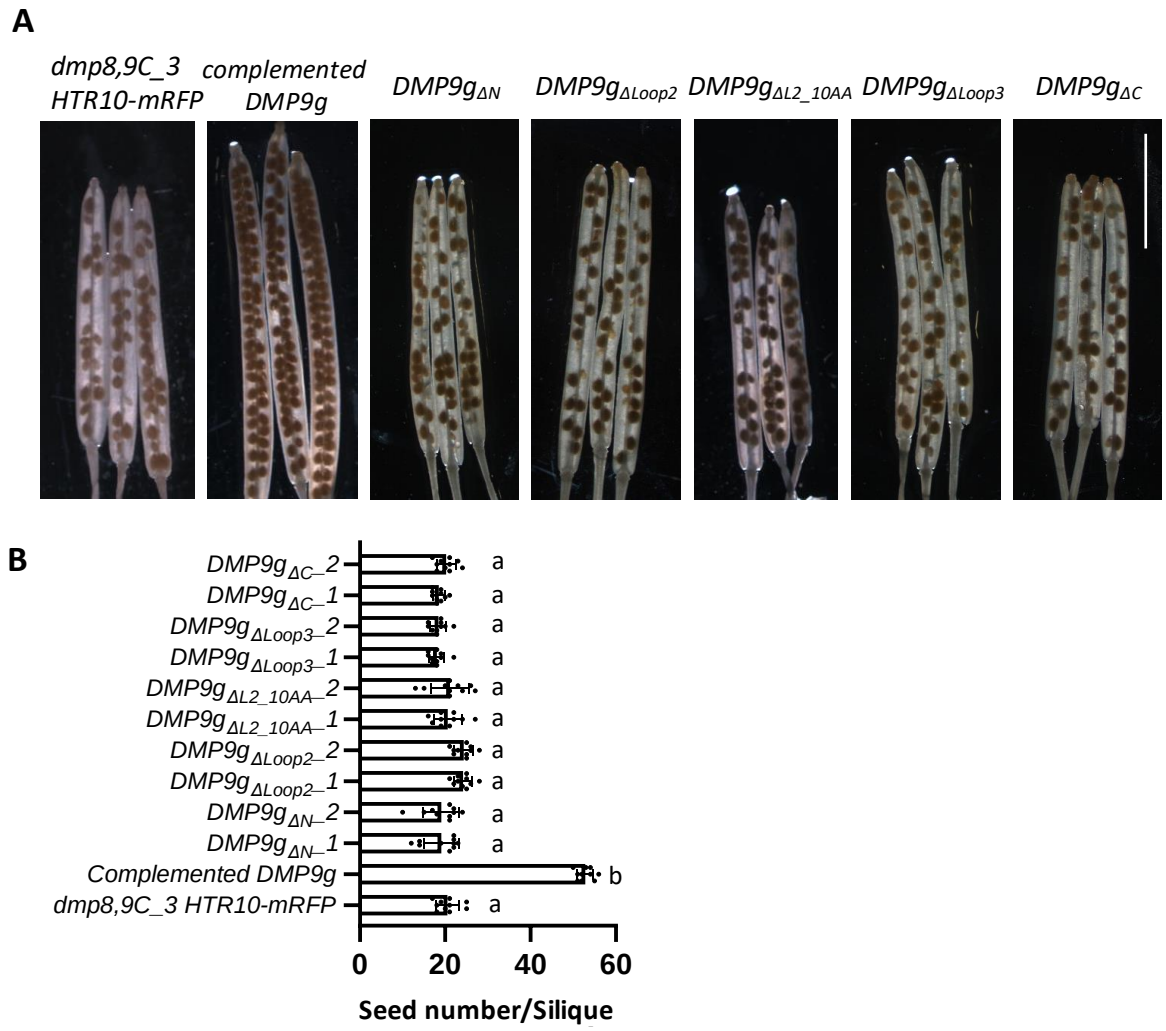


Figure 34 Complementation of the *dmp8,9C_3 HTR10-mRFP* mutant with deletion constructs. A, Representative images of cleared siliques from the uncomplemented *dmp8,9* double mutant (*dmp8,9C_5 HTR10-mRFP*), the mutant complemented with a Cas9-insensitive *DMP9* genomic fragment (*complement DMP9*), and five truncated versions of DMP9 (*DMP9g$_{\Delta N}$*, *DMP9g$_{\Delta Loop2}$*, *DMP9g$_{\Delta L2_10AA}$*, *DMP9g$_{\Delta Loop3}$*, *DMP9g$_{\Delta C}$).* **B,** Quantification of seed numbers in almost mature, cleared siliques from independent transgenic T1 lines. Data are the means ± s.d. of n = 10 siliques per genotype. Significant differences (P < 0.0001 by two-sided Student's t-test) are indicated by different letters. Scale bar = 5 mm.

The *dmp8,9C_3 HTR10-mRFP* mutant showed an average of 20.5 ± 2.6 seeds per silique, small and flat seeds included. The number of developed seeds significantly increased to an average of 52.8 ± 1.9 seeds per silique in the T1 line heterozygous for the *DMP9* genomic fragment (*Complemented DMP9g*), indicating successful rescue of the phenotype. However, all tested deletion constructs failed to complement the *dmp8,9* double mutant as the seed set in the T1 progeny was not significantly higher from the *dmp8,9C_3 HTR10-mRFP* line, with mean values ranging from an average 18 ± 1.7 (*DMP9g$_{\Delta Loop3}$_1*) to 24.1 ± 4.4 (DMP9g$_{\Delta L2_10AA}$_2) seeds per silique (Figure 34B).

4.6.10 *Amborella trichopoda DMP9-like* partially rescues the *dmp8,9* phenotype

Proteins containing the plant-specific DUF679 can be found in all phyla of the Viridiplantae and the protein family expanded specifically in the angiosperm lineage (see 4.6.2). To address whether the gamete fusion-promoting function observed for *Arabidopsis thaliana* DMP8 and DMP9 is an evolutionary conserved feature of DUF679 containing membrane proteins, rescue of the *dmp8,9C_3 HTR10-mRFP* mutant with *DMPs* from other species was performed. The single *DMP* from the unicellular algae *Chlamydomonas reinhardtii* (*CrDMP*; *Cre07.g312850*; InterPro accession: A0A2K3DIH4) was chosen for being the most ancient *DMP* (see Figure 18) and because of the large evolutionary distance to *Arabidopsis thaliana DMPs*, as green algae diverged from streptophytes including land plants over 1 billion years ago (Merchant *et al.*, 2007). *Amborella trichopoda* is holding a key position in the angiosperm evolutionary tree as the sister species to all other flowering plants, and diverged from them over 200 million years ago (Li *et al.*, 2019). The *Amborella* genome codes for 5 *DMP* genes and the *DMP-like* gene which was found within the *Arabidopsis thaliana* DMP8/9/10 subclade in the phylogenic analysis (see Figure 18) was chosen for complementation (AmTrDMP; *model.AmTr_v1.0_scaffold00010.477;* InterPro accession: W1NG81). *CrDMP* and *AmTrDMP* genomic fragments excluding putative promoter regions were PCR amplified from genomic DNA and cloned under transcriptional control of the sperm-specific *Arabidopsis thaliana HTR10* promoter (Ingouff *et al.*, 2007). The resulting constructs *HTR10p:CrDMP* and *HTR10p:AmTrDMP* were used to transform *dmp8,9C* double mutants by floral dip. Rescue of the phenotype was evaluated in two independent T1 lines each, by quantification of seed numbers in almost mature, cleared siliques (Figure 35).

The construct *HTR10p:CrDMP* failed to rescue the phenotype as the average seed number in two independent transgenic lines ranged from 18.9 ± 2.4 to 18.9 ± 3.6 seeds per silique and was not significantly different from the non-complemented control line *dmp8,9C_3 HTR10-mRFP* (20.5 ± 2.8 seeds per silique). When expressed under control of the *Arabidopsis thaliana HTR10* promoter, the *Amborella trichopoda DMP* partially rescued the phenotype and average seed number per silique significantly increased to 34.2 ± 2 and 32.2 ± 3.5 in two independent *HTR10p:AmTrDMP* lines, as compared to the non-complemented control (Figure 35B).

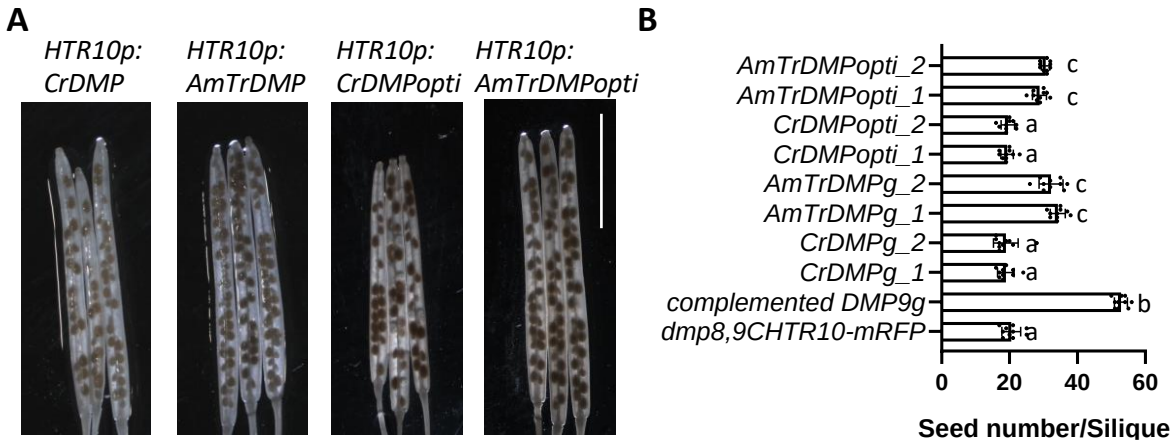


Figure 35 Complementation of the *dmp8,9C_3 HTR10-mRFP* mutant with *Chlamydomonas reinhardtii* and *Amborella trichopoda DMP*. A, Representative images of cleared siliques from the *dmp8,9C_3 HTR10-mRFP* mutant complemented with *Chlamydomonas reinhardtii DMP (CrDMP)* and *Amborella trichopoda DMP* (*AmTrDMP*) under transcriptional control of the sperm-specific *HTR10* promoter. Native genomic sequences as well as codon optimized sequences for expression in *Arabidopsis thaliana* (*opti*) were used for complementation. **B,** Quantification of seed numbers in almost mature, cleared siliques from independent transgenic lines. Data are the means ± s.d. of n = 10 siliques per genotype. Significant differences ($P < 0.0001$ by two-sided Student's t-test) are indicated by different letters. Scale bar = 5 mm.

Recent results have demonstrated that codon usage can affect protein abundance before translation, e.g. by modulating mRNA stability and assembly of the translation machinery (Wu *et al.*, 2019). Considering the large evolutionary distance, particularly between *Arabidopsis thaliana* and *Chlamydomonas reinhardtii,* codon usage of the *CrDMP* and *AmTrDMP* coding sequences was optimized for expression in *Arabidopsis thaliana*. Codon optimized coding sequences fused to the efficient *Arabidopsis thaliana HSP18.2* terminator (Nagaya *et al.*, 2010) were generated by gene synthesis and cloned to be expressed by the *HTR10* promoter (*HTR10p:CrDMPopti* ; *HTR10p:AmTrDMPopti*). Results were similar to complementation with the non-optimized constructs. The number of seeds per silique for *HTR10p:CrDMPopti* ranged from 19.3 ± 1.8 to 19.5 ± 2.0 in two independent *HTR10p:CrDMPopti* lines and was not significantly different from the non-complemented *dmp8,9C_3 HTR10-mRFP* line or the complementation with original *CrDMP*. The two independent *HTR10p:AmTrDMPopti* lines displayed a seed number per silique ranging from 28.8 ± 2.1 to 30.6 ± 1.3. Although seed set in *HTR10p:AmTrDMPopti* complemented lines significantly increased as compared to the non-complemented mutant, no significant difference between *HTR10p:AmTrDMP* and *HTR10p:AmTrDMPopti* was observed (Figure 35B). Thus, codon optimization of *AmTrDMP* for expression in *Arabidopsis thaliana* did not improve the partial complementation obtained by the original *Amborella DMP9-like* sequence any further.

4.6.11 Overexpression of DMP9-GFP in tobacco epidermis cells induces membrane remodeling

The loss-of-function studies in this work indicate a role for *Arabidopsis thaliana* DMP8 and DMP9 after gamete attachment but before plasma membrane merging, however the biochemical role of DMP8 and DMP9 and their molecular mode of action remained elusive. While functional studies on DMPs are scarce, Kasaras *et al.* (2012) suggested a role of *Arabidopsis thaliana* DMP1 in membrane remodeling of the ER and tonoplast, based on results from transient overexpression of DMP1-GFP in *Nicotiana benthamiana* leaf epidermis cells. When DMP1-GFP was overexpressed, membrane remodeling events were observed, affecting the structure of the ER and the tonoplast, characterized by bulbs, invaginations and foamy structures, ultimately leading to cell death after breakdown of the entire ER network and the vacuole (Kasaras *et al.*, 2012).

To investigate whether DMP9 also alters membrane shape or integrity *in vivo, Agrobacterium tumefaciens* harboring a *35Sp:DMP9-GFP* construct were infiltrated into *Nicotiana benthamiana* leaves for transient protein overexpression followed by CLSM analysis. *35Sp:DMP1-GFP* (a gift from Reinhard Kunze) was used as control (Figure 36A). DMP1-GFP behaved as described by Kasaras *et al.* (2012): 2 days post infiltration (dpi) GFP signal uniformly decorated the membrane of the tonoplast, and cytoplasmic strands penetrating the vacuole became visible. 4 dpi the signal at the vacuolar membrane became patchy and aggregates with strong GFP signal intensity appeared. Furthermore, spherical structures and foamy aggregates were detected. Tobacco leaf cells expressing DMP9-GFP showed a very similar behavior, but first spheres and membrane meshwork appeared already 2 dpi. The spongy fluorescent structures were rather coarse and vesicular, as compared to *35Sp:DMP1-GFP* and seemed to nucleate from membrane patches with strong GFP signal intensity (Figure 36A).

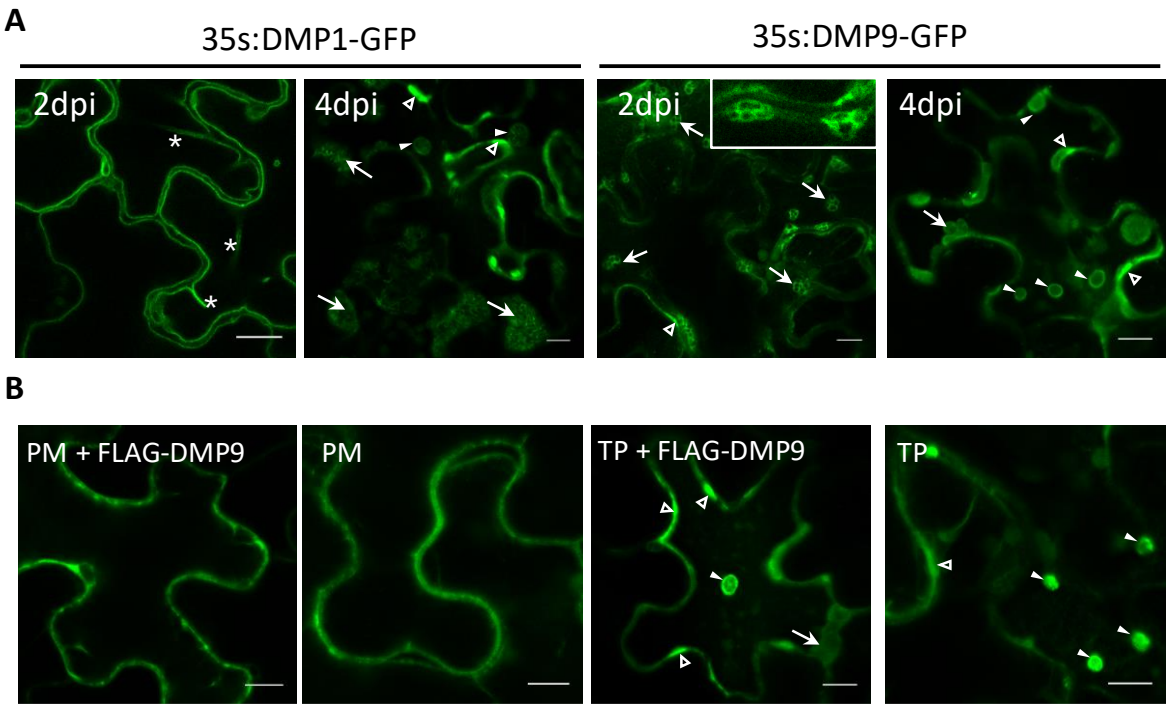


Figure 36 Transient overexpression of DMPs in *Nicotiana benthamiana* epidermis cells. A, Overexpression of C-terminal GFP fusion proteins by the CaMV 35S promoter. 2 days post infiltration (dpi) DMP1-GFP localized to the tonoplast with visible cytoplasmic strands (asterisks) penetrating the cytosol. 4 dpi spherical structures (arrowheads), patches with high GFP signal intensity (hollow arrowheads) and foamy structures became visible. Overexpression of DMP9-GFP showed a similar phenotype, but spheres and foamy structures appear earlier (2dpi). The foamy aggregates have a coarser and rather vesicular structure and seem to nucleate from membrane patches with higher signal intensity (inset in 2dpi). **B,** Co-expression of FLAG-DMP9 with organelle markers 3 dpi. Co-expression with the plasma membrane (PM) marker (*35Sp:pm-gk*) showed no alteration of plasma membrane shape or integrity, as compared to the control (PM marker only). Co-Expression of FLAG-DMP9 with a tonoplast (TP) marker (*35Sp:vac-gk*) revealed spheres, membrane patches and foamy structures, that also appeared when the tonoplast marker was infiltrated without FLAG-DMP9. Representative images are shown. Scale bars = 10 µm.

The similar appearance of DMP1-GFP and DMP9-GFP raised the question, if the signal observed in *35Sp:DMP9-GFP* infiltrated tobacco leaf epidermis cells localizes to the plasma membrane. To exclude DMP9 aggregation or mistargeting artifacts caused by the C-terminal GFP, co-infiltration experiments with N-terminally FLAG-tagged DMP9 (*35Sp:FLAG-DMP9*) and previously described plasma membrane (PM; *35Sp:pm-gk*) and tonoplast (TP; *35Sp:vac-gk*) markers (Nelson *et al.*, 2007) were performed (Figure 36B). No differences were observed between the samples co-infiltrated with *35Sp:FLAG-DMP9* and the PM marker and the marker alone. The signal uniformly decorated the periphery of the epidermal cells. Co-infiltration of FLAG-DMP9 and the TP marker resulted in bulbs and patches and spherical structures. However, a very similar pattern was also observed in the control, where only the tonoplast marker was infiltrated. These results suggest that the membrane remodeling events observed for *35Sp:DMP1-GFP* and for *35Sp:DMP9-GFP* are artifacts. The signal observed for DMP9-GFP is therefore likely not plasma membrane-

derived, but fusion of the bulky GFP-tag to the C-terminus of DMP9 leads to protein mistargeting to the vacuole. The membrane remodeling events observed for samples infiltrated with the TP marker only suggest that the spherical and foamy structures are artifacts of tonoplast-localized fusion proteins, possibly due to overexpression, aggregation or misfolding.

4.6.12 Expression and subcellular location of HAP2-YFP in *dmp8,9* sperm cells

The fusion defective phenotype observed for *dmp8,9* sperm cells was strongly reminiscent of sperm cells devoid of the fusogen HAP2, where sperm are also able to attach to the female gametes but fail to fuse (Mori *et al.*, 2014). This suggests a functional link between DMP8/9 and HAP2. To address the question whether HAP2 expression or subcellular localization is altered in sperm cells lacking functional *DMP8* and *DMP9,* the homozygous marker line *HAP2g:YFP* (Besser *et al.*, 2006) was crossed with *dmp8-10C_14 HTR10-mRFP.* Homozygous *dmp8,9* progeny with a similarly reduced seed set to that of the mother plant *dmp8-10C_14 HTR10-mRFP* was selected for analysis. Pollen tubes were germinated *in vitro* and HAP2-YFP signal intensity and subcellular distribution was monitored by CLSM (Figure 37).

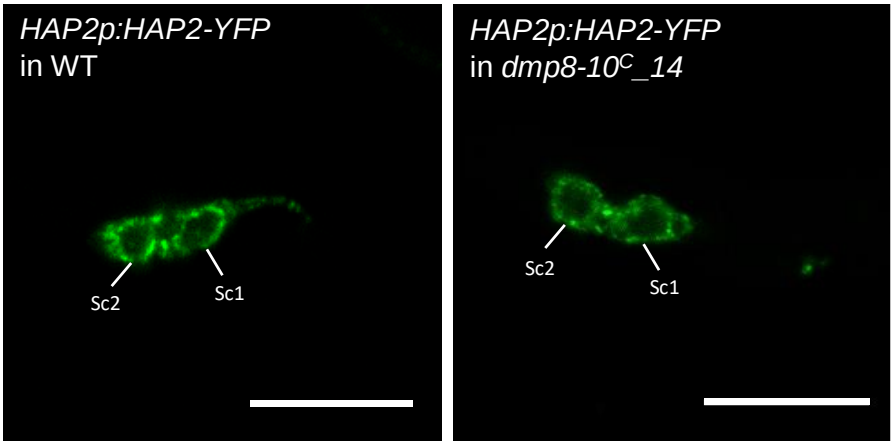


Figure 37 Expression and subcellular location of HAP2-YFP is not altered in sperm cells lacking functional *DMP8* and *DMP9*. Homozygous *HAP2p:HAP2-YFP* was crossed with *dmp8-10C_14*. Pollen was collected from F2 plants homozygous for *HAP2p:HAP2-YFP* and *dmp8-10C_14*, respectively, and germinated *in vitro*. For sperm cell imaging by CLSM, identical settings were applied for laser intensity and detection. Representative images are shown. HAP2-YFP was mainly detected in endomembrane compartments, both in the wild type (WT) and *dmp8-10C_14* background. Signal intensities of fluorescent HAP2-YFP were not noticeably altered (*n* = 30 sperm cells, each). Note that seed set in the F2 plant *HAP2p:HAP2-YFP dmp8-10C_14* used for imaging of sperm cells was similar to that of the mother plant *dmp8-10C_14* shown in Figure 23. Sc = sperm cell. Scale bars = 10 µm. Modified from Cyprys *et al.*, 2019.

HAP2-YFP was mainly detected in endomembrane compartments, both in the wild type and in the homozygous *dmp8-10C_14 HTR10-mRFP* background. Signal intensities and appearance of HAP2-YFP was very similar, suggesting that HAP2 expression and subcellular localization is not altered in the *dmp8-10C* mutant background, at least when the sperm cells are still within the pollen tube.

4.6.13 Delivery of *dmp8,9* sperm cells can trigger EC1 secretion

The role of egg cell-secreted EC1 proteins in sperm activation makes them key players ensuring successful double fertilization. While it has been shown that EC1-GFP is stored in vesicles within the egg cell in unfertilized ovules and gets secreted only upon sperm cell delivery (Sprunck *et al.*, 2012), the precise timepoint and stimulus for triggered EC1 secretion remained elusive.

To exclude that fusion defects observed for *dmp8,9* sperm cells are the consequence of failed EC1 secretion by the egg cell, leading to non-activated sperm cells with the fusogen HAP2 not exposed to the sperm surface, an assay to monitor EC secretion was established: A homozygous marker line for *EC1.1p:EC1.1-GFP* (Sprunck *et al.*, 2012) was pollinated with either control (*HTR10-mRFP*) pollen or *dmp8,9C_3 HTR10-mRFP* pollen. EC1-GFP secretion was assessed by observing GFP signal intensity and subcellular distribution within the egg cell or zygote in pistils dissected 18 HAP. At this timepoint, double fertilization is successfully accomplished in a wild type situation (Figure 38).

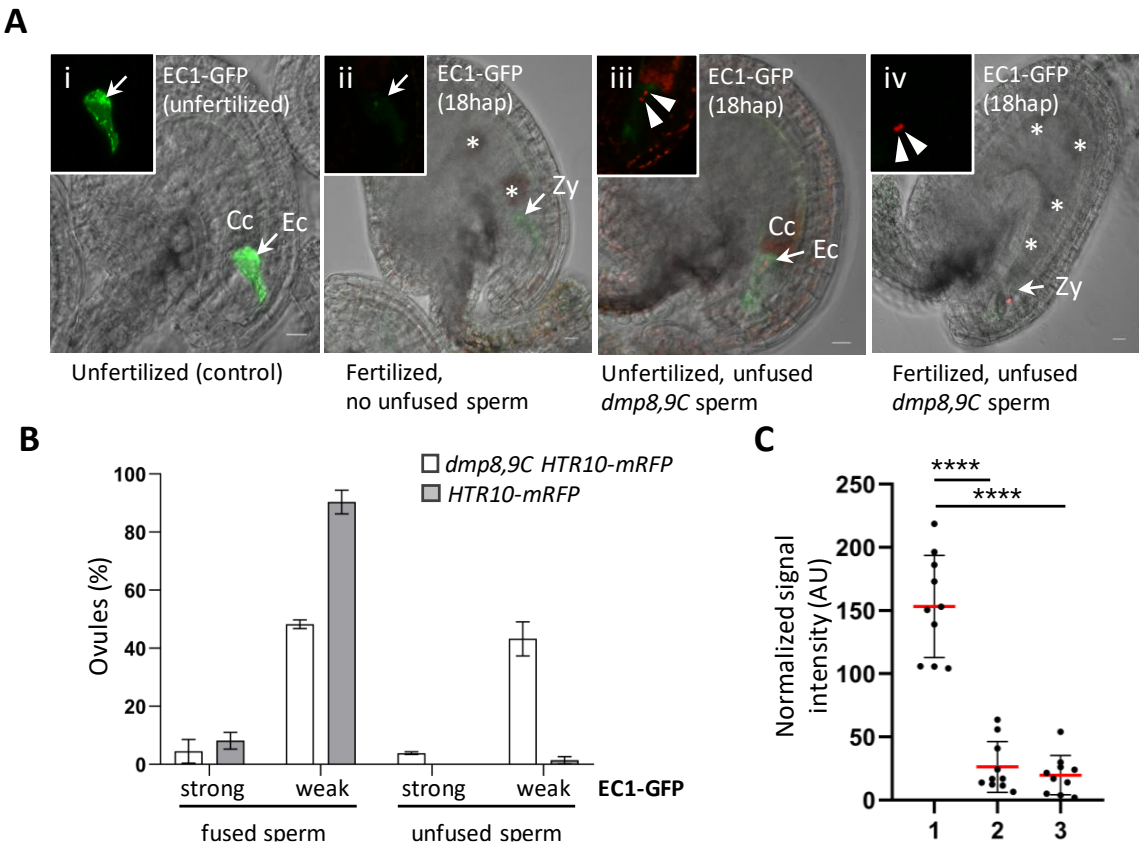


Figure 38 EC1-GFP secretion is also triggered by unfused *dmp8 dmp9* sperm cells. A, Representative images of egg cells and zygotes expressing EC1-GFP before and after fertilization, imaged by confocal laser scanning microscopy (CLSM). (i) EC1-GFP in ovule before fertilization. (ii) to (iv) EC1-GFP signals 18 HAP with *dmp8,9C HTR10-mRFP* pollen. EC1-GFP signals are much weaker and more diffuse in the zygote (ii), as well as in egg cells with adjacent unfused *dmp8,9* sperm cells (iii, iv). For all experiments, identical settings were used for microscope, fluorophore excitation and detection. **B,** Quantitative assessment of EC1-GFP fluorescence, 18 HAP with *dmp8,9C HTR10-mRFP* pollen or *HTR10-mRFP* pollen (control). Data are the means ± s.d. of n = 3 independent experiments and a total number of 178 ovules targeted by *dmp8,9C HTR10-mRFP* and 132 ovules targeted by *HTR10-mRFP pollen,* respectively. **C,** Quantification of relative EC1-GFP fluorescence intensity in individual egg cells, zygotes and egg cells which failed to fuse with *dmp8,9* sperm cells. Compared to unfertilized egg cells (1), EC1-GFP intensities are significantly reduced in zygotes (2) and in egg cells with adjacent unfused *dmp8,9C* sperm cells (3). Data are the means ± s.d. of n = 10 cells per group. Significant differences (two-sided Student's t-test) are indicated (****: P < 0.0001). Normalized signal intensity was calculated for a single optical section through the center of the cell, by dividing the fluorescence intensity measured for the cell divided by its area. Scale bars = 10 µm.

