UNIVERSITÀ DEGLI STUDI DI MILANO

Scuola di Dottorato in Scienze Biologiche e Molecolari

XXV Ciclo

Molecular Analysis of Double Fertilization Process in Arabidopsis

Marta Adelina Miranda Mendes

PhD Thesis

Scientific tutor:

Lucia Colombo

Academic year: 2012-2013

Molecular Analysis of the Double Fertilization Process in Arabidopsis

SSD: BIO 01 BIO18

This thesis was performed at Dipartimento di Bioscienze,

Plant development group,

University of Milan

Look always to the bright side of life

Contents

Part I

Abstract

Fertilization and seed formation are key events in the life cycle of flowering plants. The seed represents an elaborated functional unit, whose main purpose is to propagate the plant's offspring. The first step in seed development is the formation of ovules and subsequently the achievement of a successful double fertilization process.

In our lab we have discovered that the MADS-box domain protein complex formed by SEEDSTICK (STK) and SEPALLATA3 (SEP3), responsible to maintain ovule identity, have as a direct target a member of the REM family, *VERDANDI (VDD, REM20).* With the combination of Bio-informatics studies and ChIP-qPCR experiments using specific STK and SEP3 antibodies, we were able to identify REM11, as the second direct target of the STK-SEP3 complex. The phenotype of the *rem11* mutant is very similar to the one described for *vdd-1/+* demonstrating that REM11 plays a similar function in the fertilization process. To better understand the fertilization defect observed in these mutants we have used a new technique developed to observe fertilization *in vivo,* by visualizing the gametes with a combination of markers for sperm cells and for the female gamete. In the *rem11* or in *vdd-1* gametophytes the synergid cell seemed to lose identity. Although the pollen tubes reached the micropyle, the two sperm cells didn't migrate toward the egg and center cells. These results showed the important and the direct involvement of these two genes in the control of synergid driven processes. Ultimately we discovered that genes responsible for the pollen tube attraction like the transcription factor *MYB98,* are correctly expressed in the mutants whereas genes, probably responsible for the degeneration process are miss-expressed. In summary, we can say that, two very different processes are regulated by the synergid cells: 1) the attraction of the pollen tube and 2) the synergid degeneration (apoptosis). We discover that the second step is specifically controlled by VDD and

REM11, two proteins that by yeast-2-hybrid experiments were able to interact. Based on these results we have decided to study if other REM transcription factors could interact with REM11 and VDD.

In conclusion STK-SEP3 MADS-box complex are able to directly regulate a REM transcription factor complex that has a very important and specific role during the double fertilization process. To understand how VDD-REM11 complex regulate synergid degeneration we have performed a RNA sequencing experiment comparing wild-type mature carpels to mutant ones. Exciting targets have been discovered and discuss in this thesis.

I have also studied the regulation of *VDD* transcription by STK-SEP3 complex. In *VDD* regulatory region three CArG boxes were identified and by Chromatin Imunoprecipitation experiments, we have showed that the ovule identity proteins STK and SEP3 bind to the first and third CArG boxes. We have performed in vivo and in vitro experiments showing that the STK-SEP3 complex is needed to form short-range loops in *VDD* promoter. For years evidences based on *in vitro* biochemical assays and yeast experiments shown that MADS box proteins form multimeric complexes. New evidences for the quartet-floral model were obtained, analyzing the activation of *VDD* promoter by STK-SEP3 multimeric complex.

Least but not the last, I have also analyzed the interaction of *ARABIDOPSIS BSISTER* (*ABS*) with *STK*, showing that they have a function in the regulation of the endothelium development, the inner most integument layer of the mature ovule that we demonstrated to be required to the double fertilization process.

Introduction

The importance of Double Fertilization process

Fertilization and seed formation are key events in the life cycle of flowering plants. The seed represents an elaborated functional unit whose main purpose is to propagate the plant's offspring. The first step in seed development is the formation of ovules and the subsequent steps will culminate in a successful double fertilization process. The detailed study of this process is very importance because influence directly different human needs such as protection of the biodiversity and the assurance of sustainable agricultural systems to feed the world population. Additionally is important to improve our knowledge on the molecular mechanism controlling plant sexual reproduction to ensure and increase seed production.

Sexual reproduction requires the delivery of the sperm nuclei, via the pollen, to the embryo sac, where fertilization occurs and the new diploid sporophyte is formed.
The plant life cycle in the angiosperms is characterized by the alternation of generations between a diploid sporophyte and a haploid gametophyte. The sporophyte produces spores, which then develop into gametophytes. The gametophytes in its turn produce either the male and female gametes. In contrast to lower plant species, in which the gametophyte is the dominant, free-living generation, gametophytes of angiosperms are smaller and less complex than the sporophyte and are formed within specialized organs of the flower.

Gametophytes Development

Ovule development and the female gametophyte

The ovule is the source of the female gametophyte (embryo sac) and the progenitor of the seed. Ovules provide structural support to the female gametes and enclose them until the development of the seeds, upon fertilization. The development of the ovule can be divided in two steps, the specification of the megasporocyte plus the production of a functional megaspore (megasporogenesis) and the formation of the embryo sac (megagametogenesis) (Reiser and Fischer 1993; Shi and Yang 2011).

In most flowering plants like for *Arabidopsis thaliana*, the model specie used in our studies, ovules initiate their development as illustrated in Figure 1. Initially the ovule is composed by three basic structures: the nucellus, a middle zone were the two integuments will arise and a funiculus (Figure 1A). The nucellus is derived from the apical portion of the ovule primordium and functions as the megasporangium. Shortly after ovule initiation, a single subdermal nucellar cell enlarges and displays a prominent nucleus. This cell represents the megasporocyte, and typically occupies a position directly below the apex of the nucellus. The megasporocyte gives origin to four megaspores upon meiosis where only one does not degenerated and gives rise to the megagametophyte. Only the megaspore at the chalaza pole, also called the functional megaspore, undergoes three rounds of nuclear division to form a coenocytic, eight-nucleated embryo sac. Subsequently, nuclear migration, polar nuclear fusion, and cellularization take place to yield ultimately a seven-celled embryo sac composed of two synergids, one egg cell, one diploid central cell, and three antipodals (Figure 1C-D) (Drews and Koltunow 2011a; Reiser and Fischer 1993).

The integuments are initiated at the base of the nucellus during megasporogenesis (Figure 1B). The inner integument is of dermal origin, whereas the outer integument is usually derived from both dermal and sub-dermal layers. The two integuments are considered to have distinct evolutionary origins. Periclinal divisions in the integuments generate an increase in the number of cell layers, whereas anticlinal divisions and cell elongation are responsible for wild-type growth parallel to the nucellus.

Megagametogenesis involves first two series of mitoses without cytokinesis, followed by cellularization of the nuclei and then cell differentiation. The two rounds of mitosis without cytokinesis lead to a four-nucleate coenocyte with two nuclei at each pole. During a third mitosis, phragmoplasts and cell plates form between sister and non-sister nuclei; this is the beginning of the cellularization process and the female gametophyte cells quickly become completely surrounded by cell walls. During and after cellularization, one nucleus from each pole (the polar nuclei) migrates toward the center of the developing female gametophyte and they fuse. These events result in a seven-celled structure consisting of three antipodal cells, one central cell, two synergid cells, and one egg cell (Figure 1D). The central cell has two haploid nuclei. If the female gametophyte is unfertilized, the antipodal cells eventually might undergo cell death; however, at the time of fertilization, the female gametophyte most likely is a seven-celled structure (i.e., the antipodal cells are present (Christensen et al. 1998; Drews and Koltunow 2011b; Schneitz et al. 1995).

The egg and central cells are polarized such that the nuclei of both cells lay very close to each other. This feature is important for double fertilization because these two nuclei are the targets of the two sperm nuclei. Furthermore, in the regions where the egg, synergid, and central cells meet, the cell walls are absent or discontinuous and the plasma membranes of these cells are in direct contact with

each other to facilitate direct access of the sperm cells to the fertilization targets because the pollen tube releases the two sperm cells into one of the synergid cells.

The synergid cell wall is further specialized. At the micropylar pole, the synergid cell wall is thickened and extensively invaginated, forming a structure referred to as the "filiform apparatus". The filiform apparatus greatly increases the surface area of the plasma membrane in this region and contains a high concentration of secretory organelles, suggesting that it may facilitate transport of substances into and out of the synergid cells. Based on cytological staining properties in species other than Arabidopsis, the filiform apparatus appears to be composed of a number of substances including cellulose, hemicellulose, pectin, callose, and protein. The filiform apparatus has at least two functions associated with the fertilization process. First, the synergid cells secrete pollen tube attractants through the filiform apparatus and the pollen tube enters the synergid cell by growing through the filiform apparatus, suggesting that it is important for pollen tube reception (Drews and Koltunow 2011b; Kasahara et al. 2005; Punwani et al. 2008).

The antipodal cells in *Arabidopsis* have no dramatic specializations and no known function. In other species, the antipodal cells contain finger-like cell wall projections resembling the filiform apparatus. These observations suggest that the antipodal cells indeed have a function and that they may function as transfer-cells, transporting substances from the surrounding ovule cells into the female gametophyte (Drews and Koltunow 2011b).

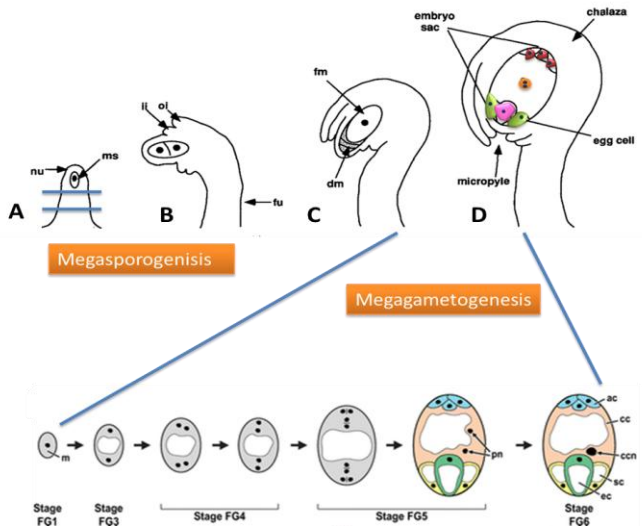


Figure 1. Polygonum-type embryo sac development, *Arabidopsis thaliana*

A-B) Megasporogenisis process, in (A) are visible three distinct parts, from the top to the bottom: nucellus (nu) showing a single megasporocyte (ms), one or two integuments, and a funiculus (fu). .
(B) Ovule after both integuments had initiated their growth. At this time, the megasporocyte has undergone the first meiotic division. ii, inner integument; oi, outer integument.
(C) Ovule after meiosis. The functional megaspore (fm) at the chalazal end has expanded, and the nonfunctional megaspores are degenerated. dm, degenerate megaspores.
(D) Ovule after megagametogenesis. The mature embryo sac contains seven cells and eight nuclei. From (C) to (D) detailed phases of the megametogenisis. Megaspore with a single nucleus (stage FG1). This nucleus undergoes two rounds of mitosis, producing a four-nucleate coenocyte, with two nuclei at each pole separated by a large central vacuole (stage FG4). During a third mitosis, phragmoplasts and cell plates form between sister and non-sister nuclei and the nuclei become completely surrounded by cell walls (Stage FG5). During cellularization, the polar nuclei migrate toward the center of the female gametophyte and fuse before fertilization. These events produce a seven-celled structure consisting of three antipodal cells, one central cell, two synergid cells, and one egg cell. (Stage FG6). White areas represent vacuoles and black circles/ovals represent nuclei.
Abbreviations: ac, antipodal cells; cc, central cell; ccn; central cell nucleus; ch, chalazal region of the ovule; ec, egg cell; f, funiculus; fg, female gametophyte; fm, functional megaspore; ii, inner integument; m, megaspore; mp, micropyle; oi, outer integument; pn, polar nuclei; sc, synergid cells.
Adapted from (Christensen et al. 1998; Drews and Koltunow 2011b; Reiser and Fischer 1993)

Integument development

As the embryo sac develops, the integuments continue to enlarge, typically overgrowing the nucellus. The amount of ovule curvature varies with the extent of differential growth of the integuments and funiculus; the degree of curvature forms a basis for classification of ovule morphology. Thus, the mature anatropous ovule shows extensive curvature such that the long axis of the nucellus is parallel to the axis of the funiculus (Figure 1D) (Schneitz K 1995).

The embryo sac is then in direct contact with the inner integument. In these situations, the innermost cell layer of the inner integument differentiates into a unique cell layer named the endothelium. Radial cell expansion, endopolyploidy, and prominent nuclei are observed in the endothelial cells (Reiser and Fischer 1993). The ovule is connected to the ovary wall by the funiculus, a stalk-like structure extending from the lowermost part of the chalaza to the placenta. Usually, a single vascular strand runs through the funiculus from the placenta terminating at the base of the embryo sac. The mature ovule displays a polarity with respect to the axis determined by the location of the chalaza and micropyle. The chalaza is defined as the region extending from the base of the integuments to the point of attachment of the funiculus. The micropyle is located at the point where the integuments terminate and is the site where pollen tubes enter the ovule (Figure 1D).

The *inner no outer* (*ino*) mutant description suggests that proper integument formation is also necessary to stimulate megagametogenesis progression *ino* ovules do not develop the outer integument; however, the inner integument seems to develop normally. *ino* embryo sacs are also gametophytically defective, since megagametogenesis cannot proceed after FG5 (Christensen et al. 1998) indicating that both integuments are important in Arabidopsis to promote female gametophyte

development. Other mutants have been described in great detail and reviewed by (Bencivenga et al. 2011), showing the integuments importance.

Male gametophyte

Male reproductive development begins with the formation of the stamen structure composed by the anther and by a filament that supports it. Two distinct sequential phases take place inside the anther, as described before for the embryo sac development, forming the male gametophyte: the microsporogenesis and the microgametogenesis (Figure 2). In microsporogenesis, diploid pollen mother cells undergo meiotic division to produce tetrads of haploid microspores, each tetrad is enclosed by a thick callose wall (Scott 2004). This stage is completed when distinct unicellular microspores are freed from tetrads by the activity of a mixture of enzymes secreted by the tapetum.

Each microspore undergoes one asymmetrical division known as Pollen Mitosis I (PMI) which gives rise to a small germ cell and a large vegetative cell. At this point, another generation starts, the male gamethophytic generation. The two cells of this bicellular pollen grain have strikingly different fates. The germ cell that represents the male germline is subsequently engulfed within the cytoplasm of the larger vegetative cell and creates a novel cell-within-a-cell structure. This swallowing up process involves degradation of the hemispherical callose wall that separates the newly formed vegetative and germ cells. The fully engulfed germ cell forms a spindle-like shape that is maintained by a cortical cage of bundled microtubules (Giampiero Cai 2006; Palevitz 1989). Asymmetric division at PMI is essential for the correct cellular patterning of the male gametophyte, since the resulting two daughter cells each one harbors a distinct cytoplasm and possesses unique gene expression profiles that confer their distinct structures and cell fates (Twell 1998). Induction of symmetrical division at PMI has demonstrated that

vegetative cell gene expression is the default developmental pathway and that the division asymmetry is critical for correct germ cell differentiation (Eady et al. 1995). The large vegetative cell has dispersed nuclear chromatin, nurtures the developing germ cell and will give rise to the pollen tube following a successful pollination. During pollen maturation, the vegetative cell accumulates carbohydrate and/or lipid reserves along with transcripts and proteins that are required for rapid pollen tube growth. The smaller germ cell has condensed nuclear chromatin and continues through a second mitosis, named the Pollen Mitosis II (PMII), to produce twin sperm cells. In species that shed tricellular pollen, such as *Arabidopsis thaliana*, PMII takes place within the pollen grain prior to anthesis. This is in contrast with the majority of species that shed their pollen in a bicellular state, such as Lilium longiflorum, with PMII taking place in the growing pollen tube. Following PMII, a physical association between the sperm cells and the vegetative nucleus is established that is referred to as the male germ unit. At the end of pollen grain development, a dehydration phase takes place where disaccharides, proline and glycine-betaine work as osmoprotectants to protect vital membranes and proteins from damage (Schwacke et al. 1999). Mature pollen grains must be released from anther locules, this is achieved through a process called anther dehiscence, which involves opening of the anther wall. This requires the degeneration of specific anther tissues called septum and stomium (Borg et al. 2009).

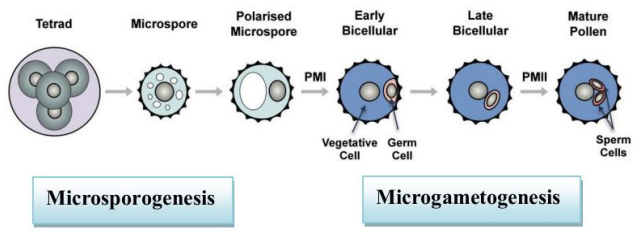


Figure 2. Schematic representation of Pollen development, adapted from (Borg et al. 2009).

Double fertilization process

The first time that was suggested that each male cell fuses with a cell in the female gametophyte, was in the year of 1898. In St. Petersburg, Sergius Nawaschin described this event in a report to the Imperial academy of Sciences. He carefully described the fate of the second sperm nucleus in *Lilium martagon* and *Fritillaria tenella*, and proposed that it fuses with the two polar nuclei to initiate endosperm development (Nawaschin S.G.)

However, some observations later occurred to be less pertinent, such as the idea that the male nuclei are released naked in the embryo sac. This *report* had an important impact. As a first consequence, the French botanist Léon Guignard decided to publish similar observations done for *Lilium martagon,* confirming "la double copulation sexuelle" i.e. double fertilization (Guignard Leon. 1899). Moreover Léon Guignard draw amazing pictures, one example is depicted in Figure 3, an embryo sac of *Lilium martagonthat*. These discoveries opened a new world were this phenomenon may not be an exception among flowering plants but may be rather common to all.

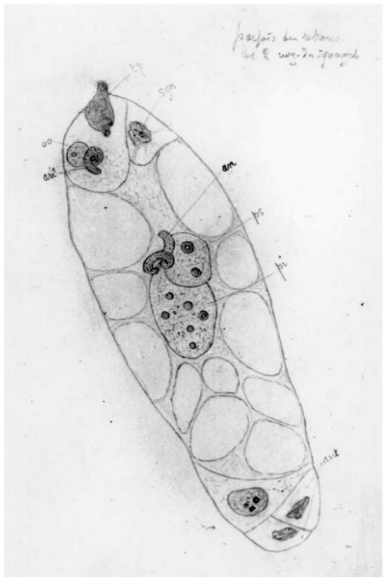


Figure 3. Original draw by Léon Guignard
Embryo sac of *Lilium Martagonat* when double fertilization is occurring. The embryo sac is made of the egg cell (oo), synergids (syn), the central cell with its two polar nuclei (ps and pi), and antipodal cells (ant). It has received a pollen tube (tp) and contains two male gamete nuclei (an) one being adjacent to the egg nucleus and the other close to the polar nuclei. Magnification:×400 approximately
Adapted from (Faure 2001).

From several years, the double fertilization process has been studied mainly due to the importance of this process for the maintenance of the next plant generations and in our lives.

The process has been carefully described and comprises a series of precise steps that end with the fusion of male and female gametes as proposed several years before.

Plant gametes are, in contrast with animal gametes, not direct products of meiosis, but differentiate within multicellular haploid generations, the male gametophyte (MG) and the female gametophyte (FG) respectively as they have been described in the previous paragraphs. Moreover to succeed in double fertilization, the two non-motile sperm cells must be transported through the pistil into the female gametophyte, which strictly depends on the directional growth of the pollen tube formed by the vegetative cell of the male gametophyte.

Pollen tube germination and penetration of the Female gametophyte

The initiation of pollen tube growth requires adhesion of the pollen grain to a receptive female stigma and hydration, providing first barriers for the species, restricted interaction between the MG and the female reproductive tract (Hiscock and Allen 2008). Once germinated, the tip-growing pollen tube penetrates the stigma tissue and navigates along the female reproductive tract towards the ovule, assisted by complex communication with the surrounding female sporophytic tissues, until finally arrives to the ovule (Figure 4A) (Rea and Nasrallah 2008).The synergid cells of the female gametophyte were shown to be most important cells in the first steps of the double fertilization process. Many experiments were done to show these meticulous synchronized set of events controlled by the synergids. First laser ablation experiments in *Torenia fournieri* showed that at least one viable synergid cell is necessary for pollen tube attraction (Higashiyama et al. 2001).

In *Arabidopsis thaliana*, the *myb98* mutant was shown to have a defective filiform apparatus and consequently fails to attract pollen tubes, indicating that synergid differentiation is also necessary for pollen attraction in other species (Kasahara et al. 2005).

Pollen tube attraction experiments in *Torenia* showed that a species-specific, signal is produced from the synergids to attract the pollen tube (Higashiyama et al. 2006).

Small cysteine-rich proteins (CRPs) are involved in many different phases of pollen-pistil interactions, from self-incompatibility, to pollen tube growth and guidance (Higashiyama 2010). In *Torenia*, CRPs named LUREs were shown to be secreted toward the micropylar ends of synergids and localized in the filiform apparatus in order to mediate short-range micropylar PT guidance (Okuda et al. 2009). In maize, CRPs were also identified to be highly expressed from the synergid cells (Cordts et al. 2001). Furthermore, a secreted protein, ZmEA1 was shown to be involved in micropylar pollen tube guidance in maize (Márton et al. 2005). The synergid cells secrete small and species-specific proteins such as URE1/2 and ZmEA1 (Zea mays EGG APPARATUS 1), which are involved in the last phase of pollen tube attraction, guiding tube growth through the micropyle into the FG (Márton et al. 2005).

The synergid cell death is required for fertilization.

Entering the micropyle, the pollen tube goes through a micropylar domain known as the filiform apparatus, this zone possesses cells with a thick cell wall and finger-like projections into the synergid cytoplasm (Higashiyama 2002). The pollen tube enters one synergid cell and burst (Figure 4B/C). Pollen tube growth arrest and burst indicate that species-specific cell recognition and signaling mechanisms also exist between the receptive synergid and a compatible pollen

tube, which may trigger synergid cell death and subsequent sperm cell delivery inside this synergid cell. The released pollen tube content mingles with the cytoplasm of the degenerating synergid and spreads, as the synergid membrane disintegrates, into a narrow space between the egg and central cell plasma membranes (Sandaklie-Nikolova et al. 2007). Typically, only one pollen tube penetrates each ovule. Rapid termination of pollen tube attractant(s) synthesis and/or secretion or degradation after successful fertilization may explain why the FG loses its ability to attract further pollen tubes (Hamamura et al. 2011).

Recently it has been shown that the two sperm cells should fuse with the female gametes (egg and central cells) to determine that the synergids stop producing the attractants (Beale et al. 2012; Kasahara et al. 2012).

The synergid cell that receives the PT undergoes cell death, however it is not yet understood if the cell death takes place before or upon the penetration of the pollen tube (Sandaklie-Nikolova et al. 2007). *Arabidopsi*s mutants like *gfa2* with defective synergid cell death remain unfertilized (Christensen et al. 2002). GFA2 is a J-domain-containing protein required for mitochondrial function (Christensen et al. 2002), suggesting that synergid cell death in Arabidopsis requires functional mitochondria, as is the case for cell death in animals (Morais Cardoso et al. 2002).

It is clear that synergid cell death is a highly controlled and coordinated process as the synergid plays an active role in the developmental events immediately before gamete fusion. Synergid cell death could be important for allowing the PT to enter the synergid, for example, by reducing its turgor pressure to allow the explosive discharge of the sperm cells (Rotman et al. 2003), and/or for setting up an environment which allows PT reception and the delivery of the sperm cells to the egg and central cell (Fu et al. 2000). Whether synergid cell death is a cause or a consequence of PT rupture remains controversial in *Arabidopsis*. Sandaklie-Nikolova et al. showed that

synergid degeneration scored by loss of GFP marker expression in the cytoplasm and change in synergid shape occurred after PT arrival at the female gametophyte but more than 100 min before PT discharge. In contrast, a recent report (Hamamura et al. 2011) discusses the loss of a nuclear-expressed synergid GFP marker occurred simultaneously with PT discharge, suggesting that break-down of the receptive synergid is dependent on PT discharge.

Arabidopsis gametophytic mutants defective in pollen tube growth arrest have been identified, including *feronia*, *fer* (Huck et al. 2003), *sirène ,srn (Rotman et al. 2003)*, *lorelei*, *lre* (Capron et al. 2008) (Tsukamoto and Palanivelu 2010), s*cylla*, *syl* (Rotman et al. 2008), *nortia, nta (Kessler et al. 2010)*, and *abstinence by mutual consent amc*, (Boisson-Dernier et al. 2008). *fer* and *srn* are allelic (Escobar-Restrepo et al. 2007). The *fer/srn, lre, nta*, and *syl* are female gametophyte-specific mutations; in these mutants, wild-type pollen tubes enter mutant female gametophytes but fail to cease growth and rupture. This results in a pollen tube overgrowth phenotype. The *amc* mutation, by contrast, affects both gametophytes. *amc* mutants also exhibit a pollen tube overgrowth phenotype but do so only when both gametophytes carried the mutation. This suggests that AMC may be necessary for cell death in both the PT and the synergid, perhaps by modulating the production of reactive oxygen species (ROS) (Boisson-Dernier et al. 2008). Additional ultrastructure analysis of *fer* embryo sacs indicates that the penetrated synergid displays electron-dense material (Huck et al. 2003), which is typical of the degenerating cell (Mansfield et al. 1991). However, this may be a very late marker for synergid cell death and all current methods for detecting synergid degeneration rely on indirect observations (i.e. loss of GFP signal and fixation/staining protocols that detect late stages of cell death). A definitive analysis of the role of synergid degeneration in PT reception awaits the development of more sophisticated tools for monitoring early stages of cell death in this specialized cell type.

As discussed above, FER/SRN, LRE, and AMC may be part of a pathway leading to ROS production in the receptive synergid cell in response to pollen tube contact, suggesting that synergid cell death may result from ROS in the synergid cell, as occurs in other cells in plants (Van Breusegem and Dat 2006). Together, these observations support a model in which pollen tube-synergid contact induces a physiological cell death program within the synergid cell.

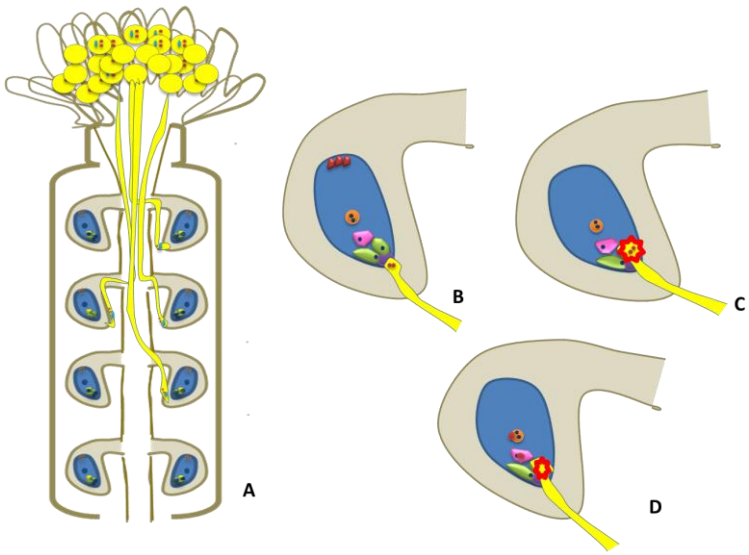


Figure 4. Schematic Representation of the double fertilization process.

(A) The pollen grains (male gametophyte) land in the stigma of a carpel hydrate and navigate until find an ovule and the female gametophyte. (B)The pollen tube arrives to the ovule enter the mycropyle enters into the synergid cell and burst (C). After bursting the two sperm cells migrate and one fertilizes the egg cell, whereas the other fertilizes the central cell.

The MADS-box protein control organ differentiation and organ function.

Modern biology textbooks contain a simple model that explains how a few genes act together to specify the four organs types to make perfect flower. Known as the ABC model (Fig. 5), it was conceived in the early 1990s, based on a series of celebrated homeotic mutants in two model species, *Arabidopsis* and *Antirrhinum* (Coen and Meyerowitz 1991). Perfect flowers contain four types of floral organ arranged in four concentric rings, known as whorls. The four organ types are sepals (outermost or whorl 1), petals (whorl 2), stamens (whorl 3) and carpels (innermost or whorl 4) (Causier et al. 2010).

The ABC model proposed that three functions, A, B and C, specify the organs that form the four whorls of the flower. The A, B and C functions are supposed to act in two adjacent whorls, which overlap with each other. Each whorl is defined by the expression of a unique function or combination of functions (Figure 5) (Causier et al. 2010; Coen and Meyerowitz 1991). The isolation of novel floral mutants in *Arabidopsis*, and other species, has led to an expansion of the ABC model to include the D and E functions. The D function, which specifies ovule identity in combination with the C function (Colombo et al. 1995; Favaro et al. 2003) and the E function, represents an important modification of the ABC model (Figure 5). Factors that widely affect the activity of the organ identity genes were first identified in tomato (TM5) and petunia (FBP2). Silencing of these related MADS-box genes resulted in a phenotype that suggested a decreased influence of B and C functions on floral development (FERRARIO et al. 2003; Pnueli et al. 1994). Later, three genes belonging to the TM5/FBP2 group were identified in Arabidopsis and named SEPALLATA 1(SEP1), SEP2 and SEP3. The *sep1 sep2 sep3* triple mutant

has a similar phenotype to the *tm5* and *fbp2* mutant lines, with all floral organs being replaced by sepals (Pelaz et al. 2000).

The first ABC genes to be cloned were the Antirrhinum B function gene *DEF* (Sommer et al. 1990) and the Arabidopsis C function gene *AG* (Yanofsky et al. 1990), the products of which shared a high degree of homology with the DNA-binding domains of two known transcription factors identified in yeast (MCM1) and animals (SRF). These four proteins became then the founding members of a very important family of transcription factors known as the MADS-box proteins, **M**CM1, **A**G, **D**EF, **S**RF (Schwarz-Sommer et al. 1990).

MADS-box factors have subsequently been shown to be key regulators of plant developmental processes, and in *Arabidopsis* at least 107 MADS-box genes have been identified (Parenicova et al. 2003). The plant MADS-box family can be divided into two large families: the type I class, which group with the human SRF protein, and the type II class that groups with yeast MEF2 (Alvarez-Buylla et al. 2000; Parenicova et al. 2003). The ABC MADS-box genes belong to the type II class and are characterized by four distinct domains. From the amino-terminal end are: the MADS-domain, the Intervening domain (I), the K-domain, and the C-domain. Together, the MADS-box and I-domain form the minimal DNA-binding domain. Plant MADS-box factors bind DNA as homo- or heterodimers, or in higher order complexes. The I- and K-domains mediate the interactions between MADS-box proteins (Causier et al. 2010).

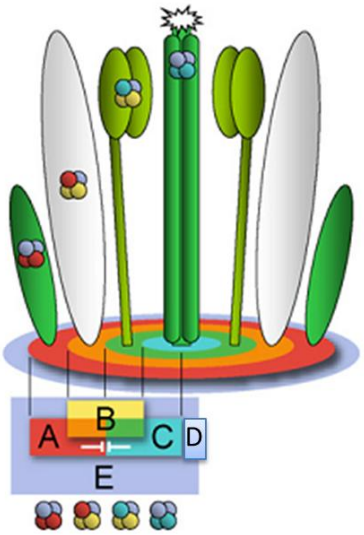


Figure 5. *ABCDE* model

Adapted from (Causier et al. 2010)

The ability of MADS-domain proteins to bind DNA as dimers is reflected by the dyad symmetry of their binding sites that are found within promoter and enhancer sequences (Shore and Sharrocks 1995). Nurrish and Treisman (1995) studied MADS-domain protein binding sites and showed that they bind to the consensus sequence "CC $(A/T)_6$ GG" named CArG box. Evidence based on *in vitro* biochemical assays and interaction studies in yeast showed that plant MADS-domain proteins form mainly heterodimers which are thought to assemble into multimeric complexes (de Folter et al. 2005; Egea-Cortines et al. 1999; Honma and Goto 2001; Riechmann et al. 1996) Figure 6.

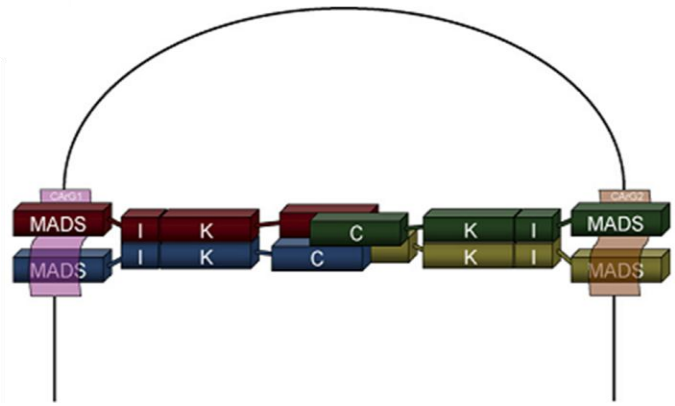


Figure 6. Proposed mode of action for MADS box complexes.

Adapted from(Causier et al. 2010)

MADS-box C-D-E Type factors

In *Arabidopsis*, the C-D type MADS-box transcription factor encoding genes *SEEDSTICK (STK), SHATTERPROOF1(SHP1)* and *SHP2* act redundantly to determine ovule identity. This was demonstrated by the phenotype of the *stkshp1shp2* triple mutant, in which ovules develop as carpelloid structures. A genetic titration experiment demonstrated that SEP genes (E factors) are also necessary for ovule formation, because the ovule phenotype of the *sep1 SEP1sep2sep3* mutant is very similar to that of the *stkshp1shp2* triple mutant (Favaro et al. 2003). This finding, together with data from protein interaction studies, suggested that STK, SHP1 and SHP2, along with the SEP factors, assemble in protein complexes in a similar manner to the floral organ identity MADS-box complexes (Honma and Goto 2001). In 2010 the first direct target of STK-SEP3 complex was published and was a gene belonging to the poorly characterized REM family, *VERDANDI* (*REM20, VDD)* (Matias-Hernandez et al. 2010). The *vdd-1/+*

mutant shows defects during the fertilization process resulting in semi-sterility. Analysis of the *vdd-1/+* mutant female gametophytes indicates that antipodal and synergid cell identity and/or differentiation are affected. These results provided new insights into the regulatory pathways controlled by the ovule identity factors and the role of the downstream target gene *VDD* in female gametophyte development (Matias-Hernandez et al. 2010).

The B3 superfamily, The REM Family

All members of the B3 superfamily factors contain 110 amino acid region, named the B3 domain. This domain was initially named B3 because it was the third basic domain in the maize transcription factor VIVIPAROUS1 (VP1) (McCarty et al. 1991). The first and second basic domains (B1 and B2) are specific to the VP1-like proteins, however the B3-domain has been found in several transcription factors. The B3 domain of VP1 encodes a sequence-specific DNA binding activity (Suzuki et al. 1997). Since its initial discovery in *VP1*, the B3 domain has been found in 118 genes in Arabidopsis and 91 in rice. B3 genes are also present in green algae, mosses (Marella et al. 2006), liverworts, ferns and gymnosperms. The B3 superfamily include several gene families, such as the **LAV** (LEAFY COTYLEDON2 [LEC2]–ABSCISIC ACID INSENSITIVE3 [ABI3]–VAL), **ARF** (AUXINRESPONSEFACTOR), **RAV** (RELATED TO ABI3 andVP1) and **REM** (REPRODUCTIVEMERISTEM) families (Swaminathan et al. 2008). At the moment the ARF family and the LAV family, are very well studied families, on the other side almost nothing is known from RAV and REM family's members.

Interestingly, it has been shown that B3 domains from distinct families bind to different DNA sites. Yet, these proteins share a common structural framework for DNA-recognition. Analysis, by NMR spectroscopy, of the structure of the B3

domain of the At1g16640 protein from Arabidopsis (Waltner et al. 2005), a member of the REM family, revealed that it has the same novel fold as RAV1 with seven-stranded b-sheet arranged in an open barrel and two short a-helices. Nevertheless, this particular gene (At1g16640) has a remarkably distinct amino acid sequence from others in the superfamily. This has raised doubts to whether this domain has the ability to bind to DNA. However, it has been showed that *VRN1* (VERNALIZATION1), a member of the REM family, binds DNA in vitro in a non-sequence-specific manner (Levy et al. 2002).

Bio-informatics studies (Romanel et al. 2009) of protein modeling showed that the B3 domains contain virtually identical tertiary structures. The structural model for the B3 domain of VRN1/REM5 suggests that the domain's characteristic fold is maintained, despite the putative DNA-binding loops being greatly reduced. Taken together, these findings suggest that the B3 domain's characteristic fold may have the basic pre-requisites to associate with DNA, while the loops might confer sequence specificity. In general, the comparison of exons/introns structure and sequences of the REM family members have showed that the basic structure of the genes are very similar in species such as *Arabidopsis*, rice and moss suggesting a high conservation rate during plant evolution.

Among the REM family member in Arabidopsis, only two of them have been characterized. *VRN1/REM5* (class VI) is a gene involved in vernalization mediated epigenetic silencing of FLC (Levy et al. 2002; Sung and Amasino 2004) (Mylne et al. 2006). The second REM gene with a known function is *VDD*/*REM20* (class VII) and the detailed phenotype was described above.

The phylogeny of the REM proteins also reveals a very active and dynamic process of gene duplication. This process resulted in the portrait of the REMs in plants, a large number of genes with a remarkable variability among them. Genome or tandem duplication may explain the emergence of the large number of *REM* genes, however what causes their maintenance as active genes in the genome is still an

open question. It has been suggested that, sub-functionalization and/or neo-functionalization play a role in the maintenance of most of the duplicated regulatory genes in Arabidopsis (Duarte et al. 2006). On the other hand, (Wellmer et al. 2006) suggested that the functional redundancy during early flower development may have increased the genetic buffering so that duplicated genes are retained by positive selection. They identified, by global analysis of gene expression, a significant enrichment of transcription factor families with closely related members expressed during *Arabidopsis* flower development (Wellmer et al. 2006). The maintenance of the REM family gene members may be a combination of the sub-functionalization and/or neo-functionalization as well as the genetic buffering processes. The elucidation of this complex gene family will passed for the study of the functional importance of REM genes during flower development.

Aim of the project

My PhD had three defined aims, all of them based on the analysis of the MADS-box STK-SEP3 complex role in plant sexual reproduction.

As a first aim I wanted to discover if the STK-SEP3 complex had other direct targets involved in the fertilization process. Secondly, I wanted to study in more detail *VDD* the first direct target of the complex. The detailed analyses of *vdd-1/+* help us to better understand the role of this REM member in the synergid differentiation and in the double fertilization process.

Finally, characterize the molecular regulation of *VDD* transcription by the STK-SEP3 complex using an *in vitro* and *in vivo* integrated approach.

Results (to be submitted)

REM11, the second direct target of STK-SEP3 is controlling synergid function interacting with VDD (REM20)

Abstract

REM11 is the second direct target of STK-SEP3 complex identified. This gene belongs to the REM family as *VDD*. Little is known about the REM family, until now only two members, of the 45 members, have been functionally characterized.

Hereby we show that *REM11* has an important role during the double fertilization process, as shown before for the *VDD*. The same phenotype as shown in *vdd-1/+* mutant plants was discovered now in the *REM11_RNA*i plants. *REM11_RNA*i mutant embryo sacs developed with 7 cells and eight nuclei as in wild type however, the synergids seem to lose their identity when crossed with specific cell markers, as shown previously for *vdd-1/+*.

The synergids play the major role during the double fertilization process they are responsible for the attraction and reception of the pollen tube. When the pollen tube penetrate the embryo sac, one of the synergid cells initiates to degenerate (apoptosis) the pollen tube arrests its growth, bursts, and releases the two sperm cells to ensure double fertilization.

VDD and REM11 have been shown to have a very important role in the correct function of the synergids and by yeast-two-hybrid experiments were shown to interact. Detailed examination showed that the synergids in *REM11_RNA*i and *vdd-1/+* mutants are still able to produce the attractants responsible for pollen tube attraction however they do not initiate the degeneration process and so the delivery of the sperm cells is compromised. A gene responsible for the pollen tube attraction, like the transcription factor MYB98 is correctly expressed in the mutants. Two very different processes are strictly coordinate by the synergids, attraction and degeneration being controlled by the new identified complex VDD-REM11.

