First posted online on 6 November 2020 as 10.1242/dev.194274
Access the most recent version at http://dev.biologists.org/lookup/doi/10.1242/dev.194274

The RNA dependent DNA methylation pathway is required to restrict *SPOROCYTELESS/NOZZLE* expression to specify a single female germ cell precursor in Arabidopsis

Marta A. Mendes[1], Rosanna Petrella[1], Mara Cucinotta[1], Edoardo Vignati[1], Stefano Gatti[1], Sara C. Pinto[4], Dayton C. Bird[3], Veronica Gregis[1], Hugh Dickinson[2], Matthew R. Tucker[3], Lucia Colombo[1*]

[1] Dipartimento di Bioscienze, Università degli Studi di Milano, Via Celoria 26, 20133 Milano, Italy
[2] Department of Plant Sciences, University of Oxford, South Parks Road, Oxford OX1 3RB, UK
[3] School of Agriculture, Food, and Wine, The University of Adelaide, Waite Campus, Urrbrae, SA 5064, Australia
[4]Departamento de Biologia, Faculdade de Ciências da Universidade do Porto, 4169-007 Porto.
*Corresponding Author: lucia.colombo@unimi.it.

Abstract

In higher plants the female germline is formed from the megaspore mother cell (MMC), a single cell in the pre-meiotic ovule. Previously, it was reported that mutants in RNA dependent DNA methylation - RdDM - pathway might be involved in restricting female germline to a single nucellus cell. We show that the DRM methyltransferase double mutant drm1drm2 also presents also presents ectopic enlarged cells, consistent with supernumerary MMC-like cells. In wild-type ovules MMC differentiation required SPOROCYTELESS/NOZZLE (SPL/NZZ), as demonstrated by the spl/nzz mutant failing to develop an MMC. We address the poorly understood upstream regulation of SPL/NZZ in ovules, showing that the *RdDM* pathway is important to restrict SPL/NZZ expression. In *ago9, rdr6* and *drm1drm2* mutants, SPL/NZZ is ectopically expressed, suggesting that the mutiple MMC-like cells observed might be due to the ectopic expression of SPL/NZZ. We show that the ovule identity gene SEEDSTICK, directly regulates the *AGO9* and *RDR6* expression in the ovule and therefore indirectly SPL/NZZ expression. A model is presented describing the network required to restrict SPL/NZZ expression to specify a single MMC.

Development • Accepted manuscript

Introduction

Plants alternate a sporophytic, diploid phase with a highly reduced gametophytic, haploid phase in which the gametes are formed. The female germline in angiosperms initiates in the first phase of ovule development through the differentiation of a distal sub-epidermal cell named the megaspore mother cell (MMC), that quickly becomes morphologically distinguishable from the surrounding sporophytic cells by its size. The MMC undergoes meiosis to form four haploid megaspores that, in Arabidopsis, are typically arranged in a linear tetrad. Only one of the four megaspores survives to form the functional megaspore (FM), while the remaining three degenerate through programmed cell death. Following megasporogenesis the FM follows the *Polygonum*-type pattern of megagametogenesis involving two mitoses without cytokinesis, resulting in a four-nucleate syncytium with two nuclei at each pole, and a third mitosis after which cell plates are formed between all eight nuclei. Prior to cell plate formation, one of the nuclei from each pole migrates toward the centre of the developing embryo sac and fuses with the other to form the homodiploid nucleus of the central cell. The result is a seven-celled structure consisting of three antipodal cells in the chalazal pole, one diploid central cell, one egg cell and two synergid cells in the micropylar pole. The mature embryo sac is surrounded by two sporophytic cell layers termed integuments (Christensen et al., 1997; Yadegari and Drews, 2004).

At the molecular level, the early phases of ovule development are controlled by a number of transcription factors, hormones and small RNA-related pathways (reviewed in Pinto et al. 2019)). Ovule identity in Arabidopsis is redundantly controlled by the MADS-box transcription factors SEEDSTICK (STK) and SHATTERPROOF 1 and 2 (SHP1, SHP2) (Favaro et al., 2003; Pinyopich et al., 2003; Brambilla et al., 2007). STK is expressed in many sporophytic cell types including the nucellus, chalaza and integuments, and has been shown to play key roles in multiple stages of ovule and seed development (Matias-Hernandez et al., 2010; Mizzotti et al., 2012; Mendes et al., 2013; Mizzotti et al., 2014; Mendes et al., 2016; Balanzà et al., 2016; Ezquer et al., 2016; Cucinotta et al., 2016; Herrera-Ubaldo et al., 2019; Di Marzo et al., 2020; Petrella et al., 2020).

After ovule initiation, the SPL/NZZ transcription factor was previously described to be essential for MMC differentiation, since in the *spl/nzz* mutant the large majority of ovules (~99%) do not form an MMC (Yang et al., 1999; Schiefthaler et al., 1999). In contrast to *SPL*/*NZZ*, genes involved in the RNA Dependent DNA methylation (*RdDM*) pathway appear to limit the differentiation of multiple MMC-like cells in the pre-meiotic ovule. In Arabidopsis, DNA methylation is initially catalysed by a 24-nucleotide (nt) siRNA-dependent *RdDM* pathway involving ARGONAUTE (AGO) proteins and DOMAINS REARRANGED METHYLASEs (DRM1 and DRM2). CG and CHG methylation are then maintained by DNA METHYLTRANSFERASE1 (MET1) and CHROMOMETHYLASE3 (CMT3), respectively (Law and Jacobsen 2010). CHH methylation is maintained either through the *RdDM*

pathway (Matzke et al., 2015)or by the CMT2 DNA methyltransferase (Zemach et al., 2013; Stroud et al., 2014). Involvement of *RdDM* in cell patterning in the developing ovule was proposed by Olmedo-Monfil et al. (2010) who showed that, in Arabidopsis*, ARGONAUTE9* (*ago9-2*) and *RNA-DEPENDENT RNA POLYMERASE6* (*rdr6-11*) mutants show additional MMC-like cells in pre-meiotic ovules (47% and 46% respectively vs 8% in wild-type).

Here we confirmed the role of methylation in preventing the formation of multiple MMC-like cells, by showing that the *drm1drm2* double mutant displays a similar phenotype to that of *ago9-2* and *rdr6-11* lines. Furthermore, we explored the possible mechanism by which the *RdDM* pathway is linked to MMC specification. Surprisingly, investigation of MMC specification and development in the *stk* mutant background revealed 46% of ovules to contain two or more putative MMC-like cells. Quantitative-PCR confirmed that *RDR6* and *AGO9* expression was down-regulated while *SPL/NZZ* was up-regulated in *stk* ovules. Furthermore, chromatin immunoprecipitation experiments (ChIP) using a *STK_GFP* fusion protein showed *AGO9* and *RDR6* to be direct targets of STK. Remarkably, expression of the functional *SPL_GFP* fusion protein was found to be restricted to a few cells of the L1 layer at the tip of ovule primordium in the wild-type nucellus, but was ectopically expressed in the L1 layer of mutants representing the two major effectors of *RdDM*, *ago9* and *drm1drm2*. Thus, *SPL/NZZ* is expressed in L1 layer but is required for MMC differentiation in L2, so the pathway controlled by SPL/NZZ acts in a non-cell autonomous manner and the *RdDM* pathway regulates its expression. A model describing the control of female germline precursor specification in the Arabidopsis pre-meiotic ovule is proposed, bringing together data both from this study and previous investigations.

Results

drm1drm2 and *stk* mutants show additional MMC-like cells

Differentiation of the MMC is of pivotal importance for the progression of female germline development and specification of a single MMC is under a strict molecular control. In the *ago9-2* mutant it has been described that around 47% of the ovules develop more than one MMC-like cell (Olmedo-Monfil et al., 2010). It is known that the *RdDM* pathway can regulate gene expression via AGO-mediated mRNA degradation, or cytosine DNA methylation of target genes by DRM1/DRM2 methylases (Matzke and Mosher, 2014). Therefore, we analysed MMC differentiation in the double *drm1drm2* mutant by DIC microscopy clearing and Feulgen staining (confocal analysis; Braselton et al., 1996). Approximately 65% (n= 359) of the *drm1drm2* pre-meiotic ovules displayed multiple MMC-like cells (Fig.1C to E) when compared to the wild-type (Fig.1A and B) showing that the full *RdDM* pathway, including cytosine methylation, is pivotal for the specification of a single MMC. The formation of an MMC in the nucellus is part of the basic ovule developmental program, which is heavily influenced by ovule identity genes such as *STK*. Examination of pre-meiotic ovules in *stk*

lines showed double MMC-like cells to be present in 46% of ovules (n= 497) significant when compared to wild-type situation (Fig.1F to H). Significantly, this level of multiple MMC development is broadly similar to that reported by Olmedo-Monfil et al. (2010) for *ago9-2 and rdr6-11* lines, and reported here for *drm1drm2* ovules. More images of ovule primordium for each mutant line are presented in Figure S1. We have also counted the number of fertilized and unfertilized ovules in the mature siliques of *stk* and *drm1drm2* mutants, wild-type and *stk* produced similar ratios, whereas ~18% of ovules were unfertilised in *drm1drm2* siliques (Fig. S2).

***RdDM* gene expression is directly regulated by STK**

To identify a possible connection between STK and *RdDM* pathway components, such as *RDR6, AGO9, DRM1* and *DRM2*, we performed real-time PCR with RNA extracted from wild-type and *stk* mutant inflorescences (Fig. S3). Quantitative real time-PCR (qRT-PCR) was used to assess the relative level of gene expression in the *stk* mutant compared to wild-type. This confirmed that *STK* was significantly down-regulated in the *stk* background, consistent with previous reports (Pinyopich et al., 2003) (Fig. S3). Transcript levels of *AGO9*, *RDR6* and *DRM1* also showed significant down-regulation, although *DRM2* expression was unchanged compared to wild type (Fig. S3).

To further assess the possibility that STK might directly regulate expression of the *RdDM* pathway components, we performed Chromatin immunoprecipitation (ChIP) experiments using a *STK_GFP* line (Mizzotti et al., 2014) followed by qRT-PCR. We searched for CArG boxes in the putative regulatory regions of *RDR6*, *AGO9*, *DRM1* and *DRM2,* allowing a maximum of two mismatches in the consensus sequence (as described by Mendes et al., 2013). In the putative promoter and 5' untranslated region of *RDR6*, seven CArG boxes were identified and successively divided in four regions; two of these were enriched (P1 and P3; Fig. 2A and C) in *STK*_GFP immuno-precipitated chromatin when compared with wild-type chromatin. In the *AGO9* promoter and 5' untranslated region we identified 6 putative CArG boxes that were divided into 4 regions; we could confirm an enrichment in region P3 (Fig. 2A and B). Using the same criteria, in the case of *DRM1* and *DRM2*, no CArG box-like sequences were identified in the promoter regions nor the 5' UTR that were suitable for analysis. Taken together, these experiments suggest that STK is a direct regulator of *RDR6* transcription with two enriched regions, and an activator of *AGO9* with one region enriched.

The identity of putative ectopic MMC-like cells in *stk* and *RdDM* mutants

The *pKNU::nlsYFP* transcriptional marker specifically marks cells possessing MMC identity (Tucker et al., 2012; Fig. 3A) and was therefore used to define the identity of the supernumerary MMC-like cells in *stk* and *RdDM* mutants. To enable accurate comparisons, the number of ovules/cells showing YFP fluorescence in each mutant background was compared to wild-type segregants from the same cross. In *drm1drm2* and *stk* mutants two patterns were observed in each ovule: one of the two

putative MMCs expressed the YFP marker, or none of the putative MMCs expressed the marker (Fig. 3B and C). Indeed, in *stk* only 62% of the analysed ovules (n=468) were positive for YFP, while in wild-type segregants 82% of the ovules showed YFP expression (n=452), a statistically significant difference (p>0,01, t-student*; Table S1). A similar pattern was observed for *drm1drm2* double mutant ovules where 61% (n=572) generated YFP signal compared with 94% in wild-type (n=152); again this difference is statistically significant (p>0,01, t-student*; Table S1). Interestingly, 80% (n=350) of *rdr6-11* mutant ovules showed YFP signal in only one of the multiple MMC-like cells, while the remaining ovules were negative. This indicates that in the analysed *drm1drm2*, *stk* and *rdr6* mutants, only one of the two (or more) MMC-like cells acquires MMC identity, and in some cases the wild-type MMC loses aspects of its identity (as determined by the *KNU* promoter). In *ago9-2* mutant lines, only one cell expressed YFP signal in 91% (n=683) of ovules, while in wild-type segregants 93% of the ovules (n=688) showed expression. Interestingly, this mutant did generate signals in more than one cell with the second enlarged cell emitting a very faint signal in 2% of the cases; however, a similar situation was also found in the rare wild-type ovules showing multiple putative MMC-like cells, where 2% of total MMCs presented the double signal (Fig. 3 G and H). The analysis of *pKNU:nlsYFP* is summarized in Table S1.

While the extra MMC-like cells failed to activate the *KNU* promoter, we also asked whether they shared physical features with the MMC, and/or if they entered meiosis. Callose accumulates in the MMC cell wall before meiosis (Fig. S4), while during meiosis it accumulates in the walls of the four megaspores with a specific pattern. Thus, MMC expansion and progression of meiotic division was observed by assessing the pattern of callose deposition before (Fig. S4) and after meiosis (Fig. 3) in *stk, ago9*, *rdr6*, and *drm1drm2* with respect to wild-type. Aniline blue staining highlighted remarkably similar staining patterns in the MMC wall of wild-type (Fig. S4A), *stk* (Fig. S4B), *ago9* (Fig. S4C), *rdr6* (Fig.S4D) and *drm1drm2* (Fig. S4E) pre meiotic ovules (n>200 for each background) and in no cases was callose detected in more than one enlarged cell. During meiosis in wild-type ovules (Fig. 3D), it was possible to clearly distinguish the formation of two defined *septa* that correlate with the second meiotic division (n=20). In *stk* and *drm1drm2* as well as in *ago9-2 and rdr6-11* mutants only one of the two enlarged cells undergo meiosis (*stk* n=15*, drm1drm2* n=18*, ago9*-2 n=10 mutant and *rdr6-11* n=10; Fig. 3E, F and I).