Unfertilized EC1-GFP expressing ovules, dissected from emasculated pistils 48 hours after emasculation served as a reference. As expected, strong granular GFP signal within the egg cell was observed in unfertilized egg cells (Figure 38A(i)). After hand pollination with *dmp8,9C_3 HTR10-mRFP* mutant pollen, different phenotypes were observed: In 48.3 ± 1.5 % of ovules EC1 was secreted and double fertilization was completed successfully, as indicated by dividing endosperm nuclei and a visible zygote. While some residual EC1-GFP signal was detectable in developing zygotes, it was much weaker and diffuse and mainly localized to the endoplasmic reticulum surrounding the zygote nucleus (Figure 38A(ii)).

43.3 % ± 5.9 % of ovules with unfused *dmp8,9C* sperm cells also showed a rather weak and diffuse GFP signal within the unfertilized egg cell, indicating successful EC1-GFP secretion (Figure 38A(iii)). Furthermore, ovules with an unfused *dmp8,9C* sperm pair but visible zygote and proliferating endosperm nuclei were observed, suggesting that the sperm pair delivered by a first pollen tube failed to fertilize the female gametes, whereas the sperm pair transported by the second pollen tube fused successfully (Figure 38A(iv)). GFP signal in those ovules was also weak and diffuse. In pistils pollinated with control (*HTR10-mRFP*) pollen 90.4 ± 4.1 % of the ovules were successfully double fertilized and showed weak residual EC1-GFP signal in the early zygote.

These results suggest that EC1-GFP is also secreted when fusion-defective *dmp8,9* sperm cells are delivered to the ovule. To asses EC1 secretion in a quantitative fashion, GFP signal intensity in individual egg cells with adjacent unfused *dmp8,9* sperm cells and in developing zygotes was quantified via ImageJ (Figure 38C). Normalized signal intensity was calculated for a single optical section through the center of the cell, by dividing the fluorescence intensity measured for the cell divided by its area. Unfertilized egg cells from emasculated pistils served as control. It became obvious that relative EC1-GFP fluorescence intensity in zygotes and egg cells with unfused *dmp8,9* sperm cells was significantly lower than in egg cells from unpollinated pistils, confirming that EC1 secretion by the egg cell can be triggered by fusion defective *dmp8,9* sperm cells.

4.6.14 Tetraspanin 9 does not accumulate at the egg-sperm interface

Tetraspanins (TETs) are a superfamily of evolutionary conserved integral membrane proteins, characterized by 4 transmembrane α-helices. Despite their similar topology to DMPs, plant TETs are characterized by larger extracellular loops 1 and 3 which get stabilized by disulfide bonds formed by highly conserved cysteine residues (Boavida *et al.*, 2013). A pivotal role for two egg-expressed TETs, CD9 and CD81 in mammalian fertilization has been demonstrated (see 3.6.1), and accordant with their role in organization of membrane microdomains and signaling complexes, local accumulation of CD9 at the egg-sperm interface has been observed prior to fusion (Ravaux *et al.*, 2018). The *Arabidopsis thaliana* genome codes for 12 *TETs* and a study on their expression and subcellular localization among reproductive tissues has identified *TET9* as plasma membrane-localized protein in the egg and central cell (Boavida *et al.*, 2013).

To investigate whether egg-expressed TET9 also accumulates at the site of egg-sperm fusion, a homozygous *TET9g-GFP* reporter line (provided by PD Dr. Stefanie Sprunck) was used to pollinate with *HTR10-mRFP* amd *dmp8,9C_3 HTR10-mRFP* mutant pollen. Ovules were dissected from pollinated pistils 8 HAP and 18 HAP and subcellular localization of TET9-GFP was analyzed in ovules with unfused *dmp8,9C_3 HTR10-mRFP* sperm cells by CLSM (Figure 39).

In unfertilized ovules, TET9-GFP was uniformly distributed and clearly labeled the plasma membranes of egg cell and central cell (Figure 39A). 8 hours after pollination with *dmp8,9C_3 HTR10-mRFP* pollen, unfused sperm cells were frequently observed adjacent to the egg cell. However, TET9-GFP signal uniformly decorated the egg cell and signal accumulation at the sperm-egg contact site was never observed (Figure 39B). On the contrary, 18 hours after pollination it seemed that TET9-GFP is depleted from the attachment site, as compared to the lateral parts and the chalazal pole of the egg cell.

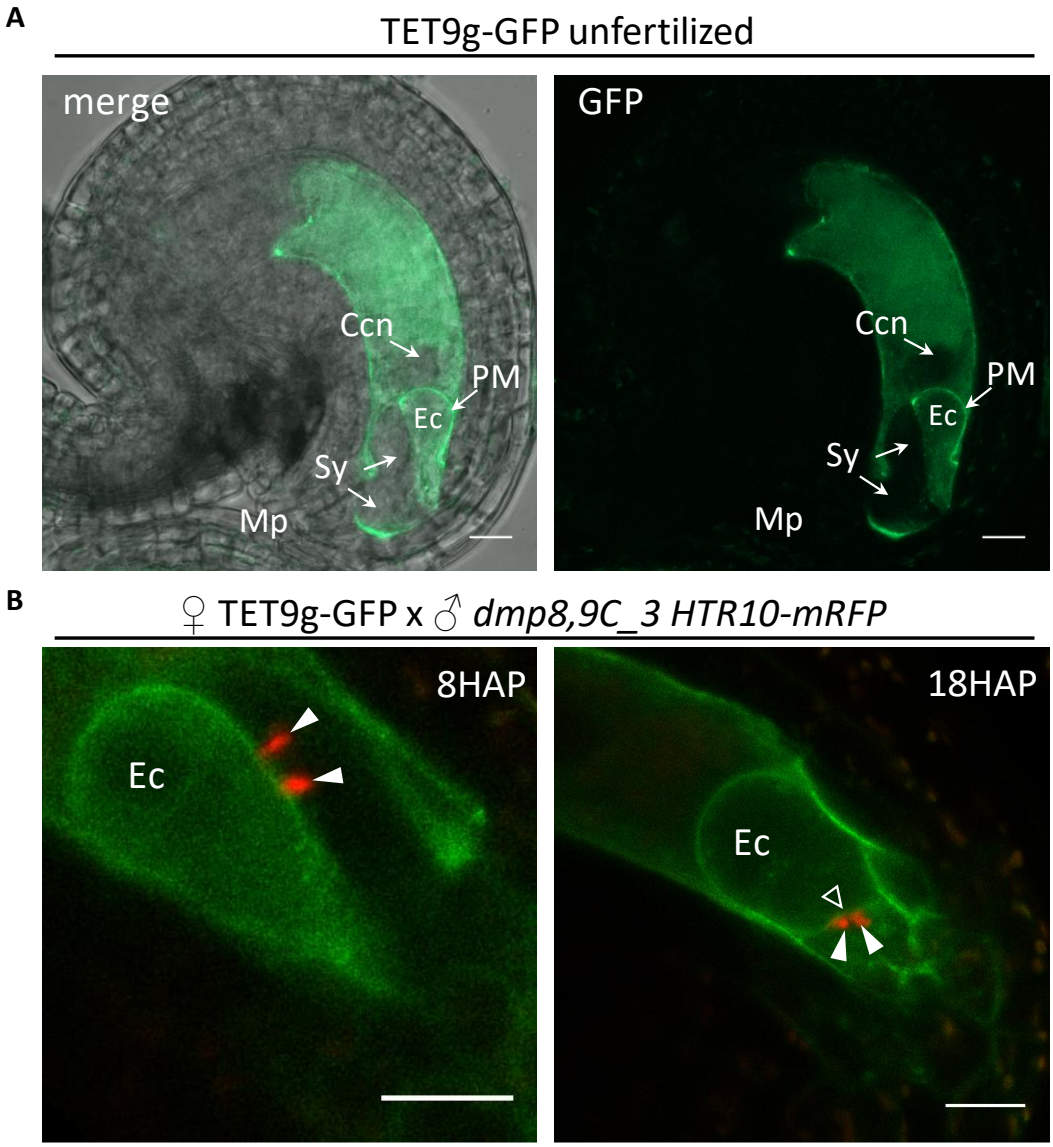


Figure 39 TET9-GFP dynamics at the sperm-egg adhesion site. A, In unfertilized ovules TET9-GFP is uniformly distributed at the plasma membrane (PM) of the egg cell (Ec). **B,** After pollination with mutant *dmp8,9C_3 HTR10-mRFP* pollen unfused sperm cells (arrowheads) were frequently observed adjacent to the egg cell. 8 hours after pollination (HAP) TET9-GFP was still uniformly distributed at the egg plasma membrane and signal accumulation at the sperm adhesion site was not detectable. 18 HAP TET9-GFP signal was depleted from the attachment site (hollow arrowhead). Ccn = central cell nucleus; Syn = synergid; Mp = micropyle; Scale bars = 10 µm.

5. Discussion

With only three players known to date, our knowledge on proteins mediating gamete interactions during double fertilization is scarce. This work aimed at identifying novel sperm-expressed membrane proteins involved in this angiosperm-specific process. Plant gamete transcriptomes and proteomes reflect the molecular character of this special cell types and can serve as useful point of origin for the identification of such proteins. However, the generation of such datasets requires to isolate the respective cell type with sufficient quantity and purity.

The female gametes of flowering plants are deeply embedded within the maternal tissue and therefore notoriously difficult to isolate (Englhart *et al.*, 2017; Flores-Tornero *et al.*, 2019). In contrast, pollen is accessible much easier and therefore a multitude of studies on sperm cell transcriptomics have been performed to so far: Sperm transcriptome data from species such as *Nicotiana tabacum*, *Oryza sativa*, *Plumbago zeylanica*, *Arabidopsis thaliana* and maize is available (Borges *et al.*, 2008; Gou *et al.*, 2009; Xin *et al.*, 2011; Russell *et al.*, 2012; Russell *et al.*, 2017; Chen *et al.*, 2017). However, these studies used semi-purified sperm cells that did not account for vegetative cell-derived contaminant transcripts. A more comprehensive study utilized fluorescence activated cell sorting (FACS) to purify maize sperm cells and generate a sperm EST database containing approximately 1,100 transcripts (Engel *et al.*, 2003). This database was employed in a cross-species comparative study which led to the identification of two sperm cell-expressed homologues in *Arabidopsis*, GEX1 and GEX2 (Engel *et al.*, 2005). A comparative transcriptomics study between *Arabidopsis thaliana* pollen, sperm cells and sporophytic tissue utilizing microarray technology showed that sperm cells have a distinct and diverse transcriptional profile (Borges *et al.*, 2008; Borges *et al.*, 2012). Modern and sensitive methods for transcriptome profiling (e.g. RNAseq) have enabled us to study transcriptomes at single-cell resolution, which was recently reported for *Arabidopsis thaliana* sperm cells (Misra *et al.*, 2019).

Compared to transcriptomic studies, proteomic approaches display the advantage that only translated mRNA sequences, which more accurately reflect the functional character of the cells, are revealed. Proteomic studies on *Arabidopsis thaliana* pollen grains are readily available, including the *Arabidopsis* pollen phosphoproteome (Holmes-Davis *et al.*, 2005; Noir *et al.*, 2005; Sheoran *et al.*, 2006; Mayank *et al.*, 2012; Mergner *et al.*, 2020). Nevertheless, these studies often employed low-throughput methods e.g. 2D electrophoresis followed by MS analysis of single protein spots. Proteomic data on plant

sperm cells is scarce as only two studies utilizing isolated rice and lily gametes exist (Abiko *et al.*, 2013; Takahashi *et al.*, 2018).

It would be interesting to compare the maize sperm proteomic datasets generated in this work with published sperm transcriptome and proteome data. Such analyses were not conducted as the focus of this work laid on functional characterization of manually selected candidates. Comparative analysis would require a sophisticated bioinformatics pipeline e.g. to annotate the older maize data (Engel *et al.*, 2003; Chen *et al.*, 2017) according to the newest maize reference genome assembly (AGPv5). Furthermore, computing of putative maize orthologs would be required to allow comparison of datasets across different species.

Obtaining sufficient amounts of pure *Arabidopsis thaliana* sperm cells is a huge obstacle and one of the main reasons why maize was used for cell isolation and proteomics studies in this work. The isolation procedure described here is rapid, easy to perform with regular molecular laboratory equipment and yields relatively pure and viable sperm cells. As proteins involved in direct gamete recognition, adhesion and fusion are thought to localize to the sperm plasma membrane, sperm cell total proteins (SC) but also membrane protein-enriched sperm microsomal fractions (SCM) were measured by LC-MS/MS, thereby increasing likelihood to capture low abundance proteins of high biological relevance. Depletion strategies such as pre-fractionation have been described as efficient method to increase proteome coverage as ion signals from low abundance analytes are often "masked" by high abundance ions of similar *m/z* and retention time (Keshishian *et al.*, 2007; Filip *et al.*, 2015). However, comparison of the SC and SCM datasets revealed a very high overlap with 88.9 % of all core protein groups present in both datasets.

Although Western Blotting suggested that sperm microsomal fractions are enriched in plasma membrane proteins, analysis of the SCM dataset revealed that a surprisingly low number (21.3 %) of all SCM proteins are predicted to be membrane-localized or -associated. Prediction of subcellular localization showed that the majority of SCM proteins are predicted to localize to plastids, the nucleus or the cytosol while localization to the plasma membrane or endomembrane compartments was predicted for only 5.2 % of all SCM proteins. In a label-free quantification approach, only 148 protein groups were found to be significantly enriched in the SCM dataset, as compared to the SC dataset. However, these were also predicted to localize mainly to the nucleus, plastids or the cytosol and none of the known sperm cell membrane proteins were found in this group. In the future, more refined membrane protein enrichment procedures should be considered. Differential centrifugation could be applied to obtain samples depleted of mitochondrial, plastidial and nuclear proteins (Nagahashi and Hiraike, 1982; Graham,

2002) which accounted for the majority of proteins identified in the SCM dataset and can be considered as contaminant. Specific isolation procedures for plasma membrane proteins have also been described in the literature: aqueous two phase partitioning separates membrane particles derived from different cellular membrane systems, based on their partition between two immiscible aqueous polymers, usually in a PEG-Dextran system and has been applied to isolate plant plasma membrane proteins (Lund and Fuglsang, 2012). This method requires a relatively large amount of starting material but could become feasible with the bulk sperm isolation procedure described in this work. Comparative proteomics after protease treatment of sperm cells in order to remove surface proteins, as compared to untreated sperm cells, could serve as alternative strategy to identify truly surface-localized sperm proteins. However, it should be noted that besides proteins localizing to the sperm surface, endomembrane-resident proteins are also of interest for direct gamete interactions, as demonstrated for HAP2, which mainly localizes to sperm endomembranes before sperm activation (Sprunck *et al.*, 2012).

Although maize was used as a model system to generate complimentary sperm transcriptomic and proteomic datasets, functional studies on selected candidates were carried out in *Arabidopsis thaliana* due to the significantly shorter generation time, well defined reproductive morphology and because protocols and marker lines to dissect the process of double fertilization in this model species are well established.

5.1 Five *Arabidopsis thaliana* candidates are sperm-expressed membrane proteins

Identifying true functional orthologs between two species can be challenging, especially for larger gene families with high amino acid sequence similarity. Combining sequence-based data such as primary sequence or protein domain architecture and publicly available transcriptomic (microarray) data in the Genevestigator software was used to identify some of the sperm-expressed *Arabidopsis thaliana* proteins and putative orthologs of maize candidates selected from the sperm proteome. As intronic sequences have been shown to play a role in transcriptional regulation (Rose *et al.*, 2008; Karve *et al.*, 2011), translational reporters using genomic sequences were chosen to investigate *Arabidopsis thaliana* sperm cell candidates. The C-terminal GFP fusion proteins also allowed to gain first insights on subcellular localization of the candidate protein.

Translational reporter lines for 7 *Arabidopsis thaliana* candidates did not show the expected signal in sperm cells, suggesting that these candidates are either not expressed in this cell type, that the GFP reporter signal is below the detection limit or that fusion of the GFP reporter to the C-terminus of the candidate protein resulted in aberrant

expression. Generation of enzymatic promoter:reporter lines (e.g. promoter:GUS) which are considered more sensitive than translational reporters, could prove useful to identify weak transcriptional activity in sperm cells for those candidates. GUS enzymatic activity as well as the chromogenic enzyme product are stable over a long period of time, leading to superior sensitivity. However, in the particular case of *Arabidopsis thaliana* sperm cells the GUS reporter seems to be unfavorable due to limited spatial resolution: in transgenic *DMP9p:GUS* lines generated by Kasaras and Kunze (2010), GUS signal was observed in the whole pollen grain, although DMP9 is a sperm cell-specific protein (Takahashi *et al.*, 2018; Cyprys *et al.*, 2019), most likely due to diffusion of the blue product from the sperm cells to the vegetative cell cytoplasm. Transcription and translation of a certain mRNA can occur in different cells or tissues, as demonstrated by cell-to-cell mobile mRNAs (Thieme *et al.*, 2015). RNA movement within the cell-in-a-cell structure of a mature pollen grain is possible, as shown by directional movement of transcripts and small interfering RNAs from the pollen vegetative cell to the sperm cells (Jiang *et al.*, 2015; Slotkin *et al.*, 2009).

It has to be noted that while the current *Arabidopsis thaliana* genome assembly (TAIR10) counts 27,206 protein-coding genes (Lamesch *et al.*, 2012), only 23,750 genes are covered by the Affymetrix ATH1 GeneChip® (Redman *et al.*, 2004). It is therefore possible that sperm-expressed *Arabidopsis thaliana* orthologs that are not present on the ATH1 chip evaded analysis. In general, expression profiling of whole gene families by (semi quantitative) RT-PCR would be helpful to narrow down a larger gene family to a smaller circle of candidates with a sperm-specific expression pattern.

Translational reporter lines for five candidates (DMP8, DMP9, LFG4, RLK2, PCR11) showed signal in sperm cells. Those proteins were predicted to localize to the plasma membrane, which matched well with the observed GFP signal at the sperm periphery and at the plasma membrane projection tethering one sperm cell to the vegetative nucleus. Although LFG4 was predicted as tonoplast-localized, the signal distribution of LFG4-GFP at the outline of the two sperm cells strongly resembled the other four candidates. However, the morphology of cellular compartments and organelles in *Arabidopsis thaliana* sperm cells is not very well defined, which is also due to the small size of the cells. Transient expression of fluorescent fusion proteins in *Nicotiana benthamiana* leaf epidermis cells is widely used for subcellular localization studies in plant molecular biology (Sparkes *et al.*, 2006), due to the large size of the cells and because multicolored fluorescent markers for diverse organelles and subcellular compartments are available (Runions *et al.*, 2007; Nelson *et al.*, 2007). In future experiments these could be used to confirm plasma membrane localization of the sperm candidates. Though, the strong promoters used in such a system and the attachment of a bulky fluorescent tag can lead to protein mistargeting,

aggregation and faulty subcellular localization *in planta* (Gao *et al.*, 2012; Brkljacic *et al.*, 2009). Therefore, the co-expression of marker proteins for subcellular compartments in sperm cells is preferable and should be considered in future experiments.

Interestingly, aberrant protein targeting also was observed when eGFP was fused to either terminus of DMP1 (Kasaras and Kunze, 2017). DMP1-GFP is artificially retained in the ER due to occlusion of a putative ER export signal at the C-terminus of DMP1. In contrast GFP-DMP1 mislocalizes to the plasma membrane because the bulky tag masks a basic tonoplast-targeting motif located at the N-terminus, which is required. While highlighting the risk of misinterpreting subcellular localization results obtained from translational fusions with bulky tags, protein mistargeting could also offer an explanation for the failed complementation of the *dmp8,9* mutants by a construct expressing *DMP9g:GFP*.

5.2 The membrane-rich sperm-sperm connection: a signaling hub during double fertilization?

Remarkably, translational reporter lines for the sperm-expressed candidates LFG4, RLK2, PCR11 and DMP8/9 showed GFP signal accumulation at the boundary of the two sperm cells, which became apparent during CLSM analysis of *in vitro* germinated pollen tubes, in which the male germ unit becomes better visible as it is transported in a stretched-out configuration with the vegetative nucleus preceding the two sperm cells. Protein enrichment at this sperm-sperm contact site was also observed for other sperm-expressed membrane proteins such as Tetraspanins TET11 and TET12 (Boavida *et al.*, 2013) and GLUTAMATE RECEPTOR 3.3 (Wudick *et al.*, 2018). Boavida *et al.* (2013) suggested the concept of a sperm-specific Tetraspanin-enriched microdomain between the two sperm cells. Tetraspanins are important regulators of cellular adhesion and signaling, and function as molecular scaffolds, thereby controlling the spatial organization of membrane proteins in large networks called the Tetraspanin web (Termini and Gillette, 2017; van Deventer *et al.*, 2017). It is therefore possible that this microdomain functions in keeping the two sperm cells physically connected while they travel towards their destination within the pollen tube, in sperm-sperm communication or in sperm separation after they have been delivered into the female gametophyte, thereby ensuring efficient fertilization of the two female gametes by one sperm cell each.

The observation that the adhesion factor GEX2 (Mori *et al.*, 2014) and the fusogen HAP2 (Mori *et al.*, 2010; Sprunck *et al.*, 2012) do not accumulate at the sperm-sperm contact site supports the concept of a microdomain that is enriched only in a specific subset of sperm membrane proteins. Stimulation of sperm pairs released from *in vitro* germinated

pollen tubes with EC1 peptides covering the conserved EC1 motifs led to a depletion of TET9-GFP from the sperm-sperm interface and two separated sperm cells became visible (Lindemeier, 2017), highlighting the dynamics of this contact site and suggesting a role of EC1 proteins in physical unpairing of the sperm cells.

Electron micrography of the *Plumbago zeylanica* male germ unit followed by 3D reconstitution (Russell, 1984) revealed membrane-rich invaginations at the sperm-sperm interface, remarkably similar to those observed at the filiform apparatus formed by the synergid cells. It is therefore conceivable that GFP signal accumulation at the sperm connection is a mere consequence of high plasma membrane density in this area. Consequently, ectopic expression of *Mus musculus* Lyn24-GFP, comprising a short 24 amino acid fragment of the murine *Src*-family tyrosine kinase Lyn with two lipidation sites in *Arabidopsis thaliana* sperm cells showed that Lyn24-GFP signal finely decorated the sperm plasma membrane but signal also accumulated at the sperm interface, where fine threads of curved membranes became apparent (Wudick *et al.*, 2018).

Whether the concept of a membrane microdomain between the two sperm cells holds true or not, identifying the exact function of proteins that specifically accumulate in this region (including PCR11, LFG4 and RLK2) will foster our understanding of its biological significance as an attachment site or as a signaling hub during double fertilization.

5.3 Functional redundancy among gene family members: a reason for the lack of phenotypes?

Sperm-specific expression and sperm membrane localization made PCR11, LFG4 and RLK2 prime candidates for loss-of-function studies. However, genome edited homozygous mutants obtained by CRISPR/Cas9 did not show an obvious fertilization-related phenotype. The seed set in these mutants was comparable to the wild type, although it can´t be guaranteed that the CRISPR/Cas9 induced indels led to a full knockout of the respective gene. Transcripts derived from CRISPR/Cas9-induced frameshift mutations that lead to formation of premature stop codons are considered to be recognized and degraded by the nonsense-mediated decay machinery, resulting in gene ablation. While CRISPR/Cas9 bears enormous potential for reverse genetics, recent reports have demonstrated formation of truncated transcript and protein after genome editing, due to post-transcriptional and -translational effects such as internal ribosomal entry, exon skipping or altered mRNA processing, resulting in transcripts that escape nonsense-mediated decay (Tuladhar *et al.*, 2019; Anderson *et al.*, 2017). A recent study reported genetic compensation, leading to differential regulation of related genes or pathway

members to compensate for the targeted loss of a specific gene after genome editing (El-Brolosy *et al.*, 2019). Careful evaluation of genome edited mutants on transcript and protein level should be performed to ensure the CRISPR/Cas9-induced indels resulted in a gene knockout.

While feasible for single-copy genes, reverse genetics approaches can be furthermore hampered by functional redundancy when studying members of larger gene families such as *LFG4, PCR11* and *RLK2*. *LFG4* encodes a protein belonging to the conserved BAX INHIBITOR 1-RELATED (BI-1) protein family (IPR006214). The *Arabidopsis thaliana* genome codes for 22 BI-1 motif-containing proteins, exhibiting strong sequence diversity, five of which have been named AtLFG1-AtLFG5 due to sequence homology to *HsLFG*, the human BI-1 member with the best-annotated function (Hu *et al.*, 2009; Weis *et al.*, 2013). BI-1 proteins are plasma membrane- or ER-resident seven-transmembrane proteins and have been identified as cell death suppressors in response to stress in animals and plants (Ishikawa *et al.*, 2011; Hückelhoven, 2004; Eichmann *et al.*, 2010). A role for LFG cell death inhibitors in sperm cells is not known and might be related to the indirect modulation of Ca^{2+} flux from the ER, as shown for BI-1 proteins from mice (Sano *et al.*, 2012). Furthermore, murine BI-1 proteins have been shown to function as a pH-sensitive Ca^{2+}-sensitive channel or pore (Bultynck *et al.*, 2012; Bultynck *et al.*, 2014). Although the search in publicly available microarray data revealed that the other four members of the *AtLFG* family show comparably low expression in sperm cells (Supplemental figure 3), the lack of a phenotype in *lfg4* mutants might nevertheless be explained by functional redundancies among the five LFG proteins or one of the other 17 BI-1 motif containing proteins in *Arabidopsis thaliana*.

Functional redundancy can also be considered for PCR11. PCR11-GFP localized to the sperm membrane, however observed signal was more granular than for the other sperm-expressed candidates described in this work and sperm-localized vesicle-like structures became apparent. PCR11 belongs to the PLACENTA-SPECIFIC 8 (PLAC8) superfamily of cysteine-rich membrane proteins, named after its initial discovery in the transcriptome of murine placenta (Galaviz-Hernandez *et al.*, 2003). In animals, PLAC8 domain containing proteins seem to participate in a multitude of cellular processes, including pathogen defense, proliferation, differentiation and apoptosis (Cabreira-Cagliari *et al.*, 2018). In plants, PLAC8 containing proteins can be divided into three classes (reviewed by Song *et al.*, 2011b). The first class (FWL proteins) modulate fruit size by negative regulation of cell division and the second class (MCA1/2) function as mechanosensitive Ca^{2+} channels. The third class of plasma membrane-localized PLAC8-containing plant proteins have putative roles in Cadmium and Zinc detoxification, hence the name PLANT CADMIUM RESISTANCE

(Song *et al.*, 2004; Song *et al.*, 2010). The identification of *PCR11* as a direct target of the transcription factor DUO1 (Borg *et al.*, 2011), a master regulator of sperm cell differentiation and specification, implies sperm cell-specific roles for PCR11, possibly in transmembrane ion transport. Notably, *Brassica juncea* PCR1 was identified as Ca^{2+} efflux transporter, and structural homology with the Mg^{2+}-transporter CorA suggests a general role of PCRs in ion transport across biological membranes, possibly by forming pores composed of homo-oligomers (Song *et al.*, 2011a; Guo *et al.*, 2010). In female gametophytic cells, Ca^{2+} signatures are known to correlate with pollen tube burst and plasmogamy (Denninger *et al.*, 2014; Hamamura *et al.*, 2014) but Ca^{2+} oscillations in sperm cells have not yet been investigated.

No fertilization-related phenotype was observed in the *pcr11* mutants generated in this work. As PLAC8 containing FWL proteins regulate cell density, organ size as well as fruit weight (Frary *et al.*, 2000; Libault *et al.*, 2010; Libault and Stacey, 2010), *pcr11* mutants were subjected to phenotyping. However, compared to the wild type, no difference in leaf size, plant height, silique size or seed number per silique was observed (Supplemental figure 4). To further address a fertilization-related function of PCR11 the generation of new *pcr11* gene-edited plant lines using sgRNAs which are located at the 5´ end of the gene are recommended. Furthermore, higher order mutants including *PCR4* and *PCR12* should be generated, since analysis of microarray data showed that *PCR4* and *PCR12* are also sperm-expressed (Supplemental figure 5).