Introduction

During plant reproduction, a series of interactions between male and female gametophytes ensures successful fusion of male and female gametes. In flowering plants, anthers give rise to the pollen grain (male gametophyte) containing the sperm cells. The female gametophyte develops inside an ovule and in many species, including *Arabidopsis*, is a seven-celled structure that includes three antipodal cells, two synergid cells, an egg and a central cell (Weterings and Russell 2004). The initiation of pollen tube growth requires adhesion of the pollen grain to a receptive female stigma and hydration, providing first barriers for the species, restricted interaction between the MG and the female reproductive tract (Rea and Nasrallah 2008). Once germinated, the tip-growing pollen tube penetrate the stigma and across the female reproductive tissues towards the ovule, assisted by complex communication between the tube cell and surrounding sporophytic tissues, until it arrives to the ovule (Rea and Nasrallah 2008). The synergid cells were shown to be important cells necessary for pollen tube attraction (Higashiyama 2002) and pollen tube burst (Sandaklie-Nikolova et al. 2007).

The two MADS-box proteins SEEDSTICK (STK) and SEPALLATA3 (SEP3) form a protein complex that directly regulates *VERDANDI* (*VDD*). *VDD* belongs to the REM family and has been shown to have a very important role in ovule development. In particular VDD is required for synergid functions and therefore for the accomplishment of double fertilization process (Matias-Hernandez et al. 2010). Using integrated approaches based on molecular biology and bio-informatics methods, *REM11,* a second STK-SEP3 target has been identified. The REM family belongs to a bigger plant family called B3 superfamily (Swaminathan et al. 2008). The B3 domain, was shown to be involved in DNA binding and additionally other domains can coexist in the multi-domain B3 proteins that are thought to mediate protein–protein interaction and/or dimerization (Romanel et al. 2009; Suzuki et al. 1997). The B3 genes have been identified also in gymnosperms,

ferns, mosses, liverworts and green algae. A recent bio-informatical analysis done by Romanel et al., described the REM family as being formed with 45 loci identified by TAIR annotation. Until now just two REM genes have been described to have a functional role in the model plant Arabidopsis thaliana, *VDD/REM20* and *VRN1/REM5*. The first one to be described was VRN1 involved in vernalization process mediated by epigenetic silencing of FLC and consequently in the flowering time process (Levy et al. 2002; Mylne et al. 2006).

To study the functional role of REM11 during plant development we created an interference line *REM11_RNAi*. Transgenic plants in which *REM11* has been down-regulated showed a high percent of ovule abortions. This high percentage of ovule abortions corresponded directly to the down-regulation of the gene. Although these transgenic plants have ovules that reach maturity, the synergids do not function in a proper way. In *REM11_RNAi* plants the synergids are able to attract the pollen tube, however upon the entrance of the pollen tube do not degenerate resulting in the fail of double fertilization. Using specific marker lines we could show that the pollen tubes once enter the micropyle do not burst. Interestingly REM11 interact with VDD to form a complex. .

Results

STK genome wide co-expression analysis, Pearson correlation

To address the existence of co-regulatory pathways and possible targets for STK, we carried out together with a bio-informatic group in our university a Pearson Correlation Coefficient (PCC) analysis of Arabidopsis Affymetrix transcriptome data from 2,000 experiments (for more details see material and methods section). All Affymetrix array experiments analysed were carried out by the same group (NASC) that have a standardized normalization process thus facilitating the creation of robust and reliable PCC matrices (Berri et al 2009). The generated matrices for *STK* using both untransformed P-lin and logarithm-transformed P-Log (see supplemental table1) expression data were analysed at different threshold values of PCC.

The predicted network showed the presence of genes that could be involved in the same signal transduction pathway as STK. A role for many of these genes has yet to be defined, but according to the Pearson co-regulatory analysis they are interesting candidates to be tested at the genetic level. As a result of this analysis REM11 scored a high value in (P-Log) and (p-Lin), and we decided to verify if was a direct target of STK.

REM11 a direct target of STK and SEP3 complex

Sequence analysis of the REM11 genomic region revealed the presence of two putative CArG boxes, the 1st CArG in the promoter region and the 2nd just after the translation starting site (Figure 1A). Chromatin immunoprecipition experiments

using STK and SEP3 native antibody followed by Quantitative real-time PCR showed an enrichment of both CArG-box regions (Figure 1B). Chromatin immune-precipitated from the *stk* single mutant was used as negative control for anti-STK experiment and wild-type leaves for anti-SEP3, because SEP3 is not expressed in leaves.

These data strongly indicates that STK and SEP3 proteins directly interact with the *REM11* genomic region. *REM11* is the second direct target of the MADS-box domain protein complex STK-SEP3 after the discovery of *VDD* (Matias-Hernandez et al. 2010).

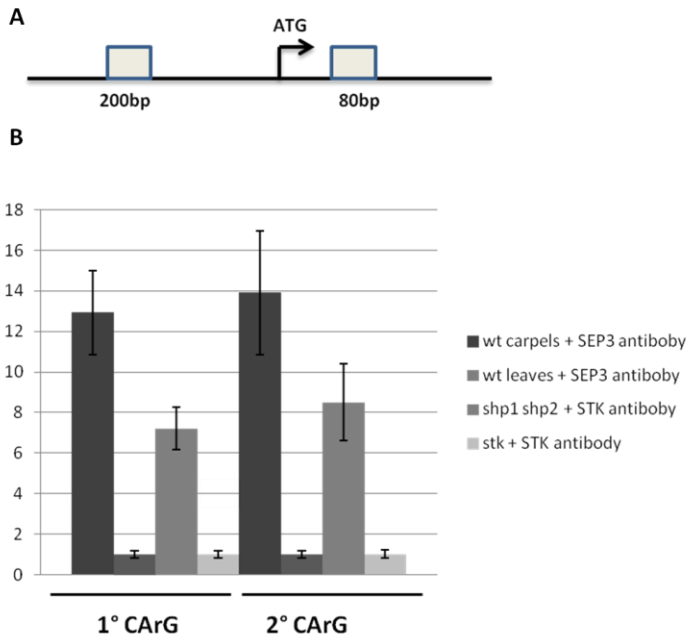


Figure 1. Quantitative Real-Time PCR on chromatin immune-precipitated with STK and SEP3 antibodies.

(A) Schematic representation of the position of the CArG boxes in the promoter region of the *REM11* gene.

(B) The ChIP enrichment were tested by quantitative real-time PCR showing that STK and SEP3 specific bind to the CArG boxes 1 and 2. The *stk* single mutant was used as a negative control in the STK-ChIP and wild-type leaves as negative control for the SEP3-ChIP assays.

REM11 expression pattern

To study the *REM11* expression profile at cellular level, *in situ* hybridization experiments were performed (Figures 2A to E). *In situ* hybridization experiments detected the *REM11* mRNA during all stages of ovule development. A stronger signal was always detected inside the carpels were the ovules were developing. Additionally in the last phases of ovule development a stronger signal was detected in mature ovule, when the embryo sac is formed (Figure 2E).

Since *REM11* is a direct target of STK, we were also interested in investigating whether the expression of *REM11* was different in *stk* and *stk shp1 shp2* mutants by *in situ hybridization*. Indeed a weaker signal was detected in the *stk* mutant carpels (Figure 2F), contrasting with no signal at all showed in the *stk shp1 shp2* triple mutant (Figure 2G). These facts demonstrate that *REM11* is directly regulated by *stk*, but in *stk* single mutant, SHP1 and SHP2 redundantly control *REM11* expression as was described also for *VDD* (Matias-Hernandez et al., 2010).

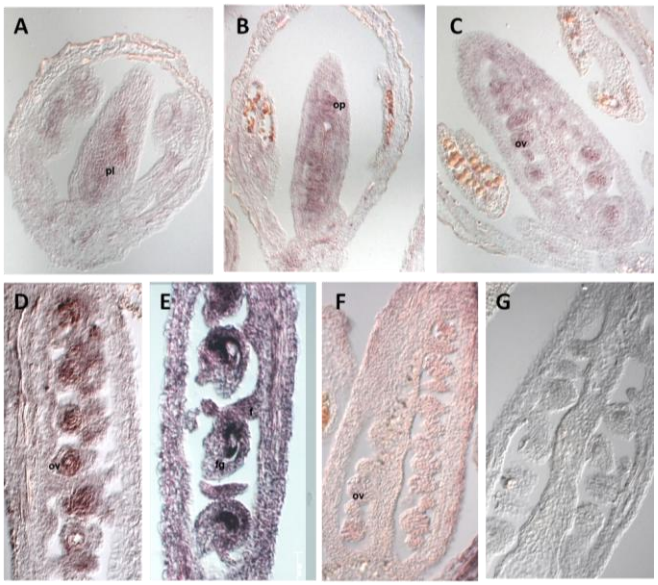


Figure 2. *REM11* in situ analysis

(A-E) *In situ* hybridization experiments performed using wild-type plants. (A-C) *REM11* gene is expressed in the early stages of ovule development. (D) *REM11* mRNA is detectable as a strong signal in the female gametophyte during later stages of ovule formation. (E) In the mature embryo sac *REM11* expression is highly detectable inside the embryo. (F) *In situ* hybridization experiment performed in *stk* single mutant. The signal is reduced.

(G) Experiment performed in *stkshp1shp2* triple mutant background, almost any signal was detected. pl-placenta; op-ovule primordial; ov-ovule; f-funiculus; fg-female gametophyte.

P35S::REM11_RNAi phenotype

To better understand the function of *REM11*, we had to construct RNA interference, *p35S::REM11_RNAi* (during the text always referred as *REM11_RNAi),* plants because no T-DNA insertion line was available at the

moment. The down-regulation of the *REM11* transcripts as shown by Quantitative Real Time-PCR (qRT-PCR) (figure 3D) demonstrated that our approach was successfully performed. Examining the immature fruits (siliques) of the transgenic lines we were able to see the presence of several aborted ovules (Figure 3A-C). The differences of number of ovule abortion was proportional to the level of down regulation of

REM11 transcripts, checked by qRT-PCR (Figure 3D). Few *REM11-1* plants have till 60% of ovule abortions (Figure 3F). The qRT-PCR showed almost no *REM11* expression for those plants.

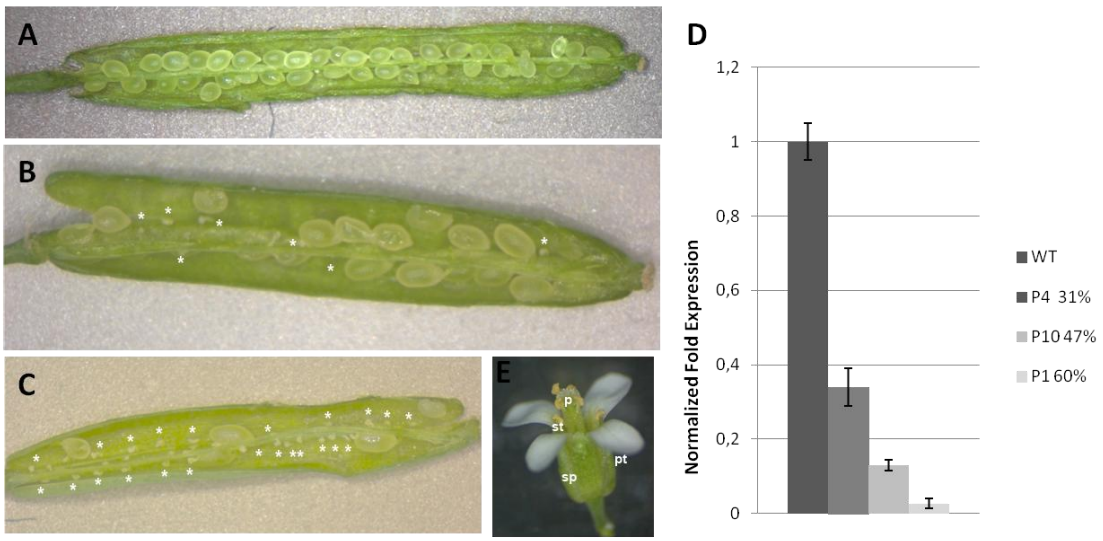


Figure 3. *REM11_RNAi* mutant analysis.

(A)Wild-type silique showing full seed set. (B-D) Siliques of *REM11_RNAi* plants containing different aborted ovules. Ovule abortion rate correlates with the down-regulation of *REM11* (D). (E) *REM11_RNAi* flower. Although the pistil presents 60% ovule abortions, all the organs were formed perfectly. *-aborted ovules; sp-sepals; pt-petals; st-stamen; p-pistil.

REM11_RNAi plants have female gametophyte defect

To test whether female, male, or both gametophytes were affected due to the down regulation of *REM11* activity, we performed reciprocal crosses between *REM11_RNAi* and wild-type plants. Crossing a wild-type female carpel with pollen from *REM11_RNAi* mutant plants show no significant increasing in the ovule abortions, comparing with the control wild-type crossed with wild-type. Interestingly when we used *REM11_RNAi* as female and wild-type pollen a high percentage of ovule abortions was observed. These results, point out to the fact that the ovule abortion phenotype is due only to a female gametophyte defect and not to a defect in the male gametophyte.

Table 1. Backcrossing between wild-type and *REM11_RNAi* plants

Back cross (female♀ x male ♂)	% Ovule abortion
wt x wt	4.3%
wt x *35S::REM11_RNAi*	6%
35S::REM11_RNAi x wt	56%

Morphological analysis and Gametophyte Cell Identity studies

In order to understand if *REM11_RNAi* mutant ovules had development defect we have performed DIC microscopy analysis of un-pollinated mature carpels. We found that the ovules in *REM11_RNAi* plants reached maturity and the embryo sac was formed with 7 cells, like to the ones in the wild type (Figure 4A). As the embryo sac seemed to be formed correctly with all the cells, to understand if the ovule abortions were due to defects in determination of female gametophyte cell identity, embryo sac cell-specific reporter constructs were introduced into the *REM11_RNAi* transgenic lines. The gene expression was analyzed in the F2 generation, were the marker was in a homozygous situation. We have used *EC1::GUS* as egg cell identity marker, (Sprunk et al.), *FIS2::GUS* as central cell identity marker (Chaudhury et al., 1997) and the promoter of the gene At1g36340 as marker for the antypodals cells identity (Figures 4B-D). No difference was found in the GUS expression in the *REM11_RNAi* transgenic lines if compared with wild type, indicating that their cell fate were not affected.

Analysis of the GUS expression using the synergid specific cell marker line (ET2634, Gross-Hardt et al., 2007) in *REM11* transgenic plants, revealed that some of the embryo sacs did not express the synergid specific marker (n = 740). Attractively the percentage of aborted ovules found per silique corresponds to the number of embryo sacs that didn't show any GUS in the synergids (Figure 4F). All the percentages and analysis are described below in the Table 2.

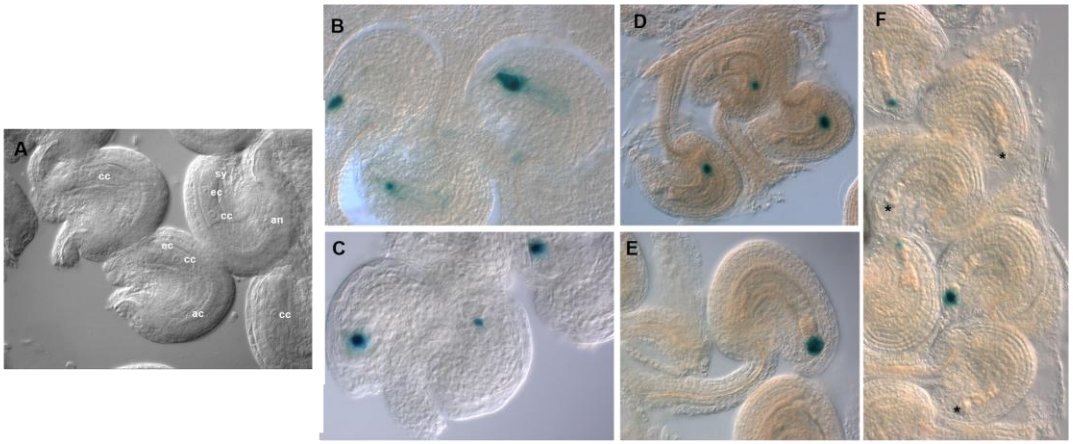


Figure 4. Embryo sac analysis

(A) DIC microscopy analysis of *REM11_RNAi* ovules. The ovules are morphologically indistinguishable from the ones of the wild-type, the embryo sac is composed by seven cells. sy- synergids; ec-egg cell; cc- central cell.

(B-E) Embryo sac cell markers expression in wild-type. (B) egg cell; (C)central central; (D) antypodals and (E) synergid; (F) Synergid marker expression in mutant background *REM11_RNAi.*

Table 2 Analysis of the Crosses between the embryo sac markers and *REM11_RNAi*

	Genotype	Ovules with signal	Ovules with no signal	Ovule abortions
Antipodal cell (F2)	wild-type	97%	3%	3%
	REM11_RNAi	92%	8%	30%
Egg cell (F2)	Wild-type	95%	5%	4%
	REM11_RNAi	95%	5%	27%
Central cell (F2)	wild-type	98%	2%	2%
	REM11_RNAi	95%	5%	37%
Synergids (F2)	wild-type	95% (n= 267)	5% (n=16)	2%
	REM11_RNAi	62,5% (n= 466)	37,5% (n=274)	37%

Double Fertilization process in *REM11_RNAi* mutant

Synergids are responsible for producing the attractants molecules, clues, for the pollen guidance, ensuring that each ovule micropyle gets a pollen tube (Higashiyama 2002). We started by analyzing if the pollen tubes were able to reach the mycropyle of each *REM11_RNAi* mutant ovule. We used aniline staining to follow the pollen tubes growth in wild-type and mutant pistils. The aniline blue is a very well-known chemical because it stains the callose present in the pollen tube cell wall (figure 5A). This experiment showed that pollen tubes targeted all the ovules entering into the embryo sac in *REM11*-RNAi plants indicating that the first part of the fertilization process wasn't affected.

To further analyze whether the pollen tube stopped or continued to growth in *REM11_RNAi* embryo sac, mutant carpels were hand pollinated with pollen from plants containing a *pLAT52:GUS* transgenic plants (Tsukamoto and Palanivelu 2010). This marker is very useful because labels the pollen tube cytosol and allows investigation of pollen growing and ultimately pollen tube burst. We could clearly see that in the wild-type plant the pollen tubes targeting all the ovules and the pollen tubes burst (blue spot) were detected in the mycropilar zone of the ovules (Figure 5C). Instead when analyzing *REM11_RNAi* carpels and despite the fact that all the ovules are reached by the pollen tube, only few of them show the pollen tube bursting signal (Figure 5C/5E). These results suggested that *REM11* is directly involved in the promotion of the pollen tube burst.

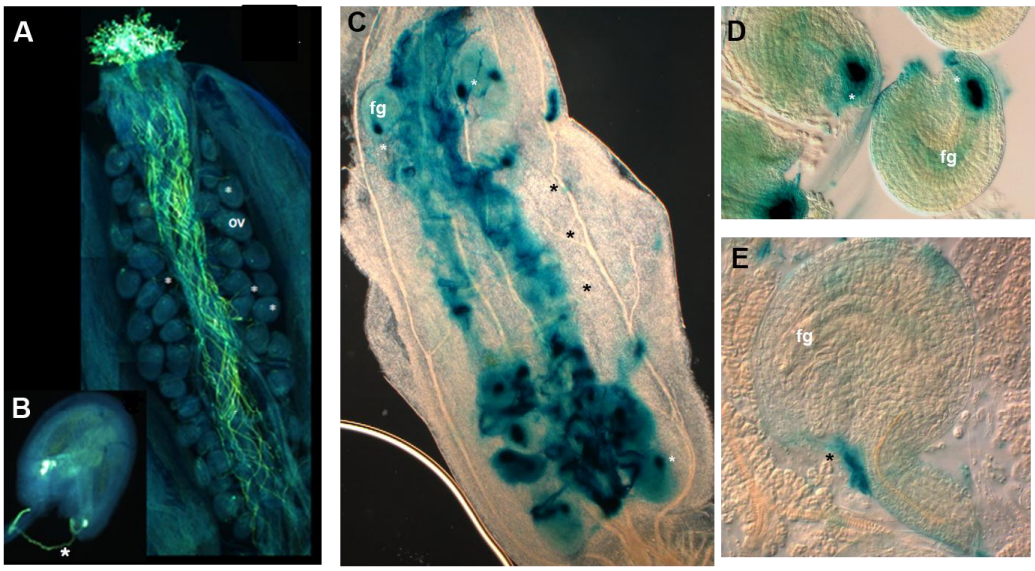


Figure 5. Double fertilization process in *REM11_RNAi*

(A) Aniline Blue Staining showed that all pollen tubes reach the mutant ovules (white asterisks). (B) Detail of a pollen tube reaching the micropyle. (C) *LAT52::GUS* crossed with mutant carpels showed that not all pollen tubes were bursting when the embryo sac was penetrated (black asterisks). In more detail a wild type ovule (C) and a mutant ovule (D) were is visible that the pollen tube arrives to the micropyle but no bursting was detected. ov-ovules; fg-female gametophyte.

REM11 interacts with VDD

REM11_RNAi mutant showed a very similar phenotype with the one described in *vdd-1/+* mutant (Matias-Hernandez et al. 2010). Both proteins belong to the same family whose members have a putative protein-protein binding domain and one protein-DNA binding domain (Romanel et al. 2009; Swaminathan et al. 2008). Due to the similarity of the mutant phenotype we wanted to check if these to proteins were able to interact. We performed a yeast-two hybrid assay (Y2H) and we were able to detect an interaction between these two proteins (Figure 6). In conclusion it seems that these two REMs are able to interact and regulate targets genes required for the correct functionality of the two synergids.

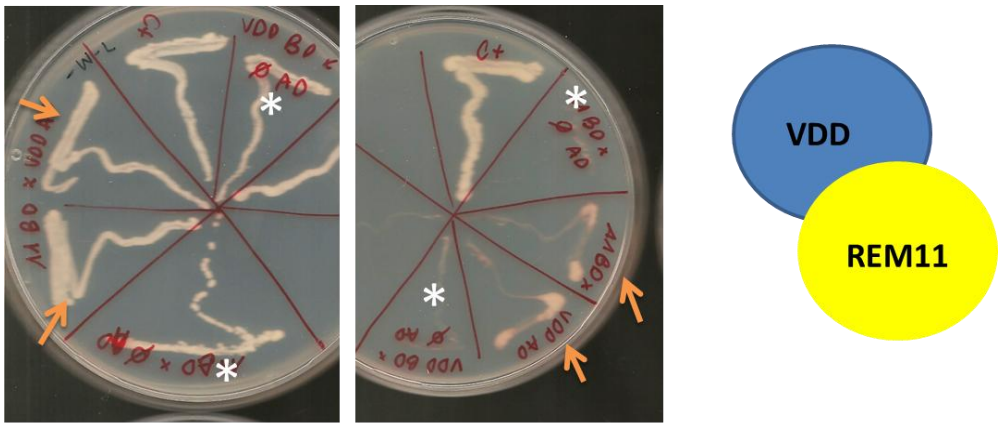


Figure 6. Y2H experiment

Arrows indicate the detection of a positive interaction when the colonies growth in -W-L-A selective medium. White (*) indicate the controls REM 11BD for empty AD and VDD AD for empty BD. C+ stands for the positive control.

The pollen tube doesn't burst the sperm cells don't migrate in *vdd-1/+ and REM11_RNAi*

To better understand the fertilization defect observed in *vdd-1/+* and in *REM11_RNAi* we have used a combination of marker lines to observe fertilization *in vivo* (Ingouff, M. et al. 2007). To test if the pollen tube was bursting we used a marker for the pollen tubed called *pLAT52::GUS* (Tsukamoto et al. 2010), marker that has the ability to stain the cytosolic content of the pollen tube that became easy to follow during growth and burst. As shown in Figure 7 the pollen tube burst is visible in wild type (blue spot). In the *vdd-1/+* mutant carpels we were able to see that not in all ovules the bursting was occurring Figure 7B as shown before for *REM11_RNAi* mutant.

As a second step and using a marker line for the sperm cells, *HTR10::HTR10_RFP (Hamamura et al. 2011)* we analysed both mutants. We were able to visualise the sperm cell nuclei fusion with red fluorescent protein. The study of these markers in the mutant backgrounds allows us to see how the sperm cells are moving and if the delivery and fusion between both male and female gametes is taking place. In the wild type plant, as the pollen tube approaches the micropyle one of the synergid cells initiates to degenerate and is penetrated by the pollen tube, which arrests its growth, bursts and releases the two sperm cells to ensure double fertilization (Figure 7C). In *vdd-1/+* and *REM11_RNAi* mutant plants the two sperm cells stay in the mycropyle area and do not migrate inside the ovule in a very high percentage, 27%, corresponding to the number of ovule abortions (Figure 7D). Another experiment was done to test whether one of the sperm cells were able to degenerate based on the fact that when one of the synergids starts the apoptosis process is produce a fluorescence signal. So we hand emasculated and pollinated wild-type and mutant carpels and we were able to see that in 95% of wild-type ovules emitted the signal. Instead only 85% emitted in the ovules in the *vdd-1/+* mutant, showing that in 25% of the embryo sac ovules was

not emitting any kind of signal (Figure 7 E and F). This data supports the fact that the two sperm cells were not migrating in the first place. Ultimately to confirm that the pollen stopped growthing after entering the embryo sac and also to confirm the non-degeneration of the synergid we planned to use a Transmission electron microscopy (TEM). Unfortunately we don't have until this moment the final results of the experiment, but we had some preliminary results that supported our hypothesis. In figure 7G is clear the fact that the two synergids stay intact (16 hours after pollination) and that the pollen tube (red arrow) stays on the side of the two synergids).

These results strongly show the important and the direct involvement of VDD and REM11 in the synergid degeneration process and in the consequent migration of the two sperm cells.

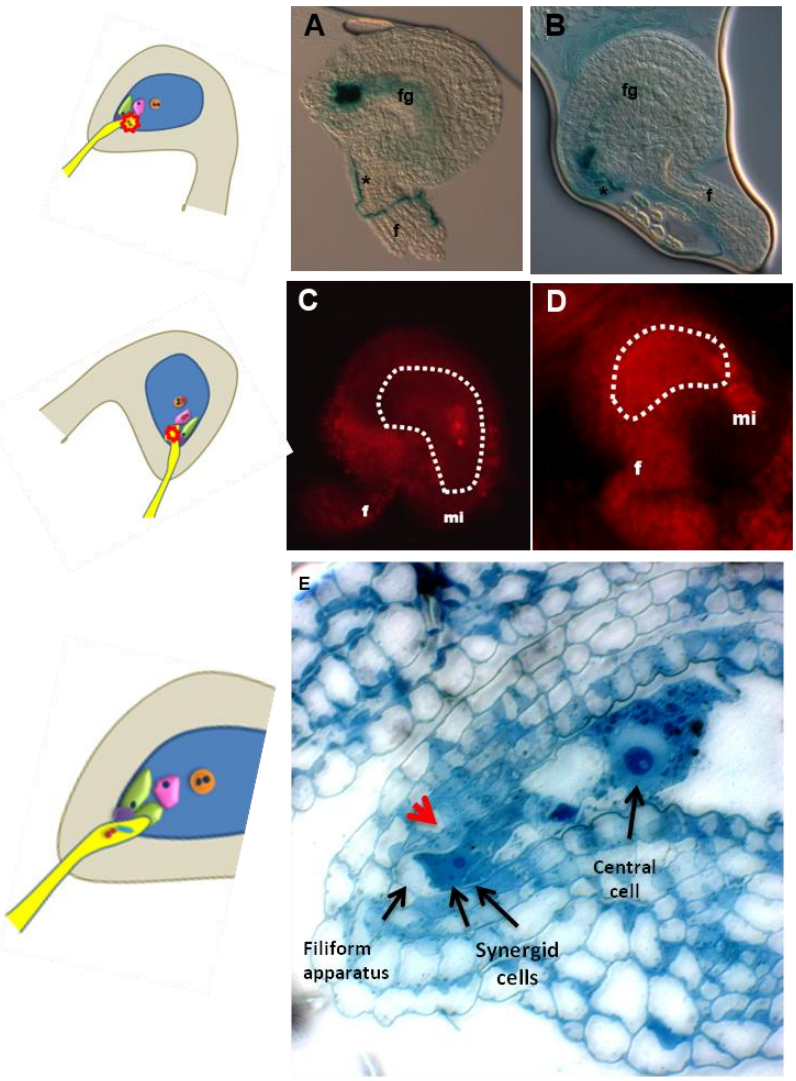


Figure 7. Detailed analysis of the double Fertilization process in *vdd-1/+* and *REM11_RNAi* mutants

(A-B) wild type ovule receives a pollen tube marker with *pLAT52::GUS* marker (A) mutant ovule receives a pollen tube marker (B).

(C) ovule without synergid degeneration (D) ovule with degenerating synergid.

(E) wild-type ovule that shows the two sperm markers fusin with egg and central cell. (F) Mutant ovule with the two sperm cells in the mycropile area. f-funiculus; fg-female gametophyte; (*)-pollen tube; mi-micropyle.

MYB98 is correctly expressed in the mutants

To better understand what was happening at a molecular level with the mutants we started our analysis by crossing *MYB98::GFP* (Kasahara et al. 2005) with our mutants. MYB98 was described has being one of the transcription factors responsible for the pollen tube attraction.

In figure 8 is clear the fact that analysing an F2 generation for the cross all the mutant ovules correctly expressed the GFP signal. This result indicates that VDD and REM11 have no influence in the expression of *MYB98*.

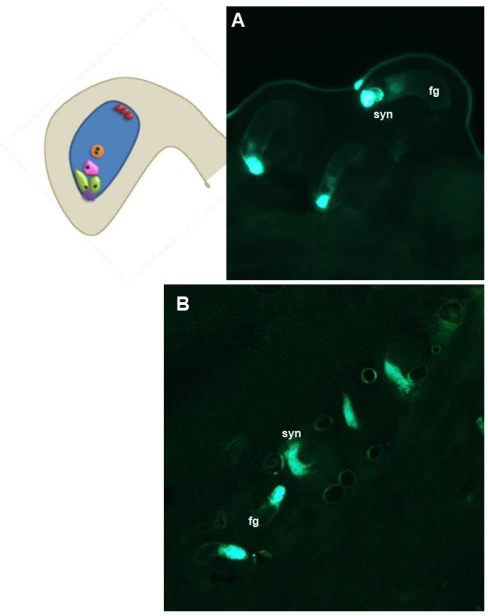


Figure 8. MYB98::GFP in a mutant background

(A) *MYB98::GFP* in a wild type situation. (B) *MYB98::GFP* crossed *REM11_RNAi* in a F2 situation, an identical result was obtained for *vdd-1*. syn-synergid; fg-female gametophyte.

RNA-Sequencing experiment, searching for VDD targets

To have a general analysis of what genes are up or down regulated in the mutant we planned an RNA-Seq experiment. The first thing that we did was to create a new mutant for VDD. *VDD* is in heterozygous being the homozygous situation lethal, therefore only 50% of the embryo sac have VDD expression down-regulated (Matias –Hernandez et al. 2010). So we constructed an RNA interference line for *VDD* and introduce it into *vdd-1* the T-DNA insertion line for *VDD*. We obtained plants almost completely sterile (Figure 9A). Finally, we verified by qRT-PCR that the levels of *VDD* transcript were down regulated respect to the wild-type. This new mutant was called *vddM.*

For the RNA-sequencing we extracted RNA from mature wild type and *vddM* mutant pistils. The RNAs were send to sequence by Illumina system and finally the data was analysed with a bioinformatical software – see Methods (Figure 9B).

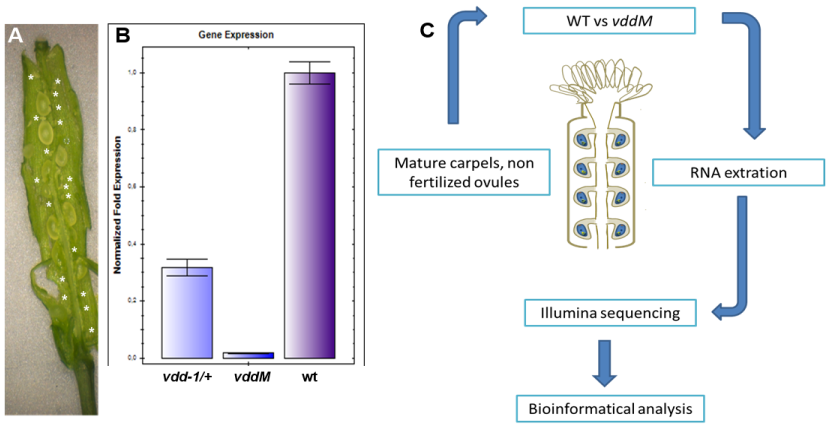


Figure 9. RNA-sequencing method.

(A) silique of *vddM* , with high 65% of ovule abortions. (B) qRT-PCR analysis showing the down-regulation of the *vddM* mutant comparing with wild-type and *vdd-1*/+ mutant. (C) RNA-sequencing method.

RNA sequencing results

Using the Bioinformatical software, AGRI: GO (http://bioinfo.cau.edu.cn/agriGO/) we could group the preliminary subset of down-regulated or up-regulated genes in *vddM* into different groups (Figure 10 and 11).

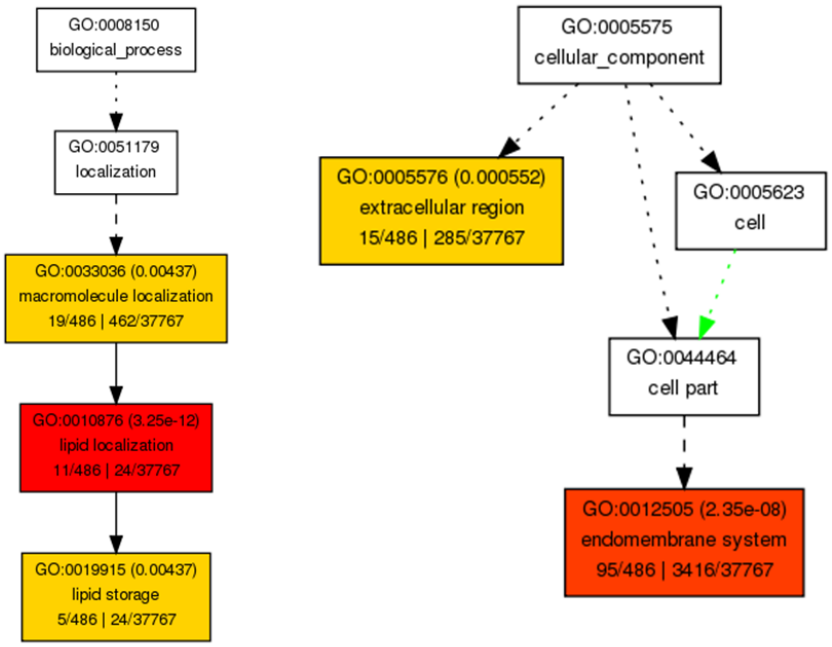


Figure 10. Down-regulated genes. GO analysis of RNA-SEQ output. In color (orange and red) the group of genes that are significantly down-regulated in *vddM* mutant.

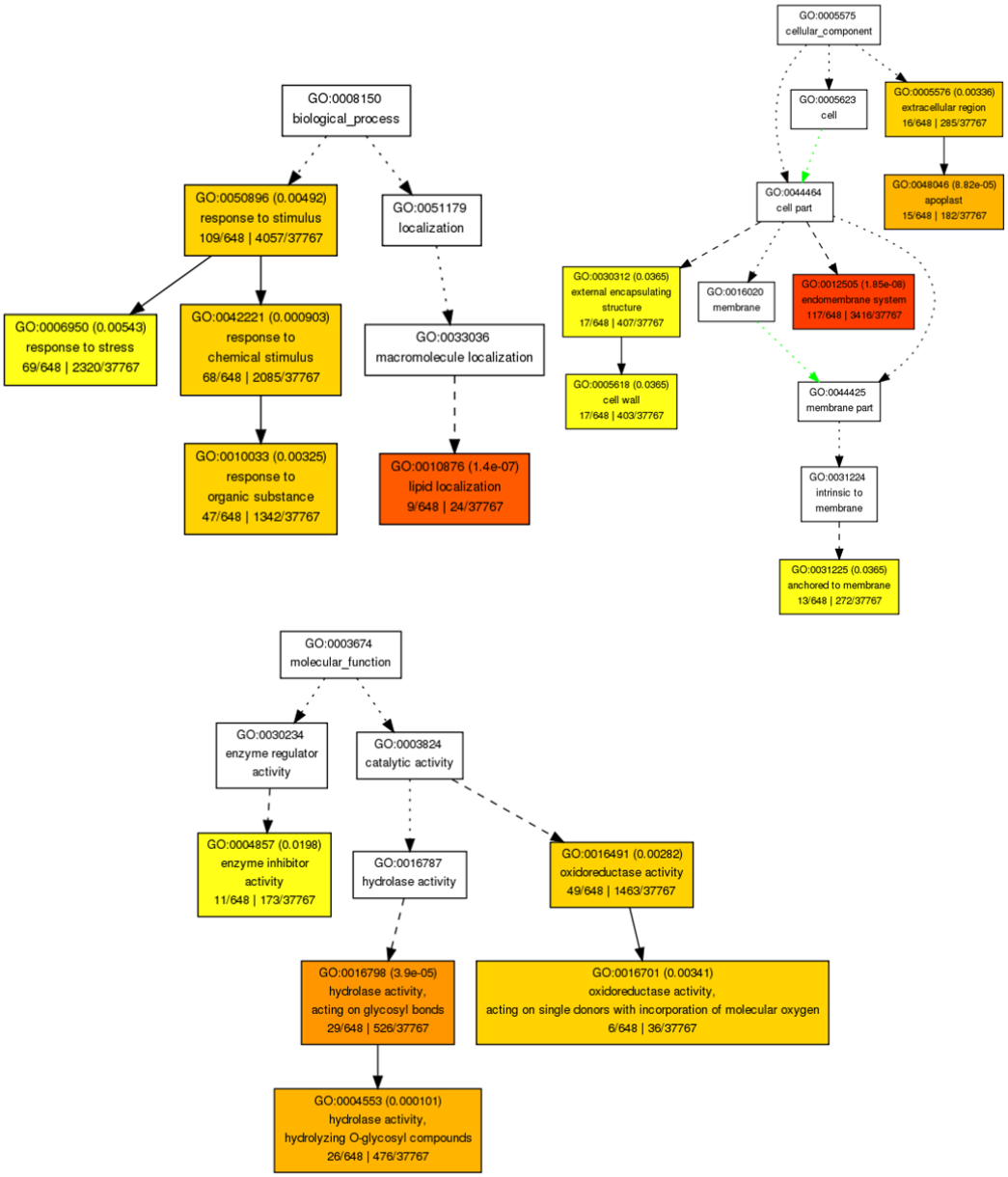


Figure 11. Up-regulated genes. GO analysis of RNA-SEQ output. In color (orange and red) the group of genes that are significantly up-regulated in *vddM* mutant.

The most down-regulated genes included the endomembrane related genes. Among the up-regulated categories we could found endomembrane related genes again plus hydrolyse related genes.

REM family characterization studies

Marta Mendes, Piero Morandini, Marcio Alves-Ferreira and Lucia Colombo

Abstract

We have previously shown that VDD and REM11 are able to interact and to form a complex required for the correct function of the synergids. VDD and REM11 belong to the REM transcription factors family that is poorly characterized. For this reason we have decided to make a genome wide characterization of this family.

REM genes co-expression analysis

We have used the Pearson Correlation Coefficient (PCC) analysis to identify clusters of co-expressed *REM* genes. This co-expression analysis might reflect possible interactions within the family. Two matrices were generated for REM-REM analysis, using both untransformed P-Lin (see supplemental table 3) and logarithm transformed P-Log (see supplemental table 2). Expression data were analyzed at different threshold values of PCC. A network was predicted between the *REM* genes, showing that they could be involved in the same transduction pathway, probably acting as protein complexes (Berri et al. 2009).

As a result of this analysis we predicted possible interactions among the REM family, the REM genes that scored a high value (>0.6) in (P-Log) and (p-Lin) were used to build a network (Figure 12). Interestingly, the Pearson Coefficient mainly identified co-regulation of *REM* genes which did not appear to be closely related in the phylogenetic tree (Romanel et al. 2009). The identification of co-regulation among phylogenetically unrelated genes might be a powerful tool to identify functional redundancy especially for those genes for which the function is still unknown.