Previous examination of the *RdDM* pathway mutants suggested that the supernumerary MMC-like cells might represent additional functional megaspores (Olmedo-Monfil et al., 2010). In particular it was shown that *pFM1::GUS* and *pFM2::GUS* were expressed a single FM/female gametophyte in wild-type ovules but accumulate in multiple cells in the *ago9* mutant (Olmedo-Monfil et al., 2010). Thus, we considered the possibility that the supernumerary MMC-like cells in *stk* ovules might express a functional megaspore marker gene. To test this we utilised the *pFM1::GUS* marker, and found that in the majority of ovules, expression appeared to be restricted to a single FM/female gametophyte in both wild-type and *stk* (Fig. S5 A-E). However, diffusion of the GUS signal was

difficult to distinguish from possible expression in extra cells. To overcome this, we utilized the fluorescent *pLC2:nlsYFP* marker (Tucker et al., 2012) which, in wild-type, is undetectable in the MMC but later shows expression in the functional megaspore and during the first mitotic divisions of megagametogenesis (97% positive signal, n= 344; Fig. 4A to C). The *ago9-2* mutant, which was previously reported to show ectopic expression of the *pFM2::GUS* functional megaspore marker (Olmedo-Monfil et al., 2010), was included as a representative of the *RdDM* pathway. No *pLC2:nlsYFP* expression was detected in ectopic enlarged cells in *stk* or *ago9-2* mutants at the MMC stage (n > 200; Fig. 4D and G), suggesting that these cells do not exhibit FM identity in pre-meiotic ovules. Surprisingly, after meiosis, expression in both mutants remained restricted to a single functional megaspore (Fig. 4E and H) and subsequently, the two nuclei produced after the first mitotic division (Fig. 4F and I). The pattern of expression was similar to that observed in wild-type ovules and showed similar staining efficiency (*stk*: 96%, n= 320; *ago9*-2: 95%, n= 440). The only variation in pattern was detected in less than 2% (9/440) of the *ago9-2* ovules where two sources of nuclear YFP signal were detected prior to FM division (Fig. S6). However, it was unclear whether this signal represented *pLC2:nlsYFP* expression in one functional megaspore and one supernumerary enlarged cell, or two megaspores within the meiotic tetrad. Taken together, these data indicate that the supernumerary enlarged cells in *stk* are unlikely to possess functional megaspore or female gametophyte identity. Moreover, inactivation of the key *RdDM* effector, *AGO9*, only leads to subtle changes in *pLC2:nlsYFP* expression, and appears insufficient to fully activate the functional megaspore program in the extra enlarged cell.

Ectopic expression of *SPL/NZZ* induces multiple MMC-like cells

Although the supernumerary MMC-like cells in *stk* and the *RdDM* mutants appear unable to fully adopt MMC or FM identity, we hypothesized that their presence in the nucellus of pre-meiotic mutant ovules might be due to ectopic expression of *SPL/NZZ*, which is required for MMC initiation (Yang et al., 1999, Schiefthaler et al., 1999). Therefore, we analysed *SPL/NZZ* expression by qRT-PCR in the *stk* mutant and in *RdDM* mutants *ago9, drm1drm2 and rdr6* (Fig. S7A). We show a consistent up-regulation of *SPL/NZZ* transcripts in closed flowers during early stages of development in *rdr6, drm1drm2* and *ago9* mutants. Furthermore, when *SPL/NZZ* expression was analysed in individual flowers enriching the first phases of megasporogenesis we could detect a significant increase in *SPL/NZZ* expression in *stk* mutant when compared to the wild-type. Taken together the results showed above further supported support the hypothesis that the presence of extra-numerary MMC-like cells in the analysed mutants might be caused by an upregulation of *SPL/NZZ* transcript (Fig. S7B).

To study the spatial expression pattern of SPL/NZZ in the *stk* and *RdDM* mutant ovules the activity of the *SPL/NZZ* promoter was investigated (Fig. S8) using a *pSPL5'::GUS:3'* construct. When wild-type lines were analysed, the promoter was seen to be active in the L1 cell layer of the extreme nucellar tip (white asterisks in Fig. S8) of pre-meiotic ovules while, in *stk*, *ago9-2* and *drm1drm2* lines, the construct was expressed ectopically throughout the distal nucellar primordium (Fig. S8). In order to investigate whether the up-regulation of *SPL/NZZ* in the *stk* mutant was due to a direct regulation by STK we analysed the regulatory region of *SPL/NZZ* searching for CArG boxes. Two CArG boxes were detected, with a maximum of two mismatches in the consensus sequence, and grouped in two regions: P1, upstream the coding sequence of *SPL/NZZ* and P2 in the 3' untranslated region, as already reported by Ito et al. (2004) (Fig. S9A). When binding of STK was tested by ChIP, neither the two regions showed enrichment in STK_GFP immuno-precipitated chromatin, when compared to the wild-type (Fig. S9B).

To determine the location of the SPL/NZZ protein in wild-type and mutant contexts we cloned the genomic region of *SPL* fused with GFP plus the regulatory region described above, creating *pSPL5'::SPL_GFP:3'* (*SPL_GFP).* To verify function of the recombinant protein we transformed *SPL_GFP* into the *spl/nzz* mutant, and confirmed that it fully complemented the fertility defects (Fig. S10). We subsequently analysed the expression of *SPL_GFP* in three independent lines and found that in pre-meiotic ovules, GFP was confined to nuclei of the L1 layer in the tip of the nucellus, similar to what was detected with the *pSPL5'::GUS:3'* transcriptional reporter. *SPL_GFP* signal was not detected in the MMC, its products or its precursors, providing compelling evidence that although SPL/NZZ is expressed in L1 is required for MMC differentiation in the L2 layer (Fig. 5 A and B). *SPL_GFP* was then introgressed into mutants of the two major *RdDM* pathway components (*ago9-2* and *drm1drm2*) to assess any changes in expression. Analysis of ovules confirmed that the recombinant protein was localised in nuclei and exhibited the same expanded pattern shown by GUS driven by the *SPL* promoter (*pSPL5'::GUS:3'*). However, subtle differences were observed in SPL/NZZ expression between the two mutant backgrounds. In *ago9-2* ovule that presented an extra MMC-like cell in the nucellus of the ovule primordia, the *SPL_GFP* fusion protein was detected in more L1 layer cells when compared to wild-type (Fig. 5E). The differences were important when the expression was compared to the wild-type like neighbour ovule with only one enlarged cell, in the confocal Z-projection the difference in the number of GFP nucleus was evident (Fig. 5F). In *drm1drm2* (Fig. 5G-I) SPL/NZZ protein was again detected in more L1 layer cells when compared to the wild-type situation and the two ovules in the Z-projection presented a clear difference in the number of GFP nucleus (Fig. 5I). With this we can conclude that the *RdDM* and DNA methylation pathways play an important role in the precise regulation of SPL/NZZ expression and activation in the L1 nucellus layer. During meiosis SPL/NZZ expression remains in the cell layers surrounding the MMC (Fig. S11).

The ectopic expression of SPL/NZZ in *ago9* and *drm1drm2* suggested that this might contribute to the formation of extra MMC-like cells in pre-meiotic ovules. To assess this hypothesis, we expressed *SPL/NZZ* under the control of the 35SCaMV promoter, which accumulates to high levels in multiple tissues including the nucellus of pre-meiotic ovules. Eight independent lines were obtained in the T1 generation (Fig. 6), all of which presented multiple MMC-like cells in pre-meiotic ovules, as shown in Figure 6F. The supernumerary enlarged cells in the nucellus were observed after clearing and DIC microscopy (Fig. 6A to C) and with confocal imaging after Feulgen staining (Fig. 6D to E). This observation supports the hypothesis that in *stk*, and in the *RdDM* mutants, the formation of extra germ line precursors are caused by broader SPL/NZZ expression.

Discussion

Absence of Methylation via *RdDM* leads to a multiple MMC-like cell phenotype

The aim of this study was to contribute to understanding of the gene network controlling formation of a single female megaspore mother cell (MMC) in the young ovule of Arabidopsis. Olmedo-Monfil *et al* (2010) reported both *ago9-2* and *rdr6-11* lines to generate additional MMC-like cells in pre-meiotic ovules, implicating *RdDM* in MMC specification. *AGO9* and *RDR6* are involved in the processing of small interfering RNAs (siRNAs) directing the cytosine methylation of complementary DNA sequences. To determine whether *AGO9* and *RDR6*-dependent siRNAs might be involved in silencing of target sequences via DNA methylation, we investigated the role of DOMAIN REARRANGEMENT METHYLASE 1 and 2 (DRM1 and DRM2), two methyltransferases that are involved solely in methylation via *RdDM* (Law and Jacobsen 2010). Here we show that ~65% of pre-meiotic ovules of *drm1drm2* lines contained multiple MMC-like cells, confirming that *RdDM* of target sequences is required for the specification of a single MMC in Arabidopsis. This finding is consistent with previous studies that showed a reduced level of methylation accompanies MMC differentiation (Ingouff et al., 2017). Similar to the MMC, a reduction in methylation has been reported during early microspore mother cell differentiation in the anther (Walker et al., 2018). DRM1 and DRM2 are also important for setting the correct methylation poise following fertilisation and DRM2 has been shown to complex with the *RdDM* effector AGO4 (Zhong et al., 2014) and to be mainly expressed in the developing embryo (Havecker et al., 2010), while both methyltransferases have been shown to be required for embryogenesis (Chow et al., 2020).

***RdDM* in the ovule is directly regulated by the key ovule identity MADS-box transcription factor STK.**

Our data also indicate that the female developmental program in Arabidopsis, including MMC specification, is coordinated by STK, a MADS-box transcription factor already implicated in a wide range of processes including the establishment of ovule identity (Matias-Hernandez et al., 2010; Mendes et al., 2016). The formation of two or more MMC-like cells in approximately 46% of the pre-meiotic ovules in *stk* mutants strikingly resembles the phenotype of the *ago9-2* and *rdr6-11* plants reported by Olmedo-Monfil et al. (2010) and that of our *drm1drm2* double mutants. We therefore searched for a link between STK and the *RdDM* pathway, finding that transcription of both *RDR6* and *AGO9* may directly be activated by STK, which we show to bind to different CArG boxes regions. This is consistent with the requirement of MADS-box TFs to bind at least to two binding sites to activate transcription (Mendes et al., 2013). Strikingly, the expression of both *AGO9* and *RDR6* is downregulated in *stk* mutant lines, confirming the functionality of these binding events.

***SPL/NZZ* expression is restricted to few cells in the nucellus by the STK/*RdDM* pathway**

Previous reports revealed that 99% of pre-meiotic *spl/nzz* ovules fail to develop a MMC (Yang et al. 1999; Wei et al. 2015), a phenotype diametrically opposed to that of the *rdr6*, *ago9*, *drm1drm2*, and *stk* plants. *SPL/NZZ* thus constitutes a strong candidate target of the STK/RdDM pathway and the qRT-PCR data described in our manuscript show *SPL/NZZ* to be overexpressed in all *RdDM* and *stk* mutant lines. In wild-type plants, *SPL/NZZ* is expressed solely in the tip ovule primordium L1 layer cells at very early stages, as clearly shown in both *pSPL5'::GUS:3'* and *pSPL5'::SPL_GFP:3'* lines. Importantly, while both promoter expression and SPL/NZZ protein synthesis occurs in these few distal cells, expression of the *SPL/NZZ* protein fusion was never detected in the single cell (in wild-type lines) destined to assume MMC identity, indicating that SPL/NZZ regulates MMC development non-cell autonomously. By contrast, striking ectopic expression of *pSPL5'::GUS:3'* occurs in the L1 layer of primordium ovules of *stk, ago9-2* and *drm1drm2* lines; furthermore, similar ectopic expression was also detected in these cells when *ago9* and *drm1drm2* mutant plants were transformed with the *pSPL5'::SPL_GFP:3'* construct confirming the presence of the SPL/NZZ protein.

Further confirmation that restriction of *SPL/NZZ* expression to only a few cells of the primordium ovule L1 layer is required for single MMC specification comes from analysis of transgenic plants expressing *SPL/NZZ* driven throughout the ovule primordium by the 35CaMV promoter. Here again, multiple MMC-like (germline precursors) cells are formed. Since evidence is accumulating that the expression pattern of *SPL/NZZ* in the ovule is regulated by RdDM-mediated methylation, a sensible next step will be to determine whether the ovular NZZ/SPL is itself a target for methylation. Evidence from global methylome analysis (Zhang et al., 2006)suggests that the *NZZ/SPL* sequence is

unmethylated, but it is unclear whether the sequencing system used would have detected differences in methylation at a local cellular context. If the *SPL/NZZ* genomic locus remains unmethylated in the ovule then an intermediate regulatory factor is likely to serve as a direct target of the RdDM pathway.

SPL/NZZ is sufficient for germline precursor specification, but not MMC function

The observation that extra MMC-like cells in *stk, ago9* and *drm1drm2* lines fail both to express a MMC identity gene and to enter meiosis indicates that while SPL/NZZ is required for initial MMC specification (i.e. germline precursor formation), it is insufficient to confer full MMC identity and function. For example, little or no expression of the MMC marker *pKNU::nlsYFP* (Tucker et al., 2012) was detectable in in the supernumerary enlarged cells of *stk*, *ago9* and *drm1drm2* ovule primordium. Also, the functional megaspore marker *pLC2:nlsYFP*, which is expressed in the post-meiotic megaspore of wild-type plants, only labels a single cell in *stk* and *ago9* ovules, indicating that the supernumerary cells formed in these lines are unlikely to be functional megaspores. Furthermore, only one of the cells in *stk* and *drm1drm2* ovules enters meiosis, as has already been shown for *ago9* and *rdr6* mutants (Olmedo-Monfil et al., 2010). In view of the disruption of early ovule development, the observation that fertility is unaffected in *stk* mutants is surprising, and only slightly reduced (by 18% $p<0.01$) in *drm1drm2* lines. The observation that the *KNU* promoter is inactive in 40% of the ovules of fertile *stk* and *drm1drm2* mutants (data in Table S1) is perplexing and may indicate that that *KNU* expression, despite being a feature of wild-type development, is not essential for entry into meiosis. Taken together, our data confirm that while the STK/RdDM/SPL pathway specifies the formation of a single female germline precursor, further factors are required for complete MMC function.