The third interesting gene-edited candidate without a noticeable phenotype in seed set was a Leucine-rich repeat Receptor-like kinase (LRR-RLK) of unknown function termed RLK2. RLKs are single pass transmembrane proteins and form the largest class of plant cell surface receptors that have been implicated in a wide range of signal transduction pathways in growth, development and defense (Walker, 1993). Despite a common architecture characterized by a single transmembrane domain and a conserved cytoplasmic serine/threonine kinase domain. The extracellular domains of plant RLKs is involved in signal perception are and very diverse (Shiu and Bleecker, 2001). The *Arabidopsis thaliana* genome codes for 610 RLKs and 223 from those belong to the subfamily of extracellular Leucine-Rich Repeats (LRR)-containing RLKs (LRR-RLKs), making LRR-RLKs the largest subfamily of plant RLKs (Shiu *et al.*, 2004; Wu *et al.*, 2016).

LRR-RLKs can be subdivided on the basis of their LRR organization, and RLK2 belongs to class “LRR-VI” that is characterized by 7 tandem LRR-repeats in their extracellular domain (Shiu and Bleecker, 2001). 11 “LRR-VI”-class RLKs are present in the genome of *Arabidopsis* thaliana (Wu *et al.*, 2016). The only two functionally characterized LRR-VI members are MALE DISCOVERER1 (MDIS1) and MALE DISCOVERER2/MORPHOGENESIS OF ROOT HAIR 1

(MDIS2/MRH1). MDIS1 and MDIS2 play a role in sensing of LURE1 pollen tube attractants and in root hair elongation during tip growth (Wang *et al.*, 2016; Sklodowski *et al.*, 2017). Expression analysis of all 11 "LRR-VI"-class RLKs using the GENEVESTIGATOR software showed that apart from *MDIS1*, *MDIS2/MRH1* and *RLK2,* none of the *Arabidopsis thaliana* LRR-VI RLKs show considerable expression in sperm (Supplemental figure 6). Nevertheless, the expression other RLK2-related RLKs in sperm cells remain to be validated and multiplex genome-editing of functionally redundant sperm-expressed RLKs might be necessary to obtain a fertilization-related phenotype.

5.4 DMP8 and DMP9 are novel fertilization regulators

Functional redundancy was also evident for the two sperm-expressed DUF679 membrane proteins DMP8 and DMP9, which were identified as novel players in *Arabidopsis thaliana* gamete interactions. The fertilization-related role for DMP8 and DMP9 was clearly attributed to a step after gamete attachment, with a greater impact on egg cell fertilization than on central cell fertilization. Nevertheless, several aspects, including the molecular mechanism by which DMP8/9 execute their fusion-promotive action remain to be addressed in the future. Siliques from *dmp8,9* mutants still show a residual seed set, indicating that despite frequent fusion failure, double fertilization was successfully completed in around 40 % of all ovules. By contrast, fusion-defective homozygous *hap2* mutants are completely sterile due to the lack of paternal transmission (Besser *et al.*, 2006; Mori *et al.*, 2006) and mutant lines must be maintained in a heterozygous state, or conditional knockout strategies have to be employed (Nagahara *et al.*, 2015). Detection of *DMP7* transcript in pollen via RT-PCR suggested that a *dmp7,8,9* triple mutant might be required to completely block male transmission. However, subsequent analysis of a fluorescent reporter line *DMP7g:GFP* revealed weak GFP fluorescence in the cytoplasm of *in vitro* germinated pollen tubes, but no signal in sperm cells. An involvement of DMP7 in gamete interaction is therefore unlikely, which was supported by generation of a *dmp8,9C* mutant in the homozygous *dmp7-1* T-DNA background, showing no further reduction in seed set (not shown).

A *dmp8,*9-like reduction of fertilization rates was observed for sperm cells lacking the adhesion factor GEX2: despite being a single-copy gene, double fertilization failed in only about 50 % of the ovules in *gex2-1* mutants (Mori *et al.*, 2014). Those results indicate the existence of attachment- and fusion-promotive factors other than GEX2, DMP8 and DMP9, and suggest that similar to animals, angiosperm gamete interactions are orchestrated by a complex repertoire of membrane-expressed proteins. Although not discovered yet, it is

conceivable that proteins present on the surface of the female gametes also contribute to gamete attachment and gamete fusion, for example as interaction partners for GEX2 and DMP8/9, respectively.

Secondary effects after CRISPR/Cas9 genome editing of *DMP8/9* can also be considered. Notably, all *Arabidopsis thaliana DMPs* except *DMP7* are single exon genes. Transcripts with a premature stop codon after genome editing could therefore possibly evade the nonsense-mediated decay machinery, leading to the formation of truncated proteins that are still functional or may interfere with endogenous DMP function. Furthermore, a genetic compensation mechanism, leading to differential regulation of related genes or pathway members is conceivable (El-Brolosy *et al.*, 2019).

Phylogenic analysis of DMPs revealed their almost exclusive appearance in the *Viridiplantae* and their expansion specifically in the angiosperm lineage, suggesting a role for DMPs in plant-specific processes. Indeed, DMP1 was initially discovered in a genome-wide expression analysis of natural and induced leaf senescence in *Arabidopsis thaliana*, and was later found to be co-expressed with the senescence marker SENESCENCE-ASSOCIATED GENE 12 *SAG12* (van der Graaff *et al.*, 2006; Kasaras and Kunze, 2010).

To address the degree of evolutionary conservation among DUF679-containing membrane proteins, an attempt to complement the *dmp8,9* mutant with the most ancient *DMP* (*CrDMP*) and a *DMP* from *Amborella trichopoda* (*AmTrDMP*), the sister species to all angiosperms, was pursued. Although transcriptional upregulation of *CrDMP* after gamete activation in *Chlamydomonas* (Ning *et al.*, 2013) suggested an evolutionary conserved function of DMPs in gametes across species, *CrDMP* failed to rescue the fertilization-defective *dmp8,9* phenotype. By contrast, mutants complemented with *AmTrDMP* showed a slight but significant increase in seeds per silique, indicating that *AmTrDMP* can at least partially account for the loss of *AtDMP8/9*.

As compared to *Arabidopsis* DMP8/9, CrDMP shows an extended C-terminus and low sequence similarity, besides the highly conserved transmembrane α-helices, indicating a lineage-specific evolutionary diversification of DMPs. Furthermore, a fifth transmembrane helix is predicted for CrDMP. An odd number of transmembrane helices would lead to a different topology from *Arabidopsis* DMPs, with the N- and C-termini of CrDMP facing opposing sides of the membrane and might explain the failed complementation by *CrDMP*. Considering the strong diversification of *DMPs* in angiosperms, a scenario in which DMPs of the basal lineages of *Viridiplantae* have a more generalized function seems likely. During evolution of flowering plants polyploidization and gene duplications occurred, offering the possibility to evolve *DMP* paralogues. It is widely accepted that the reduced selection pressure on the newly duplicated gene copy allows accumulation of neutral mutations

leading to novel functions (Ohno, 1970). The expansion of *DMPs* in the flowering plant lineage could have therefore allowed them to fulfil novel, specialized and angiosperm-specific roles related to the process of double fertilization.

As a general biochemical function of DMPs remains elusive, and the molecular mechanism by which AtDMP8 and AtDMP9 execute their fusion-promotive activity is not yet understood, it is difficult to interpret the partial complementation of the *dmp8,9* mutant by *AmTrDMP*. One explanation may be the co-evolution of *DMPs* with putative interaction partners involved in the fertilization process. The much smaller evolutionary distance and higher sequence similarity between *AtDMP8/9* and *AmTrDMP* could explain why *AmTrDMP* partially rescued the *dmp8,9* phenotype, while *CrDMP* did not. To investigate species-specific aspects of DMPs during double fertilization, rescue experiments with angiosperm *DMPs* more closely related to AtDMP8/9 should be performed. Furthermore, *dmp* loss-of-function studies in other species such as algae could foster their role as deeply conserved gamete fusion facilitators. While not feasible for *Amborella*, mutant libraries for *Chlamydomonas* are readily available (Li *et al.*, 2016) and gamete adhesion and fusion are well-dissected in this model species (reviewed by Pan *et al.*, 2003).

5.5 Are the fusion facilitators DMP8/9 involved in membrane remodeling?

*In situ g*amete attachment assays revealed that unfused *dmp8,9* sperm cells frequently attach to the egg cell. A similar assay performed with *gex2*-deficient sperm (Mori *et al.*, 2014) revealed a failure to attach, suggesting that gamete attachment and gamete fusion are distinct processes and that DMP8/9 function in a step downstream of GEX2-mediated attachment, presumably in facilitating or supporting membrane fusion.

During tight adhesion, the plasma membrane lipid bilayers of the opposing sex gametes are brought into close proximity of ~ 10 nm, which is considered the critical distance for a fusogen to execute its function in bilayer merging (Hernández and Podbilewicz, 2017). The prerequisite for membrane fusion is a complex set of bilayer rearrangements such as membrane bending, lipid rearrangements and formation of a hemifusion stalk (Jahn *et al.*, 2003) which is accomplished by a fusogen or a fusion machinery involving accessory proteins such as DMP8/9.

DMP8/9 could, for example influence biophysical membrane properties or lipid composition at the fusion site, thereby lowering the inter-bilayer forces that need to be resolved by fusogenic proteins such as HAP2. Different plasma membrane lipid properties of egg cell and central cell could also explain the bias towards central cell fertilization observed in *dmp8,9* sperm. Studies on the lipid composition of plant gamete plasma

membranes will help to unravel their role during membrane fusion. While female gametes are difficult to isolate in sufficient amounts (Englhart *et al.*, 2017), the maize sperm cell isolation procedure described in this work makes male gametes readily accessible for lipidomic analysis. Studies on the human sperm plasma membrane showed a distinct composition of lipids characterized by a high concentration of phospholipids containing poly-unsaturated fatty acids (Lenzi *et al.*, 1996). Furthermore, an inhibitory effect for plasma membrane cholesterol in human sperm capacitation was demonstrated (Cross, 1998), suggesting membrane lipid rearrangements play a functional role in mammalian sperm activation or in preparation for fusion.

A role for DMP8/9 in membrane remodeling is also conceivable, as studies on DMP1 showed that transient overexpression of DMP1-GFP fusion protein in tobacco leaf epidermis cells induced membrane remodeling of the ER and the tonoplast. Membrane bending led to the formation of spherical vesicles and foamy membrane structures that might be associated with the breakdown of cellular membrane systems during senescence (Kasaras *et al.*, 2012). This was also observed after overexpression of DMP9-GFP, including membrane patches of high GFP signal intensity from which the spherical/foamy structures seemed to nucleate, but overexpression of a tonoplast marker yielded similar results. The literature describes formation of similar spherical and foamy membranes when tonoplast-localized proteins such as TONOPLAST INTRINSIC PROTEINS (TIPs), the FAB1P/YOTB/VAC1P/EEA1 (FYVE)-domain and a tonoplast-localized H^+-Pyrophosphatase were expressed as GFP fusion proteins (Han *et al.*, 2015; Saito *et al.*, 2002; Segami *et al.*, 2014; Saito *et al.*, 2011). Segami *et al.* (2014) showed that that these bulbs are artifacts, resulting from hydrophobic interactions leading to GFP dimerization in closely opposing vacuolar membranes and they disappeared when a monomeric form of GFP was used. This suggests that membrane remodeling events after overexpression of DMP1/9-GFP are likely artifacts resulting from the attachment of a GFP to the C-terminus of the protein or from overexpression by the strong *35S* promoter.

Recent results furthermore indicated the dual targeting of AtDMP1 to either the tonoplast or the plasma membrane (Kasaras and Kunze, 2017). The N-terminus of AtDMP1 contains a basic dileucine motif, targeting the protein to the tonoplast. However, a N-terminally truncated isoform of AtDMP1 lacking this motif can be formed by alternative translation initiation and is targeted to the plasma membrane. Both isoforms can dimerize and are retained to the vacuole. The work by Kasaras and Kunze (2017) intriguingly showed that attachment of a bulky GFP tag to either the N-terminus or the C-terminus of AtDMP1 leads to mislocalization of the protein either by masking the N-terminal tonoplast-targeting motif or a C-terminal ER-export signal. In contrast to DMP1, DMP8 and DMP9 lack a true

N-terminal dibasic tonoplast targeting motif. However, the GFP signal pattern observed for *35S:DMP9-GFP* in tobacco leaf epidermis cells did not resemble plasma membrane localization but looked highly similar to the tonoplast marker. By contrast, DMP8-GFP and DMP9-GFP signal was observed at the plasma membrane projection tethering the sperm pair to the vegetative nucleus in translational reporter lines. The existence of sperm cell-specific chaperones that assist in targeting of DMP8/9 to the plasma membrane is therefore thinkable.

5.6 How do DMP8 and DMP9 facilitate HAP2-mediated gamete fusion?

Although HAP2 was suggested to act as a gamete fusogen already a decade ago (Wong and Johnson, 2010), compelling evidence was only presented very recently by solving the crystal structure of the *Chlamydomonas* HAP2 extracellular domain and by homology modelling of *Arabidopsis thaliana* HAP2: both proteins showed structural homology of HAP2 to class II viral fusion proteins and somatic fusogens (Fédry *et al.*, 2017; Valansi *et al.*, 2017). *In situ* gamete attachment assays performed with HAP2-deficient sperm showed that sperm cells are able to separate and attach to the female gametes but fail to fuse (Mori *et al.*, 2014). A very similar phenotype was observed for *dmp8,8* sperm cells, raising the possibility that DMP8/9 act corporately with HAP2 as part of the gamete fusion machinery in angiosperms.

In *semi-in* vivo grown *Arabidopsis thaliana* pollen tubes, HAP2-YFP does not localize to the sperm cell plasma membrane but is stored in granular structures of the sperm endomembrane system. Only after stimulation of sperm cells with EC1 peptides, a translocation of HAP2-YFP from endomembrane compartments to the cell surface is observed, rendering the male gametes competent for fusion with the female gametes (Sprunck *et al.*, 2012). However, the molecular mechanism by which HAP2 is translocated to the cell surface when sperm cells are activated is unknown. Sperm-specific DMP8/9 could thus function in chaperoning the fusogen to the sperm plasma membrane either by direct interaction with HAP2 or as part of a multiprotein complex forming the gamete fusion machinery.

A role for DMP8/9 in the timely activation of HAP2 may also be considered. Similar to class II viral fusion proteins, HAP2 is structurally characterized by 3 β-sheet-rich domains I-III connected via flexible linker regions, with domain II carrying fusion loops composed of non-polar amino acid residues that directly intercalate into the target membrane (Fédry *et al.*, 2017; Fedry *et al.*, 2018). It needs to be taken into account that the *Chlamydomonas* and *Arabidopsis* HAP2 ectodomains were crystallized in its post-fusion form. While the

structure of HAP2 in its pre-fusion conformation has not been resolved yet, structural data from viral class II fusion proteins are available in both conformational states and suggest that extensive domain reorganization occurs in transition from pre- to the post-fusion forms (Yin *et al.*, 2005; Yin *et al.*, 2006; Lescar *et al.*, 2001; Gibbons *et al.*, 2004). Studies on the class II fusion proteins from Flaviviruses (e.g. Dengue virus) showed that homodimers are formed as pre-fusion intermediates where the hydrophobic fusion loops of opposing monomers are buried within each other (Modis *et al.*, 2003; Allison *et al.*, 2001; Rey *et al.*, 1995). After viruses are endocytosed by a target cell, the structural rearrangements that lead to the exposure of the previously cryptic fusion loop of viral class II fusion proteins are triggered by the low pH of the endosome (Helenius *et al.*, 1980; White and Helenius, 1980). After the fusion loop of viral class II fusion proteins has been exposed, insertion of the loop into the target membrane is accompanied by the irreversible assembly of fusion protein monomers into trimers (Klimjack *et al.*, 1994; Stiasny *et al.*, 2002; Modis *et al.*, 2004).

In flowering plant sperm cells, HAP2 could benefit from hiding its fusion loops to avoid polyspermy, inadvertent sperm-sperm fusion e.g. while the male germ unit is transported within the pollen tube, or fusion of the sperm plasma membrane with the surrounding endomembrane of the pollen vegetative cell. DMP8/9 could assist in hiding the HAP2 fusion loops by stabilizing the pre-fusion form until the sperm pair has successfully separated and attached to one female gamete each or by triggering timely exposure of the fusion loop. It is not known what triggers HAP2 to expose its fusion loop, but it is plausible that DMP8/9 catalyze this structural rearrangement directly or by being part of a putative signaling pathway that is induced upon sperm activation. However, the extracellular space into which the two sperm cells are delivered after pollen tube burst is an acidic environment, and while the pH varies between species and tissues, apoplastic pH values as low as pH 4 have been reported (reviewed by Grignon and Sentenac, 1991). A mechanism similar to viral class II fusion proteins, where fusion loop exposure is triggered by a low pH, after HAP2 has been translocated from endomembrane compartments to the sperm surface, is conceivable. DMP8/9 could also assist in assembly of HAP2 homotrimers, as the mechanism governing this structural rearrangement is not known and might require accessory proteins. However, the observation that the HAP2 ectodomain behaves as a monomer in solution, but tends to trimerize upon liposome insertion without help off accessory factors argues for a scenario in which trimerization is established via intrinsic affinity of the monomers to each other, after the fusion loops of HAP2 monomers have inserted into the target membrane (Fédry *et al.*, 2017).

5.7 DMP8/9 and the formation of fusion-competent sites

It is conceivable that efficient bilayer fusion requires local clustering of HAP2 trimers at the gamete attachment site in order to generate fusogenic membrane patches. Structural studies on the Dengue Virus class II fusion protein showed that trimers with membrane-inserted fusion loops tend to cluster on liposome surfaces, often forming a continuous layer. These areas of high fusogen density are associated with stronger membrane curvature and might promote membrane fusion (Modis *et al.*, 2004). A similar scenario was observed for the Semliki Forest Virus class II fusion protein where large aggregates consisting of 40-60 trimers were formed through contacts between fusion loops in adjacent trimers (Gibbons *et al.*, 2004).

It is therefore also possible that DMP8/9 facilitate the formation of fusion competent membrane areas by local clustering of the fusogen HAP2 or other, yet unknown accessory factors of the fusion machinery. It would be interesting to investigate local accumulation of DMP8 and DMP9 at the sperm-egg attachment site, however such experiments are not easy to perform because *DMP9g:GFP* is not functional and failed to rescue the fusion defective *dmp8,9* phenotype.

Notably, mammalian sperm possess a subdomain termed equatorial segment, that remains intact following the acrosome reaction and has been predetermined to be the site for initiation of plasmogamy (Yanagimachi and Noda, 1970). The equatorial segment is enriched in proteins mediating gamete interactions such as IZUMO1, CRISP family proteins but also proteins from the TETRASPANIN (TET) superfamily (Inoue *et al.*, 2005; Rochwerger and Cuasnicu, 1992; Ito *et al.*, 2010). TETs are evolutionary conserved four-transmembrane span proteins that participate in a wide range of biological processes in multicellular organisms including cell migration, proliferation, fusion and communication. TETs function as molecular scaffolds and TET-enriched microdomains form a platform for oligomeric protein complexes involved in signaling and adhesion (reviewed by Termini and Gillette, 2017). Notably CD9 and CD81, two egg-expressed TETs, are essential for mammalian fertilization, as knock out mutants display severely reduced female fertility (Kaji *et al.*, 2000; Rubinstein *et al.*, 2006). Gamete attachment via binding of sperm IZUMO1 to egg-expressed JUNO is followed by CD9 accumulation at the sperm-egg contact site and CD9 has been implied in the formation of fusion-competent sites on the egg cell plasma membrane (Chalbi *et al.*, 2014; Jégou *et al.*, 2011).

The evolutionary conservation of TETs across multicellular organisms suggests similar roles in plants (Huang *et al.*, 2005). 17 *TETs* have been identified in *Arabidopsis thaliana*, and expression analysis across reproductive tissues revealed that TET11 and TET12 are sperm-expressed, whereas TET9 localizes to the plasma membrane of the female gametes where

it rapidly gets internalized and accumulates in the cytoplasm after fertilization (Boavida *et al.*, 2013). Nevertheless, TET9-GFP does not accumulate at the egg-sperm interface when unfused *dmp8,9* sperm cells are attached. This, and the fact that loss-of-function studies for gamete-expressed *TETs* do not yet reveal any impact on double fertilization raises the question whether TETs do contribute to the gamete fusion process in plants. Despite being evolutionary unrelated protein families, TETs are structurally related to DMPs, as members from both gene families code for relatively small proteins, characterized by four transmembrane helices, two extracellular loops and cytosolic N- and C-termini. It can therefore be speculated that plant-specific DMPs fulfil a similar function in the formation of fusion competent sites in the green plant lineage.

5.8 Species-specific aspects of gamete fusion

The role of HAP2 as a broadly conserved fusogen is underpinned by the presence of HAP2 orthologs in the genomes of all major eukaryotic taxa expect fungi (Steele and Dana, 2009; Wong and Johnson, 2010). However, the broad diversity of reproductive systems in eukaryotes imply that HAP2 has evolved to accommodate to those vast differences, e.g. in extracellular environments, extracellular matrix/cell wall or in plasma membrane composition of the gametes. A study comparing the fusion loops of HAP2 using representative members from four out of the five eukaryotic kingdoms showed adaption of HAP2 membrane insertion motifs to different fusion environments. HAP2 of flowering plants, for example, uses an amphipathic helix that presents nonpolar residues to the target membrane, while the membrane interaction surface of trypanosome HAP2 comprises three shallow nonpolar loops (Fedry *et al.*, 2018).

Complementation of the *dmp8,9* mutant with DMPs from other species performed in this work suggests that DMPs might also function in a lineage-specific manner, as the evolutionary distant *CrDMP* failed to rescue gamete fusion while complementation with *AmTrDMP* led to partial restoration of the seed set. This observation correlates with the profoundly different HAP2 fusion surfaces from those two species: *AmTrHAP2* shows a typical angiosperm membrane interaction surface characterized by a single amphipathic α-helix whereas the *CrHAP2* fusion surface is composed two smaller loops (Fedry *et al.*, 2018; Baquero *et al.*, 2019). The failed rescue by *CrDMP* might suggest lineage-specific interplay between DMP8/9 and the fusion surfaces of HAP2. Nevertheless, Fedry *et al.* (2018) showed that the HAP2 fusion helices are interchangeable across the diversity of angiosperms which argues against such a scenario. However, HAP2-DMP8/9 interaction interfaces other than the fusion loops are also conceivable. A refined crystal structure of *CrHAP2* showed species-specific features including an algae-specific cystine ladder motif

and a mucin-like domain that is highly variable in sequence and length in HAP2 orthologs across eukaryotes (Baquero *et al.*, 2019). Notably, those species-specific motifs are located laterally at the juxtamembrane region of HAP2, a part where interaction with small integral membrane proteins such as DMP8/9 or other fusion facilitators is perfectly conceivable.

Species specific aspects of HAP2 mediated gamete fusion have also been investigated by domain swap experiments, showing that the *Arabidopsis thaliana* HAP2 ectodomain tolerates exchange with a closely related sequence from the Brassicaceae *Sisymbrium irio* whereas the more distant ectodomain from rice did not restore function (Wong *et al.*, 2010). Complementation of *Arabidopsis hap2* mutants with truncated HAP2 versions has been performed in an attempt to identify functional regions of the protein, showing that the extracellular N-terminus is essential for HAP2 function (Mori *et al.*, 2010; Wong *et al.*, 2010). Data on the histidine-rich cytosolic C-terminus is ambiguous, with one study showing that this protein portion is dispensable for gamete fusion (Mori *et al.*, 2010), whereas C-terminally truncated AtHAP2 did not complement the mutant phenotype in another study (Wong *et al.*, 2010). Wong *et al.* (2010) also demonstrated that the C-terminus tolerates sequence divergence as long as a strong positive charge is maintained. More recent work in *Chlamydomonas* confirmed the importance of the HAP2 C-terminus, showing that it plays a role in targeting of the protein to the fusion site and fulfilling regulatory functions (Liu *et al.*, 2015b).

When the *dmp8,9* mutant was complemented with truncated forms of DMP9 all deletion constructs failed to restore fertilization failure, suggesting that N- and C-termini as well as the extracellular and intracellular loops of DMP9 are important for its function as fusion facilitator. The results furthermore suggest a complex interaction interface among DMP8/9 and putative interacting proteins. Nevertheless, failed rescue of the fusion failure phenotype might also be a consequence of aberrant protein transport and would require analysis, whether the truncated DMP9 versions successfully reach the sperm plasma membrane e.g. by fluorescence microscopy. As mistargeting of DMP1 after fusion of a GFP to either terminus of the protein has been reported recently (Kasaras and Kunze, 2017), N- or C-terminal GFP fusions of truncated DMP9 versions are not recommended for such an analysis. However, internal insertion of the fluorescent tag e.g. within one of the two extracellular loops, or the intracellular loop, seems plausible. Single amino acids that are conserved between DMP8/9 and their orthologs from other angiosperm species should be exchanged in future experiments to investigate their functional relevance. Notably, all three loops and the short C-terminus of DMP9 contain cysteine residues, highly conserved among AtDMPs. These cysteine residues might be required for the formation of functionally important disulfide bonds between the two extracellular loops and between

the intracellular loop and C-terminus. Indeed, the DISULFIND algorithm (Ceroni *et al.*, 2006) predicts the formation of those disulfide bonds, however with low confidence. Notably, two cysteines are located within the predicted transmembrane α-helices 1 and 3 of DMP8/9 and could serve as acylation sites. The role of transmembrane protein acylation is not well understood and could go beyond mere membrane association which is already ensured by the hydrophobic transmembrane α-helices of integral membrane proteins. Roles for acylation in protein stability and trafficking have been suggested (Linder and Deschenes, 2007). Importantly, studies have suggested that acylation of transmembrane cysteines is associated with tilting or stabilization of conformationally dynamic transmembrane helices within the lipid bilayer (Abrami *et al.*, 2008; Poschner and Langosch, 2009), a feature that could be functionally important for DMP8/9, once they are exposed to local membrane distortion, bending and deformation at the fusion site during plasma membrane merging.

5.9 DMP - EC1 dependencies and the order of events

EC1 is suggested to act as a sperm-activating molecule by inducing the redistribution of HAP2 from internal membranes to the sperm cell plasma membrane (Sprunck *et al.*, 2012). This work fosters such a role for EC1 by showing that wild type sperm cells frequently fail to attach to EC1-deficient egg cells and that EC1-GFP gets successfully secreted by the egg cell in ovules with adjacent unfused *dmp8,9* sperm cells. This suggests that EC1 acts upstream or concomitantly with GEX2-medited sperm adhesion. In contrast to HAP2-YFP, GEX2-GFP shows default plasma membrane localization before pollen tube discharge (Mori *et al.*, 2014) suggesting that EC1-mediated subcellular redistribution is not required for GEX2. However, the low frequency of sperm attached to egg cells lacking EC1 suggests that similar to mammals, flowering plants utilize several other, possibly EC1-responsive gamete attachment factors.

The trigger for EC1 secretion by the egg cell is not known. Sperm attachment, but also early events such as pollen tube reception by the synergid cell could serve as trigger and could easily be tested as a plethora of mutants defective in pollen tube reception and burst are available (reviewed by Johnson *et al.*, 2019). Mechanical forces during explosive pollen tube discharge or the molecular composition of pollen tube cytoplasmic contents serving as stimulus for EC1 secretion are also conceivable. Indeed, pollen tube contents seem to have fertilization-promotive signaling properties, as they can trigger fertilization-independent ovule enlargement, endosperm proliferation and initiate seed coat development (Kasahara *et al.*, 2016; Liu *et al.*, 2019). The recent discovery of an *Arabidopsis thaliana* mutant with pollen lacking sperm cells (Zhang *et al.*, 2017), that

successfully germinate, target an ovule and burst provides exciting opportunities to study to which extent pollen tube contents contribute to successful double fertilization.

5.10 Conclusions and future perspectives

This work has discovered DMP8 and DMP9 as novel sperm-expressed membrane proteins, acting downstream of gamete attachment. As the biochemical-mechanistic function of these fusion facilitators remains unknown, identification of DMP8/9-interacting proteins should be pursued to understand their role as part of the angiosperm gamete fusion machinery. As *hap2* mutants showed a similar fusion-defective phenotype, a modulation of HAP2 fusogenic activity by DMP8/9 is plausible and the fusogen itself should be considered a high priority candidate for protein-protein interaction studies with DMP8/9. Successful double fertilization in *dmp8,9* mutants but also in *gex2*-deficient sperm cells (Mori *et al.*, 2014) suggest functional redundancies and indicates that the protein machinery regulating flowering plant gamete interactions is of complex composition, with additional attachment- and fusion-promotive proteins that remain to be discovered. The maize sperm transcriptomic and proteomic datasets that have been generated in the course of this work will serve as valuable resources for their identification.