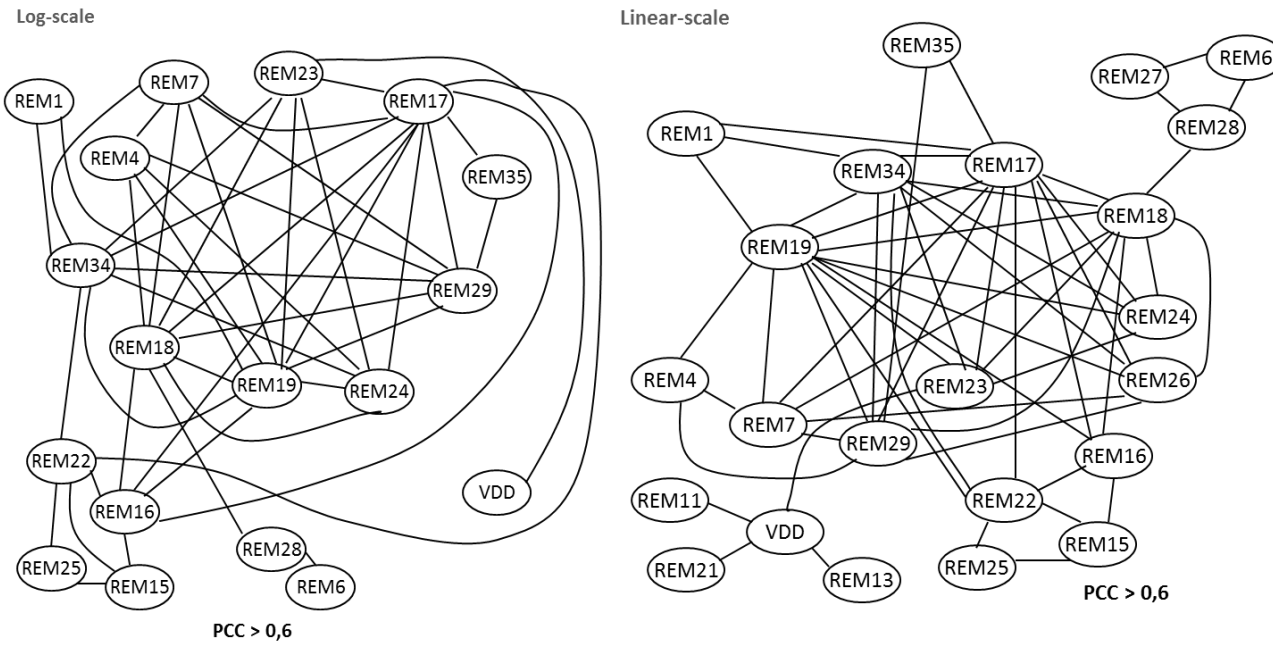


Figure 12. Proposed interaction network constructed using the Pearson coefficient analysis, based on P-Log (A) P-Lin (B).

How to test the possible REM interactions?

Based on the bio-informatical analysis we started to test the predicted REM-REM interactions. Due to the high number of interactions to be tested we designed a strategy to be able to test as many interactions as possible.

Each bait/prey pair was introduced in the α-AH109 yeast strain (Clontech), and as a control for auto activation, each bait was also co-transformed into the yeast strain with the empty AD vector, and each prey was co-transformed with the empty BD vector. The co-transformations were selected in a selective medium. Bait/prey pair colonies that grew on all selective media (−Trp-Leu-Adenine-His and

supplemented with increasing concentrations of 1 mM to 2.5 mM 3-Amino-1,2,4-triazole). System used for the co-transformations and plate organization are described in Figure (13).

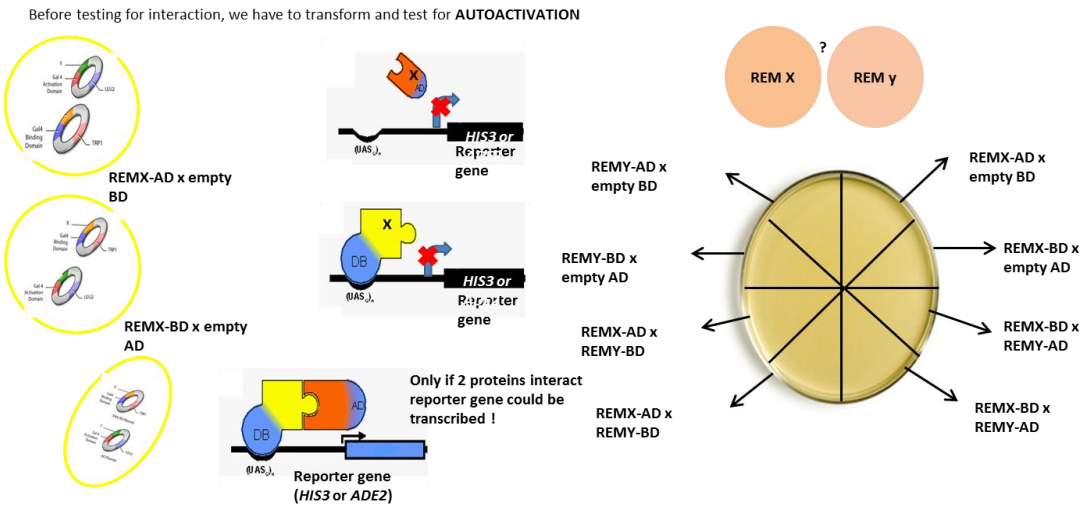


Figure 13. Strategy for Yeast transformation

(A) Co-transformation and auto-activation test of each Bait/Prey. (B) Interaction test between two REM proteins, organization of each plate.

Analysis of the interactions among REM factors

We have tested several REM proteins to verified interaction based on the co-regulation expression analysis. Two examples are depicted in figure 14 a positive interaction was found between REM23 and REM19 and between REM 23 and REM1.

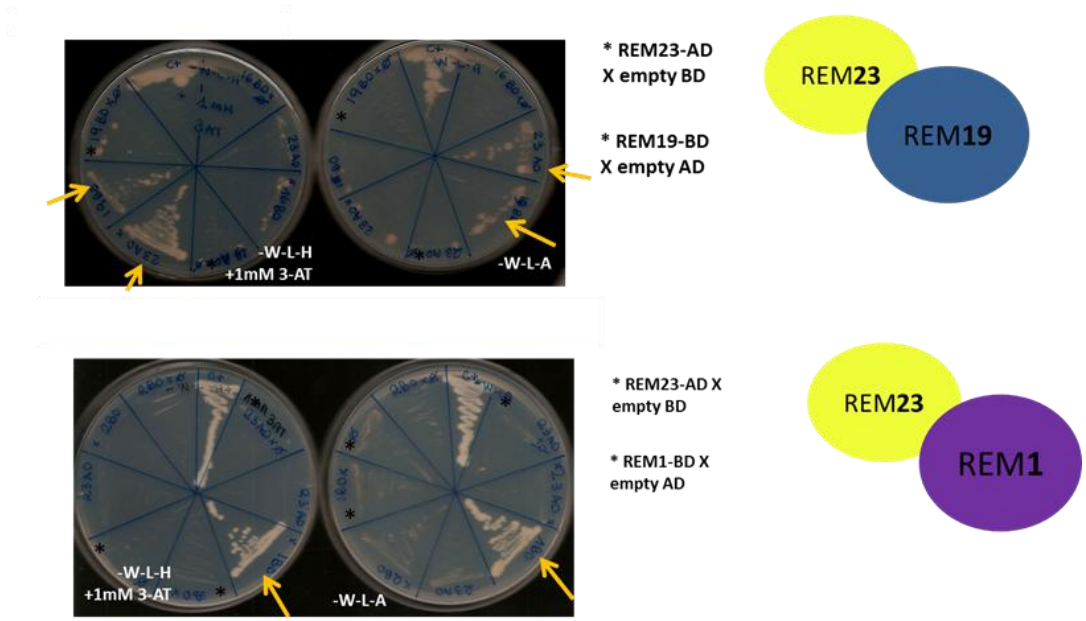


Figure 14. Example of two different interactions between REM members. These plates were left at 28ºC for one week. Arrows indicate the confirmation of the interaction.

We weren't able till now to test all the possible interactions between the REM family, but with the results obtained we could already validate some putative interaction deduced by the co-expression analysis (Figure 15). Furthermore we have confirmed some interactions with REM11 that could have a role during the double fertilization process. This type of study will be very helpful to investigate possible genetic interaction among REM transcription factors family members.

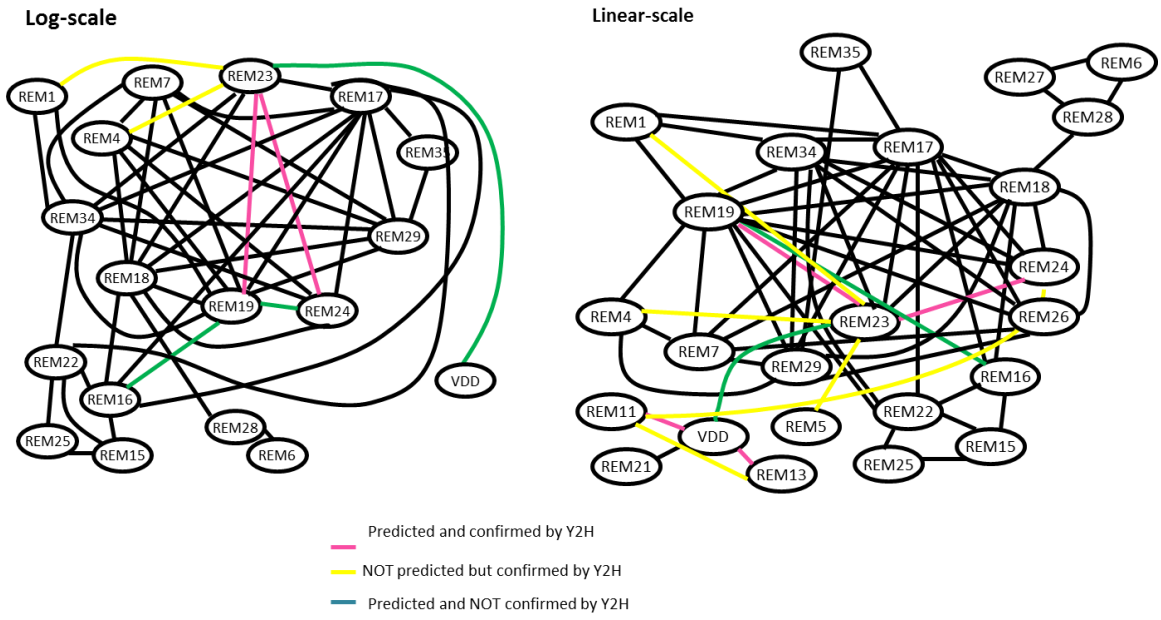


Figure15. Network according to the yeast-two hybrid results.

Discussion

VDD and REM11 control a very specific process within the synergid cell

In both mutants *vdd-1/+* and *REM11_RNAi,* we initially detected a synergid defect crossing the mutants with a specific synergid marker, till the pollen tubes were still arriving to all ovules, the attraction was not affected in the mutants. Deeper analysis using the *pLAT52::GUS* marker (Tsukamoto et al. 2010) line showed that the pollen tubes weren't bursting in all the ovules. Following them with the *HTR10::HTR10_RFP* marker (Hamamura et al. 2011) we saw that the two sperm cells were located at the micropylar region of the ovule. Additionally, evidences showed that not all the synergids were degenerating as supposed. Altogether these results showed the importance and the direct involvement of VDD and REM11 in the synergid degeneration process. Both mutants presented a female gametophyte defect that resulted in the correct arresting of the pollen tube in the micropylar zone however the synergids do not degenerate and consequently the two sperm cells do not migrate inside the embryo sac resulting in a double fertilization failure.

Detailed experiments of TEM will help to understand were the pollen tube is exactly stopping. The first results showed that the ovules had 16 hours after pollination had two synergids intact and that the pollen tube stayed just in the border of the synergids supporting our previous results. The synergids were extensively described as being responsible for the pollen tube attraction (Higashiyama 2002; Higashiyama et al. 2006; Okuda et al. 2009). Also mutants were shown to be defective in the pollen tube attraction, for instances *myb98*, this mutant have been shown to have a defective in attracting the pollen tubes (Kasahara et al. 2005). Our mutants presented defective synergids that were able to attract the pollen tubes. We introduced in our mutants *pMYB98::GFP* line and we were able to see that this gene was correctly expressed in the synergids. This result

strongly suggests that there are two processes controlled by the synergid cells and that have two transcriptional pathways completely different. VDD and REM11 are so directly involved in the degeneration pathway and aren't involved in the pollen tube attraction.

We demonstrate that the pollen tube arrest once penetrate in the embryo sac and the synergids don't initiate to degenerate. FERONIA is involved in the arrest of the pollen tube, therefore the *fer* mutant presents an overgrowth of the pollen tube. In *fer* mutant (Huck et al. 2003) the synergid penetrated by the pollen tube degenerates suggesting that there might be two check points required to the pollen tube arrest. First the pollen tube penetrate the embryo sac in the mycropilar zone and arrests its growth, the synergid initiates to degenerate and the pollen tube arrest its growth again inside and burst.

Another controversial studies published recently (Beale et al. 2012; Kasahara et al. 2012) showed that the two male gametes should fuse with the two female gametes in order to stop the production the attractants by the synergids, the non fusion of the gametes lead that different pollen tubes enter the same ovule. Curiously in our situation this is not true, we have only one pollen tube that gets attract in even if the fusion of the gametes was not happening. This situation clearly shows that an extra attraction of pollen tubes should be regulated by other processes and not only by the fusion of the male and female gametes.

The only gene published that may have a role similar to *VDD* and *REM11* is the mitochondrial chaperone *GFA2* (Christensen et al. 2002). *gfa2* mutant presented a defect in the synergid degeneration, the authors didn't explain the mechanism in detail. We are now planning different experiments to try to see a link between, GFA2, VDD and REM11.

Furthermore a detailed analysis of RNA-Seq done with *vddM* mutant will provide new information about genes that will for sure be involved in the degeneration/apoptosis pathway of the synergid cells. Preliminary analysis showed

that endo-membrane related genes were in a high percentage up and down regulated groups, maybe these genes will have a role in the apoptosis process.

REM interactome

The REM interactome will bring new and important information to REM family that could be a very important family during ovule development. The complexes will give clues that some genes maybe have a redundant role together during plant development making easy to choose which mutants to work and which genes to study. For instances we already found another REM that interact with VDD-REM11 complex, REM13 seems to interact with both of them being a good choice to study in the near future. A detailed co-expression analysis will provide information about other genes families that could interact with the REM's. These data will be the base for future studies for the functional characterization of this family and its role during ovule development.

Material and Methods

Plant Material and Growth Conditions

Arabidopsis thaliana wild-type (ecotype Columbia), mutant and embryo sac marker plants were grown at 22°C under short-day (8 h light/16 h dark) or long-day (16 h light/8 h dark) conditions. The Arabidopsis stk, shp1shp2 and stk shp1shp2 mutants were kindly provided by M. Yanofsky (Pinyopich et al.,2003).

ChIP and Quantitative Real-Time PCR Analysis

In REM11 genomic region we found two putative CArG boxes allowing 1 mismatch. The first one 100 base pairs before (5'ctattaatgg3') the translation starting site and the second 80 base pairs after (5'cttattttgg3'). Primers were designed specifically to test possible enrichments 1st CArG FW 5'gggccttagcgataccttgg3'; 1st CArG rev 5'gtgatttgatctaaaggtgttggcc3'; 2nd CArG FW 5'gaacacaagaggtttttcacttctctg3'; 2nd CArG rev 5'ccagatcatcaccggattcactagg3'. Enrichment folds were detected using a SYBR Green assay (Bio-Rad, http://www.bio-rad.com/). The real-time PCR assay was performed in triplicate using a Bio-Rad C1000 Thermal Cycler optical system. ChIP-qPCR experiments and relative enrichments were calculated as reported before (Gregis et al 2008 and Hernandez et al 2010).

Generating *REM11* and *VDD* RNA interference lines

To make *p35::REM11_RNAi* and *p35S::VDD_RNAi* constructs we used the Arabidopsis vector pFGC5941 for dsRNA production was obtained from ABRC (stock no. CD3-447).

For *REM11* a 247-bp fragment of *REM11* cDNA (position 364–611) was amplified by PCR using primers AtP_2757 (5′-ggggacaagtttgtacaaaaaagcaggctacatctggaaaaacttggat -3′) and AtP_2758 (5′ggggaccactttgtacaagaaagctgggtgatcatcaccggattcacta -3′). For *VDD* a 197 bp fragment was obtained using ATP_2783-5’ggggacaagtttgtacaaaaaagcaggctattctttgcccacaaccagag3’ and ATP_2784-5’ggggaccactttgtacaagaaagctgggtctctttcttccataatctgacc3’.

The fragments were amplified using with Phusion High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA) and purified using the GeneJET Gel Extraction Kit (Thermo Scientific). The amplified fragments were then cloned into a PDONOR 207 (Invitrogen) and then pFGC5941, following the Gateway system (invitrogen). Latter on Agrobacterium-mediated transformation of Arabidopsis plants was performed using the floral dip method (Clough and Bent, 1998). Transgenic plants were selected with 10 ng/μL BASTA.

Cytological assays

The gametophytic cell identity reporter lines used encode a nuclear localization signal that is in-frame with the GUS reporter gene. The egg cell-specific marker was kindly provided by Stefanie Sprunck (unpublished data). The synergid cell-specific marker was kindly provided by Ueli Grossniklaus (Institute of Plant Biology,University of Zurich, Switzerland) (Gross-Hardtet al., 2007). The central cell-specific marker was kindly provided by Rita Gross-Hardt (Department of Developmental Genetics, University of Tubingen, Germany) (Mollet al., 2008). The antipodal cell-specific marker, kindly provided by Rita Gross-Hardt, was generated as described by Yu et al. (2005): the promoter of At1g36340 was amplified using primers 5’-agtgaggcgcgcctgatcattaagtttaggggt-3’ and 5’-

agtgattaattaattacgagaaatcaccaaac-3', and cloned upstream from the NLS_GUS reporter into pGIIBar binary vector (Gross-Hardt et al., 2007) (cloning details are available uponrequest). The synergid cell marker *pMYB98::MYB98_GFP (Kasahara et al. 2005).*

For female gametophyte cell identity determination, marker lines were used as female and pollinated with *REM11_RNAi or vdd-1/+* pollen. In the F1generation, heterozygous plants were self-fertilized, and the presence/absence of the *REM11_RNAi* was analysed in the F2generation by PCR. The presence of the reporter genes was analysed by GUS staining, confirming the correct expression in wild-type back-ground. For GUS staining, flowers were emasculated and harvested12 h after pollination as described by Liljegren et al. (2000). Samples were incubated in chloral hydrate:glycerol:water solution8:1:2, dissected and observed using a Zeiss Axiophot D1 micro-scope equipped with DIC optics.

To analyze ovule development in *REM11_RNAi* plants, flowers at different developmental stages were cleared and analyzed as described previously (Brambilla et al., 2007).

Expression analysis by Quantitative real-time RT-PCR

Quantitative real-time RT-PCR experiments were performed on cDNA obtained from inflorescences. Total RNA was extracted using the LiCl method (Verwoerd et al., 1989). DNA contamination was removed using the Ambion TURBO DNA-free DNase kit according to the manufacturer's instructions (http://www.ambion.com/). The treated RNA was subjected to reverse transcription using the ImProm-IITM reverse transcription system (Promega).*REM11* transcripts were detected using a Sybr Green Assay (iQ SYBR Green Supermix; Bio-Rad)

with the reference gene UBIQUITIN. The real-time PCR assay was conducted in triplicate and was performed in a Bio-Rad iCycler iQ Optical System (software version 3.0a).

Relative enrichment of *REM11* transcripts was calculated normalizing the amount of mRNA against UBIQUITIN fragment. Diluted aliquots of the reverse-transcribed cDNAs were used as templates in quantitative PCR reactions containing the iQ SYBR Green Supermix (Bio-Rad).

The difference between the cycle threshold (Ct) of REM11 and that of UBIQUITIN (DCt = Ct REM11 2 Ct UBIQUITIN) was used to obtain the normalized expression of VDD, which corresponds to 2 2DCt. The expression of *REM11* was analyzed by the following primers:
REM11 Forward, 5' gaaaggcggtatctggatga 3' And REM11 Reverse, 5'ccttgacaaagatgcaacca 3'. The expression of UBIQUITIN was analyzed using the following primers: UB forward, 5'ctgttcacggaacccaattc-3', and ub reverse, 5'-ggaaaaaggtctgaccgaca-3'.

Double fertilization analysis

Pollen tube guidance, reception and burst analysis

For in vivo pollen tube guidance experiments, pistils were hand-emasculated and pollinated after 24 h with wild-type pollen. After 16–18 h, pistils were carefully isolated from the plants and fixed in a solution of acetic acid and absolute ethanol (1:3), cleared with 8 N sodium hydroxide and labelled with aniline blue (Sigma, http://www.sigmaaldrich.com/). For in vivo pollen tube reception and burst of the tubes, wild-type and double mutant pistils were emasculated and crossed after 24 h with pLAT52:GUS pollen. After 16–18 h, the pistils were carefully collected and stained for GUS activity (Liljegren et al., 2000). Samples were incubated in

clearing solution (Brambilla et al., 2007), dissected under a Leica MZ6 stereo microscope, and observed using a Zeiss Axiophot D1 microscope equipped with differential interference contrast (DIC) optics (http://www.zeiss.com/). Images were captured using an Axiocam MRc5 camera (Zeiss) with axiovision software (version 4.1).

Sperm cell migration analysis

For sperm cell migration experiments, pistils were emasculated and crossed after 24 h with the pHTR10:HTR10-RFP marker line. Pistils were collected after 16–18 h, samples were dissected under a Leica MZ6 stereo microscope, and images were obtained using a Zeiss Axiophot D1 microscope equipped with DIC optics and a rhodamine filter set.

In-situ hybridization analysis

For in situ hybridization analysis, Arabidopsis flowers were fixed and embedded in paraffin as described previously (Huijser et al., 1992).

Sections of plant tissue were probed with digoxigenin-labeled VDD antisense RNA corresponding to nucleotides 240 to 557. Hybridization and immunological detection were performed as described previously (Coen et al., 1990).

Sections of plant tissue were hybridized with digoxigenin-labelled *REM11* antisense probe, amplified using primers atp_2759 (5'-acatctggaaaaacttggatc-3') and atp_2760 (5'- gatcatcaccggattcactag -3').

RNA extraction, cDNA library preparation, and sequencing for RNA-seq

Total RNA was extracted from wild-type and *vddM* mutant mature carpels with the Qiagen Kit according to the manufacturer's instructions. DNA contaminations were removed using the PROMEGA RQ1 RNase-Free DNase according to the manufacturer's instructions. RNA quality integrity was analyzed by electrophoresis gel and was validated on a Bioanalyzer 2100 (Aligent, Santa Clara, CA); RNA Integrity Number (RIN) values were greater than 7 for all samples. In order to confirm that in *vddM* mutant samples *VDD* was not expressed, VDD expression was checked by real time PCR with primer RT 795 (5'gggaaggtcatggcaagtta3') and RT 796 (5' ccatctgcctcgaatatggt3').

Sequencing libraries were prepared according to the manufacturer's instructions (Illumina TruSeq mRNA-seq kit) and sequenced with the Illumina Lane single-read 50bp. The processing of fluorescent images into sequences, base-calling and quality value calculations were performed using the Illumina data processing pipeline (version 1.8). Raw reads were filtered to obtain high-quality reads by removing low-quality reads containing more than 30% bases with Q < 20. Finally, a quality control on the raw sequence data was performed using FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/.)

Pearson coefficient

Pearson correlation values were calculated essentially as described by (Toufighi *et al*.) for the 'Expression Angler'. To this purpose a Visual C++ based program was developed (P. Morandini, L. Mizzi, unpublished) to calculate the correlation value from the data obtained with the ATH1 GeneChip from Affymetrix and deposited at the NASC array database http://affy.Arabidopsis.info/narrays/experimentbrowse.pl webcite as of September 2008. For the calculation of Pearson coefficient from log

values, data were simply transformed into log before calculating the correlation value. From such values, networks of Arabidopsis genes were produced using the program dot http://citeseer.ist.psu.edu/gansner93technique.html website. The input text file for dot was prepared using a script that filtered *WRKY* genes with a reciprocal coefficient of 0.6 or higher from the complete table of Pearson coefficient. Intensity of arrow colours is proportional to the coefficient between each pair of *WRKY* genes. A more detailed explanation of the method used is reported in Menges *et al*. in the section *Global expression correlation analysis* in *Methods*.

Mapping of short reads and assessment of gene expression analysis for RNA-Seq

Evaluation and treatment of raw data was performed on the commercially available CLC Genomics Workbench v.4.7.1 (http://www.clcbio.com/genomics/). After trimming, the resulting high-quality reads were mapped onto the the *Arabidopsis* genome (TAIR10). Approximately, 25M reads of each sample that uniquely mapped with ≤2 mismatches were used for further analyses. The read number of each gene model was computed based on the coordinates of mapped reads. A read was counted if any portion of that read's coordinates were included within a gene model. As CLC Genomics Workbench v.4.7.1 distributes multireads at similar loci in proportion to the number of unique reads recorded and normalized by transcript length, we included in the analysis both unique reads and reads that occur up to 10 times to avoid undercount for genes that have closely related paralogs (Mortazavi et al., 2008). Gene expression value was based on reads per kilobase of exon model per million mapped reads (RPKM) values (Mortazavi et al., 2008). The data was normalized using a basic quantile approach. The fold change and differential expression values between wt and *stk* mutant was derived using the normalized RPKM values of the corresponding transcripts. To obtain statistical

confirmation of the differences in gene expression among the wt and *stk* mutant, we compared the RPKM-derived read count using standard t-test for two group comparisons. A threshold value of P =0,05 was used to ensure that differential gene expression was maintained at a significant level (5%) for the individual statistical tests. We estimated that statistical analysis was reliable when applied to genes showing a differential expression RPKM value ≥2 (i.e. five mapped reads per kilobase of mRNA). Differential expression was estimated and we calculated FDR, and estimated FC in terms of RPKM**.**. Transcripts that exhibited an FDR ≤ 0.05 and an estimated absolute FC ≥ 1,5 were determined to be significantly differentially expressed.

Yeast-two-hybrid screen

cDNA of each candidate gene was used to perform confirmation of the predicted interactions. RNA was extracted from Arabidopsis inflorescences containing flowers in several stages of development using the RNeasy Mini Kit (Qiagen). cDNA was synthesized with the Revert Aid First Strand cDNA Synthesis Kit (Thermo Scientific, Waltham, MA, USA). Full-length cDNA of these genes was amplified by PCR (Table 3) with Phusion High-Fidelity DNA Polymerase (New England Biolabs, Ipswich, MA, USA) and purified gel using the GeneJET Gel Extraction Kit (Thermo Scientific). The cDNA sequence of each gene was individually cloned into the pGADT7 vector (Clontech) via Gateway (Invitrogen).

Each bait/prey pair was introduced in the α-AH109 yeast strain (Clontech), and as a control for autoactivation false-positives, each bait was also co-transformed into the yeast strain with the empty AD vector, and each prey was co-transformed with the empty BD vector. Bait/prey pair colonies that grew on all selective media at 28 °C at least for one week (−Trp-Leu-Adenine-His and supplemented with increasing concentrations of 1 mM to 2.5 mM 3-Amino-1,2,4-triazole) were considered positive for interaction.

Tabel 3. Primers used to clone REM cDNA

Name	Sequence (5'-3')
REM1-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGTGGAATCGGAGCTAAG
REM1-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgCTACTTTCTTCCTCGTCTTTTCCT
REM2-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGTGGTCGAAAAACAAAC
REM2-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTCAGCTCATATCTTTGGAGCT
REM3-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGCCATGGCCTATGAG
REM3-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTTATGCTTTCCTCTTCCTTAGTAAC
REM4-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGCAAATGGATTCCGCAC
REM4-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTCACACATACTCATTGACACGAAAG
REM5-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGCCACGCCCTTTCTTC
REM5-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTCAGACGTACTCGTTGACTCG
REM11-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTcGATGAACACAAGAGGAAATTACTCTAATG
REM11-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTCATCCGCTGATAATCTTGAC
REM12-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGTACTACCCATTTCTTTCACC
REM12-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTCATCCACTGATAATGTGAATCTTG
REM13-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTcGATGGAAGAGAACTCAAGTAAAAAAAAGC
REM13-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTCAGGCCTTGATAATGTGGAC
REM14-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGCTGACACTGGTGAGG
REM14-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTTACTGATGTTTGTCTTGATCGTC
REM15-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGTGTTAGAAAATGGGTTTC
REM15-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTCAATTCTTCACGAAGTGCA
REM16-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGCTGACGATAGCGAAT
REM16-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTCAAGCATCTCTACGCAAGAC
REM17-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGCCTCTTCAATGGCTAG
REM17-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgCTACAGATCCTTACCGCGGA
REM18-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGTGGGCAACGAAGTCG
REM18-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTCAAGTGCTTGCTGCTGC
REM19-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGCTTCCGGTGATGTAC
REM19-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTTAATAATAGTCACTATCACTTGCATC
REM20 (VDD)-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTcGATGGTGAAAAACAAAGCTTTTTTTG
REM20 (VDD)-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgCTATTCTTTGGAGACTTTCACACG
REM21-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGAAGAACAAAGCGTTTGGTC
REM21-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTTATTTAGAGACTCTAACAAGCATCTC
REM22-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGGTAAGAGTAGTAACATAGTTTTTGATC
REM22-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTCAATCGATCATGCACTGATCG
REM23-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGCTAGAAACAGTGACAATG
REM23-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTTAGGGTTTTGAGACAAAGAGA
REM24-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGCTACGAACGCTGGT
REM24-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTTAGGGTTTGGAGACAACTAGAAG
REM25-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGCAAGTAACCCTGATTTTC
REM25-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTTATGGCTTTGAGACAACTAGAAG
REM26-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGTTGCCTAAGAATGCTATTG
REM26-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgCTAGTCTTTCTTGACGAAATGC
REM27-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGTTGCCTAGGAATGCTACTG
REM27-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTCATGTTTTATTGATGAAGTGCACC
REM28-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGCGGATCAATCTCTTCTAC
REM28-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgCTAAGACAACCTTGAGGCATAGTT
REM29-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGCGAATTCAAGGATTTATCCG
REM29-RW	GGGGACCACTTTGTACAAGAAAGCTGGGTgCTAATTAACATTTTTGTATATAGTCTCCAC
REM34-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGCGGATCCACCACAT
REM34-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTCAAACCAGATTACTGCTGAGG
REM35-FW	GGGGACAAGTTTGTACAAAAAAGCAGGCTccATGGATGATCCAGCAATTTC
REM35-RV	GGGGACCACTTTGTACAAGAAAGCTGGGTgTTACTTGAGGATTTTGTTGATTTC

References

Alvarez-Buylla, E. R., S. Pelaz, et al. (2000). An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. PNAS 97(10):5328-33.

Amien, S., I. Kliwer, et al. (2010). "Defensin-Like ZmES4 Mediates Pollen Tube Burst in Maize via Opening of the Potassium Channel KZM1." PLoS Biol 8(6): e1000388.

Beale, K. M., A. R. Leydon, et al. (2012). "Gamete Fusion Is Required to Block Multiple Pollen Tubes from Entering an Arabidopsis Ovule." Curr Biol.

Berri, S., P. Abbruscato, et al. (2009). "Characterization of WRKY co-regulatory networks in rice and Arabidopsis." BMC Plant Biology 9(1): 120.

Bencivenga S, Colombo L, Masiero S. 2011 "Cross talk between the sporophyte and the megagametophyte during ovule development. Sex Plant Reprod. 24(2):113-21."

Boisson-Dernier, A., S. Frietsch, et al. (2008). Curr. Biol. 18: 63.

Borg, M., L. Brownfield, et al. (2009). "Male gametophyte development: a molecular perspective." J Exp Bot 60(5): 1465-1478.

Brambilla, V., R. Battagila, et al. (2007). Plant Cell 19(1): 2544.

Capron, A., M. Gourgues, et al. (2008). Plant Cell 20: 3038.

Causier, B., Z. Schwarz-Sommer, et al. (2010). "Floral organ identity: 20 years of ABCs." Seminars in Cell & Developmental Biology 21(1): 73-79.

Christensen, C. A., S. W. Gorsich, et al. (2002). "Mitochondrial GFA2 is required for synergid cell death in Arabidopsis." Plant Cell 14(9): 2215-2232.

Christensen, C. A., S. Subramanian, et al. (1998). "Identification of gametophytic mutations affecting female gametophyte development in Arabidopsis." Dev Biol 202(1): 136-151.

Coen, E. S. and E. M. Meyerowitz (1991). "the war of the whorls - genetic interactions controlling flower development." Nature 353(6339): 31-37.

Colombo, L., R. Battaglia, et al. (2008). "Arabidopsis ovule development and its evolutionary conservation." Trends Plant Sci.

Colombo, L., J. Franken, et al. (1995). The petunia MADS box gene FBP11 determines ovule identity. Plant Cell 7(11):1859-68.

Cordts, S., J. Bantin, et al. (2001). "ZmES genes encode peptides with structural homology to defensins and are specifically expressed in the female gametophyte of maize." Plant J 25(1): 103-114.

de Folter, S., R. G. H. Immink, et al. (2005). "Comprehensive interaction map of the Arabidopsis MADS box transcription factors." Plant Cell 17(5): 1424-1433.

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

Duarte, J. M., L. Cui, et al. (2006). "Expression Pattern Shifts Following Duplication Indicative of Subfunctionalization and Neofunctionalization in Regulatory Genes of Arabidopsis." Molecular Biology and Evolution 23(2): 469-478.

Eady, C., K. Lindsey, et al. (1995). "The Significance of Microspore Division and Division Symmetry for Vegetative Cell-Specific Transcription and Generative Cell Differentiation." Plant Cell 7(1): 65-74.

Egea-Cortines, M., H. Saedler, et al. (1999). "Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in Antirrhinum majus." Embo Journal 18(19): 5370-5379.

Egea-Cortines, M., H. Saedler, et al. (1999). "Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in Antirrhinum majus." EMBO J 18(19): 5370-5379.

Escobar-Restrepo, J. M., N. Huck, et al. (2007). Science 317: 656.

Faure, J. E. (2001). "Double fertilization in flowering plants: discovery, study methods and mechanisms." C R Acad Sci III 324(6): 551-558.

Favaro, R., A. Pinyopich, et al. (2003). "MADS-box protein complexes control carpel and ovule development in Arabidopsis." Plant Cell 15(11): 2603-2611.

Ferrario S, Immink RG, et al. (2003). The MADS box gene FBP2 is required for SEPALLATA function in petunia. Plant Cell 15(4):914-25.

Fu, Y., M. Yuan, et al. (2000). "Changes in actin organization in the living egg apparatus of Torenia fournieri during fertilization." Sexual Plant Reproduction 12(6): 315-322.

Giampiero Cai, M. C. (2006). "The microtubular cytoskeleton in pollen tubes: Structure and role in organelle trafficking. ." Pollen Tube: Cellular and Molecular Perspective 3: 157-175.

Guignard Leon. (1899). "Sur les anthérozoïdes et la double copulation sexuelle chez les végétaux angiospermes." C. R. Acad. Sci. 128

Gutierrez-Cortines, M. E. and B. Davies (2000). "Beyond the ABCs: ternary complex formation in the control of floral organ identity." Trends in Plant Science 5(11): 471-476.

Hamamura, Y., C. Saito, et al. (2011). "Live-cell imaging reveals the dynamics of two sperm cells during double fertilization in Arabidopsis thaliana." Curr Biol 21(6): 497-502.

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

Higashiyama, T. (2010). "Peptide Signaling in Pollen-Pistil Interactions." Plant Cell Physiol 51(2):177-89.

Higashiyama, T., R. Inatsugi, et al. (2006). "Species preferentiality of the pollen tube attractant derived from the synergid cell of Torenia fournieri." Plant Physiol 142(2): 481-491.

Higashiyama, T., S. Yabe, et al. (2001). "Pollen tube attraction by the synergid cell." Science 293(5534): 1480-1483.

Hiscock, S. J. and A. M. Allen (2008). "Diverse cell signalling pathways regulate pollen-stigma interactions: the search for consensus." New Phytol 179(2): 286-317.

Honma, T. and K. Goto (2001). "Complexes of MADS-box proteins are sufficient to convert leaves into floral organs." Nature 409(6819): 525-529.

Huck, N., J. M. Moore, et al. (2003). "The Arabidopsis mutant feronia disrupts the female gametophytic control of pollen tube reception." Development 130(10): 2149-2159.

Kasahara, R. D., D. Maruyama, et al. (2012). "Fertilization recovery after defective sperm cell release in Arabidopsis." Curr Biol 22(12): 1084-1089.

Kasahara, R. D., M. F. Portereiko, et al. (2005). "MYB98 is required for pollen tube guidance and synergid cell differentiation in Arabidopsis." Plant Cell 17(11): 2981-2992.

Kelley, D. R. and C. S. Gasser (2009). "Ovule development: genetic trends and evolutionary considerations." Sex Plant Reprod 22(4): 229-234.

Kessler, S. A., H. Shimosato-Asano, et al. (2010). Science 330: 968.

Levy, Y. Y., S. Mesnage, et al. (2002). "Multiple Roles of Arabidopsis VRN1 in Vernalization and Flowering Time Control." Science 297(5579): 243-246.

Mansfield, S. G., L. G. Briarty, et al. (1991). The mature embryo sac. Can. J. Bot. 69: 447.

Marella, H. H., Y. Sakata, et al. (2006). "Characterization and functional analysis of ABSCISIC ACID INSENSITIVE3-like genes from Physcomitrella patens." The Plant Journal 46(6): 1032-1044.

Márton, M. L., S. Cordts, et al. (2005). "Micropylar pollen tube guidance by egg apparatus 1 of maize." Science 307(5709): 573-576.

Matias-Hernandez, L., R. Battaglia, et al. (2010). "VERDANDI Is a Direct Target of the MADS Domain Ovule Identity Complex and Affects Embryo Sac Differentiation in Arabidopsis." Plant Cell 22(6): 1702-1715.

McCarty, D. R., T. Hattori, et al. (1991). "The Viviparous-1 developmental gene of maize encodes a novel transcriptional activator." Cell 66(5): 895-905.

Menges M, Dóczi R, et al. 2008: Comprehensive gene expression atlas for the Arabidopsis MAP kinase signalling pathways. New Phytol. 179(3):643-62.

Morais Cardoso, S., R. H. Swerdlow, et al. (2002). "Induction of cytochrome c-mediated apoptosis by amyloid beta 25-35 requires functional mitochondria." Brain Research 931(2): 117-125.

Mylne, J. S., L. Barrett, et al. (2006). "LHP1, the Arabidopsis homologue of HETEROCHROMATIN PROTEIN1, is required for epigenetic silencing of FLC." PNAS 103(13): 5012-5017.

Nawaschin S.G. "Resultate einer Revision der Befruchtungsvor-gaenge beiLilium MartagonundFritillaria tenella, ." Bul. Acad. Imp. Sci. St.Petersburg 9 1898 377–382.

Nurrish, S. J. and R. Treisman (1995). "Dna-Binding Specificity Determinants In Mads-Box Transcription Factors." Molecular and Cellular Biol 15(8): 4076-4085.

Okuda, S., H. Tsutsui, et al. (2009). "Defensin-like polypeptide LUREs are pollen tube attractants secreted from synergid cells." Nature 458(7236): 357-361.

Palevitz, B. A., and Cresti, Mauro. (1989). "Cytoskeletal changes during generative cell-division and sperm formation in tradescantia-virginiana." Protoplasma 150: 54-71.

Parenicova, L., S. de Folter, et al. (2003). "Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: New openings to the MADS world." Plant Cell 15(7): 1538-1551.

Pelaz, S., G. S. Ditta, et al. (2000). "B and C floral organ identity functions require SEPALLATA MADS-box genes." Nature 405(6783): 200-203.

Pnueli, L., D. Hareven, et al. (1994). "The TM5 MADS Box Gene Mediates Organ Differentiation in the Three Inner Whorls of Tomato Flowers." The Plant Cell Online 6(2): 175-186.

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

Rea, A. C. and J. B. Nasrallah (2008). "Self-incompatibility systems: barriers to self-fertilization in flowering plants." Int J Dev Biol 52(5-6): 627-636.

Reiser, L. and R. L. Fischer (1993). "The Ovule and the Embryo Sac." Plant Cell 5(10): 1291-1301.

Riechmann, J. L., M. Q. Wang, et al. (1996). "DNA-binding properties of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS." Nucleic Acids Research 24(16): 3134-3141.