STK-mediated _RdDM_ restriction of _SPL/NZZ expression_, and a new model for female germline precursor formation

SPL/NZZ acts as an adaptor-like transcriptional repressor in Arabidopsis, potentially forming a bridge between TOPLESS (TPL)/related (TPR) proteins and CIN-like TCP transcription factors (Wei *et al*. 2015). When bound to SPL in this way TPL proteins are proposed to repress TCP activity to promote MMC development (Wei *et al*. 2015) and, in the absence of SPL/NZZ, over-expression of TCP genes to result in failure of megasporogenesis and abnormal ovule development. However, constitutive repression of TCP sequences alone is not sufficient for FM development, pointing to the involvement of additional factors (Wei et al., 2015). Thus, in the ovule primordia of *RdDM* mutant lines, it is presumably expansion of ectopic SPL/NZZ expression into an extended field of L1 cells that results in development of subtending supernumerary MMC-like cells (female germline percursors). Why the ectopic SPL/NZZ ectopic expression fails to develop additional fully functional MMCs is unclear. Our

observation of multiple MMC-like cells being formed when SPL is expressed throughout the ovule confirms that SPL expression itself does not inhibit MMC-like cell formation as occurs in some lateral control systems. However, the L1 is a highly specialised cell layer with a transcriptome very different from the L2 and it may simply be that additional factors necessary for MMC formation are absent from these cells. Previously it was suggested that SPL is involved in auxin homeostasis, since in the *spl/nzz* mutant, PIN1 (auxin transporter) expression was compromised, suggesting that SPL is important for PIN1 expression (Bencivenga et al., 2012). Auxin accumulates at the nucellus tip similar to the expression patter on SPL/NZZ protein and this interaction could be linked to MMC specification (Ceccato et al., 2013). SPL/NZZ has also been proposed to act in concert with the homeodomain transcription factor WUSCHEL (WUS), to promote MMC differentiation. Indeed, the *wus* mutant lacks a primary female germline cell in approximately 12% of ovules. However it still unknown how the two transcription factors interact (Groß-Hardt et al., 2002). *WUS* and/or hormone related pathways may be related to the fully MMC specification.

In an attempt to integrate our data with the findings of Olmedo-Monfil et al. (2010) we have developed a new model (Fig. 8) for the genetic and epigenetic control for specification of the single MMC in Arabidopsis.

The model proposes that STK activates in the lower L1 layer cells RdDM gene expression, siRNAs present in those cells act either via mRNA cleavage/repression or methylation to suppress SPL/NZZ expression and synthesis. On the other hand, SPL/NZZ is only synthesised and expressed at the tip of ovule primordium/L1 layer. The specification of female germline precursor can be achieved through a number of possibilities; presumably based in the movement of an effector molecule, possibly through the plasmodesmata, to the subtending L2 cell to initiate MMC expansion and eventual MMC specification. The nature of this effector, its mode of operation and its target(s) in the L2 cell remain to be determined. Interestingly, it has been reported that mutations in homologs of DRMs and AGO genes in rice and maize influenced MMC differentiation and/or the determination of the female germline precursor (Singh et al., 2011). Taken together these evidences suggested that our model might also be extended to crops. Taken in account that, a further comprehension of the mechanisms determining MMC specification could have an important role for future noteworthy crop improvement.

Material and methods

Plant material and Growth conditions

All plants *Arabidopsis thaliana* ecotype Columbia-0 (wild-type), mutants and marker lines (table 1) were grown in soil obtained from a ratio of 5:1:1 potting soil: vermiculite: perlite respectively. Plants were grown first at 21°C / 18°C under short day conditions (SD - 8h light/ 16h dark) until the bolting of the rosette leaves and successively at long day conditions (LD - 8h dark/ 16h light) until flowering with 70% humidity in both periods.

Generation of marker lines in different mutant backgrounds.

The *pKNU:nlsYFP* is expressed in the megaspore mother cell (MMC) in the pre-meiotic ovules. The characterized *pKNU:nlsYFP* Col wild-type was crossed to stk, ago9, rdr6 and drm1drm2 mutants, and at least three homozygous F3 plants were analysed for expression in each combination. The *pLC2:nlsYFP* marker (pAt5g40730:nls-vYFP) is expressed in the functional megaspore and during subsequent stages of megagametogenesis in the Columbia background (Tucker *et al*., 2012). The characterised *pLC2:nlsYFP* Col wild-type line was crossed to *ago9* and *stk* mutants, and three homozygous F3 plants were analysed for expression in each combination.

RNA extraction

Gene expression was evaluated in closed flower till stage 15 (according to Smyth et al., 1990) Tissue samples were collected using tweezers and placed immediately on liquid nitrogen. Total RNA was extracted with the RNeasy Mini Kit (Qiagen) following the manufacturer's instructions manual. RNA was treated with TURBO™ DNase (Invitrogen), and retro-transcribed using Superscript III and Oligo(dT)12–18 Primer (Invitrogen).

qPCR optimization and conditions

Oligonucleotides for expression analysis of *STK (AT4G09960), AGO9 (AT5G21150), SPL (AT4G27330), RDR6 (AT3G49500) DRM1 (AT5G15380) and DRM2* (AT5G14620) were designed de novo (table 2). Primer specificity tests and qPCR experiments were performed

as described in Burton et al. (2008) with the following qPCR cycling conditions: 3min at 95°C followed by 45 cycles of 10 s at 95°C, 1 s at 55oC, 30 s at 72°C, and 15 s at the optimal acquisition temperature (see table 2). The transcript levels of genes encoding AtCyclophilin (AT2G36130), AtActin (AT5G43500), AtTublin (AT1G50010), and AtGAPdH (AT3G26650) were used as controls (Li et al., 2014). Normalization was carried out using these control genes as described by Burton et al., 2004). The results are expressed as arbitrary units that represent the number of copies per microliter of cDNA, normalized against the geometric means of the three control genes that vary the least with respect to each other (Vandesompele et al., 2002). The standard error was calculated from the average expression level of each biological replicate. For the statistical analysis a two-tailed, unpaired, homogeneity of variance Student's t-test was performed, □□= 0.05. Three biological replicates were analysed for the whole inflorescence experiment. At least 9 technical replicates were analysed in each experiment.

Microscope analysis: Differential interference contrast (DIC) and Confocal.

To analyse cleared tissue, the differential interference contrast (DIC) microscopy (Immersion Zeiss® Axiophot D1 10X, 20X, 40X and 100X) was used. It is an optical microscopy illumination technique used to enhance the contrast in unstained and transparent samples. In particular, DIC was used to observe the percentage of single MMC and multiple MMCs in finger-like ovules in mutant lines and wild-type. Then, the statistical relevance is analysed using the Student's *t*-test. Pictures were acquired with Zeiss® Axiocam MRc5 camera and Axiovision (version 4.1) software. To analyse pre meiotic ovules even more in detail, a modification of the protocol described by Braselton et al. (1996) for confocal laser scanning microscopy was used. This treatment stains both nuclei and cell walls. SR staining solution, containing 0.1 % (v/v) SR2200 (Renaissance Chemicals; stock solution of the supplier was considered as 100 %) in water was prepared freshly prior to use. For images of ovules, they were manually dissected out of the pistil and collected in a drop of staining solution on a microscope slide and mounted under a coverslip. Images were obtained with Nikon A1 confocal microscopes. SR2200 was excited with a 405-nm laser line and emission recorded between 415 and 476 nm (405/415–476), similar to DAPI settings. The different mutant phenotypes were analysed by three different students in two different labs.

Callose Staining and microscopy

Callose analysis prior to meiosis where performed using aniline blue staining. Whole flowers were collected at stages corresponding to ovule stage 2-I (Schneitz et al., 1995) and dissected under a

Zeiss dissection microscope. Samples were collected from wild-type and at least three confirmed homozygous lines for each mutant. Carpels were gently sliced open in 20µl of Aniline Blue staining solution (0.005% [w/v] Aniline Blue diammonium salt (Sigma Aldrich, Catalogue No. 415049) in PBS) as per Rodkiewicz (1970) and Reiser and Fischer (1993). The ovules were gently detached from the placenta and released into the solution. A further 20µl of Aniline Blue solution was pipetted onto the sample, immediately covered with a coverslip and transferred to a fluorescent Axio Imager M2 microscope for vieweing under UV light (CFP filter; Zeiss Filter set 47: 436/480 nm). Auto-fluorescence was used to highlight the ovule outline in the dsRED channel (Zeiss Filter set 43: 545/605 nm) and the DIC images were captured with Normarksi optics. YFP signal (for pLC2:nlsYFP and pKNU:nlsYFP) was detected using a YFP filter (Zeiss Filter set 46: 500/535 nm). Images were processed with ZEN imaging software and Adobe Photoshop.

Callose staining during meiosis was performed using Renaissance-staining solution; containing 0.1 % (v/v) SR2200 (Renaissance Chemicals; stock solution of the supplier was considered as 100 %) in water was prepared freshly prior to use. For imaging ovules were manually dissected out of the pistil and collected in a drop of staining solution on a microscope slide and mounted under a coverslip. Images were obtained with Nikon A1 confocal microscopes. SR2200 was excited with a 405-nm laser line and emission recorded between 415 and 476 nm (405/415–476), similar to DAPI settings.

Chromatin immunoprecipitation studies (ChIP)

ChIP assays were performed as described by Gregis et al. (2009) using for STK-GFP the commercial antibody GFP:Living Colors full-length (Clontech). Quantitative Real-Time PCR assays were performed to determine the enrichment of the fragments. The detection was performed in triplicate using the iQ SYBR Green Supermix (Bio-Rad) and the Bio-Rad iCycler iQ Optical System (software version 3.0a), with the primers listed below. ChIP-quantitative Real-Time PCR experiments and relative enrichments were calculated as reported by Matias-Hernandez et al. (2010). We employed the following formulas to calculate the fold enrichment: dCT.tg = CT.i-CT.tg and dCT.nc =CT.i-CT.nc. Ct.tg is target gene mean value, Ct.i is input DNA mean value, and Ct.nc is ACTIN 7 (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 using dSD.tg = sqrt((SD.i)^2+ (SD.tg^2)/sqrt(n) and dSD.nc = sqrt((SD.i)^2+ (SD.nc^2)/sqrt(n), n = number of replicate per sample. Fold-change over negative control 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 is obtained as follows: FC = 2^(ddCT) and FC.error = ln(2)*ddSD*FC. All the experiments were performed in three biological replicates.

6. Cloning SPL5’::GUS:3’ and SPL5’::SPL_GFP:3’

First, *SPL* locus was cloned into pDONR207 (Life Technologies) and successively transferred to pBGWFS7 destination vector (Thermofisher scientific); the expression vector was used to amplify the *SPL* genomic region fused to GFP reporter gene. Successively, the obtained fragment was cloned into pDONR207 (Life Technologies). The putative promoter region of the gene plus the 5’-UTR and the 3’UTR were cloned and successively transferred to pDONR201 P4-P1r and pDONR221 P2r-P3, respectively. By a Multisite Gateway approach, we obtained *pSPL:SPL_GFP-3’UTR*, combining the obtained donor vectors (described above) and transferring them into pH7M34GW destination vector (Thermofisher scientific). The primers used are listed in the table 4.

7. Cloning 35S CaMV::SPL/NZZ

To construct *35S::SPL*, SPL CDS was amplified with the primers listed below and cloned into the Gateway® destination vector pB2GW7. All constructs were verified by sequencing and used to transform wild-type plants using the floral dip method (Clough and Bent, 1998).

ATP_4196 GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCGACTTCTCTCTTCTTC

AtP_4197 GGGGACCACTTTGTACAAGAAAGCTGGGTTTAAAGCTTCAAGGACAAATC

Competing interests

The authors declare no competing or financial interests.

Author contributions

Conceptualization: M.A.M, H.D., M.R.T, L.C.; Formal analysis and investigation: M.A.M., R.P., M.C., E.V., S.G., S.P., D.C.B., V.G. Data Curation: R.P, M.C, E.V., S.G. Writing - original draft preparation: M.A.M., L.C; Writing - review and editing: M.A.M., R.P., H.D., M.R.T., L.C.; Funding acquisition: M.A.M., M.R.T., L.C.; Resources: L.C.; Supervision: M.A.M., M.R.T., L.C.

Funding

M.A.M. was supported by Linea2 - PSR2015-1718LCOLO_12, University of Milan. M.R.T was supported by Australian Research Council (FT140100780 & DP180104092); R.P was supported by

the Doctorate School in Molecular and Cellular biology; L.C. was supported by Ministero dell'Istruzione, dell'Università e della Ricerca MIUR-PRIN 2012 and SEXSEED, Horizon2020- MSCA RISE 2016 - European Union (EU)- project n. 690946.

Acknowledgments

We thank Chao Mao for technical support and A. Movilli, A. Cavalleri and M. Di Marzo for helpful suggestions and stimulating discussions. Part of this work was carried out at NOLIMITS, an advanced imaging facility established by the Università degli Studi di Milano.

References

Balanzà, V., Roig-Villanova, I., Di Marzo, M., Masiero, S., and Colombo, L. (2016). Seed abscission and fruit dehiscence required for seed dispersal rely on similar genetic networks. *Development* **143**:3372–3381.

Bencivenga, S., Simonini, S., Benková, E., and Colombo, L. (2012). The transcription factors BEL1 and SPL are required for cytokinin and auxin signaling during ovule development in Arabidopsis. *Plant Cell* **24**:2886–2897.

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 Online* Advance Access published 2007, doi:10.1105/tpc.107.051797.

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.

Burton, R. A., Shirley, N. J., King, B. J., Harvey, A. J., and Fincher, G. B. (2004). The CesA Gene Family of Barley. Quantitative Analysis of Transcripts Reveals Two Groups of Co-Expressed Genes. *Plant Physiol.* **134**:224–236.

Ceccato, L., Masiero, S., Sinha Roy, D., Bencivenga, S., Roig-Villanova, I., Ditengou, F. A., Palme, K., Simon, R., and Colombo, L. (2013). Maternal Control of PIN1 Is Required for Female Gametophyte Development in Arabidopsis. *PLoS One* **8**:e66148.

Chow, H. T., Chakraborty, T., and Mosher, R. A. (2020). RNA-directed DNA Methylation and sexual reproduction: expanding beyond the seed. *Curr. Opin. Plant Biol.* **54**:11–17.

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.

Clough, S. J., and Bent, A. F. (1998). Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis thaliana. *Plant J.* **16**:735–43.

Cucinotta, M., Manrique, S., Guazzotti, A., Quadrelli, N. E., Mendes, M. A., Benkova, E., and Colombo, L. (2016). Cytokinin response factors integrate auxin and cytokinin pathways for

female reproductive organ development. *Dev.* **143**:4419–4424.

