UNIVERSITÀ DEGLI STUDI DI MILANO

Molecular and Cellular Biology PhD School

XXXII Ciclo

BASIC PENTACYSTEINE PROTEINs: understanding their roles and molecular mechanisms during reproductive development in *Arabidopsis thaliana*

Rosanna Petrella

PhD Thesis

Scientific tutor: Veronica Gregis

Academic year: 2018/2019

SSD: BIO/18;BIO/01

Thesis performed at University of Milan, Bioscience Department

Part I

1.Abstracts

1.1 Studio del ruolo delle BASIC PENTACYSTEINE PROTEINs nello sviluppo del fiore e del frutto in *Arabidopsis thaliana*

Il mio lavoro di dottorato si basa prevalentemente sullo studio di una classe di fattori trascrizionali, le BASIC PENTACYSTEINE PROTEINS (BPCs) in *Arabidopsis thaliana*. I BPCs possono legare il DNA al livello di specifiche sequenze consenso, le C-boxes, per regolare l'espressione di geni target.

Il primo obiettivo della mia tesi è quello di identificare il meccanismo molecolare di azione dei BPCs nella regolazione del gene d'identità dell'ovulo *SEEDSTICK (STK)* durante lo sviluppo del fiore. A tale scopo, sono stati condotti esperimenti di immunoprecipitazione della cromatina (ChIP) in diversi background, in cui è stata valutato il legame dei BPCs alla sequenza regolatrice di *STK*. Inoltre, l'analisi dell'espressione del gene omeotico nel quintuplo mutante *bpc (bpcV)* ha confermato una sua regolazione da parte dei BPCs e svelato il ruolo ridondante dei BPCs di classe I e classe II. I nostri risultati indicano che i BPCs possano regolare *STK*, interagendo con il fattore MADS-domain SVP, la cui attività è necessaria per la corretta espressione del gene omeotico. Inoltre, il mio lavoro ha permesso di identificare un altro regolatore di *STK*, LIKE HETEROCHROMATIN PROTEIN 1 (LHP1), un componente del POLYCOMB REPRESSIVE COMPLEX 1 (PRC1). I dati, presentati in questa tesi, suggeriscono che i BPCs possano formare un super complesso proteico per regolare *STK* durante lo sviluppo del fiore.

I BPCs sono espressi in tutti i tessuti della pianta durante lo sviluppo. Il gineceo del quintuplo mutante *bpc* è caratterizzato da difetti nello sviluppo del septum e del transmitting tract, una struttura che riveste un ruolo fondamentale durante la fecondazione. Il ruolo dei BPCs nello sviluppo del gineceo è stato studiato attraverso diversi esperimenti, con lo scopo di identificare i loro target e partners per localizzarli nel network di regolazione dello sviluppo del septum. Inoltre, l'analisi trascrittomica sul *bpcV* ha permesso di identificare un possibile contributo dei BPCs nella regolazione dei livelli di auxina e citochinine, due fitormoni, la cui attività è fondamentale per il corretto sviluppo del gineceo. In conclusione, il lavoro presentato nella mia tesi approfondisce i meccanismi molecolari che governano lo sviluppo riproduttivo nelle piante.

1.2 Studying the role of BASIC PENTACYSTEINE PROTEINs, SHORT VEGETATIVE PHASE and LIKE HETEROCHROMATIN PROTEIN1 in the regulation of *SEEDSTICK* expression during flower development

The GAGA binding proteins BASIC PENTACYSTEINE PROTEINS (BPCs) are transcription factors, present in several plant species. The Arabidopsis BPCs family is composed of seven members that have been divided in three different classes, based on their protein sequence similarity (Meister et al., 2004). BPCs are ubiquitously expressed in the plant and analysis of the mutants reveal a redundant function in different stages of plant development (Monfared et al., 2011). BPCs directly bind *SEEDSTICK (STK)* promoter on C-boxes, whose mutations affects its expression in the flower (Simonini et al., 2012). Another important regulator of *STK* is the MADS-domain factor SHORT VEGETATIVE PHASE (SVP); it represses *STK* expression during the first stages of flower development, by direct binding to its promoter on specific MADS-domain binding sites, named CArG-boxes (Simonini et al., 2012; Gregis et al., 2013). Here we show that MADS-domain binding sites on the *STK* promoter region are important for the correct spatial and temporal expression of the homeotic gene. We also proved that BPCs of class II and SVP can form hetero-dimers. Our data clarify the role of BPCs in regulating *STK* expression, mostly shedding light in the molecular mechanisms by which BPCs and SVP act. We found that BPCs of class I and II act redundantly and together with SVP to repress *STK* expression in the flower; these factors can interact with a member of POLYCOMB REPRESSIVE COMPLEX1 (PRC1), LIKE HETEROCHROMATIN PROTEIN1 (LHP1). We propose that the complex acts maintaining and spreading the H3K27me3 repressive mark to the regulatory region of *STK* in reproductive tissues. We further characterize the role of LIKE HETEROCHROMATIN PROTEIN1 (LHP1) in *STK* expression. Our results investigate the involvement of a protein complex in which BPCs, MADS-domain factors and LHP1 cooperate to regulate the expression of homeotic genes during development and pave the way for a more comprehensive understanding of gene regulation in plants.

1.3 Unravelling the role of BASIC PENTACYSTEINE PROTEINS during gynoecium development in *Arabidopsis thaliana*

The GAGA binding protein BASIC PENTACYSTEINE PROTEINS (BPCs) belong to a class of transcription factors, present in several plant species. The Arabidopsis BPCs family consists of seven members that have been divided in three different classes, based on their protein sequence similarity (Meister et al., 2004). BPCs are ubiquitously expressed in the plant and analysis of the mutants reveal a redundant function in different stages of plant development (Monfared et al., 2011). Mutation of the five BPCs affects a wide range of processes (i.e. fruit development). Studying molecular mechanisms involved in fruit development is fundamental because fruits are important for biological and agronomic reasons: the better understanding of this process may give an advancement of knowledge for future noteworthy crop improvement; besides, fruit provides an excellent model for studying organogenesis in plants. Analyses of the quintuple *bpc* mutant *(bpcV)* reveal that the fruit, smaller in size, has defects in the formation of the septum, an important structure that divides the ovary in two halves and play a crucial role during fertilization; in particular, we detect a defect in the fusion of the septum, following by a lack of differentiation of the transmitting tract cells, important for the correct growth and guidance of the pollen tube from the style to the ovary. Our transcriptomic data identify *SPATULA (SPT)* as a putative BPCs target, whose expression is strongly downregulated in the *bpcV* mutant. Furthermore, our genome-wide transcriptomic analysis deepens the role of BPCs of class I and II in septum and transmitting tract development and investigate their involvement in the regulation of phytohormones auxin and cytokinin pathways, whose correct balance is pivotal to assure the correct gynoecium development. Our results confirm a role of BPCs in gynoecium development and provide insight into the transcriptional regulatory networks that control fruit development in Arabidopsis.

2. Aim of the thesis

Organogenesis involves different complex transcriptional networks that must be strictly modulated throughout development. Homeotic genes play key roles in plant, animal, and fungal development, since they are necessary to define organ identity. The Drosophila GAGA Factor (GAF) has been identified as regulator of homeobox genes expression as well as a variety of developmental processes. Even though there are no homologs of GAF in plants, BRR/BPC proteins appeared to be functionally convergent. In Arabidopsis, BASIC PENTACYSTEINE PROTEINs (BPCs) have been characterized as regulators of different targets, whose activity has key roles during plant development; despite the increasing knowledge on this transcription factory family, the molecular mechanisms by which BPCs act is far to be completely understood. Our work wants to unveil and update the state of the art on this intriguing family of plant regulators.

We want to clarify the involvement of BPCs in the regulation of the previously reported target, the homeotic gene *SEEDSTICK (STK),* whose activity determines ovule identity. It is an excellent model to study homeotic genes regulation since its expression is finely confined to ovule and placenta; also, considering *STK* key role during ovule and seeds development, studying how it is regulated in Arabidopsis could further give an improvement in controlling seeds production in crops. To better understand the role of BPCs in *STK* regulation during flower development, we directly checked the spatial and temporal expression of the homeotic gene in different *bpc* mutant backgrounds, thus determining the contribution of the two classes of BPCs in the modulation of *STK* expression. BPCs accounted for several activities as the Drosophila counterpart GAF, corroborating the hypothesis that they might interact with different partners with specific and dedicated functions throughout plant development. Hence, we further investigated their relationship with the MADS-domain factor SHORT VEGETATIVE PHASE (SVP), a known regulator of *STK* expression in the flower, adopting different approaches. We gained insight into the complex in which BPCs and SVP might act by evaluating their interaction with different protein interaction assays; also, we explored the molecular processes that drive *STK* expression in the inflorescence by analysing the mechanism of binding of the TFs to the promoter region in several distinct genetic backgrounds. Recently, BPCs have been physically associated with Polycomb response elements (PREs), stressing their involvement in mediating epigenetic regulation of their targets, most likely

by the recruitment of POLYCOMB REPRESSIVE COMPLEX (PRC) proteins. Fascinatingly, *STK* showed a high coverage of H3K27me3 repressive mark on its locus. Hence, we evaluated BPCs contribution in the establishment of the H3K27me3 on *STK* regulatory region by analysing the deposition of the repressor mark at the homeotic gene locus in the *bpc* quintuple mutant background. Also, the role of LIKE HETEROCHROMATIN PROTEIN1 (LHP1), a member of the POLYCOMB REPRESSIVE COMPLEX1 (PRC1), in *STK* regulation was investigated.

BPCs are ubiquitously expressed in the plant and *bpc* multiple mutant showed a pleiotropic phenotype, thus suggesting their key roles throughout plant development. Intriguingly, our *bpc* quintuple mutants revealed some defective traits in gynoecium development. The gynoecium is the female reproductive part of the flower and its correct formation is necessary for reproduction success. Therefore, we further explored the role of BPCs in gynoecium formation, focusing in septum development and successively in transmitting tract differentiation. To fulfil this aim, we performed a morphological analysis of the multiple *bpc* mutant inflorescences during the whole gynoecium development. Hereafter, new BPCs targets and interactors were searched and successively analysed. Finally, the genetic pathway in which BPCs might act for the fulfilment of a correct gynoecium development and the processes that they might control, were further investigated by a transcriptomic analysis performed on the reproductive tissues from the *bpc* quintuple mutant background. In conclusion, here my PhD thesis provides new insights into the molecular mechanisms of gene regulation in plant, mostly by unravelling the role of BPCs during flower and fruit development in *Arabidopsis thaliana*.

3.Introduction

3.1 *Arabidopsis thaliana* as a model species

All the experiments in my thesis are performed on the flowering plant *Arabidopsis thaliana,* a small dicotyledon, belonging to the Brassicaceae family (Figure 1). It is widely used as a model plant in molecular biology, functional genomics, biochemistry and physiology for several reasons: small size, allowing the growth in laboratories and greenhouses; short life cycle (six to eight weeks); high seed yield; self-pollination; easily and efficiently transformed by *Agrobacterium tumefacens* infection; extended mutant collections available, and last but not least, a relatively small (135 Mb) and fully sequenced and annotated genome (Arabidopsis Genome Initiative, 2000). Arabidopsis genome is organized in five chromosomes and it encodes for cc.27000 coding genes, many of them already been characterized. Even though the molecular and functional characterization of this model plant species genes has no direct agricultural application, modern breeding has taken advantage of the massive amount of literature available. Genetic comparisons between *Arabidopsis* and crop species are common: genes that have been characterized in Arabidopsis indeed can be used to find their homologs in other species. On top of that, fundamental mechanisms can be understood in the model plant and used to further our knowledge on crops.

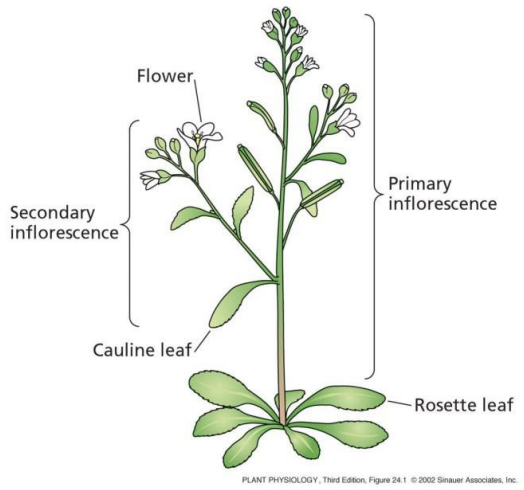


Figure 1. Arabidopsis thaliana (Plant Physiology)

3.2 The transition to flower

The life cycle of Arabidopsis is characterized by several phases: embryogenesis, germination, vegetative growth (leaf production, rosette growth) and reproductive growth (inflorescence emergence, flower development and silique ripening) (Boyes et al., 2001), as depicted in Figure 2.

During embryogenesis, the formation of two types of apical meristems occurs: the root apical meristem (RAM) and the shoot apical meristem (SAM). During the vegetative phase, the SAM produces leaves.

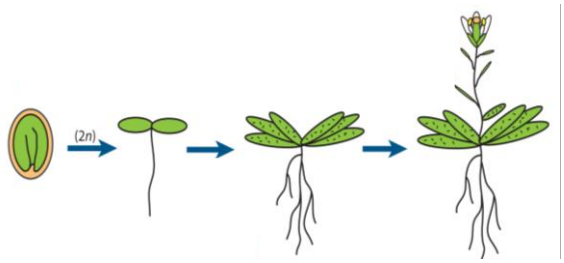


Figure 2. Schematic representation of Arabidopsis growth phases. From left to right: embryogenesis, germination, vegetative growth and reproductive growth (Mozgoya, 2015)

A combination of endogenous and environmental stimuli (photoperiod, temperature, light quality, vernalization, autonomous pathway and GA pathway) led to the switch from the vegetative to the reproductive phase, characterized by the elongation of the stem (the so-called bolting) and the conversion of the SAM in the inflorescence meristem (IM) (Weigel and Jürgens, 2002). From its flanks the floral meristem (FM) emerges (Figure 3); it is important to acknowledge that the IM is indeterminate, since it can grow indefinitely whereas the FM is determinate, terminating with the production of the female reproductive organs. The FM fate is firstly established by the action of floral meristem identity genes and later on specified by the floral-organ identity genes activity (Fletcher, 2002).

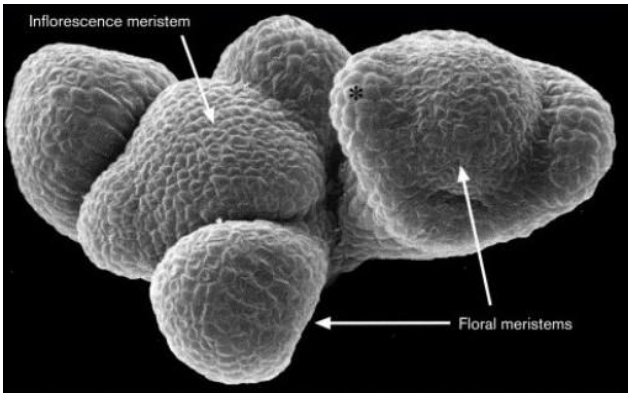


Figure 3. Scanning electron microscopy image of inflorescence and floral meristems; asterisk indicates sepal primordium (Lenhard and Laux, 1999).

3.3 The flower and the ABCDE model

Arabidopsis flowers are composed of four distinct organ types that arise in concentric structures, the whorls (Figure 4).

The outer two whorls contain non-reproductive organs: the sepals in the first whorl and petals in the second whorl. The inner two whorls contain reproductive organs: stamens in the third whorl, with anthers, that produce pollens, and carpels in the central fourth whorl which fuse to form the gynoecium; it harbours the ovules, which contain and protect the female gametophyte. After fertilization, the gynoecium and the ovules will develop in the fruit, the silique and in the seeds, respectively (Alvarez-Buylla et al., 2010).

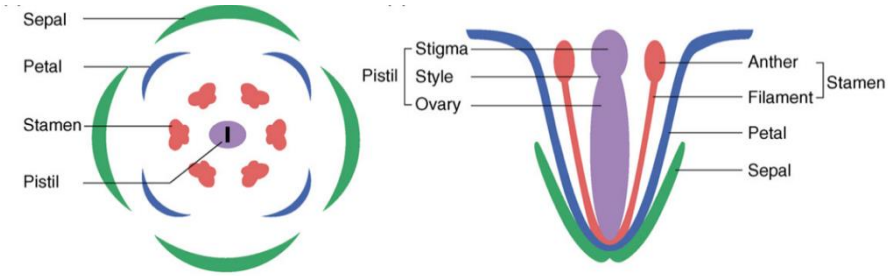


Figure 4. Schematic representation of Arabidopsis flower; at the left, transversal section; at the right, longitudinal section (Feng et al., 2007).

Through analysis of floral homeotic mutants in *Arabidopsis thaliana* and *Antirrhinum majus,* the ABC model has been formulated: it assesses that different classes of genes (class A, B and C) interact to specify floral organ identity (Coen et al., 1990). ABC model has been later extended, including other two gene classes: D and E (Colombo, 1995; Pelaz et al., 2000; Pelaz et al., 2002) (Figure 5a).

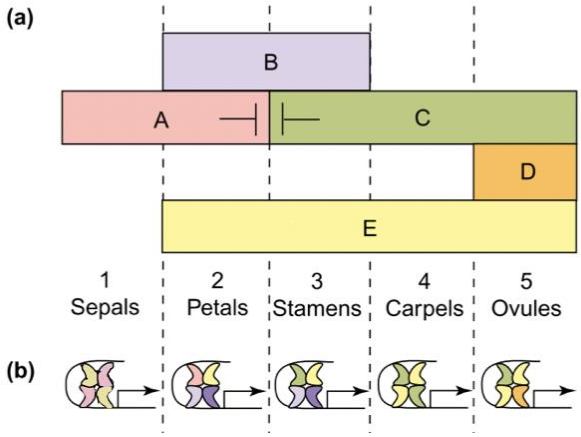


Figure 5. (a): ABCDE model for floral organ patterning in Arabidopsis thaliana. (b): floral quartet complexes (Ferrario et al., 2004).

Apart from the class A gene *APETALA2 (AP2)*, the floral organ identity genes encode for MADS-domain transcription factors (described in detail in the section below). ABCDE proteins activity in the specification of floral organs follows the floral quartet model (FQM), assessing that tetrameric MADS-domain factor complexes defined floral organs identity, as showed in Figure 5b (Honma and Goto, 2001).

3.3.1 MADS-domain factors

MADS-domain proteins have been characterized in animals, fungi, plants and algae (Smaczniak et al., 2012); they present a highly conserved DNA-binding domain of 56 amino acids, the so-called MADS-box. The acronym came from the first letter of the genes that were first characterized as MADS-domain in the relative organism, reported below:

- *MCM1* of *Saccharomyces cerevisiae*
- *AG* of *Arabidopsis thaliana*
- *DEF* of *Antirrhium majeus*
- *SRF* of *Homo sapiens*

In Arabidopsis thaliana the MADS-domain factors are specifically expressed throughout plant development to orchestrate the morphogenesis of different organs (Parenicová, 2003; Smaczniak et al., 2012a; Hugouvieux and Zubieta, 2018) . The MADS-domain family counts 107 members divided into two types (SRF-like and MEF2-like) and five subclasses (M□, M□, M□, M□ and MIKC) (Parenicová, 2003; Lai et al., 2019); a gene duplication followed by divergence is at the base of the division into type I and type II (Schwarz-sommer et al., 1992; Alvarez-Buylla et al., 2000). Flowering is one of the most characterized processes in plant development for its agronomical relevance; as matter of fact, the most studied factors are those involved in floral organ identity determination, belonging to the MIKC subclass (Alvarez-Buylla et al., 2000). They present a modular structure, made of different structural domains, illustrated in Figure 6:

- the highly conserved MADS domain (M): at the N-terminus, it is involved in DNA binding as well as in dimerization and nuclear localization; it recognizes and binds specific consensus sequences on the DNA, the CArG-boxes ($CC[A/T]_6GG$) (Schwarz-sommer et al., 1992);
- intervening domain (I): it is a weakly conserved domain that contributes to dimerization and partner selection;
- keratin-like domain (K): it is a domain involved in protein dimerization and formation of multimeric complexes;
- C-terminal domain (C): it is a variable domain, involved in transcriptional activation and multimeric complex formation, along with the K domain (Egea-Cortines et al., 1999; Yang et al., 2003; Immink et al., 2010; Theißen et al., 2016).

Figure 6. Modular structure of a plant type II MADS domain protein, with at the N-terminus the conserved DNA binding MADS motif, followed by the intervening region (I), the K-box, and the least conserved C-terminal domain. (Immink et al., 2010).

Previously, a mechanism of action of MADS-domain factors has been proposed: they associate with each other, forming quaternary complexes that bind two CArG-boxes, located in closed proximity, inducing a loop in the DNA (Honma and Goto, 2001; Theißen and Saedler, 2001; Theißen et al., 2016). According to this model, MADS-domain proteins can recruit transcriptional co-factors and chromatin remodelling proteins, acting as activators or a repressors (Smaczniak et al., 2012) (Figure 7).

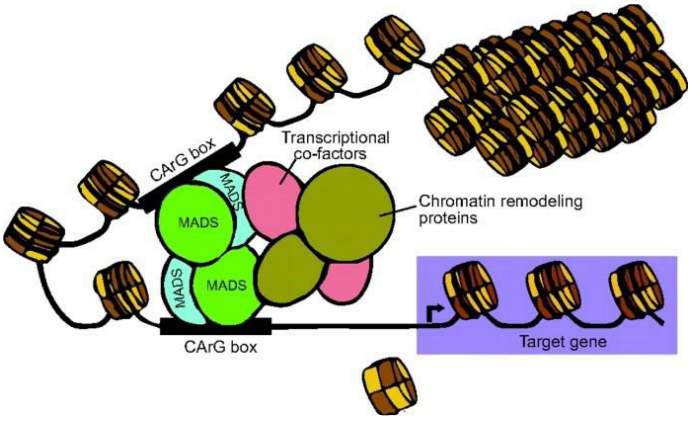


Figure 7. Model of action for the MADS-domain complex. In the following model, two MADS-domain factors interact to form a quaternary complex, binding two CArG-boxes and bending the DNA. MADS-domain proteins can recruit transcriptional co-factors and chromatin remodelling proteins to modulate the expression of target genes (Smaczniak et al., 2012)

3.4 Gynoecium development

As reported above, through the action of homeotic genes, in the centre of the flower, carpel identity is determined and gynoecium development starts. A correct gynoecium initiation and development are fundamental for a correct fruit formation (Reyes-Olalde et al., 2013). The Arabidopsis gynoecium is characterized by two congenitally fused carpels, whose fusion occurs vertically at their margins. Its inner tissue possess meristematic characteristics and it is called the carpel margin meristem (CMM); it is indeed located at the margins of the carpels and it will generate the "marginal" tissues of the gynoecium (the placenta, the ovules, the septum, the transmitting tract, the style and the stigma). The septum derives from the post genital fusion of the two developing CMMs, around stage 9 of flower development (Figure 8D). At the same time, at the flanks of the CMMs the placenta tissue starts to form. The gynoecium reaches maturity at stage 12, when the transmitting tract develops in the middle of the septum (Figure 8G). After anthesis (Figure 8H), the gynoecium is fertilised, and it differentiates into the silique.

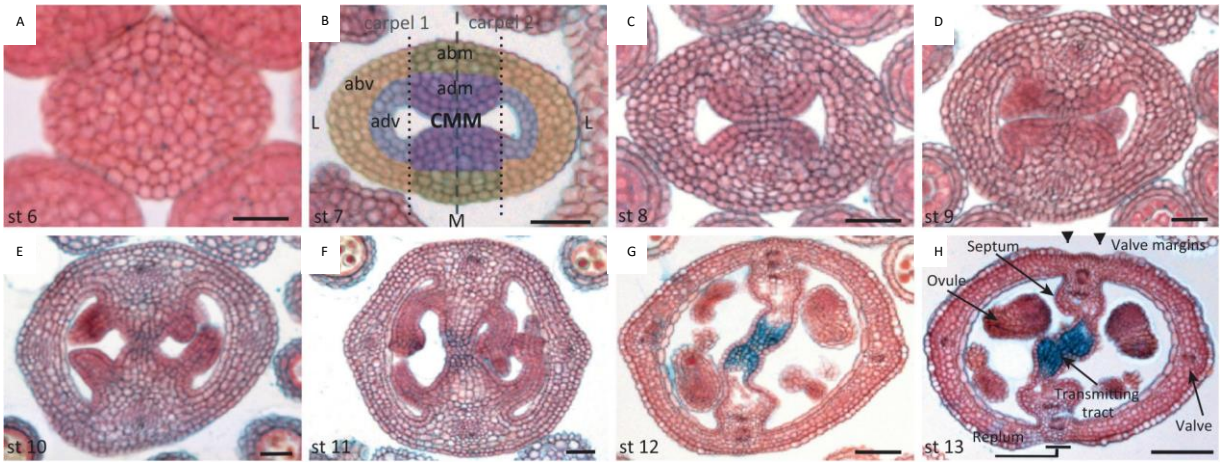


Figure 8. Transverse section of Arabidopsis gynoecium stained with neutral red (to visualize cell walls) and Alcian blue (to visualize acidic polysaccharides, major components of the extracellular matrix of the transmitting tract). Stages from 6 to 13 of flower development are represented from A to H. In B, gynoecium domains have been artificially colored: abv: abaxial valve in yellow, adv: adaxial valve in blue, abm: abaxial margin in purple, adm: adaxial margin in green, CMM: carpel margin meristem, L: lateral domains, M: medial domain. In H the carpel tissues derived from the CMMs are listed (Ovules, septum, valve margin, vales and transmitting tract). Scale bars: A = 10µm, B-F= 25µm, G = 50µm, H = 100µm (Reyes-Olande et al., 2013).

3.4.1 Networks controlling gynoecium development

The marginal tissues of the gynoecium derive from the CMMs, whose correct formation is necessary for a successful reproduction; thus, its development must be strictly regulated. The knowledge about the factors orchestrating gynoecium correct differentiation is still fragmented and poor. Recently, several transcription factors have been characterized for their important role in marginal tissues development, essential for reproduction. In particular, *SPATULA (SPT)* has been shown to have a role in septum and sequentially transmitting tract differentiation. It encodes for a basic-helix-loop-helix (bHLH) transcription factor (Heisler et al., 2001). As matter of fact, the siliques of the *spt-2* mutant is shorter and wider toward the apex, compared to a wild-type fruit (hence the spatula-like appearance); the mutant is characterized by a reduced growth of the gynoecium at anthesis and reduced development of stigmatic papillae at the tip (Figure 9A). The two carpels are often unfused on one or both sides in the upper most region; the gynoecium shows some defects in septum fusion, leading a failure in the formation of the transmitting tract, as shown in Figure 9D; although *spt-2* mutant has severe defects in the development of the pollen tract tissues (transmitting tract, style and stigma), fertilization still occurs and a small number of seeds are produced only in the upper region (Alvarez and Smyth, 1999).

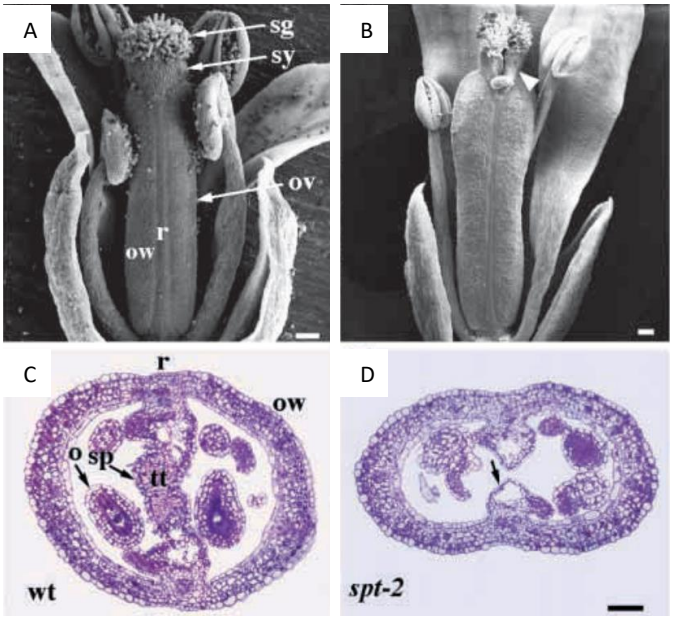


Figure 9. Structure of wild-type and spt-2 mutant gynoecium. A and B: scanning electron microscopy of flowers at anthesis. C and D: Transverse section of the gynoecium at anthesis. A and C: wild-type fruit. B and D: spt-2 mutant. In B the white arrow indicates the two unfused carpels. In F the black arrow shows the lack of transmitting tract cells. Scale bars in A and B = 100 µm, C and D = 50 µm. sg: stigma; sy: style; ov: ovary; ow: ovary walls; r: replum; o: ovules; sp: septum; tt: transmitting tract. (Alvarez et al., 1999).

SPATULA pattern of expression defines its activity during gynoecium formation; it is indeed expressed in the marginal and pollen tract tissues throughout their development (Heisler et al., 2001). *SPT* acts partially redundantly with the bHLH transcription factor *ALCATRAZ* (*ALC*), a known regulator of the separation layer during fruit dehiscence (Rajani and Sundaresan, 2001). The two bHLH transcription factors, arising from an ancestral duplication event, share a high sequence homology. Their redundant role in gynoecium development is further supported by the increased severity of

gynoecium disruption (the gynoecium is never closed, and the stigmatic tissue is almost completely lost) and valve margin defects in the *spt-2 alc-1* double mutant, compared to the single *spt-2* mutant; in contrast, *alc-1* mutant does not show any phenotype (Figure 10). As matter of fact, the expression profiles of *SPT* and *ALC*, in the gynoecium, overlap (Groszmann et al., 2011).

Figure 10. Stage 17 siliques of alc-1, spt-2 and spt-2 alc-1 mutants. The mutant alc-1 has no defects related to gynoecium development (wild-type like). In the spt-2 alc-1 double mutant the gynoecium is more unfused than in the spt-2 single mutant. Scale bar = 1mm (Groszmann et al., 2011)

Interestingly, the NAC genes *CUP-SHAPED COTYLEDON1* and *2* (*CUC1* and *CUC2*) resulted to be *SPT* interactors. They act redundantly with each other, since the single mutants do not show any phenotype (Aida, 1997). Firstly characterized for their major role in the separation of cotyledons during embryogenesis, they are also involved in the separation of sepals and stamens during flower development. On top of that, they play a key role in gynoecium development, as suggested by analysis of the *cuc1 cuc2* double mutant: it indeed fails to form the septum and ovule development is severely impaired (Figure 11) (Ishida et al., 2000).

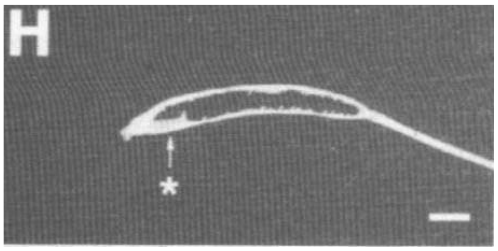


Figure 11. Septum in cuc1 cuc2. Scale bar=1mm (Ishida et al., 2011)

The relationship among *SPT*, *CUC1* and *CUC2* is quite intricated. In the apical region *SPT* promotes carpel fusion partially by suppressing *CUC1* and *CUC2*. In *cuc1 spt* and *cuc2 spt* double mutants, the cleft at the top of the gynoecium, observed in the single *spt* mutant, is largely suppressed and the style is open only at the top (Figure 12, upper row); whereas, within the basal region, the three genes act together to differentiate carpel margin-derived organs. As matter of fact, *cuc1 spt* and

cuc2 spt double mutants showed a longer unfused septum region, compared to *spt* single mutant (Figure 12, bottom row) (Nahar et al. 2012). Besides transcription regulatory factors, hormones play a pivotal role during gynoecium development. Recent reports have indeed showed that the phytohormones auxin and cytokinins are necessary for the correct differentiation of the carpel marginal tissues; their cross-talk may define the different regions of the gynoecium (Reyes-Olalde et al., 2013; Deb et al., 2018). On top of that, transcriptional regulators and hormones can play synergistical activities during gynoecium development; in fact, several transcription factors can connect the distinct phytohormones networks for a correct gynoecium formation (Zúñiga-Mayo et al., 2019). As matter of fact, Reyes-Olalde et al. (2017b) interestingly reported that cytokinin signaling modulates auxin biosynthesis and transports in the marginal tissues and that *SPT* has a key role in the mediation of those signals. Those observations suggest that deciphering the intricate networks between regulatory factors and phytohormones signalling and activity could give a noteworthy improvement on our knowledge on mechanisms controlling gynoecium development.

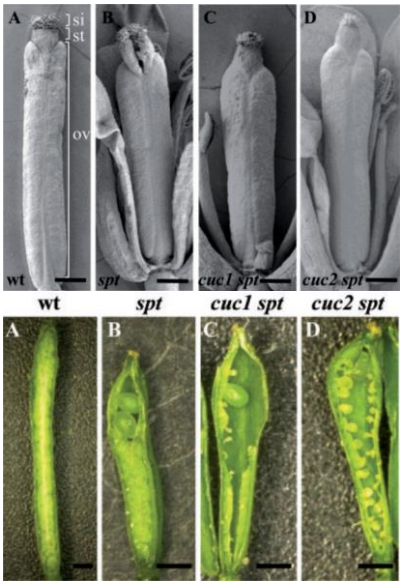


Figure 12. A-D upper row: Scanning electron microscopy of A: wild-type, B: spt, C: cuc1 spt, D: cuc2 spt at stage 14 gynoecia. si: sitgma, st: style, ov: ovary. Scale bars = 350 µm. A-D bottom row: A: wild-type, B: spt, C: cuc1 spt, D: cuc2 spt siliques. In spt (B) the upper part of the septum is unfused, while in cuc1 spt (C) and cuc2 spt (D) the fusion is missing along the whole apical-basal axis of the ovary. Scale bar = 250 mm (Nahar et al., 2012).

3.5 The ovule

Ovules are complex structures that develop from the placenta, inside the carpels; after fertilization they will originate the seeds. From their proximal-distal axis we can distinguish three elements: the funiculus, which attaches the ovule to the placenta, for passage of nutrients; the chalaza, precursor of the integuments; the nucellus, that is protected by the integuments; here the megaspore mother cell will differentiate to form the embryo sac (Colombo et al., 2008). Ovule development has been extensively studied in Arabidopsis and genes that determine the identity of the ovule have been identified. A gene that play a pivotal role in the correct development of ovules is *SEEDSTICK (STK); it* is a MADS-box gene, belonging to the class D along with *SHATTERPROOF1 (SHP1)* and *SHATTERPROOF2 (SHP2).* The *stk shp1 shp2* triple mutant is indeed characterized by defects in the development of ovules and by their conversion into either leaf-like or carpel-like structures (Figure 13) (Favaro et al., 2003; Pinyopich et al., 2003).

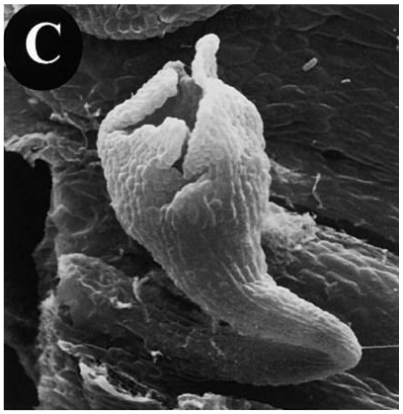


Figure 13. Ovules of stk shp1 shp2 triple mutant plant (Favaro et al., 2003)

3.6 The fruit

The fruit can be defined as the mature ovary that forms a specialized structure to protect the developing seeds; at maturity it dries and dehisces, releasing the seeds. The fruit develops from the gynoecium upon fertilization. In Arabidopsis it consists of two congenitally fused carpels, since the gynoecium arises as a single primordium (Seymour et al., 2013).

3.6.1 Fruit structure

Arabidopsis mature ovary and subsequently the fruit, is composed of four different regions, from top to bottom: stigma, style, ovary and gynophore (Roeder and Yanofsky, 2006) (Figure 14).

- stigma: located at the top of the gynoecium, it consists of a single layer of elongated papillary cells, that interact with the pollen and let it germinate; the stigma is the first member of the transmitting tract, a set of cells whose extracellular matrix is rich of polysaccharides, forming a pathway for the correct growth and guidance of the pollen tubes through the ovary;

- style: situated below the stigma, it is a vascular tissue surrounding the transmitting tract;

- ovary: it is the central part of the fruit and protects the developing seeds; the ovary is composed by several tissues:

    - valves: also called seedpod walls, they derive from the fused carpels of the gynoecium and they are joined to a central septum. They lie on the lateral sides of the ovary and serve to protect the seeds. During dehiscence, the valves separate from the replum and fall out in order to release the seeds;

    - valve margins: they are specialized lateral organs that form at the valve-replum interface. Upon maturation, they allow the dispersion of the seeds, due to the separation of the valves from the replum;

- replum and septum: they divide the ovary in two halves. Ovules and funiculi arise from the placenta, which lies along each of the inner sides of each replum. The septum divides the fruit stretching from one replum to the other. In the centre of the septum, transmitting tract cells form a continuous tract, that starts at the style, for pollen tube growth;

- gynophore: bottom element of the fruit, it supports the ovary; it also called internode.

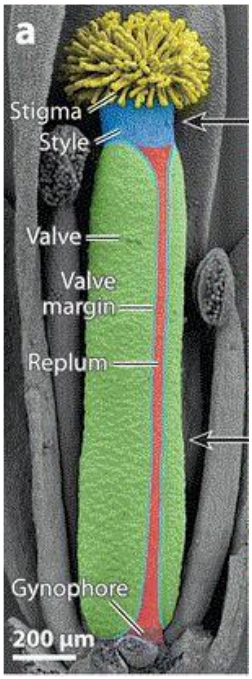


Figure 14. Arabidopsis thaliana gynoecium (Seymour et al., 2013).

3.7 Regulation of homeotic genes

Transcription factors (TFs) are the master regulators of gene expression; they act at multiple levels to orchestrate plant developmental process. TFs can bind specific motifs on the DNA, and they can cooperate with genetic and epigenetic processes to fulfil their regulatory function (Busch et al., 1999; Orphanides et al., 1999). In fact, it has been shown that different class of TFs can interact with different partners, establishing complexes that specifically act on a sub-set of targets, whose regulation is fundamental for a correct development (Smaczniak et al., 2012b; Khan et al., 2018). As mentioned before, floral organ identity is finely controlled by homeotic genes, acting as TFs; their mutation causes indeed floral organ homeosis (Murai, 2013). For their importance in establishing organ identity, the temporal and spatial modulation of homeotic genes expression is precisely regulated throughout development by several mechanisms (Drews et al., 1991; Gustafson-Brown et al., 1994; Goto and Meyerowitz, 1994; Liu and Meyerowitz, 1995; Gregis et al., 2008; Chen et al., 2018). Despite this, the molecular mechanisms for their regulation are yet to be completely uncovered.

3.7.1 *SEEDSTICK*: a model to study the regulation of homeotic genes

SEEDSTICK (STK) is a MADS-box gene, belonging to the class D along with *SHATTERPROOF1 (SHP1)* and *SHATTERPROOF2 (SHP2)*. They act redundantly for the determination of ovule identity (Favaro et al., 2003; Pinyopich et al., 2003).

On top of that, *STK* controls the development of the funiculus, a stalk connecting the ovule to the placenta. The mutant *stk* is characterized by a longer and bigger funiculus and a failure in releasing the small seeds from the mature silique, which appear shorter compared to the wild-type (Pinyopich et al., 2003).

The homeotic gene ***STK*** provides an excellent model to study gene regulation in plants, because, in its expression is precisely restricted to developing ovules and placenta during the first stages of gynoecium development; in mature ovules is detected strongly in the funiculus and weakly in integuments that will later form the seed coat, as shown in Figure 15 (Brambilla et al., 2007; Mizzotti et al., 2014). Previously, two transcriptional regulators of *STK* have been characterized: the MADS-domain factor SHORT VEGETATIVE PHASE (SVP) and the GAGA binding factors BASIC PENTACYSTEINE PROTEINs (BPCs) (Kooiker et al., 2005; Simonini et al., 2012; Gregis et al., 2013).

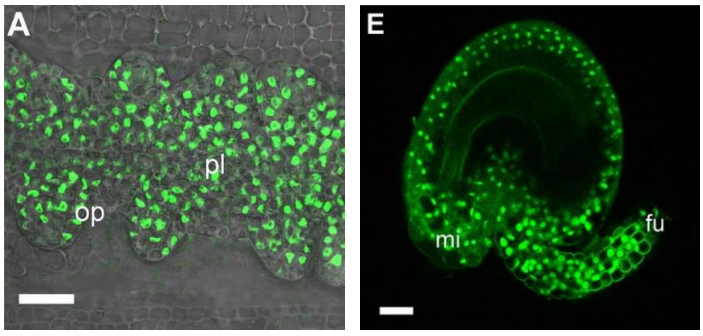


Figure 15. pSTK:STK-GFP expression during ovule development. A: early ovule development, STK is expressed in the placenta and ovule primordia, scale bar = 50 µm. E: mature ovule, after fertilization, STK is detected in the outer integuments and the funiculus, bar = 40µm. op: ovule primordia; pl: placenta; mi: micropyle; fu: funiculus (Mizzotti, 2014).

3.7.2 GAGA BINDING PROTEINs (GBPs)

GAGA binding proteins (GBPs) represent a family of proteins, known to be involved in chromatin structure and dynamics, by interacting with chromatin associated proteins (CAPs) (Berger and Dubreucq, 2012). GBPs can bind DNA sequences that present a $(GA)_n$ repetition. The first GAGA protein was identified in Drosophila and named GAGA factor (dGAF or Trl) (Soeller et al., 1993). It has been reported its role as regulator of homeobox genes (Farkas et al., 1994); a genome-wide mapping to identify in vivo target of dGAF resulted in a list of approximately 250 target genes, indicating its involvement in numerous pathways (Steensel et al., 2002). Furthermore, Berger et al. (2011) interestingly showed that GAGA motifs are present in both transcriptionally active and inactive regions, as well as in introns and exons; also, GAGA sequences were found in the PRE (POLYCOMB REPRESSIVE ELEMENT) sequences, suggesting their role in the recruitment of the Polycomb group (PcG) (Ringrose and Paro, 2007).

In plants, the first GBP was identified in soybean (*Glycine max*), as regulator of the *Gsa1* gene, involved in chlorophyll biosynthesis. Removal of the GAGA element disrupted ability of GBP to bind *Gsa1* promoter, while insertion of the GAGA motif in a non-specific DNA conferred GBP-binding activity on that DNA fragment (Sangwan and O'Brian, 2002). GBPs have been then identified in both monocotyledons and dicotyledons; among them the monocot species *Oryza sativa* (rice) and *Hardeum volgare* (barley) and the dicot species tomato (*Lycopersicon esculetum),* cucumber (*Cucumis sativus*) and Arabidopsis (Sangwan and O'Brian, 2002; Santi et al., 2003; Meister et al., 2004; Wanke et al., 2011; Monfared et al., 2011; Berger et al., 2011; Berger and Dubreucq, 2012; Simonini et al., 2012; Simonini and Kater, 2014; Hecker et al., 2015; Mu et al., 2017a; Mu et al., 2017b; Roscoe et al., 2019; Cong et al., 2018).

3.7.3 BASIC PENTACYSTEINE PROTEINs (BPCs)

In Arabidopsis, a family of plant-specific GBPs have been identified and named BASIC PENTACYSTEINE proteins (BPCs). BPCs were first identified as regulators of *INNER NO OUTER (INO)*, a gene required for the correct formation of the outer integument during ovule development (Meister et al., 2004).

BPCs achieve several fundamental activities during plant growth, as suggested by the severe defects observed in *bpc* multiple mutants (described in detail in section 3.7.3.1). They are part of a small, plant-specific family of only seven members (BPC1, 2, 3, 4, 5, 6, 7). BPCs present a conserved C-terminal region of 98 residues (42 aa are identical in all the BPCs). This region is basic and includes an unusual arrangement of five conserved cysteines (hence the name BPCs); it shows specific DNA-binding activity (Meister et al., 2004). As mentioned above, BPCs are able to bind DNA sequences that present $(GA)_n$ repeats. Binding sites for BPC proteins are called C-boxes and present the following consensus sequence: RGARAGRRA (Kooiker et al., 2005). Based on their protein sequence homology, BPCs have been subdivided into three classes (Meister et al., 2004) , as shown in Figure 16:

- Class I: BPC1, BPC2, BPC3
- Class II: BPC4, BPC5, BPC6
- Class III: BPC7

Due to the presence of an in-frame stop codon in its sequence, BPC5 has been considered a pseudogene, unable to produce a full length-protein; for this reason, it is usually excluded from BPCs analysis.

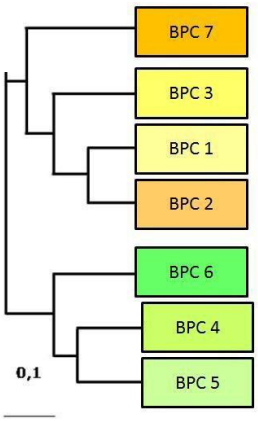


Figure 16. Phylogenetic tree of BPCs.

3.7.3.1 The *bpc* quintuple mutant

As reported above, BASIC PENTACYSTEINE PROTEINS (BPCs) are plant specific GAGA BINDING PROTEINS (GBPs) of *Arabidopsis thaliana*. They are ubiquitously expressed in the plant throughout

its development (Meister et al., 2004; Monfared et al., 2011); GUS activity assays showed that their expression patterns overlap in some tissue but are distinct in others (Monfared et al., 2011). Mutant alleles in all BPCs have been identified and characterized; all the mutations are caused by T-DNA insertions, with the only exception of *bpc3* mutant: in this case, a point mutation causes the formation of an early stop codon. Mutant alleles *bpc1-2, bpc2*, *bpc4, bpc5* and *bpc6* fail to produce full-length transcripts and thus, they can be considered knock-out mutants whereas the *bpc1-1* is classified as a knock-down mutant. BPCs have redundant activity; in fact, single mutants do not show any phenotype. The *bpc4 bpc6* double mutant (mutant in class II BPCs) present a wild-type like phenotype, with no drastic effects on plant development and fitness. In contrast, mutations in class I BPCs led to a pleiotropic phenotype; the *bpc1-2 bpc2 bpc3* triple mutant (Figure 17, centre) is indeed characterized by defects in the development of reproductive organs, inflorescence and flower structure, inducing a partial rather significant sterility, as analysed by Simonini and Kater (2014) Higher order mutant combinations between class I and class II alleles display more severe defects. The *bpc1-1 bpc2 bpc3-1 bpc4 bpc6* quintuple mutant (Figure 17, right) shows severe defects in both vegetative and reproductive development (i.e. defects in ovule and seeds development, impaired secondary roots formation and reduced ethylene response) (Monfared et al., 2011).