Successful double fertilization is required for seed formation. As seeds form the basis for most of our food and livestock feed, our understanding of molecular events mediating gamete interactions during double fertilization will form an important contribution to help securing food reserves for the growing human population in face of a changing world climate. Recent discoveries have shown that the loss of *ZmDMP*, a putative DMP8/9 ortholog can trigger haploid induction in maize, and that *dmp8,9* mutants can induce maternal haploids *in vivo* in the dicotyledonous *Arabidopsis thaliana* (Zhong *et al.*, 2019b; Zhong *et al.*, 2020). Doubled haploid technology has revolutionized traditional maize breeding by accelerating the process of obtaining homozygous elite lines (Ren *et al.*, 2017). These findings intriguingly demonstrate, how our molecular understanding of flowering plant gamete interactions can support modern plant breeding, possibly extending double haploid breeding to dicots species, many of whom are of agronomical relevance.

6. Material and Methods

6.1 Material

6.1.1 Reagents and buffers

Laboratory grade chemicals and reagents were purchased from Sigma-Aldrich or Carl Roth Laborbedarf GmbH, if not stated otherwise. BioPak-grade water was used to prepare buffers and solutions. For RNA work, RNAse-free plasticware and filter-tips were used, and water was treated with DEPC prior to use: Water with 0.1% (v/v) DEPC was stirred vigorously for 1 hour followed by autoclaving (121°C, 21 min). Surfaces and pipettes were cleaned with RNaseZAP™ (Sigma-Aldrich).

6.1.2 Oligonucleotides

Oligonucleotides were purchased from MWG Eurofins Genomics and dissolved in BioPak-grade water to generate 100 pM stock solutions. Oligonucleotide stocks were stored at - 20°C and working solutions were generated by 1:10 dilutions of the stock solutions in BioPak-grade water. A list of oligonucleotides used in this work can be found in Supplemental Table 3.

6.1.3 Enzymes

Enzymes for standard molecular biology work (e.g. restriction digest, PCR) were purchased either from Thermo Fisher Scientific or NEB. Cellulase was purchased from Worthington Biochemical Corporation (LS002601) and Pectolyase Y-23 (P8004.0001) was purchased from Duchefa Biochemie.

6.1.4 Antibodies

For Western Blotting purposes, antibodies were diluted in TBS-T with 5% (w/v) skimmed milk powder, as indicated:

Table 5 Antibodies used in this work

Antibody	Manufacturer	Dilution (Western Blot)
α-GFP	Roche #11814460001	1:10.000
α-H^+-ATPase plasma membrane	Agrisera AS07 260	1:5000
α-UGPase	Agrisera AS05 086	1:2500
α-mouse HRP conjugate	Sigma A4416	1:8000
α-rabbit HRP conjugate	Sigma A0515	1:8000

6.1.5 Microscopes

Table 6 Microscopes used in this work

Microscope	Stage	Objectives
Fluorescence microscope	Zeiss Axio Imager.M2 with ApoTome.2	Apochromat 10x/0.45 DRY Apochromat 20x/0.8 DRY Apochromat 40x/1.4 OIL DIC Apochromat 63x/1.4 OIL DIC
Inverted fluorescence microscope	Nikon Eclipse TE200-S	Plan Fluor 4x/0.13 LWD 20x/0.40 Plan Fluor 20x/0.45 LWD 40x/0.55 Plan Fluor 40x/0.60
Confocal Laser Scanning Microscope	Leica SP8	HC PL FLUOTAR 10x/0.30 DRY HC PL FLUOTAR L 20x/0.40 DRY HC PL APO CS2 20x/0.75 IMM HC PL APO CS2 40x/1.30 OIL HC PL APO CS2 63x/1.40 OIL HC PL APO CS2 63x 1.30 GLYC
Confocal Laser Scanning Microscope	Zeiss LSM 880 with Airyscan	W PL APO 40x/1.0 DIC M27 WATER
Stereo Microscope	Zeiss Stemi 305	n.a.
Stereo Microscope	Zeiss Discovery.V8	PlanApoS 0.63x FWD 81mm

6.1.6 Software, databases and online tools

Table 7 Software and databases used in this work

Software	Version	Reference
Aramemnon	8.1	Schwacke *et al.*, 2003
Arabidopsis eFP Browser	2.0	Winter *et al.*, 2007
Breaking-Cas	n.a.	Oliveros *et al.*, 2016
CLC Main Workbench	5.5	CLCbio
DSDecodeM	1.0	Liu *et al.*, 2015a
GENEVESTIGATOR®	V3	Hruz *et al.*, 2008
HMMTOP	2.0	Tusnády and Simon, 2001
ImageJ	1.50e	Schneider *et al.*, 2012
InterPro Database	Release 75.0	Mitchell *et al.*, 2019
LAS AF	V 3.1.0 build 8587	Leica Microsystems
MEGA7	7.0.26	Tamura *et al.*, 2007

Microsoft Office Package	Office 365	Microsoft Corporation
NEB Tm Calculator	1.12.0	NEB
Origin 2019	9.6.0.172	OriginLabs
Phytozome	V12.1	Goodstein *et al.*, 2012
Primer-BLAST	n.a	Ye *et al.*, 2012
Prism	8	GraphPad Software
SignalP	4.0	Petersen *et al.*, 2011
TAIR	V10	Huala *et al.*, 2001
SubaCon	4	Hooper *et al.*, 2014
WoLF Psort	1.0	Horton *et al.*, 2007
Zen Pro	3.1	Carl Zeiss Microscopy

6.1.7 Bacteria strains

Table 8 Bacteria strains used in this work

Organism	Strain	Antibiotic resistance	Purpose
Escherichia coli	DH5α	none	Cloning strain
Agrobacterium tumefaciens	C58C1	none	*Arabidopsis thaliana* transformation by floral dip Transient protein expression by tobacco leaf infiltration
Agrobacterium tumefaciens	GV3101-pMP90	Rifampicin Gentamycin Tetracyclin	*Arabidopsis thaliana* transformation by floral dip

6.2 Computed-based methods

For various *in silico* analyses the tools listed in Table 7 were employed.

6.2.1 Phylogenic analysis

Full length amino acid sequences of proteins containing the conserved DUF679 from selected species (*Arabidopsis thaliana, Araucaria cunninghamii, Amborella trichopoda, Chlamydomonas reinhardtii, Exidia glandulosa, Physcomitrella patens, Solanum lycopersicum, Sphaerobolus stellatus, Volvox carteri, Zea mays*) were retrieved from InterPro Database (IPR007770). Truncated sequences were not included. The protein alignment and the phylogenic tree (neighbor-joining algorithm, 1,000 bootstrap replicates) was constructed in MEGA7. The *Arabidopsis thaliana* DMPs protein alignment was generated with CLC Main Workbench.

6.2.2 Design of CRISPR/Cas9 sgRNAs

Gene-specific sgRNA binding sites were determined with the Breaking-Cas online tool (Oliveros *et al.*, 2016) using the *Arabidopsis thaliana* genome deposited in EnsemblGenomes as reference and nuclease setting “SpCas9; PAM:NGG”.

6.2.3 4.2.3 Data analysis and visualization

Data analysis and visualization was performed with Micosoft Excel, Origin 2019 or GraphPad Prism8. Null hypotheses were tested in two-sided student´s t-tests assuming equal variances (confidence interval $\alpha = 0.05$). The beads-on-a-string model of DMP9 was constructed with Protter (Omasits *et al.*, 2014).

6.3 Organism-based methods

6.3.1 Cultivation of bacteria

Escherichia coli were grown at 37°C in LB medium (5 g/L yeast extract, 10 g/L NaCl, 10 g/L tryptone) supplemented with the appropriate sterile-filtered antibiotic (Kanamycin: 50 µg/ml, Spectinomycin: 100 µg/ml).

Agrobacterium tumefaciens were grown at 28°C in YEP medium (10 g/L yeast extract, 10 g/L peptone, 5 g/L NaCl) supplemented with the appropriate sterile-filtered antibiotic (Kanamycin: 50 µg/ml, Spectinomycin: 250 µg/ml, Tetracyclin: 10 µg/ml, Gentamycin: 2 µg/ml, Rifampicin: 10 µg/ml). Agrobacteria for transient expression in *Nicotiana benthamiana* were cultured in DYT medium (16 g/L Tryptone, 10 g/L yeast extract, 5 g/L NaCl, pH 7).

6.3.2 Preparation of competent cells

Chemically competent *Escherichia coli* were prepared according to Inoue *et al.* (1990). Chemically competent *Agrobacterium tumefaciens* were prepared with a $CaCl_2$-method based on Dagert and Ehrlich, 1979.

6.3.3 Bacteria transformation

Escherichia coli were transformed by heat shock (42°C, 90 sec), followed by incubation with LB medium in a shaking incubator (200 rpm, 37°C, 45 min). After pelleting, cells were plated on LB agar plates supplemented with the appropriate antibiotic.

For *Agrobacterium tumefaciens* transformation, plasmid DNA (~100 ng) was added to the cells, and they were snap-frozen in liquid nitrogen. The cells were thawed in a water bath (37°C), followed by incubation with YEP medium in a shaking incubator (200 rpm, 28°C, 120 min). After pelleting, cells were plated on YEP agar plates supplemented with the appropriate antibiotics.

6.3.4 Plant growth

Arabidopsis thaliana (ecotype *Col-0*) plants were grown on soil in climate chambers (temperature: 20°C, humidity: 65%, illuminance: 8,000 Lux) under short day (8 h light / 16 h dark photoperiod) conditions to promote vegetative growth or under long day (16 h light/ 8 h dark photoperiod) conditions to stimulate flowering. An overview of marker lines used in this work can be found in Table 9.

Table 9 *Arabidopsis thaliana* marker lines used in this thesis

Marker line	Marker for	Reference
DD65p:GFP	Central cell cytoplasm	Steffen *et al.*, 2007
EC1.1p:tag-RFP-RemA	Egg cell plasma membrane	Cyprys *et al.*, 2019
EC1p:NLS-3xGFP	Egg cell nucleus	Englhart *et al.*, 2017
EC1p:EC1-GFP	Egg cell	Sprunck *et al.*, 2012
HAP2p:HAP2-YFP	Sperm cell membrane	Besser *et al.*, 2006
HTR10p:HTR10-mRFP	Sperm cell nucleus	Ingouff *et al.*, 2007
TET9p:TET9-GFP	Egg cell, central cell plasma membrane	PD Dr. Stefanie Sprunck

For *in vitro* plant growth, *Arabidopsis thaliana* seed were surface sterilized by subsequent washes (1. Wash: 70% (v/v) ethanol; 2. Wash: 0.5% (v/v) Mucasol, 0.02% (w/v) sodium hypochlorite) followed by 5 washes with sterile water. Seeds were dried on a filter paper under the sterile bench and sawn out on square petri dishes containing 2.2 g/L ½ MS medium + MES buffer + Vitamins (Duchefa Biochemie) and 1% (w/v) Phytoagar (Duchefa Biochemie). Plates were stratified in the dark (24 h, 4°C) and grown under long day conditions. For selection of transgenic plants, the MS medium was supplemented with the appropriate antibiotic (Kanamycin: 50 µg/ml, Hygromycin B: 30 µg/ml).

Plants transformed with the *bar* gene (Phosphinotricin-Acetyltransferase) as selection marker were sown on trays and selected by spraying with BASTA® (Bayer Crop Science) with a concentration of 200 mg/l glufosinate ammonium supplemented with 0.1% (v/v)

Tween-20 three days after germination. Spraying was repeated two more times with an interval of two days.

Zea mays (inbred line B73) was cultivated in a greenhouse under 16 h light at 26°C, 8 h dark at 17°C, relative air humidity varying from 50-70% and illuminance of 24,000 lux under metal-halide lamps. Seeds were soaked in water for 24h and subsequently put on soil (90% Einheitserde "SPVM", 10% sand) and directly transferred into the greenhouse. After approx. 1 month of growth plants were fertilized using 4 blocks of Plantosan Compact fertilizer.

6.3.5 Generation of stable transgenic *Arabidopsis thaliana* lines

Agrobacterium-mediated transformation of *Arabidopsis* plants was performed by floral dip (Clough and Bent, 1998), using a modified infiltration solution: 2.2 g/L (w/v) MS medium including MES buffer and Vitamins (Duchefa Biochemie), 2% (w/v) sucrose and 0.001% (v/v) Silwet L77 (Lehle Seeds). Plants were bagged 16 days after dipping and dried for at least two weeks before harvesting seeds.

6.3.6 *In vitro* pollen tube germination

Arabidopsis thaliana pollen tubes were germinated in petri dishes (Ø 35 mm) on solid pollen germination medium (0.175 % low melt agarose) supplemented with 10 µM Epibrassinolide (Sigma-Aldrich) according to Vogler *et al.*, 2014.

6.3.7 Aniline blue staining

Aniline blue staining of *Arabidopsis thaliana* pistils was performed as described by Coimbra *et al.*, 2010. Stained pistils were observed on a Zeiss Axio Imager.M2 epifluorescence microscope using the DAPI filterset.

6.3.8 DAPI staining

For DAPI (4,6-diamidino-2-phenylindole) staining, pollen grains were incubated in staining solution (2.5 µg/ml DAPI, 0.01 % (v/v) Tween-20 and 50 mM sodium phosphate buffer, pH 7.5) in the dark for 10 min before imaging by fluorescence microscopy. DAPI staining of maize sperm cells was performed in the same solution, omitting Tween-20.

6.3.9 Feulgen staining

Feulgen staining of *Arabidopsis thaliana* ovules was performed according to Barrell and Grossniklaus (2013) with dissected ovules (see 6.3.14). All incubation steps were carried out in incubation baskets (INTAVIS, 100 µm mesh) fitted into 24-well cell culture plates. Dehydration steps were performed on an orbital shaker (20 rpm).

6.3.10 Emasculation and hand pollination

An *Arabidopsis thaliana* inflorescence was gently fixed on a microscope slide with duct tape and all open flowers, developing siliques and premature buds were removed with forceps under a dissecting microscope. 2-3 closed buds of flower stage 11 (Smyth *et al.*, 1990) were emasculated by removing all parts of the flower except the pistil. Emasculated pistils were left for 1-2 days to mature before pollination with pollen from freshly opened anthers by dabbing the anthers on the stigma.

6.3.11 Reciprocal crossing

Pistils (*Col-0* or *dmp8-10C*) were emasculated (see 6.3.10) and cross-pollinated using a clean forceps. Almost mature siliques (~ 16 days after pollination) were harvested and cleared as described in 6.3.12.

6.3.12 Clearing of *Arabidopsis thaliana* siliques

Mature but unopened siliques were cleared in ethanol/acetic acid (3:1) over night, followed by overnight incubation in 70% ethanol. Finally, siliques were transferred into sterile water and stored at 4 °C before imaging.

6.3.13 Bulk pollen harvest

Mature *Arabidopsis thaliana* pollen was bulk-harvested from plants grown in trays. A vacuum cleaner was modified with three subsequent nylon meshes (80 micron, 35 micron, 10 micron mesh) for large scale pollen isolation, as described by Johnson-Brousseau and McCormick (2004). Pollen was collected in a 1.5 ml reaction tube, snap-frozen in liquid nitrogen and stored at -80°C until further use.

6.3.14 Clearing of *Arabidopsis thaliana* ovules

Ovules were dissected from the pistil under a stereo microscope using hypodermic needles (0.90x40 mm) and placed into a drop of clearing solution. Clearing of unfertilized ovules and early seeds was performed in chloral hydrate solution (20 g chloral hydrate, 2.5 ml glycerol, 10 ml H_2O) while older and aborted seeds were cleared in Hoyer´s solution (Anderson, 1954). Ovules were cleared for 1 to 4 hours before imaging by DIC microscopy.

6.3.15 Gamete attachment assays

Pistils of the egg cell plasma membrane marker line *EC1.1p:tagRFP-RemA* or the central cell marker line *DD65p:GFP* were emasculated and hand pollinated (see 6.3.10) 2 days later either with mutant pollen (*dmp8,9C_3 HTR10-mRFP* or *dmp8–10C_3 HTR10-mRFP*) or with *HTR10-mRFP* control pollen. For *in situ* sperm cell adhesion assays, dissected ovules were incubated 20–24 h later in a mixture of cell wall-degrading enzymes according to Mori *et al.*, (2014) except that Worthington Cellulase (LS002601) was used and incubation time in the enzyme mixture was reduced to 10-15 minutes. After protoplasting, ovules were transferred into 1× PBS (pH 7.4) and imaged by confocal microscopy.

6.3.16 Quantification of EC1 secretion

Pistils of the marker line *EC1p:EC1-GFP* were emasculated and hand pollinated (see 6.3.10) two days later with either *dmp8,9C HTR10-mRFP* pollen or control pollen (*HTR10p:HTR10-mRFP*). 18 HAP pistils were dissected, and ovules were placed into 1x PBS for imaging by confocal microscopy, using identical microscope and laser settings for all experiments.

EC1-GFP fluorescence intensity in unfertilized egg cells, zygotes and egg cells which failed to fuse with *dmp8,9C* sperm cells was quantified via ImageJ: the outline of the egg cell/zygote was manually selected using the freehand selection tool, signal intensity and area were calculated using the measure-tool. Intensity values were divided by area, resulting in normalized signal intensity values.

6.3.17 Confocal laser scanning microscopy (CLSM)

CLSM was performed with either a Leica SP8 or with a Zeiss LSM 880 (see 6.1.5). The 488 nm argon laser line was used for excitation of GFP and YFP and emission was detected from 500 to 535 nm (GFP) and from 510 to 550 nm (YFP). Red fluorescent proteins (mRFP and tagRFP-T) were excited using a DPSS 561 nm laser and emission was detected from

566 to 637 nm. For the imaging of Feulgen-stained pollen and ovules excitation at 488 nm was used and emission was detected from 570 to 700 nm.

6.3.18 Transient protein expression in *Nicotiana benthamiana*

Agrobacterium tumefaciens strain C58C1 was transformed with expression clones and cultured in DYT medium supplemented with the appropriate antibiotics until an OD of 0.8. 1 ml of the culture was pelleted by centrifugation and resuspended in 1 ml of infiltration buffer (5% (w/v) saccharose, 0.01% (v/v) Silwet L77, 450 µM Acetosyringone). For co-infiltration with an organelle marker (Nelson *et al.*, 2007) an equal volume of OD-adjusted resuspended cultures was mixed and incubated at 4°C for 1 hour. *Agrobacteria* carrying the p19 silencing suppressor (Voinnet *et al.*, 2003) were always co-infiltrated. Leaves of 4-week old *Nicotiana benthamiana* plants grown in a greenhouse were infiltrated with the culture using a 2 ml syringe.

6.3.19 Maize sperm cell isolation

Sperm cells were isolated from *Zea mays* pollen using density gradient centrifugation on a discontinuous Percoll gradient. The procedure was modified after Dupuis *et al.* (1987). The evening before isolation, maize plants were gently shaken to remove old pollen grains. On the following morning fresh pollen was harvested after anthesis by placing the tassels into a paper bag followed by gentle shaking. The harvested pollen grains were sieved over a tea strainer to remove spikelets and other crude plant material. Sieved pollen was transferred into a plastic petri dish, a wet filter paper was mounted into the lid and sealed with parafilm to form a moist chamber. Pollen was allowed to hydrate by incubation in the moist chamber (120 min, room temperature). Afterwards the hydrated pollen grains were ruptured by osmotic shock: the filter paper was removed and a 0.55 M Mannitol solution was added to the petri dish (1 ml/0.1 g dry pollen weight) and the mixture was incubated on a platform shaker (70 rpm) for 1 hour. Afterwards the lysate was filtered through a cell strainer (pluriStrainer® 20 µm) to remove empty and unruptured pollen grains, resulting in a yellowish filtrate.

While pollen was hydrating, a discontinuous Percoll density gradient was carefully layered in a 30 mL Corex centrifugation tube using a peristaltic pump (Flow rate: 60 ml / h) starting with the bottom layer. The layers were (from bottom to top): 5 mL 30% (v/v) Percoll in 0.55 M Mannitol, 6 mL 20% (v/v) Percoll in 0.55 M Mannitol, 6 mL 15% (v/v) Percoll in 0.55 M Mannitol. Afterwards, the filtrate was careful layered on top of the Percoll gradient using a pipette. After layering, the centrifugation tubes require careful handling to not

distort the gradient. Centrifugation was performed in a Sorvall Lynx 4000 centrifuge equipped with a TH13 swing-out rotor (12,000 xg, 4°C, 1h). It is important to turn off the centrifuge brake (deceleration profile "0") to avoid distortion of the gradient. After centrifugation distinct turbid layers became visible at the interphases. After removal of the top layers, the sperm-enriched 20/30% interphase was carefully aspirated with a pipette, transferred into a 50 ml Falcon tube and washed with 10 volumes of 0.55 M Mannitol (4°C) by inverting 10 times. Sperm cells were concentrated by centrifugation (2,500 xg , 4°C, 10 minutes) and became visible as a faint white pellet. The majority of the supernatant was removed with a pipette and the pellet was gently resuspended in the remaining 50-150 µl. This concentrated sperm cell suspension was subjected to cell counting with a haemocytometer (Marienfeld "Neubauer improved"). After counting, sperm cells were snap-frozen in liquid nitrogen and stored at -80°C until further use.

6.3.20 FDA staining of maize sperm cells

Sperm cell viability was assessed by a FDA staining procedure modified after Li (2011). 0.1 µl of fresh FDA stock solution (2 mg/ml, dissolved in acetone) was mixed with 100 µl of sperm cell suspension by gentle pipetting. 10 µl of the solution was spread on a microscope slide and a cover slip was carefully placed on top. Imaging was performed using a Zeiss Axio Imager.M2 epifluorescence microscope and the GFP filterset.

6.3.21 Isolation of maize sperm total proteins and membrane proteins

Sperm cells were lysed by snap freezing in liquid nitrogen, and the lysate was cleared by centrifugation to remove starch granules and cellular debris (10,000 xg, 4°C). Protein concentration in the supernatant was determined by Bradford assay, the supernatant was heated in reducing NuPAGE LDS sample buffer (70°C, 10 minutes) and served as total protein sample.

Microsomal fractions were prepared from the supernatant by ultracentrifugation. 1 ml of MicFrac buffer (200 mM Sucrose, 50 mM HEPES-KOH pH 7.5, 5% (v/v) Glycerol, 50 mM Sodium Pyrophosphate, 1 mM Sodium Molybdate, 25 mM Sodium Fluoride, 10 mM EDTA, 0.5% (w/v) Polyvinyl Pyrrolidone, 3 mM DTT, 1x Protease inhibitor cocktail) was added to 50-100 µl of cleared lysate and mixed by vortexing. Ultracentrifugation (100,000 xg, 4°C, 1 h) was performed in a Beckmann Optima™ Max-XP ultracentrifuge equipped with a Beckmann MLA 130 rotor. The microsomal pellet was washed in MicFrac buffer and resuspended in resuspension buffer (30 mM Tris pH 7.5, 3 mM $MgCl_2$, 1 mM DTT, 35% (v/v) glycerol) by vigorous pipetting. Protein concentrations were estimated using

Bradford assay (Bio-Rad). The Suspensions were heated in 1xSDS sample buffer (70°C, 10 minutes) and served as microsomal protein sample. The supernatant was heated in 1xSDS sample buffer (95°C, 10 minutes) and served as soluble protein sample.

6.4 Molecular biology methods

Standard methods of molecular biology were performed according to Sambrook *et al.* (1989) using molecular biology grade chemicals. Nucleic acid concentrations were measured using a NanoDrop2000 (Thermo Fisher Scientific).

6.4.1 Agarose gel electrophoresis

DNA fragments were separated on 0.8 – 1.5 % (w/v) agarose gels in a TAE buffer system (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) stained with ethidium bromide (1 µl per 100 ml gel). DNA fragments were isolated from agarose gels using the QIAquick gel extraction kit (Qiagen).

6.4.2 Polymerase chain reaction (PCR)

Routine PCR (Colony-PCR, genotyping of CRISPR/Cas9 mutants, RT-PCR) was performed in 20 µl reactions with *Taq* polymerase (Thermo Fisher Scientific) using KCl buffer according to the manufacturer´s instruction. For cloning, PCR was performed in 50 µl reactions with proofreading polymerases (Phusion® or Q5®; NEB) according to the manufacturer´s instructions. PCR fragments were purified using the QIAquick PCR purification kit (Qiagen). Annealing temperatures were determined with the NEB Tm calculator online tool.

6.4.3 DNA isolation

Plasmid DNA from *Escherichia coli* was isolated using the NucleoSpin® Plasmid miniprep kit (Machery-Nagel). *Arabidopsis thaliana* genomic DNA was extracted from leaf samples according to Edwards *et al.* (1991). For high molecular weight DNA a CTAB-based protocol was applied (Gelvin and Schilperoort, 1994). *Amborella trichopoda* genomic DNA was kindly provided by Dr. Maria Flores-Tornero. *Chlamydomonas reinhardtii* genomic DNA was a gift from Dr. Wolfgang Mages. All DNA samples were solubilized or eluted in BioPak-grade water.

6.4.4 Gateway® Cloning

All entry clones were generated with the pENTR™/D-TOPO™ cloning Kit (Thermo Fisher Scientific) according to the manufacturer´s instructions. Only half reactions were used, and incubation time was increased to up to 5 hours. All entry clones were verified by sequencing.

Expression clones were generated by recombination between an entry vector and the appropriate Gateway®-compatible destination vector. Recombination was carried out using the LR Clonase™ II Enzyme mix (Thermo Fisher Scientific) in half reactions.

Translational reporters were generated by amplifying genomic fragments including putative promoter regions, but lacking the stop codon from *Arabidopsis thaliana* genomic DNA, followed by generation of entry clones and LR recombination into the destination vector pGWB550 (Nakagawa *et al.*, 2007). Length of putative promoter regions and oligonucleotides used for amplification can be found in Supplemental table 3.

35Sp:DMP9-GFP for transient overexpression in *Nicotiana benthamiana* leaf epidermis cells was generated by amplifying the *DMP9* CDS without stop codon using primer pair #449/#348 followed by cloning into pENTR™/D-TOPO™ and LR recombination into vector pH7FWG2 (Karimi *et al.*, 2002). *35Sp:FLAG-DMP9* was generated by amplifying the DMP9 CDS with stop codon using primer pair #448/#348. The FLAG-tag was introduced via the forward primer #448. The resulting PCR fragment was introduced into pENTR™/D-TOPO™ followed by LR recombination into vector pH2GW7 (Karimi *et al.*, 2002).

35Sp:DMP1-GFP (Kasaras *et al.*, 2012) was kindly gifted by Prof. Dr. Reinhard Kunze (FU Berlin).

6.4.5 Functional complementation of *dmp* mutants

The *dmp8,9C* mutant was complemented with a 2,498 bp genomic fragment including promoter and terminator regions amplified from genomic DNA with primer pair #352/#447 and cloned into pENTR™/D-TOPO™ resulting in vector *DMP9g/pENTR*. Silent nucleotide substitutions were introduced to the *DMP9* sgRNA binding site using the Q5® Site-Directed Mutagenesis Kit (NEB) and primer pair #354/#355, resulting in vector *DMP9g$^{Cas9\text{-}ins}$/pENTR*. The Cas9-insensitive DMP9 fragment was introduced in the binary vector pKGW,0 (Karimi *et al.*, 2002) using Gateway™ LR Clonase™ II. Genotyping of complementation lines was performed with primer pair #225/#398

DMP9 deletion constructs were generated by site-directed mutagenesis using the Q5® Site-Directed Mutagenesis Kit (NEB) according to the manufacturer´s instructions, using

the plasmid *DMP9g*$^{Cas9\text{-}ins}$*/pENTR* as template. Primers were designed using the NEBaseChanger online tool: *DMP9ΔN* (#406/#407), *DMP9ΔLoop2* (#408/#409), *DMP9ΔLoop2-10AA* (#421/#422), *DMP9ΔLoop3* (#424/#424), *DMP9ΔC* (#410/#411). Deletion constructs were introduced in either pKGW,0 or pBGW,0 (Karimi *et al.*, 2002) using Gateway™ LR Clonase™ II.