Romanel, E. A. C., C. G. Schrago, et al. (2009). "Evolution of the B3 DNA Binding Superfamily: New Insights into REM Family Gene Diversification." PloS one 4(6).

Rotman, N., M. Gourgues, et al. (2008). “A dialogue between the SIRENE pathway in synergids and the fertilization independent seed pathway in the

central cell controls male gamete release during double fertilization in Arabidopsis." Mol. Plant 1: 659.

Rotman, N., F. Rozier, et al. (2003). "Female control of male gamete delivery during fertilization in Arabidopsis thaliana." Curr Biol 13(5): 432-436.

Sandaklie-Nikolova, L., R. Palanivelu, et al. (2007). "Synergid cell death in Arabidopsis is triggered following direct interaction with the pollen tube." Plant Physiol. 144(4):1753-62.

Schneitz K, H. M., Pruitt RE. (1995). "Wild-type ovule development in Arabidopsis thaliana: a light microscope study of cleared whole-mount tissue. ." Plant Journal 7: 731–749.

Schneitz, K., M. Hulskamp, et al. (1995). "Wild-Type Ovule Development In Arabidopsis-Thaliana - A Light-Microscope Study Of Cleared Whole-Mount Tissue." Plant Journal 7(5): 731-749.

Schwacke, R., S. Grallath, et al. (1999). "LeProT1, a transporter for proline, glycine betaine, and gamma-amino butyric acid in tomato pollen." Plant Cell 11(3): 377-392.

Schwarz-Sommer, Z., P. Huijser, et al. (1990). "Genetic Control of Flower Development by Homeotic Genes in Antirrhinum majus." Science 250(4983): 931-936.

Scott, R. J., Spielman, M., and Dickinson, Hugh G. (2004). "Stamen structure and function." Plant Cell 16: S46-S60.

Shi, D. and W. Yang (2011). "Ovule development in Arabidopsis: progress and challenge." Curr Opin Plant Biol 14(1): 74-80.

Shore, P. and A. D. Sharrocks (1995). "The Mads-Box Family Of Transcription Factors." European Journal of Biochemistry 229(1): 1-13.

Sommer, H., Beltr, et al. (1990). Deficiens, a homeotic gene involved in the control of flower morphogenesis in Antirrhinum majus: the protein shows homology to transcription factors. Embo Journal 9(3): 605–613.

Sung, S. and R. M. Amasino (2004). "Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3." Nature 427(6970): 159-164.

Suzuki, M., C. Y. Kao, et al. (1997). "The conserved B3 domain of VIVIPAROUS1 has a cooperative DNA binding activity." The Plant Cell Online 9(5): 799-807.

Swaminathan, K., K. Peterson, et al. (2008). The plant B3 superfamily. Trends Plant Sci. 13(12):647-55.

Theißen, G., J. Kim, et al. (1996). "Classification and phylogeny of the MADS-box multigene family suggest defined roles of MADS-box gene subfamilies in the morphological evolution of eukaryotes." Journal of Molecular Evolution 43(5): 484-516.

Toufighi K, Brady S, et al 2005 "The Botany Array Resource: e-Northerns, expression Angling, and promoter analyses." Plant Journal 43:153-163.

Tsukamoto, T. and R. Palanivelu (2010). “A role for LORELEI, a putative glycosylphosphatidylinositol-anchored protein, in Arabidopsis thaliana double fertilization and early seed development.”Plant Journal 62(4):571-88.

Twell, D., Park, S.K., and Lalanne, E. (1998). " Asymmetric division and cell-fate determination in developing pollen. ." Trends in Plant Science 3: 305-310.

Van Breusegem, F. and J. F. Dat (2006). Reactive oxygen species in plant cell death. Plant Physiol. 141: 384.

Waltner, J. K., F. C. Peterson, et al. (2005). "Structure of the B3 domain from Arabidopsis thaliana protein At1g16640." Protein Science 14(9): 2478-2483.

Wellmer, F., M. Alves-Ferreira, et al. (2006). "Genome-wide analysis of gene expression during early Arabidopsis flower development." Plos Genetics 2(7): 1012-1024.

Weterings, K. and S. D. Russell (2004). "Experimental Analysis of the Fertilization Process." The Plant Cell Online 16(suppl 1): S107-S118.

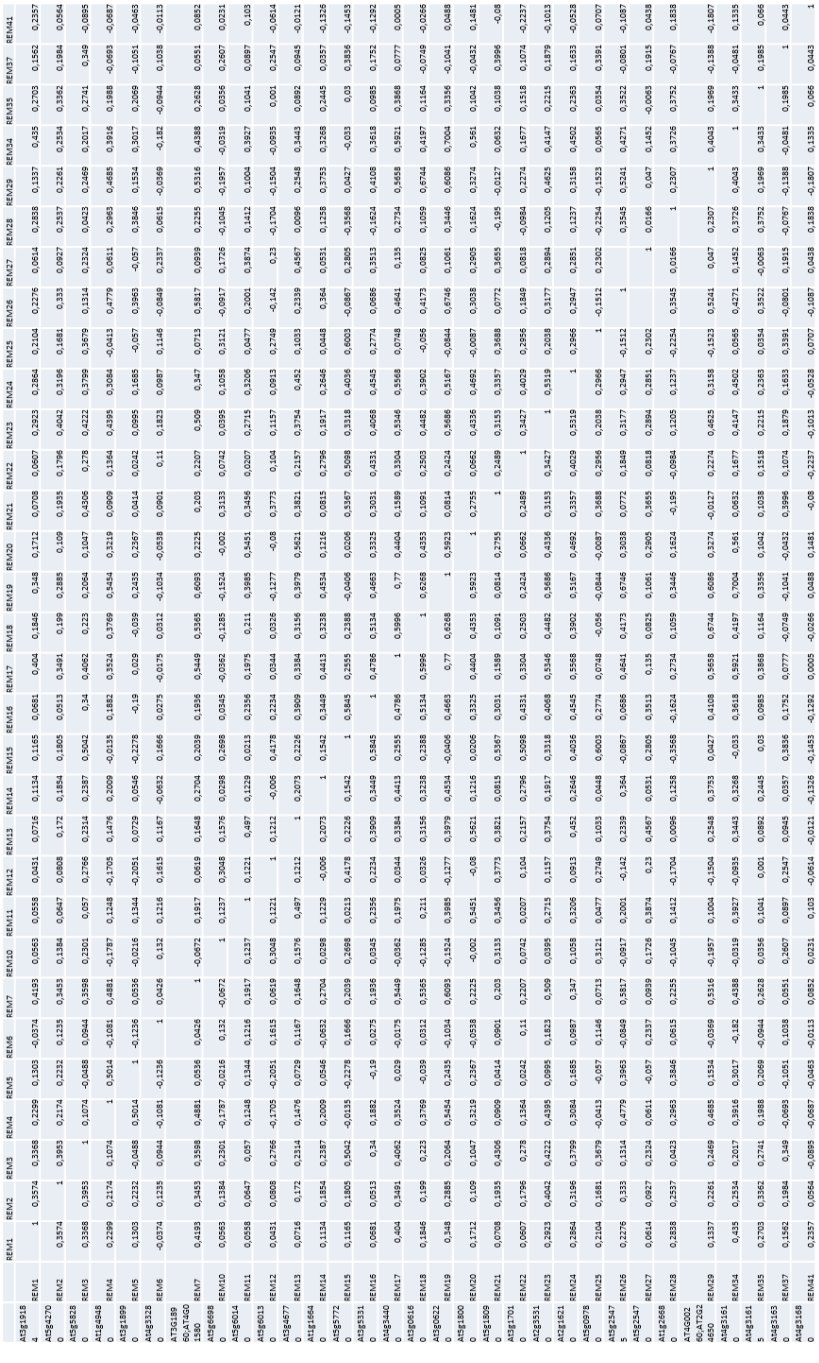
Yanofsky, M. F., H. Ma, et al. (1990). “The protein encoded by the *Arabidopsis* homeotic gene *agamous* resembles transcription factors.” Nature 346(1): 35.

Supplemental information

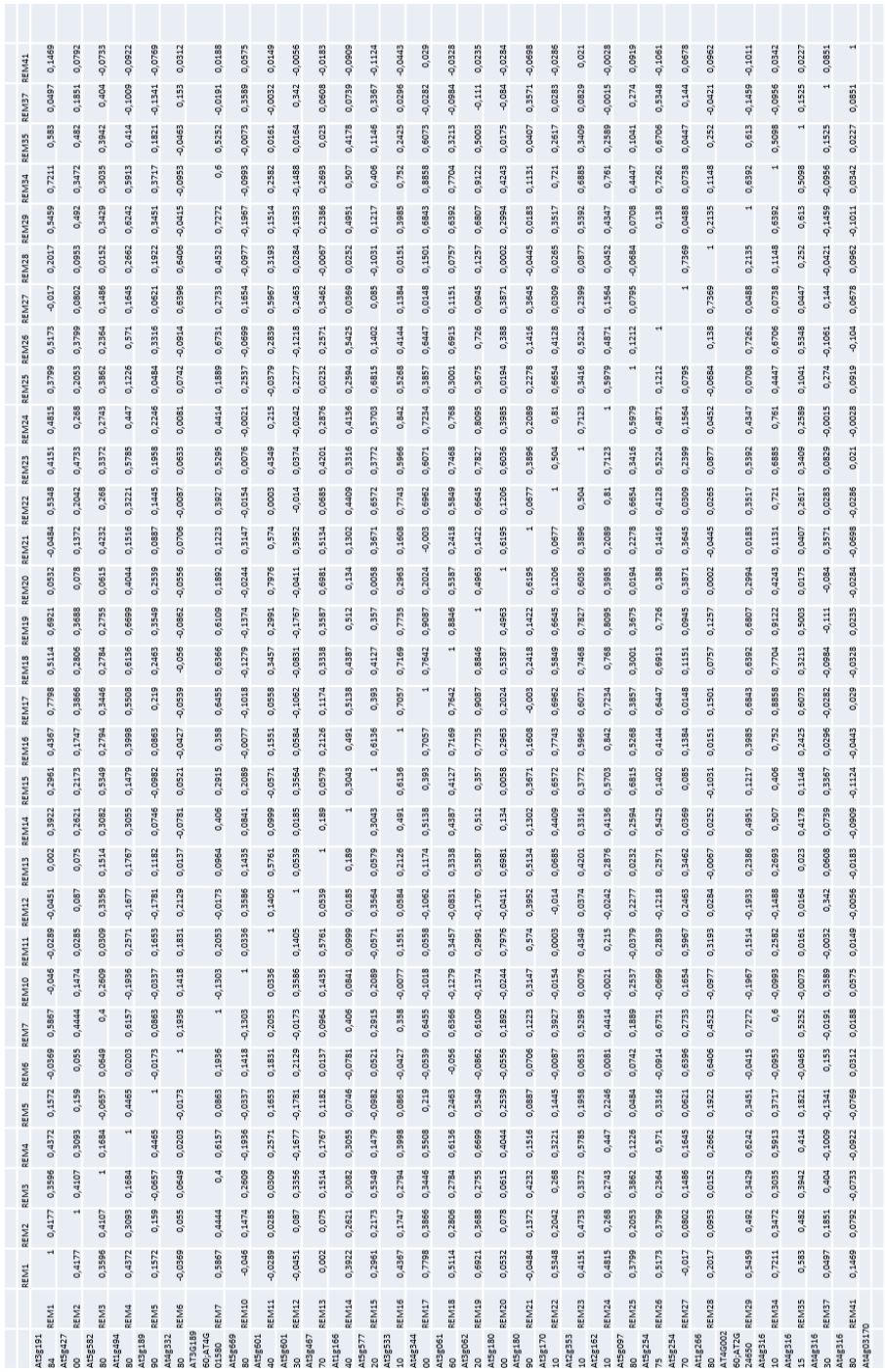
Supplemental table1. Correlation analysis of STK vs all genome, p(LIN) and p(LOG). person coef. > 0.6 .

255014_at	AT4G09960	MADS-box protein STK (AGL11)		
			Lin	Log
255014_at	AT4G09960	MADS-box protein (AGL11)	1	1
247666_at	AT5G60140	transcriptional factor B3 family protein, contains Pfam profile PF02362: B3 DNA binding domain, REM 11	0,9163	0,58
251555_at	AT3G58780	agamous-like MADS box protein AGL1 / shatterproof 1 (AGL1) (SHP1), identical to SP:P29381 Agamous-like MADS box protein AGL1 (Protein Shatterproof 1) {Arabidopsis thaliana}	0,8873	0,4815
263306_at	AT2G12480	serine carboxypeptidase S10 family protein, similar to Serine carboxypeptidase II precursor (Carboxypeptidase D) (CP-MII) (SP:P08818) (Hordeum vulgare)	0,8555	0,675
254418_at	AT4G21480	glucose transporter, putative, similar to glucose transporter (Sugar carrier) STP1, Arabidopsis thaliana, SP:P23586; contains Pfam profile PF00083: major facilitator superfamily protein	0,8555	0,4801
257089_at	AT3G20520	glycerophosphoryl diester phosphodiesterase family protein, contains Pfam PF03009: Glycerophosphoryl diester phosphodiesterase family; similar to glycerophosphodiester phosphodiesterase (GI:1399038) (Borrelia hermsii)	0,8358	0,5358
263988_at	AT2G42830	agamous-like MADS box protein AGL5 / floral homeodomain transcription factor (AGL5), identical to SP:P29385 Agamous-like MADS box protein AGL5 {Arabidopsis thaliana}	0,8175	0,607
250071_at	AT5G18000	transcriptional factor B3 family protein, contains Pfam profile PF02362: B3 DNA binding domain	0,8116	0,5102
250499_at	AT5G09730	glycosyl hydrolase family 3 protein, beta-xylosidase, Hypocrea jecorina, EMBL:Z69257	0,8062	0,5083
249241_at	AT5G42230	serine carboxypeptidase S10 family protein, similar to Serine carboxypeptidase II chains A and B (SP:P08819) (EC 3.4.16.6) (Triticum aestivum (Wheat))	0,8018	0,5594
259802_at	AT1G72260	thionin (THI2.1), identical to thionin (Arabidopsis thaliana) gi:1181531:gb:AAC41678	0,8004	0,4965
246025_at	AT5G21150	PAZ domain-containing protein / piwi domain-containing protein, similar to SP:O04379 Argonaute protein (AGO1) {Arabidopsis thaliana}; contains Pfam profile: PF02170 PAZ (Piwi Argonaut and Zwille), PF02171 Piwi domain	0,7784	0,5172
250360_at	AT5G11360	expressed protein	0,7778	0,481
254394_at	AT4G21630	subtilase family protein, contains similarity to subtilase; SP1 GI:9957714 from (Oryza sativa)	0,7702	0,4913
245640_at	AT1G25330	basic helix-loop-helix (bHLH) family protein, contains Pfam profile: PF00010 helix-loop-helix DNA-binding domain	0,7594	0,5046
256636_at	AT3G12000	S-locus related protein SLR1, putative (S1), identical to S-locus related protein SLR1 homolog (At S1) GI:246209 Arabidopsis thaliana); contains Pfam profiles PF01453: Lectin (probable mannose binding), PF00954: S-locus glycoprotein family	0,7589	0,5536
254777_at	AT4G12960	gamma interferon responsive lysosomal thiol reductase family protein / GILT family protein, similar to SP:P13284 Gamma-interferon inducible lysosomal thiol reductase precursor {Homo sapiens}; contains Pfam profile PF03227: Gamma interferon inducible lyso	0,7542	0,6419
249856_at	AT5G22980	serine carboxypeptidase III, putative, similar to serine carboxypeptidase III from Oryza sativa SP:P37891, Matricaria chamomilla GI:6960455, Hordeum vulgare SP:P21529, Triticum aestivum SP:P11515; contains Pfam profile PF0450 serine carboxypeptidase	0,7451	0,444
263204_at	AT1G05480	SNF2 domain-containing protein / helicase domain-containing protein, low similarity to SP:Q9U7E0 Transcriptional regulator ATRX homolog {Caenorhabditis elegans}; contains PFam profiles PF00271: Helicase conserved C-terminal domain, PF00176: SNF2 family N	0,739	0,3045
257629_at	AT3G26140	glycosyl hydrolase family 5 protein / cellulase family protein, contains Pfam profile: PF00150 cellulase (glycosyl hydrolase family 5)	0,7356	0,6582
259799_at	AT1G72290	trypsin and protease inhibitor family protein / Kunitz family protein, similar to water-soluble chlorophyll protein (Raphanus sativus var. niger) GI:16945735, BnD22 drought induced protein (Brassica napus) GI:17813; contains Pfam profile PF00197: Trypsin	0,735	0,5901
254786_at	AT4G12890	gamma interferon responsive lysosomal thiol reductase family protein / GILT family protein, similar to SP:P13284 Gamma-interferon inducible lysosomal thiol reductase precursor {Homo sapiens}; contains Pfam profile PF03227: Gamma interferon inducible lyso	0,7294	0,4625
253461_at	AT4G32170	cytochrome P450, putative, cytochrome p450, Arabidopsis thaliana, PID:G2252844	0,7273	0,3881
265020_at	AT1G24540	cytochrome P450, putative, similar to GB:AAB87111, similar to ESTs dbj:D41610, gb:T20562 and emb:Z26058	0,7206	0,5989
250364_at	AT5G11400	protein kinase-related, contains eukaryotic protein kinase domain, INTERPRO:IPR000719	0,7106	0,5157
249005_at	AT5G44630	Encodes a sesquiterpene synthase involved in generating all of the group B sesquiterpenes found in the Arabidopsis floral volatile blend. Strongly expressed in intrafloral nectaries.	0,695	0,4197
245389_at	AT4G17480	palmitoyl protein thioesterase family protein	0,6837	0,2814
255073_at	AT4G09090	glycosyl hydrolase family protein 17, similar to glucan endo-1,3-beta-glucosidase precursor SP:P52409 from (Triticum aestivum)	0,678	0,2401
259124_at	AT3G02310	developmental protein SEPALLATA2 / floral homeotic protein (AGL4) (SEP2), identical to developmental protein SEPALLATA2 / floral homeotic protein AGL4 GB:P29384 (Arabidopsis thaliana), Pfam HMM hit: SRF-type transcription factors (DNA-binding and dimeriz	0,6706	0,5648
252555_at	AT3G45940	alpha-xylosidase, putative, strong similarity to alpha-xylosidase precursor GI:4163997 from (Arabidopsis thaliana)	0,6589	0,291
246531_at	AT5G15800	developmental protein SEPALLATA1 / floral homeotic protein (AGL2) (SEP1), identical to developmental protein SEPALLATA1 / floral homeotic protein (AGL2 / SEP1) SP:P29382 from (Arabidopsis thaliana)	0,6583	0,5143
248593_at	AT5G49180	pectinesterase family protein, contains Pfam profile: PF01095 pectinesterase	0,652	0,6737
264499_at	AT1G30795	hydroxyproline-rich glycoprotein family protein, contains proline-rich extensin domains, INTERPRO:IPR000694	0,6426	0,5075
246939_at	AT5G25390	encodes a member of the ERF (ethylene response factor) subfamily B-6 of ERF/AP2 transcription factor family. The protein contains one AP2 domain. There are 12 members in this subfamily including RAP2.11.	0,6398	0,5031
251394_at	AT3G60900	fasciclin-like arabinogalactan-protein (FLA10)	0,6389	0,3034
251979_at	AT3G53140; AT3G53130	[AT3G53140, O-diphenol-O-methyl transferase, putative, similar to GI:6688808 (Medicago sativa subsp. x varia), caffeic acid O-methyltransferase (homt1), Populus kitakamiensis, EMBL:PKHOMT1A];[AT3G53130, cytochrome P450 family protein, similar to Cytoch	0,6334	0,4116
247796_at	AT5G58782	dehydrodolichyl diphosphate synthase, putative / DEDOL-PP synthase, putative, similar to GI:796076	0,6305	0,3032
259089_at	AT3G04960	expressed protein, low similarity to SP:P32380 NUF1 protein (Spindle poly body spacer protein SPC110) {Saccharomyces cerevisiae}	0,6298	0,4024
266169_at	AT2G38900	serine protease inhibitor, potato inhibitor I-type family protein, similar to SP:P24076 Glu S.griseus protease inhibitor (BGIA) {Momordica charantia}, SP:P20076 Ethylene-responsive proteinase inhibitor I precursor (Lycopersicon esculentum); contains Pfam	0,6295	0,4095
247035_at	AT5G67110	basic helix-loop-helix (bHLH) family protein, contains Pfam profile: PF00010 helix-loop-helix DNA-binding domain	0,6272	0,3945
262681_at	AT1G75890	family II extracellular lipase 2 (EXL2), EXL2 (PMID:11431566); similar to anter-specific proline-rich protein (APG) SP:P40602 (Arabidopsis thaliana)	0,6243	0,2256
249760_at	AT5G23960	Encodes a sesquiterpene synthase involved in generating all of the group A sesquiterpenes found in the Arabidopsis floral volatile blend. Strongly expressed in the stigma.	0,6243	0,4561
246526_at	AT5G15720	GDSL-motif lipase/hydrolase family protein, similar to family II lipase EXL3 (GI:15054386) and EXL2 (GI:15054384) (Arabidopsis thaliana); contains Pfam profile PF00657: Lipase/Acylhydrolase with GDSL-like motif	0,6239	0,4121
260990_at	AT1G12180	expressed protein	0,6238	0,3444
254391_at	AT4G21590	bifunctional nuclease, putative, similar to bifunctional nuclease (Zinnia elegans) gi:4099833/gb:AAD00694. Activated by AGAMOUS in a cal-1, ap1-1 background. Expressed in the floral meristem and during stamen development.	0,6209	0,1806
265878_at	AT2G42410	zinc finger (C2H2 type) family protein, contains Pfam profile: PF00096 zinc finger, C2H2 type	0,6174	0,4691
251025_at	AT5G02190	encodes an aspartic protease, has an important role in determining cell fate during embryonic development and in reproduction processes. The loss-of-function mutation of PCS1 causes degeneration of both male and female gametophytes and excessive cell dea	0,6118	0,3808
254336_at	AT4G22050	aspartyl protease family protein, contains Pfam profile: PF00026 eukaryotic aspartyl protease	0,6104	0,427
263203_at	AT1G05490	C protein immunoglobulin-A-binding beta antigen-related, contains weak similarity to C protein immunoglobulin-A-binding beta antigen (Streptococcus agalactiae) gi:18028989:gb:AAL56250	0,6064	0,3189
249497_at	AT5G39220	hydrolase, alpha/beta fold family protein, contains Pfam profile: PF00561 alpha/beta hydrolase fold	0,6055	0,57
264444_at	AT1G27360	squamosa promoter-binding protein-like 11 (SPL11), identical to squamosa promoter binding protein-like 11 (Arabidopsis thaliana) GI:5931665; contains Pfam profile PF03110: SBP domain	0,6032	0,2834
255644_at	AT4G00870	basic helix-loop-helix (bHLH) family protein, similar to the myc family of helix-loop-helix transcription factors; contains Pfam profile PF00010: Helix-loop-helix DNA-binding domain; PMID: 12679534	0,6032	0,4881
252306_at	AT3G49270	expressed protein	0,6006	0,6003

Supplemental table2. Correlation analysis REM vs REM using P(LOG) person coef. > 0.6 .



Supplemental table3. Correlation analysis REM vs REM using P(LIN)



	REM1	REM2	REM3	REM4	REM5	REM6	REM7	REM10	REM11	REM12	REM13	REM14	REM15	REM16	REM17	REM18	REM19	REM20	REM21	REM22	REM23	REM24	REM25	REM26	REM27	REM28	REM29	REM34	REM35	REM37	REM41
At5g191 84	REM1	1	0,4177	0,3596	0,4372	0,1572	-0,0369	0,5857	-0,046	-0,0289	-0,0451	0,002	0,3922	0,2961	0,4367	0,7798	0,5114	0,6921	0,0532	-0,0484	0,5348	0,4151	0,4815	0,3799	0,5173	-0,017	0,2017	0,5459	0,7211	0,583	0,1469
At5g427 00	REM2	0,4177	1	0,4107	0,3093	0,159	0,055	0,4444	0,1474	0,0285	0,087	0,075	0,2621	0,2173	0,1747	0,3866	0,2806	0,3688	0,078	0,1372	0,2042	0,4733	0,268	0,2053	0,3799	0,0802	0,0953	0,492	0,3472	0,482	0,1851	0,0792
At5g582 80	REM3	0,3596	0,4107	1	0,1684	-0,0657	0,0649	0,4	0,2609	0,0309	0,3356	0,1514	0,3082	0,5349	0,2794	0,3446	0,2784	0,2755	0,0615	0,4232	0,268	0,3372	0,2743	0,3862	0,2364	0,1486	0,0152	0,3429	0,3035	0,3842	0,404	-0,0733
At1g484 80	REM4	0,4372	0,3093	0,1684	1	0,4465	0,0203	0,6157	-0,1936	0,2571	-0,1677	0,1767	0,3055	0,1479	0,3998	0,5508	0,6136	0,6699	0,4004	0,1516	0,3221	0,5785	0,447	0,1226	0,571	0,1645	0,2662	0,6242	0,5913	0,414	-0,1009	-0,0922
At5g189 90	REM5	0,1572	0,159	-0,0657	0,4465	1	-0,0173	0,0863	-0,0337	0,1653	-0,1781	0,1182	0,0746	-0,0682	0,0863	0,219	0,2463	0,3549	0,2359	0,0887	0,1445	0,1958	0,2246	0,0484	0,3316	0,0621	0,1922	0,3451	0,3717	0,1821	-0,1341	-0,0769
At5g332 80	REM6	-0,0369	0,055	0,0649	0,0203	-0,0173	1	0,1936	0,1418	0,1831	0,2129	0,0157	-0,0781	0,0521	-0,0427	-0,0539	-0,056	-0,0862	-0,0556	0,0706	-0,0087	0,0833	0,0081	0,0742	-0,0914	0,5396	0,6406	-0,0415	-0,0953	0,153	0,0312
AT3G4189 0,5G,AT4G 01580	REM7	0,5867	0,4444	0,4	0,6157	0,0863	0,1936	1	-0,1303	0,2053	-0,0173	0,0964	0,406	0,2915	0,358	0,6455	0,6366	0,6109	0,1892	0,1223	0,3827	0,5285	0,4414	0,1889	0,6731	0,2733	0,4523	0,7272	0,6	0,5252	-0,0191	0,0188
At5g569 80	REM10	-0,046	0,1474	0,2609	-0,1936	-0,0337	0,1418	-0,1303	1	0,0336	0,3596	0,1435	0,0841	0,2089	-0,0077	-0,1018	-0,1279	-0,1374	-0,0244	0,3147	-0,0154	0,0076	-0,0021	0,2537	-0,0699	0,1654	-0,0977	-0,1957	-0,0993	-0,0073	0,3589	0,0575
At1g601 40	REM11	-0,0289	0,0285	0,0309	0,2571	0,1653	0,1831	0,2053	0,0336	1	0,1405	0,5761	0,0999	-0,0571	0,1551	0,0558	0,3457	0,2991	0,7976	0,574	0,0003	0,4349	0,215	-0,0379	0,2839	0,5967	0,3193	0,1514	0,2582	0,0161	-0,0032	0,0149
At5g601 30	REM12	-0,0451	0,087	0,3356	-0,1677	-0,1781	0,2129	-0,0173	0,3586	0,1405	1	0,0539	0,0185	0,3564	0,0584	-0,1062	-0,0831	-0,1767	-0,0411	0,3952	-0,014	0,0374	-0,0242	0,2277	-0,1218	0,2463	0,0284	-0,1933	-0,1488	0,0164	0,342	-0,0056
At5g467 70	REM13	0,002	0,075	0,1514	0,1767	0,1182	0,0137	0,0964	0,1435	0,5761	0,0539	1	0,189	0,0579	0,2126	0,1174	0,3338	0,3587	0,6981	0,5134	0,0685	0,4201	0,2876	0,0232	0,2571	0,3462	-0,0067	0,2386	0,2693	0,023	0,0608	-0,0183
At1g166 40	REM14	0,3922	0,2621	0,3082	0,3055	0,0746	-0,0781	0,406	0,0841	0,0999	0,0185	0,189	1	0,3043	0,491	0,5138	0,4387	0,512	0,134	0,1302	0,4409	0,3316	0,4136	0,2594	0,5425	0,0369	0,0252	0,4951	0,507	0,4178	0,0739	-0,0909
At5g577 20	REM15	0,2961	0,2173	0,5349	0,1479	-0,0982	0,0521	0,2915	0,2089	-0,0571	0,3564	0,0579	0,3043	1	0,6136	0,393	0,4127	0,357	0,0058	0,3671	0,6572	0,3772	0,5703	0,6815	0,1402	0,085	-0,1031	0,1217	0,406	0,1146	0,3367	-0,1124
At5g533 10	REM16	0,4367	0,1747	0,2794	0,3998	0,0863	-0,0427	0,358	-0,0077	0,1551	0,0584	0,2126	0,491	0,6136	1	0,7057	0,7169	0,7739	0,2963	0,1608	0,7743	0,5966	0,842	0,5268	0,4144	0,1384	0,0151	0,3985	0,752	0,2425	0,0296	-0,0443
At4g344 00	REM17	0,7798	0,3866	0,3446	0,5508	0,219	-0,0539	0,6455	-0,1018	0,0558	-0,1062	0,1174	0,5138	0,393	0,7057	1	0,7642	0,9087	0,2024	-0,003	0,6982	0,6071	0,7234	0,3857	0,6447	0,0148	0,1501	0,6843	0,8858	0,6073	-0,0282	0,029
At5g061 60	REM18	0,5114	0,2806	0,2784	0,6136	0,2463	-0,056	0,6366	-0,1279	0,3457	-0,0831	0,3338	0,4387	0,4127	0,7169	0,7642	1	0,8846	0,5387	0,2418	0,5849	0,7422	0,7468	0,3001	0,6913	0,1151	0,0757	0,6392	0,7704	0,3213	-0,0984	-0,0328
At5g062 20	REM19	0,6821	0,3688	0,2755	0,6699	0,3549	-0,0862	0,6109	-0,1374	0,2691	-0,1767	0,3587	0,512	0,357	0,9087	0,8846	1	0,4963	0,1422	0,6645	0,7827	0,8095	0,3985	0,3675	0,726	0,0945	0,1257	0,6807	0,8122	0,5003	-0,111	0,0235
At5g180 00	REM20	0,0532	0,078	0,0615	0,4044	0,2539	-0,0556	0,1892	-0,0244	0,7976	-0,0411	0,6981	0,134	0,0058	0,2963	0,2024	0,5387	0,4963	1	0,6195	0,1206	0,6036	0,3985	0,0194	0,388	0,3871	0,0002	0,2994	0,4243	0,0175	-0,084	-0,0284
At5g180 90	REM21	-0,0484	0,1372	0,4232	0,1516	0,0887	0,0706	0,1223	0,3147	0,574	0,3952	0,5134	0,1302	0,3671	0,1608	-0,003	0,2418	0,1422	0,6195	1	0,0677	0,3896	0,2089	0,2278	0,1416	0,3645	-0,0445	0,0183	0,1131	0,0407	0,3571	-0,0698
At1g170 10	REM22	0,5348	0,2042	0,268	0,3221	0,1445	-0,0087	0,3927	-0,0154	0,0003	-0,014	0,0685	0,4409	0,6572	0,7743	0,6962	0,5849	0,6645	0,1206	0,0677	1	0,504	0,81	0,6654	0,4128	0,0309	0,0265	0,3517	0,721	0,2617	0,0283	-0,0286
At1g353 10	REM23	0,4151	0,4733	0,3372	0,5765	0,1958	0,0633	0,5295	0,0076	0,4349	0,0374	0,4201	0,3316	0,3772	0,5966	0,6071	0,7468	0,7827	0,6036	0,3866	0,504	1	0,7123	0,3416	0,5224	0,2399	0,0877	0,5392	0,6885	0,3409	0,0829	0,021
At2g162 10	REM24	0,4815	0,268	0,2743	0,447	0,2246	0,0081	0,4414	-0,0021	0,215	-0,0242	0,2876	0,4136	0,5703	0,842	0,7234	0,768	0,8099	0,3985	0,2089	0,81	0,7123	1	0,5979	0,4871	0,1564	0,0452	0,4347	0,761	0,2589	-0,0015	-0,0028
At5g097 80	REM25	0,5799	0,2053	0,3862	0,1226	0,0484	0,0742	0,1889	0,2537	-0,0379	0,2277	0,0232	0,2594	0,6815	0,9268	0,3857	0,9001	0,3675	0,0194	0,2278	0,6854	0,3416	0,5979	1	0,1212	0,0795	-0,0684	0,0708	0,4447	0,1041	0,274	0,0919
At4g254 70	REM26	0,5173	0,3799	0,2364	0,571	0,3316	-0,0914	0,6731	-0,0699	0,2839	-0,1218	0,2571	0,5425	0,1402	0,4144	0,6447	0,6913	0,726	0,388	0,1416	0,4128	0,5224	0,4871	0,1212	1			0,138	0,7262	0,6706	0,5348	-0,1061
At4g254 70	REM27	-0,017	0,0802	0,1486	0,1645	0,0621	0,6396	0,2733	0,1654	0,5967	0,2463	0,3462	0,0369	0,085	0,1384	0,0148	0,1151	0,0845	0,3871	0,3645	0,0309	0,2399	0,1564	0,0795		1	0,7369	0,0488	0,0738	0,0447	0,144	0,6678
At2g266 80	REM28	0,2017	0,0953	0,0152	0,2662	0,1922	0,6406	0,4523	-0,0977	0,3193	0,0284	-0,0067	0,0252	-0,1031	0,0151	0,1501	0,0757	0,1257	0,0002	-0,0445	0,0265	0,0877	0,0452	-0,0684	0,138	0,7369	1	0,2135	0,1148	0,252	-0,0421	0,0962
AT4G002 60,AT2G 24650	REM29	0,5459	0,492	0,3429	0,6242	0,3451	-0,0415	0,7272	-0,1967	0,1514	-0,1933	0,2386	0,4951	0,1217	0,3985	0,6843	0,6392	0,6807	0,2994	0,0183	0,3517	0,5392	0,4347	0,0708	0,7262	0,0488	0,2135	1	0,6392	0,613	-0,1459	-0,1011
At4g316 10	REM34	0,7211	0,3472	0,3035	0,5913	0,3717	-0,0953	0,6	-0,0993	0,2582	-0,1488	0,2693	0,507	0,406	0,752	0,8858	0,7704	0,9122	0,4243	0,1131	0,721	0,6885	0,761	0,4447	0,6706	0,0738	0,1148	0,6392	1	0,5098	-0,0956	0,0342
At4g316 15	REM35	0,583	0,482	0,3842	0,414	0,1821	-0,0463	0,5252	-0,0073	0,0161	0,0164	0,023	0,4178	0,1146	0,2425	0,8073	0,3213	0,5003	0,0175	0,0407	0,2817	0,3409	0,2589	0,1041	0,5348	0,0447	0,252	0,613	0,5098	1	0,1525	0,0227
At4g316 30	REM37	0,0497	0,1851	0,404	-0,1009	-0,1341	0,153	-0,0191	0,3589	-0,0032	0,342	0,0608	0,0739	0,3567	0,0296	-0,0282	-0,0884	-0,111	-0,084	0,3571	0,0283	0,0829	-0,0015	0,274	-0,1061	0,144	-0,0421	-0,1459	-0,0956	0,1525	1	0,0851
At4g316 80	REM41	0,1469	0,0792	-0,0733	-0,0922	-0,0769	0,0312	0,0188	0,0575	0,0149	-0,0056	-0,0183	-0,0809	-0,1124	-0,0443	0,029	-0,0328	0,0235	-0,084	-0,0958	-0,0286	0,021	-0,0028	0,0919	-0,104	0,0678	0,0962	-0,1011	0,0342	0,0227	0,0851	1
At4g03170																																

Part II (submitted and accepted)

MADS-domain Transcription Factors Mediated Short-Range DNA Looping is Essential for Target Gene Expression in Arabidopsis

Submitted to *Plant Cell*

The MADS box SEEDSTICK and ARABIDOPSIS Bsister play a maternal role in fertelization and seed development (2012).

Published in *the plant journal*, 70(3):409-20.

MADS-domain Transcription Factors Mediated Short-Range DNA Looping is Essential for Target Gene Expression in Arabidopsis

Marta Adelina Mendes[a,2], Rosalinda Fiorella Guerra[a,2], Markus Christian Berns[b], Carlo Manzo[c], Simona Masiero [a], Laura Finzi[c], Martin M. Kater[a,1] and Lucia Colombo [a,d,1]

[a]Dipartimento di BioScienze, Università degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy

[b]Max Planck Institute for Plant Breeding Research, D–50829 Cologne, Germany

[c]Department of Cell Biology, Emory University, 400 Dowman Dr, Atlanta, GA 30322, USA

[d]CNR, Istituto di Biofisica, Università di Milano, Via Celoria 26, 20133 Milano, Italy

[1]To whom correspondence may be addressed. Email: lucia.colombo@unimi.it or martin.kater@unimi.it

[2]These authors contributed equally to this work

Running title: Target gene regulation by MADS proteins

Estimated length (using Plant Cell page calculator): 9.6 pages

The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is Lucia Colombo (lucia.colombo@unimi.it).

Abstract

MADS-domain transcription factors are key regulators of development in eukaryotes. In plants the homeotic MIKC MADS-factors that control floral organ identity have been studied in great detail. Based on genetic and protein-protein interaction studies, a “floral-quartet model” was proposed that describes how these MADS-domain proteins assemble into higher order complexes to regulate their target genes. However, despite the attractiveness of this model and its general acceptance in the literature solid in vivo proof for this model has never been provided. To provide deeper insight in the mechanisms of transcriptional regulation by MADS-domain factors we studied how SEEDSTICK (STK) and SEPALLATA3 (SEP3) directly regulate the expression of the REM transcription factor-encoding gene *VERDANDI* (*VDD*). Our data show that STK-SEP3 dimers can induce loop formation in the *VDD* promoter by binding to two nearby CArG-boxes and that this is essential for promoter activity. Our in vivo data show that the size and position of this loop, determined by the choice of CArG element usage, is essential for correct expression. Our studies provide solid in vivo evidence for the floral-quartet model.

Introduction

MADS-box genes encode transcriptional regulators involved in diverse and important biological functions. They have been identified in yeast, insects, nematodes, lower vertebrates, mammals and plants. These transcription factors contain a conserved DNA binding and dimerization domain named the MADS-domain (Schwarzsommer et al. 1992). In plants MADS-box genes have been highly amplified during evolution, for instance in Arabidopsis 107 MADS-box genes and in rice 75 MADS-domain encoding genes have been identified (Arora et al. 2007; Parenicova et al. 2003).

The ability of MADS-domain proteins to bind DNA as dimers is reflected by the dyad symmetry of their binding sites that are found within promoter and enhancer sequences (Shore and Sharrocks 1995). Nurrish and Treisman (1995) studied MADS-domain protein binding sites and showed that they bind to the consensus sequence "$CC(A/T)_6GG$" named CArG box. Evidence based on in vitro biochemical assays and interaction studies in yeast showed that plant MADS-domain proteins form mainly heterodimers which are thought to assemble into multimeric complexes (de Folter et al. 2005; Egea-Cortines et al. 1999; Honma and Goto 2001; Riechmann et al. 1996). Many of these studies have been done using MADS-domain factors that control floral organ identity in Arabidopsis, which modes of action have been described in the combinatorial genetic ABC model (Coen and Meyerowitz 1991). Importantly, the ABC MADS-domain factors are for their function dependent on another group of MADS-domain transcription factors indicated as "class E factors", which are encoded by four largely redundant *SEPALLATA* genes (*SEP1-4*) (Ditta et al. 2004; Pelaz et al. 2000). Class E factors establish interactions between A, B and C class factors and their combined ectopic expression (A, B and E or B, C and E) resulted in the homeotic conversion of leaves into petals or stamens (Honma and Goto 2001; Pelaz et al. 2001). These

studies resulted into the formulation of a 'floral quartet model' which suggests that the MADS-domain proteins form higher order (quartet) complexes to establish floral organ identity (Theissen and Saedler 2001).

Similar results were obtained for the factors that control ovule development in Arabidopsis. The three MADS-box genes *SEEDSTICK (STK), SHATTERPROOF1 (SHP1)* and *SHP2* redundantly control ovule identity, since in the *stk shp1 shp2* triple mutant ovules are converted into carpel-like structures (Pinyopich et al. 2003). Interestingly, the *SEP1/sep1 sep2 sep3* triple mutant (only one allele of *SEP1* is active) phenocopied the *stk shp1 shp2* triple mutant showing that the SEP proteins are also important for the development of ovules (Favaro et al. 2003). The role of SEP proteins in the formation of ovules is likely to favour the formation of active complexes since yeast three-hybrid studies showed that SEP3 was able to bridge interactions among STK, SHP1 and SHP2.