Ezquer, I., Mizzotti, C., Nguema-Ona, E., Gotté, M., Beauzamy, L., Viana, V. E., Dubrulle, N., Costa de Oliveira, A., Caporali, E., Koroney, A.-S., et al. (2016). The Developmental Regulator SEEDSTICK Controls Structural and Mechanical Properties of the Arabidopsis Seed Coat. *Plant Cell* **28**:2478–2492.

Favaro, R., Pinyopich, A., Battaglia, R., Kooiker, M., and Borghi, L. (2003). MADS-Box Protein Complexes Control Carpel and Ovule Development in Arabidopsis. *Plant Cell* **15**.

Gregis, V., Sessa, A., Dorca-Fornell, C., and Kater, M. M. (2009). The Arabidopsis floral meristem identity genes AP1, AGL24 and SVP directly repress class B and C floral homeotic genes. *Plant J.* **60**:626–637.

Groß-Hardt, R., Lenhard, M., and Laux, T. (2002). WUSCHEL signaling functions in interregional communication during Arabidopsis ovule development. *Genes Dev.* **16**:1129–1138.

Havecker, E. R., Wallbridge, L. M., Hardcastle, T. J., Bush, M. S., Kelly, K. A., Dunn, R. M., Schwach, F., Doonan, J. H., and Baulcombe, D. C. (2010). The arabidopsis RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. *Plant Cell* **22**:321–334.

Herrera-Ubaldo, H., Lozano-Sotomayor, P., Ezquer, I., Di Marzo, M., Chávez Montes, R. A., Gómez-Felipe, A., Pablo-Villa, J., Diaz-Ramirez, D., Ballester, P., Ferrándiz, C., et al. (2019). New roles of NO TRANSMITTING TRACT and SEEDSTICK during medial domain development in Arabidopsis fruits. *Development* **146**:dev172395.

Ingouff, M., Selles, B., Michaud, C., Vu, T. M., Berger, F., Schorn, A. J., Autran, D., Van Durme, M., Nowack, M. K., Martienssen, R. A., et al. (2017). Live-cell analysis of DNA methylation during sexual reproduction in arabidopsis reveals context and sex-specific dynamics controlled by noncanonical RdDM. *Genes Dev.* **31**:72–83.

Ito, T., Wellmer, F., Yu, H., Das, P., Ito, H., Alves-Ferreira, M., Riechmann, J. L., and Meyerowitz, E. M. (2004). The homeotic protein AGAMOUS controls microsporogeniesis by regulation of SPOROCYTELESS. *Nature* **430**:356–360.

Marzo, M. Di, Herrera-Ubaldo, H., Caporali, E., and Mendes, M. A. (2020). SEEDSTICK

Controls Arabidopsis Fruit Size by Regulating Cytokinin Levels and FRUITFULL. *CellReports* **30**:2846-2857.e3.

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.

Matzke, M. A., and Mosher, R. A. (2014). RNA-directed DNA methylation: An epigenetic pathway of increasing complexity. *Nat. Rev. Genet.* **15**:394–408.

Matzke, M. A., Kanno, T., and Matzke, A. J. M. (2015). RNA-Directed DNA Methylation: The Evolution of a Complex Epigenetic Pathway in Flowering Plants. *Annu. Rev. Plant Biol.* **66**:243–267.

Mendes, M. A., Guerra, R. F., Berns, M. C., Manzo, C., Masiero, S., Finzi, L., Kater, M. M., and Colombo, L. (2013). MADS domain transcription factors mediate short-range DNA looping that is essential for target gene expression in Arabidopsis. *Plant Cell* **25**:2560–2572.

Mendes, M. A., Guerra, R. F., Castelnovo, B., Silva-Velazquez, Y., Morandini, P., Manrique, S., Baumann, N., Groß-Hardt, R., Dickinson, H., and Colombo, L. (2016). Live and let die: a REM complex promotes fertilization through synergid cell death in *Arabidopsis*. *Development* **143**:2780–2790.

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

Mizzotti, C., Ezquer, I., Paolo, D., Rueda-Romero, P., Guerra, R. F., Battaglia, R., Rogachev, I., Aharoni, A., Kater, M. M., Caporali, E., et al. (2014). SEEDSTICK is a Master Regulator of Development and Metabolism in the Arabidopsis Seed Coat. *PLoS Genet.* **10**:e1004856.

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

Petrella, R., Caselli, F., Roig-Villanova, I., Vignati, V., Chiara, M., Ezquer, I., Tadini, L., Kater, M. M., and Gregis, V. (2020). BPC transcription factors and a Polycomb Group protein confine the expression of the ovule identity gene *SEEDSTICK* in Arabidopsis. *Plant J.* Advance Access published February 4, 2020, doi:10.1111/tpj.14673.

Pinto, S. C., Mendes, M. A., Coimbra, S., and Tucker, M. R. (2019). Revisiting the Female Germline and Its Expanding Toolbox. *Trends Plant Sci.* **24**:455–467.

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.

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.

Singh, M., Goel, S., Meeley, R. B., Dantec, C., Parrinello, H., Michaud, C., Leblanc, O., and Grimanelli, D. (2011). Production of viable gametes without meiosis in maize deficient for an ARGONAUTE protein. *Plant Cell* **23**:443–458.

Stroud, H., Do, T., Du, J., Zhong, X., Feng, S., Johnson, L., Patel, D. J., and Jacobsen, S. E. (2014). Non-CG methylation patterns shape the epigenetic landscape in Arabidopsis. *Nat. Struct. Mol. Biol.* **21**:64–72.

Su, Z., Zhao, L., Zhao, Y., Chen, P., Qin, Y., and Correspondence, X. C. (2017). The THO Complex Non-Cell-Autonomously Represses Female Germline Specification through the TAS3-ARF3 Module. *Curr. Biol.* **27**:1597-1609.e2.

Tucker, M. R., Okada, T., Hu, Y., Scholefield, A., Taylor, J. M., and Koltunow, A. M. G. (2012). Somatic small RNA pathways promote the mitotic events of megagametogenesis during female reproductive development in arabidopsis. *Development* **139**:1399–1404.

Vandesompele, J., De Preter, K., Pattyn, F., Poppe, B., Van Roy, N., De Paepe, A., and Speleman, F. (2002). Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol.* **3**:research0034.1.

Walker, J., Gao, H., Zhang, J., Aldridge, B., Vickers, M., Higgins, J. D., and Feng, X. (2018).

Sexual-lineage-specific DNA methylation regulates meiosis in Arabidopsis. *Nat. Genet.* **50**:130–137.

Wei, B., Zhang, J., Pang, C., Yu, H., Guo, D., Jiang, H., Ding, M., Chen, Z., Tao, Q., Gu, H., et al. (2015). The molecular mechanism of SPOROCYTELESS/NOZZLE in controlling Arabidopsis ovule development. *Cell Res.* **25**:121–134.

Yadegari, R., and Drews, G. N. (2004). Female gametophyte development. *Plant Cell* **16**.

Yang, W. C., Ye, D., Xu, J., and Sundaresan, V. (1999). The SPOROCYTELESS gene of Arabidopsis is required for initiation of sporogenesis and encodes a novel nuclear protein. *Genes Dev.* **13**:2108–2117.

Zemach, A., Kim, M. Y., Hsieh, P. H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S. L., and Zilberman, D. (2013). The arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. *Cell* **153**:193–205.

Zhang, W., Sun, Y., Timofejeva, L., Chen, C., Grossniklaus, U., and Ma, H. (2006). Regulation of Arabidopsis tapetum development and function by Dysfunctional Tapetum1 (DYT1) encoding a putative bHLH transcription factor. *Development* **133**:3085–3095.

Zhong, X., Du, J., Hale, C. J., Gallego-Bartolome, J., Feng, S., Vashisht, A. A., Chory, J., Wohlschlegel, J. A., Patel, D. J., and Jacobsen, S. E. (2014). Molecular mechanism of action of plant DRM de novo DNA methyltransferases. *Cell* **157**:1050–1060.

Figures

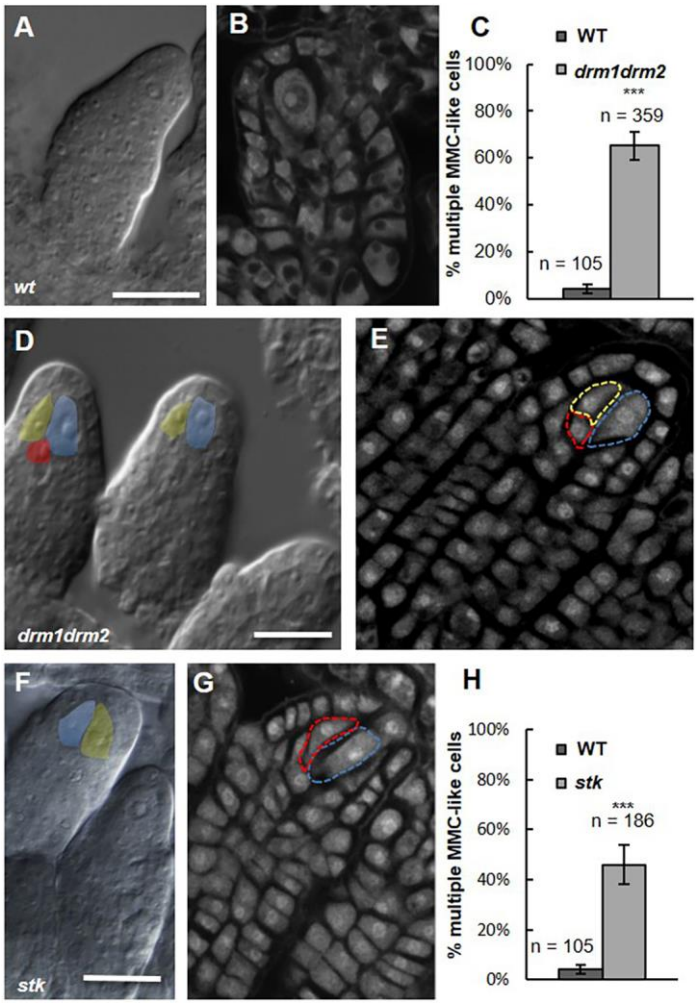


Figure 1. *drm1drm2* and *stk* mutants present multiple MMC-like cells in pre-meiotic ovules.

DIC and confocal imaging of pre-meiotic ovules; in wild-type (A-B) a single enlarged cell is detected in the nucellus, corresponding to the MMC (in light-blue); Multiple enlarged cells were detected in the nucellus in *drm1drm2* double mutant (D-E) and in *stk* single mutant (F-G); scale bar= 15 μm; different colours mark the multiple enlarged cells. Confocal sections were acquired after Feulgen staining. (C) and (H) Graphic representation of *drm1drm2* and *stk* multiple nucellar enlarged cells, respectively. Asterisks (***) indicate P< 0.001 in a Student's t-test; wild-type pre-meiotic ovules were screened as control.

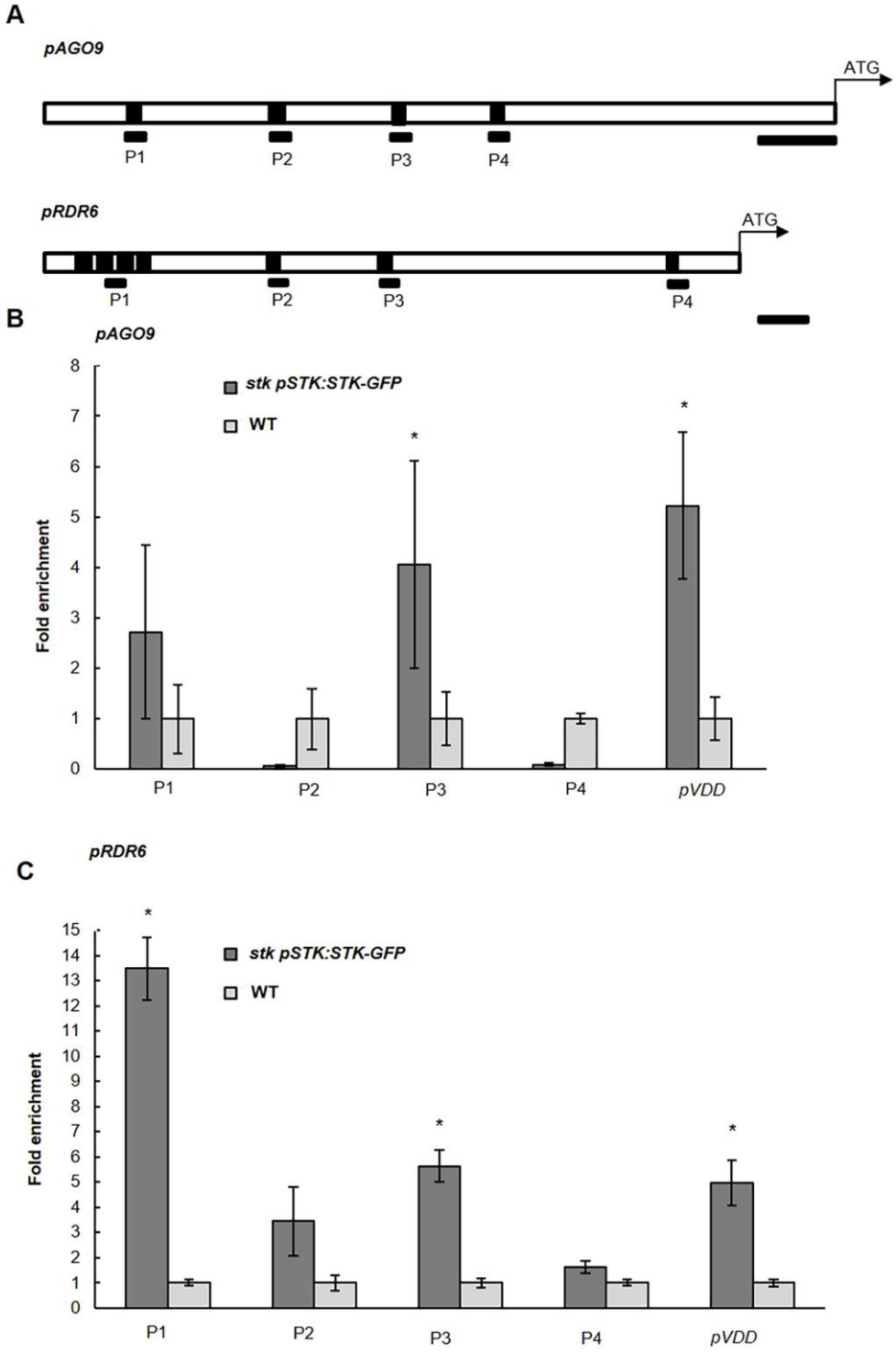


Figure 2. STK directly binds the regulatory region of RDR6 and AGO9.
(A) Schematic diagram of the AGO9 and RDR6 putative promoter and 5' untranslated regions, respectively, indicating the regions analysed by chromatin immunoprecipitation (ChIP; black bars) followed by qRT-PCR; each black box represents a CArG box; scale bar =350 bp, top and 50 bp, bottom. (B-C); Quantitative Real-Time PCR analysis of ChIP assay using chromatin extracted from *stk* mutant complemented with *pSTK::STK_GFP* and wild-type (as a negative control), testing the CArG boxes regions on *pAGO9* (B) and *pRDR6* (C); the promoter region of *VDD* was tested as positive enrichment (Matis-Hernandez et al., 2010). For the IP, commercial antibodies against GFP were used; Error bars represent the

propagated error value using three replicates. ChIP results of one representative experiment are shown. Positive binding site fragments were considered only if they were enriched compared with the controls in at least three independent experiments. Asterisks (*) indicate significant differences compared to the wild-type and represent a p-value < 0.05 in a Student's t-test.