Complementation constructs with *DMPs* from *Chlamydomonas reinhardtii* (*CrDMP;* Cre07.g312850) and *Amborella trichopoda* (*AmTrDMP;* model.AmTr_v1.0_scaffold00010.477) were generated by PCR amplification of the coding sequence including stop codon and endogenous terminator, but excluding the promoter, from genomic DNA with primer pair #415/#416 and #413/#414 followed by cloning into pENTR™/D-TOPO™. Gateway™ LR Clonase™ II recombination with the vector pB-HT10p-[GW]-GFP (Sprunck *et al.*, 2012) resulted in constructs expressing untagged *CrDMP* or *AmTrDMP* under transcriptional control of the sperm-specific *Arabidopsis thaliana HTR10* promoter.

Gene synthesis of *CrDMP* and *AmTrDMP* coding sequences codon-optimized for expression in *Arabidopsis thaliana,* fused to the *Arabidopsis thaliana HSP18.2* terminator (Nagaya *et al.*, 2010) and cloning into the pUC57 vector was performed by GenScript Biotech Corporation. Coding sequence with terminator was PCR amplified from these vectors with primer pairs #428/#429 and #427/#429 and cloned into pENTR™/D-TOPO™, followed by Gateway™ LR Clonase™ II recombination into the vector pB-HT10p-[GW]-GFP.

6.4.6 Cloning of CRISPR/Cas9 constructs

CRISPR/Cas9 mutants were generated with an egg cell-specific promoter-controlled Cas9 vector system (Wang *et al.*, 2015). Cloning was performed as following: for single sgRNAs complimentary oligonucleotides with 4bp Golden Gate overhangs were designed and phosphorylated by PNK (NEB) according to the manufacturer´s instruction. Oligonucleotides were annealed by heating in a metal heat block (95°C, 5 min) followed by slow cooling to room temperature. Annealed oligonucleotides were incorporated into pHEE401E via cycle ligation as described by Wang *et al.*, 2015. For constructs harboring 2 sgRNAs, a 4-primer PCR was set up, using vector pHEE2A-CHILI as a template. The resulting 626 bp product was purified and incorporated into pHEE401E via cycle ligation. The constructs were used to transform wildtype *Arabidopsis thaliana,* the sperm cell nuclear marker *HTR10p:HTR10-mRFP* (Ingouff *et al.*, 2007) or the T-DNA insertion line *dmp8-1* (SALK_131115C). sgRNA binding sites and primers used for cloning and genotyping of genome-edited plant lines are summarized in Table 10.

Table 10 Sequences and primers for genome editing via CRISPR/Cas9. sgRNA binding sites (PAM site "NGG" underlined), primers used for cloning of CRISPR/Cas9 constructs and primers for genotyping of transformed plants for genome editing events are indicated.

single sgRNA constructs			
Construct	**sgRNA target sites**	**cloning primers**	**genotyping primers**
DMP8C	TCACTGCCTATGCTTCCGGCAGG	#377/#378	#280/#281
DMP9C	CAGCGGCGGCTCCTGACGGGAGG	#370/#371	#282/#283
LFG4C	GAGAGCCCCGAACTCCGATGGGG	#261/#262	#272/#273
PCR11C	CGCGATTTGGACTTGGTCATCGG	#324/#325	#274/#332
RLK2C	GAGGTGGACTTAGATTAGCGAGG	#328/#329	#330/#331
double sgRNA constructs			
DMP8,9C	DMP8: TCACTGCCTATGCTTCCGGCAGG DMP9: CAGCGGCGGCTCCTGACGGGAGG	#373/#374/#375/#376	DMP8: #280/#281 DMP9: #282/#283
DMP8-10C	DMP8/9: TGAACACCGCGAGTCCACGTGG DMP10: GTCCCGGTCGGCAATAGGTTGG	#265/#266/#267/#268	DMP8: #280/#281 DMP9: #282/#283 DMP10: #284/#285

6.4.7 Identification of genome editing events

Genomic DNA was isolated from leaves and a region surrounding the sgRNA binding site was amplified via PCR with the genotyping primers indicated in Table 10. The PCR fragments were purified using ExoSAP-IT™ (Thermo Fisher Scientific) and sent for sequencing. Genome editing events were identified by assembling the obtained sequencing results to a reference genomic sequence obtained from TAIR. Superimposed chromatograms were deciphered using the DSDecodeM online tool (Liu *et al.*, 2015a).

6.4.8 Genotyping of T-DNA insertion lines

Arabidopsis thaliana T-DNA insertion lines were genotyped with the primers indicated in Table 11 to obtain homozygous mutants. The exact insertion site was determined by

sequencing of the border primer PCR fragment and assembly to the reference genomic sequence obtained from TAIR.

Table 11 *Arabidopsis* T-DNA insertion lines and primers used for genotyping of transgenic plants. LP = forward primer, RP = reverse primer, BP = T-DNA border primer.

Line	ID	Insertion	Primers (LP/RP/BP)
dmp7-1	SALK_039261	476 bp 3´ of ATG	#399/#400/#444
dmp8-1	SALK_131115C	63 bp 3´ of ATG	#440/#441/#444
dmp9-1	GABI_475G02	24 bp 5´ of 5´-UTR	#442/#443/#214
dmp10-1	SALK_059670	272 bp 3´ of ATG	#445/#446/#444

6.4.9 mRNA isolation, reverse transcription and RT-PCR

Mature *Arabidopsis thaliana* pollen harvested with a vacuum cleaner (see 6.3.13) was used for mRNA isolation with the Dynabeads mRNA DIRECT Micro Purification Kit (Thermo Fisher Scientific) as reported by Gebert *et al.*, 2008. After DNAse I treatment, mRNA was used for first-strand cDNA synthesis with a oligo$(dT)_{18}$ primer and RevertAid H Minus Reverse Transcriptase (Thermo Fisher Scientific). RT–PCR was performed with 1 µl cDNA each as a template and gene-specific primers designed with Primer-BLAST. Following primers were used: *DMP1* (#383/#384)*, DMP2* (#385/#386)*, DMP3* (#387/#388)*, DMP4* (#389/#390)*, DMP5* (#391/#392)*, DMP6* (#393/#394)*, DMP7* (#395/#396)*, DMP8* (#356/#357)*, DMP9* (#358/#359)*, DMP10* (#360/#361).

6.4.10 Measurement of protein concentrations

Protein concentrations were determined by Bradford assay (Bradford, 1976) using Protein Assay Dye Reagent Concentrate (Bio-Rad) and a NanoDrop2000 (Thermo Fisher Scientific) according to the manufacturer´s instructions. Bovine Serum Albumin Standard (Pierce™) was used to generate a reference curve.

6.4.11 SDS-Polyacrylamide gel electrophoresis (SDS-PAGE)

Protein samples were separated by SDS-PAGE on self-cast polyacrylamide gels (Mini-PROTEAN®, Bio-Rad) in a Tris-Glycine buffer system (Laemmli, 1970) or in a Tris-Tricine buffer system modified after Schägger and Jagow (1987). Silver staining of polyacrylamide gels was performed in a procedure modified after Swain and Ross (1995).

6.4.12 Western Blot

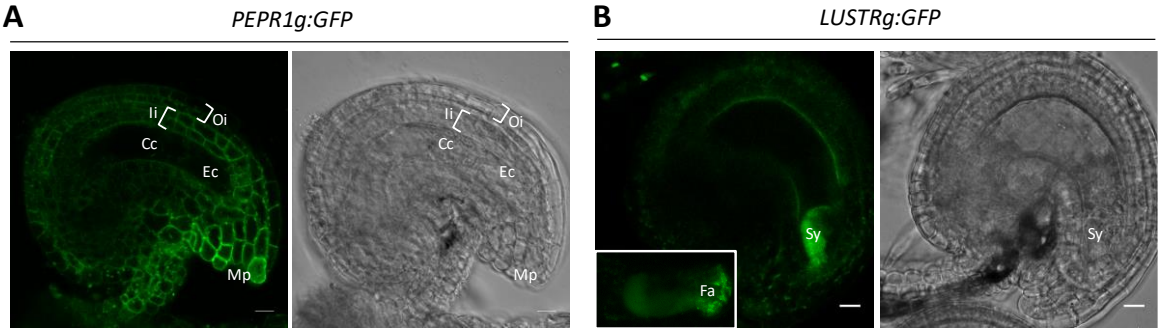
Polyacrylamide gels were blotted on a nitrocellulose membrane (Amersham™ Protran™ 0.2 µm NC, GE Healthcare) using a Mini Trans-Blot® Tank Blot Cell (Bio-Rad). Membranes were blocked in TBS-T containing 5% (w/v) skimmed milk powder and probed with an appropriate primary antibody overnight (4°C, 50 rpm). After washing, membranes were probed with an HRP-conjugated secondary antibody for at least 2 hours at room temperature. Chemiluminescence was detected on a FluorChem™ FC2 ECL-imager (Alpha Innotech) using SuperSignal™ West Femto Chemiluminescence Substrate (Thermo Fisher Scientific). Antibodies used in this work can be found in Table 5.

6.4.13 Mass Spectrometric Analysis

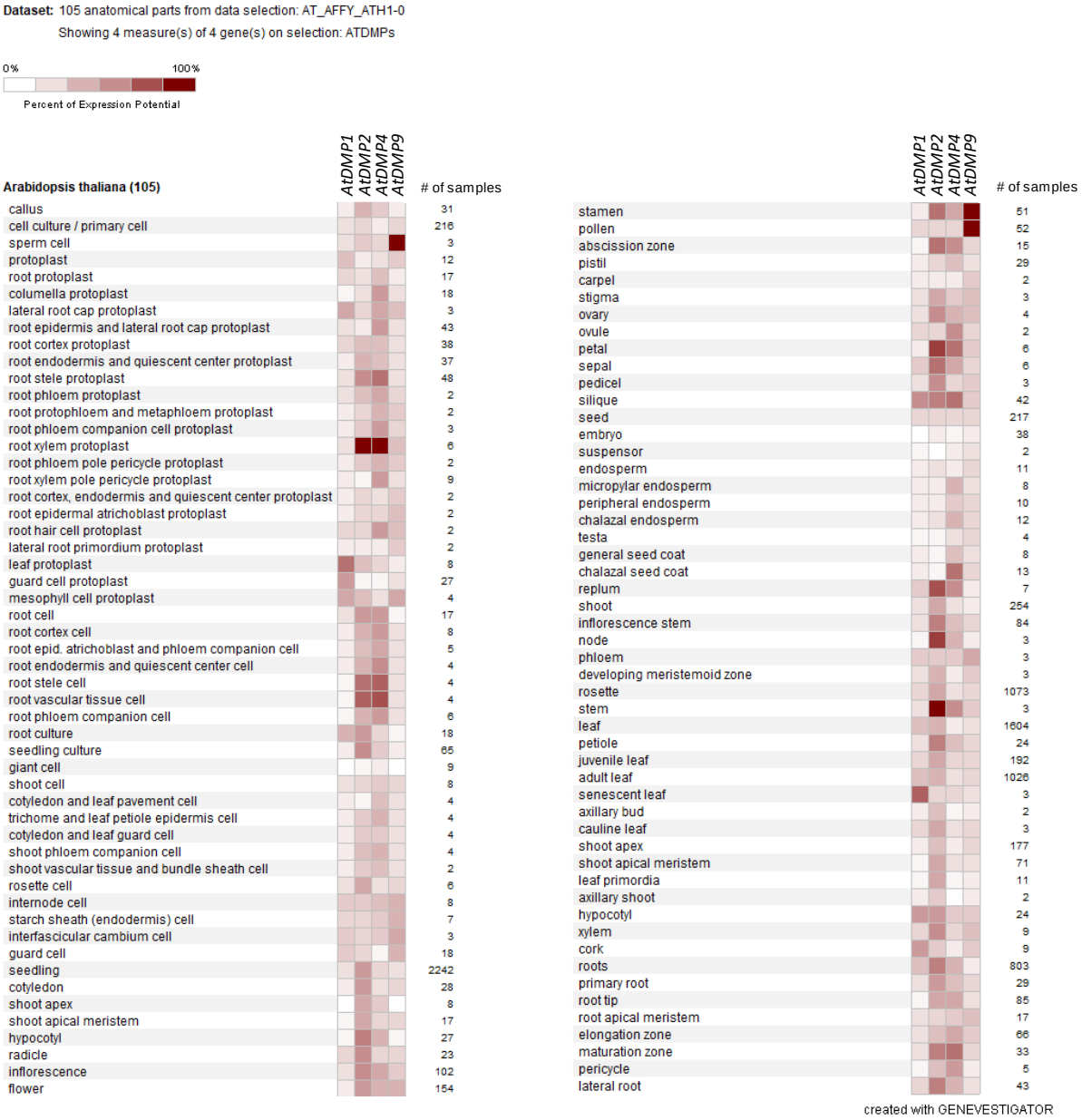
Mass Spectrometric analysis of sperm cell total proteins and microsomal fractions was performed by Dr. Julia Mergner (Technical University of Munich). 3 biological replicates were measured for both experiments. Protein samples (12 µg sperm cell total protein per replicate, 2 µg sperm cell microsomal protein per replicate) were separated on NuPAGE gels (Thermo Fisher Scientific) followed by reduction (10 mM DTT) alkylation (55 mM Chloracetamide) and tryptic in-gel digestion. LC-MS/MS was performed with a Q-Exactive HF mass spectrometer coupled to a Dionex Ultimate 3000 LC (Thermo Fisher Scientific). Peptide fractionation was performed using a 60 min linear gradient 4 to 32 % of solvent B (Solvent B: 0.1% formic acid and 5 % DMSO in acetonitrile; Solvent A: 0.1 % formic acid and 5% DMSO in H_2O). The setup consisted of a trap column with 75 µm x 2 cm, packed with 5 µm particles of ReprosilPur ODS-3 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 106 AGC target, 50 ms maximum injection time, 360-1,300 m/z scan range, profile mode) and MS2 (15,000 resolution, 50 ms maximum injection time, top 20, 1.7 m/z isolation width, 20 s dynamic exclusion, centroid mode). Protein identification was performed with Maxquant software and the *Zea mays* B73 protein database obtained from Ensembl Genomes (Zea_mays.AGPv4.pep.all.fasta; false discovery rate 1%). Comparative analysis of the datasets was performed by Dr. Julia Mergner using Perseus software package (Tyanova *et al.*, 2016).

7. Supplement

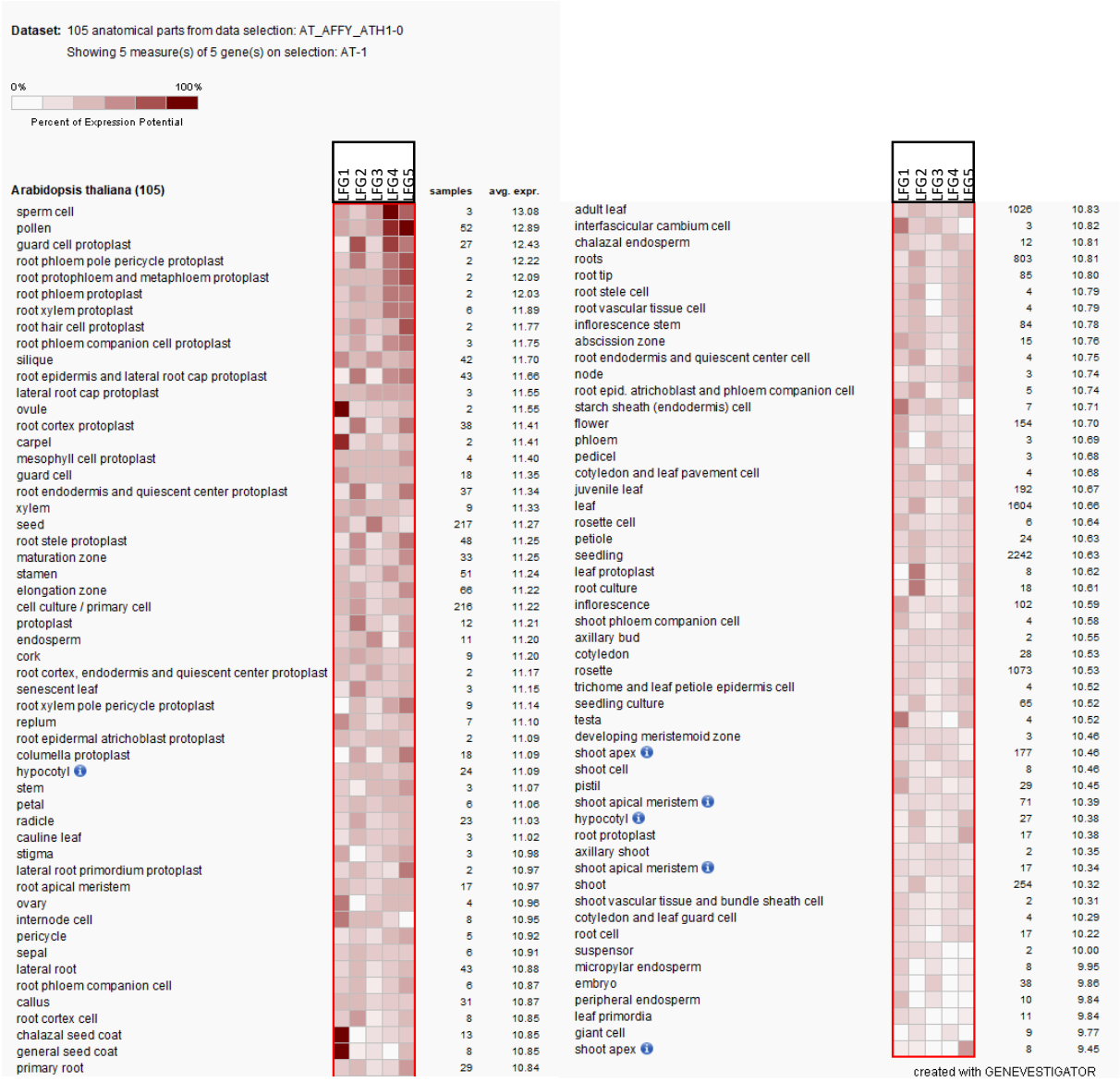
7.1 Supplemental figures



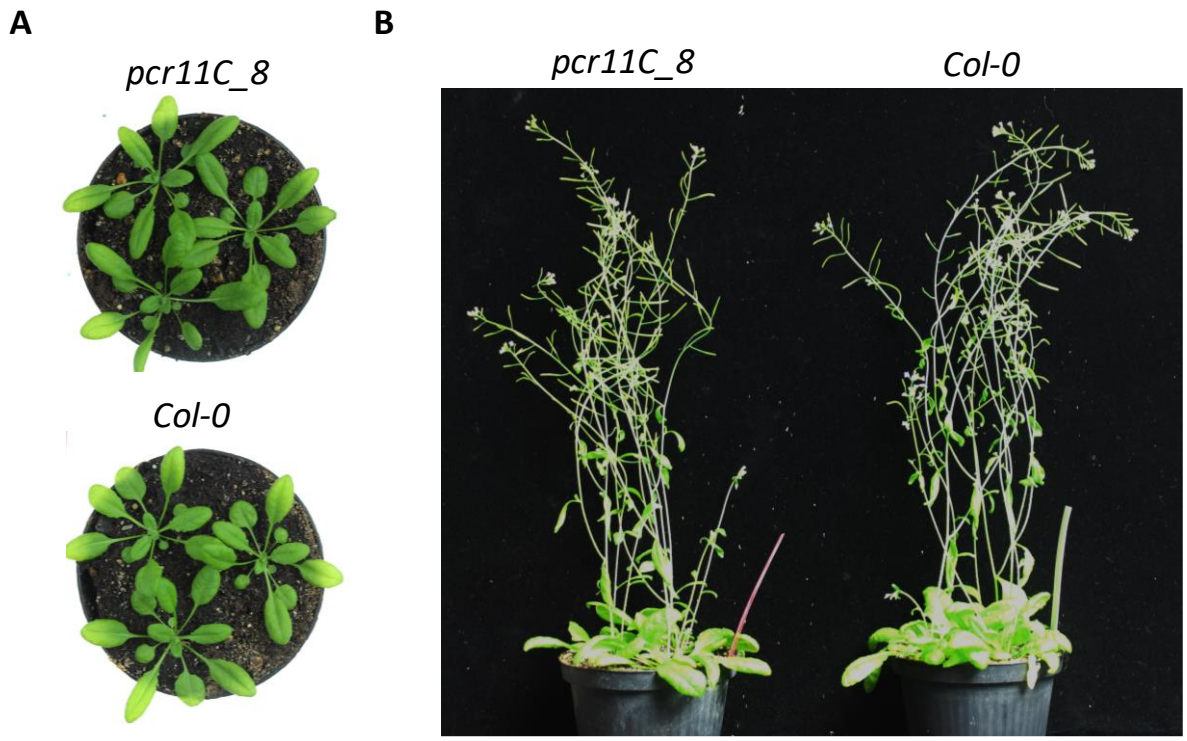
Supplemental figure 1 Localization of candidate GFP-fusion proteins in *Arabidopsis* female gametophytes. **A**, PEPR1-GFP was absent from the embryo sac but present in the plasma membrane of the sporophytic cells surrounding the embryo sac (inner and outer integuments). Signal became stronger towards the micropylar pole. **B,** LUSTR-GFP was detected in synergid cells. Inlay in (**B**) shows aclose-up of the synergid cell. LUSTR-GFP accumulates in granular structures at the filiform apparatus. Cc = central cell; Ec = egg cell; Fa = filiform apparatus; Oi = outer integuments; Ii = inner integuments; Mp = micropyle; Sy = synergid cell. Scale bars = 10 µm.



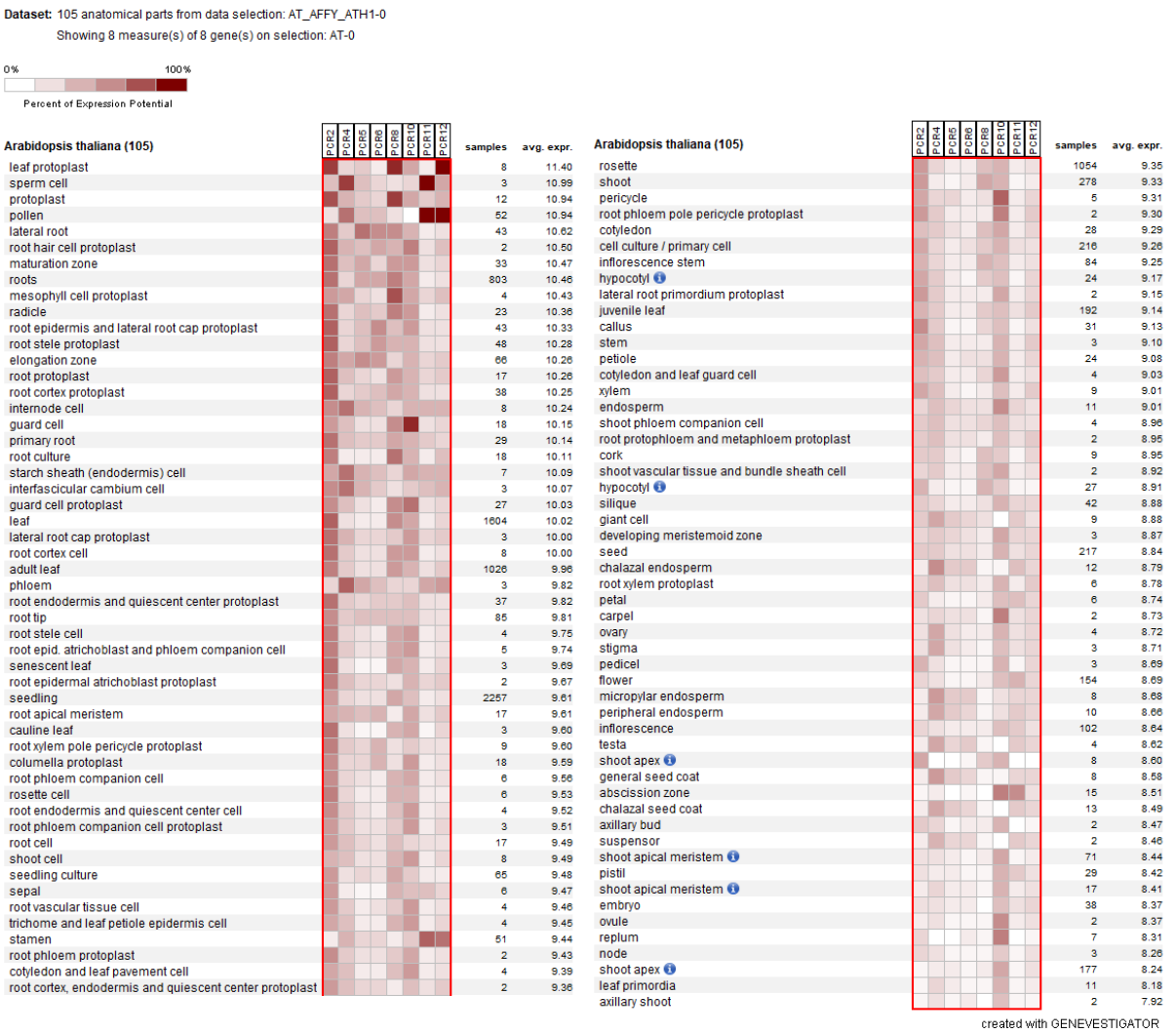
Supplemental figure 2 Microarray-based expression data for *DMP1*, *DMP2*, *DMP4*, and *DMP9*. Expression data were generated by GENEVESTIGATOR®. Note that other *Arabidopsis DMP* gene family members (*DMP3*, *DMP5*, *DMP6*, *DMP7*, *DMP8*, *DMP10*) are not represented on the ATH1 GeneChip®. Expression data were generated by GENEVESTIGATOR.



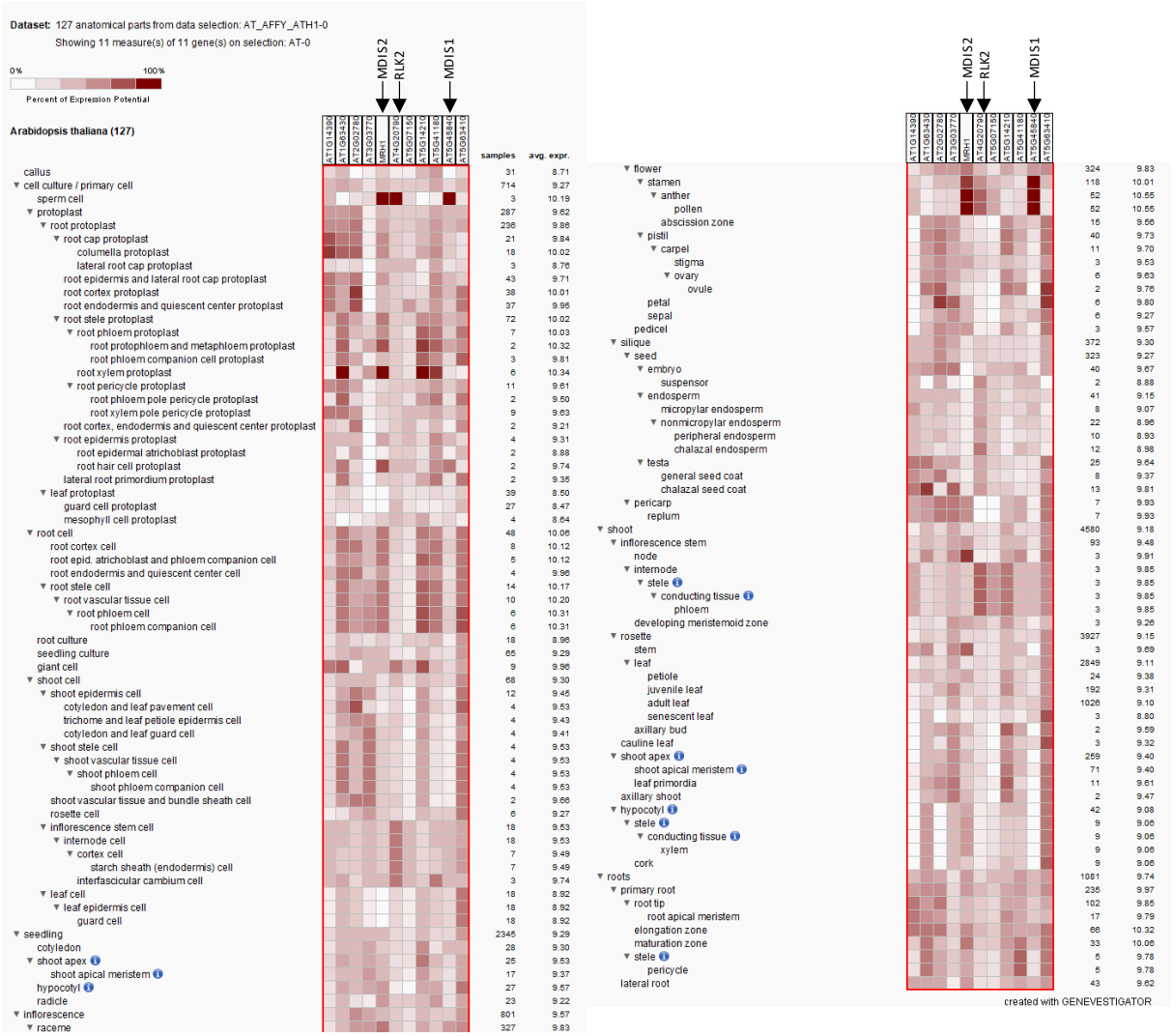
Supplemental figure 3 Microarray-based expression data for *LFG1-LFG5*. Expression data were generated by GENEVESTIGATOR.