Recently, we have identified *VERDANDI* (*VDD*), a gene belonging to the *REPRODUCTIVE MERISTEM* (*REM*) family (Romanel et al. 2009), as a target of the ovule identity factors STK, SHP1, SHP2 and SEP3 (Matias-Hernandez et al. 2010). *VDD* transcripts are present in the same tissues as these ovule identity genes and silencing of the ovule identity genes, *STK*, *SHP1* and *SHP2*, led to the complete absence of *VDD* expression during ovule development. Analysis of the *VDD* mutant revealed that this gene is important for female gametophyte cell identity determination (Matias-Hernandez et al. 2010).

Studies demonstrated that MADS-domain protein complexes often interact with DNA by contacting multiple nearby CArG box sequences, separated by less than 300 base pairs (Egea-Cortines et al. 1999; Liu et al. 2008). In the regulatory region of *VDD* three CArG boxes were identified within a region of 500 bp and by ChIP analysis it was shown that the first and third box were bound by both STK and SEP3 (Matias-Hernandez et al. 2010).

Here we describe the use of a combination of biophysical, molecular and in vivo approaches to study the regulation of *VDD* promoter activity by MADS-domain ovule identity factors. In particular, we characterized in vitro and in vivo the interactions of STK and SEP3 with the three CArG boxes and investigated the importance of these interactions for the expression of *VDD*. Our study provides a deeper insight in the mode of action of MADS-domain proteins in the regulation of their target genes.

Results

SEP3 and STK together mediate DNA looping in the *VDD* promoter region

STK and SEP3 have shown to form dimers that probably form tetrameric complexes (Favaro et al. 2003; Melzer et al. 2008). Furthermore, they are regulating the expression of *VDD* through direct binding to its promoter region (Favaro et al. 2003; Matias-Hernandez et al. 2010). The *VDD* promoter region contains three CArG boxes within 1000 bp upstream of the ATG start codon (Matias-Hernandez et al. 2010). Cooperative binding of the tetramers (composed of two SEP3-STK heterodimers) to two of the three adjacent CArG boxes would induce the formation of loops within the promoter region, which might have important regulatory functions. To investigate whether SEP3 and STK are indeed able to mediate interactions between elements in the *VDD* promoter region, Tethered Particle Motion (TPM) analysis (Dunlap et al. 2011; Finzi and Dunlap 2003; Nelson et al. 2006; Pouget et al. 2004) was performed using a *VDD* promoter fragment of 697 bp containing all three CArG boxes in the same arrangement found in vivo. TPM is a powerful, single-molecule technique, which is particularly appropriate to monitor protein-induced DNA conformational changes such as

looping, bending, large-scale compaction (Finzi and Gelles 1995; Guerra et al. 2007; Zaremba et al. 2010; Zurla et al. 2009).

In principle, binding to CArG boxes and STK-SEP3 protein-protein cooperative interactions could result in three possible loops: between CArG box 1 and 2, between CArG box 2 and 3, or between CArG box 1 and 3 (Supplemental Figure S1A).

To facilitate the correct interpretation of the TPM data, we first made a calibration curve using DNA tethers that have lengths that are predicted to be similar to each of the possible looped *VDD* promoter fragments (Supplemental Figure S1B). Therefore, we made tethers 243, 355 and 575 bp long. After analysing 20 tethers for each DNA fragment, including the 697 bp fragment, we fitted the cumulative frequency distribution of the data with a Gaussian curve (Supplemental Figure S1C). The centre of the peak of each Gaussian curve indicates the average $\sqrt{\langle \rho^2(t) \rangle_{4s}}$ value (TPM signal) for each DNA length. The four values were then plotted as a function of DNA length and these data were well in agreement with a published calibration curve obtained by Nelson et al. (2006) (Supplemental Figure S1D).

Subsequently, the effects of STK, SEP3 and STK-SEP3 heterodimers binding to the 697bp *VDD* promoter tether were studied by TPM (Figure 1). Furthermore, we also tested STK and SEP3 binding to tethers in which one or all CArG boxes were deleted.

When only STK was added to the tether (Figure 1B) no loop formation was observed. However, the unlooped tether was shorter than when no protein was added (compare the position of the curve with the calibration curve C shown at the top of the panel). This suggests that the addition of the STK protein resulted in a shortening of the tether. This seems to be unrelated to binding of the MADS-domain proteins to the CArG boxes since this shortening of the tether was also observed when we used a tether without CArG boxes (Figure 1D). When only

SEP3 was added to the tether we obtained a more complex pattern. Important to notice is that we observed a curve (blue) that arose as a shoulder of the free non-looped magenta Gaussian on the right. However, this shoulder was observed in all the experiments and is clearly a background effect of SEP3 in these TPM experiments. The green very small curve is difficult to explain but might be due to loop formation between CArG box1 and 2.

Clearer are the results when both STK and SEP3 were added to the tether. The non-specific blue curve caused by SEP3 is much less pronounced in these experiments, which is probably due to the fact that SEP3 interacts with STK and this heterodimer does not seem to cause this non-specific curve. The control experiments using the tether without CArG boxes (Figure 1D) showed that no loop formation is possible without CArG boxes. In Figure 1E is shown that two different loops were obtained when STK and SEP3 were added to the wild type tether. Our interpretation is that the most left red curve is due to loop formation between CArG box1 and 3 whereas the green curve might be due to loop formation between CArG box1 and 2. The experiments with the tethers that contain a single CArG box deletion (Figure 1F-H) showed that when CArG box1 or 3 were deleted no loop formation is possible. Only when CArG box2 was deleted the loop between CArG box1 and 3 was obtained. Interestingly, the putative loop between CArG box1 and 2 as observed using the wild-type tether (Figure 1E; green curve) was only established when CArG box1 and 3 were both present suggesting that in these in vitro TPM experiments the binding to CArG box1 and 3 might somehow facilitated loop formation between CArG1 and 2. In conclusion these experiments suggest that CArG box1 and 3 are the boxes in the *VDD* promoter that are mainly involved in the formation of loops induced by STK and SEP3.

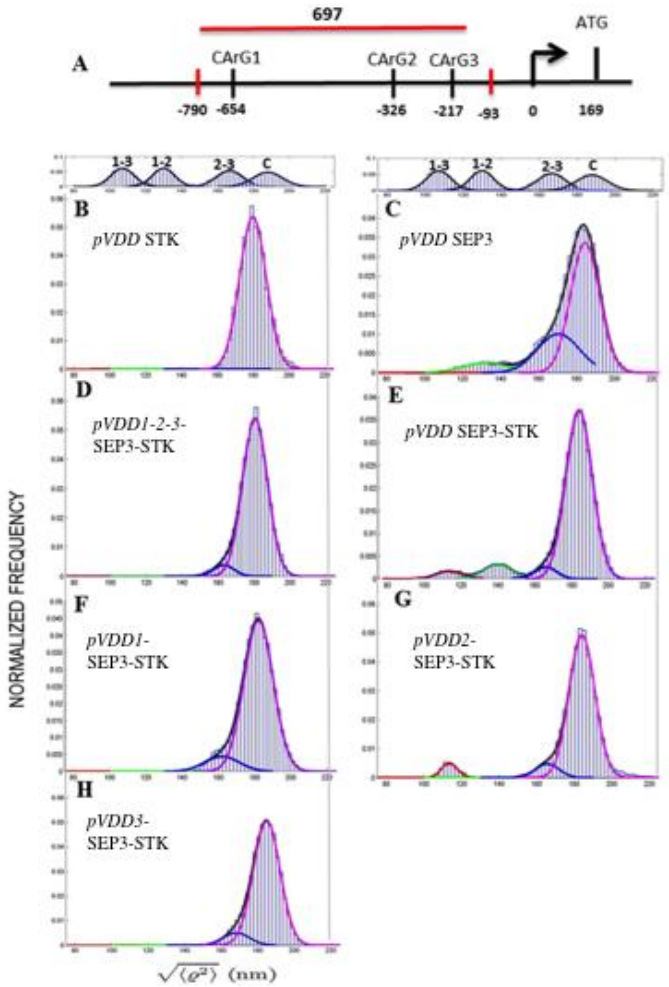


Figure 1. TPM analysis of STK and SEP3 interactions with the *VDD* promoter

(A) Schematic representation of the *VDD* promoter. Positions of the CArG boxes are relative to the translation start site. (B) *pVDD* in the presence of STK protein (C) *pVDD* in the presence of SEP3 (D) Negative control of TPM experiment, *pVDD* with the CArG boxes deleted (*pVDD-1-2-3*) in presence of both proteins STK and SEP3 (E) *pVDD* in the presence of both proteins STK and SEP3 (F) *pVDD* with the first CArG box deleted (G) with the second CArG deleted and (H) with the third CArG, deleted in presence of both proteins.

The histograms were normalized to the total number of events and to the bin width (2 nm).

The role of the CArG boxes in the regulation of *VDD* expression

The TPM analysis showed that regulators of *VDD* induced loops into the putative promoter region by binding to two CArG boxes. To investigate the importance of the CArG boxes for the expression of *VDD* we performed promoter analysis in which we mutated single or combinations of the three CArG boxes, changing the $[A/T]_6$ into $[G/C]_6$.

In order to validate the reporter gene expression profiles we first performed *VDD* in situ hybridization expression analysis. This showed that *VDD* transcripts were first detected at stage 2-I of ovule development (Schneitz et al. 1995)Figure 2A). During subsequent stages of ovule development (until stage 3-VI) *VDD* expression was observed throughout all tissues of the ovules (Figure 2B-D). After fertilization a strong *VDD* hybridization signal was observed in embryos at the globular stage, but at heart stage embryos *VDD* expression almost disappeared (Figure 3A-B).

To evaluate the importance of the CArG boxes for controlling *VDD* expression a putative promoter fragment of 1221 bp upstream of the *VDD* translation start site was cloned in frame with the *uidA* reporter gene that encodes for β-glucuronidase (GUS). This *pVDD::GUS* construct was used for Arabidopsis transformation. We generated more than 80 transgenic lines for this construct and 92% of these plants showed similar expression profiles whereas 8% did not show GUS activity. The GUS expression profile during ovule development perfectly matched the *VDD* expression that was observed by in situ hybridization experiments (Figure 2E-H). In globular stage embryos GUS expression was observed whereas at heart stage no GUS activity could be detected (Figure 3C-D). The in situ profiles confirm the expression in globular stage embryos however at heart stage they showed some residual *VDD* expression (Figure 3A-B).

Since the *pVDD::GUS* reporter construct drives GUS expression similar to the endogenous *VDD* gene, we used this *VDD* promoter fragment to generate new *uidA*

reporter gene constructs in which single or combinations of the three CArG boxes were mutated. These constructs were all used to transform Arabidopsis plants and at least 80 independent transgenic plants were obtained for each construct. In plants that contained the reporter construct with a single mutated CArG box, expression profiles changed depending on which CArG box was mutated. Mutations in the second CArG (*pVDDm2::GUS*) box did not affect the expression profile of the reporter gene (Figure 2M-P). However, when the first CArG was mutated the expression of the reporter gene (*pVDDm1::GUS)* was only detected in developing stage 3-VI ovules (Figure 2I-L). When CArG box3 was mutated (*pVDDm3::GUS*) GUS expression was visible at stage 2-I, restricted to the chalaza zone (Figure 2Q) of the ovule and expression levels at later stages were lower than in wild-type (Figure 2R-T).

We also analyzed reporter constructs in which two or all three CArG boxes were mutated (*pVDDm1-2::GUS; pVDDm1-3::GUS; pVDDm2-3::GUS; pVDDm1-2-3::GUS*). In all these transgenic plants no GUS expression was observed showing that the presence of two CArG boxes are essential for *VDD* promoter activity. An example of the obtained results is shown for *pVDDm1-2::GUS* in Figure 2U-X.

These experiments were all done with mutated CArG boxes. However in the TPM analyses described above we used promoter fragments in which CArG boxes were deleted. To verify the consistency of these data we also prepared reporter gene constructs in which CArG boxes were completely deleted as described in the TPM experiments. This showed that exactly the same results were obtained as when mutated CArG boxes were used (Supplemental Figure S2).

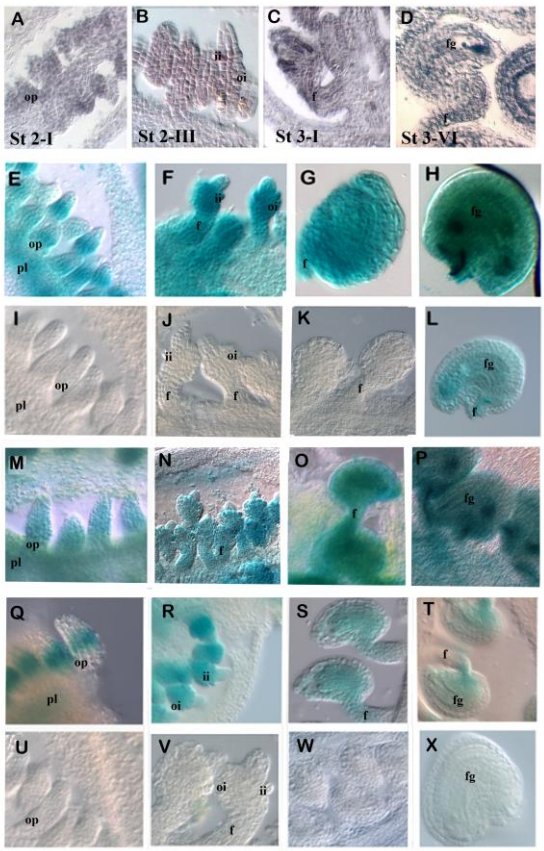


Figure 2. *VDD* expression and promoter analysis during ovule development

(A-D) In situ hybridization analysis of *VDD* during ovule development. (A) ovule development stage 2-I; (B) stage 2-III; (C) 3-I and (D) stage 3-VI (these stages should be used as reference for the next lines). (E-H) *pVDD::GUS* transgenic plants showed a similar expression pattern as observed by the in situ hybrization experiment. (I-L) GUS expression in ovules of *pVDDm1::GUS* lines. (M-P) GUS expression in ovules of *pVDDm2::GUS* lines. (Q-T) GUS expression in ovules of *pVDDm2::GUS* lines. (U-X) Absence of GUS expression as observed in the *pVDDm1-2::GUS.* Absence of GUS expression was also observed in *pVDDm1-3::GUS; pVDDm2-3::GUS; pVDDm1-2-3::GUS* lines. pl-placenta; op- ovule primordium; f-funiculus; ii- inner integument; oi- outer integument; fg- female gametophyte

Interestingly, during seed development the importance of the CArG boxes for *VDD* expression was shown to be different. Whereas during ovule development there seems to be flexibility in the use of the CArG boxes, in seeds this showed to be different. Only inactivation of CArG box2 did not result in a complete loss of *VDD* promoter activity during seed development (Figure 3E-F), whereas inactivation of CArG box1, CArG box3, and all other combinations did eliminate GUS expression during seed development (Figure 3G-H), showing that the presence of both CArG box1 and 3 are critical for correct *VDD* expression in developing seeds.

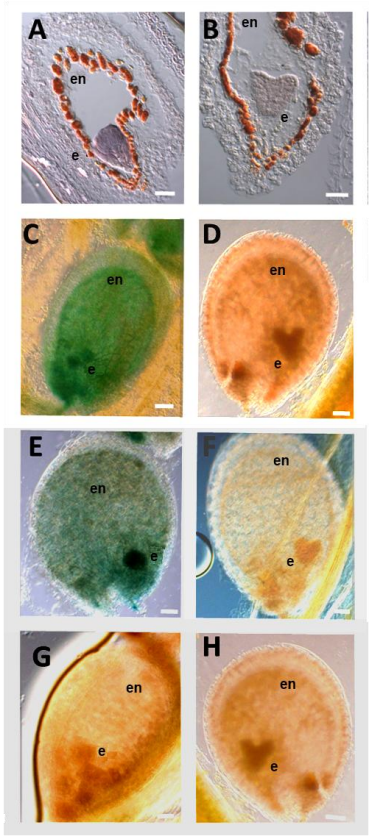


Figure 3. *VDD* expression and promoter analysis during seed development

(A-B) *VDD* in situ hybridization analysis in developing seeds with globular (A) and heart stage (B) embryos. (C-D) GUS expression in seeds of *pVDD::GUS* lines. (E-F) GUS expression in seeds of *pVDDm2::GUS* lines. (G-H) Absence of GUS expression as observed in *pVDDm1::GUS.* Absence of GUS expression was also observed in plants containing the following constructs *pVDDm3::GUS; pVDDm1-2::GUS; pVDDm1-3::GUS; pVDDm2-3::GUS; pVDDm1-2-3::GUS*..

e- embryo; en- endothelium

In vivo Binding of STK and SEP3 to the Three CArG Boxes in the *VDD* Promoter

Previously published ChIP data showed that CArG box1 and CArG box3 in the *VDD* promoter region are directly bound by SEP3 and STK whereas no binding to CArG box2 was observed (Matias-Hernandez et al. 2010). Repeating this experiment resulted in exactly the same observation (Figure 4A and B). Subsequently, we performed ChIP assays combined with real-time PCR analysis using chromatin extracted from unfertilized flowers from reporter lines that contain *VDD* promoter constructs with one of the three CArG boxes mutated. Specific primers for the mutated CArG boxes were used in order to discriminate binding to the endogenous promoter from binding to the exogenous DNA constructs. These experiments showed that when CArG box1 was mutated (*pVDDm1::GUS*) CArG box2 was used by STK and SEP3 (Figure 4C and D). A similar result was obtained when we performed ChIP analysis using chromatin extracted from inflorescences of the *pVDDm3::GUS* reporter line, in this case CArG box1 and 2 were bound by STK and SEP3 (Figure 4G and H). As expected when CArG box2 was mutated the MADS-domain factors bound normally to CArG box1 and 3 (Figure 4E and F).

Interestingly, in plants containing the reporter line in which CArG box1 and 3 were mutated (*pVDDm1-3::GUS*) no enrichment was observed on any of the CArG boxes suggesting that binding to CArG box1 or 3 facilitates binding of the SEP3-STK dimer to CArG box2. We also performed ChIP experiments using plants containing the reporter lines *pVDDm1-2::GUS* and *pVDDm2-3::GUS.* Also these experiments showed that the single non-mutated CArG box is never enriched. These experiments evidence that in these floral tissues, SEP3 and STK were only able to bind the *VDD* promoter when two CArG boxes were available (Supplemental Figure S3 A-F).

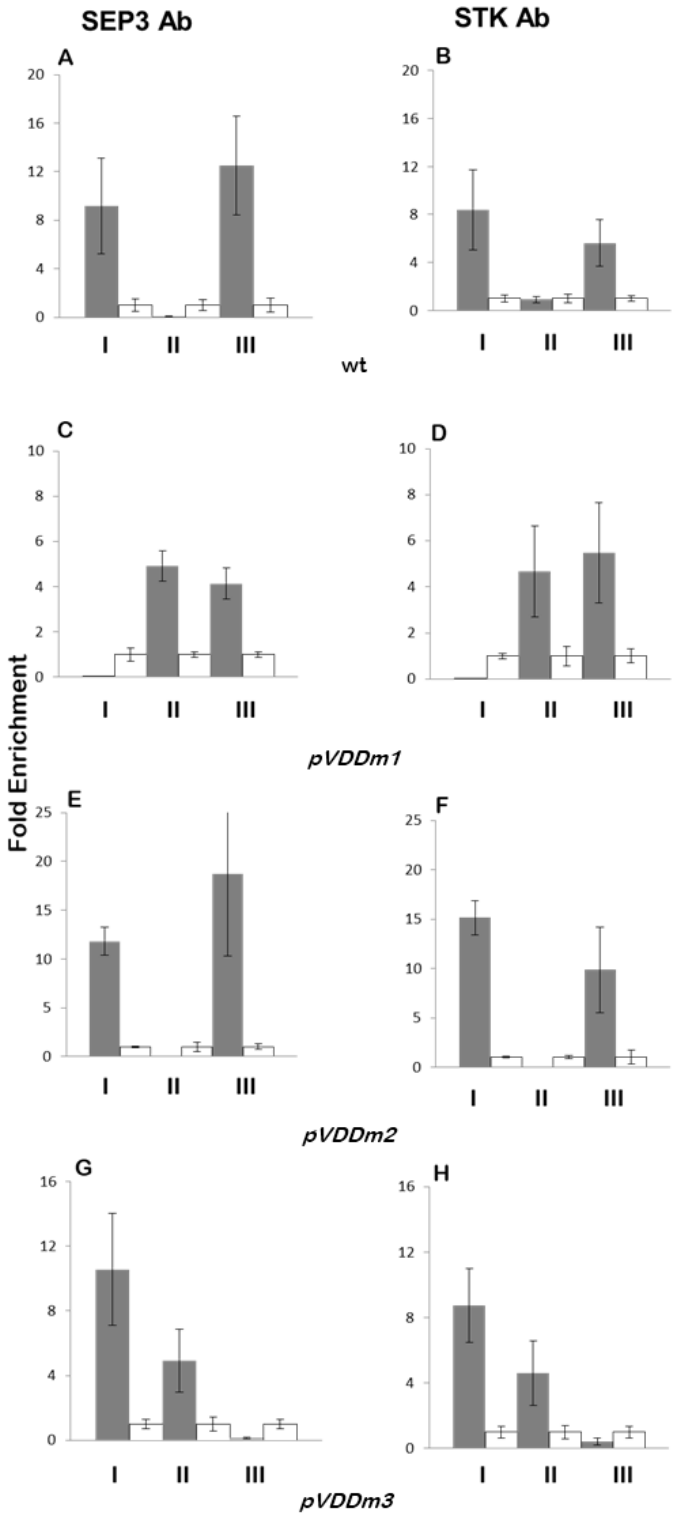


Figure 4. In vivo binding of SEP3 and STK to CArG boxes in the *VDD* promoter region

ChIP experiments using SEP3 (panel A, C, and G) and STK (panel B, D, F and H) antibodies to investigate binding to the CArG boxes (I, CArG-box1; II, CArG-box2; III, CArG-box3) in the *VDD* promoter. Negative controls (white bars) for SEP3 ChIP assays were done using wild-type leaf tissues and for STK ChIP negative controls flowers of the *stk* mutant were used. (A and B) SEP and STK binding to the endogenous VDD promoter. (C-H) ChIP assays to test SEP3 and STK binding to the CArG boxes in the heterologous *VDD* promoter of the (C and D) *pVDDm1::GUS* construct containing lines, (E and F)

pVDDm2::GUS construct containing lines and (G and H) *pVDDm3::GUS* construct containing lines. Fold enrichments were calculated over the negative controls. Error bars represent the propagated error value using three replicates.

Conservation of the *VDD* CArG boxes in species related to *Arabidopsis thaliana*

There are three CArG boxes in the *VDD* promoter region but only CArG box1 and 3 seem to be important for proper *VDD* expression in ovules and seeds. The question therefore arose if there might be conservation of all three boxes or just two of them. We investigated by a shadowing approach, using orthologous promoters of *Arabidopsis lyrata, Arabis alpina, Brassica rapa, Capsella rubella* and *Thellungiella halophile,* if there is conservation of the position of all three CArG boxes in these species (Figure 5). This analysis showed that in *Arabidopsis lyrata* all three CArG boxes are located in the same position suggesting that in the genus Arabidopsis the regulatory mechanism to control *VDD* expression is probably conserved.

CArG box3 is located in a highly conserved region in 6 species analysed whereas CArG box1 and 2 are not (Figure 5A). The conservation in CArG box sequences also confirms this, since CArG box3 is the only one that is most conserved in sequence between these species (Figure 5B). However, if we strictly consider a consensus sequence of $CC(A/T)_{6\text{-}8}GG$ (Nurrish and Treisman,1995; Wang et al., 2004) and allow only one mismatch than CArG box3 is only conserved in *Arabidopsis lyrata* and *Capsella rubella*. Searching the promoter sequences of the *VDD* orthologs of all these species showed that in *Arabis alpina* and *Thellungiella halophile* CArG boxes that full-fill the consensus sequence could be identified in the same region where the three CArG boxes are in *A. thaliana* and *A. lyrata*. However, position and spacing of these are different.

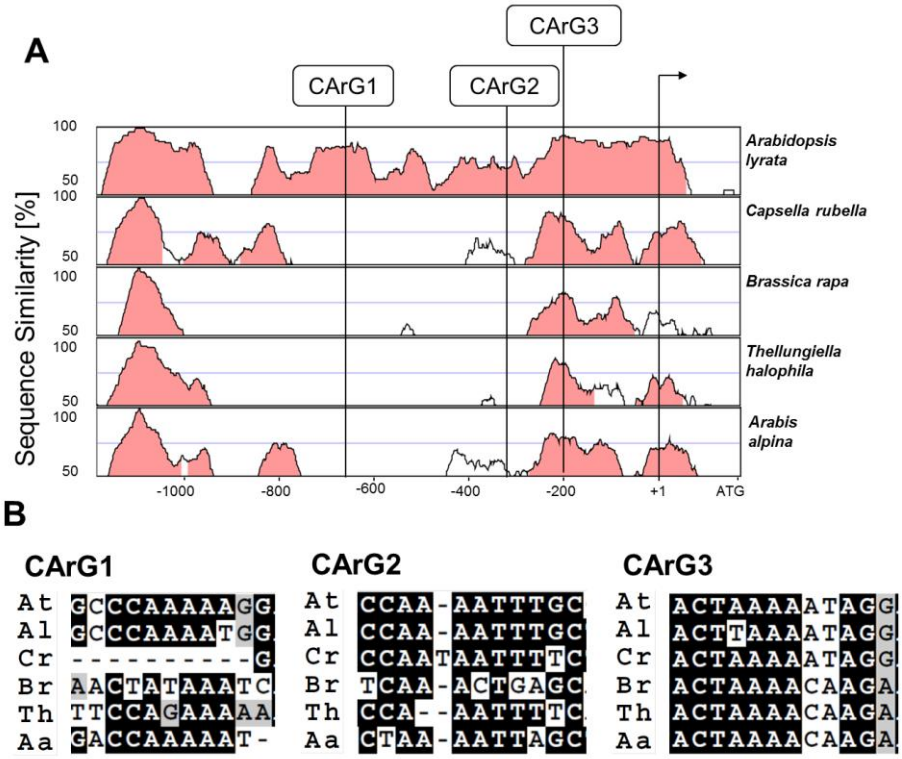


Figure 5. Phylogenetically conserved regions in the *VDD* promoter and multiple sequence alignment analysis

(A) Pairwise alignments of the *Arabidopsis thaliana VDD* promoter to orthologous sequences of *A. lyrata*, *C. rubella*, *B. rapa*, *T. halophila* and *A. alpina*, respectively, shown as VISTA plots. Light-red color indicates regions where a sliding window of at least 75 bp has >70% identity. Vertical lines indicate the position of the three CArG boxes within the *A. thaliana VDD* promoter and its orthologs, relative to the transcriptional start site (arrow). (B) Multiple sequence alignment of the three putative CArG boxes found within the same position of the *VDD* promoter of Arabidopsis and orthologous genes in the 5 *Brassicaceae;* not strictly considering the $CC(A/T)_{6\text{-}8}GG$ (allowing one mismatch) rule. At = *Arabidopsis thaliana*, Al =

Arabidopsis lyrata, Cr = *Capsella rubella*, Br = *Brassica rapa*, Th = *Thellungiella halophila*, Aa = *Arabis alpina*

Discussion

MADS-domain proteins controlling flower development have shown to interact which each other forming preferentially heterodimers (de Folter et al. 2005). The current model, based on biochemical and genetic studies predicts that these floral homeotic MADS-domain protein dimers bind to two CArG boxes as a quartet and establish DNA loops in the promoters of target genes (Egea-Cortines et al. 1999; Melzer and Theissen 2009; Smaczniak et al. 2012). The class E or SEP proteins are important in this model for the establishment of these higher order MADS-domain protein complexes (Honma and Goto 2001; Pelaz et al. 2001).

Our TPM in vitro experiments using a fragment of the *VDD* promoter region containing three adjacent CArG boxes strongly support the idea that loop formation can be established by a STK-SEP3 MADS-domain complex. Especially CArG box1 and 3 seem to be most important for the establishment of loop formation. This is in agreement with ChIP experiments that showed that in vivo STK and SEP3 only use CArG box1 and 3 in the *VDD* promoter (Matias-Hernandez et al. 2010). Interesting is the fact that when we tested by TPM assays the *VDD* promoter with a single CArG box deletion only loop formation was observed when CArG box 2 was deleted. Using tethers in which CArG box1 or 3 were deleted no loop formation was observed. This indicates that loop formation between CArG box1 and 2 or between 2 and 3 is not possible in these in vitro assays. An exception to this rule might be when all three CArG boxes are available and the binding to CArG box1 and 3 somehow also favours the establishment of a loop between CArG box1 and 2. These TPM data are not completely in agreement with our in

vivo data that suggest that the SEP3-STK dimer was also using CArG box 2 in the absence of CArG box1 or 3.

Loop formation in the DNA facilitated by specific cis-elements has shown to be important to establish the interaction of distantly related enhancers for correct regulation of transcription, as for instance described for the intensively studied cis-regulatory region of the *lac* operon of *E. coli* (Lee et al. 1992) or the *Abdominal-B* (*Abd-B*) gene of *Drosphila melongaster* (Cleard et al. 2006; Ho et al. 2011). However, most of the loop formations that have been studied intensively are related to long range DNA looping. In the case of the *VDD* promoter region CArG box1 and 3 are only 444 bp apart. The function of these short-range loops has been poorly studied and understood. The general idea is that DNA looping is a conformational state in which cis elements are brought in close vicinity to each other and create locally a high concentration of transcription factors close to the transcription start site of genes to initiate transcription (Dekker et al. 2002). If this is true then this means that long and short range looping events might in principal have the same function. A study on short-range loop formation in the murine iNOS promoter region also points in this direction (Guo et al. 2008). Our studies in flowers showed that when one of the three CArG boxes was mutated in the *VDD* promoter transcriptional activation of the *VDD* gene was still occurring. However, the promoter was inactive when two of the three CArG boxes were eliminated. This suggests that loop formation is essential for the transcriptional activation of *VDD*. Furthermore, the loop size or its location also seemed to be critical. This became clear from the ChIP and in vivo expression studies using the reporter lines. The ChIP experiments showed that normally CArG box1 and 3 are used. However, when one of these two CArG boxes was mutated, CArG box2 was occupied by SEP3 and STK. This shows that all three boxes have affinity for the SEP3-STK dimer but there seems to be an affinity difference between them, with CArG box2

having the lowest affinity. When CArG box2 was used in combination with CArG box1 or 3 a change in the DNA loop position and/or size is expected to occur (Figure 5). This change in the predicted loop structure showed to have an effect on the expression of the *VDD* gene as evidenced by the reporter gene studies. Loop formation between CArG box1 and 2 or between CArG box2 and 3 did activate expression of the reporter gene but timing and the domain of its expression were altered. These results also suggest that there might be a mechanistic difference between long-range and short-range loop formation. Our results point to the fact that the size and position of the loop is important. For instance a loop between CArG box2 and 3 has the same position relative to the transcription start site but the loop is smaller (Figure 6). It is difficult to imagine that loop size is critical for long-range loop formation, where loops can be thousands of basepairs. Therefore, short-range loops might also be important to give locally structure to the chromatin to recruit or stabilize specific transcriptional complexes. We can of course not exclude that the stability of the MADS-domain protein complex on CArG box2 is less stable and that therefore transcription of *VDD* is deregulated.

The importance of loop formation for *VDD* promoter activity was further strengthened by studying the reporter constructs in which two out of three CArG boxes were mutated. Not only reporter gene expression was completely lost in ovules but also SEP3 and STK lost their ability to bind to the remaining CArG box in floral tissues since no enrichment was found by ChIP analysis on these fragments. This illustrates that binding of SEP3 and STK was only possible when two of the three CArG boxes were available supporting the cooperative assembly of the MADS quartet on the *VDD* promoter.

Our data strongly suggest that loop formation between two CArG boxes is important for *VDD* promoter activity. Nevertheless based on our in vivo results an alternative hypothesis to explain our observations might also be considered. It could be that for *VDD* promoter activity SEP-STK dimers have to bind to at least

two of the three CArG boxes without the necessity that they also loop the DNA. However, our TPM studies and evidence coming from other studies strongly support the looping hypothesis (Egea-Cortines et al., 1999; Melzer and Theissen, 2009; Melzer et al., 2009; Smaczniak et al, 2012).

Another important consideration is that STK and SEP3 might bind as homodimers to the CArG boxes. STK homodimers were never observed in yeast two-hybrid assays (Favaro et al., 2003; de Folter et al., 2005) although this technique cannot completely exclude such possibility. Our ChIP analyses show that both STK and SEP3 bound to both CArG box1 and 3. Furthermore, the TPM data show that SEP3 or STK by themselves were unable to induce loop formation between CArG box1 and 3. Taking these data together suggests that *VDD* promoters on which STK (or SEP3) homodimers are bound to both CArG boxes will probably result inactive. Therefore, a model in which STK and SEP3 homodimers regulate *VDD* expression is not so attractive, also when considering that SEP3 and STK have high affinity for each other in the yeast assay and that the formation of this heterodimer seems to be highly conserved in plants (Favaro et al., 2002).

The shadowing experiments showed that in *Arabidopsis lyrata* all three CArG boxes were in the same position as in *Arabidopsis thaliana* suggesting that the regulatory mechanism that we describe here is at least conserved in the genus Arabidopsis. Remains of course the question why three CArG boxes are conserved when only two of them seem to be used. This could of course suggest that other MADS-domain proteins bind to box2, however, under the controlled greenhouse conditions that we used, these interactions seem not to be important for correct *VDD* expression since mutations in CArG box2 did not alter the expression profile of *VDD*. This could of course be different when plants are grown under more unfavourable climatic conditions. In the more distantly related species it is difficult to conclude whether *VDD* orthologs are regulated in a similar way. When we only

strictly consider the CArG consensus sequence than we can in some promoters find alternative binding sites although they are not exactly in the same positions as observed in Arabidopsis. Of course it might be that there is more flexibility in the CArG box sequence that still allows SEP3 and STK binding. This would open up the possibility of a wider conservation of this regulatory mechanism in more distantly related species. ChIP and reporter gene studies might clarify these interesting questions in the future.

Interestingly, during seed development both CArG box1 and 3 seemed to be essential for *VDD* expression. During this phase of development CArG box2 did not seem to be able to compensate for the loss of one of these two CArG boxes. This suggest that the composition of the MADS-domain complexes that bind to these CArG boxes during seed development are different and that these do not have enough affinity for CArG box2. This is supported by the observation that *VDD* is highly expressed in the embryo a tissue where *STK* mRNAs were never detected by in situ hybridization.

In conclusion, a combination of in vitro and in vivo data strongly support the hypothesis that MADS-domain protein dimers composed of SEP3 and STK (or SHP1/SHP2, which are considered to be redundant with STK in the control of *VDD*) can bind the DNA at nearby CArG boxes and that by forming higher order (quartet) complexes they loop the DNA. This loop formation is important for target gene expression and that both the size and position of these small loops influence gene expression. This is the first in vivo example that shows the importance of MADS domain quartets for target gene regulation and the importance of loop formation for gene expression in plants.

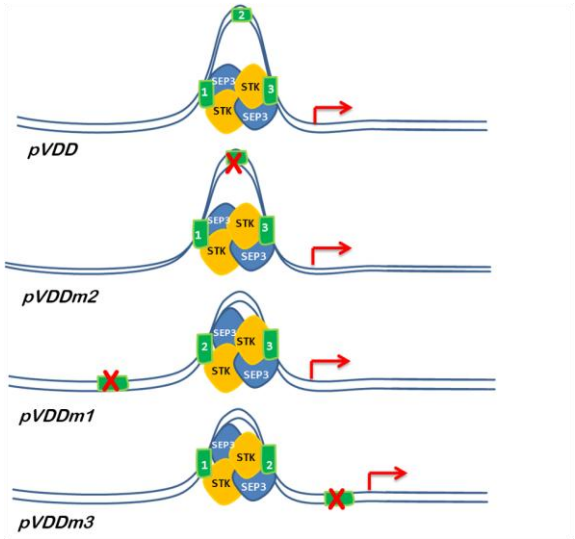


Figure 6. Schematic representation of STK-SEP3 MADS-domain complex on the *VDD* promoter.

(A) Illustration of loop formation in the *VDD* promoter, (B) *VDD* promoter with the second CArG mutated, (C) first CArG mutated and (D) third CArG mutated (D). Cartoon clarifies the putative changes in loop size and position in respect to the transcription start site (red arrow).

Material and Methods

Plant material, growth condition

Arabidopsis thaliana wild-type (ecotype Columbia) and *stk-2* mutant plants (Pinyopich et al. 2003) were grown at 22 °C under short-day (8 h light/16 h dark) or long day (16h light/8 h dark) conditions.

Plasmid constructions and Arabidopsis transformation

For *VDD* promoter analysis 1221 bp upstream of the *VDD* translation start site were amplified by PCR and fused to the *GUS* reporter gene. Seven other constructs were cloned containing different combinations of site specifically mutagenized CArG boxes (Table S1). All the constructs were made using Gateway technology (Invitrogen). In the first amplification step we used the Gateway vector *pDONOR207* and then recombined into pBGWFS7 (Karimi et al. 2002). Wild-type plants were transformed with all constructs using the floral dip method (Clough and Bent 1998). Seeds of the transformed plants were harvested upon maturation, the seeds germinated on soil and the transgenic plants were selected by spraying with 0,1% BASTA herbicide.

Cytological assays

For in situ hybridization analysis, *Arabidopsis* flowers were fixed and embedded in paraffin as described previously (Huijser et al., 1992). Sections of plant tissue were probed with digoxigenin-labeled *VDD* antisense RNA corresponding to nucleotides 240 to 557 (Matias-Hernandez et al. 2010). Hybridization and immunological detection were performed as described previously (Coen et al., 1990).

All *GUS* assays were performed overnight as described previously (Liljegren et al. 2000). Samples were incubated in clearing solution, dissected, and observed using a Zeiss Axiophot D1 microscope equipped with DIC optics. Images were captured on an AxiocamMRc5 camera (Zeiss) using the Axiovision program (version 4.1). For each construct we analysed more than 80 independent transformants.

Chromatin Immunoprecipitation assays

For ChIP experiments chromatin was extracted from wild type, *stk* mutant and transgenic (*pVDDm1::GUS, pVDDm2::GUS, pVDDm3::GUS*) flowers (max. flower developmental stage 12 and before fertilization occurs; Smyth et al, 1990).

Wild-type plants were grown under SD conditions for 2 weeks and chromatin was extracted. The *stk* single mutant and wild-type leaves were used as negative controls for STK and SEP3 ChIPs, respectively. STK and SEP3 binding to the DNA fragment was considered only when they were significantly enriched compared to the controls in at least three independent experiments, for further details see Text S1.

Quantitative real-time RT-PCR

Enrichment folds were detected using a SYBR Green assay (Bio-Rad, http://www.bio-rad.com/). The real-time PCR assay was performed in triplicate using a Bio-Rad C1000 Thermal Cycler optical system. For ChIP experiments, relative enrichment was calculated as described in Text S1. Primers used for ChIP experiments are listed in Table S2.

SEP3 and STK purification

SEP3 coding sequence was amplified using 5' ccatatgggaagagggagagtagaattg 3' and 5'cgctcgagaatagagttggtgtcataagg 3'. STK coding sequence was amplified using 5' cccatatgggaagaggaaagatagaaataaag 3'and 5' ccctcgagtccgagatgaagaattttcttg 3'. The two fragments have been digested with NdeI and XhoI and cloned in pET-23 (+) Novagen (Madison, WI, USA). The recombinant proteins have been produced in *E.coli* BL21(DE3) Novagen (Madison, WI, USA). The cultures were grown at 37° C at $A_{600=nm}$ 0.6 for STK and $A_{600=nm}$ 0.8 for Sep3. IPTG (Isopropyl ß-D-thio-galactopyranoiside supplied by Roche, Germany) was added to a final concentration of 0.1 mM for protein induction afterwards cultures were incubated at 18°C for 15 hours. The His-tagged recombinant proteins were purified using

affinity Ni-NTA Agarose columns (Qiagen). SEP3 was soluble and purified in native condition following (Bellorini et al. 1997). STK accumulated in the inclusion bodies which were solubilised with 20 mM Tris, 500 mM NaCl, 10mM Imidazole and 6M urea pH 8 (4 ml per gram of wet cells) stirring for 30' at RT. The surnatant was loaded on a Ni-NTA column and the bound recombinant STK proteins were solubilised with a linear 6-0 M urea gradient.