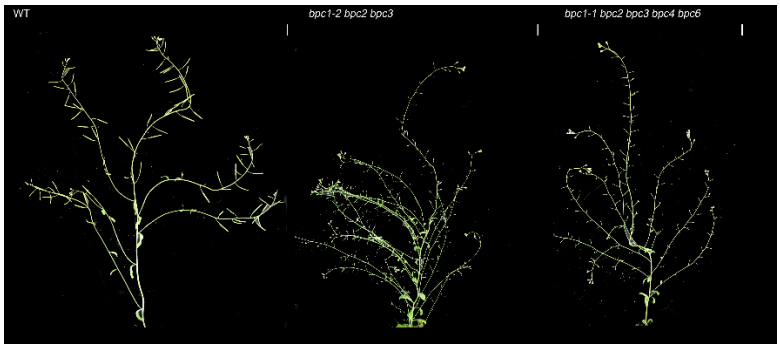


Figure 17. Phenotype of wild-type, bpc1-2 bpc2 bpc3 and *bpc1-1 bpc2 bpc3 bpc4 bpc6 (Monfared et al., 2011; Simonini and Kater, 2014).*

In contrast, mutation of *BPC7* do not show any phenotypic effects, even in combination with other BPCs mutant alleles, suggesting that the sole member of class III BPC proteins might not play an essential role during plant development (Monfared et al., 2011).

Given the broad expression of BPC proteins throughout the plant and the pleiotropic defects that arise in high order BPC mutants, BPCs are likely to act in combination with a wide variety of process-specific regulatory factors, to activate or repress gene expression. This idea is supported

by the evidence that BPCs are able to induce conformational changes in the DNA, presumably by mechanisms that involve the formation of multimeric protein complexes (Kooiker et al. 2005; Simonini et al., 2012; Theune et al., 2019).

Recently, new roles for BPCs have been explored. They can recruit Polycomb repressive complexes (PRCs) to GAGA-motifs (as described in detail in section 3.8.2) (Hecker et al., 2015; Xiao et al., 2017); in particular, it has been shown that both BPCs of class I and class II could interact or colocalize with hallmark proteins of the repressor complexes (Hecker et al., 2015; Mu et al., 2017; Xiao et al., 2017; Wu et al., 2019) suggesting an active role for BPCs in PRC-mediated silencing of a subset of targets. Therefore, studying the molecular mechanisms in which BPCs and PRCs act could deepen our knowledge on gene regulation in plants.

Furthermore, an intriguingly correlation with hormone-mediate signalling have been determined. Many genes of the brassinosteroid signalling pathway are targeted by the member of BPCs of class II, BPC6 (Shanks et al., 2018; Theune et al., 2019); as matter of fact, the *bpc4 bpc6* double mutant shows brassinosteroid dependent phenotypes during root growth (Theune et al., 2019). Also, mutations in BPCs of class I and II affects cytokinin perception in roots (Mu et al., 2017); thus, Shanks et al. (2018) reported a role for BPCs in cytokinin response pathway, already suggested by Simonini and Kater (2014) findings in inflorescence. Similarly, also ethylene response is perturbed in *bpc* multiple mutants throughout plant growth (Monfared et al., 2011).

Collectively, the data above stress that furthering the knowledge on the links between BPCs and phytohormones could help decoding their role in plant development.

3.7.4 The role of BPCs of class I and AP1-SVP-SEU-LUG repressor complex in the regulation of *STK* during flower development

To identify the regulatory elements that control the ovule- and placenta-specific expression of ***STK***, Kooiker et al. (2005) cloned and fused a 2.8-kb sequence around the ATG of the homeotic gene to the β-glucoronidase (GUS) reporter gene. This fragment contains a region of 1.4 kb upstream of the transcription start site and a region of 1.4 kb containing the 5′UTR and the first intron of 1.3 kb, suggesting the importance of the first intron for the correct spatial and temporal expression of *STK*.

Also, they performed a one-hybrid assay in order to identify putative regulators of the homeotic gene. BPC1 resulted to be associated with several regions with $(GA)_n$ repetitions on the promoter of *STK*: seven consensus regions for BPCs, named C-boxes (RGARAGRRA) have been identified, along with five non consensus sequences with no affinity for BPCs binding due to a mismatch (hence called NC-boxes).

BPCs ability to bind C-boxes on the promoter region of *STK* has been tested and confirmed by EMSA gel shift assay *in vitro* (Kooiker et al., 2005) and Chromatin Immunoprecipitation assay (ChIP) *in vivo* (Simonini et al., 2012). As reported above, class I and II BPCs act redundantly with each other, as suggested by the analysis of the higher order *bpc* mutants (Monfared et al., 2011); as matter of fact, in the *bpc 1-1* mutant no deregulation of *STK* expression was detected, since the gene was still confined to ovule and placenta, even though a slightly increase of GUS signal has been registered (Kooiker et al., 2005).

Simonini et al. (2012) further investigated the role of C-boxes in the regulation of *STK*; they confirmed that those regions are important for the correct spatial and temporal expression of the homeotic gene, as depicted in Figure 18; they indeed showed that a complete disruption of *STK* specific pattern of expression occurred when all the C-boxes were mutated *(STKpro_GAm7*:GUS); the activity of *STK* promoter was indeed no longer confined to ovule and placenta but registered also in other floral organs as well as in the first phases of flower initiation and development.

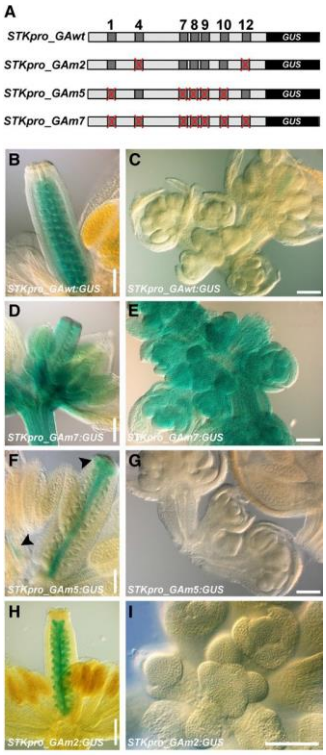


Figure 18. (A) Schematic representation of the STK promoter-GUS constructs used in this study: dark boxes represent C-boxes, and the mutated ones are indicated with a red cross. (B) to (I) GUS staining performed on inflorescences (at left, mature flowers; at right, inflorescence and floral meristems and young flowers) of STKpro_GAwt:GUS [(B) and (C)], STKpro_GAm7:GUS [(D) and (E)], STKpro_GAm5:GUS [(F) and (G)], and STKpro_GAm2:GUS [(H) and (I)] plants. Bars in (B), (D), (F), and (H) = 200 μm; bars in (C), (E), (G), and (I) = 100 μm.(Simonini et al., 2012).

The Tethered Particle Motion (TPM) analysis revealed that BPCs of class I induce DNA conformational changes in the *STK* promoter, by binding the DNA at multiple sites (Kooiker et al., 2005). To test the hypothesis whether BPCs of class I could form complexes, interacting with each other, Simonini et al. (2012) performed a yeast two hybrid assay, confirmed by Bimolecular Fluorescence assay (BiFC); the results that BPC1 interacts with BPC2 and BPC3, while BPC2 binds to BPC3 and forms homodimers corroborates the hypothesis that BPCs can bend the DNA, inducing loops at the promoter of *STK* by binding DNA at different C-boxes and forming homo- and heterodimers, as proposed in Figure 19.

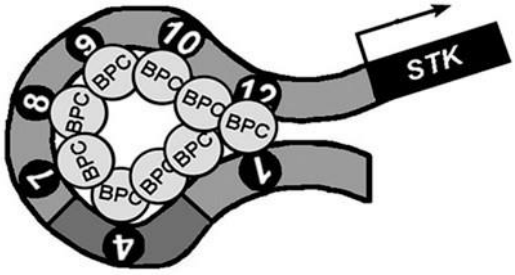


Figure 19. Proposed model for BPCs activity on STK promoter (Simonini et al., 2012).

Another important regulator of *STK* expression during flower development is the MADS-domain factor SHORT VEGETATIVE PHASE (SVP). SVP directly binds *STK* regulatory region at different sites (Gregis et al., 2013); in particular, ChIP-sequencing assay showed that the highest enrichment of SVP binding was in a region containing CArG-boxes, surrounded by C-boxes, named Region B and illustrated in Figure 20 (Simonini, 2012; Gregis et al., 2013).

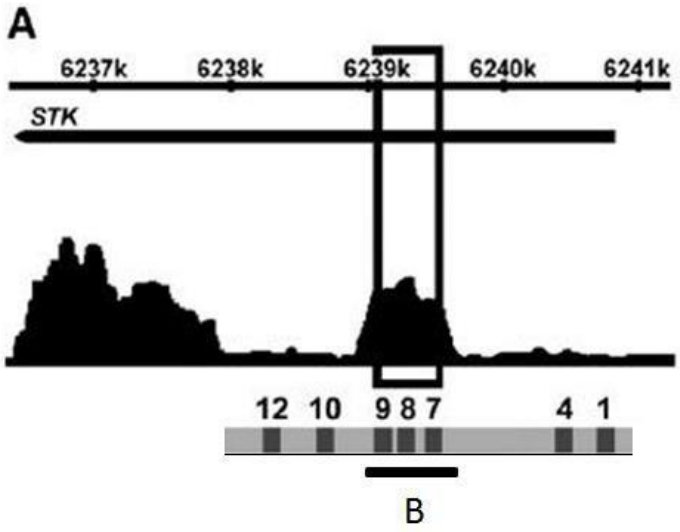


Figure 20. Schematic representation of SVP binding to the STK locus as identified by ChIP-sequencing. The fragment (Region B) present 3 C-boxes, surrounded by CArG-boxes (not shown) (Simonini et al., 2013; Gregis et al., 2013).

SVP has been first identified and characterized as flowering time regulator; in particular, it is a repressor of floral transition (Hartmann et al., 2000). Later, it acts redundantly with other two MADS-domain factors, APETALA1 (AP1) and AGAMOUS-LIKE24 (AGL24) in establishing floral

meristem identity in the early stages of flower development (Gregis et al., 2006; Gregis et al., 2008; Liu et al., 2009; Gregis et al., 2009).

In-situ hybridisation analysis, performed to further investigate the role of SVP in the regulation of *STK*, confirmed its redundant activity with AP1 and AGL24; in fact, in the *agl24 svp ap1-12* triple mutant *STK* expression was detected in the ovules and placental tissue as wild-type flowers, but also in floral meristems and young flowers (Figure 21) (Simonini et al., 2012).

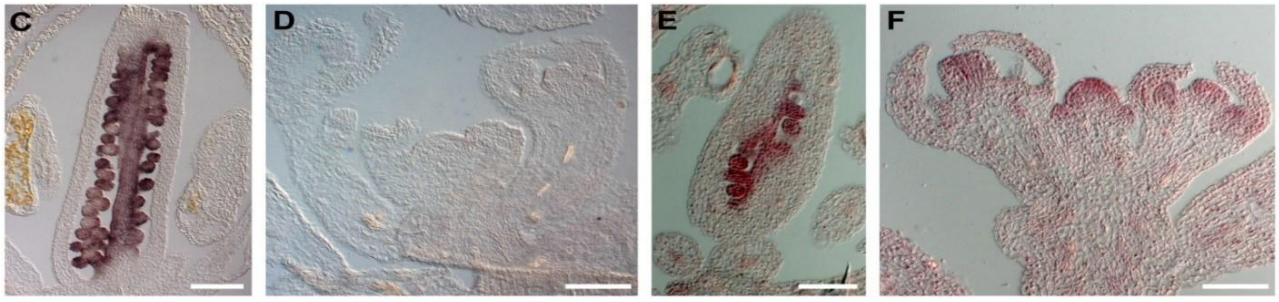


Figure 21. In situ hybridization with a STK-specific antisense probe in wild-type (C and D) and agl24 svp ap1-12 triple mutant (E and F) inflorescences. Scale bars: C and E = 200 µm, D and B and F = 100 µm (Simonini et al., 2012).

Other two factors that might play an important role in the regulation of *STK* expression are the corepressors *SEUSS (SEU)* and *LEUNIG (LUG)*. They have been identified as negative regulators of the class C gene *AGAMOUS (AG)*, whose confined expression is fundamental for a correct flower development (Bowman et al., 1991; Liu and Meyerowitz, 1995; Immink et al., 2010). As reported in Gregis et al. (2006), the triple mutant *svp agl24 ap1*,severely affected in flower development, is characterized by all four floral whorls conversions caused by the ectopic expression of class B and C genes. A similar phenotype has been observed in *seu* and *lug* single and double mutant background, suggesting a putative interaction among those factors. It has indeed been proved that AP1-AGL24 and AP1-SVP dimers interact with the SEU-LUG corepressor to repress AG expression in the flower (Gregis et al., 2006).

The observation that CArG-boxes (MADS-box consensus sites) and C-boxes (BPCs consensus sites) are very closely located in the regulatory region of *STK*, led to the hypothesis that BPCs might interact with the AP1-SVP-SEU-LUG repressor complex to repress *STK* expression during the first stages of flower development.

Physical interaction between AP1-SVP-SEU-LUG and BPCs of class I have been shown by yeast two-hybrid assay (Figure 22) and confirmed by BiFC; in particular, this analysis revealed that BPC1

interacts with SEU, AP1 and SVP, BPC3 interacts only with SVP, whereas BPC2 does not interact with any protein of the complex (Simonini et al., 2012).

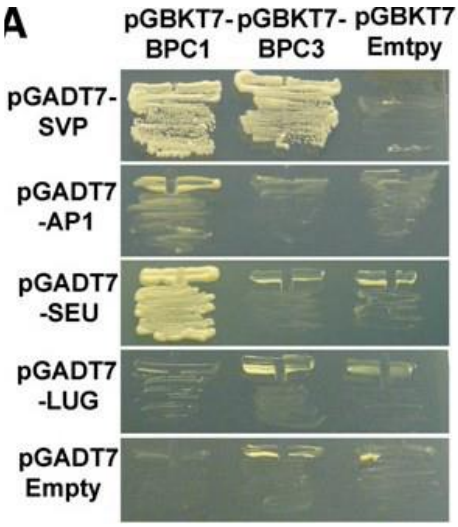


Figure 22. Yeast two-hybrid assays testing interactions between AP1-AD, SVP-AD, SEU-AD, LUG-AD, and BPC1-BD, BPC3-BD on −W−L−H +5 mM of 3-AT selective media (Simonini et al., 2012)

The hypothesis that BPCs might stabilize or facilitate the binding of SVP to the DNA was further analysed by ChIP. As shown in figure 23, the mutation of the 7 C-boxes on the promoter of *STK* (*STKpro_GAm7*:GUS) affected SVP binding to the DNA; surprisingly, when 5 out of the 7 C-boxes (*STKpro_GAm5*:GUS) were mutated SVP was still able to bind the promoter of the homeotic gene, confirming that C-boxes are important to facilitate the binding of the SVP repressor to the CArG boxes in the *STK* promoter.

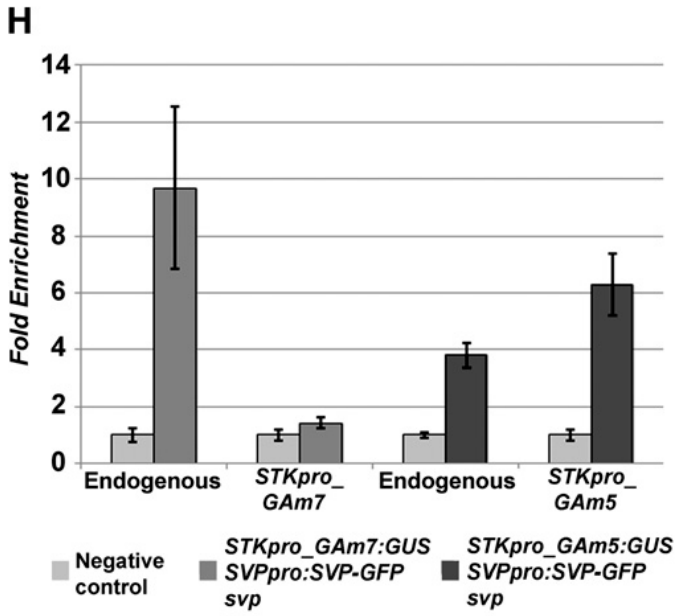


Figure 23. Quantitative Real-Time PCR analyses of the ChIP assays to analyze SVP binding to endogenous wild-type and mutated versions of fragment B of STKpro_GAm7:GUS and STKpro_GAm5:GUS promoters.

3.8 Epigenetic regulation of homeotic genes

Recently it has been reported a novel role of BPCs and MADS-domain factors in the regulation of their target expression, by recruitment of histone-modifying complexes, like the Polycomb group (PcG) proteins (Liu et al., 2009; Hecker et al., 2015; Mu et al., 2017b; Xiao, 2017; Roscoe et al., 2019). PcG proteins repress gene expression and regulate plant development by epigenetically modifying chromatin. As a matter of fact, the dynamic change of chromatin structure must be strictly regulated; the chromatin structure offers an opportunity for gene regulation since DNA elements such as promoters, enhancers and transcriptional start sites have to be exposed in order to be transcriptionally functional. The packing state of chromatin is actively interchanged by chromatin remodelling complexes and deposition of histone modifications (Clapier and Cairns, 2009).

3.8.1 The Polycomb Group (PcG)

Polycomb group (PcG) components have been first identified in Drosophila as regulators of homeobox genes, whose expression is pivotal to determine a proper body segmentation pattern (Lewis, 1978). PcG proteins participate to cell fate determination by maintaining homeotic genes

in a silent state; as expected, mutations in PcG genes commonly result in homeotic mutations. Homologues of Drosophila PcG proteins have been discovered in vertebrates and plants, suggesting a similar role in fate determination (Goodrich, 1997). PcG proteins form multi-protein complexes, classified in two groups: POLYCOMB REPRESSIVE COMPLEX1 (PRC1) and POLYCOMB REPRESSIVE COMPLEX2 (PRC2). Both complexes are necessary for gene silencing; they act in a sequential manner to repress the expression of target genes, by inducing covalent histone modifications. PRC2 introduces a trimethylation on the lysine 27 residue on histone H3 (H3K27me3) whereas PRC1 can recognize the histone repressor marker and monoubiquitinates the lysine 118 residue on histone H2A (H2AK118ub).

In plants several developmental pathways are controlled by PcG proteins activity, such as seed development, flowering time, vernalization and organ identity (Makarevich et al., 2006; Pien and Grossniklaus, 2007; Kim and Sung, 2014)

POLYCOMB REPRESSIVE COMPLEX2 (PRC2):

In Drosophila, the PRC2 is composed of four core components. Even though each subunit has a distinct role, they finely cooperate to mediate histone methyltransferase activity on H3K27. Also, PRC2 members reinforce self-propagation of the H3K27me3 repressive mark. In Arabidopsis, three different PRC2 complexes have been identified: the FERTILIZATION INDEPENDENT SEED (FIS) complex that controls seed development; the FERTILIZATION INDIPENDENT ENDOSPERM (FIE), for the endosperm formation; the VERNALIZATION (VRN) complex, which mediates the vernalization response; the EMBRYONIC FLOWERING (EMF) complex, involved in flower development. The three PRC2 share some common components, but each of them is characterized by distinct proteins, specific for the different complexes: CURLY LEAF (CLF), SWINGER (SWN), MEDEA (MEA), VERNALIZATION 2 (VRN2) and EMBRYONIC FLOWER 2 (EMF2) (Schatlowski et al., 2008; Kim and Sung, 2014; Oliva et al., 2016).

As mentioned above, many developmental processes are controlled by PcG proteins; as matter of fact, mutations in PRC2 components result in a de-repression of the expression of developmental genes. As reported by Goodrich (1997), *CLF* is required for a negative regulation of the homeotic gene *AG* in leaves, inflorescence stems and flowers; thus, the *clf* mutant has a phenotype that highly resemble the one reported for transgenic plants in which *AG* is constitutively expressed. H3K27me3 dependence on PRC2 complexes was confirmed by analysis on several PRC2 mutants;

ChIP assays showed a loss of H3K27me3 deposition in *clf swn* mutants and strongly reduced in *vrn2 emf2-10* mutants (Lafos et al., 2011).

Interestingly, whole genome analysis of H3K27me3 distribution has revealed that around 4400 Arabidopsis genes are marked with the repression mark. Protein-coding genes made up most of the targets, while transposable elements and heterochromatic regions are mostly excluded. In particular, many developmentally important transcription factors show a H3K27me3 coverage, suggesting the importance of epigenetic processes to modulate their expression. In Arabidopsis, H3K27me3 regions are considerably shorter than those in Drosophila, with the repression mark confined to single genes. Analysis of H3K27me3 pattern in several plant tissues during different developmental stages revealed a time and tissue-specific distribution of the repressive mark. Thus, H3K27me3 deposition is a dynamic process that respond to developmental and environmental cues. The presence of H3K27me3 correlated with repression, as non-expressed genes are preferential targets of the mark (Zhang et al., 2007; Lafos et al., 2011).

POLYCOMB REPRESSIVE COMPLEX1 (PRC1):

In Drosophila, the PRC1 is divided in two groups of proteins, based on their biochemical function: one group, known as writer group, possesses the catalytic activity to mono-ubiquitinate histone H2A; the other group, known as reader group, recognizes certain modified histone marks. The proteins of the reader group can bind to the H3K27me3 mark, catalysed by the PRC2, with their N-terminal chromo-domain. This interaction physically links a functional crosstalk between PRC2 and PRC1 (Kim and Sung, 2014).

Several homologs of PRC1 have been found and characterized in Arabidopsis; despite this, their functions and the regulatory mechanisms by which they act are yet poorly defined (Merini and Calonje, 2015; Merini et al., 2017; Yang et al., 2017). Among them, the chromo-domain protein LIKE-HETEROCHROMATIN PROTEIN1 (LHP1)/TERMINAL FLOWER2 (TFL2) (henceforth called LHP1) displays a high binding affinity for H3K27me3, as it associates with numerous H3K27me3-marked regions (Turck et al., 2007). Interestingly, Exner et al. (2009) showed that the chromo-domain of LHP1 is fundamental for H3K27me3 binding, since its mutation affects its binding affinity to the repressor mark and reactivate target genes expression, concluding that LHP1 is essential to maintain H3K27me3-mediated repression.

Chromosome-wide analysis of LHP1-associated regions by ChIP-chip experiments revealed that LHP1 associates with defined short domains along the genome. Similarly to H3K27me3 region's length, LHP1 domains are restricted within genes and their flanking regions. Analysis of the target genes demonstrates that LHP1 is involved in the regulation of a wide range of biological processes; thus, *lhp1* mutants show overlapping phenotypes with several PcG mutants, even though the phenotypes are less severe, suggesting that other factors must mediate PcG repression (Turck et al., 2007; Zhang et al., 2007; Exner et al., 2009; Kim and Sung, 2014). Indication of a physical interaction between PRC2 and PRC1 confirms the tight interplay of their activities. *In vivo* and *in vitro* experiment showed that LHP1 can interact with EMF and MSI1, two components of the PRC2 complex. LHP1 is associated to CLF, another member of the PRC2 (Derkacheva et al., 2013; Wang et al., 2016); also, LHP1 can interact with several PRC1 members (Xu and Shen, 2008). These observations suggest that LHP1 might act as a "bridge" for an efficient crosstalk among PRC1 and PRC2 members.

3.8.2 Role of PRCs in the regulation of BPCs and SVP targets

BPCs have been recently associated to the epigenetic gene expression regulation, operated by the PcG proteins. Specific genomic sequences recruit the PRC2 complex on target genes; these *cis*-regulatory regions, called Polycomb response elements (PREs), are enriched in GA repeats and class I BPCs physically interact and recruit the PRC2 complex on these PREs (Xiao et al., 2017). In seedlings, BPC proteins bind to the PRC2 member *SWINGER* and repress the expression of *ABSCISIC ACID INSENSITIVE 4* (*ABI4*), by directly recruiting the Polycomb-Repressive complex2 on the *ABI* promoter (Mu et al., 2017b). BPCs of class II member BPC6 directly interact with the PRC1 component LHP1 for its recruitment to GAGA-motif elements; together they seem to form a scaffold for the sequential attachment of PRC2 (Hecker et al., 2015); also, BPCs of class I have been proved to form complex with FIE and CLF (Mu et al., 2017; Xiao et al., 2017).

Interestingly, SVP can interact and recruit LHP1 on *SEP3* locus to mediate H3K27me3 deposition in order to maintain the homeotic gene in a silent state (Liu et al., 2009). All these findings suggest a possible role of BPCs, SVP and LHP1 in the regulation of gene expression, probably by influencing the H3K27me3 mark deposition at the target locus.

4. Results

4.1 Studying the role of BPCs, SVP and LHP1 in the regulation of *STK* expression during flower development

The first aim of my PhD project is to study the transcriptional and epigenetic regulation of *STK* during flower development in *Arabidopsis thaliana,* unravelling the role of BPCs, MADS-domain factor SVP and PRC1 member LHP1.

4.1.1 *bpc1-2 bpc2 bpc3 bpc4 bpc6* shows ectopic expression of *STK* during flower development

Simonini et al. (2012) reported that *bpc1-2 bpc2 bpc3* mutant does not show any deregulation of *STKpro_GAwt:GUS* construct, suggesting that BPCs of class II might play an important role for the homeotic gene regulation in the flower. To better understand the contribution of BPCs of class I and class II in the regulation of *STK* expression during flower development we generated the *bpc1-2 bpc2 bpc3 bpc4 bpc6* quintuple mutant (henceforth called *bpcV*) in which the complete knockout allele of BPC1 (Monfared et al., 2011; Simonini and Kater, 2014) is combined with the other mutants' alleles for both class I and class II BPCs. Monfared et al. (2011) already described different combinations of BPCs mutants but all of them contain the *bpc1-1* allele which still retain a part of its transcript.

In wild-type plants, *STK* expression is confined in the ovules and the placenta and never registered in flowers before stage 8 either in inflorescence and floral meristems (Figure 1A and B). The mutation of all the BPCs of class I (Figure 1C and D) or class II (Figure 1E and F) does not affect *STK* expression in the flower. In contrast, in *bpcV* mutant the expression of *STK* has been observed not only in ovules and placenta, but also in developing petals (Figure 1H), floral meristems and young flowers, especially in the tissues corresponding to the organ primordia (Figure 1G). Also, the expression of the histone H4 gene is maintained in *bpcV* mutant as in the wild-type (Fobert et al., 1994) confirming the integrity of the tissue; likewise, no signal was registered using the *STK*-specific sense probe. These controls confirmed that the observed *STK* deregulation in the flower was indeed dependent on BPCs quintuple mutant combination (Figure S1, Manuscript #1).

Our results address for the first time an important role to BPCs of class II for the regulation of *STK* during flower development, suggesting that the two classes of transcription factors (class I and class II BPCs) might modulate the expression of their target redundantly. Examples in literature about their redundant activity have been already provided by Mu et al. (2017b) in which authors showed that mutation of BPCs of class I and II increased *ABI4* expression in roots. Simonini et al. (2012) performed several protein interaction studies, showing that BPCs of class I can interact with each other; also, BPCs of class II can form homo- and heterodimers and they can interact with members of class I (Wanke et al., 2011). Placed in this scenario, our results suggest that BPCs protein-protein interactions could be essential for the correct expression of *STK* and BPCs of class I and class II can act synergistically and redundantly to regulate the expression of their targets throughout plant development.

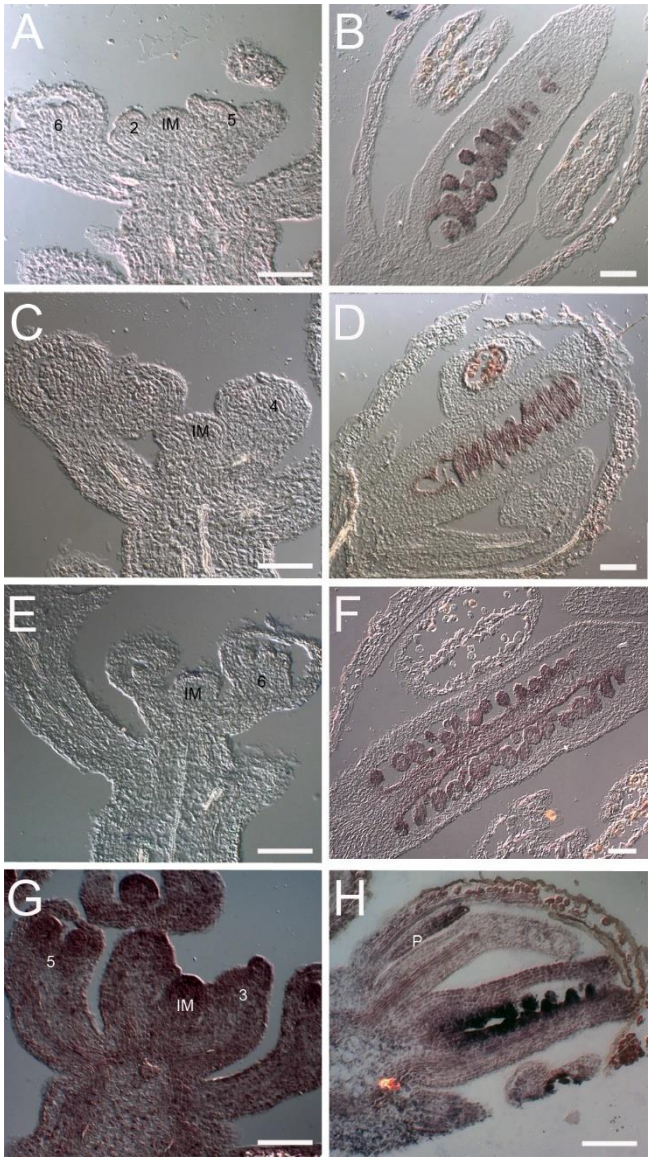


Figure 1. In-situ hybridisation on wild-type [(A) and (B)], bpc1-2 bpc2 bpc3 [(C) and (D)], bpc4 bpc6 [(E) and (F)] and bpcV [(G) and (H)] inflorescences using a STK-specific antisense probe (Brambilla et al., 2007). IM: inflorescence meristem; p: petal; numbers represent flower stages. Scale bars=50 µm.

4.1.2 Overexpression of *STK* affects plant development

As previously reported, BPCs of class I and II are ubiquitously expressed in the plant and their mutation affects plant development (Monfared et al., 2011). Our *bpcV* mutant, indeed, is smaller compared to the wild-type (Figure 2A, left and centre, respectively) and shows several vegetative (i.e. smaller rosette and cauline leaves, Figure 2B and 2C, respectively) and reproductive defects. In particular, the mutations of BPCs of class I and II causes a drastic phenotype in the silique. In the wild-type, from 3 to 12 days after pollination (dap), the silique elongates to reach its maximum

length, as showed in Figure 2D, top; in contrast, in *bpcV* plants (Figure 2D, centre) we observed no elongation and a reduced silique length.

On the other hand, the MADS-box gene *STK* is necessary for a correct ovule and seed development (Favaro et al., 2003; Pinyopich et al., 2003; Mizzotti et al., 2014; Ezquer et al., 2016). Recently, Herrera-Ubaldo et al. (2019) exploited a new role for the homeotic gene, along with *NO TRANSMITTING TRACT (NTT)*, in gynoecium development. Thus, deregulation of *STK* observed in our *bpcV* mutant might cause some effects throughout development. To further investigate this hypothesis, we analysed a line that constitutively express *STK* (Favaro et al., 2003). Interestingly, also this plant appeared shorter compared to the wild-type (Figure 2A, right); it is affected in both vegetative and reproductive development, showing several defects that phenocopies our *bpcV* mutant (Figure 2B, C and D).

Pinyopich et al. (2003) first analysed *stk* mutant, showing smaller seeds compared to the wild-type; also, *STK* is involved in the regulation of the flavonoid pathway (Mizzotti et al., 2014), who plays a pivotal role in the determination of seed size in Arabidopsis (Doughty et al., 2014); as matter of fact, several genes involved in flavonoid biosynthesis and transport were differentially expressed in the mutant *stk*, compared to the wild-type. To analyse the effect of the constitutive expression of *STK* on seed size, we analysed the seed area size in *35S:STK* and *bpcV* mutant; several controls have been adopted: first, we confirmed the reduced area size of *stk* mutant (Pinyopich et al., 2003); also, we used *arf2-8* mutant, since it shows bigger seeds compared to the wild-type (Schruff et al., 2006). interestingly, we registered a bigger seed area size in both *bpcV* and *35S:STK* line (Figure 2E), even though the effect on seed size is more severe in *bpc* quintuple mutant than in *35:STK* lines, suggesting an additional role for BPCs in seed development.

Collectively, these results confirm that the constitutive expression of *STK* affects both vegetative and reproductive phases of plant development. Also, the similar phenotype we observed in *bpcV* and *35S:STK* line (i.e. bigger seeds and defects in silique development) suggest that BPCs might regulate *STK* during fruit and seed development. It would be interested to check whether the mutation of BPCs of class I and class II affected the expression of the homeotic gene in other reproductive tissues, besides the flowers.

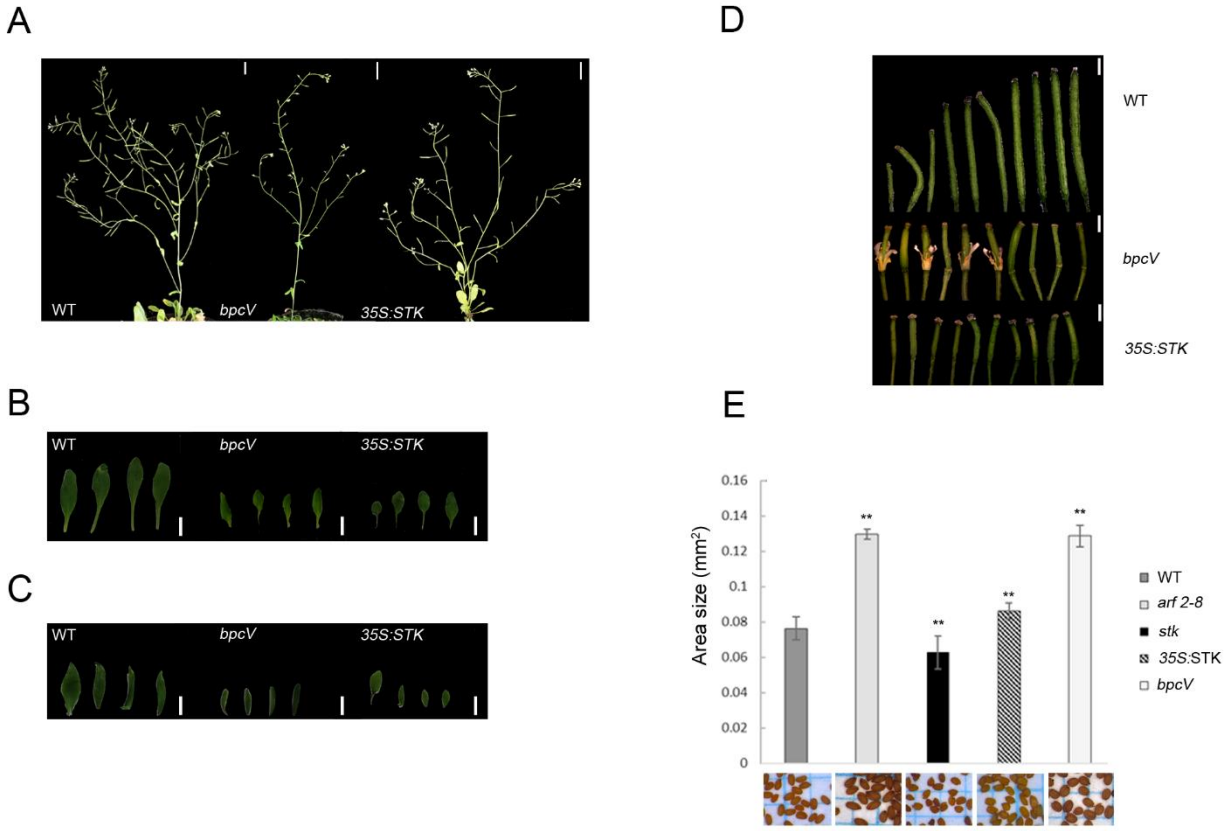


*Figure 2. Overexpression of STK affects vegetative and reproductive development. (A) From left to right: wild-type, bpcV and 35S:STK plants; plants were photographed six weeks after sowing; scale bars=1 cm. (B) Rosette leaves morphology and length: rosette leaves in wild-type, bpcV and 35S:STK (from left to right); (C) Cauline morphology and length: cauline leaves in wild-type, bpcV and 35S:STK (from left to right); scale bars= 0.5 cm. (D) Fruit morphology and length in wild type, bpcV and 35S:STK (from top to bottom); scale bars=1.5 mm.(C) (D) Average seeds area size of wild-type, arf2-8 (Schruff et al., 2006), stk (Pinyopich et al., 2003), 35S:STK and bpcV; error bars represent the standard error mean of replicates; ANOVA and post-hoc Tukey HSD ((honestly significant difference) test were used, **P < 0.01 for WT vs other genotypes comparison.In the lower row, seeds of the analysed genotypes are shown.*

4.1.3 BPCs of class II interact with SVP in vivo

The ability of class I BPCs to form homo- and heterodimers has already been proved; also, BPC1 and BPC3 can interact with SVP and AP1, forming a regulatory complex that has been proposed to repress *STK* expression in reproductive meristems (Simonini et al., 2012).

The newly presented data corroborate the hypothesis that also BPCs of class II have an important role in regulating *STK* expression by forming complexes and stabilising the binding of SVP to its promoter region. We showed that also the two BPCs of class II (BPC4 and BPC6) interact with the MADS-domain factor SVP *in vivo* by yeast two-hybrid, Co-immuno precipitation (CoIP) and

Bimolecular Fluorescence Complementation (BiFC) assays (Figure 3A, B and C, respectively). The last result was unexpected, since their interaction was indeed mostly registered outside the nuclei (Figure 3C). The necessity for dimerization to get nuclear translocation has been already reported for MADS-domain factors in Arabidopsis (McGonigle et al., 1996) and petunia (Immink et al., 2002). We tested the hypothesis of whether the interaction of SVP with the other MADS-domain factor AP1 was necessary for SVP-BPC4 and SVP-BPC6 dimers to enter the nuclei by co-transformed tobacco (*Nicotiana benthamiana)* leaves with an AP1-RFP construct in combination with our protein of interests, cloned in BiFC vectors. This experiment clearly showed that the presence of AP1 is sufficient for the correct translocation of the BPC4- and BPC6-SVP dimers to the nucleus (Figure 3D). Our results confirmed the participation of AP1 in the repressor complex SVP and BPCs are part of, for the regulation of *STK*; this hypothesis has been already suggested by two observations: first, AP1 acts redundantly with SVP (and AGL24) to repress *STK* expression; second, it can interact with BPC1 (Simonini et al., 2012). Interestingly, we registered that, likewise the redundant factor AP1, also the MADS-domain factor AGL24 trigger the translocation of BPC4- and BPC6-SVP dimers to the nucleus (data not shown), corroborating the activity of MADS-domain factors in protein complexes.

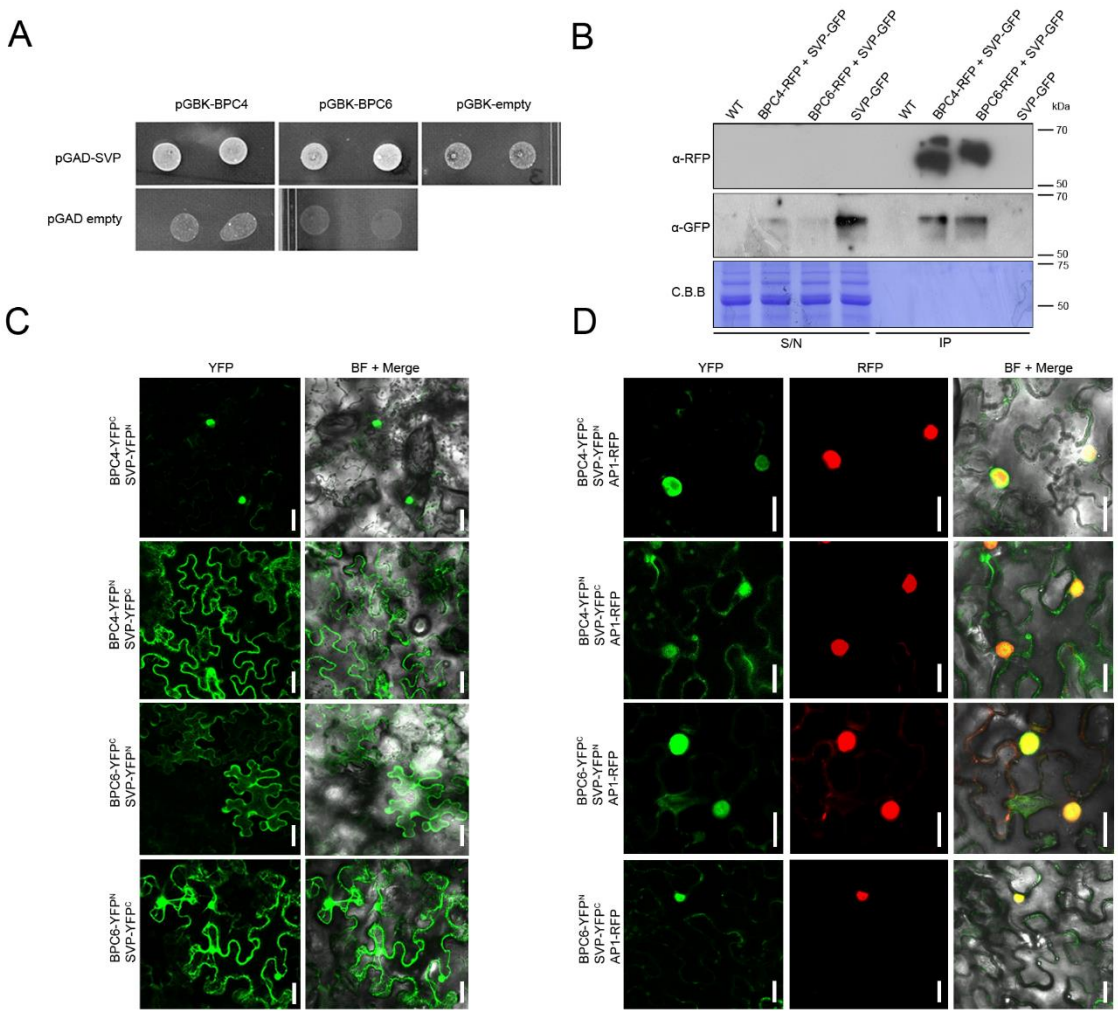


Figure 3. Class II BPCs interact with SVP in vivo. (A) Yeast two-hybrid assay for SVP and BPCs of class II: positive interactions on selective media –W-L-H +5mM 3-AT. (B) Co-immunoprecipitation assays. Nicotiana benthamiana leaves were infiltrated with constructs carrying SVP-GFP together with BPC4-RFP and BPC6-RFP, as described in materials and methods. Immunoprecipitation step was performed using RFP-trap on total protein leaf extract. Samples were probed with GFP and RFP antibodies. S/N: supernatant; IP: immunoprecipitation. (C) Bi-molecular fluorescence complementation (BiFC) assay. Nicotiana benthamiana epidermis cells were transiently transformed with the indicated YN and YC fusions. In the first and the second column, yellow fluorescence and the merging in the bright field were shown, respectively. (D) Bi-molecular fluorescence complementation (BiFC) assay. N.benthamiana epidermis cells were transiently transformed with the indicated YN and YC fusions and AP1-RFP construct. In the first, the second and the third column yellow fluorescence, red fluorescence and the merging between the two channels in the bright field were shown, respectively. Scale bars=40 μm.

4.1.4 Molecular mechanism of SVP-BPCs binding to the regulatory region of *STK*

To gain more insights into the molecular mechanisms by which SVP and BPCs complex act we performed several ChIP assays in distinct genetic background.

First, we registered no binding of BPCs to the C-boxes on the promoter of *STK* (illustrated in Figure 4A) in the *svp agl24 ap1-12* triple mutant in the region containing the C-box 12 (figure 4B). These

results demonstrate an important role of SVP in determining the binding of BPCs of class I to the promoter of *STK.*

Second, we showed that BPCs of class I are not necessary for SVP binding to the DNA. We registered indeed an enrichment when binding of SVP was tested in *pSVP*:SVP-GFP *pSTK_GAm5*:GUS *bpc1-2 bpc2 bpc3 svp* (Figure 4C). Also, this result suggests that C-box 4 and 12 play an important role in determining the binding to the DNA of the MADS-domain factor SVP; as mentioned before (Introduction, section 3.7.4), the *pSTK-Gam5:*GUS presented only 5 GAGA boxes mutated and almost no effect on reporter gene regulation; also, the C-boxes 4 and 12 (upstream and downstream to SVP binding region, respectively) were not altered.

Last, we registered enrichment when binding to the consensus regions for SVP was tested in the *bpcV* mutant background (Figure 4D). These results show that BPCs of class I and class II are not necessary for SVP binding to the DNA regulatory region of *STK*.

Our data corroborate the hypothesis that BPCs are most likely co-factors that stabilize the binding of the MADS-domain repressor complex to the DNA, probably by inducing changes into the *STK* promoter region, as showed by the previous reported in vitro TPM analysis (Kooiker et al., 2005); also, they support the idea that SVP, might act as a pioneer factor to modulate the expression of its targets.

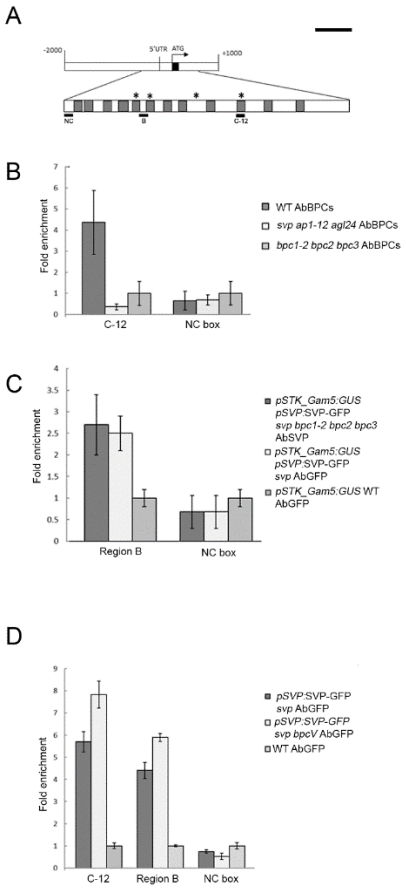


Figure 4. ChIP experiment on different mutant backgrounds. (A) Schematic diagram of the STK locus indicating the regions analysed by chromatin immunoprecipitation (ChIP; black bars). Black boxes, exons; white boxes, promoters and introns; asterisks, C-boxes; grey boxes, CArG-boxes; scale bar=500 bp. (B) Quantitative Real-time PCR analysis of ChIP assay using chromatin extracted from svp ap1-12 agl24, wild-type (as a positive control), and bpc1-2 bpc2 bpc3 (as a negative control) testing C-12 region and NC box. Antibodies against BPCs of class I were used. (C). Real-time PCR analysis of ChIP assay using chromatin extracted from pSVP:SVP-GFP pSTK_Gam5:GUS svp bpc1-2 bpc2 bpc3, pSTK_Gam5:GUS pSVP::SVP-GFP svp (as a positive control) and pSTK_Gam5:GUS wild-type (as a negative control), testing Region B and NC box. For the IP, commercial antibodies against GFP were used. (D) Real-time PCR analysis of ChIP assay using chromatin extracted from pSVP:SVP-GFP svp bpcV, pSVP:SVP-GFP svp (as a positive control) and wild-type (as a negative control), testing Region B, C-12 and NC box. 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 significantly enriched compared with the controls in at least three independent experiments.