Supplemental figure 4 Phenotyping of the *pcr11C* mutant. Leaf size (**A**) and plant height (**B**) of the mutant were comparable to the wild type (*Col-0*).



Supplemental figure 5 Microarray-based expression data for 8 *Arabidopsis thaliana PCRs*. Besides *PCR11*, *PCR4* and *PCR12* show considerable expression in sperm cells. Note, that *PCR1, PCR3, PCR7* and *PCR9* are not represented on the ATH1 GeneChip®. Expression data were generated by GENEVESTIGATOR.



Supplemental figure 6 Microarray-based expression data for all 11 *Arabidopsis thaliana* type "LRR-VI" LRR-RLKs. Apart from RLK2, MDIS1 and MDIS2, "LRR-VI"-RLKs show weak expression in sperm cells. Expression data were generated by GENEVESTIGATOR.

7.2 Supplemental tables

Supplemental table 1 Genome editing events in CRISPR/Cas9 lines generated in this work. Abbreviations: C = CRISPR/Cas9 edited plant line; HTR10-mRFP = line generated in the sperm nuclear marker line HTR10g:mRFP (Ingouff *et al.*, 2007); ho = homozygous; trans-he = trans-heterozygous; WT = wild type; +, insertion; -, deletion.

Description	Line no.	Genotype *DMP8*	Genotype *DMP9*	Genotype *DMP10*
dmp9C HTR10-mRFP	#6	WT	+A / +T (trans-he); 125 bp after ATG	WT
dmp9C HTR10-mRFP	#8	WT	+A (ho); 125 bp after ATG	WT
dmp9C, dmp8-1	#8	T-DNA insertion	+T / +A (trans-he); 125 bp after ATG	WT
dmp9C, dmp8-1	#10	T-DNA insertion	+A (ho); 125 bp after ATG	WT
dmp9C, dmp8-1	#13	T-DNA insertion	+A / +G (trans-he); 126 bp after ATG	WT
dmp8,9C HTR10-mRFP	#3	+T (ho); 122 bp after ATG	+T / +A (trans-he); 125 bp after ATG	WT
dmp8,9C HTR10-mRFP	#5	+C (ho); 120 bp after ATG	+T / +C (trans-he); 125 bp after ATG	WT
dmp8,9C HTR10-mRFP	#6	+T / +G (trans-he); 122 bp after ATG	+T (ho); 125 bp after ATG	WT
dmp8-10C	#3	+G (ho); 419 bp after ATG	+G (ho); 422 bp after ATG	-AT (ho); 53 bp after ATG
dmp8-10C	#10	-G (ho); 418 bp after ATG	+G (ho); 422 bp after ATG	-AT (ho); 53 bp after ATG
dmp8-10C in HTR10-mRFP	#14	large deletion (ho)	+C / -14bp (trans-he); 431 bp after ATG	+A / +C (trans-he); 53 bp after ATG

Description	Line no.	Genotype
lfg4C	#6	+A (ho); 105bp after ATG
lfg4C	#13	+T (ho); 104bp after ATG
lfg4C	#14	+C (ho); 104bp after ATG
pcr11C	#8	+C (ho) 387bp after ATG
rlk2C	#5	deletion (ho) 203bp after ATG

Supplemental table 2 Top 20 *GEX2* co-expressed genes encoding transmembrane, membrane-anchored or secreted proteins. Co-expression analysis was performed using GENEVESTIGATOR. *GEX2* co-expressed genes encode known transmembrane proteins of sperm cells such as GCS1/HAP2 (GENERATIVE CEL SPECIFIC 1/HAPLESS 2), GEX1 (GAMETE EXPRESSED 1), and TET11 (TETRASPANIN 11) but also hypothetical proteins and proteins with unknown function.

Pearson correlation coefficient	Gene Symbol	Gene Identifier	Measure	AramTm Con*	SUBA4**	Description [Source; Accession]
1	**GEX2**	**AT5G49150**	**248645_at**	**TMX**	**PM**	**GAMETE EXPRESSED 2 [UniProtKB/Swiss-Prot; F4K4R6]**
0,8997	SSP, BSK12	AT2G17090	263577_at		Cyt***	BRASSINOSTEROID-SIGNALING KINASE 12, SHORT SUSPENSOR [TAIR; AT2G17090]
0,8754		AT3G45220	252614_at	TMX	PM	SERPIN (Serine protease inhibitor)-Z4 [UniProtKB/Swiss-Prot; Q9M1T7]
0,8741	HAP2	AT4G11720	254883_at	TMX	PM	HAPLESS 2 [UniProtKB/Swiss-Prot; F4JP36]
0,8733		AT5G16100	246491_at	TMX	Nuc	RWP-RK domain protein [Araport11; AT5G16100]
0,8719	CRK43	AT4G28670	253787_at	TMX	PM	Cysteine-rich receptor-like protein kinase 43 [UniProtKB/Swiss-Prot; Q9M0G5]
0,8678		AT4G23960	254183_at	GPI	PM	Probable F-box protein At4g23960 [UniProtKB/Swiss-Prot; Q9T0B0]
0,8638	DMP9	AT5G39650	249468_at	TMX	PM	Protein of unknown function (DUF679) [TAIR; AT5G39650]
0,8539	PAB7	AT2G36660	265205_at		PM	Polyadenylate-binding protein 7 [UniProtKB/Swiss-Prot; Q9ZQA8], PAB7
0,8511		AT3G60760	251384_at		PM/Vac	Hypothetical protein [Araport11; AT3G60760]
0,851	GEX1	AT5G55490	248086_at	TMX	PM	GAMETE EXPRESSED 1 [UniProtKB/Swiss-Prot; Q681K7]
0,8464	PCR11	AT1G68610	262282_at	TMX	PM	PLANT CADMIUM RESISTANCE 11 [UniProtKB/Swiss-Prot; Q9SX24]
0,843	TIP5-1	AT3G47440	252440_at	TMX	Vac	Probable aquaporin TIP5-1 [UniProtKB/Swiss-Prot; Q9STX9]
0,8415		AT5G37970	249598_at		ER	S-adenosyl-L-methionine-dependent methyltransferase [TAIR; AT5G37970]
0,8334	TET11	AT1G18520	261670_at	TMX	PM	Tetraspanin-11 [UniProtKB/Swiss-Prot; Q9LPR6]
0,8247		AT2G22340	264051_at	TMX	PM	Transmembrane protein [Araport11; AT2G22340]
0,8174	PAM74	AT5G59650	247682_at	TMX	PM	Leucine-rich repeat protein kinase family protein [TAIR; AT5G59650], PAM74
0,811	BGAL16	AT1G77410	246386_at		Extra	Beta-galactosidase 16 [UniProtKB/Swiss-Prot; Q8GX69], BGAL16
0,8107		AT5G03000	250963_at	TMX	PM	Putative F-box/kelch-repeat protein [UniProtKB/Swiss-Prot; Q9LYY5]
0,8103		AT2G02061	266111_at	TMX	Chloro	Nucleotide-diphospho-sugar transferase family protein [TAIR; AT2G02061]
0,8098		AT1G30740	263217_at		Extra	FAD-binding Berberine family protein [TAIR; AT1G30740]

*Consensus transmembrane (TM) alpha helix prediction (AramTmCon) by ARAMEMNON

**SUBcellular location database for Arabidopsis proteins (SUBA4)

***TAIR: internal leaflet of plasma membrane (cytoplasmic side of plasma membrane)

Abbreviations: Chloro, chloroplast; Cyt, cytoplasm; ER, endoplasmatic reticulum; Extra, extracellular; GPI, glycosylphosphatidylinositol anchor; Nuc, nucleus; PM, plasmamembrane; TMX, transmembrane helix; Vac, vacuole.

Supplemental Table 3 Oligonucleotides used in this work.

Description	Sequence (5´ → 3´)	Oligo Stock ID
LUSTR-REV-STOP	GAGTTCAAAATCATCCTCCTTC	#124
PEPR1p_GW_FW	CACCAATTGATAACACAAAAGTTTCGG	#125
PEPR1_REV-STOP	CCGAACTGAATCAGAGGAG	#127
LUSTRp_GW_FW	CACCGGGTCTCTCTTCACTTTTG	#142
SH2Ap_GW_FW	CACCAAACCAAACCATCTCGAAAG	#143
SH2A_REV-STOP	GATGCCTCTTATAATCCTGAAAAA	#145
CRK46p_GW_FW	CACCGTTGGAGAGAATTTTCAATGAGCAA	#207
CRK46p_REV-STOP	CATGGAATCATGAAGGAAAGGTG	#209
DMP9p_GW_FW	CACCGAATGAGAATGAAGAAAGAGTTGTTTCGTAT	#223
DMP9p_REV-STOP	ACCAGTCATGCAACCAACACC	#224
DMP9_Gen_FW	TTGATGTTCGAAATGGTTCTTCCATC	#225
LFG4p_GW_FW	CACCTACTGTATTTTGTTGTAGCATCTGTTATCGTTG	#228
LFG4_REV-STOP	CTCGGCAGCTCTGAAAATGGTA	#229
PCR11p_GW_FW	CACCGTTCTTCTCCATCCACACAAACTTG	#251
RCR11_REV-STOP	AACCAATCTAGACATGGGTGCTTG	#252
LFG4_CRISPR_FW	ATTGAGAGCCCCGAACTCCGATG	#261
LFG4_CRISPR_REV	AAACCATCGGAGTTCGGGGCTCT	#262
DT1_BsaI_FW_DMP8/9	ATATATGGTCTCGATTGTGAACACCGCGAGTCCACGGTT	#265
DT1_F0_DMP8/9	TGTGAACACCGCGAGTCCACGGTTTTAGAGCTAGAAATAGC	#266
DT2_REV0_DMP10	AACACCTATTGCCGACCGGGACAATCTCTTAGTCGACTCTAC	#267
DT2_BsaI_REV_DMP10	ATTATTGGTCTCGAAACACCTATTGCCGACCGGGACAA	#268
RLK1p_GW_FW	CACCCCTCTGCTTTGTGATTAGAAGCTGTTCT	#270
RLK1_REV-STOP	CCTTGCCCTCGGAAACAACT	#271
LFG4_CGT_FW	CATTCGTCTGATCAGAATCCATTGCTT	#272
LFG4_CGT_REV	TTGGTTACCTATAAGAGGCGTGATG	#273
PCR11_CGT_FW	AAAAGTTCAAACAGATCACACCATGAACC	#274
DMP8_CGT_FW	CGAAAACTATTGAGTCACAAAACACAGAGA	#280
DMP8_CGT_REV	GATCACTAATAAAACGATGGATAGTAGTCAGTG	#281
DMP9_CGT_FW	AACACAACATACGAAAAACTGGTAAATCAC	#282
DMP9_CGT_REV	CTCCTCCTATCAAACAAGTCTAGATAGAATC	#283
DMP10_CGT_FW	GCATGCATGCATTACGTATAGAAGAAC	#284
DMP10_CGT_REV	TCAATAAGACTAGGCAAAGCTCTTTAC	#285
SNAP30p_GW_FW	CACCTATCTTTTAGTATACATATCTCACTTCCTATATGCAC	#286
SNAP30_REV-STOP	TTTGGAAAGCAAGTGACGTG	#287
RLK2p_GW_FW	CACCTTCTTATGTATTAATTGGGATTTAGACTTATAGAAAA	#288
RLK2_REV-STOP	ATGATGTCGTGAACTAGACTCATACAAT	#289
HVA22Lp_GW_FW	CACCCTTACTCCAAGATGAAATAGATTTTAGAGAGAGAGT	#290
HVA22L_REV-STOP	TGCCGCTGGCTGAGTTCT	#291
PCR11_CRISPR_FW	ATTGGCGATTTGGACTTGGTCAT	#324

PCR11_CRISPR_REV	AAACATGACCAAGTCCAAATCGC	#325
RLK2_CRISPR_FW	ATTGAGGTGGACTTAGATTAGCG	#328
RLK2_CRISPR_REV	AAACCGCTAATCTAAGTCCACCT	#329
RLK2_CGT_FW	ATGGCGTCTTCTCTGTTCTTCATC	#330
RLK2_CGT_REV	ACGGTTAAAGGACAGTACTTCCAG	#331
PCR11_CGT_REV	TGTTAAACCAATCTAGACATGGGTGC	#332
DMP8p_GW_FW	CACCAAATGAGATTGGTTTTGGGATTACTACT	#337
DMP8p_REV-STOP	TGTAGACATGCATCCGACACC	#338
DMP10p_GW_FW	CACCAATAGCACTTTCATAAAAGTATTTTGACACTTTA	#339
DMP10p_REV-STOP	ACGAATGTCTGAAATTCCGATG	#340
DMP9_CDS_REV+STOP	TTAACCAGTCATGCAACCAACACC	#348
DMP9p_compl_FW	CACCGAATGAGAATGAAGAAAGAGT	#352
DMP9_muta_FW	GTTTTCATGAAACCGCCGCCTCCAGAG	#354
DMP9_muta_REV	TGCTAGACCCCGTGGCGTCACGAAACCG	#355
DMP8_RT_FW	CCGGTGACGGATGAGAGGTA	#356
DMP8_RT_REV	CCGACACCGTATCGACTGG	#357
DMP9_RT_FW	TCCGGTGACTGATGATAGGT	#358
DMP9_RT_REV	CCAGTCATGCAACCAACACC	#359
DMP10_RT_FW	ACGACTATCCAGACGAAGGT	#360
DMP10_RT_REV	ATGCCTCTCCGTTTTGAGGG	#361
DMP9_CRISPR_FW	ATTGAGCGGCGGCTCCTGACGGG	#370
DMP9_CRISPR_REV	AAACCCCGTCAGGAGCCGCCGCT	#371
DT1_BsaI_FW_DMP8	ATATATGGTCTCGATTGCACTGCCTATGCTTCGGGCGTT	#373
DT1_F0_DMP8	TGCACTGCCTATGCTTCCGGCGTTTTAGAGCTAGAAATAGC	#374
DT2_REV0_DMP9	AACCCCGTCAGGAGCCGCCGCTCAATCTCTTAGTCGACTCTAC	#375
DT2_BsaI_REV_DMP9	ATTATTGGTCTCGAAACCCCGTCAGGAGCCGCCGCTCAA	#376
DMP8_CRISPR_FW	ATTGCACTGCCTATGCTTCCGGC	#377
DMP8_CRISPR_REV	AAACGCCGGAAGCATAGGCAGTG	#378
DMP1_RT_FW	CGGTGAATGCTCCACAGGTA	#383
DMP1_RT_REV	TGTACCAAACACAGCCAGCA	#384
DMP2_RT_FW	TTCAGGCGTAGGAGACCTCA	#385
DMP2_RT_REV	GGGAACACCGTGAAAACAGC	#386
DMP3_RT_FW	CCACAGCTAAGGAGGCAGAC	#387
DMP3_RT_REV	GCGTGAGGAGCACTATGGTT	#388
DMP4_RT_FW	TCTTTTCTCTCGCCTCGCTC	#389
DMP4_RT_REV	ATCGCATTGACCGCCATTTG	#390
DMP5_RT_FW	CCGGAGCTCGACGAGTTATC	#391
DMP5_RT_REV	AGAGGAAAAGGAGAACGGCG	#392
DMP6_RT_FW	CGGGGACAGTTCTTGCATTC	#393
DMP6_RT_REV	CTTCAGCTGATGGTTCGGGA	#394
DMP7_RT_FW	TGGAGGAGACGAAGCAGTCA	#395
DMP7_RT_REV	TCCATCACTAATAGCCCACTCC	#396
RB_REV	TGAGCTCTCCCATATGGTCGA	#398

dmp7-1_LP	CCTGAACAAACAACTCAACGG	#399
dmp7-1_RP	TAACTGCGGAACAGAAAGCTG	#400
DMP9_ΔN_FW	ACATCAATGCTCGTCAAC	#406
DMP9_ΔN_REV	CATTTTCGTGTGTTTCTC	#407
DMP9_ΔLoop2_FW	AGGTATAAGTTGACGGTTAATG	#408
DMP9_ΔLoop2_REV	GTCGGTAAAATGGAAGAAG	#409
DMP9_ΔC_FW	TAATTTATTTTCACTTTTCTGTTTTTTG	#410
DMP9_ΔC_REV	GGTCGGAAAAACAAGAAAC	#411
AmTrDMP_GW_FW	CACCATGGAAGAGATAGCGGGAGTG	#413
AmTrDMP_REV	GAGCAGAGCCGATTAGAATTGCAAT	#414
CrDMP_GW_FW	CACCATGGACGCCATCAC	#415
CrDMP_REV	TGCGTAGCGGCATGACAA	#416
DMP9_ΔL2_10AA_FW	GATAGGTATAAGTTGACGGTTAATG	#421
DMP9_ΔL2_10AA_REV	AACATCTCCGCCACCAAA	#422
DMP9_ΔLoop3_FW	ATGGATCAAGTTATGGAGAG	#423
DMP9_ΔLoop3_REV	TCGATCCGAAAACGCAATC	#424
AmTrDMP_opti_GW_FW	CACCATGGAAGAAATAGCAGGAGTGGCTGTGA	#427
CrDMP_opti_GW_FW	CACCATGGATGCTATCACCAGGGTTTCTTACTACTTT	#428
HSP18_2t_REV	CTTATCTTTAATCATATTCCATAGTCCATACCATAGCACATAC	#429
dmp8-1_LP	CAAGAAACAAAGCACTGCAAAC	#440
dmp8-1_RP	CCTTTGGTTATGCGAAATGAG	#441
dmp9-1_LP	AAACACCAAAACGCTCATCAC	#442
dmp9-1_RP	AAGAAATCCACCCCAAAACAC	#443
LBb1.3	ATTTTGCCGATTTCGGAAC	#444
dmp10-1_LP	CGTGTCCCAACTTGTTTATGG	#445
dmp10-1_RP	AAGATCAATCATGGAATTTGGG	#446
DMP9t_compl_REV	TTAATTGCATAGTTTGAGATCTTAACAAAAAAAGTAC	#447
FLAG_DMP_GW_FW	CACCATGGATTACAAGGATGATGATGATAAGGAGAAAACAGAGGAAA GCGTC	#448
DMP9_CDS_GW_FW	CACCATGGAGAAAACAGAGGAAAGC	#449

8. References

Abiko, M., Furuta, K., Yamauchi, Y., Fujita, C., Taoka, M., Isobe, T., and Okamoto, T. (2013). Identification of proteins enriched in rice egg or sperm cells by single-cell proteomics. PloS one **8** (7): e69578.

Abrami, L., Kunz, B., Iacovache, I., and van der Goot, F.G. (2008). Palmitoylation and ubiquitination regulate exit of the Wnt signaling protein LRP6 from the endoplasmic reticulum. Proceedings of the National Academy of Sciences of the United States of America **105** (14): 5384–5389.

Alandete-Saez, M., Ron, M., Leiboff, S., and McCormick, S. (2011). Arabidopsis thaliana GEX1 has dual functions in gametophyte development and early embryogenesis. The Plant journal for cell and molecular biology **68** (4): 620–632.

Allison, S.L., Schalich, J., Stiasny, K., Mandl, C.W., and Heinz, F.X. (2001). Mutational evidence for an internal fusion peptide in flavivirus envelope protein E. Journal of virology **75** (9): 4268–4275.

Almeida, E.A.C., Huovila, A.-P.J., Sutherland, A.E., Stephens, L.E., Calarco, P.G., Shaw§, L.M., Mercurio, A.M., Sonnenberg, A., Primakoff, P., Myles, D.G., and White, J.M. (1995). Mouse egg integrin α6β1functions as a sperm receptor. Cell **81** (7): 1095–1104.

Anderson, J.L., Mulligan, T.S., Shen, M.-C., Wang, H., Scahill, C.M., Tan, F.J., Du, S.J., Busch-Nentwich, E.M., and Farber, S.A. (2017). mRNA processing in mutant zebrafish lines generated by chemical and CRISPR-mediated mutagenesis produces unexpected transcripts that escape nonsense-mediated decay. PLoS genetics **13** (11): e1007105.

Anderson, L.E. (1954). Hoyer's Solution as a Rapid Permanent Mounting Medium for Bryophytes. The Bryologist **57** (3): 242.

Baquero, E., Fedry, J., Legrand, P., Krey, T., and Rey, F.A. (2019). Species-Specific Functional Regions of the Green Alga Gamete Fusion Protein HAP2 Revealed by Structural Studies. Structure (London, England 1993) **27** (1): 113-124.e4.

Barrell, P., and Grossniklaus, U. (2013). Examining female meiocytes of maize by confocal microscopy. Methods in molecular biology (Clifton, N.J.) **990**: 45–52.

Beale, K.M., Leydon, A.R., and Johnson, M.A. (2012). Gamete fusion is required to block multiple pollen tubes from entering an Arabidopsis ovule. Current biology CB **22** (12): 1090–1094.

Besser, K. von, **Frank, A.C., Johnson, M.A., and Preuss, D.** (2006). Arabidopsis HAP2 (GCS1) is a sperm-specific gene required for pollen tube guidance and fertilization. Development (Cambridge, England) **133** (23): 4761–4769.

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

Boavida, L.C., Qin, P., Broz, M., Becker, J.D., and McCormick, S. (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 physiology **163** (2): 696–712.

Boer, H.J. de, **Eppinga, M.B., Wassen, M.J., and Dekker, S.C.** (2012). A critical transition in leaf evolution facilitated the Cretaceous angiosperm revolution. Nature communications **3**: 1221.

Borg, M., Brownfield, L., Khatab, H., Sidorova, A., Lingaya, M., and Twell, D. (2011). The R2R3 MYB transcription factor DUO1 activates a male germline-specific regulon essential for sperm cell differentiation in Arabidopsis. The Plant cell **23** (2): 534–549.

Borg, M., Brownfield, L., and Twell, D. (2009). Male gametophyte development: a molecular perspective. Journal of experimental botany **60** (5): 1465–1478.

Borg, M., Rutley, N., Kagale, S., Hamamura, Y., Gherghinoiu, M., Kumar, S., Sari, U., Esparza-Franco, M.A., Sakamoto, W., Rozwadowski, K., Higashiyama, T., and Twell, D. (2014). An EAR-Dependent Regulatory Module Promotes Male Germ Cell Division and Sperm Fertility in ArabidopsisCWOPEN. The Plant cell **26** (5): 2098–2113.

Borges, F., Gardner, R., Lopes, T., Calarco, J.P., Boavida, L.C., Slotkin, R.K., Martienssen, R.A., and Becker, J.D. (2012). FACS-based purification of Arabidopsis microspores, sperm cells and vegetative nuclei. Plant Methods **8** (1): 44.

Borges, F., Gomes, G., Gardner, R., Moreno, N., McCormick, S., Feijó, J.A., and Becker, J.D. (2008). Comparative transcriptomics of Arabidopsis sperm cells. Plant physiology **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. Analytical biochemistry **72**: 248–254.

Brkljacic, J., Zhao, Q., and Meier, I. (2009). WPP-domain proteins mimic the activity of the HSC70-1 chaperone in preventing mistargeting of RanGAP1-anchoring protein WIT1. Plant physiology **151** (1): 142–154.

Brodribb, T.J., and Feild, T.S. (2010). Leaf hydraulic evolution led a surge in leaf photosynthetic capacity during early angiosperm diversification. Ecology letters **13** (2): 175–183.

Brownfield, L., Hafidh, S., Borg, M., Sidorova, A., Mori, T., and Twell, D. (2009). A plant germline-specific integrator of sperm specification and cell cycle progression. PLoS genetics **5** (3): e1000430.

Bultynck, G., Kiviluoto, S., Henke, N., Ivanova, H., Schneider, L., Rybalchenko, V., Luyten, T., Nuyts, K., Borggraeve, W. de, **Bezprozvanny, I., Parys, J.B., Smedt, H.** de, **Missiaen, L., and Methner, A.** (2012). The C terminus of Bax inhibitor-1 forms a Ca2+-permeable channel pore. The Journal of biological chemistry **287** (4): 2544–2557.

Bultynck, G., Kiviluoto, S., and Methner, A. (2014). Bax inhibitor-1 is likely a pH-sensitive calcium leak channel, not a H+/Ca2+ exchanger. Science signaling **7** (343): pe22.

Cabreira-Cagliari, C., Dias, N.d.C., Bohn, B., Fagundes, D.G.D.S., Margis-Pinheiro, M., Bodanese-Zanettini, M.H., and Cagliari, A. (2018). Revising the PLAC8 gene family: from a central role in differentiation, proliferation, and apoptosis in mammals to a multifunctional role in plants. Genome **61** (12): 857–865.

Ceroni, A., Passerini, A., Vullo, A., and Frasconi, P. (2006). DISULFIND: a disulfide bonding state and cysteine connectivity prediction server. Nucleic acids research **34** (Web Server issue): W177-81.

Chalbi, M., Barraud-Lange, V., Ravaux, B., Howan, K., Rodriguez, N., Soule, P., Ndzoudi, A., Boucheix, C., Rubinstein, E., Wolf, J.P., Ziyyat, A., Perez, E., Pincet, F., and Gourier, C. (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 (Cambridge, England) **141** (19): 3732–3739.

Charzynska, M., and Lewandowska, E. (1990). Generative Cell Division and Sperm Cell Association in the Pollen Grain of Sambucus nigra. Annals of Botany **65** (6): 685–689.

Chaudhury, A.M., Ming, L., Miller, C., Craig, S., Dennis, E.S., and Peacock, W.J. (1997). Fertilization-independent seed development in Arabidopsis thaliana. Proceedings of the National Academy of Sciences of the United States of America **94** (8): 4223–4228.

Chen, J., Strieder, N., Krohn, N.G., Cyprys, P., Sprunck, S., Engelmann, J.C., and Dresselhaus, T. (2017). Zygotic Genome Activation Occurs Shortly after Fertilization in Maize. The Plant cell **29** (9): 2106–2125.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. The Plant journal for cell and molecular biology **16** (6): 735–743.

Cohen, D.J., Busso, D., Da Ros, V., Ellerman, D.A., Maldera, J.A., Goldweic, N., and Cuasnicu, P.S. (2008). Participation of cysteine-rich secretory proteins (CRISP) in mammalian sperm-egg interaction. The International journal of developmental biology **52** (5-6): 737–742.

Coimbra, S., Costa, M., Mendes, M.A., Pereira, A.M., Pinto, J., and Pereira, L.G. (2010). Early germination of Arabidopsis pollen in a double null mutant for the arabinogalactan protein genes AGP6 and AGP11. Sexual plant reproduction **23** (3): 199–205.

Crepet, W.L., and Niklas, K.J. (2009). Darwin's second 'abominable mystery': Why are there so many angiosperm species? American journal of botany **96** (1): 366–381.

Cross, N.L. (1998). Role of cholesterol in sperm capacitation. Biology of reproduction **59** (1): 7–11.

Cyprys, P. (2015). Membrane proteomics and the triggered activation of maize sperm cells,. Master Thesis.

Cyprys, P., Lindemeier, M., and Sprunck, S. (2019). Gamete fusion is facilitated by two sperm cell-expressed DUF679 membrane proteins. Nature Plants **5** (3): 253–257.

Dagert, M., and Ehrlich, S.D. (1979). Prolonged incubation in calcium chloride improves the competence of Escherichia coli cells. Gene **6** (1): 23–28.

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

Desiderio, U.V., Zhu, X., and Evans, J.P. (2010). ADAM2 interactions with mouse eggs and cell lines expressing α4/α9 (ITGA4/ITGA9) integrins: implications for integrin-based adhesion and fertilization. PloS one **5** (10): e13744.

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

Dresselhaus, T., Sprunck, S., and Wessel, G.M. (2016). Fertilization Mechanisms in Flowering Plants. Current biology CB **26** (3): R125-39.