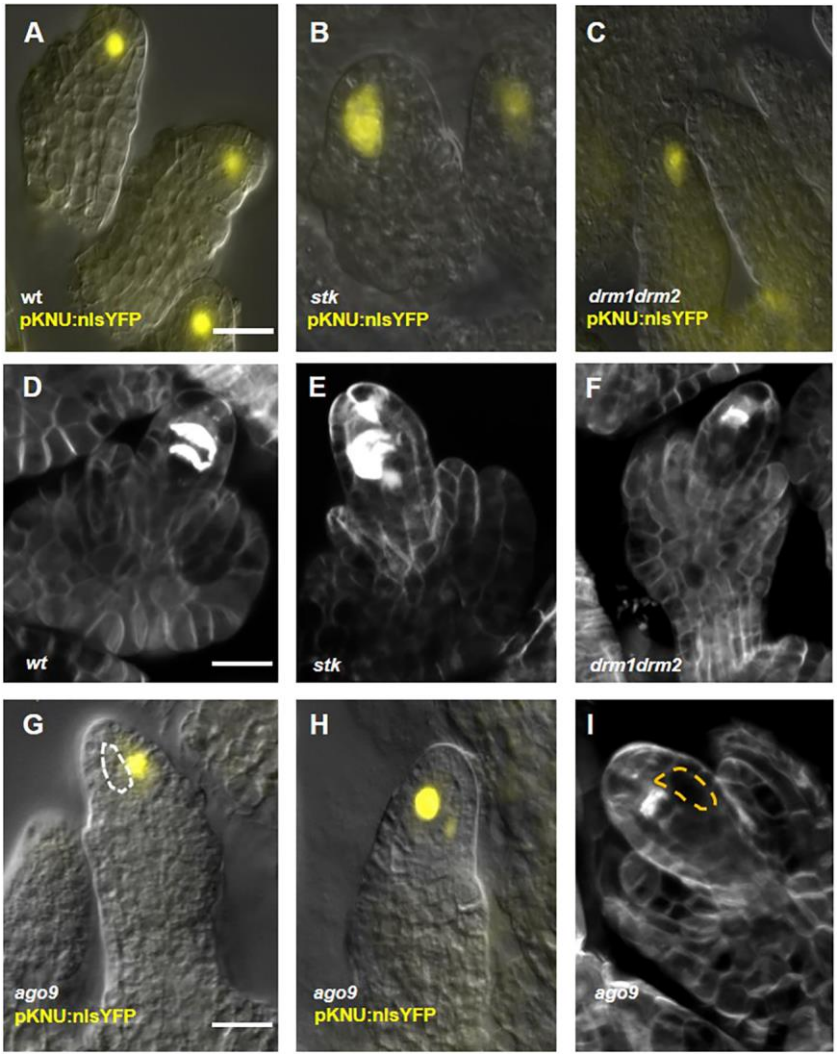


Figure 3. *pKNU:nlsYFP* expression analysis and meiosis staining.
pKNU:nlsYFP marker in wild-type (A), *stk* (B), *drm1drm2* (C) and *ago9-2* (G-H). Some examples are provided: in all the mutant background analysed, even though two enlarged cells can be observed, only one resulted positive for YFP expression (B) and (G); also, several ovules did not show any YFP signal (C); (D-F) and (I) Callose staining in wild-type, *stk, drm1drm2 and ago9* background, respectively; the analysis showed that only one cell enters meiosis, scale bar=15 µm; dashed lines highlight the extra enlarged cells.

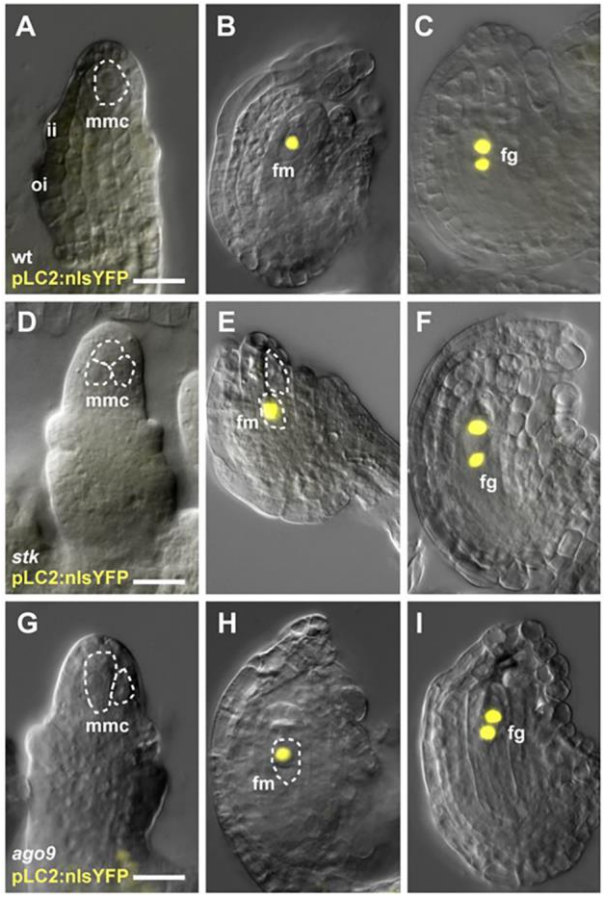


Figure 4*. pLC2::nlsYFP* expression analysis.
(A-C) wild-type; (D-F) *stk* and (G-J) *ago9-2*. Ovules were examined for *pLC2:nlsYFP* expression at the MMC stage (A,D,G) and after meiosis (B, C, E, F, H, I, J). Expression patterns were generally indistinguishable among wild-type, *ago9-2* and *stk*, despite the presence of more than one MMC-like in the mutant backgrounds. fm, functional megaspore; fg, female gametophyte; ii, inner integument; mmc, megaspore mother cell; oi, outer integument. Dashed lines highlight the mmc, fm or ectopic enlarged cells; scale bar= 15µm.

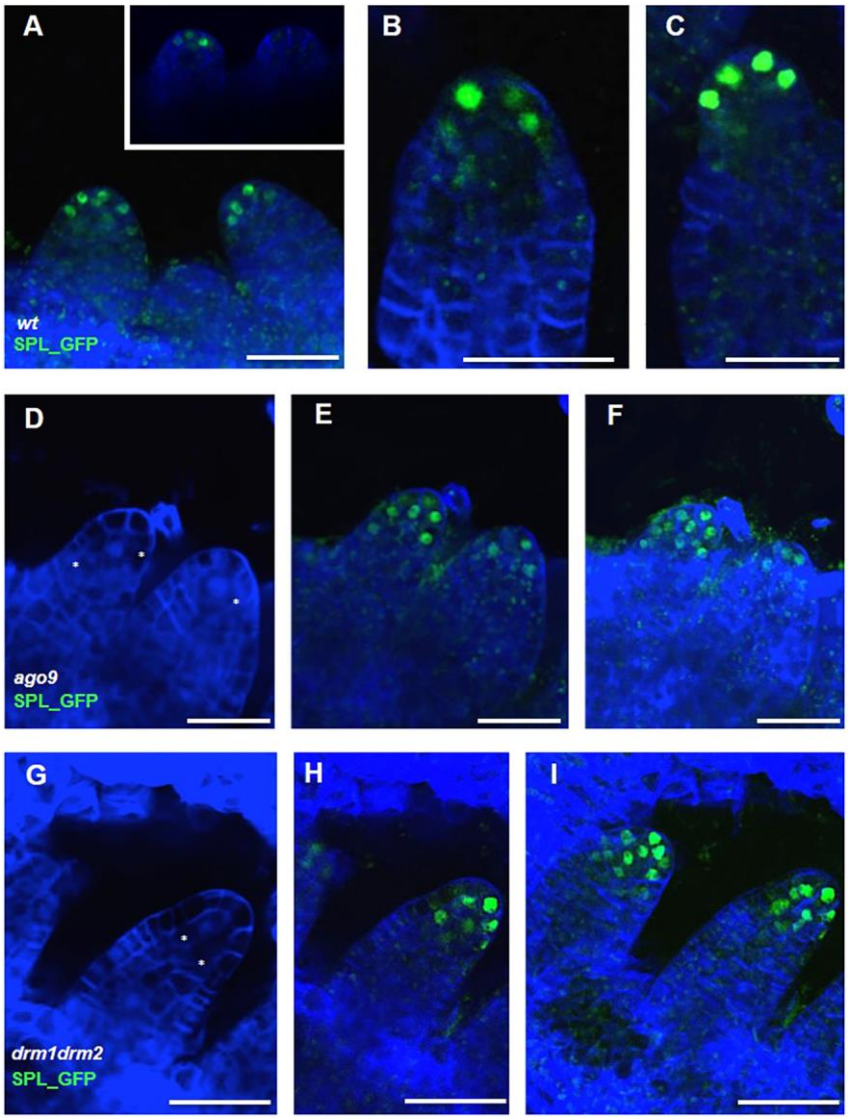


Figure 5. ***SPL_GFP* expression analysis.** Analysis of SPL/NZZ protein localization using *pSPL_5'UTR::SPL-GFP_3'UTR* . In wild-type SPL expression is confined to the tip of ovule primordium/L1 layer (A- C) and pre-meiotic ovules, detail of ovule primordium emergence in A. (D) Renaissance staining of *ago9-2* ovule primordium (E) single stack of *SPL_GFP* in *ago9-2* (F) z-stack projection of E. (G) Renaissance staining of *drm1drm2* ovule primordia (H) single stack of *SPL_GFP* in *drm1drm2* (I) z-stack projection of I. SPL expression domain resulted expanded to the lower nucellus layers in the analysed mutants; scale bar= 15 μm; Renaissance staining was used to mark the cell walls and the nuclei (white asterisks), blue in the pictures.

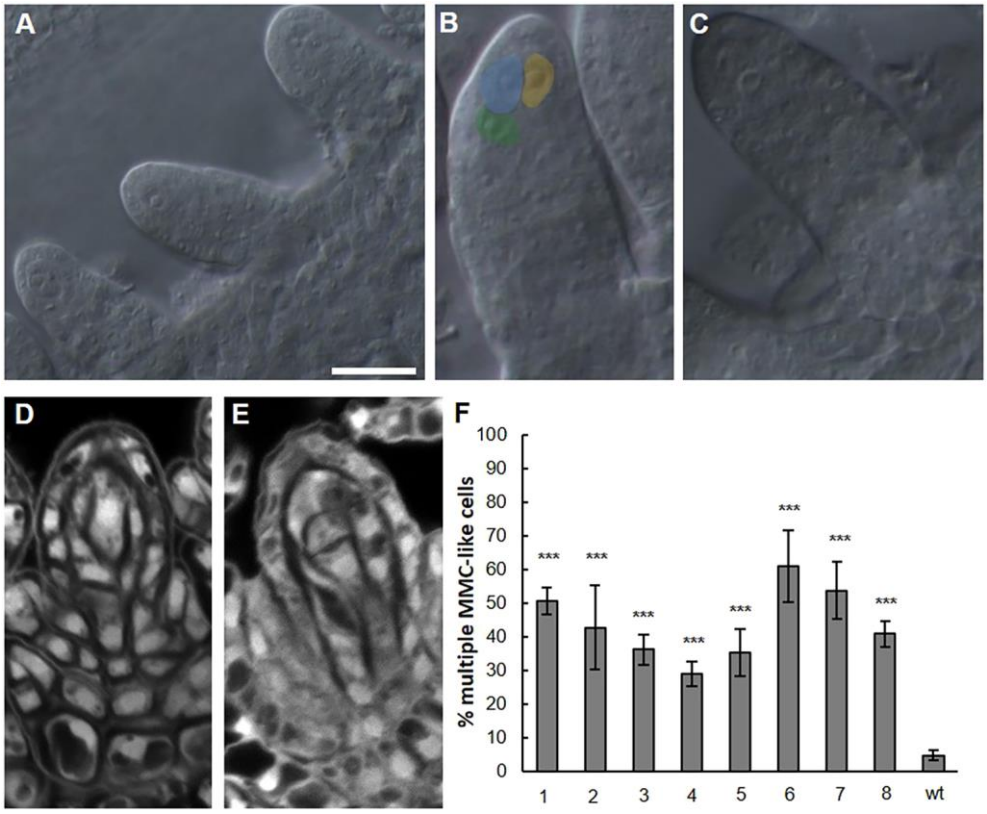


Figure 6. Constitutive expression of *SPL/NZZ* induced the formation of multiple MMC-like cells. DIC microscopy (A-C) and confocal microscopy (D-E) of *35S::SPL/NZZ* lines; multiple enlarged cells can be observed in the pre meiotic ovule nucella, highlighted with different colors. Confocal sections were acquired after Feulgen staining; scale bar= 15 μm (F) Graphic representation of the percentage of multiple enlarged cells in eight *35S::SPL/NZZ* independent lines; Asterisks (***) indicate $P<0.001$ in a Student's t-test; wild-type pre-meiotic ovules were screened as control.