4.1.5 CArG-boxes drive the correct temporal and spatial expression of *STK* and they are necessary for SVP and BPCs binding to the DNA

We further decided to investigate the role of CArG-boxes in determining SVP binding on the promoter of *STK* and successively allowing the correct spatial and temporal expression of the MADS-domain protein. The mutated version of *STK* promoter in which 11 out of the 12 CArG-boxes were altered (henceforth called *pSTK*-CArGm:GUS) was analysed. These putative MADS-domain consensus regions have been mutated considering the following criteria: firstly, by

preserving the DNA conformation, designing only transitions of 4 or 5 bases for each consensus; secondly, by avoiding the mutation of GAGA boxes; lastly, by preventing the formation of new CArG-boxes.

Interestingly, we could detect a strong deregulation when the activity of the *STK* promoter was analysed in *pSTK-CArGm*:GUS plants. In particular, GUS expression was observed in inflorescence and floral meristem (Figure 5E and G) and in all the floral organs (sepals, petals, carpels and anthers) (Figure 5F). These results strongly support the idea that the CArG-boxes on the *STK* promoter are important for the correct expression of the MADS-box gene, but also suggest the possibility that other MADS-domain factors, expressed in different tissues compared to *SVP*, could be involved in *STK* regulation. In fact, since SVP is specifically expressed in floral meristem (Gregis et al., 2008), other MADS-domain factors could bind these regions to regulate *STK* expression throughout flower development. The MADS-domain factor AGL24 regulate *STK* in the inflorescence and floral meristem, acting redundantly with SVP and AP1 (Simonini et al., 2012); some observations suggest that AGL24 might be a regulator of *STK* expression later on in flower development: the MADS-box gene is indeed expressed in several tissues in the young flowers, including carpel and stamen primordia and later on in the pollen and the adaxial surface of the gynoecium (Yu et al., 2004). A putative repressor of *STK* before flower transition is the MADS-domain factor SUPRESSOR OF OVEREXPRESSION of COSTANS1 (SOC1). It acts redundantly with AGAMOUS LIKE24 (AGL24) to activate the floral meristem identity gene *LEAFY (LFY)*, as previously reported (Lee and Lee, 2010). Also, SOC1, along with AGL24 and SVP, repress *SEPALLATA3 (SEP3*), a co-activator of class B and C genes that plays an important role in the correct differentiation of floral meristem (Liu et al., 2009); SOC1 can also directly binds *SHP2* locus (Immink et al., 2012). As mentioned above, SHP2 acts redundantly with SHP1 and STK to control ovule identity (Pinyopich et al., 2003). The ABCDE model has been widely characterized in the last decades; it assesses that different classes of genes, most of them belonging to the MADS-box family, interact to specify floral organ identity (Coen and Meyerowitz, 1991). This observation provides a likely scenario in which they could repress and precisely confine *STK* expression to ovule and placenta to assure a correct development of the flower.

Svp agl24 double mutants showed a severe phenotype in flower development only when plants grew at 30°C; in contrast, only a mild phenotype was detected in normal condition, suggesting that the high temperatures might destabilise the MADS-domain complex and its binding to the DNA (Gregis et al., 2006) Interestingly, it has been proved a temperature-sensitive effects also for

AP1 and BELLRINGER (Bowman et al., 1993; Bao et al., 2004). Our analysis revealed that, out of 39 plants with the *pSTK_CArGm:GUS* construct, 14 showed altered *STK* expression; in order to investigate the hypothesis that high temperature might alter the binding of SVP (and their partners AP1 and AGL24) on DNA, thus leading to a major deregulation in *STK* expression, we analysed the activity of the GUS reporter line in plants growing at 30°C. Our analysis showed no difference, compared to the data obtained in normal condition (22°C). In particular, the high temperature did not alter the reporter gene expression of all the plants that showed no deregulation at 22°C.

Furthermore, we could confirm the importance of CArG-boxes in determining the binding of SVP and BPCs to the *STK* promoter (Figure 5H and 5I, respectively). Collectively, these results prove that the binding of SVP is necessary for the recruitment of BPCs on *STK* promoter, corroborating the hypothesis that BPCs may have a role in the stabilisation of MADS-domain complex to the DNA.

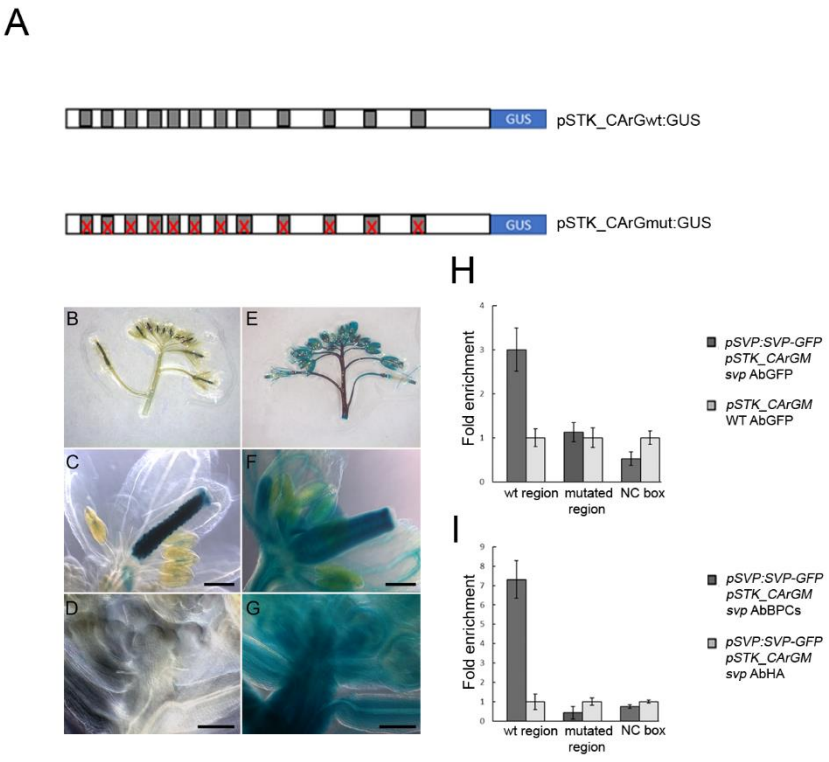


Figure 5. Mutation of CArG-boxes avoids SVP and BPCs binding to STK promoter. (A)Schematic representation of the STK promoter versions generated: dark grey squares represent CArG-boxes wild-type and mutated (crossed).(B)-(G) GUS staining on inflorescences from pSTK_CArGwt:GUS (B-D) and pSTK-CArGm:GUS (E-G): whole inflorescence [(B) and (E)]; mature flower [(C) and (F); inflorescence meristem (IM), floral meristems (FM) and young flowers [(D) and (G)]; scale bars in (C), (D), (F) and (G)=100 μm. (H) Quantitative Real-time PCR analysis of ChIP assay using chromatin extracted from pSVP:SVP-GFP pSTK-CArGm svp and pSTK-CArGm as a negative control, testing Region B and NC box (indicated in Figure 4). For the IP, antibodies against GFP have been used. (I) Quantitative Real-time PCR analysis of ChIP assay using chromatin extracted from pSVP:SVP-GFP pSTK-CArGm testing Region B and NC box. For the IP, antibodies against Class I BPCs have been used; for negative control commercial antibodies against HA was 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 significantly enriched compared to the controls in at least three independent experiments.

4.1.6 SEU and LUG act redundantly to repress *STK* expression in the flower

Putative regulators of *STK* are the corepressors SEUSS (SEU) and LEUNIG (LUG). To address the question whether SEU and LUG are important for driving the correct expression of the homeotic gene, we crossed *lug-3* and *seu-1* mutants with *pSTK*:GUS plants. In subsequent generations, plants homozygous for the *seu-1* and *leu-3,* respectively and with *pSTK*:GUS construct were selected for GUS staining. The expression of GUS was confined in ovule and placenta in *seu-1* and *lug-3* single mutants, as in the wild-type (figure 6D and 6F, respectively); moreover, a strong blue staining was observed in the unproperly fused carpels in *leu-3* mutant (figure 6F).

To test whether SEUSS and LEUNIG act redundantly for the correct *STK* expression, we crossed *seu-1* with *pSTK*:GUS with *leu-3* with *pSTK*:GUS construct. In subsequent generations, plants homozygous for the *seu-1* and *lug-3* mutations with the *pSTK*:GUS construct were selected. *lug-3 seu-1* double mutant shows a very severe phenotype, due to ectopic expression of class B and C gene (Figure 6G). *STK* expression is never registered in first stages of flower development in *pSTK:GUS* wild-type plants, as depicted in Figure 6I. Interestingly, in *lug-3 seu-1* double mutant background, the activity of the *STK* promoter was detected in the first stages of flower development (Figure 6L) till mature flowers (Figure 6N). Despite its defects, here, there is no ectopic expression of *STK*; in particular, we noticed a blue staining in a concentric structure, related to the region of the whorl 4. Even though there are some ovule-like structures, showing activity of the *STK* promoter, *seu-1 lug-3* mutant is sterile. These results suggested that LUG and SEU act redundantly for the temporal regulation of *STK* expression, as suggested by the early reported activity of the GUS line in the young flowers. It has been previously shown that they can form a complex with SVP and AP1 (through SEU-AP1 interaction) to repress the expression of *AGAMOUS* (*AG*), a carpel and stamen identity gene (Gregis et al., 2013). Despite this, *SVP* is specifically expressed in inflorescence and floral meristems and act redundantly with AP1 and AGL24 for the repression of *STK* expression (Simonini et al., 2012). Being stated that, our results suggest that SEU and LUG could regulate the timing of *STK* expression by SVP-independent mechanisms.

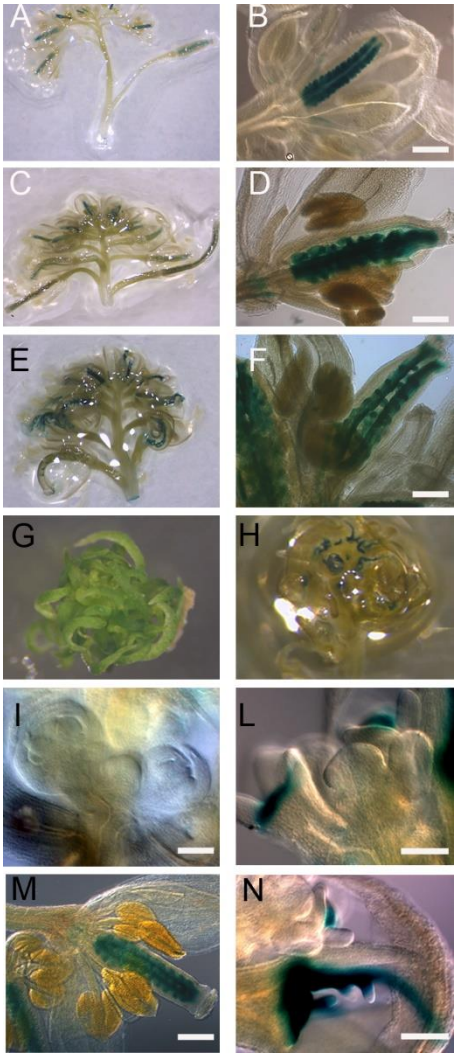


Figure 6. Figure 6. GUS staining on inflorescences from pSTK:GUS wild-type [(A) and (B)], pSTK:GUS seu-1 [(C) and (D)] and pSTK:GUS lug-3 [(E) and (F)]. (G) seu-1 lug-3 double mutant inflorescences; Gus staining of inflorescences of pSTK:GUS seu-1 lug-3 [(H), (L) and (N)] and pSTK:GUS wt [(I) and (M)]. Scale bars= 100 μm.

4.1.7 The regulation of *STK* is influenced by epigenetic modifications

Recent works revealed an intriguing correlation among BPCs and Polycomb group proteins, already investigated in animals; in fact, it has been known that dGAF can cooperate with Polycomb Group factors (PcG) to repress gene expression (Horard et al., 2000). Hecker et al. (2015) and Mu et al., (2017b) showed that BPCs can interact with proteins belonging to PRC1 and PRC2, respectively, to repress the expression of their target genes. Besides, the former also proved that class II BPCs colocalized with PRC2 components in vivo.

Intriguingly, both SVP and BPC6 are able to interact with LHP1 (Liu et al., 2009; Hecker et al., 2015). LHP1 is a component of the PRC1 complex and recognize locus marked by H3K27me3 in vivo acting as part of a mechanism that represses the expression of PRC2 targets (Turck et al., 2007). Also, it

has been recently reported that LHP1 could directly interact with several PRC2 members (Derkacheva et al., 2013; Wang et al., 2016), most likely to facilitate recruitment of PRC2 to target genes. Modifications such as the trimethylation of histone H3 have important impacts on gene transcriptional activity and chromatin organization; in fact, the presence of H3K27me3 is mainly correlated with gene silencing in animals and plants and H3K27me3 targets are enriched for genes with tissue-specific expression patterns in Arabidopsis; this evidence suggests that this epigenetic mark is modulated in response to developmental cues (Zhang et al., 2007). As we have already mentioned, the locus of *STK* shows strong coverage of H3K27me3 deposition in seedlings (Turck et al., 2007; Lafos et al., 2011); also, the localization of LHP1 on these regions has been reported in seedlings (Turck et al., 2007; Zhang et al., 2007). The data above suggests that *STK* expression could be regulated by LHP1 via H3K27me3 maintenance on its locus. All these results came from high throughput analysis, both performed in seedlings; thus, we decided to investigate whether the same results would be obtained in inflorescence tissue as well.

First we showed that the *STK* locus is still decorated with the H3K27me3 mark in the reproductive tissue (Figure 7B). Second, we addressed the question of whether the deregulation of *STK* observed in *bpcV* background correlates with alterations in the H3K27me3 epigenetic mark by analysing its deposition at the *STK* locus in inflorescences. In particular, we tested two regions in *STK* locus: region 1 is located in the H3K27me3 – enriched region published by Li et al. (2015) (spanning -2627 upstream *STK* transcriptional start site to +2050 pb downstream *STK* transcriptional start site); region 2 is localized 3 pb downstream the stop codon of the gene (Figure 7A). Interestingly, in *bpcV* inflorescences we detected a reduction of H3K27me3 deposition in both the two regions analysed, compared to the wild-type (Figure 7B). These results were consistent with the observed deregulation of *STK* expression in the same background (Figure 1G and H) and suggest an active role of BPCs in the establishment or maintenance of the repressor mark.

As previously reported, several members of BPCs family could interact with PRC2 components complex, directly involved in the deposition of H3K27me3 repressive mark (Mu et al., 2017b; Xiao et al., 2017). Our results showed that mutation of BPCs of class I and II affects the level of the repressor mark on *STK* locus. Thus, BPCs might also be involved in the recruitment of the PRC2 to the DNA. To test this hypothesis, we analysed the expression of *STK* in different mutants of PRC2 members: CURLYLEAF (CLF), FERTILIZATION-INDEPENDENT ENDOSPERM (FIE) and SWINGER (SWN), as previously reported (Introduction, section 3.8.1); we could not detect any statistically significant difference, compared to the wild-type (Figure 7C). This result suggests that other PRC2

factors might be specifically involved in the deposition of H3K27me3 on *STK* locus. Previously, a participation of EMBRYONIC FLOWER complex (EMF) in the repression of *AG, PISTILLATA (PI) and SEPALLATA3 (SEP3)* expression in the flower has been reported (Yoshida et al., 2001; Kinoshita et al., 2001; Chanvivattana et al., 2004; Calonje et al., 2008); interestingly, Simonini et al. (2012) showed BPCs binding to the promoter of *AG* and *SEP3* homeotic genes. Also, BPCs of class II have been shown to localize with the PRC2 factor VERNALIZATION2 (VRN2). Being stated that, it would be worth to check the expression of *STK* in these PRC2 components mutant backgrounds.

Compared to animals, PRC1 has evolved novel functions in plants, acquiring distinct roles in gene regulation; thus, it can act independently from PRC2 members to modulate its targets expression (Calonje, 2014). As matter of fact, PcG targets can be regulated by the two complexes via independent mechanisms (Yang et al., 2013). Our results showed that none of the PRC2 members we considered are involved in the regulation of *STK* expression. On top of that, the fascinating association of LHP1 to BPCs, SVP and *STK* (Turck et al., 2007; Liu et al., 2009; Hecker et al., 2015) suggests that the PRC1 member LHP1 could play an important role in the modulation of the expression of the homeotic gene during flower development, further analysed below.

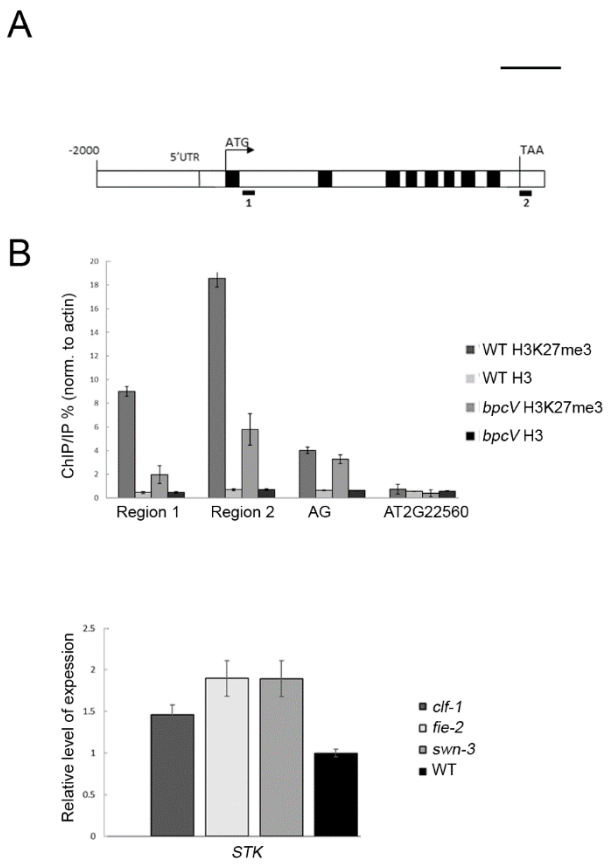


Figure 7. Epigenetic regulation of STK. (A) Schematic representation of the STK genomic region. Grey boxes indicate exonic regions. Black bars indicate the regions analysed by chromatin immunoprecipitation (ChIP); Region 1 is located in the H3K27me3 – enriched region published by Li et al. (2015) spanning -2627 upstream STK-transcriptional start site to +2050 pb downstream STK-transcriptional start site), whereas Region 2 is localized 3 pb downstream the stop codon of the gene. Black arrow indicates the STK-transcription start site. Scale bar= 500 bp. (B) Quantitative Real-Time PCR analysis determining the levels of H3K27me3 across the STK locus. Quantitative Real-Time PCR analysis of STK sequences in precipitated chromatin was used to infer the methylation of histone H3 at lysine 27 (H3K27me3) and total histone H3 representation (histone H3 density). Ct values were used to calculate the IP/IN signal. ChIP enrichments are presented as the percentage (%) of bound/input signal normalized to actin levels in the relative regions. We tested the efficiency of IP on histone modifications by quantifying the presence of the H3K27me3 mark in AG region which carries the mark H3k27me3, reported in Li et al. (2015). H3K27me3 mark in AT2G22560 was used as negative control for H3K27me3 mark (Li et al. 2015). Error bars indicate standard deviations from four biological replicates.(C) Expression analysis of STK by quantitative Real-time PCR in clf-1, fie-2, swn-3, and wild-type inflorescences. The expression of STK in the wild-type was set to 1.

Mutation in the LHP1 locus results in pleiotropic effects due to the mis regulation of several genes during plant development (Larsson et al., 1998). To explore the role of *LHP1* in the regulation of *STK* in the flower, its expression was analysed in *lhp1* background: the following mutant, much smaller compared to the wild-type, shows a strong early flowering phenotype (Larsson et al., 1998).

In line with our hypothesis, the MADS-domain factor *STK* was deregulated in the *lhp1* mutant background, compared to the wild-type as shown by *in-situ* hybridisation assay. In fact, *STK*

expression was detected in inflorescence and floral meristem as well as in the first stages of flower development (Figure 8C). In the mature flower, *STK* expression was confined to placenta and ovules (Figure 8D). Compared to our *in-situ* hybridisation results in the *bpcV*, it might suggest that *STK* regulation by BPCs could act at different levels; BPCs could indeed modulate *STK* expression in the flower also via distinct processes with no PRC involvement. We could confirm our *in-situ* hybridisation assay by checking the expression of *STK* by quantitative Real-Time PCR; as shown in Figure 8E, the MADS-box gene is upregulated in *lhp1* background.

Furthermore, we validated the association of LHP1 to the *STK* locus in inflorescences; we tested both Regions 1 and 2 (illustrated in Figure 7A) and we detected enrichment in Region 2, close to the 3' UTR of the homeotic gene. Interestingly, Gregis et al. (2013) showed that SVP preferentially binds the 3'UTR of its targets. This result supports the idea that SVP, interacting with LHP1, could recruit the PRC1 factor to the *STK* locus for a successful repression of the gene; it would be interesting to check whether the mutation of SVP and its redundant partners *AP1* and *AGL24* would affect the association of LHP1 and the deposition of H3K27me3 mark on the *STK* locus, specifically on Region 2.

Collectively, our results provide new insights into the role of BPCs in the recruitment of PRCs members on the regulatory region of target genes. Our findings indeed suggest that SVP and BPCs of class II can recruit LHP1 on the of *STK* locus and act redundantly with the class I members to regulate the deposition of the H3K27me3 mark. Also, the observed overexpression of *STK* in the *lhp1* background confirmed a direct role of the PRC1 member LHP1 in the regulation of *STK.* Different combinations of PRC proteins distinctively regulate different programs, throughout plant development; also, mutation in *LHP1* affects H3K27me3 deposition to the DNA (Wang et al., 2016). Since the mutation of PRC2 members that we tested did not affect *STK* expression, LHP1 might interact with other PRC2 proteins to repress *STK* expression in the flower, by maintaining and spreading the H3K27me3 mark to the homeotic gene locus (Veluchamy et al., 2016). Derkacheva et al. (2013) interestingly reported that the PRC1 member can interact with the EMF complex in vivo. Furthermore, it has been previously showed that LHP1 is physically associated with VRN2 (Hecker et al., 2015). Thus, the deposition and the maintenance of the H3K27me3 on *STK* locus could be caused by an interplay between LHP1 and those PRC2 proteins.

Several BPCs targets have been discovered in the recent years. Most of them are associated with PRC mediated silencing: the KNOX gene *BREVIPEDICELLUS (BP)* is repressed by BPCs throughout

flower development (Simonini and Kater, 2014); its expression is directly regulated by the recruitment of the EMBRYONIC FLOWER (EMF) complex by ASIMMETRIC LEAVES 1 and 2 (AS1 and AS2), that triggers H3K27me3 deposition onto the DNA (Lodha et al., 2013). Intriguingly, the core subunit of the EMF complex, MULTICOPY SUPRESSOR OF IRA1 (MSI1) interacts with LHP1, unravelling a novel role for the PRC1 factor; in fact, it suggests that LHP1 is necessary for PRC2 recruitment to the target genes. Also, *FUS3* has recently been characterised as a BPCs target (Roscoe et al., 2019; Wu et al., 2019) and already reported to be a target of PRCs (Makarevich et al., 2006; Zhang et al., 2007; Bouyer et al., 2011; Yang et al., 2013; Xiao et al., 2017).

Liu et al. (2009) proposed that SVP is required to recruit the PRC1 factor LHP1 to the promoter of *SEP3*. In accordance to this hypothesis, H3K27me3 deposition to *SEP3* locus is reduced in *lhp1* background. Likewise *SEP3,* we previously characterized *AGAMOUS (AG)* as a target of BPCs and SVP, (Gregis et al., 2009; Simonini et al., 2012), whose expression is deregulated in *lhp1 bpc4 bpc6* triple mutant, confirming a LHP1-BPCs of class II interplay at the protein level (Hecker et al., 2015). It would be interesting to analyse the expression of the MADS-domain factor *AG* in our *bpcV* mutant to check whether BPCs of class I and II act redundantly to regulate its expression, as we reported for *STK*.

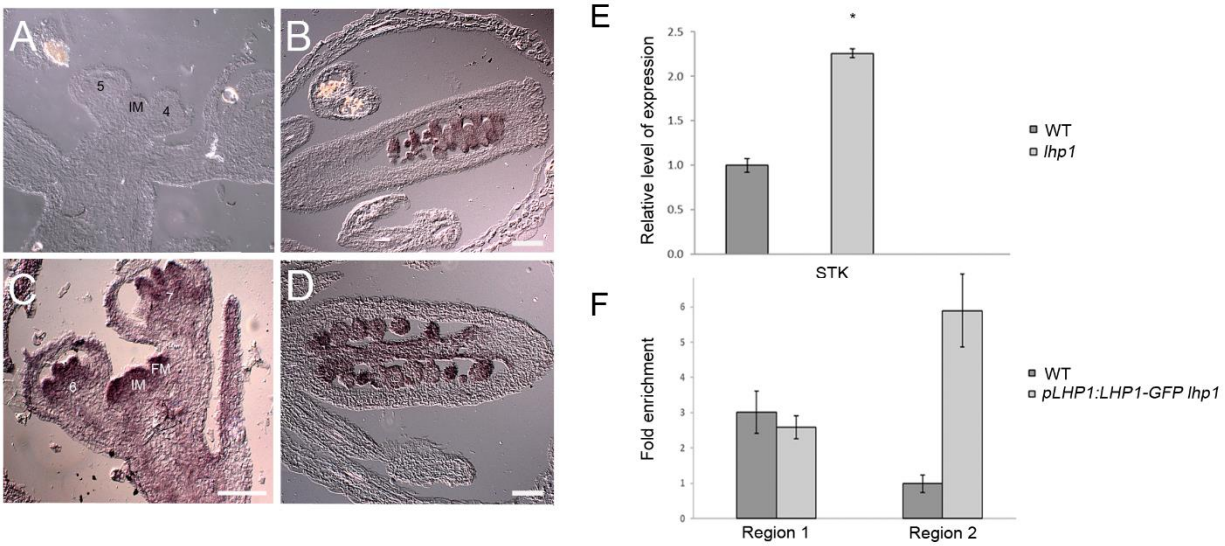


Figure 8. LHP1 directly regulates STK during flower development. (A)-(D) In situ hybridization on wild-type [(A) and (B)] and lhp1 [(C) and (D)] inflorescences using a STK-specific antisense probe (Brambilla et al., 2007). IM: inflorescence meristem; numbers represent flower stages; scale bars=50 µm. (E) Expression analysis of STK by Real-time PCR in lhp1 and wild-type inflorescences. The expression of STK was normalized to that of ubiquitin and the expression level in wild-type was set to 1. Asterisk indicates P < 0.05 in a Student's t-test. (F) Quantitative Real-time PCR analysis of ChIP assay using chromatin extracted from pLHP1:LHP1-GFP lhp1 (Kotake et al, 2003) and wild-type (as a negative control), testing the Region 1 and Region 2 (Figure 7A).

For the IP, commercial antibodies against SVP 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 significantly enriched compared with the controls in at least three independent experiments.

4.1.8 Genome-wide analysis of BPCs and MADS-domain factor binding site locations

Understanding the molecular mechanisms in which BPCs and SVP complex act is important since it is possible that the mechanism used by these factors can be accounted for the regulation of many other genes throughout plant development. This hypothesis is based on the following observations: first, many genes contain C-boxes and CArG-boxes in their putative promoter regions; second, BPCs are ubiquitously expressed in plants while MADS-domain factors are specifically expressed in all the fundamental developmental stages; and lastly, *bpc* multiple mutants show multiple severe phenotypic effects. As matter of fact, we registered a highly significant overlapping between regions bound by BPCs and MADS-domain factors, from DAP-seq data (Figure 9).

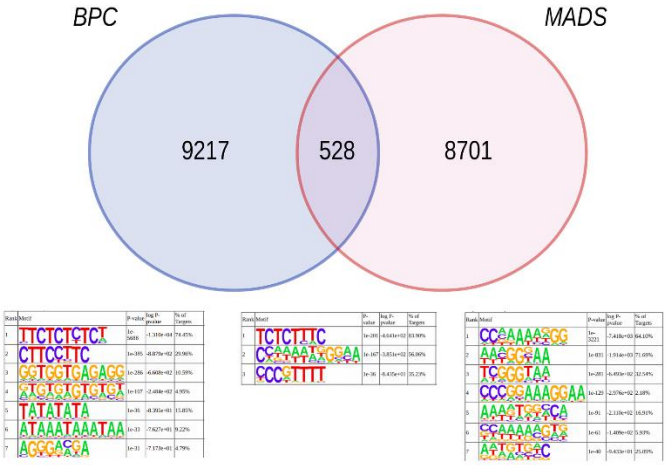


Figure 9. Analysis of BPC and MADS-box transcription factor families binding sites. Venn diagram displaying the number of DAP-seq peaks and the common number of peaks associated to the BPC and MADS transcription factor families according to our analysis of the data by O'Malley et al (see methods). Enriched motifs, as recovered by Homer (p-value<=1e-30), are displayed underneath.

Berger et al. (2011) identified three cis-elements required for LEAFY COTYLEDON2 (LEC2) repression: GAGA binding sites (bound by BPCs), CArG-boxes (MADS-domain consensus regions)

and PRE regions (associated to PRC). Hence, we subsequently cross-referenced coincident MADS-BPCs DAP-seq peaks (correlated to putative targets of MADS-BPCs complex) with H3K27me3 ChIP-seq peaks (Lafos et al., 2011). We obtained a list of 93 candidates (DataS2, Manuscript #1), that will be analysed to find putative targets that might be co-regulated by MADS-domain factors and BPCs *in vivo* via H3K27me3 deposition to the gene locus. Some interesting genes emerged from the list; in example, the already mentioned BPCs targets BP and FUS3 (Simonini et al., 2012; Roscoe et al., 2019; Wu et al., 2019), suggesting a conservative mechanism for target regulation. Hecker et al. (2015) showed that most of the genes decorated with the repressive marks resulted to be upregulated in the *lhp1* mutant background; thus, our results support the hypothesis that BPCs and MADS-domain factor could repress their target expression via LHP1 recruitment.

Collectively these findings, corroborate the hypothesis that MADS-BPCs-LHP1 complex (illustrated in Figure 10), regulating the expression of subset of genes, play a paramount role in the regulation of fundamental phases of plant development; also, they stress that a better understanding of their synergistical activity is an important key to decode gene regulation in plant.

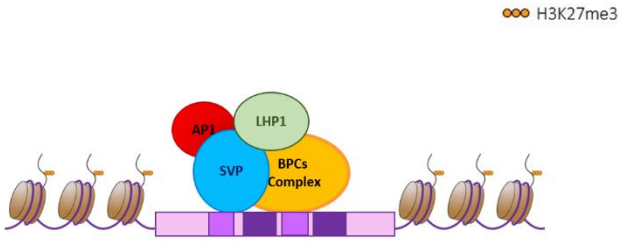


Figure 10. Model of the protein complex formed to represses gene expression during flower development. BPCs and SVP bind C-boxes (in dark purple) and CArG-boxes (in light purple) respectively, and recruit LHP1 to a subset of gene loci.

4.2. Unravelling the role of BPCs in gynoecium development

The second aim of my PhD project is to investigate the role of BPCs of class I and class II during septum development and successively in transmitting tract differentiation, mainly by discovering new putative targets and interactors of BPCs during gynoecium development.

4.2.1 Mutation of BPCs affects septum development and transmitting tract differentiation

BPCs are ubiquitously expressed throughout plant development; thus, their mutation caused a pleiotropic phenotype, as reported by Monfared et al. (2011). As described above (Results, section 4.1.1), the *bpc 1-2 bpc2 bpc3 bpc4 bpc6 (bpcV),* that carries all the knock-out alleles, has been generated in our laboratory. A first phenotypical analysis reveals a wide range of vegetative and reproductive defects, as shown In Figure 1A. The plant, smaller in size, is characterized by a reduced seed yield. Analysis of the silique revealed that the fruit, shorter than the wild-type is impaired in septum development (Figure 1B and Figure 1C, bottom). It is worth to mention that also the *bpc1-2 bpc2 bpc3* mutant shows occasionally the same defects describe above in septum development, as illustrated in Figure 1D). As described in the Introduction (Section 3.4), the septum is a structure that divides the siliques in two halves. Later in gynoecium development, some cells at its centre differentiate in the transmitting tract, that connecting to the style, forms a continuous tract for pollen tube growth and guidance. Thus, it plays a fundamental role in reproduction and defects in its development could affect the reproductive success of the plant. The septum originates from the CMMs at stage 8; by stage 10 it is fully fused and flanked by the placenta (Roeder and Yanofsky, 2006). To determine at which distinct stage of flower development the defect occurred, a morphological analysis was conducted on inflorescences from wild-type (as control), *bpc 1-2 bpc2 bpc3* and *bpcV*. The samples were transversally sectioned and stained with Alcian blue to detect the cells of the transmitting tract, whose extracellular matrix is polysaccharides-rich, whereas neutral red was used as counterstain. Both *bpc 1-2 bpc2 bpc3* and *bpcV* mutants showed some defect in CMMs fusion and later in septum formation at stage 10 (Figure 1E, second and third column), but only the latter lacks transmitting tract cells differentiation at stage 12 (Figure 1E, third column); *bpc1-2 bpc2 bpc3* indeed presents a strong blue staining, even though the tissue morphology is severely affected (Figure 1E, second column).

Our results showed for the first time that BPCs have a role in CMMs fusion and septum formation; also, they revealed that BPCs of class I and class II act redundantly for transmitting tract differentiation, as suggested by the different results in the two mutant backgrounds. The missing staining with Alcian blue in *bpcV* underlines a defect in transmitting tract differentiation. We cannot rule out the hypothesis that in the mutant background a different composition of the extracellular matrix affects the staining of the transmitting tract cells. The following speculation would explain why fertilization is severely reduced but not fully abolished. Several mutants have been characterized for their role in septum and transmitting tract development. *NO TRANSMITTING TRACT (NTT)* is required for transmitting tract differentiation; the mutant indeed showed a high sterility due to a failure in the pollen tube guidance to the ovary, mostly confined to the upper part of the pistil (Crawford et al., 2007). Despite the drastic defect, some ovules were still fertilized, suggesting that transmitting tract increases the fertilization efficiency by helping the pollen tube growth but it is not required for a successful reproduction (Crawford et al., 2007). Actually, in Arabidopsis, pistils are small structures; thus, the pollen tube journey is shorter, compared to other species. Being stated that, defects in transmitting tract development might have more drastic effects in bigger plants. In rice, the journey of the pollen tube through the style transmission tract (STT) is a long and arduous process that requires an efficient communication between the female and the male tissues (Xu et al., 2017). Therefore, the correct formation of the STT is pivotal for a correct pollen tube growth and a successful double fertilization process, that lastly affect seed-setting rate. Despite the relevance of this process, little is known about the mechanisms involved in a correct STT development. Thus, unraveling the processes governing a successful transmitting tract development in Arabidopsis could provide new insights into the molecular genetic control of grain yield in crops.

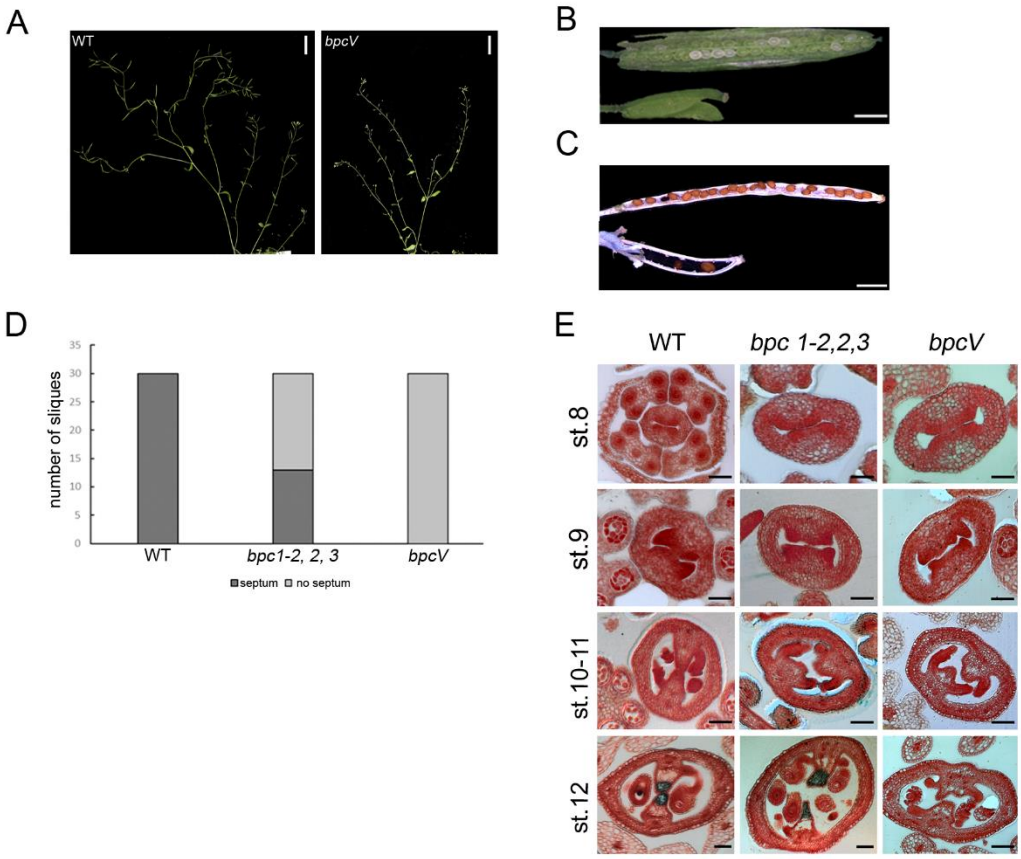


Figure 1. Mutation in BPCs affects septum development. (A) wild-type and bpcV plants; plants were photographed six weeks after sowing; scale bar= 1 cm (B) Unripe silique of wild-type (top) and bpcV (bottom); (C) Ripe silique from wild-type (top) and bpcV (bottom); scale bar= 2mm; (D)Analysis of septum in 30 siliques per wild-type, bpc1-2 bpc2 bpc3 and bpcV; (E) Alcian blue staining of trasversal sections of wild-type (first column), bpc1-2 bpc2 bpc3 (second column) and bpcV (third column) flowers at different developmental stages; neutral red was used to counterstain the cell wall; scale bars= 20µm.

4.2.2 Transcriptomic analysis of *bpcV* inflorescences deepens BPCs of class I and II role in gynoecium development

The drastic phenotype of *bpcV* showed that BPCs are involved in septum formation and transmitting tract differentiation. To gain a deeper insight into the role of BPCs of class I and II during gynoecium development, we performed a RNA-Seq on inflorescences (up to stage 12 of flower development, as reported by Smyth et al. (1990), from wild-type and *bpcV* mutant.

From the resulted transcriptomic analysis, over 2000 genes emerged as differentially expressed between *bpcV* and wild-type (Dataset 1). Most of these genes were downregulated (1275 genes),

whilst 922 appeared to be upregulated, suggesting that BPCs might act either as activators or repressors.

Most of the genes were transcription factors. Intriguingly, among them three of the main players for septum formation and/or transmitting tract development emerged: the zinc-finger factor *NTT* (mentioned in Results, section 4.2.1) and the bHLH transcription factors *SPT* (described in Introduction, section 3.4.1) and *CES/HAF (CESTA/HALF FILLED)* (Crawford and Yanofsky, 2011) resulted indeed downregulated. BPCs factors could activate their expression to define a correct gynoecium development.

GO analysis of genes that are differentially expressed in the *bpcV* mutant showed some transcripts involved in cell wall biogenesis and remodelling. Interestingly, Herrera-Ubaldo and de Folter (2018) confirmed the importance of cell wall modifications for a correct gynoecium development and a successful fertilisation process. In particular, two regions must have a dynamic cell wall composition: the septa primordia, for a correct septum formation, due to epidermal cells fusion, and the centre of the septum, for the formation of the transmitting tract. Thus, it might be possible that the defects in septum fusion and successively transmitting tract differentiation are caused by alteration in cell wall components (i.e. pectins, lignin and lipids), due to a deregulation of key factors involved in their biosynthesis and metabolism. The excretion of a special extracellular matrix (ECM) from the transmitting tract cells helps pollen tubes in their journey through the pistil (Herrera-Ubaldo and de Folter, 2018). As previously suggested, the missing Alcian blue staining registered in the *bpcV* mutant might be associated to a different composition in polysaccharides of the ECM of the transmitting tract cells. It would be interesting to check, using distinct staining (Herrera-Ubaldo and de Folter, 2018), whether the other components of their ECM were as well altered; in example, Alcian blue at different pH could stain specifically glycoproteins (at pH 1.0) or hyaluronic acid (at pH 2.5).

Tung et al. (2005) identified a set of genes specifically expressed in the transmitting tract; they were associated with either transcriptional activity, metabolism and cellular plasticity and response. Intriguingly, many of these factors are deregulated in our *bpcV* mutant, supporting our reported role for BPCs in transmitting tract development. The ethylene response factor *RAP2.6* resulted to be downregulated in our *bpc* quintuple mutant; interestingly, Swidzinski et al. (2002) showed that genes of the ethylene pathway are associated to cell death. In a wild-type situation, events of programmed cell death (PCD) are required for a successful transmitting tract

development (Crawford et al., 2007). Thus, the defect registered in our mutant might be partially caused by alteration in this pivotal step.

Furthermore, a glycoprotein membrane precursor GPI gene was likewise downregulated in *bpcV*. As reported above, the transmitting tract cells are characterized by a peculiar ECM, rich in glycoproteins and polysaccharides. Being stated that, defects in the biosynthesis or either the metabolisms of the ECM components could alter the correct formation of the transmitting tract.

In order to discover new putative direct targets of BPCs we cross-referenced the list of the deregulated genes in our RNAseq (Dataset 1) with a list of genes whose putative promoter region was bound both by BPC1 (class I BPC) and BPC6 (class II BPC), obtained from a public plant cistrome database (Malley et al., 2016; Dataset 2, first column); it is based on DNA affinity purification sequencing (DAP-seq) assay (Bartlett et al., 2017; see materials and methods for details).

The resulted list (Dataset 2, second column) presented 165 genes. GO analysis shows that most of the genes identified are involved in transcription activity; AP2/B3-like transcriptional factor family protein made up the most enriched family of TFs in the list. Intriguingly, *NTT* emerged, again, as putative target of BPCs. Thus, the GAGA binding transcription factors could directly bind the promoter of *NTT* to finely orchestrate transmitting tract formation and the correct differentiation of its cells. In fact, *NTT* has been characterized as a gene specifically required for transmitting tract development. Being stated that, the defects in *bpcV* mutant could be partially caused by lower levels of *NTT* in the maternal tissue. This hypothesis will be further explored by checking whether BPCs could bind the regulatory region of *NTT in vivo* in inflorescences.

The list of putative BPCs targets resulted to be enriched also for factors associated with catalytic and transporter activity (i.e. phosphates, hydrolases and cellulose synthases). This preliminary data suggests that BPCs might govern the correct development of the septum and the transmitting tract either by acting upstream to DNA-binding proteins, whose regulation of downstream targets might be fundamental for the fulfilment of these processes and by modulating the correct biosynthesis, metabolism and transportation of key-components of these fundamental maternal tissues.

Furthermore, many of the differentially expressed genes that apparently show no direct binding by BPCs, according to the DAP-seq, were involved in response to hormones. Fascinatingly, the first emerging pathway in the analysis of the list performed on DAVID software

(https://david.ncifcrf.gov) was the plant hormone signal transduction, depicted in Figure 2. This result showed that *bpcV* mutant are affected in hormone transport and signalling. In particular, as regards cytokinin, the *ARABIDOPSIS HISTIDINE KINASE4 (AHK4)* is downregulated in our mutant, compared to the wild-type. It plays a pivotal role as cytokinin sensor throughout plant development; its mutation causes indeed severe vegetative and reproductive defects (i.e. development of aberrant flowers), along with a complete insensitivity to cytokinin treatment (Ueguchi et al., 2001). Also, several type-A *ARABIDOPSIS RESPONSE REGULATORS (ARRs)* emerged in the downregulated genes; they respond to cytokinin levels and act negatively on its signalling (To et al., 2007).

As regards the phytohormone auxin, the transmembrane amino acid transporter *AUXIN RESISTANT1 (AUX1*) is upregulated in our mutant, thus leading to a hypersensitivity to auxin, as suggested by Swarup et al. (2004). Also, *the Auxin/Indole-3-Acetic Acid (Aux/IAA)* genes, that play a pivotal role in repressing the expression levels of genes activated by *AUXIN RESPONSE FACTORS (ARFs)* (Luo et al., 2018) were downregulated. As matter of fact, the auxin signalling was perturbed; we could register, indeed, an upregulation of *ARFs* as well as *SMALL AUXINE UPREGULATED RNAs (SAURs)* and Auxin-responsive *GH3* family protein genes.

Auxin and cytokinins play different roles during plant growth: the former, indeed, stimulates cell differentiation whereas the latter induces mitotic divisions. Therefore, they have been recently named the yin and the yang of development (Schaller et al., 2015). Their dynamic interaction, which can be antagonist or synergic, play a crucial role in gynoecium development; thus, alterations in their signalling cause drastic phenotypes (Reyes-Olalde et al., 2013; Reyes-Olalde et al., 2017b). Our preliminary analysis on transcriptomic data suggested that crosstalk between the two phytohormones might be impaired in our mutant. As matter of fact, *bpcV* could experience unbalanced levels of auxin and cytokinin; they could cause alterations in their downstream processes, thus impairing the correct formation of the gynoecium (Müller et al., 2017).

On top of that, Herrera-Ubaldo et al. (2018) developed a computational interaction map of transcription factors controlling gynoecium development, stage by stage, based on protein-protein interaction data, expression patterns and functional information of a large number of transcription factors; interestingly, some of the transcription factors identified emerged in the list of differentially expressed genes (Dataset 1) and were mostly involved in auxin and cytokinins

pathways, stressing that the cross-talk between the two phytohormones plays a pivotal role in the orchestration of gynoecium formation and tissues differentiation.