Drews, G.N., and Koltunow, A.M.G. (2011). The female gametophyte. The arabidopsis book **9**: e0155.

Dumas, C., Knox, R.B., and Gaude, T. (1985). The spatial association of the sperm cells and vegetative nucleus in the pollen grain ofBrassica. Protoplasma **124** (3): 168–174.

Dupuis, I., Roeckel, P., Matthys-Rochon, E., and Dumas, C. (1987). Procedure to Isolate Viable Sperm Cells from Corn (Zea mays L.) Pollen Grains. Plant physiology **85** (4): 876–878.

Durbarry, A., Vizir, I., and Twell, D. (2005). Male germ line development in Arabidopsis. duo pollen mutants reveal gametophytic regulators of generative cell cycle progression. Plant physiology **137** (1): 297–307.

Edwards, K., Johnstone, C., and Thompson, C. (1991). A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic acids research **19** (6): 1349.

Eichmann, R., Bischof, M., Weis, C., Shaw, J., Lacomme, C., Schweizer, P., Duchkov, D., Hensel, G., Kumlehn, J., and Hückelhoven, R. (2010). BAX INHIBITOR-1 is required for full susceptibility of barley to powdery mildew. Molecular plant-microbe interactions MPMI **23** (9): 1217–1227.

El-Brolosy, M.A., Kontarakis, Z., Rossi, A., Kuenne, C., Günther, S., Fukuda, N., Kikhi, K., Boezio, G.L.M., Takacs, C.M., Lai, S.-L., Fukuda, R., Gerri, C., Giraldez, A.J., and Stainier, D.Y.R. (2019). Genetic compensation triggered by mutant mRNA degradation. Nature **568** (7751): 193–197.

Engel, M.L., Chaboud, A., Dumas, C., and McCormick, S. (2003). Sperm cells of Zea mays have a complex complement of mRNAs. The Plant journal for cell and molecular biology **34** (5): 697–707.

Engel, M.L., Holmes-Davis, R., and McCormick, S. (2005). Green Sperm. Identification of Male Gamete Promoters in Arabidopsis1w. Plant physiology **138** (4): 2124–2133.

Englhart, M., Šoljić, L., and Sprunck, S. (2017). Manual Isolation of Living Cells from the Arabidopsis thaliana Female Gametophyte by Micromanipulation. Methods in molecular biology (Clifton, N.J.) **1669**: 221–234.

Eto, K., Huet, C., Tarui, T., Kupriyanov, S., Liu, H.-Z., Puzon-McLaughlin, W., Zhang, X.-P., Sheppard, D., Engvall, E., and Takada, Y. (2002). Functional classification of ADAMs based on a conserved motif for binding to integrin alpha 9beta 1: implications for sperm-egg binding and other cell interactions. The Journal of biological chemistry **277** (20): 17804–17810.

Faure, J.-E., Rusche, M.L., Thomas, A., Keim, P., Dumas, C., Mogensen, H.L., Rougier, M., and Chaboud, A. (2003). Double fertilization in maize: the two male gametes from a pollen grain have the ability to fuse with egg cells. The Plant journal for cell and molecular biology **33** (6): 1051–1062.

Fedry, J., Forcina, J., Legrand, P., Péhau-Arnaudet, G., Haouz, A., Johnson, M., Rey, F.A., and Krey, T. (2018). Evolutionary diversification of the HAP2 membrane insertion motifs to drive gamete fusion across eukaryotes. PLoS biology **16** (8): e2006357.

Fédry, J., Liu, Y., Péhau-Arnaudet, G., Pei, J., Li, W., Tortorici, M.A., Traincard, F., Meola, A., Bricogne, G., Grishin, N.V., Snell, W.J., Rey, F.A., and Krey, T. (2017). The Ancient Gamete Fusogen HAP2 Is a Eukaryotic Class II Fusion Protein. Cell **168** (5): 904-915.e10.

Filip, S., Vougas, K., Zoidakis, J., Latosinska, A., Mullen, W., Spasovski, G., Mischak, H., Vlahou, A., and Jankowski, J. (2015). Comparison of Depletion Strategies for the Enrichment of Low-Abundance Proteins in Urine. PloS one **10** (7): e0133773.

Flores-Tornero, M., Proost, S., Mutwil, M., Scutt, C.P., Dresselhaus, T., and Sprunck, S. (2019). Correction to: Transcriptomics of manually isolated Amborella trichopoda egg apparatus cells. Plant reproduction **32** (2): 229.

Frary, A., Nesbitt, T.C., Grandillo, S., Knaap, E., Cong, B., Liu, J., Meller, J., Elber, R., Alpert, K.B., and Tanksley, S.D. (2000). fw2.2: a quantitative trait locus key to the evolution of tomato fruit size. Science (New York, N.Y.) **289** (5476): 85–88.

Galaviz-Hernandez, C., Stagg, C., Ridder, G. de, **Tanaka, T.S., Ko, M.S.H., Schlessinger, D., and Nagaraja, R.** (2003). Plac8 and Plac9, novel placental-enriched genes identified through microarray analysis. Gene **309** (2): 81–89.

Gao, C., Yu, C.K.Y., Qu, S., San, M.W.Y., Li, K.Y., Lo, S.W., and Jiang, L. (2012). The Golgi-Localized Arabidopsis Endomembrane Protein12 Contains Both Endoplasmic Reticulum Export and Golgi Retention Signals at Its C TerminusCW. The Plant cell **24** (5): 2086–2104.

Ge, Z., Zhao, Y., Liu, M.-C., Zhou, L.-Z., Wang, L., Zhong, S., Hou, S., Jiang, J., Liu, T., Huang, Q., Xiao, J., Gu, H., Wu, H.-M., Dong, J., Dresselhaus, T., Cheung, A.Y., and Qu, L.-J. (2019). LLG2/3 Are Co-receptors in BUPS/ANX-RALF Signaling to Regulate Arabidopsis Pollen Tube Integrity. Current biology CB **29** (19): 3256-3265.e5.

Gebert, M., Dresselhaus, T., and Sprunck, S. (2008). F-actin organization and pollen tube tip growth in Arabidopsis are dependent on the gametophyte-specific Armadillo repeat protein ARO1. The Plant cell **20** (10): 2798–2814.

Gelvin, S.B., and Schilperoort, R.A., eds (1994). Plant Molecular Biology Manual (Dordrecht: Springer).

Gibbons, D.L., Vaney, M.-C., Roussel, A., Vigouroux, A., Reilly, B., Lepault, J., Kielian, M., and Rey, F.A. (2004). Conformational change and protein-protein interactions of the fusion protein of Semliki Forest virus. Nature **427** (6972): 320–325.

Goodstein, D.M., Shu, S., Howson, R., Neupane, R., Hayes, R.D., Fazo, J., Mitros, T., Dirks, W., Hellsten, U., Putnam, N., and Rokhsar, D.S. (2012). Phytozome: a comparative platform for green plant genomics. Nucleic acids research **40** (Database issue): D1178-86.

Gou, X., Yuan, T., Wei, X., and Russell, S.D. (2009). Gene expression in the dimorphic sperm cells of Plumbago zeylanica: transcript profiling, diversity, and relationship to cell type. The Plant journal for cell and molecular biology **60** (1): 33–47.

Graham, J.M., ed (2002). Subcellular fractionation: A practical approach (Oxford: IRL Press).

Grignon, C., and Sentenac, H. (1991). pH and Ionic Conditions in the Apoplast. Annu. Rev. Plant. Physiol. Plant. Mol. Biol. **42** (1): 103–128.

Grossniklaus, U., Vielle-Calzada, J.P., Hoeppner, M.A., and Gagliano, W.B. (1998). Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. Science (New York, N.Y.) **280** (5362): 446–450.

Guo, M., Rupe, M.A., Dieter, J.A., Zou, J., Spielbauer, D., Duncan, K.E., Howard, R.J., Hou, Z., and Simmons, C.R. (2010). Cell Number Regulator1 affects plant and organ size in maize: implications for crop yield enhancement and heterosis. The Plant cell **22** (4): 1057–1073.

Hamamura, Y., Nishimaki, M., Takeuchi, H., Geitmann, A., Kurihara, D., and Higashiyama, T. (2014). Live imaging of calcium spikes during double fertilization in Arabidopsis. Nature communications **5**: 4722.

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

Han, S.W., Alonso, J.M., and Rojas-Pierce, M. (2015). REGULATOR OF BULB BIOGENESIS1 (RBB1) Is Involved in Vacuole Bulb Formation in Arabidopsis. PloS one **10** (4): e0125621.

Helenius, A., Kartenbeck, J., Simons, K., and Fries, E. (1980). On the entry of Semliki forest virus into BHK-21 cells. The Journal of cell biology **84** (2): 404–420.

Hemler, M.E. (2005). Tetraspanin functions and associated microdomains. Nature reviews. Molecular cell biology **6** (10): 801–811.

Hernández, J.M., and Podbilewicz, B. (2017). The hallmarks of cell-cell fusion. Development (Cambridge, England) **144** (24): 4481–4495.

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

Higo, A., Kawashima, T., Borg, M., Zhao, M., López-Vidriero, I., Sakayama, H., Montgomery, S.A., Sekimoto, H., Hackenberg, D., Shimamura, M., Nishiyama, T., Sakakibara, K., Tomita, Y., Togawa, T., Kunimoto, K., Osakabe, A., Suzuki, Y., Yamato, K.T., Ishizaki, K., Nishihama, R., Kohchi, T., Franco-Zorrilla, J.M., Twell, D., Berger, F., and Araki, T. (2018). Transcription factor DUO1 generated by neo-functionalization is associated with evolution of sperm differentiation in plants. Nature communications **9** (1): 5283.

Holmes-Davis, R., Tanaka, C.K., Vensel, W.H., Hurkman, W.J., and McCormick, S. (2005). Proteome mapping of mature pollen of Arabidopsis thaliana. Proteomics **5** (18): 4864–4884.

Hooper, C.M., Tanz, S.K., Castleden, I.R., Vacher, M.A., Small, I.D., and Millar, A.H. (2014). SUBAcon: a consensus algorithm for unifying the subcellular localization data of the Arabidopsis proteome. Bioinformatics (Oxford, England) **30** (23): 3356–3364.

Horton, P., Park, K.-J., Obayashi, T., Fujita, N., Harada, H., Adams-Collier, C.J., and Nakai, K. (2007). WoLF PSORT: protein localization predictor. Nucleic acids research **35** (Web Server issue): W585-7.

Hou, Y., Guo, X., Cyprys, P., Zhang, Y., Bleckmann, A., Le Cai, Huang, Q., Luo, Y., Gu, H., Dresselhaus, T., Dong, J., and Qu, L.-J. (2016). Maternal ENODLs Are Required for Pollen Tube Reception in Arabidopsis. Current biology CB **26** (17): 2343–2350.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W., and Zimmermann, P. (2008). Genevestigator v3: a reference expression database for the meta-analysis of transcriptomes. Advances in bioinformatics **2008**: 420747.

Hu, L., Smith, T.F., and Goldberger, G. (2009). LFG: a candidate apoptosis regulatory gene family. Apoptosis an international journal on programmed cell death **14** (11): 1255–1265.

Huala, E., Dickerman, A.W., Garcia-Hernandez, M., Weems, D., Reiser, L., LaFond, F., Hanley, D., Kiphart, D., Zhuang, M., Huang, W., Mueller, L.A., Bhattacharyya, D., Bhaya, D., Sobral, B.W., Beavis, W., Meinke, D.W., Town, C.D., Somerville, C., and Rhee, S.Y. (2001). The Arabidopsis Information Resource (TAIR): a comprehensive database and web-based information retrieval, analysis, and visualization system for a model plant. Nucleic acids research **29** (1): 102–105.

Huang, S., Yuan, S., Dong, M., Su, J., Yu, C., Shen, Y., Xie, X., Yu, Y., Yu, X., Chen, S., Zhang, S., Pontarotti, P., and Xu, A. (2005). The phylogenetic analysis of tetraspanins projects the evolution of cell-cell interactions from unicellular to multicellular organisms. Genomics **86** (6): 674–684.

Hückelhoven, R. (2004). BAX Inhibitor-1, an ancient cell death suppressor in animals and plants with prokaryotic relatives. Apoptosis an international journal on programmed cell death **9** (3): 299–307.

Ingouff, M., Hamamura, Y., Gourgues, M., Higashiyama, T., and Berger, F. (2007). Distinct dynamics of HISTONE3 variants between the two fertilization products in plants. Current biology CB **17** (12): 1032–1037.

Ingouff, M., Sakata, T., Li, J., Sprunck, S., Dresselhaus, T., and Berger, F. (2009). The two male gametes share equal ability to fertilize the egg cell in Arabidopsis thaliana. Current biology CB **19** (1): R19-20.

Inoue, H., Nojima, H., and Okayama, H. (1990). High efficiency transformation of Escherichia coli with plasmids. Gene **96** (1): 23–28.

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

Ishikawa, T., Watanabe, N., Nagano, M., Kawai-Yamada, M., and Lam, E. (2011). Bax inhibitor-1: a highly conserved endoplasmic reticulum-resident cell death suppressor. Cell death and differentiation **18** (8): 1271–1278.

Ito, C., Yamatoya, K., Yoshida, K., Maekawa, M., Miyado, K., and Toshimori, K. (2010). Tetraspanin family protein CD9 in the mouse sperm: unique localization, appearance, behavior and fate during fertilization. Cell and tissue research **340** (3): 583–594.

Jahn, R., Lang, T., and Südhof, T.C. (2003). Membrane Fusion. Cell **112** (4): 519–533.

Jégou, A., Ziyyat, A., Barraud-Lange, V., Perez, E., Wolf, J.P., Pincet, F., and Gourier, C. (2011). CD9 tetraspanin generates fusion competent sites on the egg membrane for mammalian fertilization. Proceedings of the National Academy of Sciences of the United States of America **108** (27): 10946–10951.

Jiang, H., Yi, J., Boavida, L.C., Chen, Y., Becker, J.D., Köhler, C., and McCormick, S. (2015). Intercellular communication in Arabidopsis thaliana pollen discovered via AHG3 transcript movement from the vegetative cell to sperm. Proceedings of the National Academy of Sciences of the United States of America **112** (43): 13378–13383.

Johnson, M.A., Besser, K. von, **Zhou, Q., Smith, E., Aux, G., Patton, D., Levin, J.Z., and Preuss, D.** (2004). Arabidopsis hapless mutations define essential gametophytic functions. Genetics **168** (2): 971–982.

Johnson, M.A., Harper, J.F., and Palanivelu, R. (2019). A Fruitful Journey: Pollen Tube Navigation from Germination to Fertilization. Annual review of plant biology **70**: 809–837.

Johnson-Brousseau, S.A., and McCormick, S. (2004). A compendium of methods useful for characterizing Arabidopsis pollen mutants and gametophytically-expressed genes. The Plant journal for cell and molecular biology **39** (5): 761–775.

Kaji, K., Oda, S., Shikano, T., Ohnuki, T., Uematsu, Y., Sakagami, J., Tada, N., Miyazaki, S., and Kudo, A. (2000). The gamete fusion process is defective in eggs of Cd9-deficient mice. Nature genetics **24** (3): 279–282.

Karimi, M., Inzé, D., and Depicker, A. (2002). GATEWAY vectors for Agrobacterium-mediated plant transformation. Trends in plant science **7** (5): 193–195.

Karve, R., Liu, W., Willet, S.G., Torii, K.U., and Shpak, E.D. (2011). The presence of multiple introns is essential for ERECTA expression in Arabidopsis. RNA **17** (10): 1907–1921.

Kasahara, R.D., Maruyama, D., Hamamura, Y., Sakakibara, T., Twell, D., and Higashiyama, T. (2012). Fertilization recovery after defective sperm cell release in Arabidopsis. Current biology CB **22** (12): 1084–1089.

Kasahara, R.D., Notaguchi, M., Nagahara, S., Suzuki, T., Susaki, D., Honma, Y., Maruyama, D., and Higashiyama, T. (2016). Pollen tube contents initiate ovule enlargement and enhance seed coat development without fertilization. Science advances **2** (10): e1600554.

Kasaras, A., and Kunze, R. (2010). Expression, localisation and phylogeny of a novel family of plant-specific membrane proteins. Plant biology (Stuttgart, Germany) **12 Suppl 1**: 140–152.

Kasaras, A., and Kunze, R. (2017). Dual-targeting of Arabidopsis DMP1 isoforms to the tonoplast and the plasma membrane. PloS one **12** (4): e0174062.

Kasaras, A., Melzer, M., and Kunze, R. (2012). Arabidopsis senescence-associated protein DMP1 is involved in membrane remodeling of the ER and tonoplast. BMC plant biology **12**: 54.

Kelliher, T., Starr, D., Richbourg, L., Chintamanani, S., Delzer, B., Nuccio, M.L., Green, J., Chen, Z., McCuiston, J., Wang, W., Liebler, T., Bullock, P., and Martin, B. (2017). MATRILINEAL, a sperm-specific phospholipase, triggers maize haploid induction. Nature **542** (7639): 105–109.

Keshishian, H., Addona, T., Burgess, M., Kuhn, E., and Carr, S.A. (2007). Quantitative, multiplexed assays for low abundance proteins in plasma by targeted mass spectrometry and stable isotope dilution. Molecular & cellular proteomics MCP **6** (12): 2212–2229.

Kielian, M., and Rey, F.A. (2006). Virus membrane-fusion proteins: more than one way to make a hairpin. Nature reviews. Microbiology **4** (1): 67–76.

Klimjack, M.R., Jeffrey, S., and Kielian, M. (1994). Membrane and protein interactions of a soluble form of the Semliki Forest virus fusion protein. Journal of virology **68** (11): 6940–6946.

Laemmli, U.K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature **227** (5259): 680–685.

Lamesch, P., Berardini, T.Z., Li, D., Swarbreck, D., Wilks, C., Sasidharan, R., Muller, R., Dreher, K., Alexander, D.L., Garcia-Hernandez, M., Karthikeyan, A.S., Lee, C.H., Nelson, W.D., Ploetz, L., Singh, S., Wensel, A., and Huala, E. (2012). The Arabidopsis Information

Resource (TAIR): improved gene annotation and new tools. Nucleic acids research **40** (Database issue): D1202-10.

Lenzi, A., Picardo, M., Gandini, L., and Dondero, F. (1996). Lipids of the sperm plasma membrane: from polyunsaturated fatty acids considered as markers of sperm function to possible scavenger therapy. Human reproduction update **2** (3): 246–256.

Lescar, J., Roussel, A., Wien, M.W., Navaza, J., Fuller, S.D., Wengler, G., and Rey, F.A. (2001). The Fusion glycoprotein shell of Semliki Forest virus: an icosahedral assembly primed for fusogenic activation at endosomal pH. Cell **105** (1): 137–148.

Levin, T.C., and King, N. (2013). Evidence for sex and recombination in the choanoflagellate Salpingoeca rosetta. Current biology CB **23** (21): 2176–2180.

Leydon, A.R., Tsukamoto, T., Dunatunga, D., Qin, Y., Johnson, M.A., and Palanivelu, R. (2015). Pollen Tube Discharge Completes the Process of Synergid Degeneration That Is Initiated by Pollen Tube-Synergid Interaction in Arabidopsis. Plant physiology **169** (1): 485–496.

Li, H.-T., Yi, T.-S., Gao, L.-M., Ma, P.-F., Zhang, T., Yang, J.-B., Gitzendanner, M.A., Fritsch, P.W., Cai, J., Luo, Y., Wang, H., van der Bank, M., Zhang, S.-D., Wang, Q.-F., Wang, J., Zhang, Z.-R., Fu, C.-N., Yang, J., Hollingsworth, P.M., Chase, M.W., Soltis, D.E., Soltis, P.S., and Li, D.-Z. (2019). Origin of angiosperms and the puzzle of the Jurassic gap. Nature Plants **5** (5): 461–470.

Li, X. (2011). Pollen Fertility/viability Assay Using FDA Staining. BIO-PROTOCOL **1** (10).

Li, X., Zhang, R., Patena, W., Gang, S.S., Blum, S.R., Ivanova, N., Yue, R., Robertson, J.M., Lefebvre, P.A., Fitz-Gibbon, S.T., Grossman, A.R., and Jonikas, M.C. (2016). An Indexed, Mapped Mutant Library Enables Reverse Genetics Studies of Biological Processes in Chlamydomonas reinhardtii. The Plant cell **28** (2): 367–387.

Libault, M., and Stacey, G. (2010). Evolution of FW2.2-like (FWL) and PLAC8 genes in eukaryotes. Plant Signaling & Behavior **5** (10): 1226–1228.

Libault, M., Zhang, X.-C., Govindarajulu, M., Qiu, J., Ong, Y.T., Brechenmacher, L., Berg, R.H., Hurley-Sommer, A., Taylor, C.G., and Stacey, G. (2010). A member of the highly conserved FWL (tomato FW2.2-like) gene family is essential for soybean nodule organogenesis. The Plant journal for cell and molecular biology **62** (5): 852–864.

Lindemeier, M. (2017). Direct gamete interactions and Direct gamete interactions and EGG CELL1-mediated sperm cell activation in Arabidopsis thaliana. Dissertation.

Linder, M.E., and Deschenes, R.J. (2007). Palmitoylation: policing protein stability and traffic. Nature reviews. Molecular cell biology **8** (1): 74–84.

Liu, W., Xie, X., Ma, X., Li, J., Chen, J., and Liu, Y.-G. (2015a). DSDecode: A Web-Based Tool for Decoding of Sequencing Chromatograms for Genotyping of Targeted Mutations. Molecular plant **8** (9): 1431–1433.

Liu, X., Adhikari, P.B., and Kasahara, R.D. (2019). Pollen tube contents from failed fertilization contribute to seed coat initiation in Arabidopsis. F1000Research **8**: 348.

Liu, Y., Pei, J., Grishin, N., and Snell, W.J. (2015b). The cytoplasmic domain of the gamete membrane fusion protein HAP2 targets the protein to the fusion site in Chlamydomonas and regulates the fusion reaction. Development (Cambridge, England) **142** (5): 962–971.

Liu, Y., Tewari, R., Ning, J., Blagborough, A.M., Garbom, S., Pei, J., Grishin, N.V., Steele, R.E., Sinden, R.E., Snell, W.J., and Billker, O. (2008). The conserved plant sterility gene HAP2 functions after attachment of fusogenic membranes in Chlamydomonas and Plasmodium gametes. Genes & development **22** (8): 1051–1068.

Lund, A., and Fuglsang, A.T. (2012). Purification of plant plasma membranes by two-phase partitioning and measurement of H+ pumping. Methods in molecular biology (Clifton, N.J.) **913**: 217–223.

Malho, R., and Trewavas, A.J. (1996). Localized Apical Increases of Cytosolic Free Calcium Control Pollen Tube Orientation. The Plant cell **8** (11): 1935–1949.

Márton, M.L., Cordts, S., Broadhvest, J., and Dresselhaus, T. (2005). Micropylar pollen tube guidance by egg apparatus 1 of maize. Science (New York, N.Y.) **307** (5709): 573–576.

Maruyama, D., Völz, R., Takeuchi, H., Mori, T., Igawa, T., Kurihara, D., Kawashima, T., Ueda, M., Ito, M., Umeda, M., Nishikawa, S.-I., Groß-Hardt, R., and Higashiyama, T. (2015). Rapid Elimination of the Persistent Synergid through a Cell Fusion Mechanism. Cell **161** (4): 907–918.

Mayank, P., Grossman, J., Wuest, S., Boisson-Dernier, A., Roschitzki, B., Nanni, P., Nühse, T., and Grossniklaus, U. (2012). Characterization of the phosphoproteome of mature Arabidopsis pollen. The Plant journal for cell and molecular biology **72** (1): 89–101.

McCue, A.D., Cresti, M., Feijó, J.A., and Slotkin, R.K. (2011). Cytoplasmic connection of sperm cells to the pollen vegetative cell nucleus: potential roles of the male germ unit revisited. Journal of experimental botany **62** (5): 1621–1631.

Mecchia, M.A., Santos-Fernandez, G., Duss, N.N., Somoza, S.C., Boisson-Dernier, A., Gagliardini, V., Martínez-Bernardini, A., Fabrice, T.N., Ringli, C., Muschietti, J.P., and Grossniklaus, U. (2017). RALF4/19 peptides interact with LRX proteins to control pollen tube growth in Arabidopsis. Science (New York, N.Y.) **358** (6370): 1600–1603.

Merchant, S.S., Prochnik, S.E., Vallon, O., Harris, E.H., Karpowicz, S.J., Witman, G.B., Terry, A., Salamov, A., Fritz-Laylin, L.K., Maréchal-Drouard, L., Marshall, W.F., Qu, L.-H., Nelson, D.R., Sanderfoot, A.A., Spalding, M.H., Kapitonov, V.V., Ren, Q., Ferris, P., Lindquist, E., Shapiro, H., Lucas, S.M., Grimwood, J., Schmutz, J., Cardol, P., Cerutti, H., Chanfreau, G., Chen, C.-L., Cognat, V., Croft, M.T., Dent, R., Dutcher, S., Fernández, E., Fukuzawa, H., González-Ballester, D., González-Halphen, D., Hallmann, A., Hanikenne, M., Hippler, M., Inwood, W., Jabbari, K., Kalanon, M., Kuras, R., Lefebvre, P.A., Lemaire, S.D., Lobanov, A.V., Lohr, M., Manuell, A., Meier, I., Mets, L., Mittag, M., Mittelmeier, T., Moroney, J.V., Moseley, J., Napoli, C., Nedelcu, A.M., Niyogi, K., Novoselov, S.V., Paulsen, I.T., Pazour, G., Purton, S., Ral, J.-P., Riaño-Pachón, D.M., Riekhof, W., Rymarquis, L., Schroda, M., Stern, D., Umen, J., Willows, R., Wilson, N., Zimmer, S.L., Allmer, J., Balk, J., Bisova, K., Chen, C.-J., Elias, M., Gendler, K., Hauser, C., Lamb, M.R., Ledford, H., Long, J.C., Minagawa, J., Page, M.D., Pan, J., Pootakham, W., Roje, S., Rose, A., Stahlberg, E., Terauchi, A.M., Yang, P., Ball, S., Bowler, C., Dieckmann, C.L., Gladyshev, V.N., Green, P., Jorgensen, R., Mayfield, S., Mueller-Roeber, B., Rajamani, S., Sayre, R.T., Brokstein, P., Dubchak, I., Goodstein, D., Hornick, L., Huang, Y.W., Jhaveri, J., Luo, Y., Martínez, D., Ngau, W.C.A., Otillar, B., Poliakov, A., Porter, A., Szajkowski, L., Werner, G., Zhou, K., Grigoriev, I.V., Rokhsar, D.S., and Grossman, A.R. (2007). The Chlamydomonas genome reveals the evolution of key animal and plant functions. Science (New York, N.Y.) **318** (5848): 245–250.

Mergner, J., Frejno, M., List, M., Papacek, M., Chen, X., Chaudhary, A., Samaras, P., Richter, S., Shikata, H., Messerer, M., Lang, D., Altmann, S., Cyprys, P., Zolg, D.P., Mathieson, T., Bantscheff, M., Hazarika, R.R., Schmidt, T., Dawid, C., Dunkel, A., Hofmann, T., Sprunck, S., Falter-Braun, P., Johannes, F., Mayer, K.F.X., Jürgens, G., Wilhelm, M., Baumbach, J., Grill, E., Schneitz, K., Schwechheimer, C., and Kuster, B. (2020). Mass-spectrometry-based draft of the Arabidopsis proteome. Nature **579** (7799): 409–414.

Misamore, M.J., Gupta, S., and Snell, W.J. (2003). The Chlamydomonas Fus1 Protein Is Present on the Mating Type plus Fusion Organelle and Required for a Critical Membrane Adhesion Event during Fusion with minus Gametes. Molecular Biology of the Cell **14** (6): 2530–2542.

Misra, C.S., Santos, M.R., Rafael-Fernandes, M., Martins, N.P., Monteiro, M., and Becker, J.D. (2019). Transcriptomics of Arabidopsis sperm cells at single-cell resolution. Plant reproduction **32** (1): 29–38.