DNA constructs for TPM analysis

The *VDD* promoter (*pVDD*) fragment, containing the three CArG boxes, was obtained by PCR using a 5' digoxigenin- and a 3' biotin-labelled primer (Oligos etc. Inc., OR, USA). The *pVDDdel1-2-3*-fragment, in which the three CArG boxes were deleted, was produced by PCR with specific primers carrying CArG-box deletions. The final fragment was recombined into the pGEM-T-Easy plasmid (Promega, Madison Wisconsin, USA) and amplified with labelled primers as described above. Fragments used to obtain a calibration curve, corresponding to 243, 355, and 575 bp were amplified from *pVDD* (primers are listed in Table S3).

Tethered particle motion assay

TPM analysis was performed as described previously by (Finzi and Dunlap 2003). About 50 DNA-tethered beads were tracked for each of the following experimental conditions: i) *pVDD* incubated with SEP3, STK or both. We also tested the mutated promoter without any of the three CArG boxes, *pVDDdel1-2-3-*, in the presence of both proteins (concentration of 700 nmol each).

Phylogenetic Shadowing

Sequences from Arabidopsis lyrata, Brassica rapa and Thellungiella halophila were obtained from Phytozome (www.phytozome.net).

The Capsella rubella sequence was assembled from raw sequence reads (http://trace.ncbi.nlm.nih.gov/Traces/sra/sra.cgi). The Arabis alpina sequence was obtained from an internal genome-sequencing project at the MPIPZ Köln. Pairwise alignments and VISTA plots (Mayor et al., 2000) were made as described previously (Herrero et al., 2012), but with a calculation window of 75bp and a consensus identity of 70%. Multiple sequence alignments were performed with ClustalW (Larkin et al., 2007) and conserved cis-regulatory elements were visualized with WEBLOGO (Crooks et al., 2004). The CArG box consensus that we used was $CC(A/T)_{6-8}GG$, allowing one mismatch. However, the base preceding the (A/T)s should be a C and the base after the (A/T)s should be a G (Nurrish and Treisman,1995; Wang et al. 2004).

Acknowledgments

We thank Dr. George Coupland for the scientific discussion. M.A.M. was funded by EU-ITN (ITN-SYSFLO project) and L.F. was supported by NIH RGM084070A.

Authors contributions: M.M.K. and L.C. designed research; R.F.G. and M.A.M. performed research; S.M., L.F. and C.M. contributed new reagents/analytic tools; R.F.G., M.A.M., M.B., M.M.K. and L.C. analyzed data; and R.F.G., M.A.M., M.M.K. and L.C. wrote the paper.

REFERENCES

Arora, R., Agarwal, P., Ray, S., Singh, A.K., Singh, V.P., Tyagi, A.K., and Kapoor, S. (2007). MADS-box gene family in rice: genome-wide identification,

organization and expression profiling during reproductive development and stress. BMC Genomics **8**, 242

Bellorini, M., Lee, D.K., Dantonel, J.C., Zemzoumi, K., Roeder, R.G., Tora, L., and Mantovani, R. (1997). CCAAT binding NF-Y-YBP interactions: NF-YB and NF-YC require short domains adjacent to their histone fold motifs for association with TBP basic residues. Nucleic Acids Research **25**, 2174-2181.

Cleard, F., Moshkin, Y., Karch, F., and Maeda, R.K. (2006). Probing long-distance regulatory interactions in the Drosophila melanogaster bithorax complex using Dam identification. Nature Genetics **38**, 931-935.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. Plant Journal **16**, 735-743.

Coen, E.S., and Meyerowitz, E.M. (1991). The war of the whorls. Genetic interaction controlling flower development. Nature **353**, 31-37.

de Folter, S., Immink, R.G.H., Kieffer, M., Parenicova, L., Henz, S.R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M.M., *et al.* (2005). Comprehensive interaction map of the Arabidopsis MADS box transcription factors. Plant Cell **17**, 1424-1433.

Dekker, R.J., van Soest, S., Fontijn, R.D., Salamanca, S., de Groot, P.G., VanBavel, E., Pannekoek, H., and Horrevoets, A.J.G. (2002). Prolonged fluid shear stress induces a distinct set of endothelial cell genes, most specifically lung Kruppel-like factor (KLF2). Blood **100**, 1689-1698.

Ditta, G., Pinyopich, A., Robles, P., Pelaz, S., and Yanofsky, M.F. (2004). The SEP4 gene of Arabidopsis thaliana functions in floral organ and meristem identity. Current Biology **14**, 1935-1940.

Dunlap, D., Zurla, C., Manzo, C., and Finzi, L. (2011). Probing DNA Topology Using Tethered Particle Motion. In Single Molecule Analysis: Methods and Protocols, E.J.G. Peterman, and G.J.L. Wuite, eds., pp. 295-313.

Egea-Cortines, M., Saedler, H., and Sommer, H. (1999). Ternary complex formation between the MADS-box proteins SQUAMOSA, DEFICIENS and GLOBOSA is involved in the control of floral architecture in Antirrhinum majus. Embo Journal **18**, 5370-5379.

Favaro, R., Pinyopich, A., Battaglia, R., Kooiker, M., Borghi, L., Ditta, G., Yanofsky, M.F., Kater, M.M., and Colombo, L. (2003). MADS-box protein complexes control carpel and ovule development in Arabidopsis. Plant Cell **15**, 2603-2611.

Finzi, L., and Dunlap, D. (2003). Single-molecule studies of DNA architectural changes induced by regulatory proteins. Rna Polymerases and Associated Factors, Pt C **370**, 369-378.

Finzi, L., and Gelles, J. (1995). Measurement of lactose repressor-mediated loop formation and breakdown in single DNA-molecules . Science **267**, 378-380.

Guerra, R.F., Imperadori, L., Mantovani, R., Dunlap, D.D., and Finzi, L. (2007). DNA compaction by the nuclear factor-Y. Biophysical Journal **93**, 176-182.

Guo, H., Mi, Z., and Kuo, P.C. (2008). Characterization of short range DNA looping in endotoxin-mediated transcription of the murine inducible nitric-oxide synthase (iNOS) gene. Journal of Biological Chemistry **283**, 25209-25217.

Herrero, E., Kolmos, E., Bujdoso, N., Yuan, Y., Wang, M., Berns, M.C., Uhlworm, H., Coupland, G., Saini, R., Jaskolski, M., Webb, A., Gonçalves, J., and Davis, S.J. (2012) EARLY FLOWERING4 recruitment of EARLY FLOWERING3 in the Nucleus Sustains the Arabidopsis Circadian Clock. Plant Cell **24**: 428-443.

Huijser, P., Klien, J., Lonnig, W.E., Meijer, H., Saedler, H., and Sommer, H. (1992). Bracteomania, an inflorescence anomaly is caused by the loss of function of the MADS-box gene *SQUAMOSA* in *Antirrhinum majus*. EMBO J. **11:** 1239–1249.

Ho, M.C.W., Schiller, B.J., Akbari, O.S., Bae, E., and Drewell, R.A. (2011). Disruption of the Abdominal-B Promoter Tethering Element Results in a Loss of Long-Range Enhancer-Directed Hox Gene Expression in Drosophila. PloS one **6** (1) e 16283.

Honma, T., and Goto, K. (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. Nature **409**, 525-529.

Karimi, M., Inze, D., and Depicker, A. (2002). GATEWAY((TM)) vectors for Agrobacterium-mediated plant transformation. Trends in Plant Science **7**, 193-195.

Lee, D.H., Huo, L., and Schleif, R. (1992). REPRESSION OF THE ARABAD PROMOTER FROM ARAO1. Journal of Molecular Biology **224**, 335-341.

Liljegren, S.J., Ditta, G.S., Eshed, H.Y., Savidge, B., Bowman, J.L., and Yanofsky, M.F. (2000). SHATTERPROOF MADS-box genes control seed dispersal in Arabidopsis. Nature **404**, 766-770.

Liu, C., Chen, H., Er, H.L., Soo, H.M., Kumar, P.P., Han, J.-H., Liou, Y.C., and Yu, H. (2008). Direct interaction of AGL24 and SOC1 integrates flowering signals in Arabidopsis. Development **135**, 1481-1491.

Matias-Hernandez, L., Battaglia, R., Galbiati, F., Rubes, M., Eichenberger, C., Grossniklaus, U., Kater, M.M., and Colombo, L. (2010). VERDANDI Is a Direct Target of the MADS Domain Ovule Identity Complex and Affects Embryo Sac Differentiation in Arabidopsis. Plant Cell **22**, 1702-1715.

Mayor, C., Brudno, M., Schwartz, J.R., Poliakov, A., Rubin, E.M., Frazer, K.A., Pachter, L.S., Dubchak, I. (2000). VISTA: visualizing global DNA sequence alignments of arbitrary length. Bioinformatics **16:** 1046-1047.

Melzer, R., and Theissen, G. (2009). Reconstitution of floral quartets in vitro involving class B and class E floral homeotic proteins. Nucleic Acids Research **37**, 2723-2736.

Melzer, R., Verelst, W., and Theissen, G. (2009). The class E floral homeotic protein SEPALLATA3 is sufficient to loop DNA in floral quartet-like complexes in vitro. Nucleic Acids Research **37**, 144-157.

Melzer, S., Lens, F., Gennen, J., Vanneste, S., Rohde, A., and Beeckman, T. (2008). Flowering-time genes modulate meristem determinacy and growth form in Arabidopsis thaliana. Nature Genetics **40**, 1489-1492.

Nelson, P.C., Zurla, C., Brogioli, D., Beausang, J.F., Finzi, L., and Dunlap, D. (2006). Tethered particle motion as a diagnostic of DNA tether length. Journal of Physical Chemistry **110**, 17260-17267.

Nurrish, S.J., and Treisman, R. (1995). DNA-binding specificity determinants in MADS-BOX transcription factors. Molecular and Cellular Biology **15**, 4076-4085.

Parenicova, L., de Folter, S., Kieffer, M., Horner, D.S., Favalli, C., Busscher, J., Cook, H.E., Ingram, R.M., Kater, M.M., Davies, B., **Angenent C.G and Colombo L**. (2003). Molecular and phylogenetic analyses of the complete MADS-box transcription factor family in Arabidopsis: New openings to the MADS world. Plant Cell **15**, 1538-1551.

Pelaz, S., Ditta, G.S., Baumann, E., Wisman, E., and Yanofsky, M.F. (2000). B and C floral organ identity functions require SEPALLATA MADS-box genes. Nature **405**, 200-203.

Pelaz, S., Gustafson-Brown, C., Kohalmi, S.E., Crosby, W.L., and Yanofsky, M.F. (2001). APETALA1 and SEPALLATA3 interact to promote flower development. Plant Journal **26**, 385-394.

Pinyopich, A., Ditta, G.S., Savidge, B., Liljegren, S.J., Baumann, E., Wisman, E., and Yanofsky, M.F. (2003). Assessing the redundancy of MADS-box genes during carpel and ovule development. Nature **424**, 85-88.

Pouget, N., Dennis, C., Turlan, C., Grigoriev, M., Chandler, M., and Salome, L. (2004). Single-particle tracking for DNA tether length monitoring. Nucleic Acids Research **32** (9) e 73

Riechmann, J.L., Wang, M.Q., and Meyerowitz, E.M. (1996). DNA-binding properties of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. Nucleic Acids Research **24**, 3134-3141.

Romanel, E.A.C., Schrago, C.G., Counago, R.M., Russo, C.A.M., and Alves-Ferreira, M. (2009). Evolution of the B3 DNA Binding Superfamily: New Insights into REM Family Gene Diversification. PloS one **4**, (6) e5791.

Schneitz, K., Hulskamp, M., and Pruitt, R.E. (1995). Wild-tpe ovule development in Arabidopsis thailana - A light-microscope study of cleared whole-mount tissue Plant Journal **7**, 731-749.

Schwarz Sommer, Z., Hue, I., Huijser, P., Flor, P.J., Hansen, R., Tetens, F., Lonnig, W.E., Saedler, H., and Sommer, H. (1992). Characterization of the Antirrhinum floral homeotic MADS-box gene *DEFICIENS* - evidence for DNA-binding and auto regulation of its persistent expression throughout flower development. EMBO Journal 11, 251-263.

Shore, P., and Sharrocks, A.D. (1995). The MADS-box family of transcription factors. European Journal of Biochemistry **229**, 1-13.

Smaczniak, C., Immink, R.G.H., Angenent, G.C., and Kaufmann, K. (2012). Developmental and evolutionary diversity of plant MADS-domain factors: insights from recent studies. Development **139**, 3081-3098.

Theissen, G., and Saedler, H. (2001). Plant biology - Floral quartets. Nature **409**, 469-471.

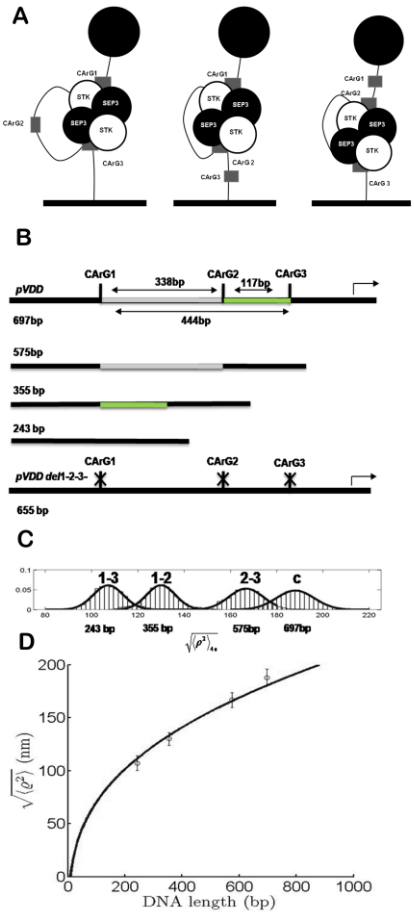
West, A.G., and Sharrocks, A.D. (1999). MADS-box transcription factors adopt alternative mechanisms for bending DNA. Journal of Molecular Biology **286**, 1311-1323.

Wang, H., Caruso, L.V., Downie, A.B. and Perry S.E. (2004). The embryo MADS domain protein AGAMOUS-Like 15 directly regulates expression of a

gene encoding an enzyme involved in gibberellin metabolism. Plant Cell **16**, 1206-1219.

Zaremba, M., Owsicka, A., Tamulaitis, G., Sasnauskas, G., Shlyakhtenko, L.S., Lushnikov, A.Y., Lyubchenko, Y.L., Laurens, N., van den Broek, B., Wuite, G.J.L., **Siksnys, V.** (2010). DNA synapsis through transient tetramerization triggers cleavage by Ecl18kI restriction enzyme. Nucleic Acids Research **38**, 7142-7154.

Zurla, C., Manzo, C., Dunlap, D., Lewis, D.E.A., Adhya, S., and Finzi, L. (2009). Direct demonstration and quantification of long-range DNA looping by the bacteriophage repressor. Nucleic Acids Research **37**, 2789-2795.



Supplemental information

Figure S1. Summary of TPM analysis

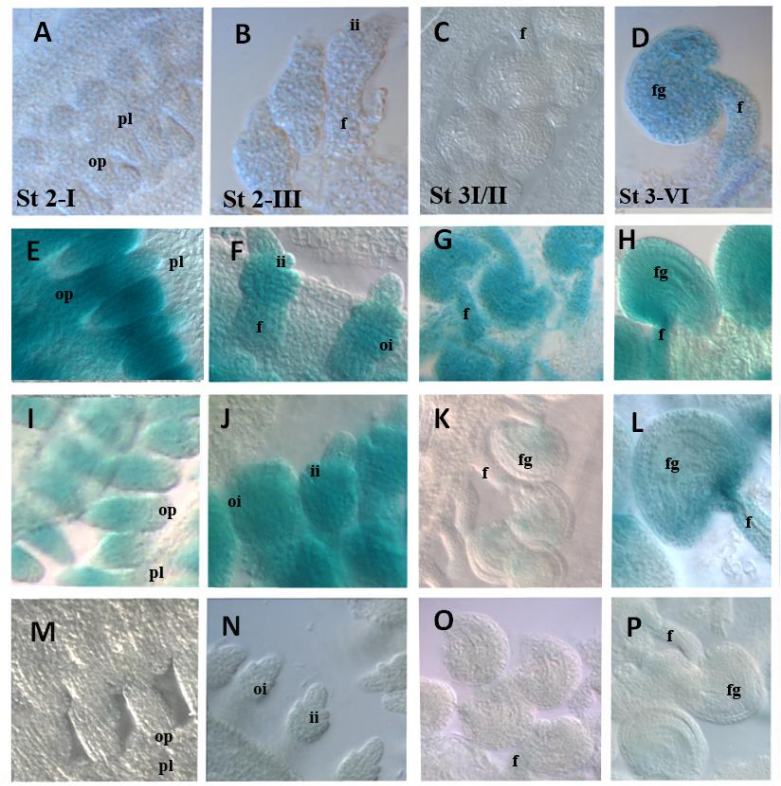
(A) Schematic representation of TPM assay. DNA looping is observed as a result of changes in the Brownian motion of the tethered bead (tb). STK-SEP3 complexes could bind to the CArG boxes making three possible loops: between CArG box 1 and 3, between CArG box 1 and 2, or between CArG box 2 and 3.

(B) DNA constructs used. Total construct length is indicated. Fragment 575 bp has the same length as expected for the *pVDD* fragment of 697bp with a loop between CArG2 and CArG3. Fragment 355bp has a length corresponding to *pVDD* containing a loop between CArG1 and CArG2 and the fragment of 243bp corresponds to *pVDD* with a loop between CArG1 and CArG3. *pVDDdel1-2-3-* is a DNA fragment without any of the three CArG boxes, and it is 655 bp long.

(C) Cumulative histograms of $\sqrt{\langle \rho^2 \rangle_{4s}}$ for the 243, 355, 575, 697 bp fragments. The data have been fitted to a Gaussian distribution. The numbers above each distribution indicate the CArG boxes, whose interaction in the wild-type fragment would produce a DNA tether of equivalent length and TPM signal. "**c**" stands for wild type control fragment. The histograms are normalized to the total number of events and to the bin width (2 nm).

(D) Calibration curve relating the expected TPM signal, $\sqrt{\langle \rho^2 \rangle_{4s}}$, to DNA length. Experimental data match the TPM signal obtained for the four DNA lengths in (A). The error bars represent the standard deviation of the data. The continuous line is the calibration curve as obtained by Nelson et al., 2006, assuming a DNA persistence length = 41 nm.

Figure S2 *pVDD* deletion studies



(A-D) GUS expression in ovules of *pVDDDel1::GUS* lines; (E-H) GUS expression in ovules of *pVDDDel2::GUS* lines; (I-L) GUS expression in ovules of *pVDDdel3::GUS* lines; (M-P) Absence of GUS expression as observed in the *pVDDdel-1-2::GUS.* The same result was obtained in *pVDDdel1-3::GUS; pVDDdel2-3::GUS; pVDDdel1-2-3::GUS* lines. pl-placenta; op- ovule primordium; f-funiculus; ii- inner integument; oi- outer integument; fg- female gametophyte

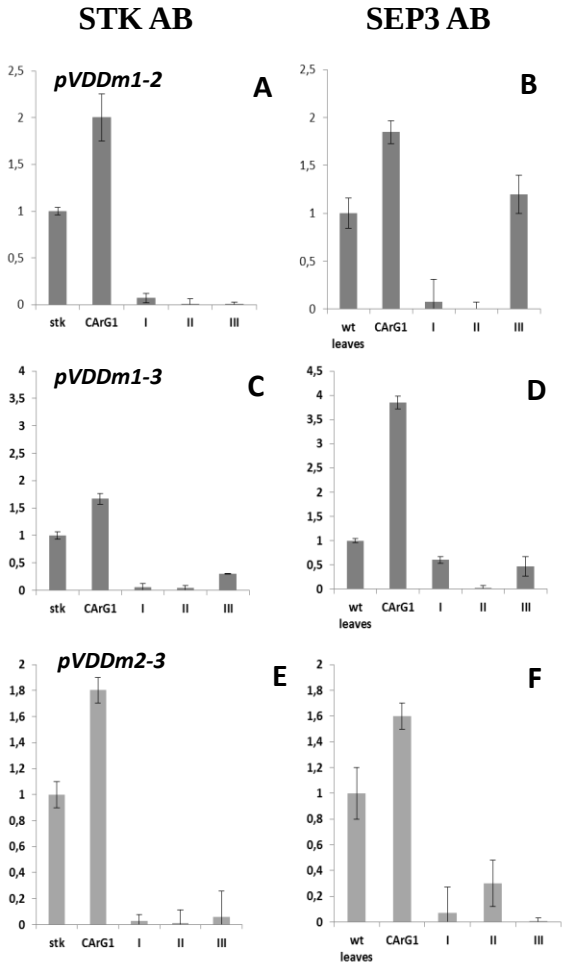


Figure S3. In vivo binding of SEP3 and STK to CArG boxes with two CArG boxes mutated in the*VDD* promoter region

ChIP experiments using STK (panel A, C, and G) and SEP3 (panel B, D and F) antibodies, to investigate the binding to different CArG boxes. As positive control we tested endogenous CArG box 1 of *VDD* promoter as negative controls, for SEP3 ChIP assays we used wild type leaf tissues and for STK ChIP negative controls flowers of the *stk* mutant were used. I, referes to CArG box1 mutated; II to CArG-box2 mutated; III to CArG box3 mutated in the *VDD* promoter. (A and B)

STK and SEP3 binding to the heterologous *pVDDm1-2*. (C-D) ChIP assays to test SEP3 and STK binding *pVDDm1-3.* (E and F) *pVDDm2-3* ChIP experiments. Error bars represent the propagated error value using three replicates.

Table S1. Primers used for mutated plasmids construction

Construct	Primer designation	Nucleotide sequence
pVDD::GUS	VDD promoter wild type forward	5’ ggggacaagtttgtacaaaaaagcaggctcccgaactttattccggata 3’
	VDD promoter wild type reverse	5’ggggaccactttgtacaagaaagctgggtcctctgcttctctccttc 3’
pVDD1-::GUS	*pVDD1-::GUS* mutated forward	5’ccaaataataaagacaagtatacattgcccCGGGGGgaaaactatagag 3’
	pVDD1-::GUS mutated reverse	5’-ctctatagttttcCCCCCGgggcaatgtatacttgtctttattatttgg-3’
pVDD2-::GUS	*pVDD2-::GUS* mutated forward	5’-gaggttttcaaatgggtgattaaccCCCCGGGgccgtgaatgtattttag-3’
	pVDD2-::GUS mutated reverse	5’-ctaaaatacattcacggCCCCGGGgggttaatcacccatttgaaaacctc;
pVDD3-::GUS	*pVDD3-::GUS* mutated forward	5’-gctgtctttttagaattcagttactGGGGGtagga atttgctctgctttttac-3’
	pVDD3-::GUS mutated reverse	5’-gtaaaaagcagagcaaattcctaCCCCCagtaactgaattctaaaaagacagc-3’

Detailed list of primers used for constructing all plasmids. CArG box consensus sequences are underlined and mutations are indicated by upper case letters. Mutated plasmids were named: *pVDDm1::GUS*, with the CArG1 mutated; *pVDDm2::GUS* , CArG2 mutated; *pVDDm3::GUS* , CArG3 mutated; *pVDDm1-2-::GUS* , CArG1 and CArG2 mutated; *pVDDm2-3::GUS* , CArG2 and CArG3

mutated; *pVDDm1-3::GUS* , CArG1 and CArG3 mutated ; *pVDDm1-2-3::GUS* with all CArG boxes mutated.

Table S2. Primers used for ChIP experiment

Primer designation	Nucleotide sequence
ACTIN forward (Liu et al., 2008)	5'-cgtttcgctttccttagtgttagct-3'
ACTIN reverse (Liu et al., 2008)	5'-agcgaacggatctagagactcaccttg-3'
wild type CArG box1 forward	5'-aacattgctttctccttccaaa-3'
wild type CArG box1 reverse	5'-gtatattcagcgtaacagatacg-3'
wild type CArG box2 forward	5'-ctacattctacagactagctag-3'
wild type CArG box2 reverse	5'-ctaaaaagacagcgtcatatttcc-3'
wild type CArG box3 forward	5'-ggaaatatgacgctgtctttttag-3'
wild type CArG box3 reverse	5'-cagaaacagcaatatgctcgtg-3'
mutated CArG box1 forward	5'-caagtatacattgccccgggg-3'
mutated CARG box1 reverse	5'-cccctattaactttatacaagc-3'
mutated CArG box2 forward	5'-cgtatctgttacgctgaatatac-3'
mutated CArG box2 reverse	5'-ctaaaatacattcacggccccgggg-3'
mutated CArG box3 forward	5'-gtctttttagaattcagttactggggg-3'
mutated CArG box3 reverse	5'-ggttagttggaaaagattccc-3'

List of primers used for Real-Time PCR. Specific primers on the three single mutations were used in order to discriminate enrichment from wild-type promoter sequences and exogenous DNA.

Table S3. Tethered particle motion fragments

Construct	Primer designation	Nucleotide sequence
pVDD	*pVDD* 697 bp forward	5'-gttttcaagatattgtcaagc-3'
	pVDD 697 bp reverse	5'-ggttagttggaaaagattccc-3'
pVDDdel1-pVDDdel1-::GUS	deleted CArG box 1 forward	5'-taataaagacaagtatac*ctatagagacacgcactagttagggttg-3'
	deleted CArG box 1 reverse	5'- caaccctaactagtgcgtgtctctatag*gtatacttgtctttatta -3'
pVDDdel2-pVDDdel2-::GUS	deleted CArG box 2 forward	5'-gttttcaaatgggtgattaacc*gtgaatgtattttagtacagtataag-3'
	deleted CArG box 2 reverse	5'- cttatactgtactaaaatacattcac*ggttaatcacccatttgaaaac -3';
pVDDdel3-pVDDdel3-::GUS	deleted CArG box 3 forward	5'-cgctgtctttttagaattca * ttgctctgctttttacgtgtctggg-3',
	deleted CArG box 3 reverse	5'-cccagacacgtaaaaagcagagcaa*tgaattctaaaaagacagcg -3'
575 bp	R-575 bp reverse	5'-ttcctatttttagtaactg-3'
355 bp	R-355 bp reverse	5'-acatgtttgaaaacttagc-3'
243 bp	R-243 bp reverse	5'-gcatatatgtatattcag-3'

Primers used for amplifying tethers for TPM experiment. Fragments used to obtain a calibration curve corresponding to 243, 355, and 575 bp were amplified from tether *pVDD* using the same *pVDD* forward and as reverse the listed ones. Stars on primers indicate the deletion point.

Text S1

ChIP experiments and Quantitative real-time RT-PCR

About 600 mg of each sample was fixed at 4°C for 20 minutes in 1% formaldehyde under vacuum. ChIP experiments were performed as version described by Dorca-Fornell et al. (2011). The STK and SEP3 polyclonal antibody were previously described (Hernandez et al, 2010). DNA enrichment was tested in triplicate using a Sybr Green Assay (iQ_ SYBR Green Supermix; Bio-Rad) and performed in a Bio-Rad C1000 Thermal Cycler optical system. Relative enrichment was calculated normalizing the amount of immune precipitated DNA against an ACTIN2/7 (ACT2/7) fragment and against total INPUT DNA. In particular, for the binding of STK to the selected genomic regions, the affinity of the purified sample obtained in the wild-type inflorescence was compared with the affinity-purified sample obtained in the *stk* single mutant background, which was used as negative control. For the binding of SEP3 to the selected genomic regions, the affinity of the purified sample obtained from wild-type inflorescences was compared with the affinity-purified sample obtained from wild-type leaf tissue, which was used as negative control. Fold enrichment was calculated using the following formulas, where Ct. tg is target gene mean value, Ct.i is input DNA mean value, and Ct.nc is actin (negative control) mean value: dCT.tg = CT.i-CT.tg and dCT.nc =CT.i-CT.nc. The propagated error values of these CTs are calculated: dSD. tg = sqrt((SD.i)^2+ (SD.tg^2)/sqrt(n) and dSD.nc = sqrt((SD.i)^2+ (SD.nc^2)/ sqrt(n), where n = number of replicates per sample. Fold-change over negative control (actin and wild-type plants) was calculated finding the “delta delta CT” of the target region as follows: ddCT = dCT.tg- dCT.nc and ddSD = sqrt((dSD.tg)^2+ (dSD.nc)^2. The transformation to linear “fold-change” values was obtained as follows: FC = 2^(ddCT) and FC.error = ln(2)*ddSD*FC.

the plant journal

SEB Society for Experimental Biology

The Plant Journal (2012) **70**, 409–420 doi: 10.1111/j.1365-313X.2011.04878.x

The MADS box genes *SEEDSTICK* and *ARABIDOPSIS* B_{sister} play a maternal role in fertilization and seed development

Chiara Mizzotti[1,†], Marta Adelina Mendes[1,†], Elisabetta Caporali[1], Arp Schnittger[2], Martin M. Kater[3], Raffaella Battaglia[3] and Lucia Colombo[1,4,*]

[1]*Dipartimento di Biologia, Università degli Studi di Milano, Via Celoria 26, 20133 Milan, Italy,*

[2]*Institut de Biologie Moléculaire des Plantes du Centre National de la Recherche Scientifique, 12 rue du Général Zimmer, 67084 Strasbourg, France,*

[3]*Dipartimento di Scienze Biomolecolari e Biotecnologie, Università degli Studi di Milano, Via Celoria 26, 20133 Milan, Italy, and*

[4]*Consiglio Nazionale delle Ricerche, Istituto di Biofisica, 20133 Milan, Italy*

Received 6 June 2011; revised 1 December 2011; accepted 6 December 2011; published online 16 January 2012.

*For correspondence (fax +39 2 50314764; e-mail lucia.colombo@unimi.it).

†These authors contributed equally to this work.

SUMMARY

The haploid generation of flowering plants develops within the sporophytic tissues of the ovule. After fertilization, the maternal seed coat develops in a coordinated manner with formation of the embryo and endosperm. In the *arabidopsis bsister* (*abs*) mutant, the endothelium, which is the most inner cell layer of the integuments that surround the haploid embryo sac, does not accumulate proanthocyanidins and the cells have an abnormal morphology. However, fertility is not affected in *abs* single mutants. *SEEDSTICK* regulates ovule identity redundantly with *SHATTERPROOF 1* (*SHP1*) and *SHP2* while a role in the control of fertility was not reported previously. Here we describe the characterization of the *abs stk* double mutant. This double mutant develops very few seeds due to both a reduced number of fertilized ovules and seed abortions later during development. Morphological analysis revealed a total absence of endothelium in this double mutant. Additionally, massive starch accumulation was observed in the embryo sac. The phenotype of the *abs stk* double mutant highlights the importance of the maternal-derived tissues, particularly the endothelium, for the development of the next generation.

Keywords: MADS-box genes, ovule development, fertilization, Arabidopsis, endothelium.

INTRODUCTION

Fertilization and seed formation are key events in the life-cycle of flowering plants. The seed can be considered as a functional unit required for the protection and propagation of the offspring represented by the embryo. The first step in seed development is the formation of ovules. In Arabidopsis, ovule primordia arise from the placenta at stage 8 of flower development, and differentiation is completed at stage 13 (Smyth *et al.*, 1990), when the embryo sac is mature and can be fertilized (Schneitz *et al.*, 1995). The MADS box gene *SEEDSTICK* (*STK*), which encodes a key regulator of ovule development, redundantly controls ovule identity, together with *SHATTERPROOF 1* (*SHP1*) and *SHP2* (Pinyopich *et al.*, 2003). In the *stk shp1 shp2* triple mutant, integuments are converted into carpelloid structures and female gametophyte development is arrested just after megasporogenesis (Brambilla *et al.*, 2007; Battaglia *et al.*, 2008). Furthermore, the funiculus of *stk* single mutant ovules is larger than wild-type, and the mature seeds do not detach from the silique once they are mature (Pinyopich *et al.*, 2003). Recently, we also showed that STK and SEPALLATA 3 (SEP3) together regulate the reproductive meristem (REM) transcription factor gene *VERDANDI* (*VDD*), which controls cell identity in the female gametophyte (Matias-Hernandez *et al.*, 2010).

The functional role of integuments during ovule and seed development has been investigated for many years. Characterization of several Arabidopsis mutants highlighted the role of integuments in many steps of ovule and seed development (for review see Bencivenga *et al.*, 2011). These data support the possibility of an interaction between the

ovule integuments and the developing gametophyte, as well as between the seed coat and the fertilization products (for review, see Nowack *et al.*, 2010). However, it is still not clear how communication between the mother plant and the next generation is established, and what the messengers are. The existence of an interaction between ovule integuments and the developing embryo sac has been proposed on the basis of sporophytic ovule mutants, which show defects in embryo sac development (for review, see Bencivenga *et al.*, 2011). For example, the Arabidopsis *inner no outer* (*ino*) mutant develops a normal inner integument but lacks the outer integument, and megagametogenesis is impaired in this sporophytic mutant (Christensen *et al.*, 1997; Schneitz *et al.*, 1997). Another sporophytic Arabidopsis mutant with gametophytic defects is *aintegumenta* (*ant*). In this mutant, the integuments are reduced or absent, and the embryo sac is not able to develop (Elliott *et al.*, 1996; Klucher *et al.*, 1996; Schneitz *et al.*, 1997).

When double fertilization takes place, one sperm cell fertilizes the egg cell, giving rise to the embryo, and the other fertilizes the diploid central cell, producing the triploid endosperm. In response to fertilization, the ovule integuments differentiate into the seed coat. The seed coat provides a protected environment for embryo development and creates a barrier to prevent damage induced by external factors such as UV radiation or pathogens (Haughn and Chaudhury, 2005). The seed coat of Arabidopsis consists of five cell layers: two originate from the outer integument, and the others derive from the inner integument. The innermost layer, the endothelium, develops from the inner integument by periclinal divisions at the four-nuclear embryo sac stage (Schneitz *et al.*, 1995). Upon fertilization, proanthocyanidins are produced in the endothelium, and their oxidation is responsible for the brown colour of Arabidopsis seeds (Debeaujon *et al.*, 2003). At the torpedo stage of embryo development, the endothelium cell layer begins to break down as the cellular endosperm expands (Andème Ondzighi *et al.*, 2008).

Flavonoid biosynthesis in the innermost layer of the seed coat is an important aspect of seed coat differentiation. Screens for seed pigmentation phenotypes have lead to identification of several genes involved in flavonoid biosynthesis. Some of the mutants that were identified map to the *TRANSPARENT TESTA 16* (*TT16*) locus, which encodes the ARABIDOPSIS B_{sister} (ABS) MADS domain protein. ABS is necessary for proanthocyanidin accumulation in the endothelium of the seed coat, with the exception of the chalaza/micropyle area. Detailed morphological analysis of the endothelium in the *abs* mutant revealed that the cells of this layer had an altered morphology, with a parenchymal appearance and occasionally flattening, suggesting loss of cell identity. Despite these morphological observations, ovule function did not appear to be affected (Nesi *et al.*, 2002).

Recently, an important role in the transport of nutrients has been proposed for the seed coat. Nutrients move centripetally from the outer to the inner integuments via the apoplast to the endosperm and to the embryo. In this process, the endothelium plays an important role in protecting and nourishing the embryo (Stadler *et al.*, 2005; Morley-Smith *et al.*, 2008).

Here we describe a new function for the Arabidopsis MADS domain transcription factor-encoding genes *STK* and *ABS*. We describe the role of STK and ABS in the maternal control of endothelium formation, as the *abs stk* double mutant completely lacks endothelium. Interestingly, the absence of endothelium is responsible for defects in the female gametophyte, reduced fertilization and defects during seed formation. Altogether, these defects lead to a severely reduced seed set. Our data show that the endothelium has a fundamental function in the interaction between the maternally derived integuments and the next generation which is protected by this sporophytic tissue.

RESULTS

STK and ABS play a maternal role in ovule and seed development

Protein interaction studies showed that a MADS box protein complex comprising ABS, SEP3 and STK was formed in yeast (Kaufmann *et al.*, 2005). Furthermore, *in situ* hybridization experiments revealed that *ABS* and *STK* mRNAs are both present in the innermost integument layer, the endothelium (Figure 1). Therefore, to identify unknown functions

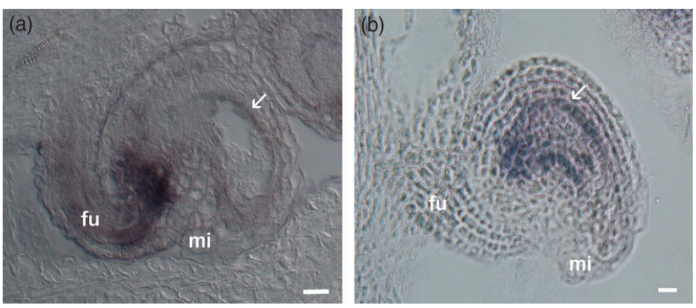


Figure 1. *In situ* hybridization.
(a) Wild-type mature ovule: *STK* mRNA is present in the endothelium (white arrow).
(b) Wild-type mature ovule: *ABS* is clearly expressed in the endothelium (white arrow).
mi, micropyle; fu, funiculus.
Scale bars = 20 μm.

for STK and to investigate genetic and molecular interactions between ABS and STK, the *abs-6 stk-2* double mutant was generated. Both single mutants are fully fertile, but the *abs stk* double mutant showed a drastic reduction in fertility due to defects during ovule and seed formation (Figure 2).

The ovules of the *stk* single mutant have bigger and longer funiculi, and upon fertilization the mature seeds are smaller than wild-type and do not detach from siliques (Pinyopich *et al.*, 2003). The cells forming the endothelium in the *abs* mutant are flatter compared with wild-type, and proanthocyanidins are not synthesized in this integument layer upon fertilization (Nesi *et al.*, 2002). However, neither ovule abortions nor seed arrest were observed in *abs* and *stk* single mutants (Nesi *et al.*, 2002; Pinyopich *et al.*, 2003). This suggests that the ovule and seed abortions observed in the *abs stk* double mutant are linked to the concomitant lack of STK and ABS activities.

As the seed phenotypes of *abs* and *stk* single mutants have a sporophytic origin (Nesi *et al.*, 2002; Pinyopich *et al.*, 2003), genetic experiments were performed to assess whether the *abs stk* double mutant phenotype also had a sporophytic nature. In the *abs stk* double mutant, we observed that 643 ovules of the 1169 analysed (55%) aborted prior to fertilization. To quantify the seed abortion phenotype, we analysed siliques of wild-type and *abs stk* plants at various days after pollination (DAP) and counted the number of seeds that developed from the fertilized ovules (Table 1). The embryos were staged as previously described by Hsu *et al.* (2010). This analysis revealed that 98.7% of the developing embryos in wild-type siliques at 5 DAP were at early heart stage, but only 66.8% of the embryos had reached this stage of development in the *abs stk* mutant. At 13 DAP, only 58.1% of the *abs stk* developing embryos had reached the bent cotyledon stage, whereas 93.5% were at this stage of development in the wild-type. Finally, when the siliques were mature, 98.5% of the seeds of wild-type plants had completed their development, compared with only 24.4% in the *abs stk* double mutant. This analysis was also performed in the *abs* and *stk* single mutants, and, as expected, no differences were observed compared to wild-type (Table 1). Furthermore, reciprocal crosses confirmed that these *abs stk* double mutant phenotypes are sporophytically controlled (Table S1).

The role of STK and ABS in embryo and endosperm development was tested using a genetic approach: double heterozygous *abs/ABS stk/STK* plants were pollinated with *abs stk* homozygous double mutant pollen. Within the

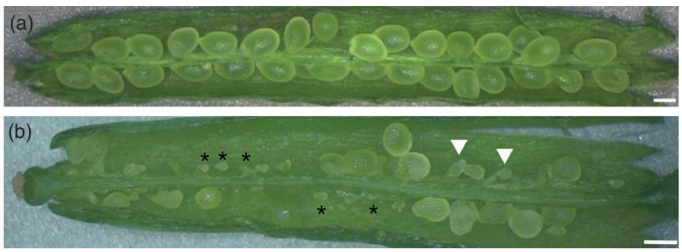


Figure 2. Wild-type and *abs stk* double mutant siliques.
(a) Wild-type silique showing full seed set.
(b) Silique of an *abs stk* double mutant containing aborted ovules (asterisk) and aborted seeds (arrowheads).
Scale bars = 200 μm.