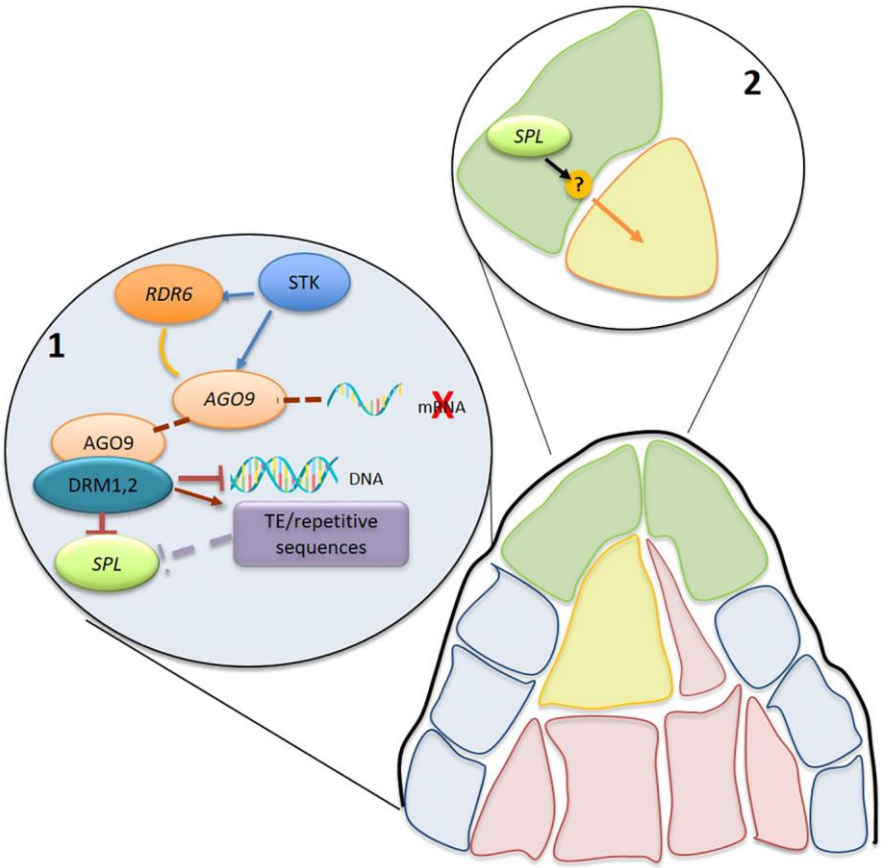


Figure 7. Proposed model of action. In the lower part of the L1 layer, STK binds directly to *AGO9* and *RDR6* promoter regions, prompting the *RdDM* pathway, which finishes on silencing TEs or repetitive sequences that on their way silence SPL/NZZ. *RdDM* pathway can also silence SPL/NZZ directly, by degrading its mRNA through *AGO9* and or by methylation of SPL/NZZ genomic region through DRM1 and DRM2 methyltransferases, that will on its way result on a repression. Altogether this leads to a direct/indirect repression of *SPL* transcription in the nucella cells, except in the apical nucella cells (L1). In the apical nucella SPL/NZZ can promote cell expansion/elongation by activation/repression of a factor that from the upper cells generates a signal for cell expansion in the cell below.

Tables

Table 1. Mutant and marker lines of Arabidopsis used in this study.

MUTANT	AT*G	ALLELE	LINE
drm1drm2	AT5G15380 (*DRM1*)	*drm1-2*	SALK_021316
	AT5G14620 (*DRM2*)	*drm2-2*	SALK_150863
ago9-2	AT5G21150	*ago9-2*	SALK_112059
spl	AT4G27330	*spl-1*	T-DNA
stk	AT4G09960	*stk-2*	TE
rdr6-11	AT3G49500	*rdr6-11*	N24285

MARKER LINE	PUBLICATION
pKNU::YFP	*Tucker et al., 2012*
pLC2::YFP	*Tucker et al., 2012*

Table 2. Primers used for qPCR experiments.

Gene name	Gene	Sequence	Amplicon size	Acquisition temperature
AGO9	AT5G21150	CACCTTTCGTTCCAGCAAAT CCGCTTGGTTTGTTGAACTT	154	79ºC
RDR6	AT3G49500		187	75ºC
SPL/NZZ	AT4G27330	CTTGGGAAGCCTTGTAGCAC AGCTCGAGCGTCAGAGAATC	148	75ºC
DRM1	AT5G15380	CCAATTCCGGATATTGCTCGTG TTGGCCCAAACACCTTTTGG	80	75ºC
DRM2	AT5G14620	TGTGCTGCTCAAATGGCTAG GCTCTCCTTTTGACTCAAACCG	106	75ºC
AtCyclophilin	AT2G36130	TGGCGAACGCTGGTCCTAATAC A GTCAGCCAAGTCAACAACTCTC TG	223	79ºC
AtGAPdH	AT3G26650	TGGTTGATCTCGTTGTGCAGGT CTC GTCAGCCAAGTCAACAACTCTC TG	262	79ºC
AtActin	AT5G43500	GAGTTCTTCACGCGATACCTCC A GACCACCTTTATTAACCCCATTT ACCA	180	79ºC
AtTubulin	AT1G50010	ATGTGGGTCAGGGTATGGAA	143	79ºC

Table 3. Primers used in ChIP experiments:

Primer name	Gene	Sequence	Amplificate size	Objective
RT_2905	AT5G21150	GTCACATCACGTTAAATCA	178	*pAGO9*
RT_2905		GTGGAAACTGTTTAACC		region 1
RT_2907	AT5G21150	AACGACGACCTGCAAAAC	300	*pAGO9*
RT_2908		TGTCATCACCTCAAATTTG		region 2
RT_2909	AT5G21150	CACAATTGTTTATTGGAACAC	137	*pAGO9*
RT_2910		CACAAAAAGAGCTATATGAAC		region 3
RT_2911	AT5G21150	GAACAACATCTTTAGCAC	220	*pAGO9*
RT_2912		CAGGTACTTAAACCGGTTATTC		region 4
RT_2293	AT3G49500	CCTTCAGTCTGTTCTCTGTTGC	180	*pRDR6*
		T		region 1
RT_2294		TGTGCTGCTCACGTGCTATT		
RT_2295	AT3G49500	CAGTCTCCGCATTGCTTCTAT	150	*pRDR6*
RT_2296		CGATCACCTAATACATTGTTGA		region 2
		TTG		
RT_2297	AT3G49500	CAATCAACAATGTATTAGGTGA	110	*pRDR6*
		TCG		region 3
RT_2298		CGCTACTCCGTTGGAAGATC		
RT_2299	AT3G49500	GATCTTCCAACGGAGTAGCG	120	*pRDR6*
RT_2300		CCCTCTGACCCCATTTCTC		region 4
RT_0045	AT5G09810	CGTTTCGCTTTCCTTAGTGTTA	134	*ACT 7*
		GCT		
RT_0046		AGCGAACGGATCTAGAGACTC		
		ACCTTG		
RT_0795	AT5G18000	GGGAAGGTCATGGCAAGTTA	170	*pVDD*
RT_0796		CCATCTGCCTCGAATATGGT		

RT_2915	AT4G27330	ccaagattaacatggaggca		*pSPL reg 1 fw*
RT_2916		ctcactatacttagctgac		*pSPL reg 1 rv*
RT_2917	AT4G27330	gtctgagtgtttaggttag		*SPL 3-UTR fw*
RT_2918		CCATTGATTTGTCCTTGAAGC		*SPL 3-UTR rv*

Table 4. Primers used for cloning

Primer name	Gene	Sequence	Amplificate size	Objective
ATP_6475	AT4G27330	GGGGACAAGTTTGTACAAAAAAGCAGGCTcaATGGCGACTTCTCTCTTCTTC	1171	SPL locus
ATP_6476	AT4G27330	GGGGACCACTTTGTACAAGAAAGCTGGGTcGAAGCTTCAAGGACAAATCAATG		
ATP_6475	AT4G27330	aacgacgacctgcaaaac	1850	SPL-GFP
ATP_6548	AT4G27330	tgtcatcacctcaaatttg		
ATP_6234	5'_AT4G27330	GGGGACAACTTTGTATAGAAAAGTTGtagaatgcaatacatggctg	3906	5'SPL (attB4)
ATP_6235		GGGGACTGCTTTTTTGTACAAACTTGtgatgatgatcttcttctcgga		5'SPL (attB1r)
ATP_6236	3'_AT4G27330	GGGGACAGCTTTCTTGTACAAAGTGGatgttttatctttctatattg	1609	3'SPL (attB2r)
ATP_6237		GGGGACAACTTTGTATAATAAAGTTGgatcgtcttcttcttcttctctgc		3'SPL (attB3)
ATP_1575	β-glucuronidase (GUS)	CCACTGTTACGTCCTGTAGAAACCCC	1809	GUS for TOPO cloning
ATP_1576	β-glucuronidase (GUS)	TCATTGTTTGCCTCCCTGCTG		GUS rev

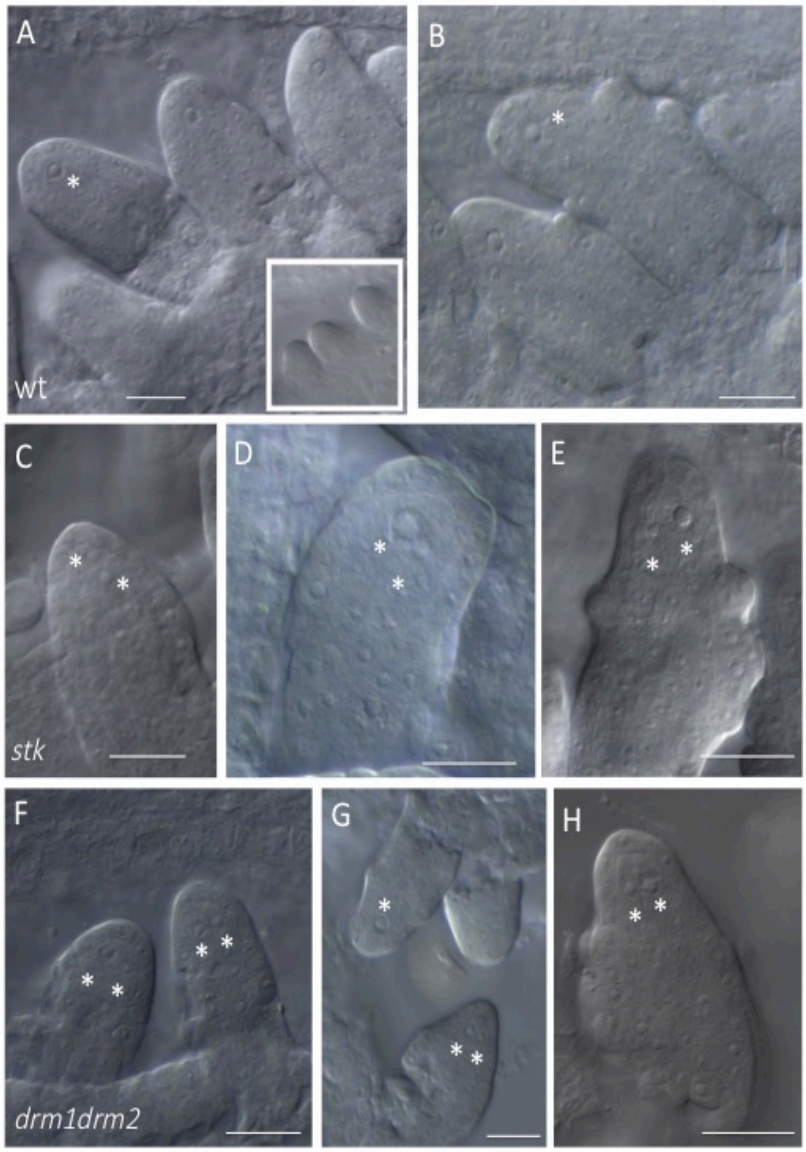


Figure S1. DIC analysis of ovule primordium. (A-B) wild-type, during the first stages of ovule primordia development any enlarged cell are detected - small detail in A, only after the ovule elongation, at finger like stage, a single enlarged cell is detectable corresponding to the MMC is detectable; (C-F) *stk* mutant; (G-I) *drm1drm2* double mutant. Multiple enlarged cells - MMC-like cells - were detected in several mutant ovules nucellus and are marked with a white asterisks.

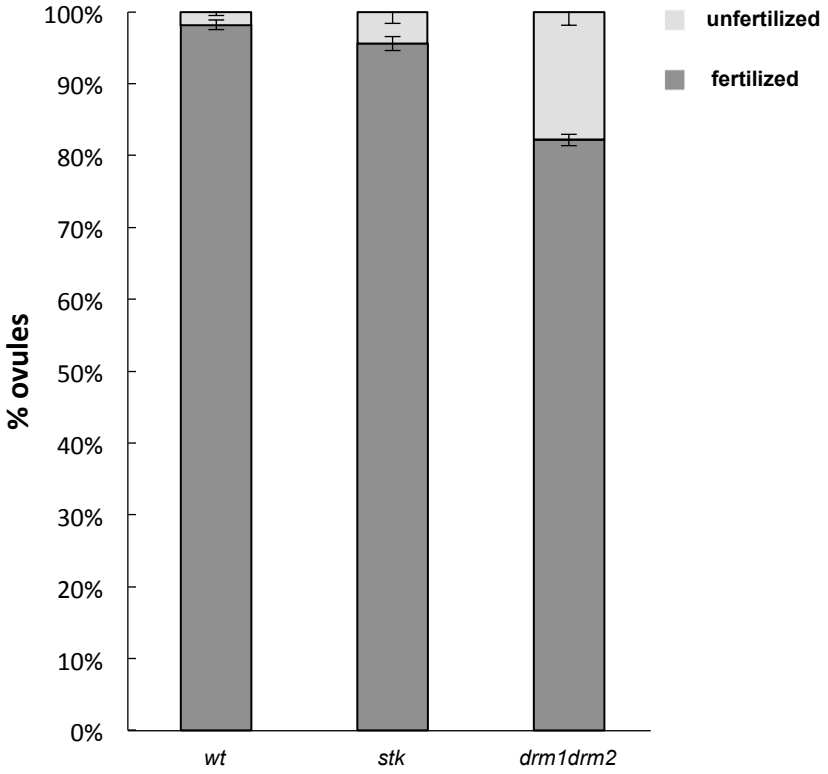


Figure S2. Fertility analyses in the selected mutants. The number of fertilized and unfertilized ovules were counted in *stk* mutant and *drm1drm2* double mutant in comparison with wild-type.