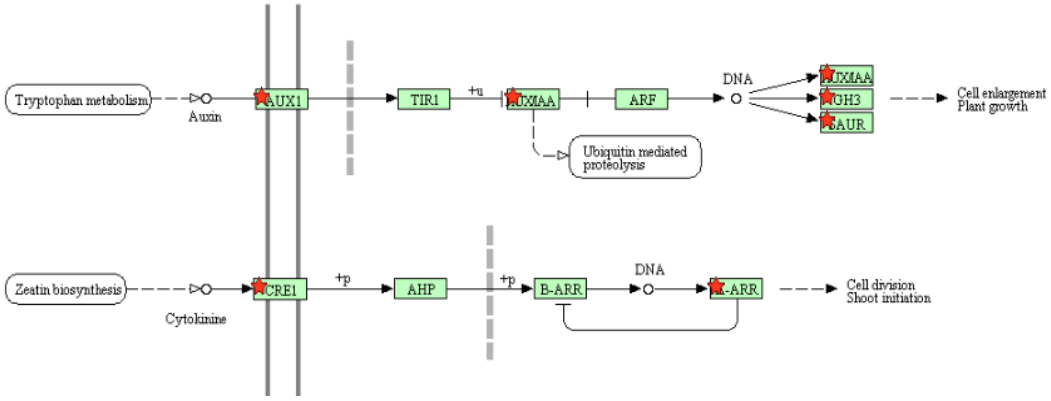


Figure 2. Representation of plant hormone signal transduction pathway. Red stars represent the family genes differentially expressed in bpcV, compared to the wild-type.

4.2.3 *SPATULA* as a putative target of BPCs during gynoecium development

The defects in septum formation and transmitting tract differentiation observed in *bpcV* might be caused by a mis-regulation of downstream factors, as supported by our RNA-seq data. As mentioned before (Results, section 4.2.2), among the genes downregulated in our *bpcV* mutant, *SPT* popped up. A quantitative Real-Time PCR could confirm the lower levels of the bHLH factor in the the *bpc* multiple mutant backgrounds (Figure 3A). Interestingly, the level of downregulation of *SPT* was proportional to the number of knockout BPCs genes; the strongest downregulation was indeed registered in *bpcV*. This result suggests that BPCs of class I and class II act redundantly to regulate *SPT* expression during gynoecium development. To determine whether the mutation of the BPCs would also affect *SPT* spatial-temporal expression a *in-situ* hybridization assay was performed. In the wild-type, *SPT* is strongly expressed in floral meristems and in the developing flowers (Figure 3B); later, its expression becomes restricted to the medial tissues, the septum and the transmitting tract (Figure 3C and D); also, *SPT* expression is detected in ovule primordia (Heisler et al., 2001). In the *bpc1-2 bpc2 bpc3* and *bpcV* mutant, the pattern of expression in the floral meristems and the early floral buds was maintained (Figure 3E and G, respectively); later on gynoecium development, its expression was confined in the ovules (Figure 3F and 3H), whereas in *bpcV* no signal was detectable in other tissues, due to the missing septum and transmitting tract

(Figure 3I), confirming that BPCs mutation does not alter *SPT* pattern of expression. A new *in-situ* assay will be performed to collect data about *SPT* expression at stage 12 in the *bpc1-2 bpc2 bpc3* background.

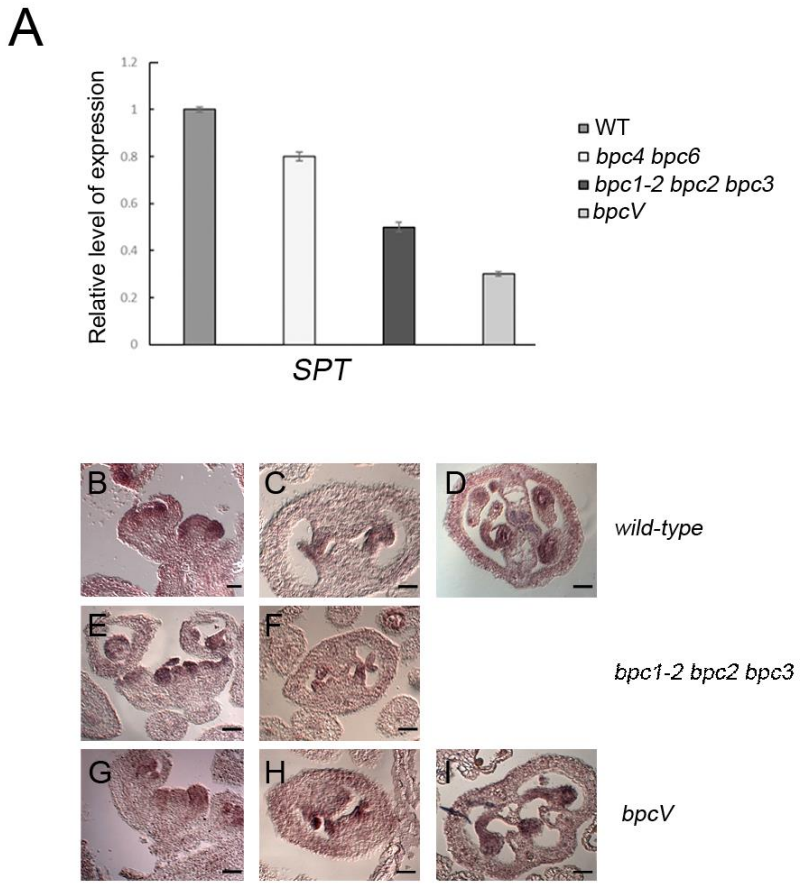


Figure 3. SPT as putative target of BPCs. (A) Expression analysis of SPT by quantitative Real-time PCR in wild-type, bpc4 bpc6, bpc1-2 bpc2 bpc3 and bpcV inflorescences; expression of SPT in wild-type was set to 1. (B)-(G) In situ hybridisation assay with a SPT-specific antisense probe on wild-type [(B)-(D)] bpc1/2 bpc2 bpc3[(E)-(F)] and bpcV [(G)-(I)] inflorescences; scale bar=20 µm.In (G) multiple "septa", derived by multiple MMCs are showed.

To investigate whether the drastic phenotype observed in *bpcV* was due to a downregulation of *SPT*, we cloned the CDS of *SPT* in a 35S CAMv promoter vector which led to the overexpression of the bHLH gene throughout the plant. The choice of this approach came from the observation that the constitutive expression of *SPT* could completely rescue the phenotype of *spt-2* siliques in terms of length, seed set and septum formation (Heisler et al., 2001). Then, *bpc1-2 bpc2 bpc3*, *bpcV* and wild-type (as positive control) plants were then transformed with the described construct. We obtained several lines and *SPT* expression was checked by quantitative Real-Time PCR. We selected 3 lines per background, showing upregulation of *SPT* expression, as illustrated in figure

4A, B and 4C. The following lines will be analysed to check whether the overexpression of the gene could recover the *bpcV* defects in septum formation and transmitting tract differentiation.

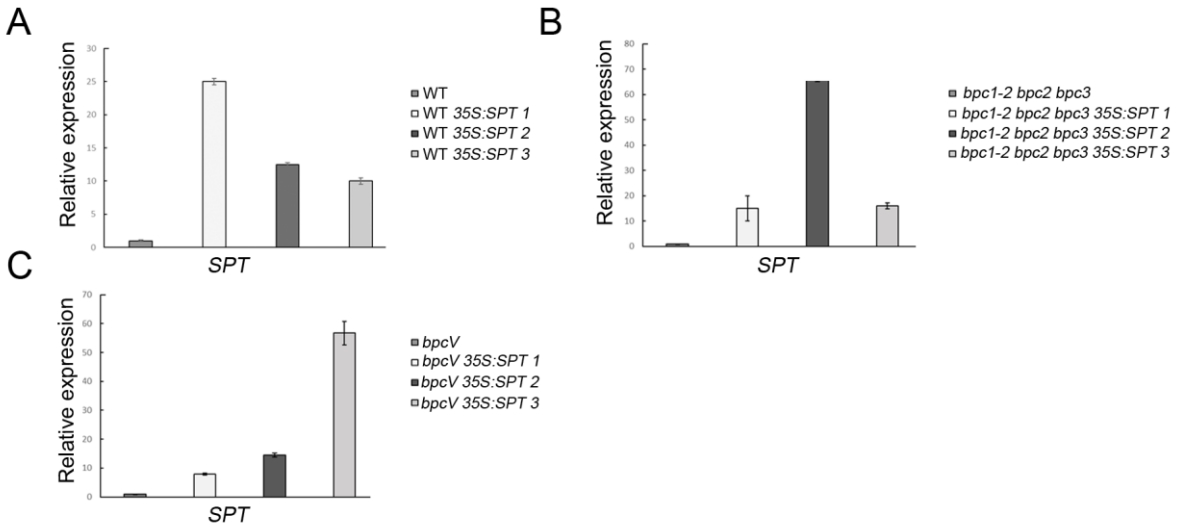


Figure 4. (A)- (C). Expression analysis of SPT by quantitative Real-time PCR in 35S:SPT wild-type (A), 35S:SPT bpc1-2 bpc2 bpc3 (B) and 35S:SPT bpcV (C); expression of SPT in wild-type, bpc1-2 bpc2 and bpc3 and bpcV inflorescences were set to 1 in (A), (B) and (C), respectively.

BPCs acts as transcription factors, binding C-boxes on DNA. From our list (Table 2) *SPT* did not emerge as putative BPCs target. The DAP-sequencing is an elegant approach for a rapid yet efficient analysis of TFs ability to bind DNA sequences. Despite its advantages, it presents several drawbacks; the most relevant one is the lack of chromatin; thus, the contribution of transcription factors complexes are not considered. For this reason, we decided to analyse BPCs binding to the regulatory region of *SPT* by ChIP, using a *35S:BPC1-RFP* line (Wu et al., 2019). The promoter region of the bHLH transcription factor is 9.2 kb long (Groszmann et al., 2010) and presents 7 C-boxes, as shown in Figure 5A; we identified three regions. No enrichment was detected in neither the three regions when BPCs binding was analyzed, as showed in Figure 5B; the promoter region of *FUSCA3 (FUS3)* was used as positive control (Wu et al., 2019). These results suggest that *SPT* is not a direct target of BPCs; alternatively, BPCs could regulate its expression, by activating/repressing intermediate regulators. The knowledge about *SPT* regulation in the gynoecium is still poor and fragmented. *SPT* promotes carpel fusion partially by repressing the NAC genes *CUC1* and *CUC2* in the apical part of the gynoecium (Nahar et al., 2012). Kamiuchi et al. (2014) showed that they play

critical role in the early stages of CMMs formation and positioning: the *cuc1 cuc2* mutant is indeed defective in the formation of either one of the CMMs, whereas plants carrying miR164-resistant forms of CUC1 and CUC2 resulted in extra CMM activity with altered positioning. Interestingly, the *bpcV* occasionally showed extra CMMs (Figure 3I). That being stated, it is tempting to speculate that BPCs might regulate CUCs expression during gynoecium development; preliminary analysis showed indeed that the putative promoter region of CUC2 presents several C-boxes. Although the RNA-seq analysis did not register any down- or upregulation of the gene, we cannot exclude the possibility that BPCs of class I and class II could still regulate its spatial-temporal expression. Thus, it would be worth to check whether the mRNA localization of CUC2 is altered in the *bpc* multiple mutants by *in-situ* hybridisation assay.

On top of that, we cannot rule out the hypothesis that the low expression of the bHLH transcription factor registered in our mutants is caused by a depletion of the tissue *SPT* is mostly expressed in. As matter of fact, the lowest expression of *SPT* was registered in *bpcV* mutant that showed the strongest defects in septum and transmitting tract development. The analysis of the *bpcV* lines, showing an ectopic expression of *SPT* could help in understanding whether the drastic phenotype observed in our mutant is in part due to a deregulation of the bHLH transcription factor; even a partial rescue in the septum fusion or either transmitting tract differentiation could confirm a direct involvement of BPCs in *SPT* regulation.

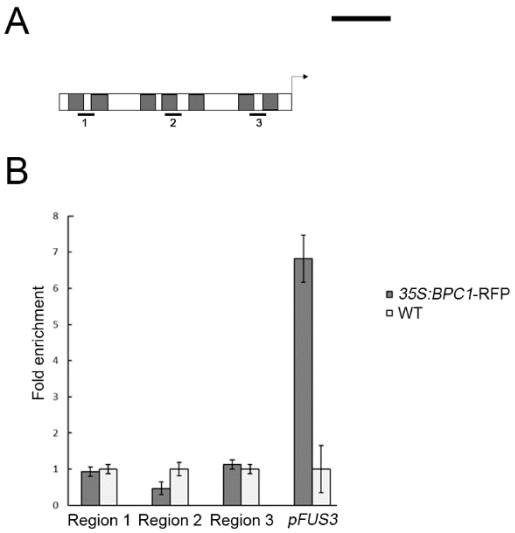


Figure 5. (A) Schematic representation of SPT promoter region; grey boxes represent C-boxes; black lines represent the regions analysed by ChIP; dark arrow represents the ATG; scale bar=1.5 kb. (B) Quantitative Real-Time PCR of ChIP assay using

chromatin extracted from 35S:BPC1-RFP and wild-type (as a negative control), testing the Region 1, Region 2, Region 3 and pFUS3 (Wu et al., 2019). For the IP, RFP trap was 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 significantly enriched compared with the controls in at least three independent experiments.

4.2.4 Finding new putative interactors of BPCs during gynoecium development

It has been shown that BPCs acts with several partners in order to regulate their target expression (Simonini et al., 2012; Hecker et al., 2015; Xiao, 2017). To gain more insight into the genetic networks BPCs might be part of during gynoecium development, a yeast two hybrid assay was conducted, screening BPCs against a set of several factors, involved in gynoecium development and phytohormones pathways (Francesca Caselli, unpublished results). The analysis identified TEOSINTE BRANCHED1-CYCLOIDEA-PCF15 (TCP15) and HECATE1 (HEC1) as putative interactors; in particular, HEC1 interacts with BPC1 while TCP15 forms heterodimers with all BPCs class I members. A BiFC assay was conducted to confirm the interactions, as showed in Figure 6; controls of the experiments are reported in Supplementary Figure 1 (Caselli et al., in press; Manuscript #2).

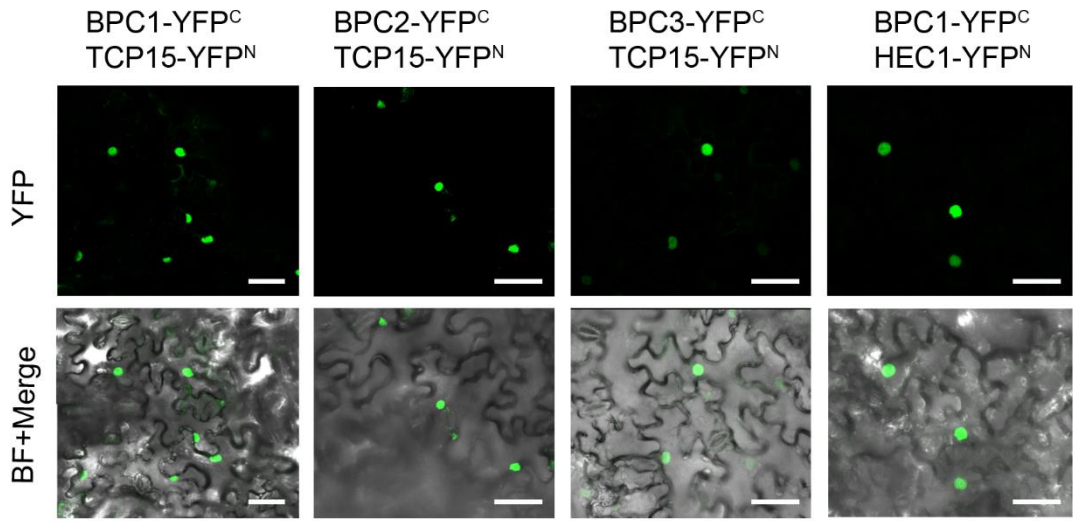


Figure 6. Bi-molecular fluorescence complementation (BiFC) assay. Nicotiana benthamiana epidermis cells were transiently transformed with the indicated YN and YC fusions. In the first and the second rows yellow fluorescence and the merging in the bright field were shown, respectively. Scale bars=30 µm.

TCP15 is a member of the TEOSINTE BRANCHED1-CYCLOIDEA-PCF (TCP) transcription factor family; it has been shown to be involved in gynoecium development (Lucero et al., 2015). In fact,

the constitutive expression of the gene affects carpels fusion. Also, *TCP15* is involved in auxin and cytokinin responses, since a deregulation of the gene or either its constitutive expression led to an alteration in the phytohormones auxin and cytokinin signalling (Uberti-Manassero et al., 2012; Lucero et al., 2015). Being stated that, TCP15 might control gynoecium development by finely regulating auxin homeostasis, which ultimately affects cytokinin responses, as reported in Figure 7 (Lucero et al., 2015).

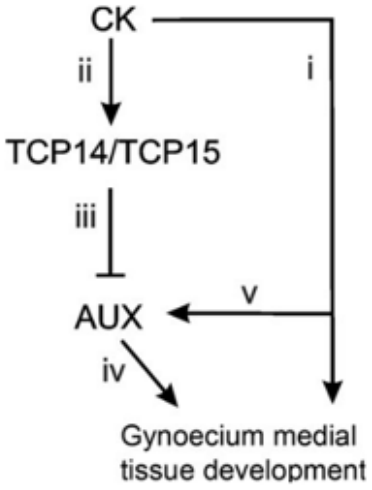


Figure 7. Schematic representation of TCP14/15 activity in gynoecium development (Lucero et al., 2015)

HEC1 is a bHLH protein that acts redundantly with HEC2 and HEC3 to control transmitting tract development (Gremski et al., 2007). Recently, it has been reported that HEC1 interacts with SPT to control carpel fusion, restricting sensitivity to cytokinin in the gynoecium; on top of that, HEC1 is tightly integrated into the auxin-signalling network at the levels of biosynthesis, transport and transcriptional response; thus, it acts as a local modulator of auxin and cytokinin responses to control gynoecium development (Schuster et al., 2015). On the other hand, SPT modulates cytokinin signalling to mediate the activation of auxin biosynthesis and transport genes at the medial domain of the gynoecium (Reyes-Olalde et al., 2017a) These observations corroborate the hypothesis that SPT might be not a BPCs target, but rather an interactor, suggesting a cooperation between bHLH transcription factors and BPCs family members. It would be interesting to check whether BPCs could interact with SPT and other members of this transcription factor family; also, stating that both BPC1 and SPT could interact with HEC1 it is tempting to speculate that they could regulate gynoecium development, acting in the same protein complex.

BPCs have been shown to be associated with phytohormones biosynthesis and perception, as previously reported (Monfared et al., 2011; Simonini and Kater, 2014; Mu et al., 2017b; Shanks et al., 2018). Our interaction analyses suggest that BPCs might cooperate with HEC1-SPT complex and TCP15 for the correct modulation of the two phytohormones balances throughout gynoecium development. The possible involvement of BPCs in auxin- and cytokinin-mediated signalling, will be further analysed by the exploitation of specific auxin and cytokinin reporter lines; in particular, we transformed both *bpc1-2 bpc2 bpc3* and *bpcV* mutants (and wild-type, as control) with the DR5v2 and the TCSn reporter line, that efficiently allow to dissect auxin and cytokinin signalling in vivo, respectively (Chen et al., 2013; Liu and Müller, 2017). The transformant progenies will be selected for the presence of the construct and the reporter lines will be analysed in order to directly investigate whether the drastic phenotype observed in our mutants could be associated to a disruption of phytohormones signalling. In parallel, auxin and cytokinin treatments will be performed on our mutants; it might be possible that an exogenous application of either the two phytohormones could partially rescue the septum and transmitting tract defects of *bpcV*; likewise, we cannot exclude the opposite scenario in which BPCs mutation could affect phytohormones perception, thus leading to an hyper- or either insensitivity to cytokinin or auxin treatments. Shanks et al. (2018) have already observed that BPCs factors have a positive effect on cytokinin sensitivity in roots; thus, this might be observed in inflorescences as well.

In conclusion, a better understanding of the participation of BPCs in phytohormones responses might be the key to unravel their role in septum formation and transmitting tract differentiation.

5. Conclusions and future perspectives

The genetic pathways that determine flower development have been extensively investigated in the last decades, using the model species *Arabidopsis thaliana*. In fact, molecular studies on Arabidopsis are a powerful tool to gain knowledge on flower development and define the molecular mechanisms that precisely drive the correct organ formation; on top of that, the knowledge acquired in the model species can be easily translated in economically relevant plants. Regulation of homeotic genes has been broadly studied in all organisms, mostly for their key roles in organ identity and differentiation. For their importance in establishing organ identity, the expression of homeotic genes is finely orchestrated during development by different mechanisms.

In plant, the five basic floral homeotic gene classes, A, B, C, D and E all encode for members of the MADS-box TF family, apart from AP2. Their expression depends on the activity of a large spectrum of regulatory mechanisms, including epigenetic and transcriptional regulation. In plants, MADS-box genes have a key role in determining the identity of flower meristem and organs and controlling flowering time; moreover, they contribute to the correct formation of embryo, ovule, fruit, leaf, and root. Despite their importance in plant development, many questions related to those mechanisms still require detailed answers. In particular, little is known about how TFs act to orchestrate homeotic gene expression throughout development.

Our work provides new insights into the molecular mechanisms of gene regulation in plant, using the ovule identity gene *STK* as model and focusing on the role of two families of TFs, BPCs and MADS-domain in orchestrating its expression throughout flower development. We assessed a role of BPCs of class II in the regulation of *STK*; our data showed indeed that they act redundantly and synergistically with class I BPCs for the correct modulation of *STK* expression in the flower. Fascinatingly, we found that the PRC1 member LHP1 is a new regulator of *STK*; in fact, we proved that it is directly associate to the locus of the homeotic gene to repress its expression, most likely by recruiting BPCs and SVP to the DNA and mediating the deposition of H3K27me3 repressor mark. Furthermore, we determined the importance of CArG-boxes in the regulation of *STK* spatial and temporal expression, confirming the pivotal role of SVP in repressing the homeotic gene in the first stages of flower development. Our results showed a direct interaction between BPCs of class II and SVP, corroborating the formation of a complex to regulate *STK* expression. Besides, the genome-wide analysis of BPCs and MADS-domain factor binding site locations suggests the possible existence of general mechanism for the regulation of homeotic gene throughout plant

development. With our work we have achieved an advancement of knowledge in the molecular mechanisms that orchestrate homeotic genes expression in plant.

Furthermore, we propose a new role for BPCs in gynoecium development; specifically, we unveil their contribution in septum formation and transmitting tract differentiation. Interestingly, our transcriptomic data pinpointed several mis-regulated genes involved in phytohormones biosynthesis, signaling and transport (i.e. ARFs and ARRs). We therefore assume that defects in the auxin and cytokinin pathways, whose cross-talk is necessary for a successful gynoecium development, partially cause the failure in septum fusion and transmitting tract differentiation, observed in *bpc* quintuple mutant. The phytohormones treatments and the analysis of the auxin and cytokinin reporter lines will help us to answer this question. Also, the deregulation of several cell wall modifier and genes encoding factors associated with transporter and catalytic activity suggest that our mutant might experience defects in cell wall composition that could consequently alter the correct development of the gynoecium. This hypothesis will be further investigated by specific analyses, allowing the staining of the distinct components of this important maternal tissue. Through a transcriptional analysis, we could identify a putative target of BPCs: the bHLH transcription factor *SPT*. Even though the GAGA BINDING FACTORs did not directly bind its promoter region, it could still be possible an "indirect" regulation, supported by the gene expression analysis. The relationship between BPCs and *SPT* will be further unraveled by the analysis of the already generated lines of *bpc* quintuple mutant constitutively expressing the bHLH transcription factor. Furthermore, we found new interactors of BPCs by protein interaction assays; our results supported the hypothesis that the GAGA binding proteins forms a complex with other TFs families (bHLH and TCP transcription factors) to regulate downstream targets, thus orchestrating gynoecium development. A successful formation of the pistil is pivotal for plant reproduction. Being stated that, the knowledge on the molecular mechanisms that led to a correct gynoecium differentiation in Arabidopsis could be successfully applied to crops and other economically relevant plants in order to improve agricultural productivity.

6. Material and methods

For the experiments showed in the Manuscript #1, material and methods are described in the manuscript itself; for the other experiments, material and methods are described below:

Arabidopsis thaliana ecotype Columbia was used in this study; the plants were directly sown on soil and kept under short-day conditions for 2 weeks (22°C, 8 h light and 16 h dark) and then moved to long-day conditions (22°C, 16 h light and 8 h dark). Seeds from the *lhp1, swn-3, fie-2 and clf-1* mutants in Columbia background were obtained from the Nottingham Arabidopsis Stock Centre.

Generation of *seu-1 lug-3* mutants and marker lines

The *pSTK*:GUS *seu-1* and *lug-3* lines were obtained by crossing the *seu-1* and *leu-3* single mutant with the *pSTK*:GUS line (Kooiker et al., 2005; Simonini et al., 2012). The *seu-1 leu-3 pSTK:GUS* line was obtained by crossing the lines described above. The *pSVP:SVP*-GFP *pSTK_Gam5:GUS svp bpc1-2 bpc2 bpc3* line was obtained by crossing the *pSVP:SVP*-GFP *pSTK:Gam5*:GUS line (Simonini et al., 2012) with *bpc1-2 bpc2 bpc3* line (Monfared et al., 2011).

Generation of *35:SPT* construct and plant transformation

The *SPT* coding sequence was first cloned into pDONR221 (Life Technologies) and subsequently transferred to pH2WG7, purchased from the Flanders Interuniversity Institute for Biotechnology (Gent, Belgium). Wild-type, *bpc1-2 bpc2 bpc3* and *bpcV* plants were transformed using the Agrobacterium tumefaciens–mediated floral dip method (Clough and Bent, 1998). Transformant plants were sown on MS medium and selected by hygromycin (20 mg/L); presence of the construct was assessed by genotyping and analysis of *SPT* expression.

Imaging and microscopy

Images of plants were acquired using a Canon EOS 6D camera whereas images of siliques were taken using a Leica® MZ 6 stereomicroscope.

C-boxes prediction and promoter analysis

The putative promoter region (5 Kb upstream the ATG) of the candidate genes were analysed in terms of C-boxes presence using the following online tolls: Match (http://www.gene-regulation.com) e Fuzznucc (http://www.bioinformatics.nl/cgi-bin/emboss/fuzznuc). Genes were considered good candidates if they have at least two C-boxes.

ChIP assay

ChIP assay was performed as described by Gregis et al. (2009) using for BPC1-RFP an anti-RFP V_HH coupled to magnetic agarose beads RFP-trap_MA® (Chromotek). Quantitative Real-time PCR assays were conducted 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 in the table. ChIP-quantitative Real-Time PCR experiments and relative enrichments were calculated as reported by Gregis et al, (2009) using the primers in the Table 3; the promoter region of *FUS3* was used as positive control (Wu et al., 2019).

Gene expression analysis

Quantitative Real-time PCR experiments were performed using cDNA obtained from inflorescences. Total RNA was extracted using lithium chloride. The Ambion TURBO DNA-free DNase kit was used to remove genomic DNA contaminations, according to the manufacturer's instructions (http://www.ambion.com/). The ImProm-IITM reverse transcription system (Promega) was used to retro-transcribe the treated RNA. Transcripts were detected using a Sybr Green Assay (iQ SYBR Green Supermix; Bio-Rad) using *UBIQUITIN* as a reference gene. Assays were done in in triplicate using a Bio-Rad iCycler iQ Optical System (software version 3.0a). The enrichments were calculated normalising the amount of mRNA against housekeeping gene fragments. The expression of different genes was analysed using specific oligonucleotides primers (listed in the Table 3).

RNA extraction, sequencing for RNA-Seq and computational analyses

Total RNA was extracted from three biological replicates (1 gr) from both wild-type and *bpcV* mutant inflorescences till stage 12 before fertilization, using the RNeasy Plant Mini kit according to the manufacturer's instructions. RNA concentrations and integrity were determined using Qubit Fluorometer and the Qubit™ RNA XR Assay Kit (ThermoFisher Scientific). Sequencing libraries were

prepared using the NEBNext Ultra II Directional RNA library Prep Kit for illumina (NEB) according to the manufacture's instruction and sequenced on HiSeq Illumina platform. Reads were mapped on the reference Arabidopsis thaliana transcriptome (TAIR, version 10) using the bowtie2 program (Langmead and Salzberg, 2012). Estimation of gene expression levels was performed using RSEM (Li and Dewey, 2011). Identification of differentially expressed genes was performed by the quasi-likelihood F-test as implemented by edgeR (Robinson et al., 2010). To gain insight into the biological processes associated with BPCs factors, we determined which GO annotation terms were over-represented, in the deregulated gene lists (Dataset 1). Gene set enrichment analysis was performed with the agriGOv2 database (Tian et al., 2017) using the Singular Enrichment Analysis (SEA). Annotation of selected DAP-seq peaks was performed by the means of the annotatePeaks program from the Homer suite (Heinz et al., 2010) using the reference TAIR10 annotation.

In- situ hybridisation assay

Arabidopsis flowers were collected, fixed and embedded in paraffin as described by Huijser et al. (1992). Plant tissue sections were probed with a *SPT* antisense RNA designed prior to my arrival in the laboratory by Irma Roig Villanova. Hybridisation and immunological detection were executed as described previously by Coen et al. (1990).

Protein-protein interactions

The yeast two-hybrid assays were performed in the yeast strains PJ69-4A and PJ69-4α. The coding sequences of *TCP15, BPC1, BPC2, BPC2* and *HEC1* were cloned in the pDEST32 (bait vector, BD; Invitrogen) and pDEST22 (prey vector, AD; Invitrogen) Gateway vectors. The bait constructs were tested for autoactivation on selective yeast synthetic dropout medium lacking Leu, Trp and His supplemented with 1, 3, 5, 10 or 15 mM of 3-aminotriazole, in order to set the screening conditions. After mating, colonies were plated on the proper selective media and allowed to grow for 5 days at 20°. For BiFC, the coding sequences of *TCP15* and *HEC1* were first cloned into pDONR207 (Life Technologies) and subsequently transferred to the pYFPN43 and pYFPC43 vectors by Gateway recombination; while the BiFC constructs for *BPC1*, *BPC2* and *BPC3* by (Simonini et al., 2012) were used; the previously described formation of REM34-REM35 heterodimers was used as positive control, whereas REM34-REM34 combination was used as negative control (Caselli et al., in press; Manuscript #2). BiFC assays were performed injecting Agrobacterium expressing viral

suppressor p19/experimental constructs as described by Belda-Palazón et al. (2012). The abaxial surfaces of infiltrated tobacco (*Nicotiana benthamiana*) leaves were imaged 3 days after inoculation.

Morphological analysis

For the analysis of the siliques, 30 siliques per genotype (wild-type, *bpc1-2 bpc2 bpc3* and *bpcV*) were sectioned and analysed for the presence of the septum at the stereomicroscope.

Samples preparation, embedding and sectioning

Inflorescences with flowers from stages 6 to 12 (Smyth et al., 1990) were collected in 50 mL falcon tubes containing 10 mL of FAA. Tubes were placed in a vacuum desiccator and vacuum was applied for 20 min. Afterward, samples were incubated at 4°C overnight. The tissue was rehydrated passing through a series of ethanol solutions (70, 50 and 30% Ethanol, 10 minutes each, two times per solution) and then transferred to ddH_20. The samples were pre-embedded in 1.5% agarose gel; this step facilitates the orientation of the inflorescences in the Teflon blocks during the polymerization phase, described below. The agarose blocks were then dehydrated passing through a series of ethanol solutions (30, 50, 70% ethanol, 2 washes of 10 minutes per solution; and 85, 95, 100% ethanol, 1 wash of 30 minutes per solution) and kept overnight at 4°C. The samples were successively embedded in acrylate according to manufacturer instructions; we used the reagent Technovit (Heraeus Kulzer, Germany). Blocks were sectioned using a microtome; 8-12 µm thick sections were then placed to glass slides and air dried. The slides were placed in a coplin jar filled with 0.5% Alcian blue 8GX Solution (dissolved in water, pH adjusted to 3.1 with Acetic Acid) for 25'. Slides were then washed with tap water, until water appears clear. The samples were transferred into 0.5% Neutral Red Solution (dissolved in water) for 5' and once again washed with distilled water. The slides were removed from the coplin jar, air dried and mounted with BioMount prior to microscope analysis.

Bibliography

Aida, M. (1997). Genes Involved in Organ Separation in Arabidopsis: An Analysis of the cup-shaped cotyledon Mutant. *Plant Cell Online* **9**:841–857.

Alvarez-Buylla, E. R., Pelaz, S., Liljegren, S. J., Gold, S. E., Burgeff, C., Ditta, G. S., Ribas de Pouplana, L., Martinez-Castilla, L., and Yanofsky, M. F. (2000). An ancestral MADS-box gene duplication occurred before the divergence of plants and animals. *Proc. Natl. Acad. Sci.* **97**:5328–5333.

Alvarez-Buylla, E. R., Benítez, M., Corvera-Poiré, A., Chaos Cador, A., de Folter, S., Gamboa de Buen, A., Garay-Arroyo, A., García-Ponce, B., Jaimes-Miranda, F., Pérez-Ruiz, R. V, et al. (2010). Flower development. *Arab. B.* **8**:e0127.

Alvarez, J., and Smyth, D. R. (1999). CRABS CLAW and SPATULA, two Arabidopsis genes that control carpel development in parallel with AGAMOUS. *Development* **126**:2377–86.

Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. *Nature* **408**:796–815.

Bao, X., Franks, R. G., Levin, J. Z., and Liu, Z. (2004). Repression of *AGAMOUS* by *BELLRINGER* in Floral and Inflorescence Meristems. *Plant Cell* **16**:1478–1489.

Bartlett, A., O'Malley, R. C., Huang, S. C., Galli, M., Nery, J. R., Gallavotti, A., and Ecker, J. R. (2017). Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.* **12**:1659–1672.

Belda-Palazón, B., Ruiz, L., Martí, E., Tárraga, S., Tiburcio, A. F., Culiáñez, F., Farràs, R., Carrasco, P., and Ferrando, A. (2012). Aminopropyltransferases Involved in Polyamine Biosynthesis Localize Preferentially in the Nucleus of Plant Cells. *PLoS One* **7**:e46907.

Berger, N., and Dubreucq, B. (2012). Evolution goes GAGA: GAGA binding proteins across kingdoms. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1819**:863–868.

Berger, N., Dubreucq, B., Roudier, F., Dubos, C., and Lepiniec, L. (2011). Transcriptional regulation of Arabidopsis LEAFY COTYLEDON2 involves RLE, a cis-element that regulates trimethylation of histone H3 at lysine-27. *Plant Cell* **23**:4065–78.

Bouyer, D., Roudier, F., Heese, M., Andersen, E. D., Gey, D., Nowack, M. K., Goodrich, J., Renou, J.-P., Grini, P. E., Colot, V., et al. (2011). Polycomb Repressive Complex 2 Controls the Embryo-to-Seedling Phase Transition. *PLoS Genet.* **7**:e1002014.

Bowman, J. L., Drews, G. N., and Meyerowitz, E. M. (1991). Expression of the Arabidopsis floral homeotic gene AGAMOUS is restricted to specific cell types late in flower development. *Plant Cell* **3**:749–58.

Bowman, J. L., Alvarez, J., Weigel, D., Meyerowitz, E. M., and Smyth, D. R. (1993). Control of flower development in Arabidopsis thaliana by APETALA1 and interacting genes. *Development* **119**.

Boyes, D. C., Zayed, A. M., Ascenzi, R., Mccaskill, A. J., Hoffman, N. E., Davis, K. R., and Görlach, J. (2001). <Arabidopsis_Growth_Stage.Pdf> **13**:1499–1510.

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**:2544–56.

Busch, M. A., Bomblies, K., and Weigel, D. (1999). Activation of a floral homeotic gene in Arabidopsis. *Science (80-.).* **285**:585–587.

Calonje, M. (2014). PRC1 Marks the Difference in Plant PcG Repression. *Mol. Plant* **7**:459–471.

Calonje, M., Sanchez, R., Chen, L., and Sung, Z. R. (2008). EMBRYONIC FLOWER1 Participates in Polycomb Group–Mediated *AG* Gene Silencing in *Arabidopsis*. *Plant Cell* **20**:277–291.

Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.-H., Sung, Z. R., and Goodrich, J. (2004). Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. *Development* **131**:5263–76.

Chen, Y., Yordanov, Y. S., Ma, C., Strauss, S., and Busov, V. B. (2013). DR5 as a reporter system to study auxin response in Populus. *Plant Cell Rep.* **32**:453–463.

Chen, D., Yan, W., Fu, L.-Y., and Kaufmann, K. (2018). Architecture of gene regulatory networks controlling flower development in Arabidopsis thaliana. *Nat. Commun.* **9**:4534.

Clapier, C. R., and Cairns, B. R. (2009). The Biology of Chromatin Remodeling Complexes. *Annu. Rev. Biochem.* **78**:273–304.

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.

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

Coen, E. S., Romero, J. M., Doyle, S., Elliott, R., Murphy, G., and Carpenter, R. (1990). floricaula: a homeotic gene required for flower development in antirrhinum majus. *Cell* **63**:1311–22.

Colombo, L. (1995). The Petunia MADS Box Gene FBP11 Determines Ovule Identity. *Plant Cell Online* **7**:1859–1868.

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

Crawford, B. C. W., and Yanofsky, M. F. (2011). HALF FILLED promotes reproductive tract development and fertilization efficiency in Arabidopsis thaliana. *Development* **138**:2999–3009.

Crawford, B. C. W., Ditta, G., and Yanofsky, M. F. (2007). The NTT Gene Is Required for Transmitting-Tract Development in Carpels of Arabidopsis thaliana. *Curr. Biol.* **17**:1101–1108.

Deb, J., Bland, H. M., and Østergaard, L. (2018). Developmental cartography: coordination via hormonal and genetic interactions during gynoecium formation. *Curr. Opin. Plant Biol.* **41**:54–60.

Derkacheva, M., Steinbach, Y., Wildhaber, T., Mozgová, I., Mahrez, W., Nanni, P., Bischof, S., Gruissem, W., and Hennig, L. (2013). Arabidopsis MSI1 connects LHP1 to PRC2 complexes. *EMBO J.* **32**:2073–2085.

Doughty, J., Aljabri, M., and Scott, R. J. (2014). Flavonoids and the regulation of seed size in *Arabidopsis*. *Biochem. Soc. Trans.* **42**:364–369.

Drews, G. N., Bowman, J. L., and Meyerowitz, E. M. (1991). Negative regulation of the Arabidopsis homeotic gene AGAMOUS by the APETALA2 product. *Cell* **65**:991–1002.

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 J.* **18**:5370–5379.

Exner, V., Aichinger, E., Shu, H., Wildhaber, T., Alfarano, P., Caflisch, A., Gruissem, W., and Ko, C. (2009). The Chromodomain of LIKE HETEROCHROMATIN PROTEIN 1 Is Essential for H3K27me3 Binding and Function during Arabidopsis Development **4**:1–10.

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.

Farkas, G., Gausz, J., Galloni, M., Reuter, G., Gyurkovics, H., and Karch, F. (1994). The Trithorax-like gene encodes the Drosophila GAGA factor. *Nature* **371**:806–808.

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**.

Fletcher, J. C. (2002). SHOOT AND FLORAL MERISTEM MAINTENANCE IN ARABIDOPSIS. *Annu. Rev. Plant Biol.* **53**:45–66.

Fobert, P. R., Coen, E. S., Murphy, G. J., and Doonan, J. H. (1994). Patterns of cell division revealed by transcriptional regulation of genes during the cell cycle in plants. *EMBO J.* **13**:616–24.

Goodrich, J. J. (1997). A Polycomb-group gene regulates homeotic gene expression in Arabidopsis **38616**:2–9.

Goto, K., and Meyerowitz, E. M. (1994). Function and regulation of the Arabidopsis floral homeotic gene PISTILLATA. *Genes Dev.* **8**:1548–60.

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 ONLINE* **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.* **56**:891–902.

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.

Gregis, V., Andrés, F., Sessa, A., Guerra, R. F., Simonini, S., Mateos, J. L., Torti, S., Zambelli, F., Prazzoli, G. M., Bjerkan, K. N., et al. (2013). Identification of pathways directly regulated by SHORT VEGETATIVE PHASE during vegetative and reproductive development in Arabidopsis. *Genome Biol.* **14**:R56.

Gremski, K., Ditta, G., and Yanofsky, M. F. (2007). The HECATE genes regulate female reproductive tract development in Arabidopsis thaliana. *Development* **134**:3593–3601.

Groszmann, M., Bylstra, Y., Lampugnani, E. R., and Smyth, D. R. (2010). Regulation of tissue-specific expression of SPATULA, a bHLH gene involved in carpel development, seedling germination, and lateral organ growth in Arabidopsis. *J. Exp. Bot.* **61**:1495–1508.

Groszmann, M., Paicu, T., Alvarez, J. P., Swain, S. M., and Smyth, D. R. (2011). SPATULA and ALCATRAZ, are partially redundant, functionally diverging bHLH genes required for Arabidopsis gynoecium and fruit development. *Plant J.* **68**:816–829.

Gustafson-Brown, C., Savidge, B., and Yanofsky, M. F. (1994). Regulation of the arabidopsis floral homeotic gene APETALA1. *Cell* **76**:131–43.

Hartmann, U., Ho, S., Nettesheim, K., Wisman, E., Saedler, H., and Huijser, P. (2000). Molecular cloning of SVP: a negative regulator of the floral transition in Arabidopsis. *plant J.* **21**:351–360.

Hecker, A., Brand, L. H., Peter, S., Simoncello, N., Kilian, J., Harter, K., Gaudin, V., and Wanke, D. (2015). The Arabidopsis GAGA-Binding Factor BASIC PENTACYSTEINE6 Recruits the POLYCOMB-REPRESSIVE COMPLEX1 Component LIKE HETEROCHROMATIN PROTEIN1 to GAGA DNA Motifs.

Plant Physiol. Advance Access published 2015, doi:10.1104/pp.15.00409.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**:576–589.

Heisler, M. G. B., Atkinson, A., Bylstra, Y. H., Walsh, R., and Smyth, D. R. (2001). SPATULA , a gene that controls development of carpel margin tissues in Arabidopsis , encodes a bHLH protein **1098**:1089–1098.

Herrera-Ubaldo, H., and de Folter, S. (2018). Exploring Cell Wall Composition and Modifications During the Development of the Gynoecium Medial Domain in Arabidopsis. *Front. Plant Sci.* **9**:454.

Herrera-Ubaldo, H., Campos, S. E., Luna-García, V., Zúñiga-Mayo, V. M., Armas-Caballero, G., DeLuna, A., Marsch-Martínez, N., and Folter, S. de (2018). An interaction map of transcription factors controlling gynoecium development in Arabidopsis. *bioRxiv* Advance Access published December 20, 2018, doi:10.1101/500736.

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.

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

Horard, B., Tatout, C., Poux, S., and Pirrotta, V. (2000). Structure of a polycomb response element and in vitro binding of polycomb group complexes containing GAGA factor. *Mol. Cell. Biol.* **20**:3187–97.

Hugouvieux, V., and Zubieta, C. (2018). MADS transcription factors cooperate: complexities of complex formation. *J. Exp. Bot.* **69**:1821–1823.

Immink, R. G. H., Gadella, T. W. J., Ferrario, S., Busscher, M., and Angenent, G. C. (2002). Analysis of MADS box protein-protein interactions in living plant cells. *Proc. Natl. Acad. Sci.* **99**:2416–2421.

Immink, R. G. H., Kaufmann, K., and Angenent, G. C. (2010). The 'ABC' of MADS domain protein behaviour and interactions. *Semin. Cell Dev. Biol.* **21**:87–93.

Immink, R. G. H., Posé, D., Ferrario, S., Ott, F., Kaufmann, K., Valentim, F. L., de Folter, S., van der Wal, F., van Dijk, A. D. J., Schmid, M., et al. (2012). Characterization of SOC1's Central Role in Flowering by the Identification of Its Upstream and Downstream Regulators. *Plant Physiol.* **160**:433–449.

Ishida, T., Aida, M., Takada, S., and Tasaka, M. (2000). Involvement of CUP-SHAPED COTYLEDON genes in gynoecium and ovule development in Arabidopsis thaliana. *Plant Cell Physiol.* **41**:60–67.

Kamiuchi, Y., Yamamoto, K., Furutani, M., Tasaka, M., and Aida, M. (2014). The CUC1 and CUC2 genes promote carpel margin meristem formation during Arabidopsis gynoecium development. *Front. Plant Sci.* **5**.

Khan, S.-A., Li, M.-Z., Wang, S.-M., and Yin, H.-J. (2018). Revisiting the Role of Plant Transcription Factors in the Battle against Abiotic Stress. *Int. J. Mol. Sci.* **19**:1634.

Kim, D., and Sung, S. (2014). Polycomb-Mediated Gene Silencing in Arabidopsis thaliana **37**:841–850.

Kinoshita, T., Harada, J. J., Goldberg, R. B., and Fischer, R. L. (2001). Polycomb repression of flowering during early plant development. *Proc. Natl. Acad. Sci. U. S. A.* **98**:14156–61.

Kooiker, M., Airoldi, C. A., Losa, A., Manzotti, P. S., Finzi, L., and Kater, M. M. (2005). BASIC PENTACYSTEINE1 , a GA Binding Protein That Induces Conformational Changes in the Regulatory Region of the Homeotic Arabidopsis Gene SEEDSTICK **17**:722–729.

Lafos, M., Kroll, P., Hohenstatt, M. L., Thorpe, F. L., Clarenz, O., and Schubert, D. (2011). Dynamic Regulation of H3K27 Trimethylation during Arabidopsis Differentiation. *PLoS Genet.* **7**:e1002040.

Lai, X., Daher, H., Galien, A., Hugouvieux, V., and Zubieta, C. (2019). Structural Basis for Plant MADS Transcription Factor Oligomerization. *Comput. Struct. Biotechnol. J.* **17**:946–953.

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**:357–359.

Larsson, A. S., Landberg, K., and Meeks-Wagner, D. R. (1998). The TERMINAL FLOWER2 (TFL2) gene controls the reproductive transition and meristem identity in Arabidopsis thaliana. *Genetics* **149**:597–605.

Lee, J., and Lee, I. (2010). Regulation and function of SOC1, a flowering pathway integrator. *J. Exp. Bot.* **61**:2247–2254.

Lenhard, M., and Laux, T. (1999). Shoot meristem formation and maintenance. *Curr. Opin. Plant Biol.* **2**:44–50.

Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. *Nature* **276**:565–570.