Table 1 Ovule abortions and developing seeds in wild type and *abs*, *stk*, *abs stk* mutants

	Number of siliques analysed	Ovule abortions (Total ovules)	Total developing seeds (Total ovules) [Fertilized ovules]	Percentage (Total ovules) [Fertilized ovules]
0 DAP WT	24	60 (1149)		(5.2)
0 DAP *abs stk*	26	643 (1169)		(55)
0 DAP *abs*	24	43 (1081)		(4)
0 DAP *stk*	24	52 (1032)		(5)
5 DAP WT	12		512 (548) [519]	(93.4) [98.7]
5 DAP *abs stk*	24		326 (1085) [488]	(30) [66.8]
5 DAP *abs*	12		479 (503) [483]	(95.2) [99.2]
5 DAP *stk*	12		538 (577) [548]	(93.2) [98.2]
13 DAP WT	12		603 (645) [645]	(93.5) [93.5]
13 DAP *abs stk*	24		282 (1077) [485]	(26.2) [58.1]
13 DAP *abs*	12		566 (594) [570]	(95.3) [99.3]
13 DAP *stk*	12		585 (628) [597]	(93.2) [98]
Mature WT	12		592 (634) [601]	(93.4) [98.5]
Mature *abs stk*	24		120 (1092) [491]	(11) [24.4]
Mature *abs*	12		548 (593) [569]	(92.4) [96.3]
Mature *stk*	12		512 (560) [532]	(91.4) [96.2]

progeny of this cross, the *abs* and *stk* mutant alleles segregate in the embryo and endosperm, but the sporophytically derived seed coat was heterozygous for both genes. In segregating progeny, 25% of the seeds are expected to have embryo and endosperm homozygous for both *abs* and *stk*, which allowed us to analyse the role of STK and ABS in embryo and endosperm development regardless of their role in the integuments. As a control, we pollinated double heterozygous *abs/ABS stk/STK* plants with wild-type pollen. When pollinating the *abs/ABS stk/STK* plants with *abs stk* homozygous double mutant pollen or wild-type pollen, we observed normal seed set, suggesting that the observed defects in the *abs stk* double mutant are sporophytically determined. Furthermore, these crosses demonstrate that the ABS and STK MADS domain proteins play a fundamental role, even before fertilization, in the molecular interaction between the integuments and the female gametophyte. After fertilization, they control communication between the seed coat and the fertilization products, i.e. the embryo and the endosperm.

To further investigate later seed developmental stages, we performed a germination assay. We tested seeds from *abs* and *stk* single mutant plants and those obtained from self fertilization of a plant heterozygous for both mutant alleles. As mentioned above, more than 90% of seeds develop normally in single mutant plants, but the percentage of mature seeds was only 11% in the *abs stk* double mutant (Table 1). The germination assay (Table 2) showed that seeds obtained from the single mutants, heterozygous double mutant and wild-type plants all had a germination rate of approximately 85%, whereas 76.7% of the *abs stk* double mutant seeds germinated. The small difference in germination rate between the double mutant seeds and the other genotypes is most likely due to the fact that some of the mature *abs stk* double mutant seeds were blocked at a very late stage of development, and even though morphologically indistinguishable, should not be considered as mature seeds. We assume that those *abs stk* double mutant seeds that normally complete their development are able to germinate like wild-type seeds.

Table 2 Total number of mature seeds that developed and germination assay in wild-type, *abs*, *stk*, *abs/ABS stk/STK* and *abs stk*

	Total number of mature seeds (total number of ovules)	Percentage	Number of germinated seeds (total number of seeds)	Percentage
Wild-type	592 (614)	96.4	642 (733)	87.5
abs	548 (593)	92.4	504 (580)	86.9
stk	512 (560)	91.4	485 (592)	81.9
abs/ABS stk /STK	552 (570)	96.8	630 (726)	86.8
abs stk	120 (1092)	11	674 (879)	76.7

Loss of the endothelium layer in the *abs stk* double mutant causes fertilization defects

To further investigate the defects that lead to the observed ovule and seed phenotypes in the *abs stk* double mutant, detailed histological, genetic and molecular analyses were performed. In the wild-type (Figure 3a) and *stk* and *abs* single mutants (Figure S1), ovule integuments and seed coat are composed of five layers. The mature ovule integuments and seed coats of *abs stk* double mutants are composed of four layers. Morphological analysis based on the shape of the cells showed that the innermost layer, the endothelium, was missing from the inner integument (Figure 3b). In addition, the remaining four integument layers showed an irregular shape compared to wild-type and the single mutants (Figure S1). To investigate how the absence of the endothelium in *abs stk* ovules could cause ovule sterility, the fertilization process and female gametophyte formation were analysed in detail. In Arabidopsis, the mature embryo sac comprises seven cells: three antipodal cells, one diploid central cell, and one egg cell surrounded by two synergids (Mansfield *et al.*, 1991). The fertilization process begins with hydration of the pollen grains and subsequent germination on the stigmatic tissue. Subsequently, the pollen tube grows and is attracted to the ovule micropyle. After entering a synergid, the pollen tube bursts to release the two sperm cells: one fuses with the egg cell and the other with the central cell of the mature female gametophyte (Drews *et al.*, 1998).

In the *abs stk* double mutant, the ovules reach maturity (Figure 3b) as in wild-type (Figure 3a). To understand whether the observed ovule abortions are due to defects in determination of female gametophyte cell identity, embryo sac cell-specific reporter constructs were introduced into the *abs stk* double mutant. *abs stk* double mutants that are homozygous for an egg cell-specific marker (S. Sprunck, Department of Cell Biology and Plant Physiology, University of Regensburg, Germany, unpublished data), a synergid cell-specific marker (Gross-Hardt *et al.*, 2007), a central cell-specific marker (Moll *et al.*, 2008) or antipodal cell-specific marker (Matias-Hernandez *et al.*, 2010) were used to determine the correct identity of the cells in the female gametophytes (n = 250). These analyses showed that cell identity was not affected in *abs stk* double mutant gametophytes (Figure 3c–j). However, the antipodal cells in the *abs stk* double mutant (Figure 3h) appeared to be closer to the micropylar end than in wild-type ovules (Figure 3g). This phenotype is probably due to the reduced size of the embryo sac in the double mutant (Figure 3b, h). Another interesting observation is that, although the identity of the central cell was not affected, the two polar nuclei did not fuse in 15% of the observed *abs stk* ovules (n = 250) (Figure 3b, j), a phenotype that was never observed in wild-type embryo sacs at this stage (Figure 3i).

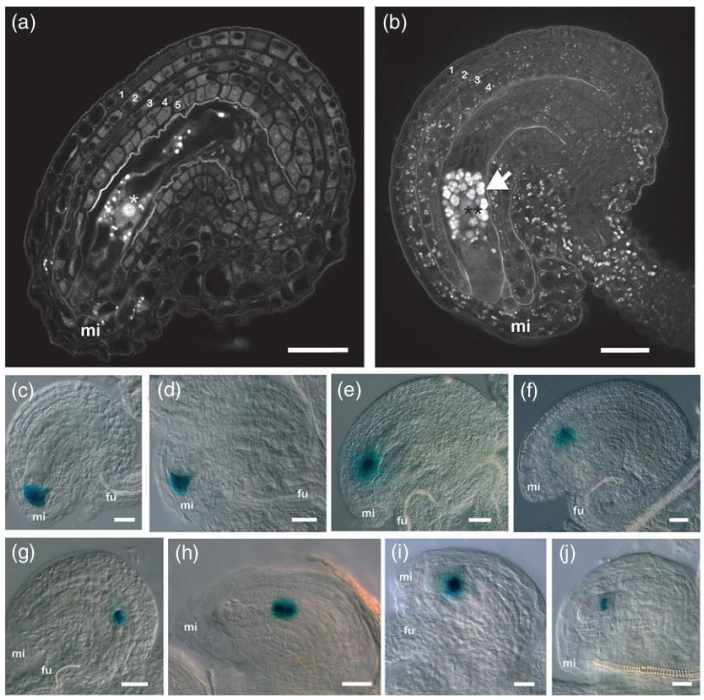


Figure 3. Morphological and molecular analysis of *abs stk* ovules.
(a, b) Confocal analysis of wild-type (a) and *abs stk* (b) ovules. The integument of the wild-type ovule (a) has five layers (1, epidermis; 2, palisade; 3 and 4, two inner integument layers; 5, endothelium), but the double mutant seed coat (b) has only four layers: the innermost layer (endothelium) is missing.
(c–j) Analysis of *abs stk* double mutant gametophyte cell identity. Crossing the *abs stk* double mutant with embryo sac specific reporter lines showed that female gametophyte cell identity was not altered in the double mutant.
(c,d) Synergid cell marker expression in wild-type ovules (c) and *abs stk* mutant ovules (d).
(e,f) Egg cell-specific marker expression in wild-type (e) and *abs stk* (f).
(g,h) Antipodal cell marker expression in wild-type ovules (g) and in *abs stk* double mutant ovules (h). There is a slight difference in the position of the antipodal cells: in the *abs stk* double mutant, the antipodal cells are closer to the micropylar region of the ovule.
(i,j) Central cell-specific marker expression in wild-type (i) and *abs stk* (j). Interestingly, the two polar nuclei do not always fuse.
mi, micropyle; fu, funiculus. The white asterisk in (a) indicates the central cell. The two black asterisks in (b) indicates unfused polar nuclei. The white arrow in (b) indicates starch. Scale bars = 30 μm (a,b) and 20 μm (c–j).

Morphological and cytological analysis of *abs stk* double mutant ovules indicated the presence of a large amount of starch in the embryo sac (Figures 3b and 4b). When we compared wild-type with *abs stk* double mutant ovules before fertilization, we observed that starch was mainly present at the micropylar region in wild-type embryo sacs (Figure 3a), as previously described by Mansfield *et al.* (1991), whereas a large amount of starch accumulated in the whole embryo sac in *abs stk* double mutant ovules (Figures 3b). After fertilization, starch was apparently used by the developing embryo and endosperm in wild-type ovules, and consequently the amount was drastically reduced (Figure 4e). In the *abs stk* double mutant, the number of ovules that showed a significant reduction in starch accumulation (Figure 4f) corresponded to the number of fertilized ovules (108/250), but the amount of starch was still high in the remaining ovules. For comparison, in the *abs* and *stk* single mutants, starch accumulation was similar to that in the wild-type (Figure S2).

To investigate whether pollen tubes were attracted by the ovules, and whether they correctly entered the micropyle, aniline blue staining of wild-type and *abs stk* mutant pistils was performed. This experiment showed that pollen tubes did indeed reach at all the *abs stk* double mutant ovules and entered the micropyle, indicating that this first part of the fertilization process was not affected (Figure 5a, b).

To analyse whether sperm cells were released into *abs stk* mutant embryo sac, *abs stk* carpels were pollinated with pollen from plants containing a *pLAT52:GUS* transgene (Tsukamoto *et al.*, 2010), which labels the pollen tube cytosol and allows investigation of pollen tube burst and release of the sperm cells (Figure 5c). This analysis confirmed that, as in the wild-type, pollen tube burst occurred in all *abs stk* ovules (Figure 5d). To further investigate the migration of the sperm cells in *abs stk* ovules, we used the *pHTR10:HTR10-RFP* reporter line (Aw *et al.*, 2010). In this reporter line, sperm cell nuclei are labelled with RFP. Migration of sperm cell nuclei in wild-type ovules was clearly observed (Figure 5e, f). Surprisingly, sperm nuclei were visible at the micropylar region after pollen tube burst in only 10% of *abs stk* mutant ovules. In all other *abs stk* mutant ovules, we were unable to observe the sperm cells after the pollen tube entered the micropyle (Figure 5g, h). The reason for this may be that the RFP signal was masked by the enormous amount of starch accumulated in *abs stk* ovules.

As pollen tubes do arrive at *abs stk* ovules, we analysed how many ovules were successfully fertilized using the *pMINI3:GUS* marker line. *MINI3* is expressed in the endosperm immediately after fertilization (Luo *et al.*, 2005). This experiment showed that only 43% of the ovules (n = 400) were fertilized in *abs stk* carpels (Figure 5i), whereas 83% of

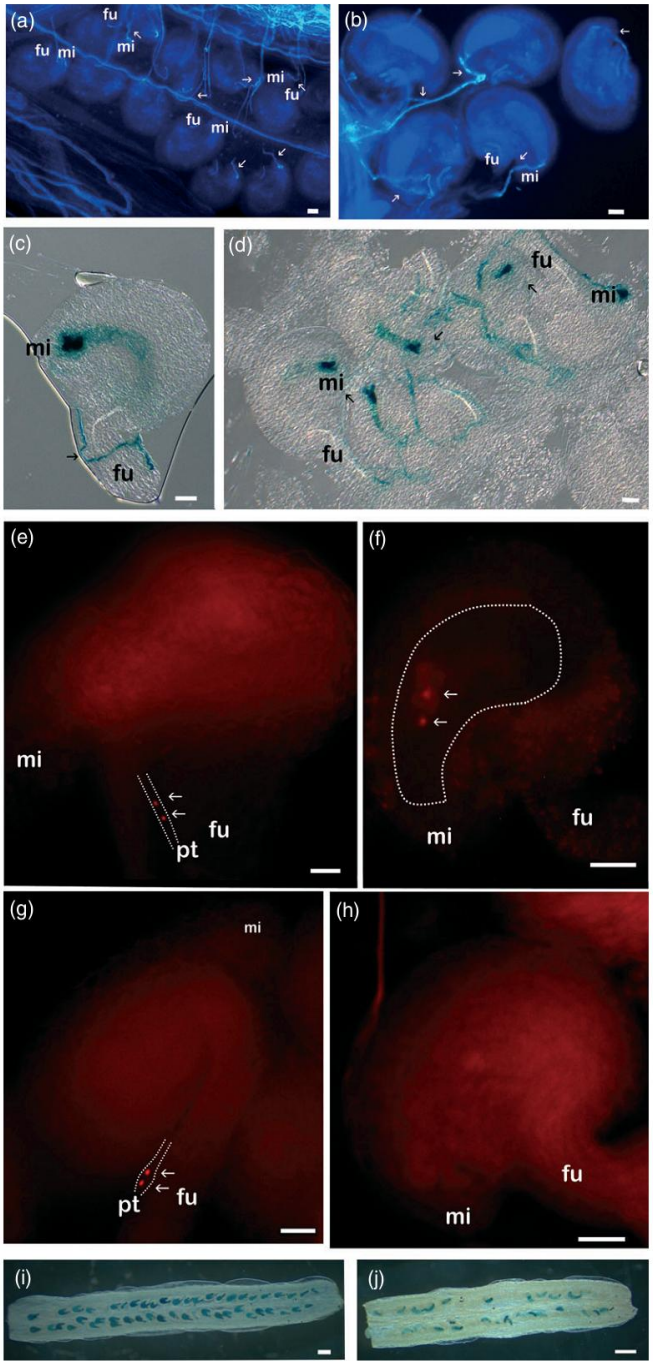


Figure 5. Pollen tube guidance, reception and burst, and sperm cell delivery/central cell fertilization.
(a,b) Pollen tube guidance and reception: pollen tubes (white arrows) arrive and enter the micropyle of wild-type (a) and *abs stk* double mutant (b) ovules.
(c,d) Fertilization of wild-type ovules (c) and *abs stk* double mutant ovules (d) with *LAT52:GUS* pollen showed that pollen tubes arrive at all the double mutant ovules (black arrows). In addition, the pollen tube burst is also evident close to the micropyle.
(e–h) After bursting of the pollen tube, sperm cell migration was followed using the *pHTR10:HTR10-RFP* marker line in wild-type (e,f) and double mutant ovules (g,h). It is possible to follow the sperm cells (white arrows) in the pollen tube (dotted line) in both wild-type (e) and the *abs stk* mutant (g). After arriving in the micropyle, the two sperm cells (white arrows) were visible inside the wild-type embryo sac (dotted line) (f), but they were almost never visible inside the double mutant embryo sac in the double mutant (h).
(i,j) Using the *pMINI3:GUS* marker line, 83% of the ovules in wild-type plants (n = 600) showed GUS expression (i), indicating that the central cell was fertilized and the triploid endosperm was formed. In the *abs stk* double mutant (j), only 43% of the ovules showed blue staining (n = 400).
pt, pollen tube; mi, micropyle; fu, funiculus. Scale bars = 20 μm (a–h) and 150 μm (i,j).

importance of sporophytic tissues for development of the haploid generation (reviewed by Bencivenga *et al.*, 2011). Interestingly, integument development was also affected in the *abs stk* double mutant. In this double mutant, the innermost integument layer, the endothelium, is missing. Therefore, this double mutant provides a useful tool to study

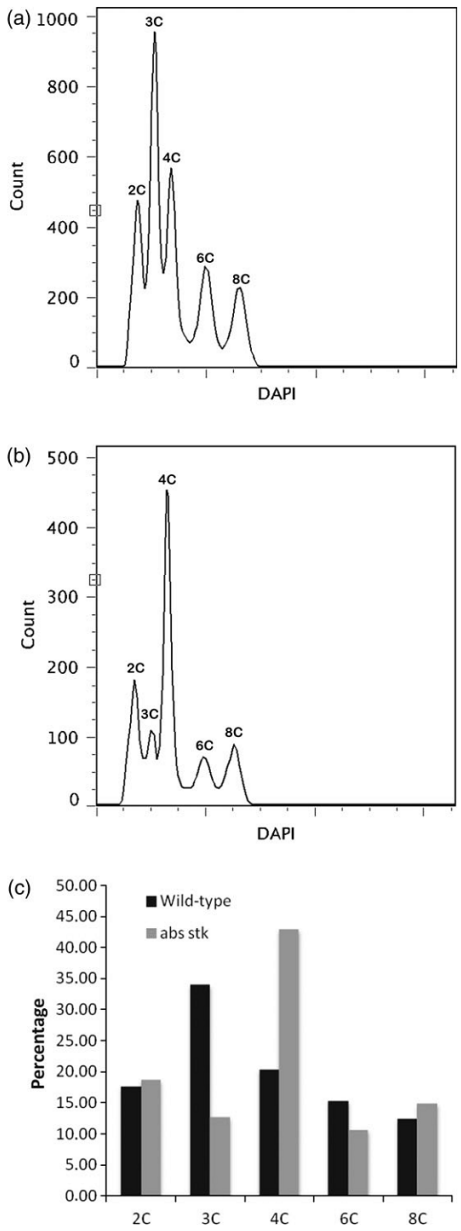


Figure 6. Flow cytometry ploidy analyses of the wild-type and *abs stk* double mutant.
(a,b) DNA profile of the wild-type (a) and *abs stk* double mutant (b), showing a mixture of diploid (2C, 4C), endo-replicated (8C, most likely starting from a diploid DNA content) and triploid (3C, 6C) nuclei. The presence of triploid nuclei in developing *abs stk* seeds, although severely reduced in comparison to wild-type, suggests that at least a proportion of the seeds derive from a successful double fertilization.
(c) Quantification of diploid (2C, 4C), endo-replicated (8C, most likely starting from a diploid DNA content) and triploid (3C, 6C) nuclei.

the function of the endothelium during seed development. Interestingly, ABS and STK appear to act redundantly in the formation of the endothelium even though they belong to different clades of the MADS box phylogenetic tree. Redundancy among distantly related MADS box genes has been described previously: for example, *AGAMOUS-LIKE 24* (*AGL24*) and *SHORT VEGETATIVE PHASE* (*SVP*) are redundant with *APETALA1* (*AP1*) during early stages of flower development (Gregis *et al.*, 2006, 2008). Comparing the primary amino acid sequences of these proteins does not explain their redundancy, but redundancy might be due to the similarity in interactions they can make with other MADS domain factors (De Folter *et al.*, 2005).

Previously, it was shown that co-suppression of two *STK* orthologues, *FBP7* and *FBP11*, in *Petunia hybrida* resulted in degeneration of the endothelium (Colombo *et al.*, 1997), leading to the formation of seeds with a reduced amount of endosperm. Genetic analysis confirmed that this seed phenotype was completely sporophytically controlled, as shown here for the *abs stk* double mutant. The importance of the endothelium in ovule development and fertilization was also suggested by characterization of the Arabidopsis *pdi5* mutant (Andème Ondzighi *et al.*, 2008). This mutant does not produce the disulfide isomerase PDI-5, which is involved in the timing of the programmed cell death in the endothelium. In the *pdi5* mutant, the endothelium cells prematurely degenerate, resulting in partial loss of fertility (Andème Ondzighi *et al.*, 2008).

Our studies using various gametophyte cell identity marker lines showed that cell identity was correctly determined in the embryo sac of the *abs stk* double mutant. Furthermore, we did not observe any arrest in gametophyte development as described for mutants with more severe integument defects. This suggests that the endothelium is not necessary for the progression of female gametophyte development. However, some effects were visible, such as the smaller size of the embryo sac and occasionally unfused central cell polar nuclei.

It has been suggested that an auxin gradient controls the identity of the gametophyte cells in the embryo sac (Pagnussat *et al.*, 2009). In this respect, it is interesting to note that cell identity was not affected in the *abs stk* double mutant female gametophyte, despite the drastic reduction in embryo sac size and massive starch accumulation. These factors may interfere with creation of the postulated auxin gradient. Future experiments using auxin-responsive reporter lines in *abs stk* double mutants may clarify this. However, the absence of the endothelium has a clear effect on fertilization of the ovules. In the *abs stk* double mutant, only 45% of the ovules were fertilized. This reduction in fertilization is probably not due to defects in fusion of the polar nuclei, as polar nuclei fusion is also affected in the *ig binding protein* (*bip1*) and *bip2* double mutant, without any effect on ovule fertilization (Maruyama *et al.*, 2010). In *Saccharomyces*

cerevisiae, *BiP* genes have been shown to encode molecular chaperone Hsp70 proteins that are located in the endoplasmatic reticulum, which regulate nuclear membrane fusion during mating. In the Arabidopsis *bip1 bip2* double mutant, the sperm nucleus fuses with only one of the two central nuclei. This causes the arrest of endosperm development and leads to seed abortion. Therefore, the seed abortions observed in the *abs stk* mutant may partly correspond to ovules in which polar fusion did not occur. However, as only 15% of the polar nuclei did not fuse, the polar nuclei phenotype could not completely explain the seed abortion phenotype (75% abortion among the fertilized ovules) observed in the *abs stk* double mutant. Furthermore, as reported in Table 1, we observed a block in seed development after the bent cotyledon stage (13 DAP). This *abs stk* double mutant phenotype could also not be explained by the formation of a diploid endosperm since in the *bip1 bip2* double mutant, seeds never reached the bent cotyledon stage (Maruyama *et al.*, 2010).

The reduced fertilization efficiency in the *abs stk* mutant is not due to defects in pollen tube guidance as aniline blue staining of pollinated *abs stk* carpels showed that the pollen tubes arrived at the micropyle of all the ovules. Furthermore, pollen tube burst was observed in all ovules, indicating that sperm cells were correctly released. However, we could not further investigate the fertilization process because we observed sperm nuclei at the micropylar region in only 10% of the double mutant ovules. In all other ovules, no sperm nuclei were visible using the *pHTR10:HTR10-RFP* reporter line. Thus, at this point, we can only speculate about the possible reasons for these observations. It may be that sperm cells degenerated or did not properly fuse with the female gametes. However, as 43% of the ovules are fertilized, it is more likely that sperm cells migrate to the egg and central cell, but the RFP signal is masked by the enormous starch accumulation observed in *abs stk* double mutant ovules. This hypothesis is supported by the fact that we observed a few times a very faint RFP signal from the nuclei in the ovule. Our flow cytometry data also suggest that fertilization of the central cell takes place in the *abs stk* double mutant. Interestingly, the flow cytometry analysis suggested that ABS and STK are also required for timely entry into mitosis. In relation to this, it is interesting to observe that the founding member of the MADS box gene family, MINI-CHROMOSOME MAINTENANCE 1 (MCM1) in yeast (and homologues in other organisms), is a key regulator of the cell cycle (Passmore *et al.*, 1988). In fact, MCM1 has been shown to control the expression of the mitosis-promoting B-type cyclins CLB1 and CLB2. Consequently, cell cycle progression is compromised in *mcm1* mutant cells which accumulate in the G_2 phase (Althoefer *et al.*, 1995), similar to what was observed for the *abs stk* double mutant. Moreover, another MADS box gene, *XAANTAL1* (XAL1)/*AGAMOUS-LIKE 12* (*AGL12*), is also involved in cell-cycle regulation, as the mutant has a significantly longer cell cycle duration (Tapia-Lopez *et al.*, 2008). This suggests a potential role for plant MADS box genes, including *ABS* and *STK*, in control of cell-cycle progression. This may therefore represent a mechanism by which cell proliferation and differentiation are linked (Harashima and Schnittger, 2010).

In the *abs stk* mutant background, loss of endothelium correlates with starch accumulation. We propose that the endothelium plays a fundamental role in regulating glucose metabolism. The reason for the reduced fertilization may be at least partially linked to this massive accumulation of starch, which could form a barrier to proper migration of the sperm cells to the egg and central cell. Carbohydrate metabolism in developing gametophytes is poorly understood, but it strongly influences plant fitness. Mutants with altered sucrose metabolism have recently been isolated that show defects during female gametophyte development and reduced fertility, for instance the *cytosolic phosphoglucomutase 2* (*cpm2*) *cpm3* double mutant (Egli *et al.*, 2010) and the *sucrose transporter 2* (*suc2*) mutant (Gottwald *et al.*, 2000). Moreover, the carbohydrate transporter gene *SUC1* was shown to be expressed in the female gametophyte (Feuerstein *et al.*, 2010). Interestingly, the Arabidopsis *GLUCOSE-6-PHOSPHATE/PHOSPHATE TRANSLOCATOR 1* (*GPT1*) gene was shown to be important for male and female gametophyte development (Niewiadomski *et al.*, 2005). In the ovaries of plants heterozygous for the *gpt1* mutation, 36% of the embryo sacs were defective in polar nuclear fusion, similar to what has been described for the *gametophytic factor 2* (*gfa2*) mutant (Christensen *et al.*, 2002; Niewiadomski *et al.*, 2005). In the *gpt1* mutant, de-regulation of metabolic pathways is probably the cause of gametophyte defects. These examples indicate that tight regulation of metabolic pathways is required in the female gametophyte, and disturbance of these causes gametophyte defects.

In conclusion, we have identified a new function for the MADS domain protein-encoding genes *ABS* and *STK*. Despite the fact that these genes belong to different clades of the MADS box phylogenetic tree, they redundantly regulate endothelium formation, as the complete absence of endothelium is detectable only in the *abs stk* double mutant and not in *abs* and *stk* single mutants. Our results show that fertilization is under sporophytic control, and that the endothelium is an important component contributing to ovule fertility. Furthermore, we also show that the endothelium is important during seed development, as it is probably essential for regulating the flux of nutrients, as evidenced by the massive accumulation of starch in the embryo sac of the *abs stk* double mutant. All these data indicate an essential role of maternal tissues for correct and balanced nutrition of the next generation that develops within the seeds.

EXPERIMENTAL PROCEDURES

Plant material and growth conditions

Arabidopsis thaliana wild-type (ecotype Columbia) and transgenic plants were grown at 22°C under short-day conditions (8 h light/16 h dark) and long-day conditions (16 h light/8 h dark). The *stk-2* allele contains a 74 nt insertion near the splice site of the 3rd intron (Pinyopich *et al.*, 2003). The *abs-6* allele contains a T-DNA insertion in the 3′ region of the 1st intron, 51 nt upstream of the start of the 2nd exon (De Folter *et al.*, 2006).

PCR-based genotyping

Identification of the *STK* wild-type and mutant alleles was performed by PCR analysis using oligonucleotides AtP204 (5′-GCT TGTTCTGATAGCACCAACACTAGCA-3′) and AtP561 (5′-GGAACT CAAAGAGTCTCCCATCAG-3′). The mutant allele gives a 399 bp DNA fragment, and the wild-type allele gives a 325 bp DNA fragment.

For identification of the *ABS* wild-type allele, oligonucleotides AtP3077 (5′-ATGGGTAGAGGGAAGATAGAGATAAAGAA-3′) and AtP3078 (5′-TTAATCATTCTGGGCCGTTGGATCGTTTT-3′) were used. The wild-type allele gives a 1899 bp DNA fragment. Identification of the *abs* mutant allele was performed using oligonucleotides AtP3079 (5′-TTTCTCCATATTGACCATCATACTCATTG-3′) and AtP3078 (5′-TTAATCATTCTGGGCCGTTGGATCGTTTT-3′). The mutant allele gives a 1200 bp DNA fragment.

Expression analysis by *in situ* hybridization

DIG-labelled RNA probes for detection and hybridization of *ABS* and *STK* were prepared as described by Ambrose *et al.* (2000). Sections of plant tissue were hybridized with digoxigenin-labelled *ABS* antisense probe, amplified using primers TT16For (5′-ATGA CGACCAGGAGCAATTG-3′) and TT16Rev (5′-AGAACTCAAGTGCAT TGTGC-3′), and with digoxigenin-labelled *STK* antisense probe, amplified using primers AtP2563 (5′-ATCTAAGAACTAAGGTAGC AGAAGT-3′) and AtP2564 (5′-TCTTAATCATTACACAACACAAA TTGC-3′).

Confocal laser scanning analysis

For integument morphological analysis, plants were emasculated, and 24 h after emasculation, pistils were fixed as described by Braselton *et al.* (1996). The samples were excited using a 532 nm laser, and emission was detected between 570 and 740 nm. The samples were observed using a Leica SP5 confocal laser scanning microscope (http://www.leica.com/).

Pollen tube guidance, reception and burst analysis

For *in vivo* pollen tube guidance experiments, pistils were emasculated and pollinated after 24 h with wild-type pollen. After 16–18 h, pistils were carefully isolated from the plants and fixed in a solution of acetic acid and absolute ethanol (1:3), cleared with 8 N sodium hydroxide and labelled with aniline blue (Sigma, http://www.sigmaaldrich.com/).

For *in vivo* pollen tube reception and burst of the tubes, wild-type and double mutant pistils were emasculated and crossed after 24 h with *pLAT52:GUS* pollen. After 16–18 h, the pistils were carefully collected and stained for GUS activity (Liljegren *et al.*, 2000). Samples were incubated in clearing solution (Brambilla *et al.*, 2007), dissected under a Leica MZ6 stereo microscope, and observed using a Zeiss Axiophot D1 microscope equipped with differential interference contrast (DIC) optics (http://www.zeiss.com/). Images were captured using an Axiocam MRc5 camera (Zeiss) with AXIOVISION software (version 4.1).

Sperm cell migration analysis

For sperm cell migration experiments, pistils were emasculated and crossed after 24 h with the *pHTR10:HTR10-RFP* marker line. Pistils were collected after 16–18 h, samples were dissected under a Leica MZ6 stereo microscope, and images were obtained using a Zeiss Axiophot D1 microscope equipped with DIC optics and a rhodamine filter set.

Cytological analysis

For cytological analysis of ovules, Lugol staining was performed. Pistils were isolated from closed and open flowers, fixed and cleared with the same solutions used for aniline blue staining, and labelled with Lugol's solution. Stained pistils were observed using a Zeiss Axi-overt 200 DIC microscope.

The gametophytic cell identity reporter lines used in here encode a nuclear localization signal that is in-frame with the GUS reporter gene. The egg cell-specific marker was kindly provided by Stefanie Sprunck (unpublished data). The synergid cell-specific marker was kindly provided by Ueli Grossniklaus (Institute of Plant Biology, University of Zurich, Switzerland) (Gross-Hardt *et al.*, 2007). The central cell-specific marker was kindly provided by Rita Gross-Hardt (Department of Developmental Genetics, University of Tübingen, Germany) (Moll *et al.*, 2008). The antipodal cell-specific marker, kindly provided by Rita Gross-Hardt, was generated as described by Yu *et al.* (2005): the promoter of At1g36340 was amplified using primers 5′-AGTGAGGCGCGCCTGATCATTAAGTTTAGGGGT-3′ and 5′-AGTGATTAATTAATTACGAGAAATCACCAAAC-3′, and cloned upstream from the NLS_GUS reporter into pGIIBar binary vector (Gross-Hardt *et al.*, 2007) (cloning details are available upon request).

For female gametophyte cell identity determination, marker lines were used as female and pollinated with *abs stk* pollen. In the F_1 generation, heterozygous plants were self-fertilized, and the presence/absence of the mutant alleles for *abs* and *stk* was analysed in the F_2 generation by PCR.

The presence of the reporter genes was analysed by GUS staining, confirming the correct expression in wild-type background. For GUS staining, flowers were emasculated and harvested 12 h after pollination as described by Liljegren *et al.* (2000). Samples were incubated in chloral hydrate:glycerol:water solution 8:1:2, dissected and observed using a Zeiss Axiophot D1 microscope equipped with DIC optics.

Central cell fertilization and ploidy analyses

To confirm the fertilization of the central cell, *abs stk* double mutant pistils were pollinated with *pMINI3:GUS* pollen (Luo *et al.*, 2005). After 48 h, pistils were collected and GUS staining was performed as described by Liljegren *et al.* (2000). Samples were incubated in clearing solution, dissected and observed using a Zeiss Axiophot D1 microscope equipped with DIC optics.

For flow cytometry analysis, seeds at 6 DAP were taken from the siliques, and crushed in 2 ml test tubes with nuclear extraction buffer (CyStain UV-precise kit, Partec, http://www.partec.com/). All samples were subsequently stained with nuclear staining solution (CyStain UV-precise kit) containing 4′,6-diamidino-2-phenylindole (DAPI) as previously described (Nowack *et al.*, 2007). The ploidy level, represented by the mean peak position in a DAPI fluorescence intensity histogram, was calibrated against the 2C nuclear DNA content peak derived from a preparation of young rosette leaves

(Nowack *et al.*, 2007). Data were presented using CYFLOW PLOIDY ANALYSER software (Partec).

ACCESSION NUMBERS

The GenBank/EMBL accession numbers for *ARABIDOPSIS* B_{sister} and *SEEDSTICK* are, At5g23260 and At4g09960, respectively.

ACKNOWLEDGEMENTS

We would like to thank Martin Yanofsky (Section of Cell and Developmental Biology, University of California at San Diego, CA, USA) for providing the *stk-2* mutant, Richard Immink (Plant Research International, Wageningen University & Research Centre, The Netherlands) for the *abs-6* mutant, Rita Gross-Hardt (Department of Developmental Genetics, University of Tübingen, Germany) for the central cell marker line and the antipodal marker line, Stefanie Sprunck (Department Cell Biology and Plant Physiology, University of Regensburg, Germany) for the egg cell marker line, Ueli Grossniklaus (Institute of Plant Biology, University of Zurich, Switzerland) for the synergid cell marker line, Ming Luo (Commonwealth Scientific and Industrial Research Organization, Plant Industry Division, Australia) for p*MINI3*:GUS, Ravishankar Palanivelu (School of Plant Science, University of Arizona, AZ, USA) for p*LAT52*:GUS, and Frédéric Berger (Department of Biological Sciences, University of Singapore, Singapore) for the p*HTR10*:*HTR10*-RFP marker. We also thank Xin'Ai Zhao (Institut de Biologie Moléculaire des Plantes du Centre National de la Recherche Scientifique, Strasbourg, France) for her help with flow cytometry. For the *STK in situ* hybridization analysis, we would like to thank Valeria Gagliardini (Institute of Plant Biology, University of Zurich, Switzerland). We also thank Barbara Ambrose (Assistant Curator of Plant Genomics, New York Botanical Garden, NY) for her help with the *ABS in situ* hybridization experiments. Finally, thanks also to Sonia Sacchi, Antonella Stillitano and Lara Grasseni for help with GUS experiments. This work was supported by an EU-ITN fellowship (ITN-SYSFLO project) to M.A.M. and a fellowship from the Università degli Studi di Milano to C.M. This work was also supported by grant 'Action Thématique et Incitative sur Programme' from the Centre National de la Recherche Scientifique and a European Research Council Starting Independent Researcher grant to A.S.

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

Figure S1. Morphological analysis of integuments.

Figure S2. Lugol staining before and after fertilization in wild-type, *abs stk*, *abs* and *stk* mutants.

Table S1. Reciprocal crosses between wild-type and the *abs stk* mutant.

Please note: As a service to our authors and readers, this journal provides supporting information supplied by the authors. Such materials are peer-reviewed and may be re-organized for online delivery, but are not copy-edited or typeset. Technical support issues arising from supporting information (other than missing files) should be addressed to the authors.

REFERENCES

Althoefer, H., Schleiffer, A., Wassmann, K., Nordheim, A. and Ammerer, G. (1995) Mcm1 is required to coordinate G_2-specific transcription in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **15**, 5917–5928.

Ambrose, B.A., Lerner, D.R., Ciceri, P., Padilla, C.M., Yanofsky, M.F. and Schmidt, R.J. (2000) Molecular and genetic analyses of the *Silky1* gene reveal conservation in floral organ specification between eudicots and monocots. *Mol. Cell* **5**, 569–579.

Andème Ondzighi, C., Christopher, D.A., Cho, E.J., Chang, S.C. and Staehelin, L.A. (2008) *Arabidopsis* protein disulfide isomerase-5 inhibits cysteine proteases during trafficking to vacuoles before programmed cell death of the endothelium in developing seeds. *Plant Cell*, **20**, 2205–2220.

Aw, S.J., Hamamura, Y., Chen, Z., Schnittger, A. and Berger, F. (2010) Sperm entry is sufficient to trigger division of the central cell but the paternal genome is required for endosperm development in *Arabidopsis*. *Development*, **137**, 2683–2690.

Baker, S.C., Robinson-Beers, K., Villanueva, J.M., Gaiser, J.C. and Gasser, C.S. (1997) Interactions among genes regulating ovule development in *Arabidopsis thaliana*. *Genetics*, **145**, 1109–1124.

Battaglia, R., Brambilla, V. and Colombo, L. (2008) Morphological analysis of female gametophyte development in the *bel1 stk shp1 shp2* mutant. *Plant Biosyst.* **142**, 643–649.

Bencivenga, S., Colombo, L. and Masiero, S. (2011) Cross talk between the sporophyte and the megagametophyte during ovule development. *Sex. Plant Reprod.* **24**, 113–121.

Brambilla, V., Battaglia, R., Colombo, M., Masiero, S., Bencivenga, S., Kater, M.M. and Colombo, L. (2007) Genetic and molecular interactions between BELL1 and MADS box factors support ovule development in *Arabidopsis*. *Plant Cell*, **19**, 2544–2556.

Braselton, J.P., Wilkinson, M.J. and Clulow, S.A. (1996) Feulgen staining of intact plant tissues for confocal microscopy. *Biotech. Histochem.* **71**, 84–87.

Christensen, C.A., King, E.J., Jordan, J.R. and Drews, G.N. (1997) Megagametogenesis in *Arabidopsis* wild type and the *Gf* mutant. *Sex. Plant Reprod.* **10**, 49–64.

Christensen, C.A., Gorsich, S.W., Brown, R.H., Jones, L.G., Brown, J., Shaw, J.M. and Drews, G.N. (2002) Mitochondrial GFA2 is required for synergid cell death in Arabidopsis. *Plant Cell*, **14**, 2215–2232.

Colombo, L., Franken, J., Van der Krol, A.R., Wittichy, P.E., Donsy, H.J.M. and Angenent, G.C. (1997) Downregulation of ovule-specific MADS box genes from petunia results in maternally controlled defects in seed development. *Plant Cell*, **9**, 703–715.

Colombo, L., Battaglia, R. and Kater, M.M. (2008) *Arabidopsis* ovule development and its evolutionary conservation. *Trends Plant Sci.* **13**, 444–450.

De Folter, S., Immink, R.G., Kieffer, M. *et al.* (2005) Comprehensive interaction map of the Arabidopsis MADS box transcription factors. *Plant Cell*, **17**, 1424–1433.

De Folter, S., Shchennikova, A.V., Franken, J., Busscher, M., Baskar, R., Grossniklaus, U., Angenent, G.C. and Immink, R.G. (2006) A B_{sister} MADS-box gene involved in ovule and seed development in petunia and Arabidopsis. *Plant J.* **47**, 934–946.

Debeaujon, I., Nesi, N., Perez, P., Devic, M., Grandjean, O., Caboche, M. and Lepiniec, L. (2003) Proanthocyanidin-accumulating cells in Arabidopsis testa: regulation of differentiation and role in seed development. *Plant Cell*, **15**, 2514–2531.

Drews, G.N., Lee, D. and Christensen, C.A. (1998) Genetic analysis of female gametophyte development and function. *Plant Cell*, **10**, 5–17.