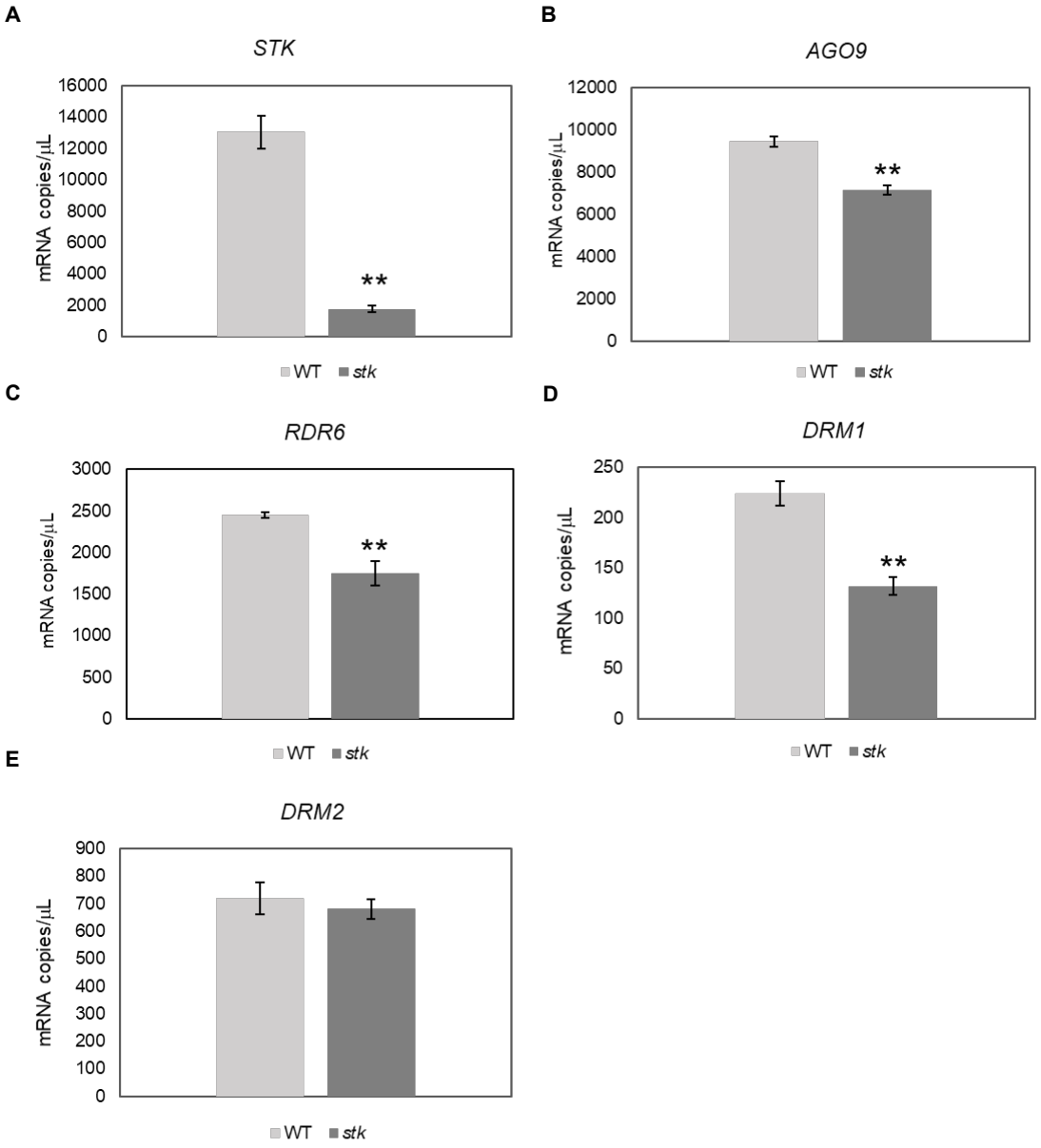


Figure S3. Gene expression analysis in wt and *stk* inflorescences and flowers. (A-E) Quantitative PCR was used to examine expression of *STK* , *AGO9*, *RDR6*, *DRM1* and *DRM2* in whole inflorescences. Asterisks (**) indicate significant differences in *stk* compared to wt and represent a *p*-value < 0.01 in a Student's t-test.

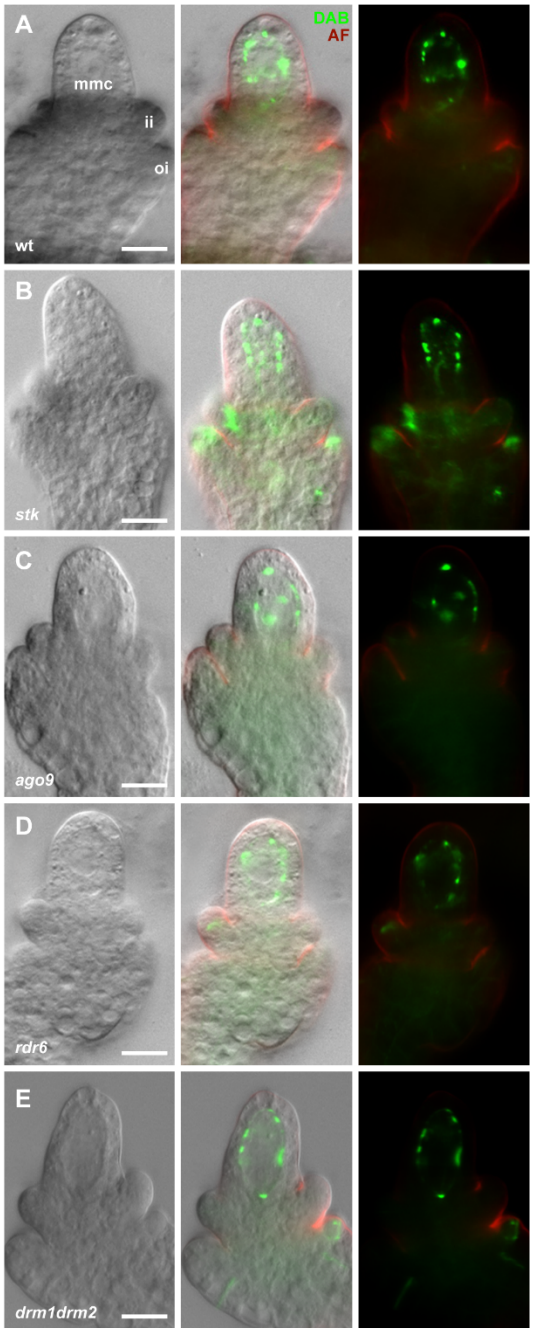


Figure S4. Callose accumulation during megaspore mother cell expansion.
Ovule from wt (A), *stk* (B), *ago9* (C) *rdr6* (D) and *drm1drm2* (E) mutant backgrounds were stained with decolourised aniline blue (DAB) and examined at stage 2-I (MMC stage); representative ovules are shown. Callose staining is indicated in green, while autofluorescence (AF) is indicated in red; ii, inner integument; mmc, megaspore mother cell; oi, outer integument. Scale bar = 12µm.

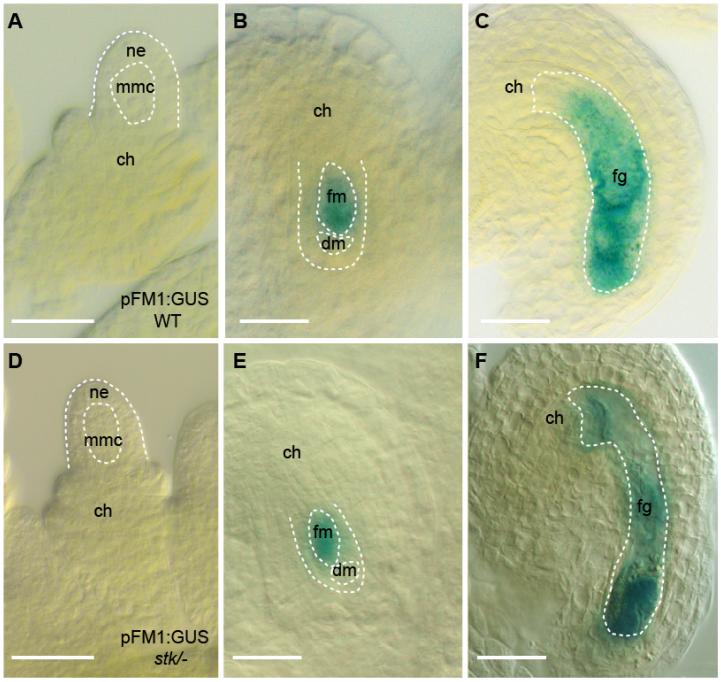


Figure S5. *pFM1::GUS* expression analysis. (A-C) In wild-type the marker is absent at the MMC stage an it is only detectable at the functional megaspore (fm) and in the developing female gametophyte (fg). (D-F) In *stk* mutant any differences were detected.

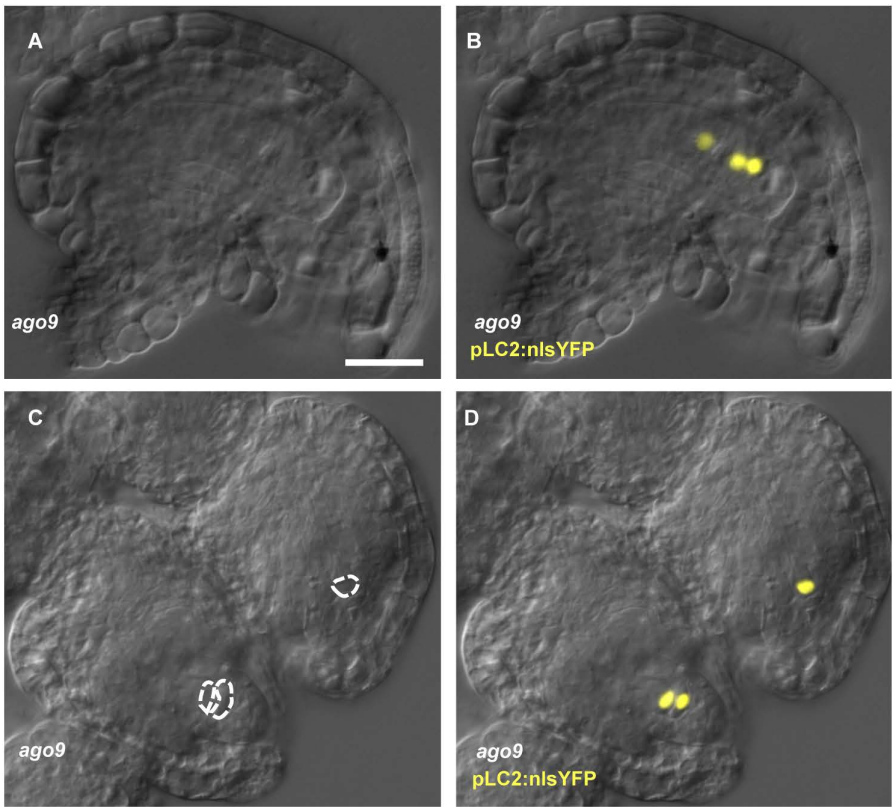
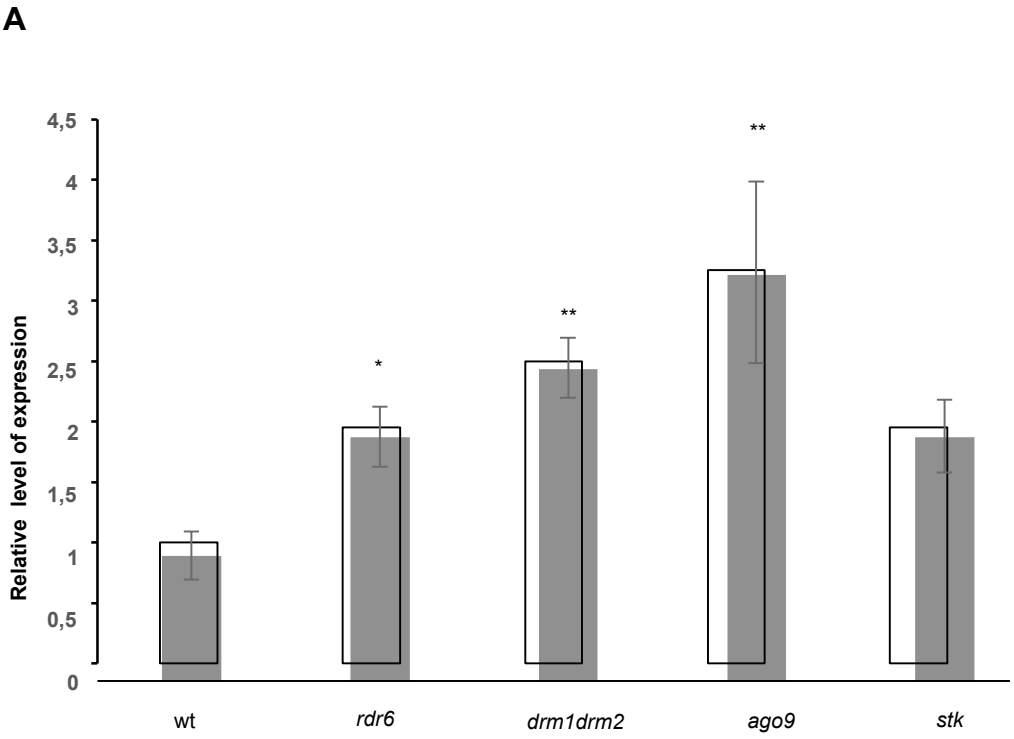


Figure S6. Callose accumulation during megaspore mother cell expansion. Ovule from wt (A), stk (B), ago9 (C) rdr6 (D) and drm1drm2 (E) mutant backgrounds were stained with decolourised aniline blue (DAB) and examined at stage 2-I (MMC stage); representative ovules are shown. Callose staining is indicated in green, while autofluorescence (AF) is indicated in red; ii, inner integument; mmc, megaspore mother cell; oi, outer integument. Scale bar = 12μm.