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

Li, C., Chen, C., Gao, L., Yang, S., Nguyen, V., Shi, X., Siminovitch, K., Kohalmi, S. E., Huang, S., Wu, K., et al. (2015). The Arabidopsis SWI2/SNF2 Chromatin Remodeler BRAHMA Regulates Polycomb Function during Vegetative Development and Directly Activates the Flowering Repressor Gene SVP. *PLOS Genet.* **11**:e1004944.

Liu, Z., and Meyerowitz, E. M. (1995). LEUNIG regulates AGAMOUS expression in Arabidopsis flowers. *Development* **121**:975–91.

Liu, J., and Müller, B. (2017). Imaging TCSn::GFP, a Synthetic Cytokinin Reporter, in Arabidopsis thaliana. In *Methods in molecular biology (Clifton, N.J.)*, pp. 81–90.

Liu, C., Xi, W., Shen, L., Tan, C., and Yu, H. (2009). Regulation of Floral Patterning by Flowering Time Genes. *Dev. Cell* **16**:711–722.

Lodha, M., Marco, C. F., and Timmermans, M. C. P. (2013). The ASYMMETRIC LEAVES complex maintains repression of KNOX homeobox genes via direct recruitment of Polycomb-repressive complex2. *Genes Dev.* **27**:596–601.

Lucero, L. E., Uberti-Manassero, N. G., Arce, A. L., Colombatti, F., Alemano, S. G., and Gonzalez, D. H. (2015). TCP15 modulates cytokinin and auxin responses during gynoecium development in Arabidopsis. *Plant J.* **84**:267–282.

Luo, J., Zhou, J.-J., and Zhang, J.-Z. (2018). Aux/IAA Gene Family in Plants: Molecular Structure, Regulation, and Function. *Int. J. Mol. Sci.* **19**.

Makarevich, G., Leroy, O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U., and Köhler, C. (2006). Different Polycomb group complexes regulate common target genes in Arabidopsis. *EMBO Rep.* **7**:947–952.

Malley, R. C. O., Carol, S., Song, L., Galli, M., Ecker, J. R., Malley, R. C. O., Huang, S. C., Song, L., Lewsey, M. G., Bartlett, A., et al. (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape Resource Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**:1280–1292.

McGonigle, B., Bouhidel, K., and Irish, V. F. (1996). Nuclear localization of the Arabidopsis APETALA3 and PISTILLATA homeotic gene products depends on their simultaneous expression. *Genes Dev.*

10:1812–1821.

Meister, R. J., Williams, L. A., Monfared, M. M., Gallagher, T. L., Kraft, E. A., Nelson, C. G., and Gasser, C. S. (2004). Definition and interactions of a positive regulatory element of the *Arabidopsis INNER NO OUTER* promoter. *Plant J.* **37**:426–438.

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.

Monfared, M. M., Simon, M. K., Meister, R. J., Roig-Villanova, I., Kooiker, M., Colombo, L., Fletcher, J. C., and Gasser, C. S. (2011). Overlapping and antagonistic activities of BASIC PENTACYSTEINE genes affect a range of developmental processes in Arabidopsis. *Plant J.* **66**:1020–1031.

Mu, Y., Liu, Y., Bai, L., Li, S., He, C., Yan, Y., Yu, X., and Li, Y. (2017a). Cucumber CsBPCs Regulate the Expression of CsABI3 during Seed Germination. *Front. Plant Sci.* **8**:459.

Mu, Y., Zou, M., Sun, X., He, B., Xu, X., Liu, Y., Zhang, L., and Chi, W. (2017b). BASIC PENTACYSTEINE proteins repress Abscisic Acid INSENSITIVE 4 expression via direct recruitment of the polycomb-repressive complex 2 in arabidopsis root development. *Plant Cell Physiol.* **58**:607–621.

Müller, C. J., Larsson, E., Spíchal, L., and Sundberg, E. (2017). Cytokinin-Auxin Crosstalk in the Gynoecial Primordium Ensures Correct Domain Patterning. *Plant Physiol.* **175**:1144–1157.

Murai, K. (2013). Homeotic Genes and the ABCDE Model for Floral Organ Formation in Wheat. *Plants (Basel, Switzerland)* **2**:379–95.

Nahar, M. A. U., Ishida, T., Smyth, D. R., Tasaka, M., and Aida, M. (2012). Interactions of CUP-SHAPED COTYLEDON and SPATULA genes control carpel margin development in arabidopsis thaliana. *Plant Cell Physiol.* **53**:1134–1143.

Oliva, M., Butenko, Y., Hsieh, T.-F., Hakim, O., Katz, A., Smorodinsky, N. I., Michaeli, D., Fischer, R. L., and Ohad, N. (2016). FIE, a nuclear PRC2 protein, forms cytoplasmic complexes in *Arabidopsis thaliana*. *J. Exp. Bot.* **67**:6111–6123.

Orphanides, G., Wu, W. H., Lane, W. S., Hampsey, M., and Reinberg, D. (1999). The chromatin-specific transcription elongation factor FACT comprises human SPT16 and SSRP1 proteins. *Nature* **400**:284–288.

Parenicová, 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 SEPALLATTA MADS-box genes. *Nature* **405**:200–203.

Pelaz, S., Gustafson-Brown, C., Kohalmi, S. E., Crosby, W. L., and Yanofsky, M. F. (2002). APETALA1 and SEPALLATA3 interact to promote flower development. *Plant J.* **26**:385–394.

Pien, S., and Grossniklaus, U. (2007). Polycomb group and trithorax group proteins in Arabidopsis. *Biochim. Biophys. Acta - Gene Struct. Expr.* **1769**:375–382.

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.

Reyes-Olalde, J. I., Zuñiga-Mayo, V. M., Chávez Montes, R. A., Marsch-Martínez, N., and de Folter, S. (2013). Inside the gynoecium: At the carpel margin. *Trends Plant Sci.* **18**:644–655.

Reyes-Olalde, J. I., Zúñiga-Mayo, V. M., Serwatowska, J., Chavez Montes, R. A., Lozano-Sotomayor, P., Herrera-Ubaldo, H., Gonzalez-Aguilera, K. L., Ballester, P., Ripoll, J. J., Ezquer, I., et al. (2017a).

The bHLH transcription factor SPATULA enables cytokinin signaling, and both activate auxin biosynthesis and transport genes at the medial domain of the gynoecium. *PLOS Genet.* **13**:e1006726.

Reyes-Olalde, J. I., Zúñiga-Mayo, V. M., Marsch-Martínez, N., and de Folter, S. (2017b). Synergistic relationship between auxin and cytokinin in the ovary and the participation of the transcription factor SPATULA. *Plant Signal. Behav.* **12**:e1376158.

Ringrose, L., and Paro, R. (2007). Polycomb/Trithorax response elements and epigenetic memory of cell identity. *Development* **134**:223–232.

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* **26**:139–40.

Roeder, A. H. K., and Yanofsky, M. F. (2006). *Fruit Development in Arabidopsis*.

Roscoe, T. J., Vaissayre, V., Paszkiewicz, G., Clavijo, F., Kelemen, Z., Michaud, C., Lepiniec, L., Dubreucq, B., Zhou, D.-X., and Devic, M. (2019). Regulation of *FUSCA3* Expression During Seed Development in Arabidopsis. *Plant Cell Physiol.* **60**:476–487.

Sangwan, I. (2002). Identification of a Soybean Protein That Interacts with GAGA Element Dinucleotide Repeat DNA. *Plant Physiol.* **129**:1788–1794.

Santi, L., Wang, Y., Stile, M. R., Berendzen, K., Wanke, D., Roig, C., Pozzi, C., Müller, K., Müller, J., Rohde, W., et al. (2003). The GA octodinucleotide repeat binding factor BBR participates in the transcriptional regulation of the homeobox gene *Bkn3*. *Plant J.* **34**:813–826.

Sarojam Rajani and Venkatesan Sundaresan (2001). The Arabidopsis myc / bHLH gene ALCATRAZ enables cell separation in fruit dehiscence Sarojam Rajani * † and Venkatesan Sundaresan * ‡. *Curr. Biol.* **11**:1914–1922.

Schaller, G. E., Bishopp, A., and Kieber, J. J. (2015). The Yin-Yang of Hormones: Cytokinin and Auxin Interactions in Plant Development. *Plant Cell* **27**:44–63.

Schatlowski, N., Creasey, K., Goodrich, J., and Schubert, D. (2008). Keeping plants in shape: Polycomb-group genes and histone methylation. *Semin. Cell Dev. Biol.* **19**:547–553.

Schruff, M. C., Spielman, M., Tiwari, S., Adams, S., Fenby, N., and Scott, R. J. (2006). The AUXIN RESPONSE FACTOR 2 gene of Arabidopsis links auxin signalling, cell division, and the size of seeds and other organs. *Development* **133**:251–61.

Schuster, C., Gaillochet, C., and Lohmann, J. U. (2015). Arabidopsis HECATE genes function in phytohormone control during gynoecium development. *Development* **142**:3343–50.

Schwarz-sommer, Z., Hue, I., Huijser, P., Flor, P. J., Hansen, R., Tetens, F., Lonnig, W., Saedler, H., and Sommer, H. (1992). Characterization of the Antirrhinum floral homeotic MADS-box gene deficiens : evidence for DNA binding and autoregulation of its persistent expression throughout flower development **1**:251–263.

Seymour, G. B., Østergaard, L., Chapman, N. H., Knapp, S., and Martin, C. (2013). Fruit Development and Ripening. *Annu. Rev. Plant Biol.* **64**:219–241.

Shanks, C. M., Hecker, A., Cheng, C.-Y., Brand, L., Collani, S., Schmid, M., Schaller, G. E., Wanke, D., Harter, K., and Kieber, J. J. (2018). Role of *BASIC PENTACYSTEINE* transcription factors in a subset of cytokinin signaling responses. *Plant J.* **95**:458–473.

Simonini, S., and Kater, M. M. (2014). Class I BASIC PENTACYSTEINE factors regulate HOMEOBOX genes involved in meristem size maintenance. *J. Exp. Bot.* **65**:1455–1465.

Simonini, S., Roig-Villanova, I., Gregis, V., Colombo, B., Colombo, L., and Kater, M. M. (2012). BASIC

PENTACYSTEINE Proteins Mediate MADS Domain Complex Binding to the DNA for Tissue-Specific Expression of Target Genes in *Arabidopsis*. *Plant Cell* **24**:4163–4172.

Smaczniak, C., Immink, R. G. H., Muiño, J. M., Blanvillain, R., and Busscher, M. (2012a). Characterization of MADS-domain transcription factor complexes in Arabidopsis fl ower development Advance Access published 2012, doi:10.1073/pnas.1112871109/-/DCSupplemental.www.pnas.org/cgi/doi/10.1073/pnas.1112871109.

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

Smyth, D. R., Bowman, J. L., and Meyerowitz, E. M. (1990). Early flower development in Arabidopsis. *Plant Cell* **2**:755–67.

Soeller, W. C., Oh, C. E., and Kornberg, T. B. (1993). Isolation of cDNAs encoding the Drosophila GAGA transcription factor. *Mol. Cell. Biol.* **13**:7961–7970.

Steensel, B. Van, Delrow, J., and Bussemaker, H. J. (2002). Genomewide analysis of Drosophila GAGA factor target genes reveals context-dependent DNA binding Advance Access published 2002.

Swarup, R., Kargul, J., Marchant, A., Zadik, D., Rahman, A., Mills, R., Yemm, A., May, S., Williams, L., Millner, P., et al. (2004). Structure-function analysis of the presumptive Arabidopsis auxin permease AUX1. *Plant Cell* **16**:3069–83.

Swidzinski, J. A., Sweetlove, L. J., and Leaver, C. J. (2002). A custom microarray analysis of gene expression during programmed cell death in Arabidopsis thaliana. *Plant J.* **30**:431–46.

Theißen, G., and Saedler, H. (2001). news and views. *Plant Biol.* Advance Access published 2001.

Theißen, G., Melzer, R., and Rümpler, F. (2016). MADS-domain transcription factors and the floral quartet model of flower development: linking plant development and evolution. *Development* **143**:3259–3271.

Theune, M. L., Bloss, U., Brand, L. H., Ladwig, F., and Wanke, D. (2019). Phylogenetic Analyses and GAGA-Motif Binding Studies of BBR/BPC Proteins Lend to Clues in GAGA-Motif Recognition and a Regulatory Role in Brassinosteroid Signaling. *Front. Plant Sci.* **10**:466.

Tian, T., Liu, Y., Yan, H., You, Q., Yi, X., Du, Z., Xu, W., and Su, Z. (2017). agriGO v2.0: a GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **45**:W122–W129.

To, J. P. C., Deruère, J., Maxwell, B. B., Morris, V. F., Hutchison, C. E., Ferreira, F. J., Schaller, G. E., and Kieber, J. J. (2007). Cytokinin Regulates Type-A *Arabidopsis* Response Regulator Activity and Protein Stability via Two-Component Phosphorelay. *Plant Cell* **19**:3901–3914.

Tung, C.-W., Dwyer, K. G., Nasrallah, M. E., and Nasrallah, J. B. (2005). Genome-Wide Identification of Genes Expressed in Arabidopsis Pistils Specifically along the Path of Pollen Tube Growth. *Plant Physiol.* **138**:977–989.

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.-L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R. A., Coupland, G., and Colot, V. (2007). Arabidopsis TFL2/LHP1 Specifically Associates with Genes Marked by Trimethylation of Histone H3 Lysine 27. *PLoS Genet.* **3**:e86.

Uberti-Manassero, N. G., Lucero, L. E., Viola, I. L., Vegetti, A. C., and Gonzalez, D. H. (2012). The class I protein AtTCP15 modulates plant development through a pathway that overlaps with the one affected by CIN-like TCP proteins. *J. Exp. Bot.* **63**:809–823.

Ueguchi, C., Koizumi, H., Suzuki, T., and Mizuno, T. (2001). Novel Family of Sensor Histidine Kinase Genes in Arabidopsis thaliana. *Plant Cell Physiol.* **42**:231–235.

Veluchamy, A., Jégu, T., Ariel, F., Latrasse, D., Mariappan, K. G., Kim, S.-K., Crespi, M., Hirt, H., Bergounioux, C., Raynaud, C., et al. (2016). LHP1 Regulates H3K27me3 Spreading and Shapes the Three-Dimensional Conformation of the Arabidopsis Genome. *PLoS One* **11**:e0158936.

Wang, H., Liu, C., Cheng, J., Liu, J., Zhang, L., and He, C. (2016). Arabidopsis Flower and Embryo Developmental Genes are Repressed in Seedlings by Different Combinations of Polycomb Group Proteins in Association with Distinct Sets of Cis-regulatory Elements Advance Access published 2016, doi:10.1371/journal.pgen.1005771.

Wanke, D., Hohenstatt, M. L., Dynowski, M., Bloss, U., Hecker, A., Elgass, K., Hummel, S., Hahn, A., Caesar, K., Schleifenbaum, F., et al. (2011). Alanine zipper-like coiled-coil domains are necessary for homotypic dimerization of plant GAGA-factors in the nucleus and nucleolus. *PLoS One* **6**:e16070.

Weigel, D., and Jürgens, G. (2002). Stem cells that make stems. *Nature* **415**:751–754.

Wu, J., Petrella, R., Dowhanik, S., Gregis, V., and Gazzarrini, S. (2019). Spatiotemporal restriction of FUSCA3 expression by class I BPC promotes ovule development and coordinates embryo and endosperm growth. *bioRxiv* Advance Access published April 19, 2019, doi:10.1101/612408.

Xiao, J. (2017). Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in Arabidopsis. *Nat. Genet.* **49**:1546–1552.

Xiao, J., Jin, R., Yu, X., Shen, M., Wagner, J. D., Pai, A., Song, C., Zhuang, M., Klasfeld, S., He, C., et al. (2017). Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in Arabidopsis. *Nat. Genet.* **49**:1546–1552.

Xu, L., and Shen, W.-H. (2008). Polycomb Silencing of KNOX Genes Confines Shoot Stem Cell Niches in Arabidopsis. *Curr. Biol.* **18**:1966–1971.

Xu, Y., Yang, J., Wang, Y., Wang, J., Yu, Y., Long, Y., Wang, Y., Zhang, H., Ren, Y., Chen, J., et al. (2017). OsCNGC13 promotes seed-setting rate by facilitating pollen tube growth in stylar tissues. *PLoS Genet.* **13**:e1006906.

Yang, Y., Fanning, L., and Jack, T. (2003). The K domain mediates heterodimerization of the Arabidopsis floral organ identity proteins, APETALA3 and PISTILLATA. *Plant J.* **33**:47–59.

Yang, J., Lee, S., Hang, R., Kim, S.-R., Lee, Y.-S., Cao, X., Amasino, R., and An, G. (2013a). OsVIL2 functions with PRC2 to induce flowering by repressing *O s LFL 1* in rice. *Plant J.* **73**:566–578.

Yang, C., Bratzel, F., Hohmann, N., Koch, M., Turck, F., and Calonje, M. (2013b). VAL- and AtBMI1-Mediated H2Aub Initiate the Switch from Embryonic to Postgerminative Growth in Arabidopsis. *Curr. Biol.* **23**:1324–1329.

Yoshida, N., Yanai, Y., Chen, L., Kato, Y., Hiratsuka, J., Miwa, T., Sung, Z. R., and Takahashi, S. (2001). EMBRYONIC FLOWER2, a novel polycomb group protein homolog, mediates shoot development and flowering in Arabidopsis. *Plant Cell* **13**:2471–81.

Yu, H., Ito, T., Wellmer, F., and Meyerowitz, E. M. (2004). Repression of AGAMOUS-LIKE 24 is a crucial step in promoting flower development. *Nat. Genet.* **36**:157–161.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y. V, Pellegrini, M., Goodrich, J., and Jacobsen, S. E. (2007a). Whole-Genome Analysis of Histone H3 Lysine 27 Trimethylation in Arabidopsis. *PLoS Biol.* **5**:e129.

Zhang, X., Germann, S., Blus, B. J., Khorasanizadeh, S., Gaudin, V., and Jacobsen, S. E. (2007b). The Arabidopsis LHP1 protein colocalizes with histone H3 Lys27 trimethylation. *Nat. Struct. &Amp; Mol. Biol.* **14**:869.

Zúñiga-Mayo, V. M., Gómez-Felipe, A., Herrera-Ubaldo, H., and de Folter, S. (2019). Gynoecium development: networks in Arabidopsis and beyond. *J. Exp. Bot.* **70**:1447–1460.

Supporting data

Dataset 1.

List of deregulated genes from RNA-seq of *bpcV* inflorescences.

Dataset 2.

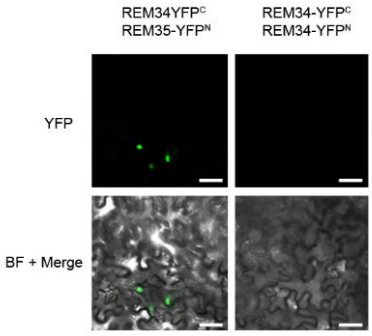
List of Putative BPCs target. First column, list of Dap-seq peaks, common to BPCs of class I and II; second column, list of deregulated genes from RNA-seq data (Dataset 1), cross-referenced with Dap-seq peaks data.

For overexpression lines	
SPT CDS AttB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTGTTGTTGGTGTAATGATATCAC
SPT CDS AttB2	GGGGACCACTTTGTACAAGAAAGCTGGGTGGGACACTGTTCAAGTAATTC
For expression analysis	
SPT fw	CCTTACTTCACCCGTGGAGATG
SPT rv	GCGTTGGAATGACCAATGTTC
UBI fw	CTGTTCACGGAACCCAATTC
UBI rv	GGAAAAAGGTCTGACCGACA
For ChIP analysis	
pSPT region 1 fw	CAGTATTAATGGTGAGACGAG
pSPT region 1 rv	ACGGTCGCATAGCTTGTAGG
pSPT region 2 fw	GTCATTTTCAAGTAATGTGTCC
pSPT region 2 rv	GTCATTTTCAAGTAATGTGTCC
pSPT region 3 fw	ATGCTACAGTAACAGCTACCTTC
pSPT region 3 rv	TTATCTCCCATCACTCTCTGC
pFUS3 fw	GCCTCTGTTTCGATCTGC
pFU3S rv	CAACCATCATTTTTCTCTCTC
ACT7 fw	CGTTTCGCTTTCCTTAGTGTTAGCT
ACT7 rv	AGCGAACGGATCTAGAGACTCACCTTG
For BiFC analysis	
HEC1 CDS AttB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGATTCTGACATAATGAAC
HEC1 CDS AttB2	GGGGACCACTTTGTACAAGAAAGCTGGGTTCATCTAAGAATCTGTGCATTGC
TCP15 CDS AttB1	GGGGACAAGTTTGTACAAAAAAGCAGGCTCGATGGATCCGGATCCGGATCAT
TCP15 CDS AttB2	GGGGACCACTTTGTACAAGAAAGCTGGGTGCTAGGAATGATGACTGGTGC

Table 1.

List of oligonucleotides used in this work

Supplementary figure



Supplementary figure 1. Bi-molecular fluorescence complementation (BiFC) assay. Nicotiana benthamiana epidermis cells were transiently transformed with the indicated YN and YC fusions. In the first and the second row yellows fluorescence and the merging in the bright field were shown, respectively. Scale bars=30 µm.

Part II

Manuscripts

Manuscript #1: BASIC PENTACYSTEINE and MADS-domain factors regulate the homeotic gene *SEEDSTICK* during flower development via LHP1 recruitment

Submitted to Plant Journal

BASIC PENTACYSTEINE and MADS-domain factors regulate the homeotic gene *SEEDSTICK* during flower development via LHP1 recruitment

Rosanna Petrella[1], Francesca Caselli[1], Irma Roig-Villanova[1,2], Valentina Vignati[1], Matteo Chiara[1], Ignacio Ezquer[1], Luca Tadini[1], Martin M. Kater[1], Veronica Gregis[1*]

[1] Università Degli Studi di Milano, Dipartimento di Bioscienze. Via Celoria 26, 20133, Milan, Italy.

[2] Centre for Research in Agricultural Genomics (CRAG), CSIC-IRTA-UAB-UB, Campus UAB, 08193 Barcelona, Spain.

Contact information: *Email of corresponding author Veronica Gregis: veronica.gregis@unimi.it

Running title: BPCs and SVP regulate *STK* via PRC

Summary

The BASIC PENTACYSTEINE (BPC) GAGA (C-box) binding proteins belong to a small plant transcription factor family. We previously reported that BPCs of class I directly bind to C-boxes in

the *SEEDSTICK (STK)* promoter and the mutagenesis of these *cis*-elements affects *STK* expression in the flower. Another key regulator of *STK* is the MADS-domain factor SHORT VEGETATIVE PHASE (SVP), which directly binds CArG-boxes to repress *STK* expression during the first stages of flower development. Here we show that BPCs of class II directly interact with SVP, and that MADS-domain binding sites in the *STK* promoter region are important for the correct spatial and temporal expression of this homeotic gene. Furthermore, we show that BPCs of class I and II act redundantly to repress *STK* expression in the flower, most likely by recruiting the POLYCOMB REPRESSIVE COMPLEX 1 and mediating the establishing and the maintenance of H3K27me3 repressive marks on the DNA. We investigate the role of TERMINAL FLOWER 2/LIKE HETEROCHROMATIN PROTEIN 1 (TFL2/LHP1) in the regulation of *STK* expression. Besides providing a better understanding of the role of BPC transcription factors in the regulation of S*TK* expression, our results suggest the existence of a more general regulatory complex composed of BPCs, MADS-domain factors and PRCs, that cooperate to regulate gene expression in reproductive tissues. We believe that our data along with the molecular model herein described could provide significant insights for a more comprehensive understanding of gene regulation in plants.

Significance statement: Both class I and II BASIC PENTACYSTEINE (BPC) proteins regulate the homeotic gene *SEEDSTICK (STK)* together with the MADS-box factor SHORT VEGETATIVE PHASE (SVP); furthermore, the component of PRC1 complex LIKE HETEROCHROMATIN PROTEIN 1 (LHP1) is involved in *STK* repression modulating H3K27me3 deposition and maintenance on the locus.

Keywords: MADS-box, BPCs, homeotic genes, STK, PRC, Arabidopsis, transcription factors

Introduction

Transcription factors (TFs) are regulators of gene expression; they act at multiple levels to orchestrate developmental processes. TFs bind specific DNA sequences and they can cooperate through genetic and epigenetic mechanisms. TFs act in multimeric complexes that can include members of different TFs families and other proteins. The composition of these complexes determines their binding specificity and their activity on target gene regulation (Martinez and Rao, 2012). Although in the last decades different classes of plant TFs have been characterised, the molecular mechanisms by which they act and the complexes they are part of, are yet to be fully understood. Recently, a new class of transcription factors, named BASIC PENTACYSTEINE/ BARLEY B RECOMBINANT (BPC/BBR), have been identified (Santi et al., 2003). They bind the RGARAGRRA consensus site, called GAGA or C-box, to regulate their target genes (Meister et al., 2004; Kooiker

et al., 2005; Simonini et al., 2012; Simonini and Kater, 2014; Hecker et al., 2015; Mu et al., 2017b; Shanks et al., 2018; Theune et al., 2019; Roscoe et al., 2019; Wu et al., 2019). BPCs have been described in different plant species including monocots (*Oryza sativa* (rice) and *Hordeum vulgare* (barley)) and dicots (*Glycine max* (soy-bean) and *Arabidopsis thaliana)* (Sangwan and O'Brian, 2002; Santi et al., 2003; Kooiker et al., 2005; Monfared et al., 2011; Berger and Dubreucq, 2012; Simonini et al., 2012; Simonini and Kater, 2014; Hecker et al., 2015; Mu et al., 2017a; Mu et al., 2017b; Xiao et al., 2017; Shanks et al., 2018; Theune et al., 2019; Roscoe et al., 2019; Wu et al., 2019). In Arabidopsis, *BPCs* are divided in three subfamilies: class I (containing *BPC1* to *BPC3*), class II (containing *BPC4* to *BPC6*), and class III (containing only *BPC7*) (Meister et al., 2004; Monfared et al., 2011). Except for *BPC5*, which is a pseudogene, all the other *BPCs* are ubiquitously expressed. Combinations of multiple *bpc* mutants show strong phenotypes with a wide range of defects, addressing an important role during plant development (Monfared et al., 2011).

Previously, we have identified the MADS-box gene *SEEDSTICK (STK)* as a direct target of BPCs belonging to the class I (Kooiker et al., 2005; Simonini et al., 2012). *STK* is specifically expressed during ovule and seed development and has a wide range of functions in these tissues (Favaro et al., 2003; Pinyopich et al., 2003; Brambilla et al., 2007; Losa et al., 2010; Mizzotti et al., 2014; Mendes et al., 2016; Balanzà et al., 2016; Ezquer et al., 2016; Herrera-Ubaldo et al., 2019).

During carpel development *STK* expression is confined to placental tissues and ovule primordia; in mature ovules it is expressed strongly in the funiculus and in integuments that will later form the seed coat (Mizzotti et al., 2014). *STK* acts redundantly with two other MADS-box factors named *SHATTERPROOF 1* (*SHP1*) and *SHATTERPROOF 2* (*SHP2*) in the determination of ovule identity (Favaro et al., 2003; Pinyopich et al., 2003). BPCs of class I form homo- and hetero-dimers and bind the promoter of *STK* at several C-boxes inducing DNA loop formation (Kooiker et al., 2005). The C-boxes are important for *STK* regulation since mutations in these sequences cause ectopic expression of the homeotic gene in the flower (Kooiker et al., 2005; Simonini and Kater, 2014). The MADS-domain factor SHORT VEGETATIVE PHASE (SVP) is another important regulator of *STK.* SVP acts redundantly with APELATA 1 (AP1) and AGAMOUS-LIKE 24 (AGL24) to repress *STK* expression during early stages of flower development, by binding directly to its promoter (Simonini et al., 2012; Gregis et al., 2013). Furthermore, BPCs of class I and SVP directly interacts to repress *STK* expression in the floral meristem, and C-boxes are important to facilitate or stabilise the binding of SVP to the *STK* promoter region (Simonini et al., 2012).

Recently, members of the BPCs family have been shown to be implicated in the recruitment of histone-modifying complexes that can inactivate gene expression, like the Polycomb Repressive Complexes (PRCs) (Hecker et al., 2015; Mu et al., 2017b; Xiao et al., 2017; Roscoe et al., 2019; Wu et al., 2019). BPCs of class II directly interact with LHP1, a plant PRC1 component that is associated with genes marked by trimethylation of histone H3 lysine 27 (H3K27me3). Interestingly, it was demonstrated that SVP can form hetero-dimers with LHP1 to modulate H3K27me3 deposition on the *SEPALLATA3* (*SEP3*) locus (Liu et al., 2009). Furthermore, BPCs can physically interact with the PRC2 subunit SWINGER (SWN) to repress the expression of their target *ABSCISIC ACID INSENSITIVE4 (ABI4)* during root development by the trimethylation of Histone H3 Lysine 27 (Mu et al., 2017b).

Here we clarify the role of BPCs of class II and SVP in the regulation of *STK* expression, mostly focusing on the molecular mechanisms they act. We show that MADS-domain binding sequences in the *STK* promoter region are important for the correct spatial and temporal expression of the ovule identity gene. Our data indicate that both BPCs of class I and II are necessary for the correct expression of *STK*, by modulating the deposition and the maintenance of H3K27me3 marks. Our results provide insights into the molecular mechanisms that drive transcription regulation in plants and investigate the involvement of a protein complex in which BPCs, MADS-domain factors and PRCs can cooperate to orchestrate the expression of homeotic genes during plant development.

Results

STK is deregulated during flower development in the *bpc1-2 bpc2 bpc3 bpc4 bpc6* mutant

To gain more insights into the role of BPCs of class I, but also class II in the regulation of *STK* expression during flower development we generated the *bpc1-2 bpc2 bpc3 bpc4 bpc6* quintuple mutant (henceforth called *bpcV*), in which the complete knockout allele *bpc1-2* is combined with the previously described mutant alleles for both class I and II BPCs (Monfared et al., 2011; Simonini and Kater, 2014) In fact, Monfared et al. (2011) already described different combinations of BPCs mutants but all of them contained the *bpc1-1* allele which does not seem to be a full knockout allele (Monfared et al., 2011; Simonini and Kater, 2014). The contribution of class I and II BPCs in the correct regulation of *STK* was analysed by *in-situ* hybridisation assay. In wild-type plants, *STK* expression is confined to ovules and the placenta and was never observed in flowers before stage 8 neither in inflorescences nor in floral meristems (Figure 1A and B). The knock-out of all the BPCs of class I (Figure 1C and D) or class II (Figure 1E and F) did not affect *STK* expression in the flower.

In contrast, in the *bpcV* mutant the expression of *STK* was observed in ovules and placenta, but also in developing petals (Figure 1H), floral meristems and young flowers; notably, the expression was detectable also in the organ primordia (Figure 1G). To confirm that the deregulation of *STK* expression registered in our *bpcV* mutant was specifically due to BPCs mutation, we hybridised both *bpcV* and *wild-type* inflorescences with H4 histone gene and specific-*STK* sense probe (Fobert et al., 1994; Favaro et al., 2003), as showed in Figure S1. The maintained expression of histone H4 gene in bpcV confirmed the integrity of the tissue (Figure S1C and D). These results clearly demonstrated the redundant role that class I and II BPCs have in the regulation of *STK* expression during flower development.

Overexpression of *STK* affects plant development

To further investigate the role of BPCs in plant development, we performed a phenotypical analysis on the *bpcV* mutant. The plant is shorter compared to the wild-type, as depicted in Figure 2A. Furthermore, it is characterized by some vegetative and reproductive defects. Development of either rosette and cauline leaves is affected, as showed in figure 2B and C, respectively; as matter of fact, they are smaller, compared to the wild-type. The knockout of the five BPCs genes caused a drastic phenotype in the silique. In the wild-type, upon a successful fertilization, from 3 to 12 days after pollination (dap), the silique elongates to reach its maximum length, as showed in Figure 2D; in contrast, in *bpcV* no elongation was registered; furthermore, the siliques appeared shorter compared to the wild-type (Figure 2D). The results above clearly suggested that BPCs mutation broadly affects plant development.

The deregulation of *STK* during flower development, observed in our mutant (Figure 1G and H) might cause some effects throughout plant development. To further investigate this hypothesis, Arabidopsis wild-type plants were transformed with a chimeric gene construct in which the CDS of *STK* was fused to the Cauliflower mosaic virus (CaMV) 35S promoter (Favaro et al., 2003). *STK* expression was checked by quantitative Real-Time PCR in three lines, where we could detect statistically significant upregulation of *STK* expression (Figure S2). The line that showed the highest upregulation of the homeotic gene (henceforth called *35S:STK*) was propagated and the following generations were selected for further analysis. Intriguingly, also this plant was shorter compared

to the wild-type (Figure 2A) and showed some defects that phenocopies our *bpcV* mutant during vegetative development (Figure 2B and C). In particular, we could detect the same defects in siliques, observed in our *bpcV* mutant (Figure 2D).

The MADS-domain factor *STK* is a master regulator of ovule and consecutively seeds development and production (Favaro et al., 2003; Pinyopich et al., 2003; Mizzotti et al., 2014; Ezquer et al., 2016). To determine whether the constitutive expression of *STK* would affect seed development, seed area was analysed in *35S:STK* plant and *bpcV* mutant; wild-type, *stk* and *arf2-8* seeds were used as controls. Our results could confirm that *stk* had smaller seeds, as previously reported by Pinyopich et al. (2003) whereas *arf2-8* seeds were much bigger (Schruff et al., 2006). Interestingly, both *bpcV* and *35S:STK* plants showed wider seeds area, compared to the wild-type and the *stk* mutant (Figure 2E). Our results support the hypothesis that BPCs might regulate *STK* expression later on development, in the gynoecium and in the seeds. Also, they address an important role to *STK* and BPCs of class I and II throughout plant development.

BPCs of class II interact with SVP

The *in-situ* analysis suggests that BPCs of class II have an important role in regulating *STK* expression. To understand whether BPCs of class II (BPC4 and BPC6) could interact with SVP, different protein interaction assays were performed.

We confirmed by yeast two-hybrid assays and bimolecular fluorescence complementation assays (BiFC) in tobacco leaves (*Nicotiana benthamiana*) that BPC4 and BPC6 can form homo- and heterodimers (Wanke et al., 2011 and Figure S3A and D).

We showed, using yeast two-hybrid assays, that both BPC4 and BPC6 can interact with SVP (Figure 3A). To confirm the interactions between SVP and the BPC4 and BPC6 factors, a Co-Immunoprecipitation (Co-IP) assay was performed, using SVP-GFP in combination with BPC4-RFP and BPC6-RFP fusion proteins, transiently co-expressed under the control of the Cauliflower mosaic virus (CaMV) promoter in *Nicotiana benthamiana* leaves. These Co-IP experiments, using the BPC-RFP-tagged proteins or SVP-GFP as bait all revealed co-precipitation of the BPC and SVP proteins, respectively (Figures 2B and Figure S3B), suggesting that BPC4 and BPC6 are able to form complex(es) with SVP *in vivo*.

Further validation of these results in planta was obtained by BiFC assays in tobacco (*Nicotiana benthamiana*) leaves. The combination SVP-YFP^N-BPC4-YFP^C showed a clear nuclear interaction

between BPC4 and SVP (Figure 3C). All the other combinations that were tested (BPC4-YFP^N SVP-YFP^C, SVP-YFP^N BPC6-YFP^C, BPC6-YFP^N SVP-YFP^C) resulted in an interaction in the cytoplasm (Figure 2C). Although this result was unexpected, Immink et al. (2002) previously showed that some MADS-domain proteins need to dimerise with another MADS-domain factor for their nuclear localisation. SVP interacts with the MADS-domain protein AP1 during floral development and therefore it might facilitate nuclear location of SVP-BPC dimers (Pelaz et al., 2002; de Folter et al., 2005). We tested this hypothesis by co-expressing SVP-BPC4 and SVP-BPC6 dimers with an AP1-RFP fusion protein in tobacco leaves. As showed in Figure 3D, the presence of AP1 is enough for the nuclear localisation of the BPC4- and BPC6-SVP dimers. To determine whether BPCs of class II could directly interact with AP1, a yeast-two hybrid assay was performed. The reported interaction between BPC6 and AP1 (Figure S3C) could not be confirmed by BIFC interaction assay, as reported in Supplementary figure 3C. Taken together this results clearly showed that AP1 is sufficient for the traslocation of the class BPCs II-SVP heterodimers to the nucleus.

Molecular mechanism of SVP-BPCs binding to the regulatory region of *STK*

To clarify the mechanism by which BPCs and SVP interact with the *STK* regulatory region, we performed a series of Chromatin Immunoprecipitation (ChIP) experiments in different mutant backgrounds.

As shown in Figure 4A, SVP binds CArG-boxes that are surrounded by several C-boxes in the regulatory region of *STK* (Simonini et al., 2012). As previously shown, SVP, AP1 and AGL24 redundantly control the identity of the floral meristem through direct repression of floral homeotic genes (Gregis et al., 2008; Gregis et al., 2009). In fact, in the *svp agl24 ap1-12* triple mutant, *STK* is ectopically expressed in the floral meristem and young flowers (Simonini et al., 2012). To determine whether SVP, AP1 and AGL24 are required for BPCs binding to the promoter of *STK*, three independent ChIP assays using specific antibodies against class I BPCs were performed. The experiments were conducted using *svp agl24 ap1-12* triple mutant inflorescences. Furthermore, inflorescences from wild-type and *bpc1-2 bpc2 bpc3* triple mutant plants were used as positive and negative control, respectively. In our ChIP experiments, no enrichment was detected in the *svp agl24 ap1-12* triple mutant in the region containing C-box 12 (Simonini et al., 2012; Figure 4B). These results demonstrated that SVP, but probably also AP1 and AGL24, are necessary for class I BPCs binding to the *STK* promoter.

Subsequently, the role of class I and class II BPCs in the control of SVP binding to the promoter of *STK* was investigated by crossing the *bpcV* mutant, described above, with *pSVP*:SVP-GFP *svp* plants. In subsequent generations, plants homozygous for the *svp* and *bpc1-2 bpc2 bpc3 bpc4 bpc6* mutations containing the *pSVP:*SVP-GFP construct were selected. ChIP experiments using commercial antibodies against GFP were performed. Inflorescences from *pSVP*:SVP-GFP *svp* Arabidopsis plants were used as a positive control, whereas wild-type was used as a negative control. An enrichment was detected when binding to the consensus regions for SVP was tested in the *bpcV* mutant background (Figure 4C). These results suggest that BPCs of class I and class II are not necessary for SVP binding to the CArG-boxes in the *STK* promoter.

Taken together, the results obtained by these ChIP assays are consistent with a model where SVP binds the *STK* promoter independently of BPCs, whereas BPCs require MADS-domain factors for the correct binding to the *STK* regulatory region.

CArG boxes drive the correct temporal and spatial expression of *STK* and are important for SVP and BPCs binding to the promoter of *STK*

To further characterise the role of SVP and in general of MADS-domain factors in *STK* regulation, we decided to perform a functional characterisation of the CArG-boxes contained in *STK* regulatory region; these regions were identified based on the MADS-domain consensus sequences located in the locus of *STK* where SVP binding was detected by ChIP-seq (Gregis et al., 2013). Considering our previous experiments using a *STK* promoter with mutated C-boxes or CArG-boxes (Simonini et al., 2012; Mendes et al., 2016), we suspected that the 12 CArG-boxes in the regulatory region of *STK* could be redundant. Therefore, a mutated version of the *STK* promoter was used in which 11 out of the 12 CArG-boxes were altered, considering the following criteria: (i) preserving the DNA conformation, introducing only 4 to 5 transitions to each consensus; (ii) avoiding the mutation of C-boxes; (iii) preventing the formation of new CArG-boxes (see Table S1). The mutagenized *STK* promoter was fused to the *uidA* reporter gene that encodes for beta-glucuronidase (GUS) and the resulting *pSTK-CArGm*:GUS construct was used to transform Arabidopsis Col-0 wild-type plants and *pSVP*:SVP-GFP *svp* plants. As a positive control, the wild-type *STK* promoter (*pSTK-CArGwt*:GUS), which drives specific expression in the placenta and all stages of ovule development, was used (Figure 5A). Out of the 39 plants transformed with the *pSTK-CArGwt*:GUS construct, 36 (92%) showed a correct spatial and temporal expression of the GUS reporter, reflecting the endogenous expression of *STK* (Figure 5B-D), whereas the other three

plants did not show any GUS activity. In contrast, out of 39 plants transformed with the *pSTK-CArGm*:GUS construct, 14 (36%) showed a strong deregulation of GUS expression whereas the remaining 64% showed a correct expression of the reporter. Interestingly, GUS expression was extended also in the inflorescence and floral meristems (Figure 5E and G) and in all the floral organs (Figure 5F). These results support the idea that the CArG-boxes in the *STK* promoter have a role for the correct expression of this MADS-domain protein, but also suggest the possibility that other MADS-box transcription factors, expressed in different tissues compared to where *SVP* is expressed, might be involved in the regulation of *STK*.

To assess whether MADS-domain binding sites on the *STK* promoter are necessary for SVP binding, we performed ChIP experiments using antibodies against GFP and inflorescences of *pSVP*:SVP-GFP *svp* plants with *pSTK-CArGm*:GUS. The wild-type endogenous *STK* promoter was used as a positive control, whereas as negative control, inflorescences of *pSTK-CArGm*:GUS plants without SVP:GFP were used. Specific primers were used to discriminate between the endogenous wild-type promoter and the mutated one (see Experimental procedures and Table S2). These experiments showed that no enrichment was detected when binding to the mutated region was tested (Figure 5H), which suggests that CArG-boxes in the promoter of *STK* are necessary for SVP binding.

To further investigate the role of SVP and BPCs in the regulation of *STK* expression, binding of class I BPCs to the *pSTK* promoter with the mutated CArG-boxes was tested using the *pSTK-CArGm*:GUS line and using antibodies against BPCs. As positive control, the endogenous region of the *STK* promoter was used, whereas as negative control antibodies against HA were used. No enrichment was detected when BPCs binding to the mutated region was tested, suggesting that mutating CArG-boxes abolished BPCs binding (Figure 5I). Collectively, these results indicate that CArG-boxes are important for the correct spatiotemporal regulation of *STK* expression. Moreover, these experiments confirm that SVP binding is necessary for the recruitment of BPCs to the *STK* promoter.

The expression of *STK* is influenced by epigenetic modifications

A novel role of BPCs in the regulation of their target expression by the recruitment of PRCs, has been recently reported (Mu et al., 2017b; Xiao et al., 2017; Shanks et al., 2018; Roscoe et al., 2019; Wu et al., 2019) The presence of H3K27me3 is mainly correlated with gene silencing; also, H3K27me3 targets are enriched for genes with tissue-specific expression patterns in Arabidopsis, suggesting that this epigenetic mark is modulated in response to developmental cues (Zhang et al., 2007). Interestingly, the

locus of *STK* shows strong coverage of H3K27me3 depositions in seedlings (Turck et al., 2007; Lafos et al., 2011; Li et al., 2015).

To understand whether the deregulation of *STK* observed in the *bpcV* background was correlated with alterations in the deposition of the H3K27me3 epigenetic mark, we analysed the relative enrichment of this mark in the *STK* locus. Two regions were tested: region 1, located in the first *STK* intron, and region 2 immediately after the stop codon of the gene, as illustrated in Figure 6A. ChIP experiments were performed using specific antibodies against H3K27me3 and analysed by quantitative Real-Time PCR. The AT2G22560 and *AGAMOUS* loci were used as negative and positive control for H3K27me3 marks, respectively (Li et al., 2015). Interestingly, in *bpcV* inflorescences a reduction of H3K27me3 deposition was detected in both the two selected *STK* regions, compared to the wild-type (Figure 6B). These results were consistent with the observed ectopic expression of *STK* in *bpcV* background (Figure 1G and H) and suggest an active role of BPCs in the establishment of repressive epigenetic marks.

LHP1 regulates *STK* expression during flower development

To further investigate the molecular mechanism by which BPCs and SVP regulate *STK* expression, we considered previously published interactions for both SVP and BPCs. Interestingly, SVP and BPC6 are both able to interact with LHP1 (Liu et al., 2009; Hecker et al., 2015). LHP1 is a component of the PRC1 complex and recognises loci marked by H3K27me3 *in vivo*, acting as part of a mechanism that represses the expression of PRC2 targets (Turck et al., 2007). Furthermore, it has been recently reported that LHP1 could directly interact with several members of the PRC2 gene family (Derkacheva et al., 2013b; Wang et al., 2017) to facilitate their recruitment to target genes.

Mutation of the LHP1 locus resulted in pleiotropic effects due to the deregulation of several genes during plant development (Larsson et al., 1998)To address the role of LHP1 in the regulation of *STK* during flower development, *STK* expression was analysed in the *lhp1* mutant background by *in-situ* hybridisation. In line with our hypothesis, the knock-out of *LHP1* affects *STK* expression in the flower; the homeotic gene is indeed detected in floral and inflorescence meristems, as well as in young flowers (Figure 7C) whereas the expression of *STK* in the mature flowers were not altered, as shown in Figure 7D. *STK* expression was further analysed in *lhp1* inflorescences by quantitative Real-Time PCR. The MADS-box gene *STK* turned out to be upregulated in the *lhp1* mutant background, as shown in Figure 7E.

LHP1 is associated to *STK* genomic region in seedlings (Turck et al., 2007). To validate the association to the *STK* locus in reproductive tissues we performed a ChIP assay, collecting inflorescences from *pLHP1*:LHP1-GFP *lhp1* plants (Kotake et al., 2003), testing Region 1 and 2. We

could confirm LHP1 association to the *STK* locus, even though enrichment was detected only in Region 2, close to the 3' UTR of the homeotic gene. Collectively, those results clearly confirmed a role of LHP1 for the regulation of *STK* expression in the flower.

Genome-wide analysis of BPCs and MADS-domain factor binding site locations

To investigate patterns of genome-wide enrichment of MADS-domain and BPCs transcription binding sites, a publicly available repository of transcription factor binding profiles was used (http://neomorph.salk.edu/dap_web/pages/index.php; (O'Malley et al., 2016). DNA affinity purification sequencing (DAP-seq) is a transcription factor (TF)-binding site discovery assay which combines next-generation sequencing of a genomic DNA library with affinity-purified TFs (Bartlett et al., 2017). Average profiles for the two families were reconstructed by using a simple consensus method (see Experimental procedures). Overlap of genomic regions associated with DAP-seq peaks of MADS-box and BPCs TFs families was used as a proxy to investigate possible interactions.

.