Egli, B., Kölling, K., Köhler, C., Zeeman, S.C. and Streb, S. (2010) Loss of cytosolic phosphoglucomutase compromises gametophyte development in Arabidopsis. *Plant Physiol.* **154**, 1659–1671.

Elliott, R.C., Betzner, A.S., Huttner, E., Oakes, M.P., Tucker, W.Q., Gerentes, D., Perez, P. and Smyth, D.R. (1996) *AINTEGUMENTA*, an APETALA2-like gene of *Arabidopsis* with pleiotropic roles in ovule development and floral organ growth. *Plant Cell*, **8**, 155–168.

Feuerstein, A., Niedermeier, M., Bauer, K., Engelmann, S., Hoth, S., Stadler, R. and Sauer, N. (2010) Expression of the *AtSUC1* gene in the female gametophyte, and ecotype-specific expression differences in male reproductive organs. *Plant Biol.* **12**, 105–114.

Gottwald, J.R., Krysan, P.J., Young, J.C., Evert, R.F. and Sussman, M.R. (2000) Genetic evidence for the *in planta* role of phloem-specific plasma membrane sucrose transporters. *Proc. Natl Acad. Sci. USA*, **97**, 13979–13984.

Gregis, V., Sessa, A., Colombo, L. and Kater, M.M. (2006) *AGL24*, *SHORT VEGETATIVE PHASE*, and *APETALA1* redundantly control *AGAMOUS* during early stages of flower development in *Arabidopsis*. *Plant Cell*, **18**, 1373–1382.

Gregis, V., Sessa, A., Colombo, L. and Kater, M.M. (2008) *AGAMOUS-LIKE24* and *SHORT VEGETATIVE PHASE* determine floral meristem identity in Arabidopsis. *Plant J.* **66**, 891–902.

Gross-Hardt, R., Kägi, C., Baumann, N., Moore, J.M., Baskar, R., Gagliano, W.B., Jürgens, G. and Grossniklaus, U. (2007) LACHESIS restricts gametic cell fate in the female gametophyte of Arabidopsis. *PLoS Biol.* **5**, 47–53.

Harashima, H. and Schnittger, A. (2010) The integration of cell division, growth and differentiation. *Curr. Opin. Plant Biol.* **13**, 66–74.

Haughn, G.W. and Chaudhury, A. (2005) Genetic analysis of seed coat development in *Arabidopsis*. *Trends Plant Sci.* **10**, 472–477.

Hsu, S.C., Belmonte, M.F., Harada, J.J. and Inoue, K. (2010) Indispensable roles of plastids in *Arabidopsis thaliana* embryogenesis. *Current Genomics*, **11**, 338–349.

Kaufmann, K., Anfang, N., Saedler, H. and Theissen, G. (2005) Mutant analysis, protein–protein interactions and subcellular localization of the Arabidopsis B_{sister} (ABS) protein. *Mol. Gen. Genomics.* **274**, 103–118.

Klucher, K.M., Chow, H., Reiser, L. and Fischer, R.L. (1996) The *AINTEGUMENTA* gene of *Arabidopsis* required for ovule and female gametophyte development is related to the floral homeotic gene *APETALA2*. *Plant Cell*, **8**, 137–153.

Liljegren, S., Ditta, G., Eshed, Y., Savidge, B., Bowman, J. and Yanofsky, M. (2000) *SHATTERPROOF* MADS-box genes control seed dispersal in *Arabidopsis*. *Nature*, **404**, 766–769.

Losa, A., Colombo, M., Brambilla, V. and Colombo, L. (2010) Genetic interaction between *AINTEGUMENTA* (*ANT*) and the ovule identity genes *SEEDSTICK* (*STK*), *SHATTERPROOF1* (*SHP1*) and *SHATTERPROOF2* (*SHP2*). *Sex. Plant Reprod.* **23**, 115–122.

Luo, M., Dennis, E.S., Berger, F., Peacock, W.J. and Chaudhury, A. (2005) *MINISEED3* (*MINI3*), a *WRKY* family gene, and *HAIKU2* (*IKU2*), a leucine-rich repeat (*LRR*) *KINASE* gene, are regulators of seed size in *Arabidopsis*. *Proc. Natl Acad. Sci. USA*, **102**, 17531–17536.

Mansfield, S.G., Briarty, L.G. and Erni, S. (1991) Early embryogenesis in *Arabidopsis thaliana*. I. The mature embryo sac. *Can. J. Bot.* **69**, 447–460.

Maruyama, D., Endo, T. and Nishikawa, S. (2010) BiP-mediated polar nuclei fusion is essential for the regulation of endosperm nuclei proliferation in *Arabidopsis thaliana*. *Proc. Natl Acad. Sci. USA*, **107**, 1684–1689.

Matias-Hernandez, L.M., Battaglia, R., Galbiati, F., Rubes, M., Eichenberger, C., Grossniklaus, U., Kater, M.M. and Colombo, L. (2010) *VERDANDI* is a direct target of the MADS domain ovule identity complex and affects embryo sac differentiation in *Arabidopsis*. *Plant Cell*, **22**, 1702–1715.

Modrusan, Z., Reiser, L., Feldmann, K.A., Fischer, R.L. and Haughn, G.W. (1994) Homeotic transformation of ovules into carpel-like structures in *Arabidopsis*. *Plant Cell*, **6**, 333–349.

Moll, C., von Lyncker, L., Zimmermann, S., Kägi, C., Baumann, N., Twell, D., Grossniklaus, U. and Gross-Hardt, R. (2008) CLO/GFA1 and ATO are novel regulators of gametic cell fate in plants. *Plant J.* **56**, 913–922.

Morley-Smith, E.R., Pike, M.J., Findlay, K., Köckenberger, W., Hill, L.M., Smith, A.M. and Rawsthorne, S. (2008) The transport of sugars to developing embryos is not via the bulk endosperm in oilseed rape seeds. *Plant Physiol.* **147**, 2121–2130.

Nesi, N., Debeaujon, I., Jond, C., Stewart, A.J., Jenkins, G.I., Caboche, M. and Lepiniec, L. (2002) The *TRANSPARENT TESTA16* locus encodes the ARABIDOPSIS BSISTER MADS domain protein and is required for proper development and pigmentation of the seed coat. *Plant Cell*, **14**, 2463–2479.

Niewiadomski, P., Silke Knappe, S., Geimer, S., Fischer, K., Schulz, B., Unte, U.S., Rosso, M.G., Ache, P., Flugge, U.-I. and Schneider, A. (2005) The Arabidopsis plastidic glucose 6-phosphate/phosphate translocator GPT1 is essential for pollen maturation and embryo sac development. *Plant Cell*, **17**, 760–775.

Nowack, M.K., Shirzadi, R., Dissmeyer, N., Dolf, A., Endl, E., Grini, P.E. and Schnittger, A. (2007) Bypassing genomic imprinting allows seed development. *Nature*, **447**, 312–315.

Nowack, M.K., Ungru, A., Bjerkan, K.N., Grini, P.E. and Schnittger, A. (2010) Reproductive cross-talk: seed development in flowering plants. *Biochem. Soc. Trans.* **38**, 604–612.

Pagnussat, G.C., Alandete-Saez, M., Bowman, J.L. and Sundaresan, V. (2009) Auxin-dependent patterning and gamete specification in the Arabidopsis female gametophyte. *Science*, **26**, 1684–1689.

Passmore, S., Maine, G.T., Elble, R., Christ, C. and Tye, B.K. (1988) *Saccharomyces cerevisiae* protein involved in plasmid maintenance is necessary for mating of MATα cells. *J. Mol. Biol.* **204**, 593–606.

Pinyopich, A., Ditta, G.S., Savidge, B., Liljegren, S.J., Baumann, E., Wisman, E. and Yanofsky, M.F. (2003) Assessing the redundancy of MADS-box genes during carpel and ovule development. *Nature*, **424**, 85–88.

Reiser, L., Modrusan, Z., Margossian, L., Samach, A., Ohad, N., Haughn, G.W. and Fischer, R.L. (1995) The *BELL1* gene encodes a homeodomain protein involved in pattern formation in the *Arabidopsis* ovule primordium. *Cell*, **83**, 735–742.

Robinson-Beers, K., Pruitt, R.E. and Gasser, C.S. (1992) Ovule development in wild-type Arabidopsis and two female-sterile mutants. *Plant Cell*, **4**, 1237–1249.

Schneitz, K., Hülskamp, M. and Pruitt, R.E. (1995) Wild-type ovule development in *Arabidopsis thaliana*: a light microscope study of cleared whole-mount tissue. *Plant J.* **7**, 731–749.

Schneitz, K., Hulskamp, M., Kopczak, S.D. and Pruitt, R.E. (1997) Dissection of sexual organ ontogenesis: a genetic analysis of ovule development in *Arabidopsis thaliana*. *Development*, **124**, 1367–1376.

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

Stadler, R., Lauterbach, C. and Sauer, N. (2005) Cell-to-cell movement of green fluorescent protein reveals post-phloem transport in the outer integument and identifies symplastic domains in Arabidopsis seeds and embryos. *Plant Physiol.* **139**, 701–712.

Tapia-Lopez, R., Garcia-Ponce, B., Dubrovsky, J.G., Garay-Arroyo, A., Perez-Ruiz, R.V., Kim, S.H., Acevedo, F., Pelaz, S. and Alvarez-Buylla, E.R. (2008) An *AGAMOUS*-related MADS-box gene, *XAL1* (*AGL12*), regulates root meristem cell proliferation and flowering transition in Arabidopsis. *Plant Physiol.* **146**, 1182–1192.

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. *Plant J.* **62**, 571–588.

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

Supporting Information

Figure S1

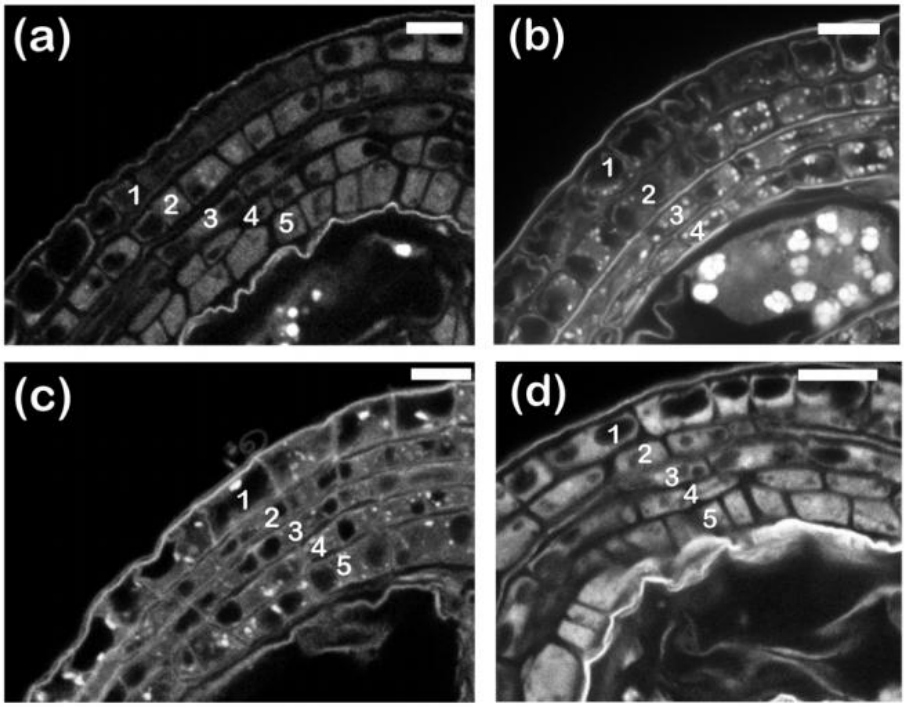


Figure S1. Morphological analysis of integuments. Wild- type (a), *abs* (c) and *stk* ovules integuments (d) are composed of five layers. The *abs stk* integuments (b) are composed of four layers: the missing layer is the endothelium. In the double mutant also the remaining four integument layers showed an irregular shape when compared to wild-type and to the single mutants. Scale bars: 10 µm.

Figure S2

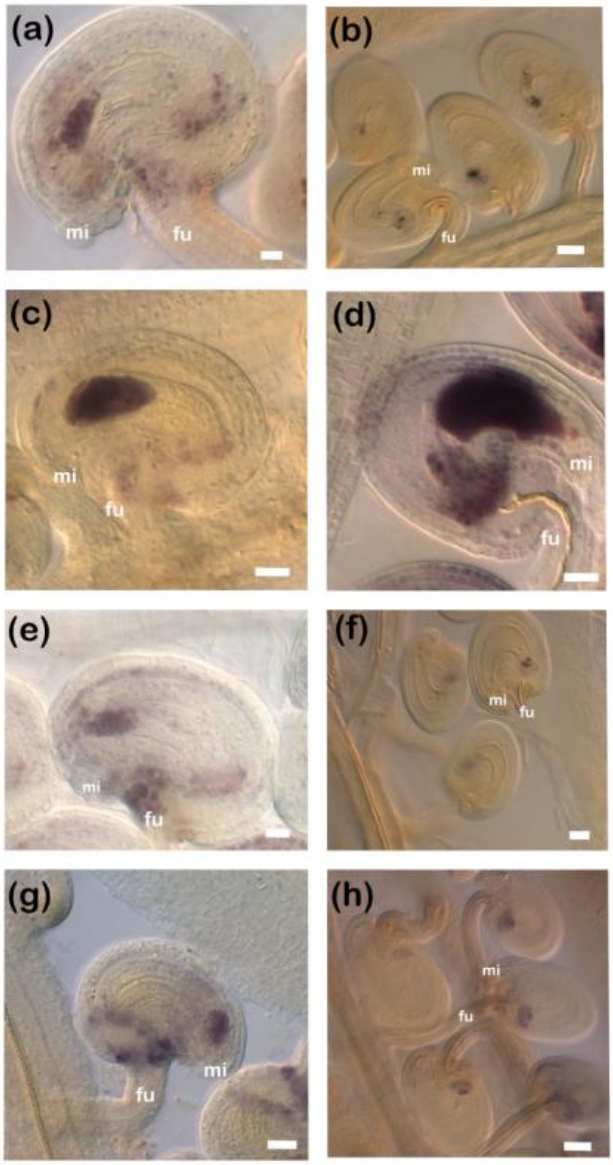


Figure S2. Lugol staining before and after fertilization in wild-type, *abs stk*, *abs* and *stk* mutants. Wild-type (a), *abs stk* (c), *abs* (e) and *stk* (g) ovules before fertilization. All the ovules showed starch accumulation. After fertilization there is no starch accumulation detectable in wild-type (b), *abs* (f) or *stk* (h) ovules, whereas in the *abs stk* double mutant (d) the starch accumulation is still visible as before. (mi, micropyle; fu, funiculus). Scale bars: 20 µm.

Table S1

Table S1.
Reciprocal crosses between wild type and *abs stk* mutant. Ovule abortions, seed abortions and developing seeds analysed at 5 DAP.

	Total ovules	Ovule abortions (Percentage)	Seed abortions (Percentage)	Developing seeds (Percentage)
WT X *abs stk*	166	10 (6%)	8 (5%)	148 (89%)
abs stk X WT	326	171 (52%)	45 (14%)	110 (34%)

Table S1. Reciprocal crosses between wild-type and *abs stk* mutant. Ovule abortions, seed abortions and developing seeds analysed at 5 DAP.

Part III

Controllo molecolare dello sviluppo dell'ovulo

Published
Accademia dei Georgofili

Molecular Analysis of the Double Fertilization Process in Arabidopsis

Marta Adelina Mendes*, Beatrice Castelnovo*, Lucia Colombo*

Controllo molecolare dello sviluppo dell'ovulo

introduzione

Gli ovuli sono strutture molto specializzate costituiti da tre regioni morfologicamente e funzionalmente distinte: la nucella, la calaza e il funicolo. La nucella funziona da megasporangio e al suo interno si sviluppa la cellula madre delle megaspore che in seguito al processo di sporogenesi e gametogenesi formerà il gametofito femminile. La calaza costituisce la porzione centrale del primordio e da questa si originano uno o due tegumenti che circondano e racchiudono il gametofito femminile mantenendo una piccola apertura, il micropilo entro cui penetrerà il tubetto pollinico durante il processo di fecondazione. Infine, l'ovulo è connesso alla parete dell'ovario tramite il funicolo (Yadegari and Drews, 2004).

In Arabidopsis, quando il gametofito femminile raggiunge la maturità è composto da sette cellule: la cellula uovo, che è fiancheggiata dalle due cellule sinergidi, la cellula centrale diploide e tre cellule antipodali (fig. 1A). Al contrario, il granulo di polline consiste di sole tre cellule: una cellula vegetativa responsabile della formazione del tubetto pollinico, e due cellule germinative (fig. 1B). La formazione della generazione sporofitica successiva dipende dall'interazione tra il gametofito femminile e quello maschile.

Nelle Angiosperme il processo di fecondazione coinvolge entrambe le due cellule spermatiche, una feconda la cellula uovo per formare lo zigote da cui si svilupperà l'embrione, mentre l'altra feconda la cellula centrale diploide per dare origine all'endosperma, che protegge l'embrione e gli fornisce nutrimento durante il suo sviluppo (Sundaresan et al., 2010).

* *Dipartimento di BioScienze, Università degli Studi di Milano*

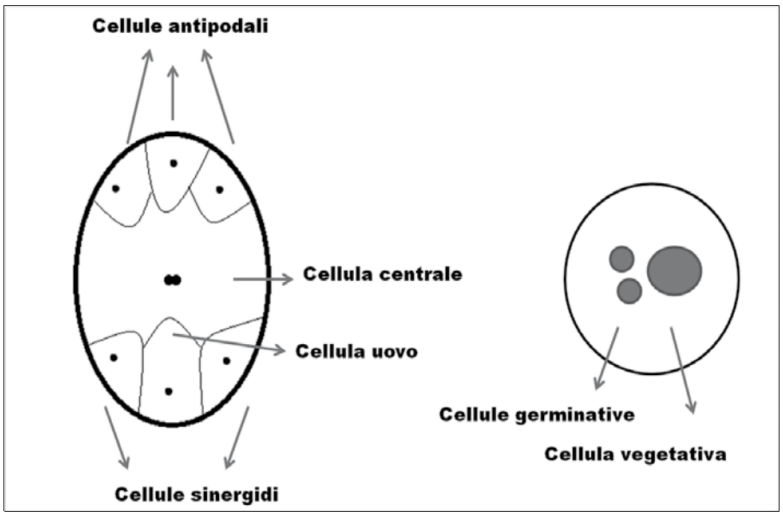


Fig. 1 *(A) Rappresentazione schematica del sacco embrionale e (B) di un granulo pollinico di «Arabidopsis thaliana»*

È noto che i membri della famiglia di fattori trascrizionali MADS-box svolgano un ruolo cruciale nello sviluppo del fiore. I geni MADS-box codificano per fattori di trascrizione che presentano tutti lo stesso dominio di legame con il DNA. Il dominio MADS (acronimo derivato dai nomi dei geni in cui esso è stato individuato per la prima volta), riconosce sempre una particolare sequenza di DNA: la CArG-box, caratterizzata dalla presenza del motivo $CC[A/T]_6GG$ In Arabidopsis sono stati identificati più di 100 geni MADS-box, che controllano funzioni fondamentali dello sviluppo quali la formazione dei fiori, e lo sviluppo dei semi e la transizione fiorale.

È stato inoltre dimostrato che i fattori MADS-box agiscono formando complessi multimerici come, ad esempio i fattori STK, SHP1, SHP2 e SEPALLATA (SEP) che promuovono lo sviluppo dell'ovulo e del carpello (Favaro et al., 2002).

Recentemente sono stati resi noti dati d'immunoprecipitazione della cromatina che hanno evidenziato come i fattori di trascrizione di tipo MADS-box si leghino a centinaia di regioni nel genoma di Arabidopsis. Questi fattori sembrano agire come regolatori globali dell'espressione genica durante i vari stadi di sviluppo del fiore (Smaczniak et al., 2011, Kaufmann et al., 2009).

Per esempio STK regola la trascrizione di geni, non solo coinvolti nella

determinazione dell'identità dell'ovulo, ma anche coinvolti nel metabolismo delle cellule di cui determina l'identità (Mizzotti et al., 2012). In questo manoscritto si presentano i risultati ottenuti relativamente alla caratterizzazione di uno dei target di STK, *VERDANDI*, un gene codificante per un fattore di trascrizione appartenente alla superfamiglia B3, coinvolto nella determinazione dell'identità delle sinergidi e della loro funzionalità (Matias-Hernandez et al., 2010).

RISULTATI

VDD è necessario per l'apoptosi delle sinergidi

Il pattern di espressione di *VDD* è stato studiato mediante esperimenti di real-time PCR e di ibridazione *in situ*. Queste analisi hanno evidenziato che *VDD* è espresso nelle infiorescenze e nel meristema fiorale. Durante la formazione dell'ovulo, *VDD* è espresso in tutti gli stadi dello sviluppo.

Allo scopo di studiare la funzione del gene *VDD*, è stato studiato l'allele mutante *vdd-1*; l'analisi di segregazione nella progenie di piante eterozigoti dei *VDDvdd-1* ha mostrato un rapporto di segregazione 1:1 con totale assenza di individui omozigote per l'allele *vdd-1*. Le silique di piante eterozigoti *VDDvdd-1* mostrano circa il 50% di ovuli non fecondati (Matias-Hernandez et al., 2010).

Per esaminare se il difetto di fecondazione osservato fosse dovuto a un difetto durante lo sviluppo del sacco embrionale, sono state eseguite analisi mediante microscopia a contrasto interferenziale su piante eterozigoti *VDDvdd-1*. L'analisi di ovuli maturi non fecondati ha evidenziato che tutti gli ovuli in piante eterozigoti *VDDvdd-1* sono morfologicamente indistinguibili dagli ovuli di piante wild type (fig. 2).

Ulteriori analisi per mezzo di incroci di piante *VDDvdd-1* con linee marker specifiche per i diversi tipi di cellule del gametofito, hanno mostrato che le identità della cellula centrale e della cellula uovo sono specificate come nel wild type, mentre circa metà dei gametofiti delle piante *VDDvdd-1* presentano cellule sinergidi e antipodali non si differenziate correttamente (Mattias-Hernandez et al, 2010).

Per verificare che la perdita dell'identità delle sinergidi comprometta l'attrazione del tubetto pollinico i pistilli di piante *VDDvdd-1* sono stati impollinati manualmente e in seguito colorati mediante blue di anilina per eviden-

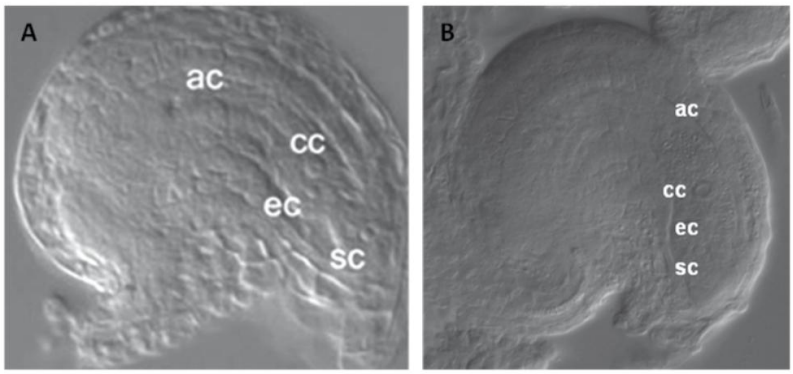


Fig. 2 *(A) Ovulo maturo di una pianta wild-type. (B) Ovulo maturo di pianta vdd-1*

ziare i tubetti pollinici. Inaspettatamente i tubetti pollinici arrivano a tutti gli ovuli delle piante *VDDvdd-1* ed entrano correttamente a livello del micropilo (fig. 3).

L'analisi morfologica delle fasi successive ha permesso di verificare che nelle piante *VDDvdd-1* solo la metà degli ovuli presenta degenerazione delle sinergide in seguito alla penetrazione del tubetto pollinico, evento necessario al fine di portare a termine il processo di doppia fecondazione (fig. 4). Questo dato suggerisce che VDD sia coinvolto nel processo di apoptosi della sinergide in seguito all'entrata del tubetto pollinico nel sacco embrionale.

DISCUSSIONE

L'importanza dei fattori di trascrizione MADS-box nei differenti stadi di sviluppo della pianta è stata discussa ampiamente negli anni passati. In questo lavoro è stata evidenziata l'importanza del complesso STK-SEP3 nella regolazione di un processo molto specifico ma estremamente importante per il processo riproduttivo.

In particolare grazie a questo studio è stato possibile comprendere la grande importanza rivestita dal complesso STK-SEP3 durante il processo di doppia fecondazione in *Arabidopsis thaliana*. Questo complesso regola direttamente *VDD*, un membro della famiglia B3, che gioca un ruolo essenziale nel processo di fecondazione.

VDD è coinvolto nella specificazione dell'identità e nella funzionalità delle sinergidi. Nel mutante *VDDvdd-1* le cellule sinergidi nel 50% dei gametofiti hanno

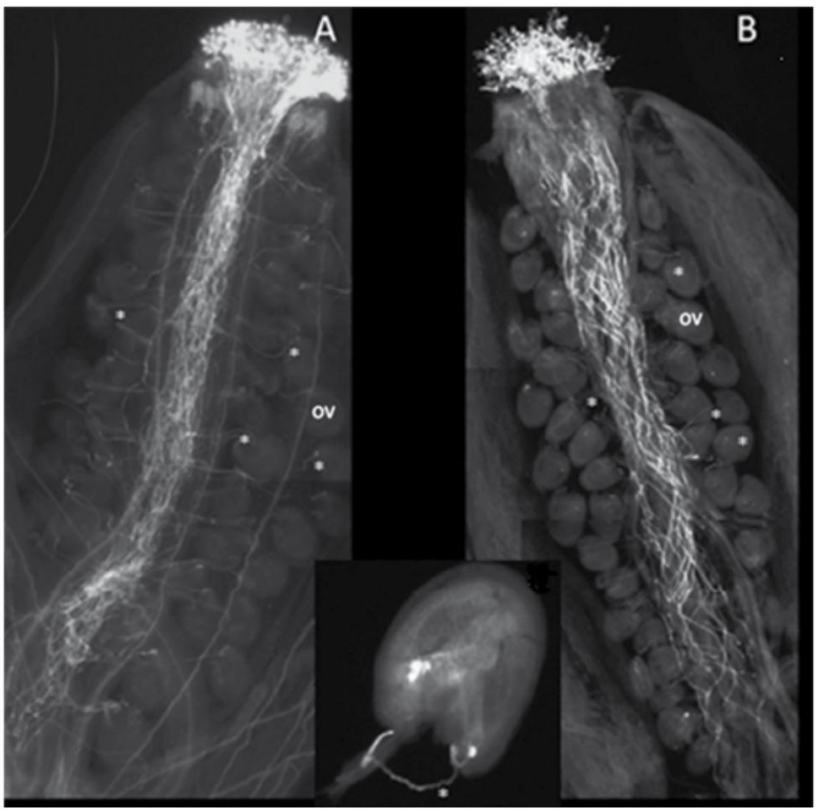


Fig. 3 *Nelle piante wild-type (A) e nelle piante vdd-1 (B) il tubetto pollinico (*) arriva a tutti gli ovuli. (C) Nel dettaglio, è possibile vedere un ovulo wild-type e il tubetto pollinico. Abbreviazioni: ov, ovulo*

perso la loro identità ma questo non ha compromesso l'attrazione del tubetto pollinico; questi risultati suggeriscono che l'attrazione del tubetto pollinico e la degenerazione delle sinergidi sono due processi completamente separati anche se entrambi regolati dalle sinergidi (fig. 5). Questo è un chiaro esempio di come i complessi di fattori MADS box siano coinvolti non solo nella determinazione dell'identità degli organi, ma anche nel regolare il funzionamento degli organi stessi.

A conferma di questo, recentemente si è scoperto che molti dei target diretti dei complessi costituiti da questi fattori di trascrizione sono, geni che codificano per enzimi fondamentali del metabolismo cellulare (Kaufmann et al., 2010; Kaufmann et al., 2009).

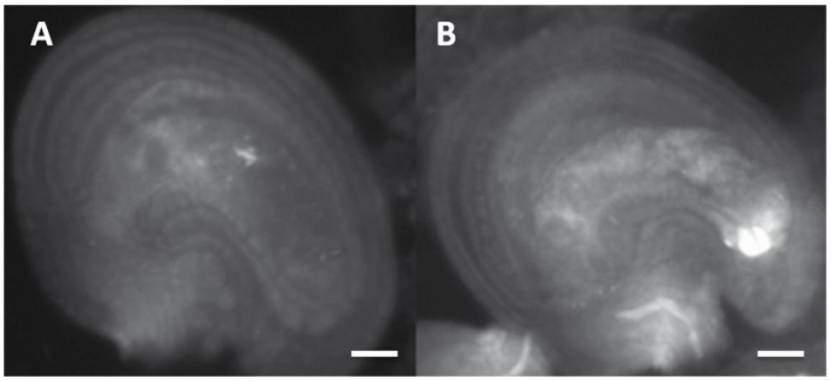


Fig. 4 *(A) Ovulo wild-type prima della degenerazione (B) Ovulo wild-type dopo la degenerazione*

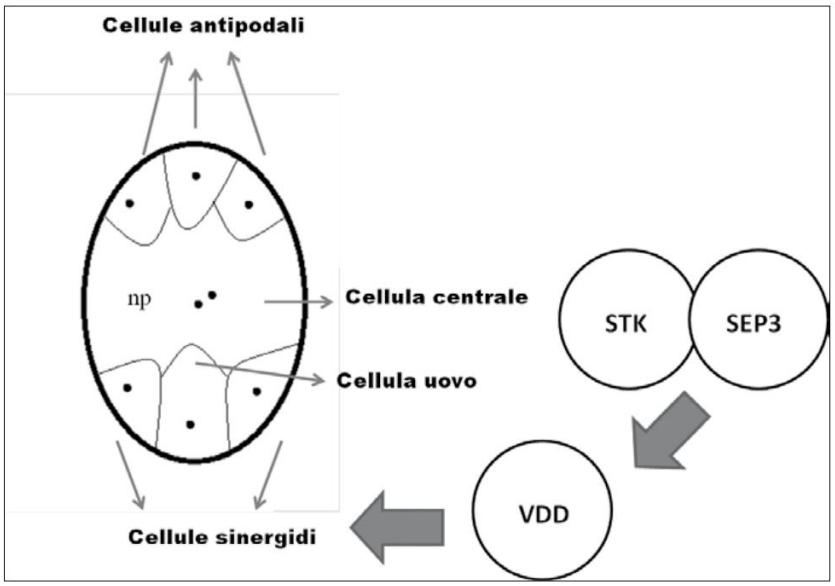


Fig. 5 *Il complesso composto dalle proteine SEEDSTICK (STK) e SEPPALLATA3 (SEP3) regola direttamente il gene VERDANDI (VDD) il quale svolge un ruolo essenziale nello sviluppo delle cellule sinergidi*

MATERIALI E METODI

Materiale vegetale e crescita delle piante

Piante wild-type di *Arabidopsis thaliana* ecotipo Columbia e piante transgeniche sono state cresciute a 22° C in condizioni di fotoperiodo lungo (16h di luce/ 8h di buio). Le linee marker per le cellule del gametofito sono quelle descritte in Matias Hernandez et al., 2010.

Tutte le linee marker per il gametofito femminile codificano per un segnale di localizzazione nucleare che è in frame con il gene reporter GUS. Il mutante *VDDvdd-1* proviene dalla collezione Syngenta Arabidopsis Insertion Library (SAIL 50_C03) e contiene l'inserzione di un T-DNA nel primo introne, 44pb a monte del secondo esone.

Analisi delgenotipo

L'identificazione dell'allele mutante *VDDvdd-1* è stata eseguita mediante analisi PCR usando i primers5'-GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC-3(TDNA)e 5'-CGAAGGAGAGAAGCAGAGATG-3'. L'allele wild-type di *VDD* è stato identificato usando il primer 5'-TGAAGTACCGGCTTCAGAGTC-3'.

Analisi di microscopia

Per analizzare lo sviluppo dell'ovulo in piante eterozigoti *vdd-1* fiori a differenti stati di sviluppo sono stati posti nella soluzione di clearing per una notte usando una soluzione composta da 160 g di cloro idrato (C-8383; Sigma-Aldrich), 100mL di acqua e 50 mL di glicerolo. Gli ovuli sono stati osservati mediante un microscopio Zeiss Axiophot D1.

Per la colorazione mediante blu di anilina, piante eterozigoti *VDDvdd-1* sono state emasculate e impollinate 24h dopo l'emasculazione. Dopo, 16-18 h i pistilli sono stati raccolti e fissati in una soluzione composta da acido acetico e etanolo assoluto (1:3), in seguito trattati con NaOH 8N e colorati con blu di anilina (Sigma).

Per l'analisi di degenerazione della cellula sinergide, piante wild type eterozigoti *VDDvdd-1* sono state impollinate con una linea marker specifica per il tubetto pollinico e delle cellule spermatiche (contenente il transgene

pAKV::H2B-GFP). Dopo 16h i pistilli sono stati raccolti e osservati al microscopio ottico Zeiss®Axiophot D1 10X, 20X, 40X e 100X a immersione munito di obiettivo a contrasto interferenziale. Le fotografie sono state scattate utilizzando una camera Zeiss Axiocam MRc5 e il programma Axiovision (versione 4.1).

RIASSUNTO

Negli ultimi anni si è compiuto uno sforzo notevole per capire quali siano i fattori che controllano lo sviluppo dell'ovulo e gli eventi che precedono e seguono il processo di doppia fecondazione, necessari alla formazione dei semi. Gli aspetti più interessanti, da un punto di vista applicativo sono il controllo del numero degli ovuli all'interno dell'ovario e la possibilità di avere un'efficiente fecondazione anche in condizioni di stress ambientali che normalmente influiscono negativamente sull'efficienza del processo stesso. La scelta di numerosi gruppi di ricerca è stata quella di focalizzarsi sullo studio dei processi di riproduzione in specie modello quali Arabidopsis o riso, tenendo conto dell'elevato grado di conservazione dei meccanismi molecolari che stanno alla base dello sviluppo degli organi riproduttivi e del processo di fecondazione.

In particolare in Arabidopsis sono stai identificati i geni che controllano lo sviluppo dell'ovulo e del gametofito femminile. I fattori trascrizionali appartenenti alla famiglia MADS-box quali STK, SHP1, SHP2 e SEP. formano un complesso multimerico in grado di regolare centinaia di geni targets coinvolti nelle diverse fasi di sviluppo dell'ovulo e del seme di Arabidopsis. Uno di questi target è *VERDANDI(VDD)*un fattore che svolge un ruolo essenziale nello sviluppo del gametofito femminile ed è coinvolto nella determinazione dell'identità delle cellule accessorie del sacco embrionale (cellule sinergidi e antipodali).

ABSTRACT

In the last years an effort has been made to understand the regulatory network that controls ovule development and the fertilization process. The events that occur before and after the double fertilization are extremely important because they assure seed set and thereby the next plant generation. One of the most important economical aspects of this process is the control of the number of ovules that are produced in each carpel and the possibility to optimize the fertilization process. The choice of various research groups to use model species like Arabidopsis or rice, is due to the fact that the molecular mechanisms controlling the fertilization process seems to be widely conserved in the plant kingdom and therefore the use of these model species provides a benefit to study this process.

In Arabidopsis three MADS-box genes, *SEEDSTICK (STK*), *SHATTERPROOF1* (*SHP1*) and *SHP2,* redundantly control ovule development. Furthermore, genetic and protein interaction studies have shown that these ovule identity factors are able to interact with the SEPALLATA (SEP) MADS domain factors. The interaction between these MADS proteins has shown to be essential for ovule development. Our lab identified the first target gene of the STK-SEP3 complex, which is *VERDANDI (VDD)*, a putative

transcription factor that belongs to the poorly characterized REM family. *VDD* has a role during female gametophyte development and is involved in the determination of the identity of the embryo sac accessory cells, which are important in the fertilization process.

BIBLIOGRAFIA

Brambilla V., Battaglia R., Colombo M., Masiero S., Bencivenga S., Kater M.M., and Colombo L. (2007): *Genetic and molecular interactions between BELL1 and MADS-box factors support ovule development in Arabidopsis*, «Plant Cell», 19, pp. 2544-2556.

Favaro R., Immink R.G.H., Ferioli V., Bernasconi B., Byzova M., Angenent G.C., Kater M.M., Colombo L. (2002): *Ovule-specific MADS-box proteins have conserved protein-protein interactions in monocot and dicot plants*, «Mol. Genet. Genomics», 268, pp. 152-159.

Kaufmann K., Muiño J.M., Jauregui R., Airoldi C.A., Smaczniak C., Krajewski P., Angenent G.C. (2009): *Target genes of the MADS transcription factor SEPALLATA3: integration of developmental and hormonal pathways in the Arabidopsis flower*, «PLoS Biology», 21, 7 (4).

Kaufmann K., Wellmer F., Muiño J.M., Ferrier T., Wuest S. E., Kumar V., Serrano-Mislata A., Madueño F., Krajewski P., Meyerowitz E. M., Angenent G.C., Riechmann J. L. (2010): *Orchestration of Floral Initiation by APETALA1*, «Science», 238.

Matias-Hernandez L., Battaglia R., Galbiati F., Rubes M., Eichenberger C., Grossniklaus U., Kater M.M., Colombo L. (2010): *VERDANDI Is a Direct Target of the MADS Domain Ovule Identity Complex and Affects Embryo Sac Differentiation in Arabidopsis*, «The Plant Cell», 22, pp. 1702-1715.

Mizzotti C., Mendes M.A., Caporali E., Schnittger A., Kater M.M., Battaglia R., Colombo L. (2012): *The MADS-box genes SEEDSTICK and ARABIDOPSIS BSISTER play a maternal role in fertilization and seed development*, «The Plant Journal», 70, pp. 409-420.

Nurrish S.J., Treisman R. (1995): *DNA binding specificity determinants in MADS-box transcription factors*, «Mol. Cell. Biol», 15, pp. 4076-85.

Smaczniaka C., Imminkb R., Muiñom J., Blanvillaind R., Busscherb M., Busscher-LangebDinhb J.P., Liuh S., Westphali A., Boereni S., Parcy F., Xuh L., Cristel C., Angenent G. and Kaufmann K. (2012): *Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development*, PNAS 1560-1565.

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

Yadegari R., Drews G.N. (2004): *Female gametophyte development*, «The Plant Cell», 16, pp. 133-141.

Acknowledgments

I want to dedicate this thesis to all the people that make me smile every day, that supported me and went with me in this long journey.
I want to start by thanking the Scientific family that helped me obtaining all the beautiful, important and wonderful results!
Lucia for all the guidance, discussion, sustain and for being a true scientific mother to me. Martin and Simona for all the discussion and advices, for being there every time I needed. To all the members of the magnificent Lab 6 former and actual for making our Lab the best in the world! I want to thank to Irma for being there every time I needed help, company, advices and to make my arriving and adaptation to Italy much more easy (Grazie Tia!).
I want to thank to my Italian teachers Chiara and Dario, not only for being amazing teachers but also for the companionship! Francesca REM for passing me all the information about REM super family, for the dancing moments in lab. To the Brazilian girl Larissa, to the Spanish girl Paloma for the social life that they offer me, and off course for all the scientific advices!
To my new sister Francesca Resentini! I want to thank her for everything that I could think and even more! For the scientific discussions, for the company, for the Italian teaching, for the road trips… For being a true sister to me when my family was not around!
I want to thank to the Spanish community for all the god moments, Vicent, Sergi, Nacho!
To my students, I really need to thank them for all the work that they help me to finish!!! To Beatrice, Marco and Chiara!
To all my Lab colleges Mara, that for sure we will have a future together in the citokinin world! Goyo for the all the lab talks! To my Portuguese home mates and lab companions Ana Marta and Sofia, for all the laughs and for talking to me in

Portuguese, it is very important when you are surrounding by Italians!!!! To all my friends and collaborators in Brasil Sarah and Marcio.
To all the former members in the lab, Fiorella for the scientific work. Dola for all the discussion and company. I want also to thank to the 5th floor and all the members of the Kater's group.
Least but not the last I want to thank to my parents, otherwise I wouldn't be here! For all the love and support! To my sister Ricardina for all the love and care. To my closest friends, that I consider as my sisters Silvia, Rute and Sofia, just for existing in my life, making it wonderful! To Walter for making me believe in myself every time that I didn't have the strength to do it anymore, for making me happy and more positive every day!

Thank you