Mitchell, A.L., Attwood, T.K., Babbitt, P.C., Blum, M., Bork, P., Bridge, A., Brown, S.D., Chang, H.-Y., El-Gebali, S., Fraser, M.I., Gough, J., Haft, D.R., Huang, H., Letunic, I., Lopez, R., Luciani, A., Madeira, F., Marchler-Bauer, A., Mi, H., Natale, D.A., Necci, M., Nuka, G., Orengo, C., Pandurangan, A.P., Paysan-Lafosse, T., Pesseat, S., Potter, S.C., Qureshi, M.A., Rawlings, N.D., Redaschi, N., Richardson, L.J., Rivoire, C., Salazar, G.A., Sangrador-Vegas, A., Sigrist, C.J.A., Sillitoe, I., Sutton, G.G., Thanki, N., Thomas, P.D., Tosatto, S.C.E., Yong, S.-Y., and Finn, R.D. (2019). InterPro in 2019: improving coverage, classification and access to protein sequence annotations. Nucleic acids research **47** (D1): D351-D360.

Mizukami, A.G., Inatsugi, R., Jiao, J., Kotake, T., Kuwata, K., Ootani, K., Okuda, S., Sankaranarayanan, S., Sato, Y., Maruyama, D., Iwai, H., Garénaux, E., Sato, C., Kitajima, K., Tsumuraya, Y., Mori, H., Yamaguchi, J., Itami, K., Sasaki, N., and Higashiyama, T. (2016). The AMOR Arabinogalactan Sugar Chain Induces Pollen-Tube Competency to Respond to Ovular Guidance. Current biology CB **26** (8): 1091–1097.

Modis, Y., Ogata, S., Clements, D., and Harrison, S.C. (2003). A ligand-binding pocket in the dengue virus envelope glycoprotein. Proceedings of the National Academy of Sciences of the United States of America **100** (12): 6986–6991.

Modis, Y., Ogata, S., Clements, D., and Harrison, S.C. (2004). Structure of the dengue virus envelope protein after membrane fusion. Nature **427** (6972): 313–319.

Mori, T., Hirai, M., Kuroiwa, T., and Miyagishima, S.-Y. (2010). The functional domain of GCS1-based gamete fusion resides in the amino terminus in plant and parasite species. PloS one **5** (12): e15957.

Mori, T., Igawa, T., Tamiya, G., Miyagishima, S.-Y., and Berger, F. (2014). Gamete attachment requires GEX2 for successful fertilization in Arabidopsis. Current biology CB **24** (2): 170–175.

Mori, T., Kuroiwa, H., Higashiyama, T., and Kuroiwa, T. (2006). GENERATIVE CELL SPECIFIC 1 is essential for angiosperm fertilization. Nature cell biology **8** (1): 64–71.

Nagahara, S., Takeuchi, H., and Higashiyama, T. (2015). Generation of a homozygous fertilization-defective gcs1 mutant by heat-inducible removal of a rescue gene. Plant reproduction **28** (1): 33–46.

Nagahashi, J., and Hiraike, K. (1982). Effects of centrifugal force and centrifugation time on the sedimentation of plant organelles. Plant physiology **69** (2): 546–548.

Nagaya, S., Kawamura, K., Shinmyo, A., and Kato, K. (2010). The HSP terminator of Arabidopsis thaliana increases gene expression in plant cells. Plant & cell physiology **51** (2): 328–332.

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

Navašin, S. (1898). Resultate einer Revision der Befruchtungsvorgänge bei Lilium Martagon und Fritillaria tenella.

Nelson, B.K., Cai, X., and Nebenführ, A. (2007). A multicolored set of in vivo organelle markers for co-localization studies in Arabidopsis and other plants. The Plant journal for cell and molecular biology **51** (6): 1126–1136.

Ning, J., Otto, T.D., Pfander, C., Schwach, F., Brochet, M., Bushell, E., Goulding, D., Sanders, M., Lefebvre, P.A., Pei, J., Grishin, N.V., Vanderlaan, G., Billker, O., and Snell, W.J. (2013). Comparative genomics in Chlamydomonas and Plasmodium identifies an ancient nuclear envelope protein family essential for sexual reproduction in protists, fungi, plants, and vertebrates. Genes & development **27** (10): 1198–1215.

Noir, S., Bräutigam, A., Colby, T., Schmidt, J., and Panstruga, R. (2005). A reference map of the Arabidopsis thaliana mature pollen proteome. Biochemical and biophysical research communications **337** (4): 1257–1266.

Ohad, N., Margossian, L., Hsu, Y.C., Williams, C., Repetti, P., and Fischer, R.L. (1996). A mutation that allows endosperm development without fertilization. Proceedings of the National Academy of Sciences of the United States of America **93** (11): 5319–5324.

Ohno, S. (1970). Evolution by Gene Duplication (Berlin, Heidelberg: Springer Berlin Heidelberg).

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

Oliveros, J.C., Franch, M., Tabas-Madrid, D., San-León, D., Montoliu, L., Cubas, P., and Pazos, F. (2016). Breaking-Cas-interactive design of guide RNAs for CRISPR-Cas experiments for ENSEMBL genomes. Nucleic acids research **44** (W1): W267-71.

Omasits, U., Ahrens, C.H., Müller, S., and Wollscheid, B. (2014). Protter: interactive protein feature visualization and integration with experimental proteomic data. Bioinformatics (Oxford, England) **30** (6): 884–886.

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

Palanivelu, R., and Tsukamoto, T. (2012). Pathfinding in angiosperm reproduction: pollen tube guidance by pistils ensures successful double fertilization. Wiley interdisciplinary reviews. Developmental biology **1** (1): 96–113.

Pan, J., Misamore, M.J., Wang, Q., and Snell, W.J. (2003). Protein transport and signal transduction during fertilization in chlamydomonas. Traffic (Copenhagen, Denmark) **4** (7): 452–459.

Pan, J., and Snell, W.J. (2000). Signal transduction during fertilization in the unicellular green alga, Chlamydomonas. Current opinion in microbiology **3** (6): 596–602.

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

Poschner, B.C., and Langosch, D. (2009). Stabilization of conformationally dynamic helices by covalently attached acyl chains. Protein science a publication of the Protein Society **18** (8): 1801–1805.

Rademacher, S., and Sprunck, S. (2013). Downregulation of egg cell-secreted EC1 is accompanied with delayed gamete fusion and polytubey. Plant Signaling & Behavior **8** (12).

Ravaux, B., Favier, S., Perez, E., and Gourier, C. (2018). Egg CD9 protein tides correlated with sperm oscillations tune the gamete fusion ability in mammal. Journal of molecular cell biology **10** (6): 494–502.

Redman, J.C., Haas, B.J., Tanimoto, G., and Town, C.D. (2004). Development and evaluation of an Arabidopsis whole genome Affymetrix probe array. The Plant journal for cell and molecular biology **38** (3): 545–561.

Ren, J., Wu, P., Trampe, B., Tian, X., Lübberstedt, T., and Chen, S. (2017). Novel technologies in doubled haploid line development. Plant biotechnology journal **15** (11): 1361–1370.

Reski, R., and Cove, D.J. (2004). Physcomitrella patens. Current biology CB **14** (7): R261-2.

Rey, F.A., Heinz, F.X., Mandl, C., Kunz, C., and Harrison, S.C. (1995). The envelope glycoprotein from tick-borne encephalitis virus at 2 A resolution. Nature **375** (6529): 291–298.

Rochwerger, L., and Cuasnicu, P.S. (1992). Redistribution of a rat sperm epididymal glycoprotein after in vitro and in vivo capacitation. Molecular reproduction and development **31** (1): 34–41.

Rose, A.B., Elfersi, T., Parra, G., and Korf, I. (2008). Promoter-Proximal Introns in Arabidopsis thaliana Are Enriched in Dispersed Signals that Elevate Gene ExpressionWOA. The Plant cell **20** (3): 543–551.

Rotman, N., Durbarry, A., Wardle, A., Yang, W.C., Chaboud, A., Faure, J.-E., Berger, F., and Twell, D. (2005). A novel class of MYB factors controls sperm-cell formation in plants. Current biology CB **15** (3): 244–248.

Rubinstein, E., Ziyyat, A., Prenant, M., Wrobel, E., Wolf, J.-P., Levy, S., Le Naour, F., and Boucheix, C. (2006). Reduced fertility of female mice lacking CD81. Developmental biology **290** (2): 351–358.

Runions, J., Hawes, C., and Kurup, S. (2007). Fluorescent protein fusions for protein localization in plants. Methods in molecular biology (Clifton, N.J.) **390**: 239–255.

Russell, S.D. (1984). Ultrastructure of the sperm of Plumbago zeylanica II. Quantitative cytology and three-dimensional organization. Planta **162** (5): 385–391.

Russell, S.D., and Cass, D.D. (1981). Ultrastructure of the sperms ofPlumbago zeylanica. Protoplasma **107** (1-2): 85–107.

Russell, S.D., Gou, X., Wong, C.E., Wang, X., Yuan, T., Wei, X., Bhalla, P.L., and Singh, M.B. (2012). Genomic profiling of rice sperm cell transcripts reveals conserved and distinct elements in the flowering plant male germ lineage. The New phytologist **195** (3): 560–573.

Russell, S.D., Jones, D.S., Anderson, S., Wang, X., Sundaresan, V., and Gou, X. (2017). Isolation of Rice Sperm Cells for Transcriptional Profiling. Methods in molecular biology (Clifton, N.J.) **1669**: 211–219.

Saito, C., Ueda, T., Abe, H., Wada, Y., Kuroiwa, T., Hisada, A., Furuya, M., and Nakano, A. (2002). A complex and mobile structure forms a distinct subregion within the continuous vacuolar membrane in young cotyledons of Arabidopsis. The Plant journal for cell and molecular biology **29** (3): 245–255.

Saito, C., Uemura, T., Awai, C., Tominaga, M., Ebine, K., Ito, J., Ueda, T., Abe, H., Morita, M.T., Tasaka, M., and Nakano, A. (2011). The occurrence of 'bulbs', a complex configuration of the vacuolar membrane, is affected by mutations of vacuolar SNARE and phospholipase in Arabidopsis. The Plant journal for cell and molecular biology **68** (1): 64–73.

Sambrook, J., Fritsch, E., and Maniatis, T. (1989). Molecular cloning: A laboratory manual Vol. 2 (S.l.: Cold Spring Harbor).

Sano, R., Hou, Y.-C.C., Hedvat, M., Correa, R.G., Shu, C.-W., Krajewska, M., Diaz, P.W., Tamble, C.M., Quarato, G., Gottlieb, R.A., Yamaguchi, M., Nizet, V., Dahl, R., Thomas, D.D., Tait, S.W., Green, D.R., Fisher, P.B., Matsuzawa, S.-I., and Reed, J.C. (2012). Endoplasmic reticulum protein BI-1 regulates Ca^{2+}-mediated bioenergetics to promote autophagy. Genes & development **26** (10): 1041–1054.

Satouh, Y., Inoue, N., Ikawa, M., and Okabe, M. (2012). Visualization of the moment of mouse sperm-egg fusion and dynamic localization of IZUMO1. Journal of cell science **125** (Pt 21): 4985–4990.

Schägger, H., and Jagow, G. von (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. Analytical biochemistry **166** (2): 368–379.

Schmidt, A., Schmid, M.W., and Grossniklaus, U. (2015). Plant germline formation: common concepts and developmental flexibility in sexual and asexual reproduction. Development (Cambridge, England) **142** (2): 229–241.

Schneider, C.A., Rasband, W.S., and Eliceiri, K.W. (2012). NIH Image to ImageJ: 25 years of image analysis. Nature methods **9** (7): 671–675.

Schwacke, R., Schneider, A., van der Graaff, E., Fischer, K., Catoni, E., Desimone, M., Frommer, W.B., Flügge, U.-I., and Kunze, R. (2003). ARAMEMNON, a novel database for Arabidopsis integral membrane proteins. Plant physiology **131** (1): 16–26.

Segami, S., Makino, S., Miyake, A., Asaoka, M., and Maeshima, M. (2014). Dynamics of vacuoles and H+-pyrophosphatase visualized by monomeric green fluorescent protein in Arabidopsis: artifactual bulbs and native intravacuolar spherical structures. The Plant cell **26** (8): 3416–3434.

Sheoran, I.S., Sproule, K.A., Olson, D.J.H., Ross, A.R.S., and Sawhney, V.K. (2006). Proteome profile and functional classification of proteins in Arabidopsis thaliana (Landsberg erecta) mature pollen. Sex Plant Reprod **19** (4): 185–196.

Shi, L., Zhu, T., Mogensen, H.L., and Keim, P. (1996). Sperm Identification in Maize by Fluorescence in Situ Hybridization. The Plant cell **8** (5): 815–821.

Shiu, S.H., and Bleecker, A.B. (2001). Receptor-like kinases from Arabidopsis form a monophyletic gene family related to animal receptor kinases. Proceedings of the National Academy of Sciences of the United States of America **98** (19): 10763–10768.

Shiu, S.-H., Karlowski, W.M., Pan, R., Tzeng, Y.-H., Mayer, K.F.X., and Li, W.-H. (2004). Comparative analysis of the receptor-like kinase family in Arabidopsis and rice. The Plant cell **16** (5): 1220–1234.

Simonin, K.A., and Roddy, A.B. (2018). Genome downsizing, physiological novelty, and the global dominance of flowering plants. PLoS biology **16** (1): e2003706.

Skinner, D.J., and Sundaresan, V. (2018). Recent advances in understanding female gametophyte development. F1000Research **7**.

Sklodowski, K., Riedelsberger, J., Raddatz, N., Riadi, G., Caballero, J., Chérel, I., Schulze, W., Graf, A., and Dreyer, I. (2017). The receptor-like pseudokinase MRH1 interacts with the voltage-gated potassium channel AKT2. Scientific reports **7**: 44611.

Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzić, M., Becker, J.D., Feijó, J.A., and Martienssen, R.A. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell **136** (3): 461–472.

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

Song, W.-Y., Choi, K.S., Kim, D.Y., Geisler, M., Park, J., Vincenzetti, V., Schellenberg, M., Kim, S.H., Lim, Y.P., Noh, E.W., Lee, Y., and Martinoia, E. (2010). Arabidopsis PCR2 Is a Zinc Exporter Involved in Both Zinc Extrusion and Long-Distance Zinc TransportW. The Plant cell **22** (7): 2237–2252.

Song, W.-Y., Choi, K.-S., Alexis, D.A., Martinoia, E., and Lee, Y. (2011a). Brassica juncea plant cadmium resistance 1 protein (BjPCR1) facilitates the radial transport of calcium in the root. Proceedings of the National Academy of Sciences of the United States of America **108** (49): 19808–19813.

Song, W.-Y., Hörtensteiner, S., Tomioka, R., Lee, Y., and Martinoia, E. (2011b). Common functions or only phylogenetically related? The large family of PLAC8 motif-containing/PCR genes. Molecules and cells **31** (1): 1–7.

Song, W.-Y., Martinoia, E., Lee, J., Kim, D., Kim, D.-Y., Vogt, E., Shim, D., Choi, K.S., Hwang, I., and Lee, Y. (2004). A Novel Family of Cys-Rich Membrane Proteins Mediates Cadmium Resistance in Arabidopsis1. Plant physiology **135** (2): 1027–1039.

Southworth, D., and Cresti, M. (1997). Comparison of flagellated and nonflagellated sperm in plants. American journal of botany **84** (9): 1301.

Sparkes, I.A., Runions, J., Kearns, A., and Hawes, C. (2006). Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. Nature protocols **1** (4): 2019–2025.

Speijer, D. (2016). What can we infer about the origin of sex in early eukaryotes? Philosophical transactions of the Royal Society of London. Series B, Biological sciences **371** (1706).

Sprunck, S. (2020). Twice the fun, double the trouble: gamete interactions in flowering plants. Current opinion in plant biology **53**: 106–116.

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

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

Steele, R.E., and Dana, C.E. (2009). Evolutionary history of the HAP2/GCS1 gene and sexual reproduction in metazoans. PloS one **4** (11): e7680.

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

Stiasny, K., Allison, S.L., Schalich, J., and Heinz, F.X. (2002). Membrane interactions of the tick-borne encephalitis virus fusion protein E at low pH. Journal of virology **76** (8): 3784–3790.

Strasburger, E. (1894). The Periodic Reduction of the number of the Chromosomes in the Life-History of Living Organisms. Annals of Botany **8** (31): 281–316.

Swain, M., and Ross, N.W. (1995). A silver stain protocol for proteins yielding high resolution and transparent background in sodium dodecyl sulfate-polyacrylamide gels. Electrophoresis **16** (6): 948–951.

Takahashi, T., Mori, T., Ueda, K., Yamada, L., Nagahara, S., Higashiyama, T., Sawada, H., and Igawa, T. (2018). The male gamete membrane protein DMP9/DAU2 is required for double fertilization in flowering plants. Development (Cambridge, England) **145** (23).

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

Takeuchi, H., and Higashiyama, T. (2016). Tip-localized receptors control pollen tube growth and LURE sensing in Arabidopsis. Nature **531** (7593): 245–248.

Tamura, K., Dudley, J., Nei, M., and Kumar, S. (2007). MEGA4: Molecular Evolutionary Genetics Analysis (MEGA) software version 4.0. Molecular biology and evolution **24** (8): 1596–1599.

Termini, C.M., and Gillette, J.M. (2017). Tetraspanins Function as Regulators of Cellular Signaling. Frontiers in Cell and Developmental Biology **5**.

Thieme, C.J., Rojas-Triana, M., Stecyk, E., Schudoma, C., Zhang, W., Yang, L., Miñambres, M., Walther, D., Schulze, W.X., Paz-Ares, J., Scheible, W.-R., and Kragler, F. (2015). Endogenous Arabidopsis messenger RNAs transported to distant tissues. Nature Plants **1** (4): 15025.

Tian, H.Q., Zhang, Z., and Russell, S.D. (2001). Sperm dimorphism in Nicotiana tabacum L. Sexual plant reproduction **14** (1-2): 123–125.

Tsukamoto, T., Qin, Y., Huang, Y., Dunatunga, D., and Palanivelu, R. (2010). A role for LORELEI, a putative glycosylphosphatidylinositol-anchored protein, in Arabidopsis thaliana double fertilization and early seed development. The Plant journal for cell and molecular biology **62** (4): 571–588.

Tuladhar, R., Yeu, Y., Tyler Piazza, J., Tan, Z., Rene Clemenceau, J., Wu, X., Barrett, Q., Herbert, J., Mathews, D.H., Kim, J., Hyun Hwang, T., and Lum, L. (2019). CRISPR-Cas9-

based mutagenesis frequently provokes on-target mRNA misregulation. Nature communications **10** (1): 4056.

Tusnády, G.E., and Simon, I. (2001). The HMMTOP transmembrane topology prediction server. Bioinformatics (Oxford, England) **17** (9): 849–850.

Tyanova, S., Temu, T., Sinitcyn, P., Carlson, A., Hein, M.Y., Geiger, T., Mann, M., and Cox, J. (2016). The Perseus computational platform for comprehensive analysis of (prote)omics data. Nature methods **13** (9): 731–740.

Valansi, C., Moi, D., Leikina, E., Matveev, E., Graña, M., Chernomordik, L.V., Romero, H., Aguilar, P.S., and Podbilewicz, B. (2017). Arabidopsis HAP2/GCS1 is a gamete fusion protein homologous to somatic and viral fusogens. The Journal of cell biology **216** (3): 571–581.

van der Graaff, E., Schwacke, R., Schneider, A., Desimone, M., Flügge, U.-I., and Kunze, R. (2006). Transcription analysis of arabidopsis membrane transporters and hormone pathways during developmental and induced leaf senescence. Plant physiology **141** (2): 776–792.

van Deventer, S.J., Dunlock, V.-M.E., and van Spriel, A.B. (2017). Molecular interactions shaping the tetraspanin web. Biochemical Society transactions **45** (3): 741–750.

Vogler, F., Konrad, S.S.A., and Sprunck, S. (2015). Knockin' on pollen's door: live cell imaging of early polarization events in germinating Arabidopsis pollen. Frontiers in plant science **6**: 246.

Vogler, F., Schmalzl, C., Englhart, M., Bircheneder, M., and Sprunck, S. (2014). Brassinosteroids promote Arabidopsis pollen germination and growth. Plant reproduction **27** (3): 153–167.

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

Walker, J.C. (1993). Receptor-like protein kinase genes of Arabidopsis thaliana. The Plant journal for cell and molecular biology **3** (3): 451–456.

Wang, T., Liang, L., Xue, Y., Jia, P.-F., Chen, W., Zhang, M.-X., Wang, Y.-C., Li, H.-J., and Yang, W.-C. (2016). A receptor heteromer mediates the male perception of female attractants in plants. Nature **531** (7593): 241–244.

Wang, Z.-P., Xing, H.-L., Dong, L., Zhang, H.-Y., Han, C.-Y., Wang, X.-C., and Chen, Q.-J. (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**: 144.

Weis, C., Hückelhoven, R., and Eichmann, R. (2013). LIFEGUARD proteins support plant colonization by biotrophic powdery mildew fungi. Journal of experimental botany **64** (12): 3855–3867.

White, J., and Helenius, A. (1980). pH-dependent fusion between the Semliki Forest virus membrane and liposomes. Proceedings of the National Academy of Sciences of the United States of America **77** (6): 3273–3277.

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G.V., and Provart, N.J. (2007). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. PloS one **2** (8): e718.

Wong, J.L., and Johnson, M.A. (2010). Is HAP2-GCS1 an ancestral gamete fusogen? Trends in cell biology **20** (3): 134–141.

Wong, J.L., Leydon, A.R., and Johnson, M.A. (2010). HAP2(GCS1)-dependent gamete fusion requires a positively charged carboxy-terminal domain. PLoS genetics **6** (3): e1000882.

Wu, Q., Medina, S.G., Kushawah, G., DeVore, M.L., Castellano, L.A., Hand, J.M., Wright, M., and Bazzini, A.A. (2019). Translation affects mRNA stability in a codon-dependent manner in human cells. eLife **8**.

Wu, Y., Xun, Q., Guo, Y., Zhang, J., Cheng, K., Shi, T., He, K., Hou, S., Gou, X., and Li, J. (2016). Genome-Wide Expression Pattern Analyses of the Arabidopsis Leucine-Rich Repeat Receptor-Like Kinases. Molecular plant **9** (2): 289–300.

Wudick, M.M., Luu, D.-T., Tournaire-Roux, C., Sakamoto, W., and Maurel, C. (2014). Vegetative and sperm cell-specific aquaporins of Arabidopsis highlight the vacuolar equipment of pollen and contribute to plant reproduction. Plant physiology **164** (4): 1697–1706.

Wudick, M.M., Portes, M.T., Michard, E., Rosas-Santiago, P., Lizzio, M.A., Nunes, C.O., Campos, C., Santa Cruz Damineli, D., Carvalho, J.C., Lima, P.T., Pantoja, O., and Feijó, J.A. (2018). CORNICHON sorting and regulation of GLR channels underlie pollen tube Ca2+ homeostasis. Science (New York, N.Y.) **360** (6388): 533–536.

Xin, H.-P., Peng, X.-B., Ning, J., Yan, T.-T., Ma, L.-G., and Sun, M.-X. (2011). Expressed sequence-tag analysis of tobacco sperm cells reveals a unique transcriptional profile and selective persistence of paternal transcripts after fertilization. Sexual plant reproduction **24** (1): 37–46.

Yanagimachi, R., and Noda, Y.D. (1970). Ultrastructural changes in the hamster sperm head during fertilization. Journal of ultrastructure research **31** (5-6): 465–485.

Yang, W.-C., Shi, D.-Q., and Chen, Y.-H. (2010). Female gametophyte development in flowering plants. Annual review of plant biology **61**: 89–108.

Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., and Madden, T.L. (2012). Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC bioinformatics **13**: 134.

Yin, H.-S., Paterson, R.G., Wen, X., Lamb, R.A., and Jardetzky, T.S. (2005). Structure of the uncleaved ectodomain of the paramyxovirus (hPIV3) fusion protein. Proceedings of the National Academy of Sciences of the United States of America **102** (26): 9288–9293.

Yin, H.-S., Wen, X., Paterson, R.G., Lamb, R.A., and Jardetzky, T.S. (2006). Structure of the parainfluenza virus 5 F protein in its metastable, prefusion conformation. Nature **439** (7072): 38–44.

Zhang, J., Huang, Q., Zhong, S., Bleckmann, A., Huang, J., Guo, X., Lin, Q., Gu, H., Dong, J., Dresselhaus, T., and Qu, L.-J. (2017). Sperm cells are passive cargo of the pollen tube in plant fertilization. Nature Plants **3**: 17079.

Zhong, S., Liu, M., Wang, Z., Huang, Q., Hou, S., Xu, Y.-C., Ge, Z., Song, Z., Huang, J., Qiu, X., Shi, Y., Xiao, J., Liu, P., Guo, Y.-L., Dong, J., Dresselhaus, T., Gu, H., and Qu, L.-J. (2019a).

Cysteine-rich peptides promote interspecific genetic isolation in Arabidopsis. Science (New York, N.Y.) **364** (6443).

Zhong, Y., Chen, B., Li, M., Wang, D., Jiao, Y., Qi, X., Wang, M., Liu, Z., Chen, C., Wang, Y., Chen, M., Li, J., Xiao, Z., Cheng, D., Liu, W., Boutilier, K., Liu, C., and Chen, S. (2020). A DMP-triggered in vivo maternal haploid induction system in the dicotyledonous Arabidopsis. Nature Plants **6** (5): 466–472.

Zhong, Y., Liu, C., Qi, X., Jiao, Y., Wang, D., Wang, Y., Liu, Z., Chen, C., Chen, B., Tian, X., Li, J., Chen, M., Dong, X., Xu, X., Li, L., Li, W., Liu, W., Jin, W., Lai, J., and Chen, S. (2019b). Mutation of ZmDMP enhances haploid induction in maize. Nature Plants **5** (6): 575–580.

Zhou, X., and Meier, I. (2014). Efficient plant male fertility depends on vegetative nuclear movement mediated by two families of plant outer nuclear membrane proteins. Proceedings of the National Academy of Sciences of the United States of America **111** (32): 11900–11905.

Ziyyat, A., Rubinstein, E., Monier-Gavelle, F., Barraud, V., Kulski, O., Prenant, M., Boucheix, C., Bomsel, M., and Wolf, J.-P. (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. Journal of cell science **119** (Pt 3): 416–424.

Acknowledgements

I want to express my sincere gratitude to my supervisor PD Dr. Stefanie Sprunck. Thank you for introducing me to the fascinating world of (plant) cell biology and for the trust and confidence you have placed in me and my work, already since the beginning of my bachelor thesis (wow…that was almost 10 years ago ☺). I very much enjoyed working in your lab and I am grateful for the freedom I had to develop my own ideas while benefiting from your immense knowledge on all things related (but not limited) to plant sex. Thanks also for correcting my thesis.

I thank Prof. Dr. Thomas Dresselhaus and Prof. Dr. Bernhard Küster for mentoring. Dear Thomas, thank you for the opportunity to do my PhD at your department, I have always felt most welcome. Many thanks to Prof. Dr. Kay Schneitz and Prof. Dr. Stephan Schneuwly for kindly accepting to act as referees of my thesis.

I want to thank all the collaborators for contributing to this project. Especially Dr. Julia Mergner for measuring sperm protein samples and Dr. Maxim Messerer for bioinformatic analysis. Dr. Maria Flores-Tornero and Dr. Wolfgang Mages are acknowledged for providing *Amborella* and *Chlamydomonas* gDNA. Thanks to Prof. Dr. Reinhard Kunze for gifting the DMP1-GFP overexpression plasmid.

I want to thank all former and present lab and department members (AG Grasser included) for contributing to a familiar and productive working atmosphere. Especially Karl, Ivan and Raffi, but also the former members Tom, Mary, Hans, Tina, Maria, Marc and Frank. Thanks for all the fun we had in and outside of the lab. I also want to thank the numerous students that supported my work as practical, bachelor or master students, especially Raphael Malka and Katharina Hoppe. Monika Kammerer, Ingrid Fuchs and Verena Pirzer are acknowledged for their excellent technical assistance and I am very grateful to our gardeners Armin Hildebrand and Uschi Wittmann for taking care of hundreds of *Arabidopsis* and maize plants. Thanks also to our secretaries Sabine Kolb and Vroni Mrosek for competent help with administrative issues.

The DFG is acknowledged for funding my work through the SFB924. Thanks also to the international PhD program Regensburg (iPUR) for funding my conference visit in Japan.

Ich möchte mich von Herzen bei meinen Eltern Elisabeth und Franz bedanken. Vielen Dank, dass ihr immer an mich geglaubt und mir diesen Weg ermöglicht habt. Viola, vielen Dank für deine Geduld, deinen Rückhalt und dein sonniges Gemüt ☺.

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.

Regensburg, 12.6.2020

Philipp Cyprys