A highly significant over-representation of overlapping peaks (p-value hypergeometric ≤ 3.5e-4) was observed which can be considered as an indication of possible direct interaction between MADS-domain and BPC proteins on a DNA target sequences. Of note, our analysis of the complete dataset of Malley et al. (2016) which provides DAP-seq data for more than 500 TFs, belonging to 41 distinct transcription factor families, suggests that overall only 4 additional families of transcription factors show significant levels of overlap with DAP-seq peaks of members of the BPC family (REM, C2C2gata, SRS and Trihelix, Data S1). As outlined in Figure S4, overlap with DAP-seq peaks associated with transcription factors of the MADS-box family accounts for 19.6% of the total number of significantly overlapped peaks. All in all we believe that these data are consistent with a model where BPCs can interact with a restricted set of TFs families, which is not limited to- but is very likely to include members of the MADS-box family. *In silico* prediction of enriched sequence motifs is largely concordant with this model (Figure 6). In fact, when *de-novo* reconstruction of enriched motifs is performed we observe: i) a strong enrichment in CArG-box like motifs in genomic regions that are bound by MADS-domain factors; ii) a strong enrichment of C-boxes like motifs in genomic regions associated with BPC DAP-seq peaks; iii) a strong enrichment of both type of motifs (CArG-boxes and C-boxes) when regions containing coincident DAP-seq peaks are considered.

To further investigate possible biological pathways regulated by MADS-BPCs complexes, genomic regions associated with overlapped MADS-box and BPCs DAP-seq peaks were annotated by the means of the annotatePeaks program from the Homer suite (Heinz et al., 2010a). A total of 519 candidate target genes was obtained, which were subjected to functional enrichment analyses using the DAVID program (Huang et al., 2009).

Coincident MADS-BPCs and DAP-seq peaks were subsequently cross-referenced with H3K27me3 ChIP-seq peaks (Lafos et al., 2011), to gain a better insight on potential MADS-BPCs target genes that are co-regulated *in vivo*. A list of 93 candidate genes was obtained (Data S2). Interestingly, functional enrichment analysis of this set of genes resulted in a significant enrichment of transcription factor encoding genes (GO term "DNA-binding transcription factor activity") for both lists: genes associated with MADS-BPCs peaks and/or with MADS-BPCs and H3K27me3. These results support the idea that MADS-BPC-PRC complexes play a pivotal role in the regulation of master players in development as shown in Data S2.

Discussion

Our knowledge on the molecular mechanisms controlling gene expression in plants is still fragmented and needs further study. Here we used the ovule identity gene *STK*, which is specifically expressed in Arabidopsis placenta, ovules and seeds, as a model system to investigate the regulation of homeotic genes expression. Previously, we showed that the MADS-domain factors SVP, AGL24 and AP1 and the class I BPC transcription factors repress *STK* expression during early stages of flower development (Kooiker et al., 2005; Simonini et al., 2012; Gregis et al., 2013). Information concerning the role of the MADS-domain factors in *STK* regulation derived by analysis of the *agl24 svp ap1-12* triple mutant, whereas the role of BPC factors in *STK* regulation was mainly investigated by mutagenesis of multiple BPCs binding sites (C-boxes) in the *STK* regulatory region. Mutation of those sites caused indeed a strong deregulation of *STK* during flower development, suggesting that C-boxes and therefore BPC factors are important for *STK* regulation (Simonini et al., 2012). Even though these data are very interesting, mutating cis-elements still provides indirect evidence. Therefore, we further investigated the role of BPC factors in the regulation of *STK* by generating the *bpc1-2 bpc2 bpc3 bpc4 bpc6* quintuple mutant (*bpcV)*. In this background the pattern of *STK* expression was also compromised and its transcripts were detected by *in-situ* hybridisation in the floral meristem and floral organs, confirming the important role of BPCs

belonging to both class I and class II in *STK* repression. Intriguingly, ChIP experiments showed that in the *bpcV* mutant background SVP could still bind the *STK* promoter, suggesting that SVP bound the *DNA* independently of BPC factors. Previously, we reported that the mutagenesis of C-boxes on the *STK* promoter affects SVP binding to the DNA. This discrepancy could not be due to the fact that nearby C-box mutations influence CArG-box affinity, since control experiments ruled this option out (Simonini et al., 2012). However, we cannot exclude the possibility that mutagenesis of C-box elements in the promoter of *STK* might introduce structural changes which can alter significantly the binding affinity of regulatory elements in the promoter of *STK*; moreover, the presence of additional unknown co-factors of SVP cannot be completely excluded. This notwithstanding it is important to stress that all our experiments demonstrate that BPC and MADS-domain factors are both essential for the correct expression of *STK* and that, binding of SVP alone is not *per se* sufficient to repress *STK* expression in the floral meristems.

We also showed that CarG-boxes in the promoter of *STK* are required for the binding of both SVP and BPCs . Mutagenesis of CArG-boxes in the *STK* promoter resulted in a strong deregulation of this homeotic gene in all tissues of the inflorescence. Considering that *SVP* is specifically expressed in floral meristems (Gregis et al., 2008), it is highly likely that other MADS-domain factors bind the CArG-boxes to regulate *STK* expression in other tissues. In the light of their expression patterns during later stages of flower development, we believe that AGL24 and AP1 should be good candidates to repress *STK* in other floral tissues (Gustafson-Brown et al., 1994; Yu et al., 2004).

Our *in-situ* hybridisation address an important role for BPCs of class II in the repression of *STK,* since we showed a deregulation of *STK* only in *bpcV* mutant background.

The analysis of the *35S:*STK line further explored the effects of the deregulation of *STK* throughout reproductive development. (Pinyopich et al., 2003) reported that *stk* presented smaller seeds compared to the wild-type. It has been later clarified that this phenotype is highly likely caused by defects in the flavonoid pathway, that ultimately controls seeds size (Doughty et al., 2014); as matter of fact, several genes involved in flavonoid biosynthesis and transport were differentially expressed in the mutant *stk*, compared to the wild-type (Mizzotti et al., 2014). The defects in seeds size registered in *bpcV* and *35S:STK* suggest that BPCs might control *STK* expression later on development, in the seed.It would be worth to check whether the mutation of BPCs of class I and II could affect the expression of the homeotic gene also later on development, in the silique.

We previously revealed that BPCs of class I can interact with each other (Simonini et al., 2012); moreover, BPCs of class II form homo and heterodimers with members of class I (Wanke et al.,

2011; Simonini et al., 2012 and Figure S3). BPC protein-protein interactions studies suggest that BPC factors of class I and II can act synergistically and redundantly to regulate the expression of their targets as we demonstrated for *STK*. An example has been provided by Mu et al. (2017b), who showed that mutations in BPCs of class I and II increased *ABI4* expression in roots. To further investigate the molecular and functional relationships of MADS-domain factors SVP and BPCs of class II, we tested their ability to form heterodimers in planta. Interestingly, here we showed that AP1 can trigger the colocalization of SVP-BPC4 (and BPC6) to the nucleus. Our results support and further clarify the role of AP1 in the regulation of *STK*.

Farkas et al. (1994) has first characterised the GAGA Associated Factor of *Drosophila melanogaster* (dGAFs). Even though GAFs and BPCs are phylogenetically unrelated, they present several similarities. BPCs can bind to (GA)n sequences (Berger and Dubreucq, 2012) to control the expression of their targets(Meister et al., 2004; Berger et al., 2011; Simonini et al., 2012; Simonini and Kater, 2014; Mu et al., 2017a; Mu et al., 2017b; Theune et al., 2019; Roscoe et al., 2019; Wu et al., 2019). They also present a highly conserved zinc finger like DNA-binding domain, similar to *Trl* of Drosophila (Wanke et al., 2011). Interestingly, cooperative binding of BPC1 proteins to GA-rich motifs in the *STK* promoter region leads to a condensation and looping of DNA (Kooiker et al., 2005), similar to what has been described for dGAF from Drosophila. Recent works in Arabidopsis revealed an intriguing interaction among BPCs and Polycomb group proteins, as it was found in animals for dGAF, which can cooperate with Polycomb Group factors (PcG) to repress gene expression (Horard et al., 2000).

PcG complexes have paramount roles in cell fate determination or cell differentiation both in plants and in animals. These proteins have been identified in Drosophila more than 40 years ago as key repressors of homeotic genes (Hox) throughout embryonic development (Lewis, 1978). Besides, the sequences and functions of PcG genes are highly conserved between animals and plants. Different works recently showed that BPCs can interact with proteins belonging to PRC1 and PRC2, suggesting that it could be a mechanism to repress the expression of their target genes (Wanke et al., 2011; Mu et al., 2017b; Xiao et al., 2017).

Our results provide new insights into the role of BPCs in the recruitment of PRC members on the regulatory region of target genes. We indeed suggest that BPCs of class II and SVP can recruit LHP1 and act redundantly with the class I members to regulate the accumulation of the H3K27me3 repressive mark on the regulatory region of *STK.* In fact, we registered a reduction of H3K27me3 in the *bpcV* mutant. In the *lhp1* background we detected increased levels of *STK,* whose expression

was localized also in the inflorescences and in the floral meristems as well as in the first floral buds. In contrast to our results in the *bpcV*, no signal was detected in other floral structures at maturity, thus suggesting that BPCs of class I and II might repress *STK* expression during flower development also via other mechanisms that do not involve LHP1 activity.

Several BPCs targets have been discovered in the recent years. Most of them are also associated with PRC mediated silencing: the KNOX gene *BREVIPEDICELLUS (BP)* is repressed by BPCs throughout flower development (Simonini et al., 2012). Its expression is directly regulated by the recruitment of the EMBRYONIC FLOWER (EMF) complex by ASIMMETRIC LEAVES 1 and 2 (AS1 and AS2), which triggers H3K27me3 deposition (Lodha et al., 2013). *BP* was also identified in our computational analysis of regions enriched in binding sites for MADS-domain and BPC family members and resulted decorated with H3K27me3 marks (Data S2). Also, *FUS3* has recently been characterised as a BPCs target (Roscoe et al., 2019; Wu et al., 2019) and already reported to be a target of PRCs (Makarevich et al., 2006; Zhang et al., 2007; Bouyer et al., 2011; Yang et al., 2013; Xiao et al., 2017) Interestingly, our computational analysis showed that its regulatory region is also bound by MADS-domain factors, suggesting a conserved mechanism for target regulation.

The MADS-domain factor SVP interacts with LHP1 and is required to recruit the PRC1 factor to the promoter of *SEPALLATA3 (SEP3)*, acting like a pioneer factor (Liu et al., 2009). In accordance to this hypothesis, H3K27me3 deposition on the *SEP3* locus is reduced in *lhp1* background. ChIP assay could confirm a direct role of LHP1 in the regulation of *STK*; we showed indeed that the member of PRC1 is associated to the locus of the homeotic gene in a region, closed to the 3'UTR. We previously reported that SVP preferentially binds the 3'UTR of its targets (Gregis et al.,2013). Thus, it is highly possible that SVP could recruit LHP1 on the *STK* locus and subsequently repress the expression of the ovule identity gene via PRC2 recruitment, as previously showed for *SEP3* (Liu et al., 2009).

PRC2 components are required for H32K27me3 deposition to the target locus (Wang et al., 2017). Three different PRC2 complexes regulate plant development by targeting a subset of genes. LHP1 has been reported to associate with several PRC2 members (Derkacheva et al., 2013; Wang et al., 2016), to mediate their recruitment to the target locus. Previously, has been reported a participation of EMF complex in the repression of *AGAMOUS (AG) and SEPALLATA3 (SEP3)* expression in the flower (Yoshida et al., 2001; Kinoshita et al., 2001; Chanvivattana et al., 2004; Calonje et al., 2008); intriguingly, Derkacheva et al. (2013) reported LHP1 direct association to the EMF complex.

We previously characterised *SEP3* and *AG,* as targets of BPCs and SVP (Gregis et al., 2009; Simonini et al., 2012); as matter of fact, both *SEP3* and *AG* are deregulated in *lhp1 bpc4 bpc6* triple mutant, confirming a LHP1-class II BPCs interplay (Hecker et al., 2015).

Considering that, it is tempting to speculate that BPCs and SVP might regulate *STK* expression by the recruitment of LHP1; the PRC2 member could interact with the EMF complex to mediate the correct deposition of the H3K27me3; also, LHP1 activity could assure the maintenance and spreading of the repressor mark to the locus of the homeotic gene throughout flower development.

All these examples suggest the existence of a protein complex (Figure 9), in which BPCs and SVP cooperate to recruit LHP1 and regulate the expression of a subset of genes during early stages of flower development.

Understanding the molecular mechanisms through which BPCs and SVP containing complexes act is important since it is likely that the mechanism by which these factors regulate *STK* can be extended to many other genes during plant development. This is based on the following observations: (i) many genes contain both C-boxes and CArG-boxes in their putative promoter regions; (ii) BPCs are ubiquitously expressed in plants while MADS-domain factors are specifically expressed in all the fundamental developmental stages; and (iii) combination of *bpc* alleles showed pleiotropic phenotypes (Monfared et al., 2011). Furthermore, Berger et al. (2011) identified three cis-elements required for LEAFY COTYLEDON2 (LEC2) repression: C-boxes, CArG-boxes and PRE-like elements, corroborating the idea that the understanding of the synergistic interaction between MADS-domain factors and BPCs is an important key to decode gene regulation in plant. Several key developmental factors that are worth to be tested as putative direct targets of both MADS-domain and BPC factors are reported in Data S2. In fact, they were identified in our computational analysis of regions enriched in binding sites for both MADS and BPC family members.

The regulatory mechanism through which BPCs act is not restricted to Arabidopsis. Several GAGA binding proteins have first been discovered in crops and several targets have already been characterised (Sangwan and O'Brian, 2002; Santi et al., 2003; Meister et al., 2004). Therefore, a better understanding of the mechanisms by which these factors act in Arabidopsis may give an advancement of knowledge for future noteworthy crop improvement.

In conclusion, our work provides a contribution to a better understanding of how genes are regulated in plants, mostly by exploring the molecular mechanism through which MADS-domain and BPC factors modulate gene expression.

Experimental procedures

Plant Material and Growth Conditions

Arabidopsis thaliana ecotype Columbia was used in this study; the plants were directly sown on soil and kept under short-day conditions for 2 weeks (22°C, 8 h light and 16 h dark) and then moved to long-day conditions (22°C, 16 h light and 8 h dark). The *agl24 svp ap1-12* triple mutant and the *pSVP*:SVP-GFP *svp* line were previously described by Gregis et al. (2008; 2009); genotyping of the *bpc1-2 bpc2 bpc3 bpc4 bpc6* mutants was done according to Simonini and Kater (2014) and Monfared et al. (2011). Seeds from the *lhp1, arf2-8* and *stk* mutant in Columbia background were obtained from the Nottingham Arabidopsis Stock Centre.

Generation of quintuple mutants and marker lines

The *bpc 1-2 bpc2 bpc3 bpc4 bpc6* quintuple mutant was obtained crossing the *bpc 1-2 bpc2 bpc3* triple mutant (Simonini and Kater, 2014) and *bpc4 bpc6* double mutant (Monfared et al., 2011); the *pSVP*:SVP-GFP *svp bpc1-2 bpc2 bpc3 bpc4 bpc6* was obtained crossing the line previously described by Gregis et al. (2009) and the *bpcV*.

Generation of *35S:STK* line

Arabidopsis plants were transformed with the chimeric gene construct in which the CDS of *STK* was fused to the Cauliflower mosaic virus (CaMV) 35S promoter (Favaro et al., 2003) using the Agrobacterium tumefaciens–mediated floral dip method (Clough and Bent, 1998). Transformant plants were sown on MS medium and selected by hygromycin (20 mg/L); presence of the construct was assessed by genotyping and analysis of *STK* expression.

STK promoter constructs and plant transformation

The mutated version of the *STK* promoter (*pSTK_CArGm*) was synthesised by Twin Helix. The synthetic DNA fragment, like the wild-type version of the *STK* promoter, were cloned in pUC57-Simple (GenScript). The two fragments were digested with AccI and KpnI and cloned in pDONR207 entry clone (Invitrogen), and successively into pGWB3 binary vector containing the GUS reporter gene. Arabidopsis plants were transformed with these constructs using the *Agrobacterium*

tumefaciens–mediated floral dip method (Clough and Bent, 1998). Transformant plants were sown on MS plates and selected with hygromycin (20 mg/L); presence of the construct was assessed by PCR.

GUS staining

GUS assays were performed as described previously by Liljegren et al. (2000). The samples were mounted in lactic acid and subsequently observed using a Zeiss Axiophot D1 microscope equipped with differential interference contrast optics. Images were captured on an Axiocam MRc5 camera (Zeiss) using the Axiovision program (version 4.1).

In-situ hybridisation assay

Arabidopsis flowers were collected, fixed and embedded in paraffin as described by Huijser et al. (1992). Plant tissue sections were probed with *STK* antisense RNA, described in Brambilla et al. (2007); *STK*-sense and H4 RNA were used as controls (Fobert et al., 1994). Hybridisation and immunological detection were executed as described previously by Coen et al. (1990).

ChIP assay

ChIP assays were performed as described by Gregis et al. (2009) using for SVP-GFP the commercial antibody GFP:Living Colors full-length (Clontech), and for BPCs a polyclonal antibody as described by Simonini et al. (2012); HA antibody Anti-HA (Roche) were used as negative control in one of the experiments. 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 in Table S2. ChIP-quantitative Real-Time PCR experiments and relative enrichments were calculated as reported by Gregis et al. (2009) All the experiments were performed in three biological replicates.

ChIP-based analysis of H3K27me3 histone modification

For ChIP-based analysis of histone modifications, the following antibodies were used for immunoprecipitation: Anti-H3K27me3 Rabbit Polyclonal Antibody (Merck 07-449) and Rabbit anti-histone H3 (Sigma-Aldrich H0164). 0,8 mg of grinded and fixed material from unfertilized flowers from wild-type and *bpcV* mutant was collected. ChIP experiments were performed in a modified version of a previously reported protocol (Mizzotti et al., 2014). The quantitative Real-Time PCR assay was conducted in triplicate on four different biological replicates, with three technical replicates for each sample, and was performed in a Bio-Rad iCycler iQ optical system (software version 3.0a). Quantitative Real-Time PCR assays were performed on input and

immunoprecipitated samples and % of input was calculated. The signal obtained after precipitation with anti-H3K27me3 antibody (as indicated in Figure 6A) was normalized to actin levels. Agamous region was used as a reference as it carries the H3K27me3 mark (Li et al., 2015). Relative enrichment of *AT2G22560* was included as negative control for the H3K27me3 mark (Li et al., 2015). Sequences of oligonucleotides used for ChIP analyses are listed in Table S2.

Yeast two-hybrid assay

The two-hybrid assays were performed at 28°C in the yeast strain AH109 (Clontech). The coding sequences of *BPC4*, *BPC6* and *SVP* were cloned into pDONR207 (Life Technologies) and successively transferred to the Gateway vector GAL4 system (pGADT7 and pGBKT7; Clontech). Yeast two-hybrid assays were performed on selective yeast synthetic dropout medium lacking Leu, Trp, Ade, and His supplemented with different concentrations of 3-aminotriazole (1, 2.5, and 5 mM of 3-AT).

BiFC assay

The *BPC4*, *BPC6* and *SVP* coding sequences were first cloned into pDONR207 (Life Technologies) and subsequently transferred to the pYFPN43 and pYFPC43 vectors by Gateway recombination; while the *AP1* coding sequence was cloned into pDONR207 (Invitrogen) and then transferred to pB7RWG2, purchased from the Flanders Interuniversity Institute for Biotechnology (Gent, Belgium); the previously described formation of VERDANDI-VALKYRIE heterodimers was used as positive control, whereas VERDANDI-VERDANDI combination was used as negative control (Figure S5A; Mendes et al. (2016); all the controls are reported in Figures S5 and S6. BiFC assays were performed injecting Agrobacterium expressing viral suppressor p19/experimental constructs as described by Belda-Palazón et al. (2012). The abaxial surfaces of infiltrated tobacco (*Nicotiana benthamiana*) leaves were imaged 3 days after inoculation.

Co-Immunoprecipitation (Co-IP) Protocol

The coding sequences of *BPC4, BPC6* and *SVP* were cloned into pDONR221 and then transferred to pB7RWG2 and pB7FWG2, both purchased from the Flanders Interuniversity Institute for Biotechnology (Gent, Belgium). *Nicotiana benthamiana* leaves were infiltrated with *Agrobacterium tumefaciens*, as previously described. 4 days after infiltration, leaf disks (16 mm diameter) were collected and homogenised in 1 ml of immunoprecipitation (IP) buffer (30 mM HEPES-KOH pH 8.0, 200 mM NaCl, 60 mM KOAc, 10 mM MgOAc, 0,5% [v/v] Nonidet P-40 and proteinase inhibitor cocktail [cOmplete™, COEDTAF-RO, Roche]). Samples were incubated in ice for 15 min to allow membrane solubilisation and subjected to a centrifugation step (10 min at

16,000 *g*). Supernatants were incubated (2 h, at 4°C) with 20 µl RFP-Trap®_MA (ChromoTek) or GFP-Trap®_MA (ChromoTek). Beads were then washed 3 times for 10 min with 1 ml of IP buffer and eluted with Laemmli sample buffer. Protein samples were fractionated on SDS–PAGE (10% [w/v] acrylamide (Schägger and von Jagow, 1987) and then transferred to polyvinylidene difluoride (PVDF) membranes. Filters were immuno-decorated with specific antibodies; the Coomassie Brilliant Blue (CBB) staining of the gel was performed as loading control. The anti-GFP antibody was purchased from Thermo Fisher Scientific while the anti-RFP antibody was obtained from ChromoTek.

Gene expression analysis

Quantitative Real-Time PCR experiments were performed using cDNA obtained from inflorescences. Total RNA was extracted using lithium chloride. The Ambion TURBO DNA-free DNase kit was used to remove genomic DNA contaminations, according to the manufacturer's instructions (http://www.ambion.com/). The ImProm-IITM reverse transcription system (Promega) was used to retro-transcribe the treated RNA. Transcripts were detected using a Sybr Green Assay (iQ SYBR Green Supermix; Bio-Rad) using *UBIQUITIN* as a reference gene. Assays were done in in triplicate using a Bio-Rad iCycler iQ Optical System (software version 3.0a). The enrichments were calculated normalising the amount of mRNA against housekeeping gene fragments. The expression of different genes was analysed using specific oligonucleotides primers (Table S2).

Seed area size

Seed area size were analysed by using SMART-GRAIN software. ANOVA and post-hoc Tukey HSD (honestly significant difference) test were used for wild-type versus other genotypes comparison.

Microscopy and imaging

Images of plants, cauline and rosette leaves were acquired using a Canon EOS 6D camera whereas images of siliques were taken using a Leica® MZ 6 stereomicroscope.

Computational analyses

Annotation of selected DAP-seq peaks was performed by the means of the annotatePeaks program from the Homer suite (Heinz et al., 2010) using the reference TAIR10 annotation Intersection between peaks coordinates was performed using the bedtools intersect program (Quinlan and Hall, 2010). The "-u" option was used in order to collapse peaks showing multiple overlaps. Identification of enriched sequence motifs and identification of closely related motifs from publicly available dataset of were performed by the means of

the findMotifsGenome utility in Homer. Statistical analyses for the identification of significant overlaps were performed by using a simple statistical test based on the hyper-geometric distribution. To delineate the binding profiles of the MADS and BPC families of transcription factors, binding profiles of family members were obtained in the form of narrowpeaks files from the

http://neomorph.salk.edu/dap_web/pages/index.php repository. Narrow-peaks files were concatenated and overlapped genomic regions were merged by the means of the bedtools merge utility. Finally, candidate binding regions showing a positive hit for the majority (that is n/2+1, if profiles for n family members were available) of the members of a family were retained to form the "consensus" family profile.

Functional enrichment analyses were performed by using the web interface of the DAVID suite (Huang et al., 2009).

Accession numbers

Sequence data from this article can be found in the Arabidopsis Genome Initiative or GenBank/EMBL databases under the following accession numbers: *STK* (AT4G09960), *BPC1* (AT2G01930), *BPC2* (AT1G14685), *BPC3* (AT1G68120), *BPC4* (AT2G21240), *BPC6* (AT5G42520), *AGL24* (AT4G24540), *SVP* (AT2G22540), *AP1* (AT1G69120), *LHP1* (AT5G17690), *NETWORK*

2D (AT2G22560), *AGAMOUS* (AT4G18960), *VERDANDI* (AT5G18000), *VALKYRIE* (AT2G24690), *AC TIN7* (AT5G09810) and *UBIQUITIN* (AT4G36800).

No Data Statement has been used in the manuscript.

Author contributions

V.G. designed the research strategy and the experiments. V.G. and R.P. performed most of the experiments. F.C. performed yeast two-hybrid analyses. T.L. designed and performed CoIP experiments. V.V. did the expression analysis on *lhp1* by quantitative Real-Time PCR and performed bioinformatics analyses, M.C. supervised and performed bioinformatics analyses. I.R.V. obtained the *bpcV* multiple mutants. I.E. performed ChIP experiments on chromatin modifications and analysis on seed size area. V.G., R.P. and M.M.K interpreted data and wrote the manuscript. All authors discussed the results and commented on the manuscript. All authors read and approved the final manuscript.

Acknowledgments

We thank V. Borrelli, A. Ravasio, A.M.Piva (Università degli Studi di Milano) and S.de Folter (Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Mexico) for technical support. S. Masiero (Università degli Studi di Milano) and R. Battaglia (CREA-GB - Fiorenzuola d'Arda PC) for helpful suggestions and valuable discussions. Part of this work was carried out at NOLIMITS, an advanced imaging facility established by the Università degli Studi di Milano.

Funding

This work was supported by the Ministero dell'Istruzione, dell'Università e della Ricerca MIUR, SIR2014 MADSMEC, Proposal number RBSI14BTZR. The PhD fellowships of R.P and F.C were supported by the Doctorate School in Molecular and Cellular Biology, Università degli Studi di Milano. R.P. and F.C. were supported by H2020-MSCA-RISE-2015 ExpoSEED Proposal Number: 691109.

Conflicts of interest

The authors declare no conflicts of interest.

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.

Bartlett, A., O'Malley, R. C., Huang, S. C., Galli, M., Nery, J. R., Gallavotti, A., and Ecker, J. R. (2017). Mapping genome-wide transcription-factor binding sites using DAP-seq. *Nat. Protoc.* **12**:1659–1672.

Belda-Palazón, B., Ruiz, L., Martí, E., Tárraga, S., Tiburcio, A. F., Culiáñez, F., Farràs, R., Carrasco, P., and Ferrando, A. (2012). Aminopropyltransferases Involved in Polyamine Biosynthesis Localize Preferentially in the Nucleus of Plant Cells. *PLoS One* **7**:e46907.

Berger, N., and Dubreucq, B. (2012). Evolution goes GAGA: GAGA binding proteins across kingdoms. *Biochim. Biophys. Acta - Gene Regul. Mech.* **1819**:863–868.

Berger, N., Dubreucq, B., Roudier, F., Dubos, C., and Lepiniec, L. (2011). Transcriptional regulation of Arabidopsis LEAFY COTYLEDON2 involves RLE, a cis-element that regulates trimethylation of histone H3 at lysine-27. *Plant Cell* **23**:4065–78.

Bouyer, D., Roudier, F., Heese, M., Andersen, E. D., Gey, D., Nowack, M. K., Goodrich, J., Renou, J.-P., Grini, P. E., Colot, V., et al. (2011). Polycomb Repressive Complex 2 Controls the Embryo-to-Seedling Phase Transition. *PLoS Genet.* **7**:e1002014.

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.

Calonje, M., Sanchez, R., Chen, L., and Sung, Z. R. (2008). EMBRYONIC FLOWER1 Participates in Polycomb Group–Mediated *AG* Gene Silencing in *Arabidopsis*. *Plant Cell* **20**:277–291.

Chanvivattana, Y., Bishopp, A., Schubert, D., Stock, C., Moon, Y.-H., Sung, Z. R., and Goodrich, J. (2004). Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. *Development* **131**:5263–76.

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.

Coen, E. S., Romero, J. M., Doyle, S., Elliott, R., Murphy, G., and Carpenter, R. (1990). floricaula: a homeotic gene required for flower development in antirrhinum majus. *Cell* **63**:1311–22.

de Folter, S., Immink, R. G. H., Kieffer, M., Parenicová, 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 ONLINE* **17**:1424–1433.

Derkacheva, M., Steinbach, Y., Wildhaber, T., Mozgová, I., Mahrez, W., Nanni, P., Bischof, S., Gruissem, W., and Hennig, L. (2013). Arabidopsis MSI1 connects LHP1 to PRC2 complexes. *EMBO J.* **32**:2073–2085.

Doughty, J., Aljabri, M., and Scott, R. J. (2014). Flavonoids and the regulation of seed size in *Arabidopsis*. *Biochem. Soc. Trans.* **42**:364–369.

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.

Farkas, G., Gausz, J., Galloni, M., Reuter, G., Gyurkovics, H., and Karch, F. (1994). The Trithorax-like gene encodes the Drosophila GAGA factor. *Nature* **371**:806–808.

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**.

Fobert, P. R., Coen, E. S., Murphy, G. J., and Doonan, J. H. (1994). Patterns of cell division revealed by transcriptional regulation of genes during the cell cycle in plants. *EMBO J.* **13**:616–24.

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.* **56**:891–902.

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.

Gregis, V., Andrés, F., Sessa, A., Guerra, R. F., Simonini, S., Mateos, J. L., Torti, S., Zambelli, F., Prazzoli, G. M., Bjerkan, K. N., et al. (2013). Identification of pathways directly regulated by SHORT VEGETATIVE PHASE during vegetative and reproductive development in Arabidopsis. *Genome Biol.* **14**:R56.

Gustafson-Brown, C., Savidge, B., and Yanofsky, M. F. (1994). Regulation of the arabidopsis floral homeotic gene APETALA1. *Cell* **76**:131–43.

Hecker, A., Brand, L. H., Peter, S., Simoncello, N., Kilian, J., Harter, K., Gaudin, V., and Wanke, D. (2015). The Arabidopsis GAGA-Binding Factor BASIC PENTACYSTEINE6 Recruits the POLYCOMB-REPRESSIVE COMPLEX1 Component LIKE HETEROCHROMATIN PROTEIN1 to GAGA DNA Motifs.

Plant Physiol. **168**:1013–1024.

Heinz, S., Benner, C., Spann, N., Bertolino, E., Lin, Y. C., Laslo, P., Cheng, J. X., Murre, C., Singh, H., and Glass, C. K. (2010). Simple Combinations of Lineage-Determining Transcription Factors Prime cis-Regulatory Elements Required for Macrophage and B Cell Identities. *Mol. Cell* **38**:576–589.

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.

Horard, B., Tatout, C., Poux, S., and Pirrotta, V. (2000). Structure of a polycomb response element and in vitro binding of polycomb group complexes containing GAGA factor. *Mol. Cell. Biol.* **20**:3187–97.

Huang, D. W., Sherman, B. T., and Lempicki, R. A. (2009). Systematic and integrative analysis of large gene lists using DAVID bioinformatics resources. *Nat. Protoc.* **4**:44–57.

Immink, R. G. H., Gadella, T. W. J., Ferrario, S., Busscher, M., and Angenent, G. C. (2002). Analysis of MADS box protein-protein interactions in living plant cells. *Proc. Natl. Acad. Sci.* **99**:2416–2421.

Kinoshita, T., Harada, J. J., Goldberg, R. B., and Fischer, R. L. (2001). Polycomb repression of flowering during early plant development. *Proc. Natl. Acad. Sci. U. S. A.* **98**:14156–61.

Kooiker, M., CA, A., A, L., PS, M., L, F., MM, K., and L., C. (2005). BASIC PENTACYSTEINE1, a GA Binding Protein That Induces Conformational Changes in the Regulatory Region of the Homeotic Arabidopsis Gene SEEDSTICK. *PLANT CELL ONLINE* **17**:722–729.

Kotake, T., Takada, S., Nakahigashi, K., Ohto, M., and Goto, K. (2003). Arabidopsis TERMINAL FLOWER 2 Gene Encodes a Heterochromatin Protein 1 Homolog and Represses both FLOWERING LOCUS T to Regulate Flowering Time and Several Floral Homeotic Genes. *Plant Cell Physiol.* **44**:555–564.

Lafos, M., Kroll, P., Hohenstatt, M. L., Thorpe, F. L., Clarenz, O., and Schubert, D. (2011). Dynamic Regulation of H3K27 Trimethylation during Arabidopsis Differentiation. *PLoS Genet.* **7**:e1002040.

Larsson, A. S., Landberg, K., and Meeks-Wagner, D. R. (1998). The TERMINAL FLOWER2 (TFL2) gene controls the reproductive transition and meristem identity in Arabidopsis thaliana. *Genetics* **149**:597–605.

Lewis, E. B. (1978). A gene complex controlling segmentation in Drosophila. *Nature* **276**:565–570.

Li, C., Chen, C., Gao, L., Yang, S., Nguyen, V., Shi, X., Siminovitch, K., Kohalmi, S. E., Huang, S., Wu, K., et al. (2015). The Arabidopsis SWI2/SNF2 Chromatin Remodeler BRAHMA Regulates Polycomb Function during Vegetative Development and Directly Activates the Flowering Repressor Gene SVP. *PLOS Genet.* **11**:e1004944.

Liljegren, S. J., Ditta, G. S., Eshed, 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., Xi, W., Shen, L., Tan, C., and Yu, H. (2009). Regulation of Floral Patterning by Flowering Time Genes. *Dev. Cell* **16**:711–722.

Lodha, M., Marco, C. F., and Timmermans, M. C. P. (2013). The ASYMMETRIC LEAVES complex maintains repression of KNOX homeobox genes via direct recruitment of Polycomb-repressive complex2. *Genes Dev.* **27**:596–601.

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**:115–121.

Makarevich, G., Leroy, O., Akinci, U., Schubert, D., Clarenz, O., Goodrich, J., Grossniklaus, U., and

Köhler, C. (2006). Different Polycomb group complexes regulate common target genes in Arabidopsis. *EMBO Rep.* **7**:947–952.

Malley, R. C. O., Carol, S., Song, L., Galli, M., Ecker, J. R., Malley, R. C. O., Huang, S. C., Song, L., Lewsey, M. G., Bartlett, A., et al. (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape Resource Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**:1280–1292.

Martinez, G. J., and Rao, A. (2012). Cooperative Transcription Factor Complexes in Control. *Science (80-.).* **338**:891–892.

Meister, R. J., Williams, L. A., Monfared, M. M., Gallagher, T. L., Kraft, E. A., Nelson, C. G., and Gasser, C. S. (2004). Definition and interactions of a positive regulatory element of the *Arabidopsis INNER NO OUTER* promoter. *Plant J.* **37**:426–438.

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., 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.

Monfared, M. M., Simon, M. K., Meister, R. J., Roig-Villanova, I., Kooiker, M., Colombo, L., Fletcher, J. C., and Gasser, C. S. (2011). Overlapping and antagonistic activities of BASIC PENTACYSTEINE genes affect a range of developmental processes in Arabidopsis. *Plant J.* **66**:1020–1031.

Mu, Y., Zou, M., Sun, X., He, B., Xu, X., Liu, Y., Zhang, L., and Chi, W. (2017). BASIC PENTACYSTEINE proteins repress Abscisic Acid INSENSITIVE 4 expression via direct recruitment of the polycomb-repressive complex 2 in arabidopsis root development. *Plant Cell Physiol.* **58**:607–621.

O'Malley, R. C., Huang, S. C., Song, L., Lewsey, M. G., Bartlett, A., Nery, J. R., Galli, M., Gallavotti, A., and Ecker, J. R. (2016). Cistrome and Epicistrome Features Shape the Regulatory DNA Landscape. *Cell* **165**:1280–1292.

Pelaz, S., Gustafson-Brown, C., Kohalmi, S. E., Crosby, W. L., and Yanofsky, M. F. (2002). APETALA1 and SEPALLATA3 interact to promote flower development. *Plant J.* **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.

Quinlan, A. R., and Hall, I. M. (2010). BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**:841–842.

Roscoe, T. J., Vaissayre, V., Paszkiewicz, G., Clavijo, F., Kelemen, Z., Michaud, C., Lepiniec, L., Dubreucq, B., Zhou, D.-X., and Devic, M. (2019). Regulation of *FUSCA3* Expression During Seed Development in Arabidopsis. *Plant Cell Physiol.* **60**:476–487.

Sangwan, I., and O'Brian, M. R. (2002). Identification of a soybean protein that interacts with GAGA element dinucleotide repeat DNA. *Plant Physiol.* **129**:1788–94.

Santi, L., Wang, Y., Stile, M. R., Berendzen, K., Wanke, D., Roig, C., Pozzi, C., Müller, K., Müller, J., Rohde, W., et al. (2003). The GA octodinucleotide repeat binding factor BBR participates in the transcriptional regulation of the homeobox gene *Bkn3*. *Plant J.* **34**:813–826.

Schägger, H., and von Jagow, G. (1987). Tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis for the separation of proteins in the range from 1 to 100 kDa. *Anal. Biochem.*

166:368–79.

Schruff, M. C., Spielman, M., Tiwari, S., Adams, S., Fenby, N., and Scott, R. J. (2006). The AUXIN RESPONSE FACTOR 2 gene of Arabidopsis links auxin signalling, cell division, and the size of seeds and other organs. *Development* **133**:251–61.

Shanks, C. M., Hecker, A., Cheng, C.-Y., Brand, L., Collani, S., Schmid, M., Schaller, G. E., Wanke, D., Harter, K., and Kieber, J. J. (2018). Role of *BASIC PENTACYSTEINE* transcription factors in a subset of cytokinin signaling responses. *Plant J.* **95**:458–473.

Simonini, S., and Kater, M. M. (2014). Class I BASIC PENTACYSTEINE factors regulate HOMEOBOX genes involved in meristem size maintenance. *J. Exp. Bot.* **65**:1455–1465.

Simonini, S., Roig-Villanova, I., Gregis, V., Colombo, B., Colombo, L., and Kater, M. M. (2012). BASIC PENTACYSTEINE Proteins Mediate MADS Domain Complex Binding to the DNA for Tissue-Specific Expression of Target Genes in *Arabidopsis*. *Plant Cell* **24**:4163–4172.

Theune, M. L., Bloss, U., Brand, L. H., Ladwig, F., and Wanke, D. (2019). Phylogenetic Analyses and GAGA-Motif Binding Studies of BBR/BPC Proteins Lend to Clues in GAGA-Motif Recognition and a Regulatory Role in Brassinosteroid Signaling. *Front. Plant Sci.* **10**:466.

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.-L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R. A., Coupland, G., and Colot, V. (2007). Arabidopsis TFL2/LHP1 Specifically Associates with Genes Marked by Trimethylation of Histone H3 Lysine 27. *PLoS Genet.* **3**:e86.

Wang, H., Liu, C., Cheng, J., Liu, J., Zhang, L., and He, C. (2016). Arabidopsis Flower and Embryo Developmental Genes are Repressed in Seedlings by Different Combinations of Polycomb Group Proteins in Association with Distinct Sets of Cis-regulatory Elements Advance Access published 2016, doi:10.1371/journal.pgen.1005771.

Wang, X., Paucek, R. D., Gooding, A. R., Brown, Z. Z., Ge, E. J., Muir, T. W., and Cech, T. R. (2017). Molecular analysis of PRC2 recruitment to DNA in chromatin and its inhibition by RNA. *Nat. Struct. Mol. Biol.* **24**:1028–1038.

Wanke, D., Hohenstatt, M. L., Dynowski, M., Bloss, U., Hecker, A., Elgass, K., Hummel, S., Hahn, A., Caesar, K., Schleifenbaum, F., et al. (2011). Alanine zipper-like coiled-coil domains are necessary for homotypic dimerization of plant GAGA-factors in the nucleus and nucleolus. *PLoS One* **6**:e16070.

Wu, J., Petrella, R., Dowhanik, S., Gregis, V., and Gazzarrini, S. (2019). Spatiotemporal restriction of FUSCA3 expression by class I BPC promotes ovule development and coordinates embryo and endosperm growth. *bioRxiv* Advance Access published April 19, 2019, doi:10.1101/612408.

Xiao, J., Jin, R., Yu, X., Shen, M., Wagner, J. D., Pai, A., Song, C., Zhuang, M., Klasfeld, S., He, C., et al. (2017). Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in Arabidopsis. *Nat. Genet.* **49**:1546–1552.

Yang, J., Lee, S., Hang, R., Kim, S.-R., Lee, Y.-S., Cao, X., Amasino, R., and An, G. (2013). OsVIL2 functions with PRC2 to induce flowering by repressing *O s LFL 1* in rice. *Plant J.* **73**:566–578.

Yoshida, N., Yanai, Y., Chen, L., Kato, Y., Hiratsuka, J., Miwa, T., Sung, Z. R., and Takahashi, S. (2001). EMBRYONIC FLOWER2, a novel polycomb group protein homolog, mediates shoot development and flowering in Arabidopsis. *Plant Cell* **13**:2471–81.

Yu, H., Ito, T., Wellmer, F., and Meyerowitz, E. M. (2004). Repression of AGAMOUS-LIKE 24 is a crucial step in promoting flower development. *Nat. Genet.* **36**:157–161.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y. V, Pellegrini, M., Goodrich, J., and Jacobsen, S. E. (2007). Whole-Genome Analysis of Histone H3 Lysine 27 Trimethylation in Arabidopsis. *PLoS Biol.*

5:e129.

Figures.

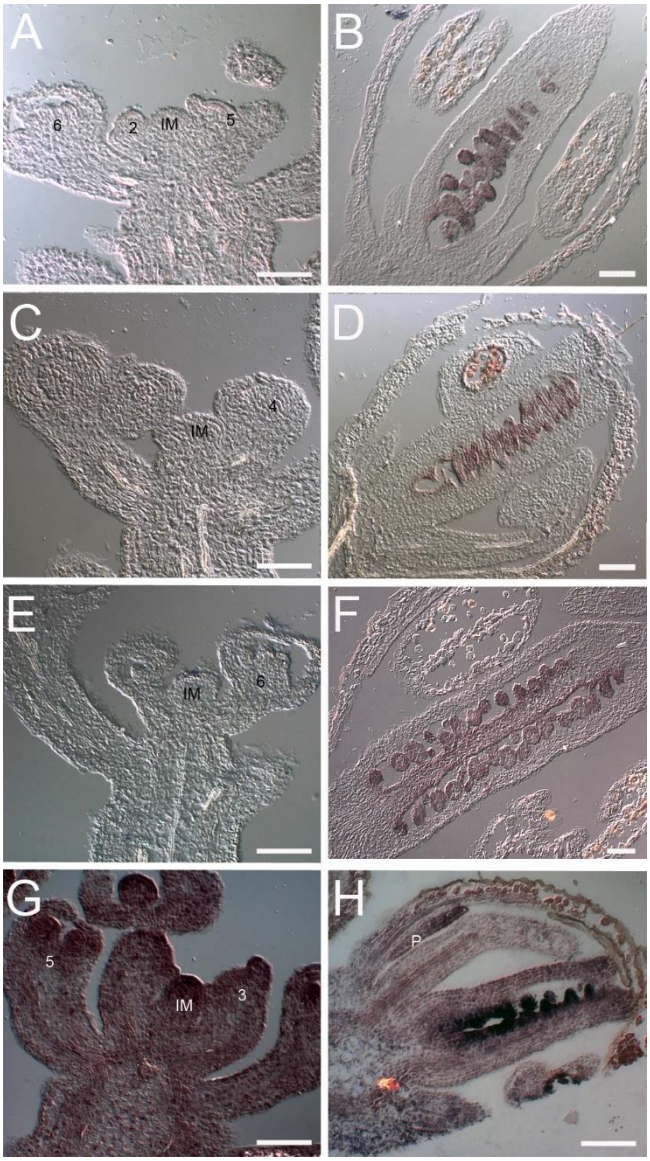


Figure 1. Mutation of BPCs affects *STK* expression in the flower.

In-situ hybridisation on wild-type [(A) and (B)], *bpc1-2 bpc2 bpc3* [(C) and (D)], *bpc4 bpc6* [(E) and (F)] and *bpcV* [(G) and (H)] inflorescences using a *STK*-specific antisense probe (Brambilla et al., 2007). IM: inflorescence meristem; P: petal; numbers represent flower stages. Scale bars=50 µm.

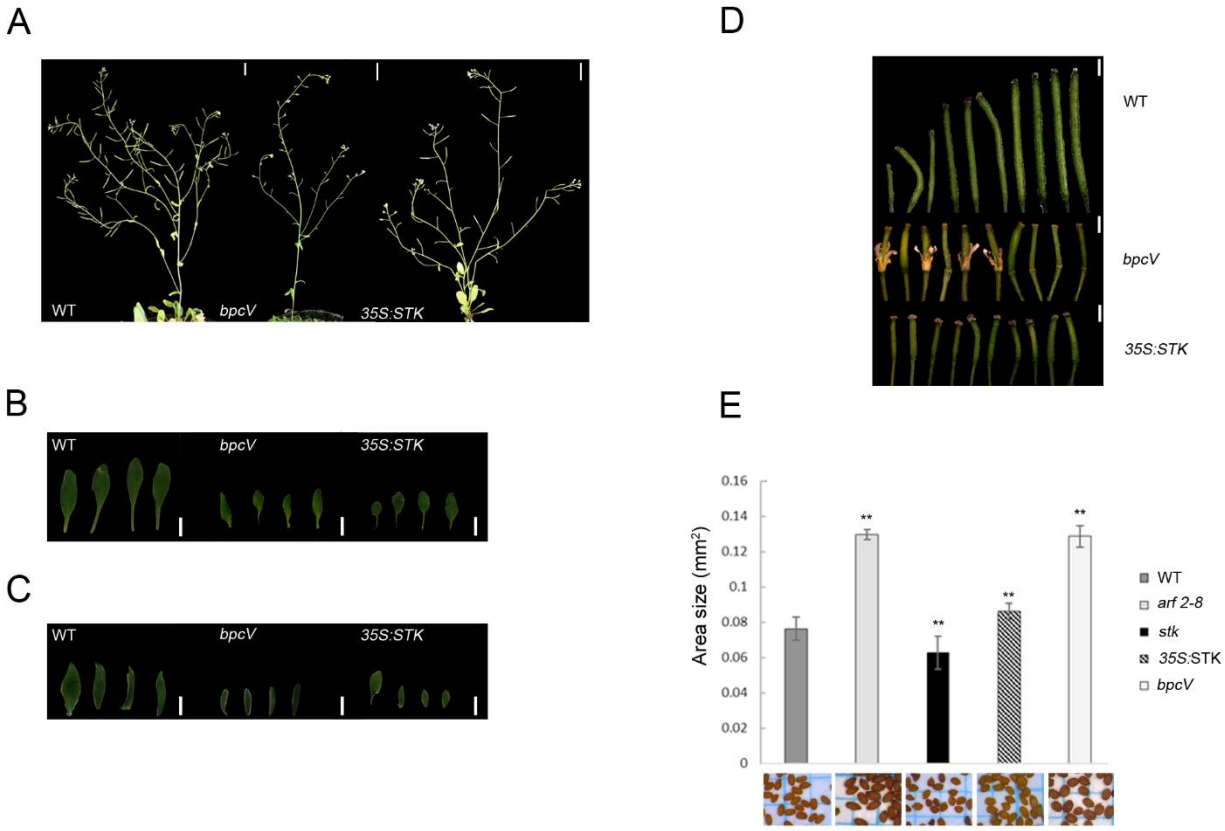


Figure 2. Overexpression of *STK* affects vegetative and reproductive development.