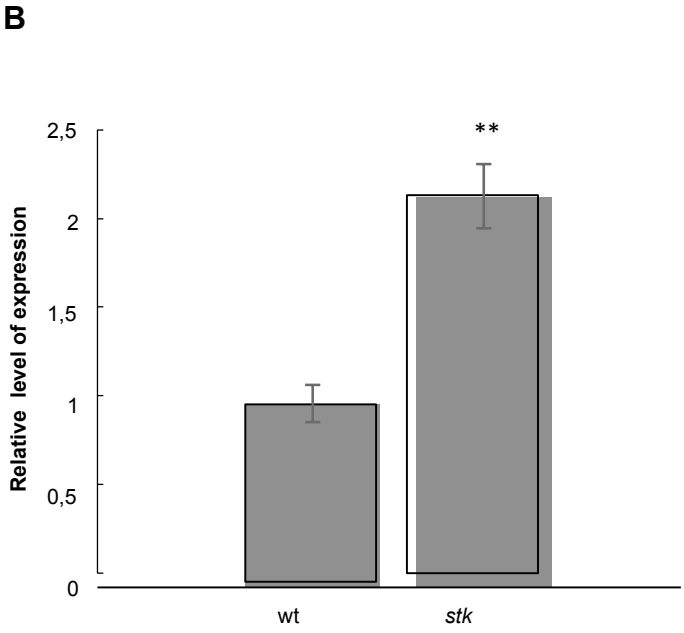


Figure S7. **qRT-PCR analysis of *SPL/NZZ* expression.** (A) Expression analysis of SPL/NZZ by qRT-PCR in wild-type, *rdr6, drm1drm2, ago9 and stk* closed flowers. The expression of *SPL* was normalized to that of ubiquitin and the expression level in wild-type was set to 1. Error bars represent the standard error mean of three biological replicates; analysis of variance (ANOVA) and post-hoc Tukey honestly significant difference (HSD) test were used; Asterisks (*) and (**) indicate significant differences in the mutants background compared to the wild-type and represent a p-value < 0.05 and p < 0.01, respectively. (B) Expression analysis of SPL/NZZ by qRT-PCR in wild-type and stk mutants in selected stages, encompassing megasporogenesis. Asterisks (**) indicate significant differences in the mutant background compared to the wild-type and represent a p-value < 0.01.

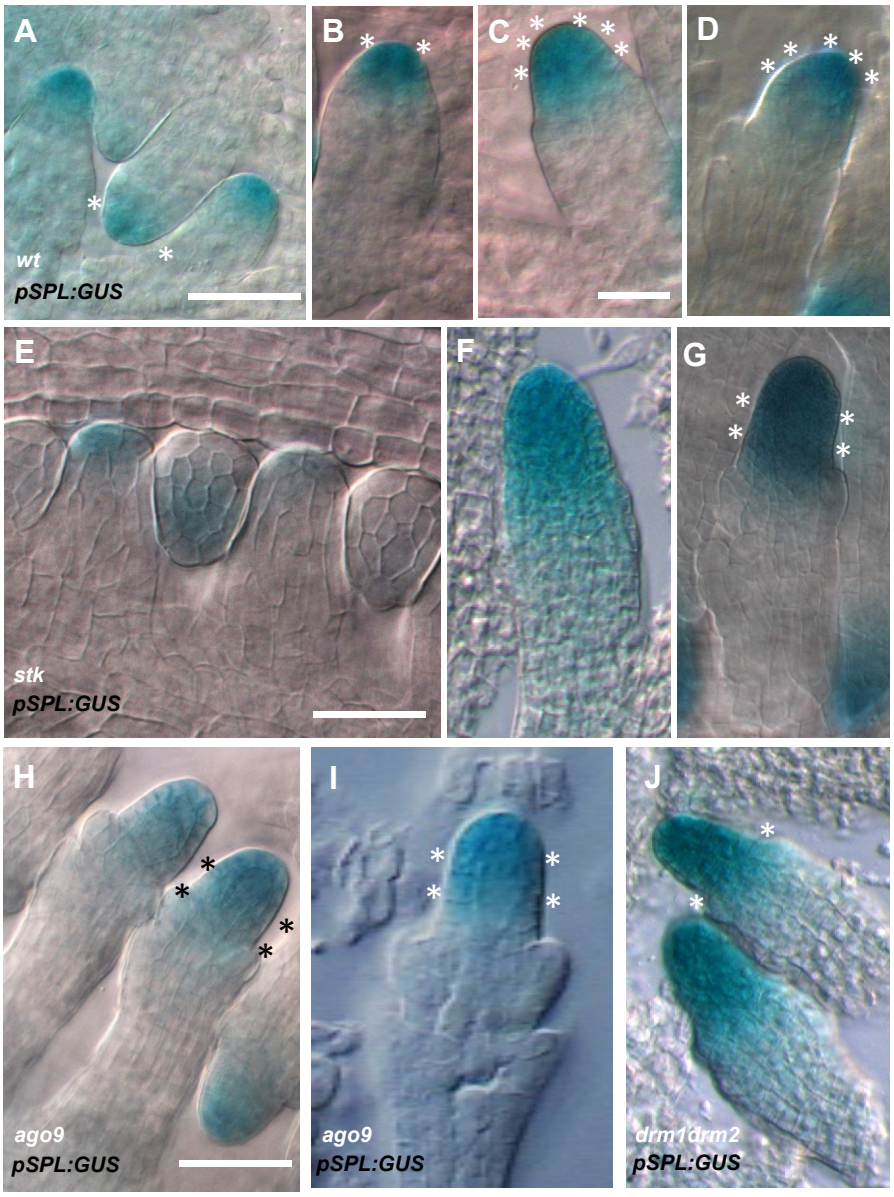


Figure S8. Mutation in STK and in the rdDM pathway affects pSPL expression domain. Analysis of *pNZZ/SPL5'::GUS:3 '* expression in wt (A-D); *stk (*E-G), *drm1drm2, ago9-2* (H-I) and *drm1drm2 (J)* mutants background; scale bar= 15 µm. Asteriks highlight the specific expression domain of pSPL, confined to L1 nucella; in the mutant backgrounds analysed we could observe an extention of the domain, no longer restricted to tip of the nucellus.

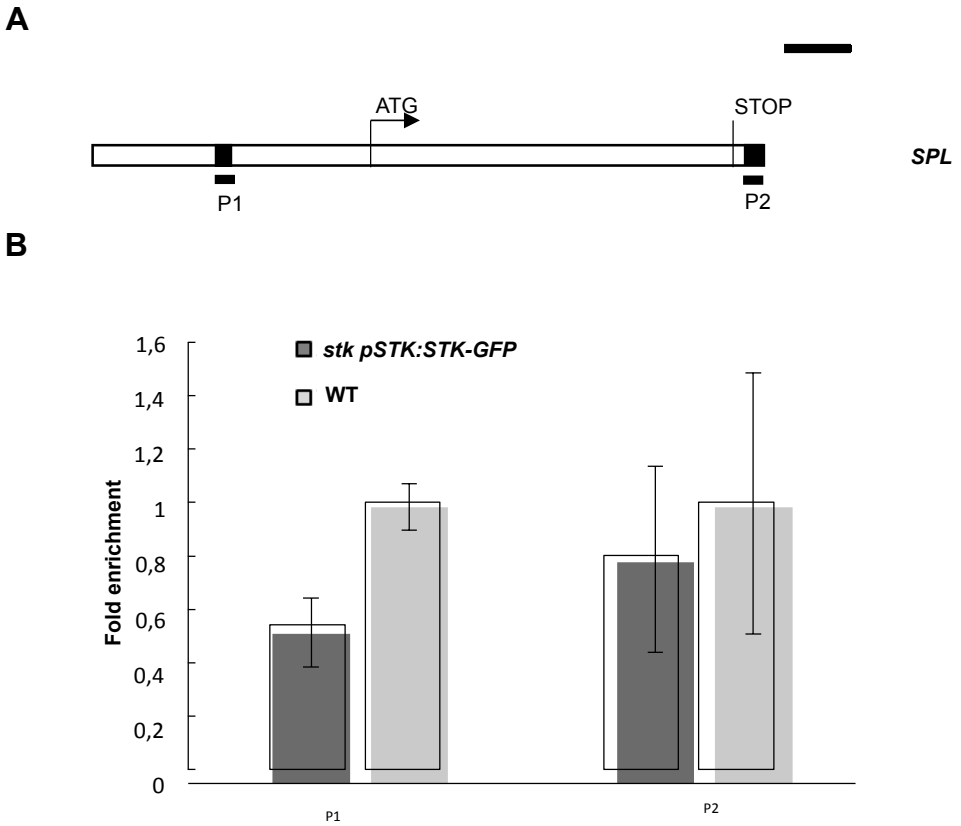


Figure S9. STK do not directly bind to the *SPL* locus. (A) Schematic diagram of the *SPL* locus, indicating the regions analysed by chromatin immunoprecipitation (ChIP; black bars); black boxes, CArG boxes; scale bar =500 bp. (B); Quantitative Real-Time PCR analysis of ChIP assay using chromatin extracted from *pSTK::STK_GFP* and wt (as a negative control), testing the CArG boxes regions on *SPL* locus. For the IP, commercial antibodies against GFP were used; Error bars represent the propagated error value using three replicates. ChIP results of one representative experiment are shown. Positive binding site fragments were considered only if they were enriched compared with the controls in at least three independent experiments.

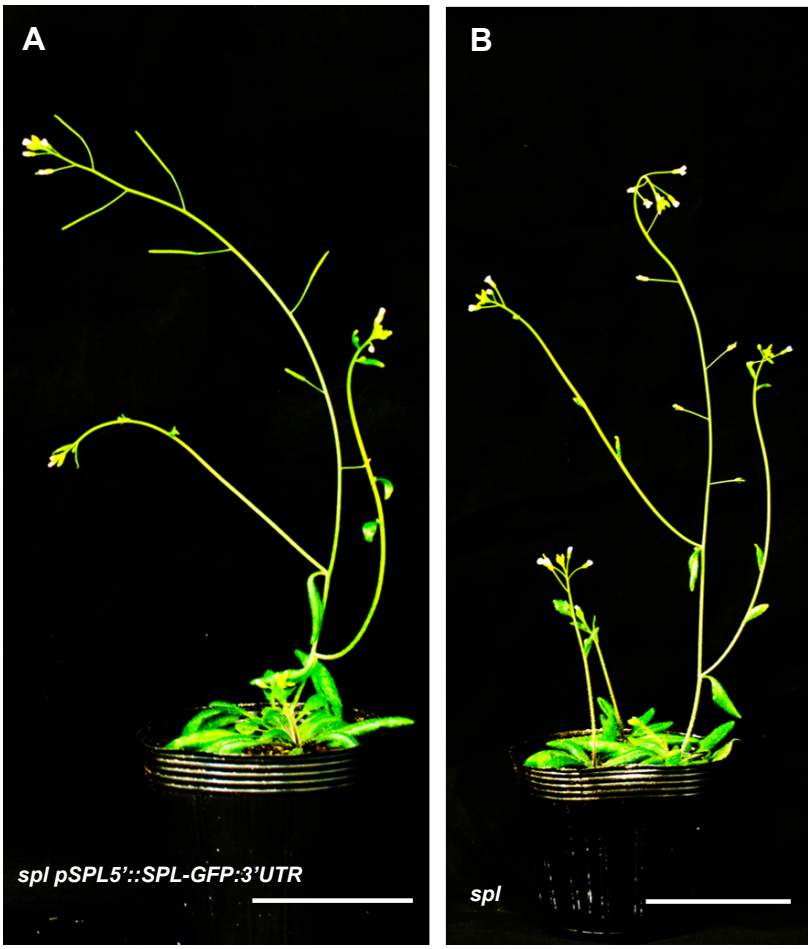


Figure S10. **Complementation analysis**. (A) *spl/nzz* mutant complemented with *pSPL5'UTR::SPL-GFP:3'UTR*; (B) *spl* mutant; scale bar=5 cm.

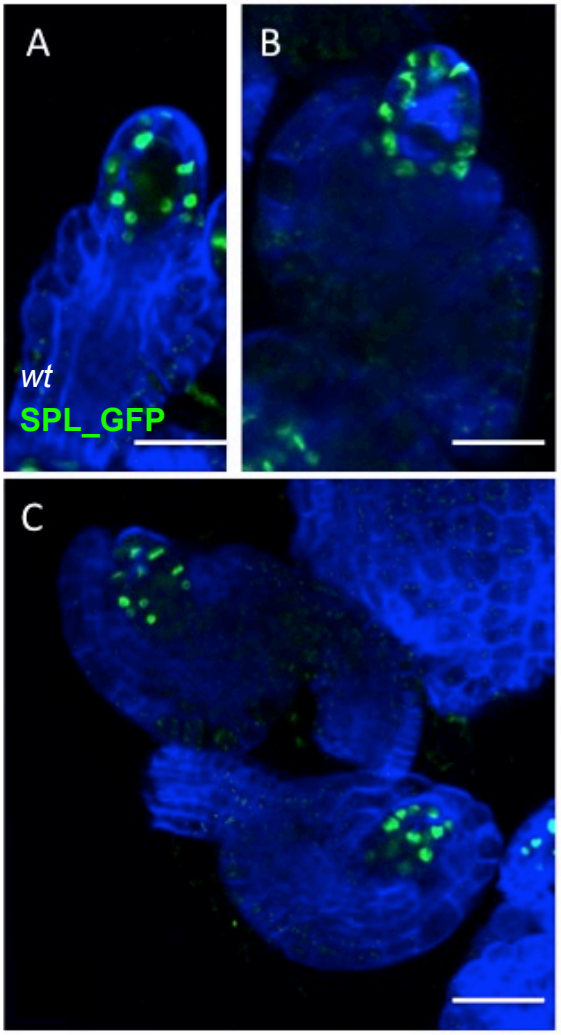


Figure S11. *SPL_GPF* expression analysis during meiosis. Analysis of SPL/NZZ protein localization using *pSPL_5'UTR::SPL-GFP_3'UTR* .(A-C) In wild-type SPL expression during meiosis is restricted to the layers that involve the MMC that goes in meiosis. In blue Renaissance staining.

Table S1. *pKNU::nlsYFP* analysis in all mutant backgrounds

Mutant	Single cell	Two cells with expression	No expression (or v. weak)	Total
stk	290	0	178	468
	62%	**0%**	**38%**	
SD	3%		2%	
wt	411	0	51	462
	89%	**0%**	**11%**	
SD	2%		2%	
drm1drm2	349	0	222	571
	61%	**0%**	**39%**	
SD				
wt	142	0	11	152
	93%	**0%**	**7%**	
ago9-2	623	12	48	683
	91%	**2%**	**7%**	
SD	3%	0.5%	2%	
wt	659	15	14	688
	96%	**2%**	**2%**	
SD	1%	0.5%	2%	
rdr6-11	280	0	70	350
	80%	**0%**	**20%**	
SD	3%		3%	
wt	99	0	11	110
	90%	**0%**	**10%**	
SD	1%		2%	