(A) From left to right: wild-type, *bpcV* and *35S:*STK plants; plants were photographed six weeks after sowing; scale bars=1 cm. (B) Rosette leaves morphology and length: rosette leaves in wild-type, *bpcV* and *35S:*STK (from left to right); (C) Cauline morphology and length: cauline leaves in wild-type, *bpcV* and *35S:*STK (from left to right); scale bars= 0.5 cm. (D) Fruit morphology and length in wild type, *bpcV* and *35S*:STK (from top to bottom); scale bars=1.5 mm.(C) (D) Average seeds area size of wild-type, *arf2-8* (Schruff et al., 2006), *stk* (Pinyopich et al., 2001), *35S*:STK and *bpcV*; error bars represent the standard error mean of replicates; ANOVA and post-hoc Tukey HSD (honestly significant difference) test were used, **P < 0.01 for wild-type versus other genotypes comparison. In the lower row, seeds of the analysed genotypes are shown.

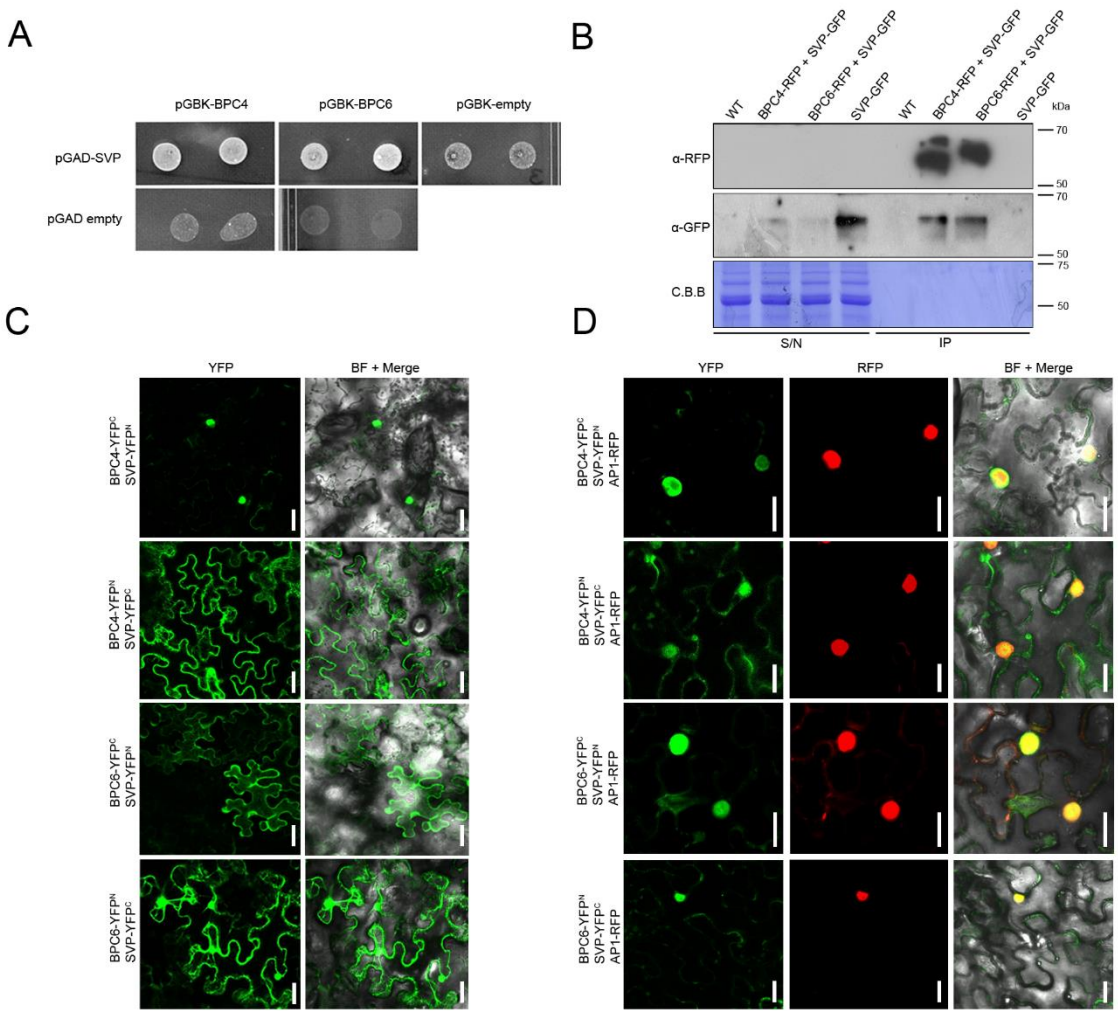


Figure 3. Class II BPCs interact with SVP *in vivo*.

(A) Yeast two-hybrid interaction assay for SVP and BPCs of class II: positive interactions on selective media –W-L-H +5mM 3-AT.

(B) Co-immunoprecipitation assays. *Nicotiana benthamiana* leaves were infiltrated with constructs carrying SVP-GFP together with BPC4-RFP and BPC6-RFP, as described in experimental procedures. Immunoprecipitation step was performed using RFP-trap on total protein leaf extract. Samples were probed with GFP and RFP antibodies. S/N: supernatant; IP: immunoprecipitation.

(C) Bi-molecular fluorescence complementation (BiFC) assay. *Nicotiana benthamiana* epidermis cells were transiently transformed with the indicated YN and YC fusions. In the first and the second column yellows fluorescence and the merging in the bright field were shown, respectively.

(D) Bi-molecular fluorescence complementation (BiFC) assay. *N.benthamiana* epidermis cells were transiently transformed with the indicated YN and YC fusions and AP1-RFP construct. In the

first, the second and the third column yellow fluorescence, red fluorescence and the merging between the two channels in the bright field were shown, respectively. Scale bars=40 µm.

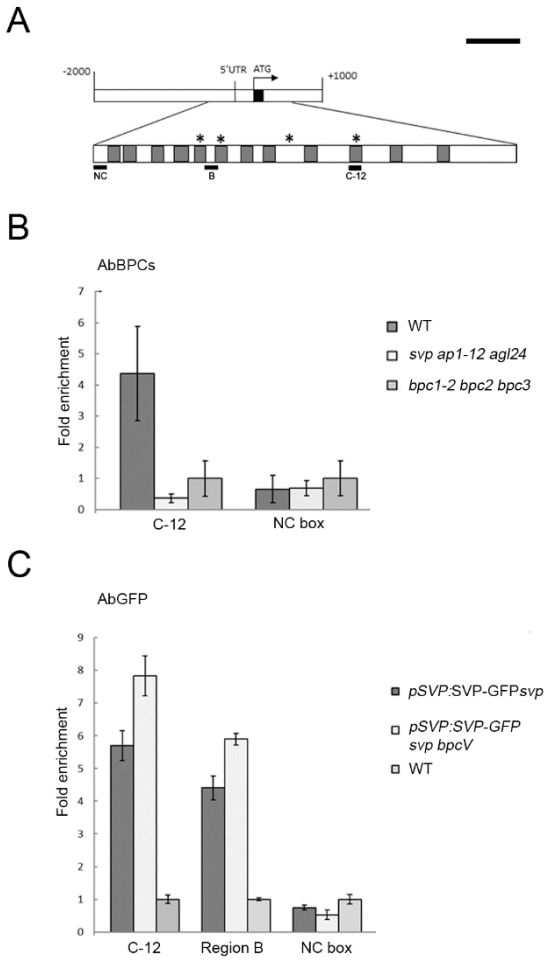


Figure 4. ChIP experiments on different mutant backgrounds.

ChIP experiments on different mutant backgrounds. (A) Schematic diagram of the *STK* locus indicating the regions analysed by chromatin immunoprecipitation (ChIP; black bars). Black boxes, exons; white boxes, promoters and introns; asterisks, C-boxes; grey boxes, CArG-boxes; scale bar=500 bp. (B) Quantitative Real-Time PCR analysis of ChIP assay using chromatin extracted from *svp ap1-12 agl24*, wild-type (as a positive control), and *bpc1-2 bpc2 bpc3* (as a negative control) testing the C-12 region and NC box. Antibodies against BPCs of class I were used. (C) Quantitative Real-time PCR analysis of ChIP assay using chromatin extracted from *pSVP*:SVP-GFP *svp bpcV*,

pSVP:SVP-GFP *svp* (as a positive control) and wild-type (as a negative control), testing Region B, C-12 and NC box. 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 significantly enriched compared with the controls in at least three independent experiments.

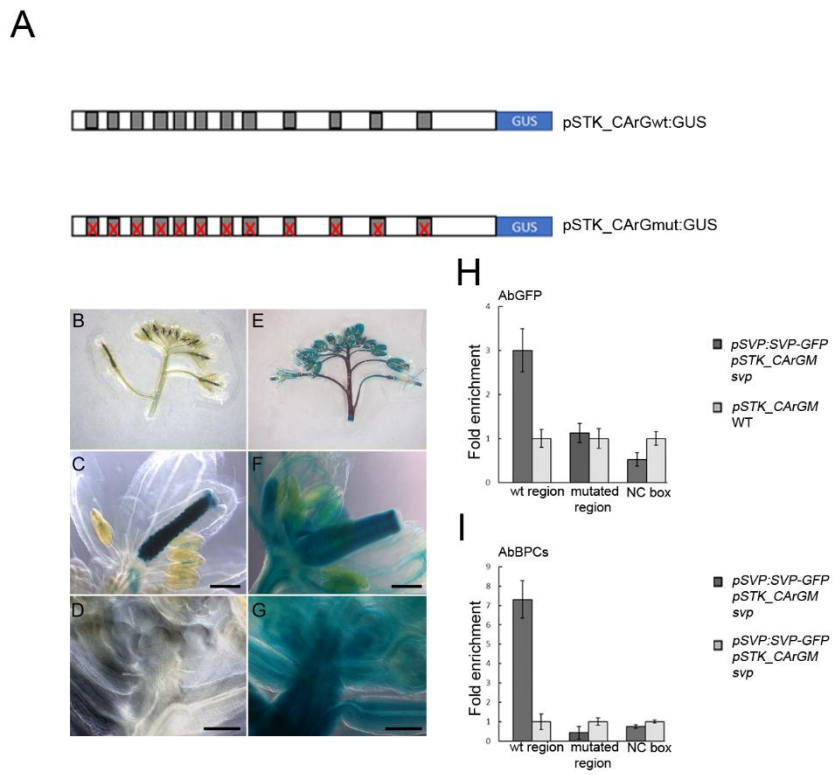


Figure 5. Mutation of CArG-boxes avoids SVP and BPCs binding to *STK* promoter.

(A)Schematic representation of the *STK* promoter versions generated: dark grey squares represent CArG-boxes wild-type and mutated (crossed). (B)-(G) GUS staining on inflorescences from *pSTK:*GUSwt (B-D) and *pSTK-CArGm:*GUS (E-G): whole inflorescence [(B) and (E)]; mature flower [(C) and (F); inflorescence meristem (IM), floral meristems (FM) and young flowers [(D) and (G)]; scale bars in (C), (D), (F) and (G)=100 µm. (H) Quantitative Real-Time PCR analysis of ChIP assay using chromatin extracted from *pSVP:SVP-GFP pSTK-CArGm svp* and *pSTK-CArGm* as a negative control, testing Region B and NC box (indicated in Figure 4). For the IP, antibodies against GFP have been used. (I) Quantitative Real-Time PCR analysis of ChIP assay using chromatin extracted from *pSVP:SVP-GFP pSTK-CArG*m testing Region B and NC box. For the IP, antibodies against Class I BPCs have been used; for negative control commercial antibodies against HA was

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 significantly enriched compared to the controls in at least three independent experiments.

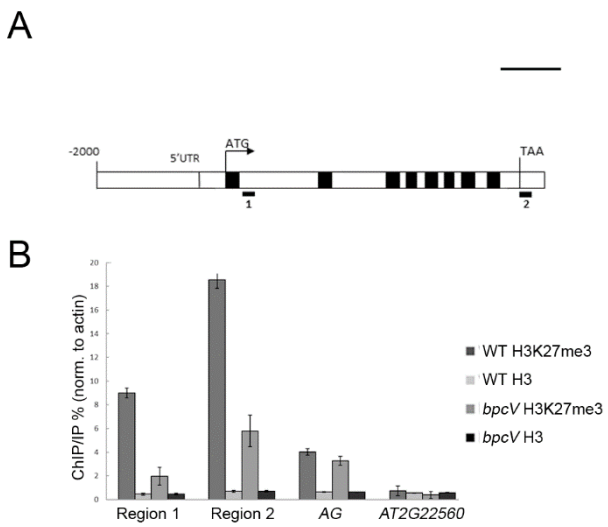


Figure 6. Epigenetic regulation of *STK.*

(A) Schematic representation of the *STK* genomic region tested in ChIP assay. Grey boxes indicate exonic regions. Black bars indicate the regions analysed by chromatin immunoprecipitation (ChIP); region1 is located in the H3k27me3 – enriched region published by (Li et al. 2015) spanning -2627 upstream *STK*-transcriptional start site to +2050 pb downstream *STK*-transcriptional start site, whereas region 2 is localized 3 pb downstream the stop codon of the gene. Black arrow indicates the *STK*-transcription start site. Scale bar= 500 bp.

(B) ChIP-quantitative Real-Time PCR determining the levels of H3K27me3 across the *STK* locus. Quantitative Real-Time PCR quantification of *STK* sequences in precipitated chromatin was used to infer the methylation of histone H3 at lysine 27 (H3K27me3) and total histone H3 representation (histone H3 density). Ct values were used to calculate the IP/IN signal. ChIP enrichments are presented as the percentage (%) of bound/input signal normalized to actin levels in the relative regions. We tested the efficiency of IP on histone modifications by quantifying the presence of the H3K27me3 mark in *AG* region which carries the mark H3k27me3, reported in Li et al. (2015). H3K27me3 mark in AT2G22560 was used as negative

control for H3K27me3 mark (Li et al. 2015). Error bars indicate standard deviations from four biological replicates.

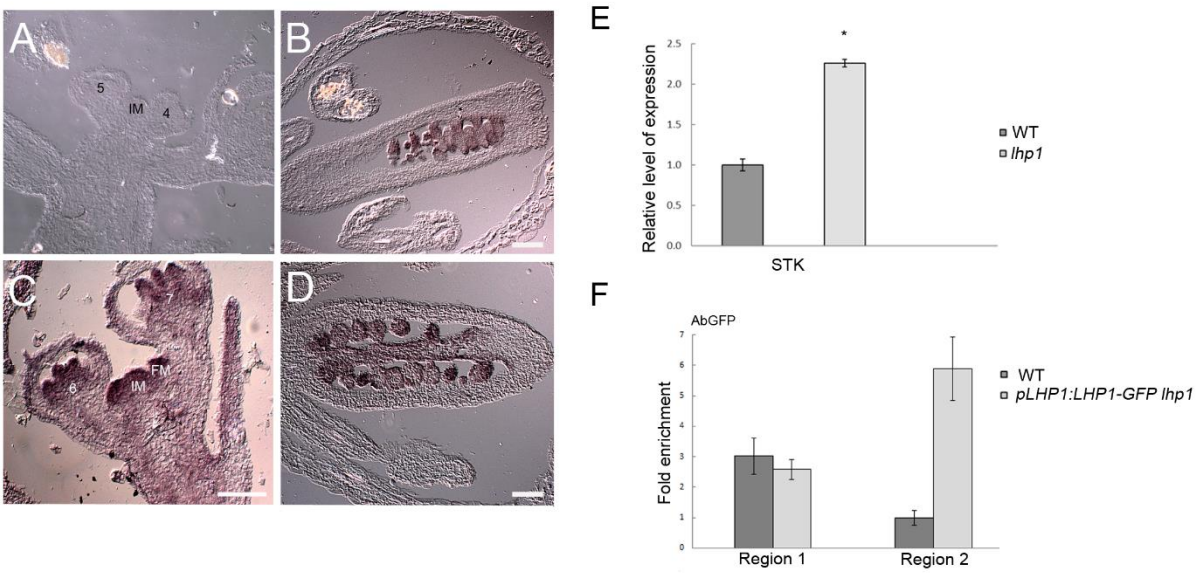


Figure 7. LHP1 directly regulates *STK* during flower development.

(A)-(D) *In-situ* hybridisation on wild-type [(A) and (B)] and *lhp1* inflorescences using a *STK*-specific antisense probe (Brambilla et al., 2007). IM: inflorescence meristem; p: petal; numbers represent flower stages; scale bars=50 µm.

(E) Expression analysis of *STK* by quantitative Real-Time PCR in *lhp1* and wild-type inflorescences. The expression of *STK* was normalized to that of ubiquitin and the expression level in wild-type was set to 1. Asterisk indicates $P < 0.05$ in a Student's t-test.

(F) Quantitative Real-Time PCR analysis of ChIP assay using chromatin extracted from *pLHP1:*LHP1-GFP (Kotake et al, 2003) and wild-type (as a negative control), testing the Region 1 and Region 2 (Figure 6A). For the IP, commercial antibodies against SVP 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 significantly enriched compared with the controls in at least three independent experiments.

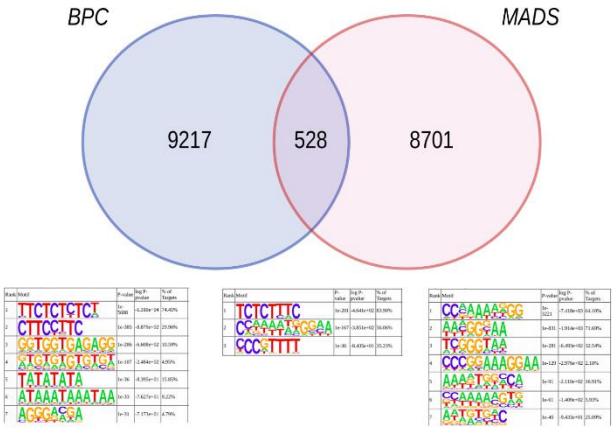


Figure 8. Analysis of BPC and MADS-box transcription factor families binding sites.

Venn diagram displaying the number of DAP-seq peaks and the common number of peaks associated to the BPC and MADS transcription factor families according to our analysis of the data by O'Malley et al (see Experimental procedures). Enriched motifs, as recovered by Homer (p-value ≤ 1e-30), are displayed underneath.

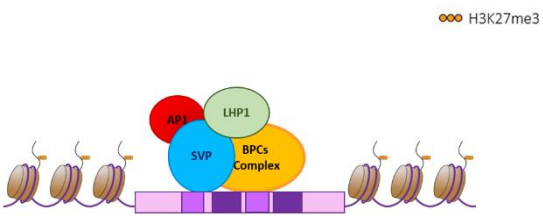


Figure 9. Model of the protein complex formed to represses gene expression during flower development.

BPCs and SVP bind C-boxes (in dark purple) and CArG-boxes (in light purple) respectively, and recruit LHP1 to a subset of gene loci.

Supplementary information (SI)

Supplementary tables.

CArG-box 1	Wild-type version	CCATCTTTGT
	Mutated version	CTGCGCTTGT
CArG-box 2	Wild-type version	CTATATATGC
	Mutated version	CTATGCGCAC
CArG-box 3	Wild-type version	ACAATTATGA
	Mutated version	ACAGCCACAA
CArG-box 4	Wild-type version	CCTTATTTTT
	Mutated version	CTCCGCTTTT
CArG-box 5	Wild-type version	CCATGAAAGA
	Mutated version	CTGCAGAAGA
CArG-box 6	Wild-type version	CCTTTCTTGT
	Mutated version	CCTTTTCCGT
CArG-box 7	Wild-type version	GCTAAAGTGG
	Mutated version	GCTAGGACAG
CArG-box 8	Wild-type version	CCAAATCTGT
	Mutated version	CCAAGCTCAT
CArG-box 9	Wild-type version	GTAATAATGT
	Mutated version	GTAACGGCAT
CArG-box 10	Wild-type version	CCATATTTCC
	Mutated version	CCATATGGTT
CArG-box 11	Wild-type version	CCAATTTTTT
	Mutated version	TTGGTTTTTT

Table S1. MADS-domain consensus regions identified in *STK* promoter and the designed mutated versions.

Genotyping	
BPC1 fw	TAGCGATCTTCTCATCGAAGC
BPC1 rv	AGTCGTACAACAAGCGGATTG
bpc 1-2 fw	TAGCGATCTTCTCATCGAAGC
bpc 1-2 rv	AGTCGTACAACAAGCGGATTG
BPC2 fw	AGCCCGGGCATGGATGACGATGGGTTTCG
BPC2 rv	AGTCGTACAACAAGCGGATTG
bpc2 fw	AGCCCGGGCATGGATGACGATGGGTTTCG
bpc2 rv	TAGCGATCTTCTCATCGAAGC
BPC3 fw	GGGGACCACTTTGTACAAGAAAGCTGGGTTTATCTGATGGTGACGAACTTATTGG
BPC3 rv	GAGTACAAAGAGAGAGAAGTCC
BPC4 fw	CCCCAGCATCAGATTAAGGA
BPC4 rv	CGTGCCTAGCCCAATAGTCT
bpc4 fw	CCCCAGCATCAGATTAAGGA
bpc4 rv	TAGCATCTGAATTTCATAACCAATCTCGATACAC
BPC6 fw	ATCTCAAATGGATGATGGTGG
BPC6 rv	TTCCCCATTTGTAGCACTGTC

bpc6 fw	ATCTCAAATGGATGATGGTGG
bpc6 rv	GCGTGGACCGCTTGCTGCAACT
SVP fw	GACCCACTAGTTATCAGCTCAG
SVP rv	AAGTTATGCCTCTCTAGGAC
svp-41 fw	GACCCACTAGTTATCAGCTCAG
svp-41 rv	AAGTTATGCCTCTCTAGGTT
AP1 fw	GGAGAGGGAAAAAATTCTTAGGGCTCCAC
AP1 rv	ATCATGACATCATGTAACCATCACTAACAGC
ap1-12 fw	TGGTTCTGCTGATCCCACTGCTCATA
ap1-12 rv	CCATACAGGAGCAAAACAGCATG
AGL24 fw	GATCCACCTTCTACTCATCTCC
AGL24 rv	CCACACACATGAAATAGATGATC
agl24-2 fw	GATCCACCTTCTACTCATCTCC
agl24-2 rv	GAGCGTCGGTCCCCACACTTCTATAC
Expression analysis by qRT-PCR	
UBI fw	CTGTTCACGGAACCCAATTC
UBI rv	GGAAAAAGGTCTGACCGACA
STK fw	ACGCGCAGAAAAGGGAGATTGAGC
STK rv	TGTCGGGATCAGAGTAAGAACCTCC
qRT-PCR ChIP assay	
ACTIN7 fw	CGTTTCGCTTTCCTTAGTGTTAGCT
ACTIN7 rv	AGCGAACGGATCTAGAGACTCACCTTG
gSTK region B fw	CTTTATAAAGGAGAAAGAAAGAGA
gSTK region B rv	CAAAGATGGGAACTTGATGAG
gSTK C-12 fw	TATCAATTTGATTTGTTTTCTCTCT
gSTK C-12 rv	CAAAGATGGGAACTTGATGAG
pSTK-CArGwt fw	TCTCTGCTAGATTCTCTTTC
pSTK-CArGwt rv	GGGAAACACAAGAAACATTA
pSTK-CArGm fw	TCTCTGCTAGATTCTCTTTC
pSTK-CArGm rv	GGGAAACACAAGAAATGCCG
gSTK region 1 fw	TCTCTGCTAGATTCTCTTTC
gSTK region 1 rv	GGGAAACACAAGAAACATTA
gSTK region 2 fw	CGTCTGCGAAAAACCGAGCT
gSTK region 2 rv	GGACCAATACCTTCATTGTACTTTGAA
AGAMOUS fw	ATGCTGAAGTCGCACTCATCGTCT
AGAMOUS rv	GAGCACGAGAAGAAGAAGAAACCTG
AT2G22560 fw	TAATGTCCCTAATGTTCCCAAA
AT2G22560 fw	CTCAGGCTTACTCAAACCCGA
NC box fw	CCTTATTTTGTTTCTTTTTACC
NC box rv	CTAAGATTGCGAGCAGTAG
Cloning yeast hybrid, BiFC and CoIP constructs	
BPC4 AttB1 fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGAGAATGGTGGTCAGTA
BPC4 AttB2 rv	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTACTTGATAGTGATGTAGCGG
BPC6 AttB1 fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGATGATGGTGGGCA
BPC6 AttB2 rv	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCATTTAATCGTAATGTAGCGG
SVP AttB1 fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCGAGAGAAAAGATTC
SVP AttB2 rv	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTAACCACCATACGGTAAGC
AP1 AttB1 fw	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGGAGGGGTAGGGTTCA
AP1 NO STOP AttB2 rv	GGGGACCACTTTGTACAAGAAAGCTGGGTTCATGCGGCGAAGCAGCCAAGG
BPC4 NO STOP AttB2 rv	GGGGACCACTTTGTACAAGAAAGCTGGGTCCTTGATAGTGATGTAGCGG
BPC6 NO STOP AttB2 rv	GGGGACCACTTTGTACAAGAAAGCTGGGTCTTTAATCGTAATGTAGCGG

SVP NO STOP AttB2 rv	GGGGACCACTTTGTACAAGAAAGCTGGGTCACCACCATACGGTAAGCCG

Table S2. Sequences of oligonucleotides used in this manuscript.

Data S1. p-values for the intersection of DAP-seq peaks of TFs families as defined by (Malley et al., 2016) with DAP-seq peaks of the BPCs family of transcription factors. Column1: transcription factor family. Column 2: p-value for the overlap with DAP-seq peaks of BPC family transcription factors. Column 3: Bonferroni adjusted p-value.

Data S2. Lists of identified genome-wide binding locations for all the MADS-box and BPCs factors, coincident DAP-seq peaks and functional enrichment analyses of associated target genes.

Supplementary Figures.

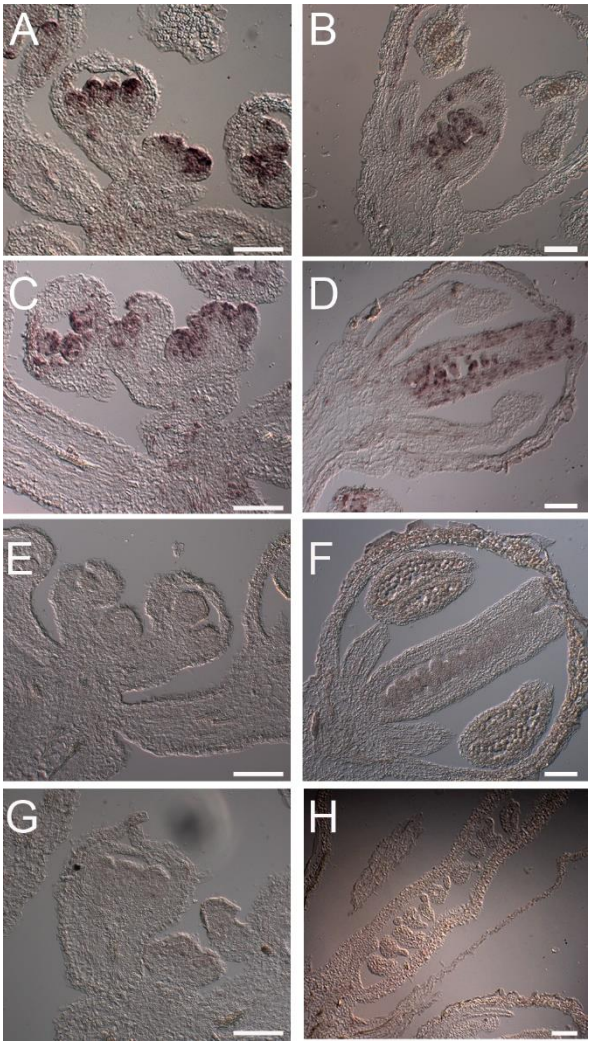


Figure S1.

In-situ hybridisation on wild-type [(A) and (B); (E) and (F)] and *bpcV* [(C) and (D); (G) and (H)] inflorescences using a histone H4 probe (Fobert et al., 1994) [(A)-(D)] and a specific *STK* sense probe [(E)-(H)] (Brambilla et al., 2007). Scale bars=50 µm.

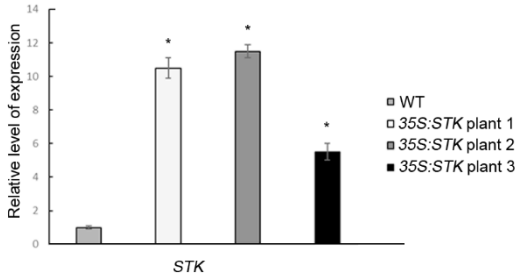


Figure S2.

Expression analysis of *STK* by quantitative Real-Time PCR in three different *35S:*STK T1 lines and wild-type inflorescences. The expression of *STK* was normalized to that of ubiquitin and the expression level in wild-type was set to 1. Asterisk indicates P < 0.05 in a Student's t-test.

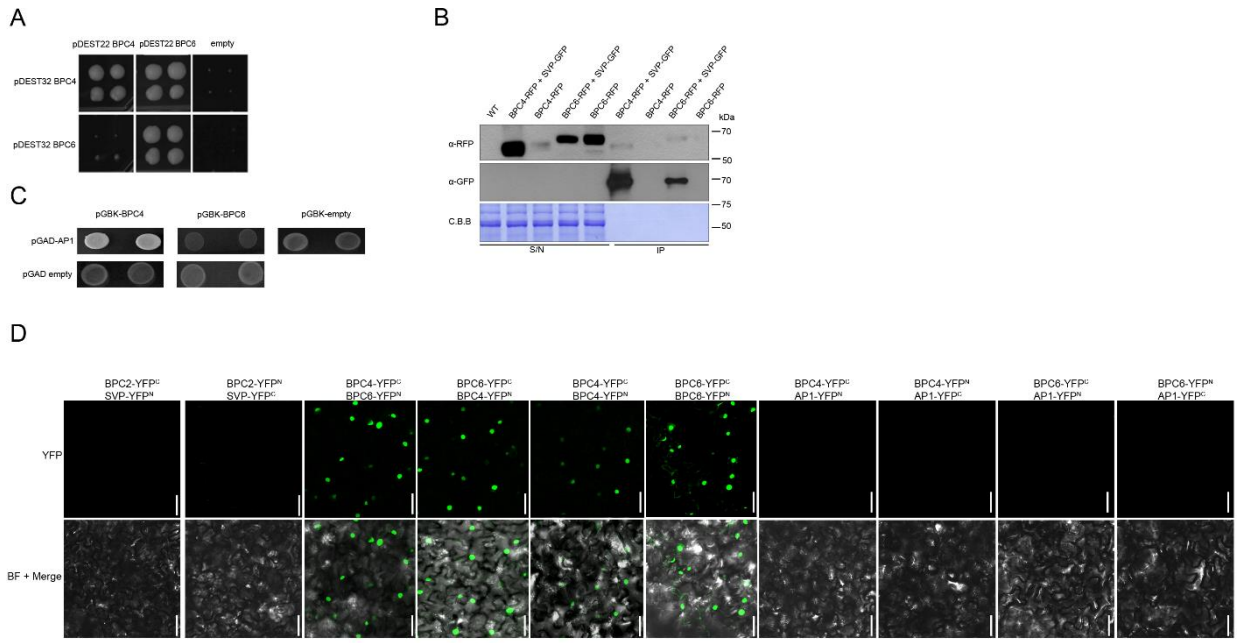


Figure S3.

(A) Yeast two-hybrid assay between BPCs of class II: positive interactions on selective media –W-L-H +5mM 3-AT. (B) Co-immunoprecipitation assays. *Nicotiana benthamiana* leaves were infiltrated with constructs carrying SVP-GFP together with BPC4-RFP and BPC6-RFP, as described in experimental procedures. Immunoprecipitation step was performed using GFP-trap on total protein leaf extract. Samples were probed with GFP and RFP antibody. S/N: supernatant; IP: immunoprecipitation. (C) Yeast two-hybrid assay between AP1 and BPCs of class II: positive interactions on selective media –W-L-H +5mM 3-AT. (D) Bi-molecular fluorescence complementation (BiFC) assay. *N.benthamiana* epidermis cells were transiently transformed with the indicated YN and YC fusions; In the first and the second row yellow fluorescence and merging in the bright field were shown, respectively. Scale bars=50 µm.

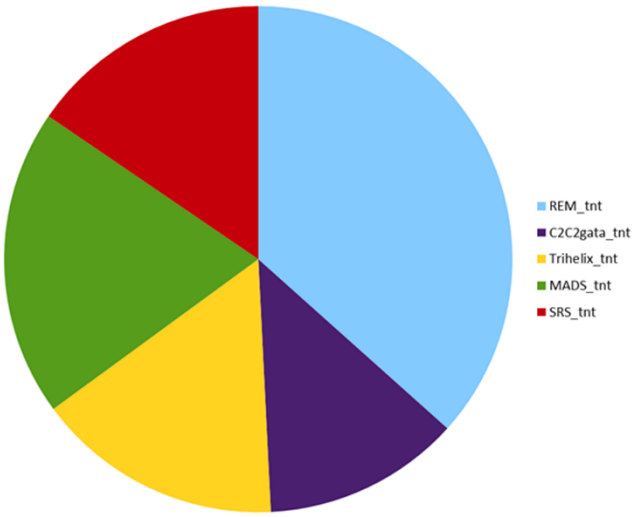


Figure S4.

Proportion of significantly overlapped DAP-Seq Peaks by TF family. Proportion, with respect to the total number of significantly overlapped peaks, of significantly overlapped DAP-seq peaks associated to each TF-family.

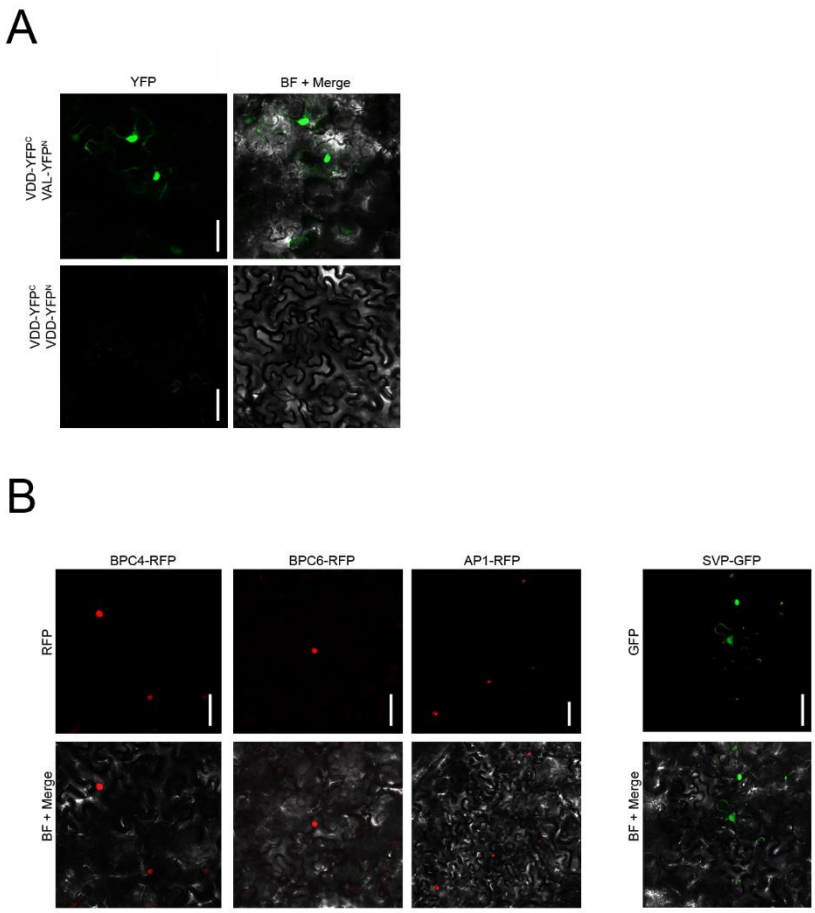


Figure S5

(A) Bi-molecular fluorescence complementation (BiFC) assay. *N.benthamiana* epidermis cells were transiently transformed with the indicated YN and YC fusions; VDD-VAL and VDD-VDD were tested as positive and negative control, respectively (Mendes et al., 2016). In the first and the second column yellow fluorescence and merging in the bright field were shown, respectively. Scale bars=40 µm. (B) *N.benthamiana* epidermis cells were transiently transformed with the indicated constructs. In the first and the second row yellow/red fluorescence and merging in the bright field were shown, respectively. Scale bars=40 µm.

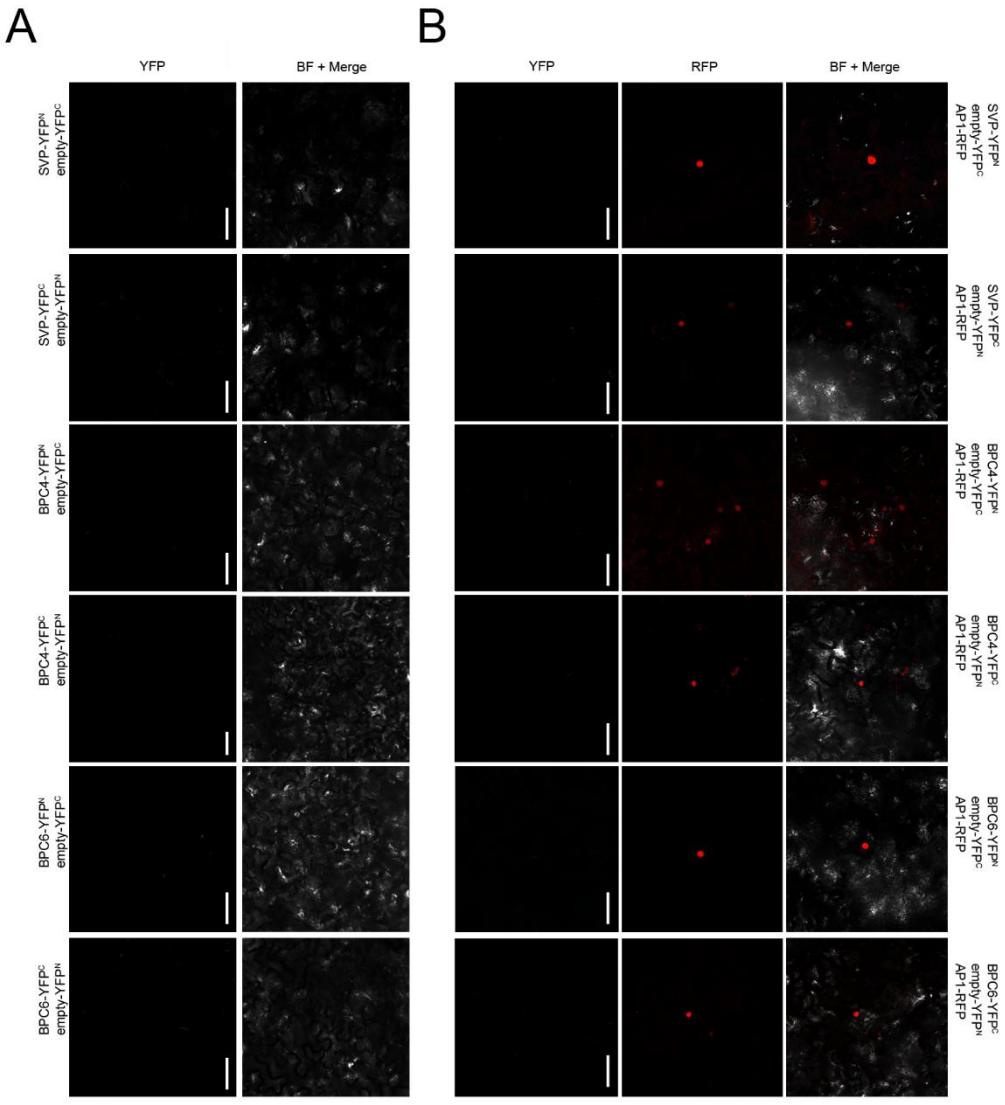


Figure S6.

(A) Bi-molecular fluorescence complementation (BiFC) assay. *N.benthamiana* epidermis cells were transiently transformed with the indicated YN and YC fusions. In the first and the second column yellow fluorescence and merging in the bright field were shown, respectively. Scale bars=40 µm. (B) Bi-molecular fluorescence complementation (BiFC) assay. *N.benthamiana* epidermis cells were transiently transformed with the indicated YN and YC fusions and AP1-RFP construct. In the first, the second and the third column yellow fluorescence, red fluorescence and the merging between the two channels in the bright field were shown, respectively. Scale bars=40 µm.

Manuscript #2: REM34 and REM35 control female and male gametophyte development in *Arabidopsis thaliana*

In press to Frontiers in Plant science

This manuscript assesses a role for REM34 and REM35 in female and male gametophyte development. I performed several BiFC assays to confirm and demonstrate REM34-REM35 heterodimers and REM35-REM35 homodimers formation *in planta*.

REM34 and *REM35* control female and male gametophyte development in *Arabidopsis thaliana*

Caselli F.[1], Beretta V.M. [1], Mantegazza O. [1], Petrella R. [1], Leo G. [1], Guazzotti A.[1], Herrera-Ubaldo H. [2], de Folter S.[2], Mendes M.A[1], Kater M.M. [1] and Gregis V. [1]*.

[1] Università degli Studi di Milano, Dipartimento di Bioscienze, Milan, Italy.

[2] Unidad de Genómica Avanzada, Laboratorio Nacional de Genómica para la Biodiversidad, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, Irapuato, Mexico.

*Correspondence

Corresponding Author

veronica.gregis@unimi.it

Keywords: gametophyte development, REM, transcriptional regulation, ovule, pollen, post-meiotic division, *Arabidopsis thaliana*.

Running title: *REM34* and *REM35* control gametophyte development in Arabidopsis

Abstract

The REproductive Meristem (REM) gene family encodes for transcription factors belonging to the

B3 DNA binding domain superfamily. In *Arabidopsis thaliana* the *REM* gene family is composed of 45 members, preferentially expressed during flower, ovule and seed development. Only a few members of this family have been functionally characterized: *VERNALIZATION1* (*VRN1*) and most recently *TARGET OF FLC AND SVP1* (*TFS1*) regulate flowering time and *VERDANDI* (*VDD*), together with *VALKYRIE* (*VAL*) control the death of the receptive synergid cell in the female gametophyte. We investigated the role of *REM34, REM35* and *REM36,* three closely related and linked genes similarly expressed in both female and male gametophytes. Simultaneous silencing by RNA interference (RNAi) caused about 50% of the ovules to remain unfertilized. Careful evaluation of both ovule and pollen development showed that this partial sterility of the transgenic RNAi lines was due to a post meiotic block in both female and male gametophytes. Furthermore, protein interaction assays revealed that REM34 and REM35 interact, which suggests that they work together during the first stages of gametogenesis.

Introduction

In higher plants, the alternation between the diploid sporophytic generation and the haploid gametophytic generation is a fundamental characteristic of their life cycle. The formation of the gametophyte from the sporophyte is the result of two sequential processes, sporogenesis and gametogenesis. Angiosperms are heterosporous plants, characterized by the production of two types of unisexual gametophytes, the megagametophyte (embryo sac) and microgametophyte (pollen). Development of both female and male gametophytes can be divided into two main steps: sporogenesis, during which meiosis occurs giving rise to haploid spores, and gametogenesis, which leads to the formation of the gametes (Berger and Twell, 2011).

In Arabidopsis, the female gametophyte develops in the gynoecium. The first step of megasporogenesis consists in the formation of the ovule primordia, in which one cell differentiates into the megaspore mother cell (MMC) or megasporocyte; the MMC sustains one meiotic division, giving rise to four haploid megaspores. Only one of them, the functional megaspore, continues its development and goes through three mitotic divisions forming a mature embryo sac composed of eight nuclei and seven cells: three antipodal cells, two medial polar nuclei, and one egg cell surrounded by two synergids (Mansfield et al. 1991).

In the anthers the microspore mother cell gives rise, through meiosis, to four microspores, which develop into mature pollen grains, containing two sperm cells surrounded by the vegetative cell (Hafidh et al., 2016).

The transition from sporogenesis to gametogenesis is directly correlated with the cell cycle transition from meiosis to mitosis. During gametogenesis, the number of mitotic divisions (two for the male and three for the female gametophyte) has to be tightly regulated and coordinated with cytokinesis. This cell division process is complex and requires the integration of different pathways such as those involved in cell cycle progression, chromatin modifications and hormonal signalling. Moreover, mitotic progression during gametogenesis is also affected when interfering with basic biological processes like organelle and ribosome biogenesis (Shi et al., 2005; Li et al., 2009; Wang et al., 2012).

In both gametophytes the retinoblastoma-related protein (RBR) plays a key role in the regulation of the cell cycle by inhibiting cell cycle entry through repressing E2F transcription factors. The *rbr* mutation results in an uncontrolled nuclear proliferation in both gametophytes (Ebel et al., 2004; Ingouff et al., 2006; Johnston et al., 2008). More recently *RBR* was also associated with the meiosis activation, when the MMC is getting reduced by meiosis and forming subsequently the functional megaspore (Zhao et al., 2017).

In all eukaryotic organisms, cell cycle progression is tightly linked to the activation and degradation of different cyclin-dependent kinases (CDKs). During both female and male gametophyte development the activity of two homologous RING finger E3 ubiquitin ligases, RHF1 and RHF2, are required for the degradation of the CDK inhibitor ICK4/KRP6, which allows the correct progression of the cell cycle. In the *rhf1 rhf2* double mutant, both female and male gametophytes fail to complete their development and are arrested in FG1 and microspore stage respectively (Liu et al., 2008).

The transcriptional activity in different cell types during plant development is dependent on epigenetic modifications, such as chromatin remodelling and histone modifications. Failure in the establishment of such modifications can cause different defects throughout the plant's life cycle. During gametogenesis, silencing of the *CHROMATIN-REMODELLING PROTEIN 11* (*CHR11*) within the embryo sac causes an arrest of nuclear proliferation from stage FG1 to FG5 (Huanca-Mamani et al., 2005). Furthermore, mutations in the histone acetyl transferase genes *HAM1* and *HAM2*

causes an arrest in the early stages of both megagametogenesis and microgametogenesis (Latrasse et al., 2008).

Genetic studies have identified a large number of loci that control gametophyte development. Molecular cloning and characterization of some of them has revealed insights in sporocyte formation, meiosis/mitosis and gametophyte development. Detailed phenotypic and molecular characterization of mutants remains a big challenge also because of the complication to work with such mutants, which often are partially sterile or even lethal (Muralla et al., 2011).

In the context of finding new players involved in the control of this process, the *REM* gene transcription factor family promises to be a good candidate since of the four members that were functionally characterized, two of them *VERDANDI* (*VDD* or *REM20*) and *VALKYRIE* (*VAL* or *REM11*) have a function in gametophyte development (Matias-Hernandez et al., 2010; Mendes et al., 2016). The other two members, *VERNALIZATION1* (*VRN1* or *REM5*) and *TARGET OF FLC AND SVP1* (*TFS1* or *REM17*) were shown to be involved in the control of flowering time (Sung et al.; Levy, 2002, Richter et al., 2019).

The expression patterns of *REM* genes were analyzed by Mantegazza et al. (2014) showing that the majority of the members of this family are preferentially expressed during flower and seed development. Through this analysis we identified *REM34*, *REM35* and *REM36*, which are mainly expressed in the reproductive meristems but also throughout different stages of flower development. *REM34*, *REM35* and *REM36* are located in a cluster, containing in total 9 *REM* genes on the fourth chromosome of Arabidopsis. *REM34*, *REM35* and *REM36* are very similar, which might indicate a possible functional redundancy.

Insertional mutants already analyzed for *REM34* and *REM36* are not complete knock-outs and showed no visible phenotype whereas no insertional mutants are available for *REM35* (Mantegazza et al., 2014). Since these genes are located in linkage on the Arabidopsis genome, it is also practically impossible to obtain multiple mutant combinations by crossing the available mutant lines.

Therefore, in this study we investigated the role of *REM34*, *REM35* and *REM36* through simultaneously downregulating them by RNA interference. Plants in which at least *REM34* and *REM35* were down-regulated showed an early arrest in the development of both female and male gametophytes. The process of mega/micro sporogenesis was not affected and meiosis was taking

place. However subsequent mitosis was not occurring after spore formation, suggesting that these genes play a role in gametogenesis progression.

Materials and methods

Plant material and growth conditions

All experiments were performed in *Arabidopsis thaliana* ecotype Columbia-0 (Col-0). Plants were grown in a controlled environment at 20-22°C either under long day conditions (16h light/8h dark) or under short day (8h light/16h dark) conditions for four weeks after germination and then transferred to long day conditions. The *suf4-1 pSUF4*:*SUF4-GUS* seeds were donated by S.D. Michaels. Tobacco plants were germinated and grown at 20-22°C under long day conditions.

RNA interference and *35S:EAR_REM34* constructs

To obtain the *REM_RNAi* construct 252, 232 and 254 base pairs long DNA fragments specific for the coding sequence of each of the genes *REM34, REM35* and *REM36* were selected (the primers used to amplify the fragments are listed in the Supplementary Table 1). The fragments specificity was checked by BLAST against the Arabidopsis genome. The three selected regions were PCR amplified, adding the *BsaI* sites to the primers, and cloned in a pENTR™ vector previously modified to function as a Golden Gate acceptor, with a single Golden Gate reaction, producing the pENTR-*RNAi_REM* vector. The Gateway LR reaction (Invitrogen™ Gateway™ recombination cloning system) was then performed to sub-clone the *RNAi_REMs* fragments into the pFGC5941 vector and used to transform Arabidopsis. Primers that were used are listed in Supplementary Table 1. The EAR motif was added to the C terminus of the *REM34* coding sequence (see primer sequences in Supplementary Table 1). The fragment was cloned into the pB2GW7 plasmid (35S) passing through the pENTRY-D-TOPO vector (Invitrogen™ Gateway™ recombination cloning system). Arabidopsis plants were transformed using the floral-dip method (Clough and Bent, 1998).

Quantitative RT-PCR

Total RNA was extracted from whole inflorescences. RNA samples were treated with DNase (TURBO DNA-free®; Ambion, http://www.ambion.com/) and retrotranscribed employing the ImProm-IITM reverse transcription system (Promega). Diluted aliquots of the cDNAs or genomic DNA were used as templates in qRT-PCRs, using the iQ SYBR Green Supermix (Bio-Rad) to detect

target synthesis. All the experiments were performed with three technical replicates for each of the three biological replicates, with the exception of the expression analysis of *REM34, REM35* and *REM36* in the T1 *REM_RNAi*, in the T1 35S:*REM34_EAR* plants and for T-DNA abundancy evaluation. Primers employed for these analyses are listed in Supplementary Table 1.

Silique length, seed number evaluation and reciprocal crosses

For each line, 10 siliques (dissected from three different plants) were measured, and seed, aborted seed and non-fertilized ovule numbers was counted. For this purpose, a Leica® MZ 6 microscope was used. For the reciprocal crosses between wild-type and *REM_RNAi* #1 plants, mature siliques as well as open flowers and buds in an advanced stage of development were removed from the inflorescence of the mother plant, along with the meristem and smallest buds. Remaining buds were emasculated by removal of all floral organs except for the ovary. Then, anthers in the correct stage of development were taken from other flowers and used to pollinate the stigma. The number of seeds and unfertilized ovules were assessed for at least five pistils for each cross and three biological replicas of the experiment were performed.

In situ hybridization analysis

In situ hybridization analysis for *REM34, REM35* and *REM36* were performed following the same protocol and employing the same probes described by Mantegazza et al. (2014). Evaluation of the expression profile in the inflorescence and flower meristems was used as a positive control.

Protein-protein interaction analysis

Yeast two-hybrid assays were performed in the yeast strains PJ69-4A and PJ69-4α (de Folter and Immink, 2011). The coding sequences of *REM34, REM35* and *REM36* were cloned in the pDEST32 (bait vector, BD; Invitrogen) and pDEST22 (prey vector, AD; Invitrogen) Gateway vector. The bait constructs were tested for autoactivation on selective yeast synthetic dropout medium lacking Leu, Trp and His supplemented with 1, 3, 5, 10 or 15 mM of 3-aminotriazole, in order to set the screening conditions. After mating, colonies were plated on the proper selective media and grown for 5 days at 20°C.

The same coding sequences were also cloned in the pYFPN43 and pYFPC43 vectors, to perform the BiFC assay. Agrobacterium, transformed with the vectors and the viral suppressor p19 construct, was used to infiltrate tobacco leaves. The abaxial surfaces of infiltrated leaves were imaged 3 d after inoculation. As positive control for the infiltration the already published VAL-VDD

interaction was tested (Mendes et al., 2016). As negative controls the constructs containing the proteins of interest were co-transformed with the empty pYFN43 and pYFC43 vectors. Furthermore, REM34 homodimerization, which was not observed in the Y2H assays, was also employed as a negative control (Supplementary Figure 5).

Female Gametophyte Characterization

Female gametophytes were cleared and analyzed as previously described by Brambilla et al. (2007). Inflorescences were prepared for observation using the following protocol: flowers were emasculated and the next day harvested. The emasculated pistils were left O/N at 4°C in a 1:9 Acetic Acid: Ethanol solution. Samples were rehydrated by subsequent washes with Ethanol 90% and 70% and then incubated O/N at 4°C in Clearing Solution (160g Chloral Hydrate, 50 g Glycerol, and H_2O to a final volume of 250 ml). Pistils at different developing states were separated from the other floral organs and opened to evaluate the female gametophyte morphology. For these experiments, a Zeiss Axiophot® microscope equipped with differential interference contrast (DIC) optics was used.

In vitro pollen germination

For this experiment, the protocol published by Bou Daher, Chebli, & Geitmann (2009) was followed applying minor modifications. Pollen grains were plated on small glass plates, containing 2,5 ml of Pollen Germination Medium (PGM:18% sucrose, 0.01% Boric Acid, 1 mM $CaCl_2$, 1 mM $Ca(NO_3)_2$, 1 mM $MgSO_4$, 0.5% Agarose pH=7). The plates were overnight incubated at 22°C, with wet paper to maintain humidity. The next day, pollen germination and growth were evaluated with a Zeiss Axiophot® microscope.

Aniline Blue Staining

Flowers were emasculated and, after 24 hours, pollinated. The pollinated ovaries were collected at two different time points: 5 and 24 hours after pollination. Samples were overnight fixed and stained in absolute ethanol/glacial acetic acid 9:1, as previously described by Mori et al. (2006). Subsequently, they were transferred into a 8M NaOH solution for 1 hour at 50°C. Finally, the carpels were washed twice with ddH_2O for 10 minutes. The staining was performed with a modified Aniline blue solution (Aniline blue 2%, glycerol 1M ddH_2O) (Takeuchi and Higashiyama, 2016). Samples were stored at 4°C for 3 hours or overnight. The observation was done under UV light (350-400 nm) with a Zeiss Axiophot® microscope

Pollen DAPI staining

Pollen was stained according to Park et al. (1998). Mature pollen was obtained by placing 3-4 open flowers in a microcentrifuge tube containing 300 µl of DAPI staining solution (0.1 M sodium phosphate (pH 7), 1 mM EDTA, 0.1% Triton X-100, 0.4 µg/ml DAPI high grade, Sigma). After brief vortexing and centrifugation, the pollen pellet was transferred to a microscope slide and observed with a Zeiss Axiophot® microscope. Pollen at earlier stages of maturation were also analyzed by dissecting single anthers. Anthers were disrupted on microscope slides and squashed in DAPI staining solution (1 µg/ml) under a coverslip.

GUS staining

GUS assays were performed as described by Resentini et al. (2017). Pistils at different developmental stages were dissected and fixed in acetone 90% and incubated O/N at 37°C. After staining, they were cleared using the protocol described above.

Alexander Staining for Pollen Grains

Staining of pollen grains was performed as described by Peterson et al. (2010). After fixation (performed with 6 alcohol:3 chloroform:1 acetic acid) the anthers were placed on a microscope slide with a few drops of staining solution (10 ml 95% alcohol, 1 ml Malachite green (1% solution in 95% alcohol), 50 ml distilled water, 25 ml glycerol, 5 ml acid fuchsin (1% solution in water), 0.5 ml orange G (1% solution in water), 4 ml glacial acetic acid and distilled water (4.5 ml) to a total of 100 ml). Samples were analysed with a Zeiss Axiophot® microscope.

CLSM analysis

For confocal imaging, the Laser Scanning Confocal Microscope Nikon A1 was used. Inflorescences were fixed as described by Braselton et al. (1996). Samples were then excited using a laser (532 nm) and emission was detected between 570 and 740 nm.

Results

RNAi mediated silencing of *REM34, REM35* and *REM36*

Since *REM34, REM35* and *REM36* are very similar and in linkage, an RNA interference approach was adopted to investigate their role during reproductive development in Arabidopsis.

Due to sequence divergency, even in the B3 DNA binding domain (Romanel et al., 2009), it was impossible to design a single artificial small interfering RNA fragment that was able to silence the three *REM* genes simultaneously. Therefore, a multiple RNA interference (RNAi) technology was used to express a single chimeric double stranded RNA that targeted the three REM genes under the control of CaMV35S (Miki et al., 2005; Bucher et al., 2006) (Figure 1A).

We selected three regions specific for the coding sequence of *REM34*, *REM35* and *REM36*. The regions selected for *REM34* and *REM36* are highly specific for the genes of interest and were expected not to have any off target in the Arabidopsis genome. The RNAi fragment that targets *REM35* has a partial complementarity with *REM36*, and, at a lower level, with *REM37,* whose expression is almost undetectable in most Arabidopsis tissues (Mantegazza et al., 2014; Klepikova et al., 2016).

Forty *REM_RNAi* T1 transgenic Arabidopsis lines were obtained. We evaluated the down-regulation of the *REM* genes in 9 different T1 lines (Figure 1B), which all showed defects in silique and gametophyte development.

Silencing of the three target genes was confirmed in the T2 generation by qRT-PCRs (Supplementary Figure 1). Furthermore, we showed that the RNA interference construct was specific for their targets by testing the expression of *REM37* and *REM39*. The latter was chosen due to the fact that *REM39* is highly expressed in the tissues where *REM34, REM35* and *REM36* are also active (Mantegazza et al., 2014; Supplementary Figure 1).

REM_RNAi lines have a reduced ovules number and seed set compared to wild-type plants

We selected three *REM_RNAi* lines (#1, #4 and #5), with different levels of silencing of *REM36,* for further investigations in the T2 generation. In line #1, *REM36* showed a downregulation of around 50%, while in line #4 and #5 *REM36* was found to be slightly upregulated compared to the wild-type

(Figure 1B).

In the T2 generation, silique length and seed number were evaluated for the three selected lines. The *REM_RNAi #T2.1* line showed a decrease of 35.3% in the silique length and a 19.4% reduction in total ovule number (Figure 2A,B and C). Furthermore, on average 66% of the ovules failed to be fertilized (Figure 2C,D). The other two *REM_RNAi* lines, #*T2*.4 and #*T2*.5, showed a similar phenotype even if the percentage of unfertilized ovules was lower, 35.3% and 45.4% respectively (Figure 2C).

The *REM_RNAi #1* line was selected to further investigate the sterility phenotype caused by the downregulation of *REM34, REM35* and *REM36.* This line was propagated to the T3 generation, where plants homozygous for the *REM_RNAi* construct were selected. Even if the RNAi construct has a dominant effect, we evaluated whether the sterility observed in the *REM_RNAi T2* segregating lines was exacerbated in plants homozygous for the construct. For this purpose, the seed set of the *REM_RNAi T3.1* homozygous line was evaluated.

Interestingly, comparing both the *REM_RNAi #T2.1* and the *REM_RNAi #T3.1* we noticed that the percentage of ovule abortion was the same, suggesting that the silencing of REMs is probably acting both at the sporophytic and gametophytic levels.

Since the two lines in which *REM36* was not downregulated displayed a milder phenotype compared to the *REM_RNAi #T2.1 and REM_RNAi #T3.1* lines, in which all three genes were downregulated, it is possible that *REM36* is partially redundant to *REM34* and *REM35* during gametophyte development. On the contrary, the ovule number was the same in all three *REM_RNAi* lines (Figure 2A,C), indicating that *REM36* is not involved in the determination of the ovule primordia number.

To further confirm that no phenotypical differences were detectable between plants homozygous and heterozygous for the T-DNA insertion, we analyzed the silique content of 10 *REM_RNAi #T2.4* and 10 *REM_RNAi #T2.5* T2 plants in which the construct was still segregating, and we found no significant differences between all the herbicide resistant plants (Supplementary Figure 2A,C). For both REM_RNAi #T2.4 and #T2.5 lines, a relative evaluation of T-DNA copies in each of the 9 plants considered, was performed. The RT-PCR analyses showed a various ammount of T-DNA amplicons which is clearly unrelated to the ovule abortions and the overall seed set observed in all the REM_RNAi #T2.4 and #T2.5 analyzed individuals (Supplementary Figure 2). The *ACTIN7* amplicon was used as normalizer and the herbicide resistance BAR gene used to estimate the abundancy of T-DNA copies.

These analyses allowed excluding the possibility that the reduced seed set was linked to the presence of a heterozygous T-DNA insertion (Curtis et al., 2009; Clark et al., 2010) and suggests that either the sporophytic silencing of *REM34*, *REM35* and *REM36* affects the gametophyte, or that the mobile siRNA diffuses from the sporophyte to the gametes (Mlotshwa et al., 2002; Melnyk et al., 2011; Skopelitis et al., 2018).

To understand if the reduced seed set was due to problems in the female or the male gametophyte, we performed reciprocal crosses between *REM_RNAi #T3.1* and wild-type plants. As a control both *REM_RNAi #T3.1* (homozygous for the T-DNA triggering the RNAi silencing) and wild-type plants were manually selfed, in order to evaluate if the manipulation of the flower was affecting the fertility of the analyzed plants (Figure 2E).

When *REM_RNAi #T3.1* pistils were pollinated with *REM_RNAi #T3.1* pollen, 73.3% of the ovules failed to be fertilized while wild-type lines manually pollinated with wild-type pollen resulted in 19.5% unfertilized ovules. When the *REM_RNAi #T3.1* line pistils were pollinated with wild-type pollen the percentage of unfertilized ovules was 78.6%, indicating a strong contribution of the female reproductive organ defects to this phenotype. Interestingly, when wild-type pistils were pollinated with *REM_RNAi #T3.1* pollen still 61.0% of the ovules were not fertilized (Figure 2E). Moreover, we observed a high variability in the number of unfertilized ovules using *REM_RNAi* pollen as shown in Figure 2E. Macroscopical inspection revealed a decrease in pollen grain number compared to wild-type anthers and a lack of adherence of the pollen to the wild-type stigma, both observations were further investigated (see below). All these considerations strongly suggest that both female and male reproductive organs are affected in the *REM_RNAi* lines.

REM34, REM35 and *REM36* are expressed in both female and male reproductive organs in adjacent sporophytic and gametophytic cells

Previously, the expression pattern of the *REM* genes was characterized in the shoot apex by *in situ* hybridization analysis, showing that *REM34, REM35* and *REM36* are expressed from the earliest stages of reproductive development of Arabidopsis in the inflorescence meristem, flower meristem and during the first stages of flower development with the exception of sepals (Franco-Zorrilla et al., 2002; Mantegazza et al., 2014).

In order to analyse the expression profiles in more detail during male and female sporophytic /gametophytic development we performed *in situ* hybridization analysis for *REM34, REM35* and *REM36* in both female and male reproductive organs. The flower stages are described accordingly to Smyth et al. (1990) and Schneitz et al. (1995).

In Arabidopsis, pollen mother cells differentiate inside the young anther and ovule primordia arise from the placenta in stage 8 of flower development and differentiation is completed at stage 13.

At stage 8/9 of flower development (Smyth et al., 1990) hybridization signals were detected for all three genes, in the anthers where the pollen mother cells differentiate and within the carpels, although in this case the signal was stronger in the placenta and ovule primordia (Figure 3A).

At stage 10 a strong signal was always detected in developing ovules and pollen (Figure 3B). Our analysis revealed that the timing of expression of the three REM genes coincided with male and female sporogenesis.

During subsequent stages of flower development, stages 11-12, both female and male initiate gametophyte development. During these stages a decrease in the signal was clearly observed in anthers (Figure 3C); during these stages pollen reach maturity and the vegetative and generative cells are differentiated after mitosis (Twell et al., 1998). In contrast, a strong signal was detected during ovule development when the surviving megaspore undergoes three rounds of mitosis and passes from stage 3I to 3V (Schneitz et al., 1995). Interestingly, when the ovule is at its very last stage of development 3-VI (Schneitz et al., 1995), a strong signal was detected in the funiculus, in the innermost integument and, inside the mature female gametophyte, in the central cell region (Figure 3D).

The expression analysis of *REM34, REM35* and *REM36* highlighted the fact that also during anther/pollen and carpel/ovule development these three *REMs* have a similar pattern of expression.

The analysis of the expression patterns of *REM34, REM35* and *REM36* combined with the phenotypes observed in the *REM_RNAi* lines denote an important role for these genes during the development and production of viable male and female structures and gametes.

In *REM-RNAi* lines the female gametophyte is unable to complete its development

The expression profile of *REM34, REM35* and *REM36* suggests that these genes play a role during ovule development. Furthermore, the reciprocal crosses showed that between 73.3% and 78.6% of the ovules in the *REM_RNAi #T3.1*, which is homozygous for the RNAi cassette, were not fertilized (Figure 2E).

Based on this evidence we hypothesized that the ovule defects in the *REM_RNAi* lines might be due to an arrest in their development. Therefore, a detailed evaluation of female gametophyte development was carried out in the *REM_RNAi #T3.1* homozygous line. In this line, 42.9% (227/529) of the ovules failed to complete their development and showed an arrest in the FG1 stage (Figure 4A). These ovules were characterized by an embryo sac containing one large cell, the functional megaspore, with a single nucleus, the rest of the ovules completed their development reaching the FG7 stage (Figure 4B,C). The same phenotype was observed in the *RNAi #T2.4* and *#T2.5* lines which both derived from hemizygous mothers (Supplementary Figure 3).

To confirm that in the *REM_RNAi* lines the defective female gametophytes were arrested in the FG1 stage, after meiosis, we crossed the *pSUF4:SUF4-GUS* marker line with the *REM_RNAi #T3.1* line. In the *pSUF4:SUF4-GUS* marker line GUS expression is not detectable during megasporogenesis, but it becomes visible after meiosis, once the functional megaspore is formed, and marks all the nuclei of the embryo sac (Resentini et al., 2017). Observing *REM_RNAi #T3.1* pistils, both wild-type like ovules, with more than one nucleus and ovules arrested in the FG1 stage, with the nucleus of the functional megaspore, expressed the *GUS* reporter (Figure 4D) suggesting that the defect in female gametophyte development was post-meiotic.

To investigate in detail the arrest at the FG1 stage, we carried out confocal laser scanning microscopy (CLSM) on *REM_RNAi #T3.1* developing ovules. The feulgen staining perfectly marked the cell wall of the ovule integuments and the embryo sac dividing nuclei, allowing the recognition of the gametophytic developmental stages. In the same *REM_RNAi #T3.1* pistil we observed ovules that normally developed until stage FG4 (Figure 4E) and those that were arrested in FG1 in which the embryo sac contains the functional megaspore and the three degenerating spores on top of it (Figure 4F).

The *REM-RNAi* lines showed a post-meiotic defect of the male gametophyte

From the analysis of wild-type carpels pollinated with *REM_RNAi* pollen, we observed that 61.1% of the ovules were not fertilized, suggesting that the male gametophyte in these lines is also defective (Figure 2E). To understand the cause of this defect we first carried out an *in vitro* pollen germination assay which showed a 30% decrease in the germination rate of the *REM_RNAi #T3.1* pollen compared to the wild type. (Figure 5A,B and Supplementary Figure 4).

The growth rate of *REM_RNAi* pollen tubes and their ability to correctly target ovules were also evaluated *in vivo* by means of Aniline Blue staining (Supplementary Figure 4). The *REM_RNAi* pollen tubes did not show any growth defect, they reached the end of the pistil in the same time as the wild-type pollen tubes, and the mature ovules were correctly targeted (Supplementary Figure 4). We noticed that, as mentioned before, the *REM_RNAi* pollen number appeared to be lower and it did not adhere well to the stigma papillae, which could explain the high variability observed in the backcrosses between wild-type pistils and *REM_RNAi* pollen (Figure 2E).

To try to understand the cause of the male sterility phenotype, pollen grains were collected from mature anthers and treated with Alexander's stain, which colors viable pollen red. While in the wild type all the collected pollen was viable, in the *REM_RNAi #T3.1* anthers 33.9% of the grains were not stained, indicating that those pollen grains were non-viable and did not appear to contain any cytoplasm (Figure 5C-E). Interestingly, the percentage of non-viable pollen grains in the *REM_RNAi* line corresponds to the decreased germination capability observed *in vitro,* suggesting that the grains which are unable to produce the pollen tube are the degenerated ones.

To investigate the pollen defect in more detail, confocal laser scanning microscopy (CLSM) was used. In Figure 5F, wild-type pollen from a mature anther is shown, intine and exine layers were very well distinguishable and inside the pollen grain, the sperm cells and the vegetative cell nuclei were stained. On the contrary, in the *REM_RNAi #T3.1* mature anthers, a high percentage of pollen grains appeared shrunken and empty, neither sperm nor vegetative cells were identified, although the intine and exine layers looked intact (Figure 5G).

To understand when the pollen grains degenerated, we visualized their nuclei with DAPI staining at different developmental stages (Figure 6A-F and Supplementary Figure 3). At the microspore stage, all *REM_RNAi #T3.1* grains were characterized by the presence of a single bright nucleus localized at the center of the cell, indicating that the pollen, like wild-type, passed through meiosis correctly (Figure 6A and D). After meiosis, in wild-type the microspores underwent a first mitotic division that produced one vegetative and one sperm nucleus (Figure 6B). Subsequently the

second round of mitosis led to the formation of the mature pollen grain, which contained two small sperm cells each with a bright and elongated nucleus and the vegetative cell (Figure 6C). Interestingly, in *REM_RNAi #T3.1* anthers, some grains were characterized by the lack of nuclei, this phenotype was detectable also at the tricellular stage (Figure 6E,F and Supplementary Figure 3).

Thus, after meiosis *REM_RNAi* anthers displayed both viable pollen grains, with two sperm cells nuclei and a distinct vegetative nucleus, and not viable pollen grains, in which no DNA is detectable (Figure 6E,F). This is similar to what was observed with the CLSM analysis.

All this evidence suggests that the degeneration of pollen grains observed in the *REM_RNAi* lines could be due to a post-meiotic block in their development, a similar defect as the one observed in the female gametophytes.

REM35 formed homodimers and heterodimers with REM34

REM transcription factors can form functional heterodimers (Mendes et al., 2016). To understand if also REM34, REM35 and REM36 could function via dimer formation yeast two-hybrid assays were performed. This approach revealed that REM35 is able to interact strongly with itself and also with REM34, while no interactions were detected with REM36 (Figure 7A and Supplementary Figure 5).

All the interactions observed in the yeast two-hybrid assays were confirmed *in vivo* with a Bimolecular Fluorescence Complementation (BiFC) assay in *Nicotiana benthamiana* leaves (Figure 7B,C and Supplementary Figure 5). This finding suggests that REM34 and REM35 could act as heterodimers.

Downregulation of *REM34*, *REM35* and *REM36* altered expression of genes involved in post-meiotic divisions

As described above, the downregulation of *REM34, REM35* and partially *REM36* resulted in a post-meiotic arrest in both female and male gametophytes, suggesting that these transcription factors could be involved in regulating mitosis progression during gametogenesis.

To elucidate the molecular mechanism causing this block, we measured the expression levels of genes that control gametogenesis by q-RTPCR. We focused on genes that, when mutated or overexpressed, cause similar defects to those observed in the *REM_RNAi* gametophytes. Those genes were divided into three categories based on the biological process in which they are involved in: ribosome biogenesis (*MDS, NLE*), cell cycle control (*RBR, KRP6*) and chromatin regulation (*HAM1, HAM2*).

MDS, which, together with *NLE,* is involved in the biogenesis of the 60S ribosomal subunit and is essential during megagametogenesis (Chantha et al., 2010), was downregulated in the *REM_RNAi #T3.1* lines. *KRP6*, a CDK inhibitor whose overexpression causes a block in mitosis progression during female and male gametophytic development, was also downregulated in the *REM_RNAi* lines.

Among the genes involved in chromatin modifications two histone acetyltransferases (HATs), *HAM1* and *HAM2,* were selected. Only *HAM2* was downregulated in the *REM_RNAi #T3.1* line (Figure 8).

These results suggest an intricate interconnection among regulators and effectors, which end up in a correct gametogenesis program.

Overexpression of the REM34_EAR chimeric protein

The genes that were downregulated in the *REM_RNAi* lines might be targets of the REM transcription factors. This suggests that REM34 and REM35 might be transcriptional activators. To investigate whether REMs transcription factors work as activators of transcription, we fused REM34 with the dominant EAR repressor domain (known as chimeric repressor silencing technology CRES-T) and transformed wild-type Arabidopsis plants with this construct. Five transgenic lines that overexpressed the *REM34_EAR* chimeric gene at different levels in the T1 generation were obtained (Figure 9A).

In the T2 generation, silique length and ovule number were measured in two independent lines (*REM34_EAR*#T2.1 and *REM34_EAR*#T2.7). In both the selected T2 *REM34_EAR* lines we observed a decrease in the silique length of 23.1% and 25.0% respectively, and the presence of 55.0% and 42.6% aborted ovules, similar to what was observed in the *REM_RNAi* lines (Figure 9B-E).

The phenotype of the aborted ovules was further evaluated in cleared mature carpels of the *REM34_EAR #T2.1* and *#T2.7* lines. We detected both ovules at FG7 stage, with the seven cells clearly distinguishable, and ovules at FG1 stage, characterized by a single cell embryo sac (Figure 9F).

To confirm also the post meiotic block in the male gametophyte, mature pollen of both *REM34_EAR #T2.1* and *#T2.7* lines were stained with DAPI. Similarly to what was observed in the *REM_RNAi* lines, some pollen grains were able to reach the tricellular stage while others appeared shrunken and degenerated, with no visible nuclei (Figure 9G).

The strong similarity between the *REM34_EAR* and the *REM_RNAi* phenotypes might suggest that the overexpression of the chimeric REM34_EAR protein was causing co-suppression of other *REM* genes. To exclude this possibility, we investigated the expression level of *REM35, REM36, REM37* and *REM39* in the *REM34_EAR #T2.1* and *#T2.7* lines. The level of expression of the endogenous *REM34* was not taken into account, as the perturbation of *REM34* expression alone did not cause any evident phenotypical defects (Supplementary Figure 6) (Franco-Zorrilla et al., 2002; Mantegazza et al., 2014). The obtained results showed that the closely related *REMs* were not affected suggesting that the expression of the REM34_EAR chimeric protein caused the observed phenotypes.

Discussion

Functional analysis of *REM* genes

The plant-specific REM family in Arabidopsis is composed of 45 genes, generated through multiple duplication events, which are mostly expressed during flower and ovule development (Romanel et al., 2009). Even if the expression pattern of these genes suggests that they could play an important role in regulating developmental processes such as shoot architecture and flower development, until now only a few of them have been associated to a function (Levy, 2002; Matias-Hernandez et al., 2010; Mendes et al., 2016). This might be due to their functionally redundancy but also because they are often in linkage on the genome.

Here we investigated the function of the linked duplicated *REM34, REM35* and *REM36* genes by a multiple RNAi approach and showed that *REM34*, *REM35* and partially *REM36,* are involved in male and female gametophytic development during post-meiotic divisions. A similar multiple RNAi approach was previously employed to silence simultaneously up to six target genes in *Arabidopsis thaliana* (Czarnecki et al., 2016).

The *REM_RNAi* construct was found to be a very efficient tool: by selecting specific gene sequences, we were able to silence the three target genes with a single construct and a single transformation event. Importantly, the construct showed to be highly specific for the three genes of interest without any obvious off-target activity. Transgenic lines showing silencing of the *REM* genes under study were all characterized by a reduced seed set and an arrest in female gametophyte development at the earliest stages of gametogenesis. Since *REM34, REM35* and *REM36* appeared to be mainly expressed in sporophytic tissues throughout Arabidopsis reproductive development, the CaMV35S promoter was chosen to drive the expression of the RNA interference fragments. The activity of the CaMV35S promoter seems to be low during female and male gametophyte development, but it has been shown that such promoter can be successfully employed to silence genes during gametophytic development (Acosta-García and Vielle-Calzada, 2004; Mendes et al., 2016). A valid hypothesis for the observed gametophyte phenotypes might be that it is caused indirectly by the silencing of *REM34, REM35* and *REM36* in the female and male sporophytic cells. However, it is also important to consider that the RNAi construct is dominant and that it can trigger a non-cell autonomous and systemic silencing signal which might be maintained throughout the different phases of plant development (Mlotshwa et al., 2002; Melnyk et al., 2011; Skopelitis et al., 2018).

Since functional redundancy is a common phenomenon in plants (Briggs et al., 2006), this kind of RNAi approach will be helpful for the functional characterization of members of highly redundant families and especially that are in linkage. Furthermore, since silencing of genes by *RNAi* is often not complete, this approach could favor the analysis of genes for whom knock-out leads to lethality or complete sterility.

REM protein interactions

Protein interaction studies revealed that REM34 and REM35 were able to interact with each other, while no interaction was found with REM36, this supports the hypothesis that *REM36* might not be able to substitute completely *REM34* and *REM35* function. Interactions between REM factors

were found before. VDD and VAL, two functionally characterized REM factors involved in synergid degeneration upon fertilization (Mendes et al., 2016), also interact with each other. Furthermore, both *VAL* and *REM35* were also able to make homodimers. These characteristics might well be a common feature for the REM family and, in perspective of the guilt-by-association principle, it would be informative to analyse all possible REM protein interactions. The same approach was shown to be extremely useful for the characterization of MADS domain transcription factor family, for which extensive protein-protein interaction studies effectively guided genetic studies and functional characterization of many of them (de Folter et al., 2005; Gregis et al., 2006; Fornara et al., 2008; Immink et al., 2009).

REM34 and *REM35* control female and male gametogenesis

We discovered that in the *REM_RNAi* lines both the male and female germ lines were able to go through meiosis correctly, but they were not able to pass the FG1 stage, suggesting a role for *REM34* and *REM35* in the control of gametogenesis in Arabidopsis.

Although the *REM* gene family was named after the specific meristematic expression of its first member *AtREM1,* named *REM34* (Franco-Zorrilla et al., 2002), our data showed that *REM34, REM35* and *REM36* are also expressed during gametophytic development and we discovered that they were expressed starting from both carpels and anther primordia specification throughout all the stages of anther and carpel development. In the carpel, the signal is strongly localized in the placenta and ovule primordia and in the developing ovules.

Indeed, our deep morphological analysis of both female and male gametophytes of the *REM_RNAi* lines showed that from 35% to 65% of the female gametophytes were unable to undergo mitosis and were arrested at the FG1 stage when the MMC acquires functional megaspore identity.

REM36 seemed to be partially redundant with *REM34* and *REM35.* Indeed, in the two lines in which the level of *REM36* expression was higher compared to the wild-type, the penetrance of the embryo sac defect was less. However, in all *REM_RNAi* lines we also observed a decrease of around 20% in the total ovules number irrespectively of the expression levels of *REM36.* Thus, *REM36* might be involved in embryo sac development together with *REM34* and *REM35* but is not controlling ovule primordia specification.

In these lines, also pollen development was affected showing the same post meiotic arrest of the embryo sac. Thus, in Arabidopsis, REM34, REM35 and partially REM36 transcription factors seem to be required post-meiotically for gametophytic development.

Further confirmation for their role during both female and male gametogenesis came from the analysis of different *35S:REM34_EAR* lines. These plants, in which *REM34* fused to the EAR repressor domain was overexpressed, showed the same post meiotic arrest both in embryo sac and pollen development, suggesting that a complex formed by REM34 and REM35 could act as a positive transcriptional regulator of gametogenesis.

Because of the redundancy and position in linkage of the three genes of interest in the genome, most of this functional study was conducted using RNA interference. This approach was found to be very effective in the silencing of *REM34, REM35* and *REM36,* but the transgenic lines cannot be easily employed for genetic studies, due to the fact that it acts dominantly and because the level of silencing of the target genes can vary between different lines and throughout subsequent generations. Despite these difficulties, the analyses performed on both segregating and homozygous lines suggest that these three genes can influence gametogenesis acting mainly at the sporophytic level. This hypothesis is also supported by the expression pattern of these genes which, as shown by the *in situ* hybridization analysis, are present in the sporophytic tissues both in pistils and anthers when gametogenesis is taking place. The observation that *REM34, REM35* and *REM36* appeared to be expressed throughout all stages of gametogenesis in the embryo sac, leaves of course the possibility open that they directly play a function in the female gametophyte. The employment of an embryo sac specific promoter could be useful in order to validate this hypothesis and to be able to better distinguish between the sporophytic and gametophytic role of *REM34, REM35* and *REM36.*

To understand how the *REM* genes under study act, we tested whether the down-regulation of *REM34, REM35* and *REM36* perturbed the expression of genes known to be involved in gametogenesis progression. These genes were classified accordingly to their biological function in three categories: cell cycle control, chromatin remodelling and ribosome biogenesis. Interestingly, we observed that several genes involved in different biological pathways were downregulated in the *REM_RNAi* lines. This observation suggests that *REM34, REM35* and, in some measures, *REM36* are involved in the control of a very early steps of gametogenesis. In particular, they regulate the expression of different targets both directly and indirectly along the genetic network that controls gametophytic development in Arabidopsis.

Among the downregulated genes, the one that stands out most is *HAM2*, a histone acetyltransferase (HAT) that, together with its homolog *HAM1*, belongs to the MYST clade of the HATs family and was shown to be involved in post-meiotic control of female and male gametophytic development (Latrasse et al., 2008). In mammals, the MYST protein family was found to be involved in many fundamental cell functions such as cell cycle progression and DNA repair (Pillus, 2008; Sapountzi and Côté, 2011). Furthermore, the human MYST4 acetylase was found to be expressed and involved in the control of gametogenesis as well (McGraw et al., 2007). In Arabidopsis the *ham1 ham2* double mutant is lethal, while keeping one the two genes heterozygous for the mutant allele resulted in a post-meiotic arrest of both female and male gametophyte development (Latrasse et al., 2008). This phenotype is similar to the one observed in the *REM_RNAi* as well as in the *35S:REM34_EAR* lines. Interestingly, *HAM1* and *HAM2* were also found to be involved in the control of flowering time via the epigenetic regulation of *FLOWERING LOCUS C* (*FLC*), which is also a target of VRN1, one of the four *REM* genes for which the function is known so far, suggesting a common mechanism throughout plant development. The artificial silencing of these two acetyltransferases causes an early flowering phenotype (Xiao et al., 2013) which was also noticed in the *REM_RNAi* lines (data not shown). The downregulation of *HAM2* and the phenotypical similarity between the *REM_RNAi* lines and the *HAM* downregulation suggest that these genes might be involved in the control of the same biological processes throughout Arabidopsis development. Further analysis will be needed to confirm the possible interaction between *REM34, REM35* and the histone acetyltransferases HAM1 and HAM2. The observed downregulation of the other analyzed target genes could be due to the general deregulation of transcription caused by the reduced expression of the chromatin remodelling factor *HAM2.*

While not much is known about the *REM* gene family, substantial information is available for other transcription factor families that are characterized by the presence of the B3 DNA binding domain. In particular, the well-characterized *Auxin Response Factor* (*ARF*) family, known to play a crucial role in regulating auxin responses, and the *Related to ABI3/VP1* (*RAV*) family, which was found to be involved in hormonal regulation during different stages of Arabidopsis development (Swaminathan et al., 2008). The plant hormone auxin was found to be involved in gametogenesis (Pagnussat et al., 2009; Panoli et al., 2015). Indeed perturbation of auxin transport in the embryo sac causes an arrest in the earliest stages of megagametogenesis (Ceccato et al., 2013). Auxin biosynthesis in the male gametophyte was also recently shown to be essential for the transition from microsporogenesis to microgamentogenesis (Yao et al., 2018). Because of the phenotypic similarities between the auxin defective mutant (Pagnussat et al., 2009; Panoli et al., 2015; Ceccato et al., 2013).and the *REM_RNAi* lines and because of the linkage between transcription factors containing the B3 DNA-binding domain and the regulation of hormonal responses, it is tempting to speculate that the role of *REM34, REM35* and *REM36* play in gametogenesis is also based on the regulation of a hormonal related processes.

In summary, we gained new information about the expression pattern and function of *REM34, REM35* and *REM36* during gametophyte development in Arabidopsis, those genes might control post-meiotic divisions in both embryo sac and pollen grains. These findings underline further the importance of *REM* genes during reproductive development in plants. Although these genes are often highly redundant and physically linked in the genome, slowly on we start to get a better understanding about their functions in plant development. Of course we just see the tip of the iceberg and still a huge amount of work has to be done to fully understand in detail the molecular and genetic mechanisms by which *REM* genes function.

Conflict of Interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Author Contributions

FC performed most of the experiments and wrote the manuscript. VB performed morphological analyses and contributed to writing the manuscript. OM and GL designed and employed the RNA interference and EAR constructs. MM designed and performed the CLS experiment and performed part of the back-crosses. RP made BiFC experiments. HHU and SFD designed the Y2H screening and helped FC with the experiment. AG designed the in-vitro pollen germination experiment and the Aniline blue analyses. MK contributed to the design of the experiments and helped writing the manuscript. VG designed the research, helped with the experiments and wrote the manuscript.

Funding

VG was supported by Ministero dell'Istruzione, dell'Università e della Ricerca MIUR, SIR2014 MADSMEC, Proposal number RBSI14BTZR. The post-doctoral fellowship of AG was supported by MIUR, SIR2014 MADSMEC, Proposal number RBSI14BTZR. The PhD fellowship of FC and RP were supported by the Doctorate School in Molecular and Cellular Biology, Università degli Studi di Milano. FC was supported by PROCROP-H20MC_RISE15LCOLO_M. RP was supported by H2020-MSCA-RISE-2015 ExpoSEED Proposal Number: 691109. Work in the SdF laboratory was financed by the Mexican National Council of Science and Technology (CONACyT) grants CB-2012-177739 and FC-2015-2/1061, and SdF acknowledges support of the Marcos Moshinsky Foundation and the European Union H2020-MSCA-RISE-2015 project ExpoSEED (grant no. 691109).

Acknowledgments

We thank Simona Masiero, Francesca Resentini, Lucia Colombo for helpful suggestions and valuable discussions. We also thank Annamaria Piva, Radha Cighetti and Francesco Gozzo from University of Milan, Toshiaki MITSUI and Marouane BASLAM Department of Applied Biological Chemistry Graduate School of

Science & Technology - Niigata University Ikarashi, Nishi-ku, Niigata, Japan and Ravishankar Palanivelu School of Plant Science-University of Arizona Tucson for their technical support.

Part of this work was carried out at NOLIMITS, an advanced imaging facility established by the Università degli Studi di Milano.

References

Acosta-García, G., and Vielle-Calzada, J.-P. (2004). A classical arabinogalactan protein is essential for the initiation of female gametogenesis in Arabidopsis. *Plant Cell* 16, 2614–28. doi:10.1105/tpc.104.024588.

Berger, F., and Twell, D. (2011). Germline Specification and Function in Plants. *Annu. Rev. Plant Biol.* 62, 461–484. doi:10.1146/annurev-arplant-042110-103824.

Bou Daher, F., Chebli, Y., and Geitmann, A. (2009). Optimization of conditions for germination of cold-stored Arabidopsis thaliana pollen. *Plant Cell Rep.* 28, 347–357. doi:10.1007/s00299-008-0647-1.

Brambilla, V., Battaglia, R., Colombo, M., Masiero, S., Bencivenga, S., Kater, M. M., et al. (2007). Genetic and Molecular Interactions between BELL1 and MADS Box Factors Support Ovule Development in Arabidopsis. *Plant Cell Online* 19, 2544–2556. 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, 84–7. Available at: http://www.ncbi.nlm.nih.gov/pubmed/9138536 [Accessed March 17, 2019].

BRIGGS, G., OSMONT, K., SHINDO, C., SIBOUT, R., and HARDTKE, C. (2006). Unequal genetic redundancies in Arabidopsis – a neglected phenomenon? *Trends Plant Sci.* 11, 492–498. doi:10.1016/j.tplants.2006.08.005.

Bucher, E., Lohuis, D., van Poppel, P. M. J. A., Geerts-Dimitriadou, C., Goldbach, R., and Prins, M. (2006). Multiple virus resistance at a high frequency using a single transgene construct. *J. Gen. Virol.* 87, 3697–3701. doi:10.1099/vir.0.82276-0.

Ceccato, L., Masiero, S., Sinha Roy, D., Bencivenga, S., Roig-Villanova, I., Ditengou, F. A., et al. (2013). Maternal Control of PIN1 Is Required for Female Gametophyte Development in Arabidopsis. *PLoS One* 8, 2–8. doi:10.1371/journal.pone.0066148.

Cecchetti, V., Celebrin, D., Napoli, N., Ghelli, R., Brunetti, P., Costantino, P., et al. (2017). An auxin maximum in the middle layer controls stamen development and pollen maturation in Arabidopsis. *New Phytol.* 213, 1194–1207. doi:10.1111/nph.14207.

Chantha, S. C., Gray-Mitsumune, M., Houde, J., and Matton, D. P. (2010). The MIDASIN and NOTCHLESS genes are essential for female gametophyte development in Arabidopsis thaliana. *Physiol. Mol. Biol. Plants* 16, 3–18. doi:10.1007/s12298-010-0005-y.

Clark, K. A., and Krysan, P. J. (2010). Chromosomal translocations are a common phenomenon in Arabidopsis thaliana T-DNA insertion lines. *Plant J.* 64, 990–1001. doi:10.1111/j.1365-313X.2010.04386.x.

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. Available at:

http://www.ncbi.nlm.nih.gov/pubmed/10069079 [Accessed March 17, 2019].

Curtis, M. J., Belcram, K., Bollmann, S. R., Tominey, C. M., Hoffman, P. D., Mercier, R., et al. (2009). Reciprocal chromosome translocation associated with TDNA-insertion mutation in Arabidopsis: genetic and cytological analyses of consequences for gametophyte development and for construction of doubly mutant lines. *Planta* 229, 731–45. doi:10.1007/s00425-008-0868-0.

Czarnecki, O., Bryan, A. C., Jawdy, S. S., Yang, X., Cheng, Z. M., Chen, J. G., et al. (2016). Simultaneous knockdown of six non-family genes using a single synthetic RNAi fragment in Arabidopsis thaliana. *Plant Methods* 12, 1–11. doi:10.1186/s13007-016-0116-8.

de Folter, S., and Immink, R. G. H. (2011). "Yeast Protein–Protein Interaction Assays and Screens," in *Methods in molecular biology (Clifton, N.J.)*, 145–165. doi:10.1007/978-1-61779-154-3_8.

de Folter, S., Immink, R. G. H., Kieffer, M., Parenicová, L., Henz, S. R., Weigel, D., et al. (2005). Comprehensive Interaction Map of the Arabidopsis MADS Box Transcription Factors. *PLANT CELL ONLINE* 17, 1424–1433. doi:10.1105/tpc.105.031831.

Ebel, C., Mariconti, L., and Gruissem, W. (2004). Plant retinoblastoma homologues control nuclear proliferation in the female gametophyte. *Nature* 429, 776–780. doi:10.1038/nature02637.

Fornara, F., Gregis, V., Pelucchi, N., Colombo, L., and Kater, M. (2008). The rice StMADS11-like genes OsMADS22 and OsMADS47 cause floral reversions in Arabidopsis without complementing the svp and agl24 mutants. *J. Exp. Bot.* doi:10.1093/jxb/ern083.

Franco-Zorrilla, J. M., Cubas, P., Jarillo, J. A., Fernández-Calvín, B., Salinas, J., and Martínez-Zapater, J. M. (2002). AtREM1, a member of a new family of B3 domain-containing genes, is preferentially expressed in reproductive meristems. *Plant Physiol.* 128, 418–27. doi:10.1104/pp.010323.

Gh Immink, R., An, I., ¤ T., De Folter, S., Shchennikova, A., Dj Van Dijk, A., et al. (2009). SEPALLATA3: the "glue" for MADS box transcription factor complex formation. *Genome Biol. Artic. R24 Genome Biol. Genome Biol.* 10. doi:10.1186/gb-2009-10-2-r24.

Gregis, W. V., Sessa, A., Colombo, L., Kater, M. M., Gregis, V., Sessa, A., et al. (2006). AGL24, SHORT VEGETATIVE PHASE, and APETALA1 Redundantly Control AGAMOUS during Early Stages of Flower Development in Arabidopsis. *PLANT CELL ONLINE* 18, 1373–1382. doi:10.1105/tpc.106.041798.

Hafidh, S., Fíla, J., and Honys, D. (2016). Male gametophyte development and function in angiosperms: a general concept. *Plant Reprod.* 29, 31–51. doi:10.1007/s00497-015-0272-4.

Huanca-Mamani, W., Garcia-Aguilar, M., Leon-Martinez, G., Grossniklaus, U., and Vielle-Calzada, J.-P. (2005). CHR11, a chromatin-remodeling factor essential for nuclear proliferation during female gametogenesis in Arabidopsis thaliana. *Proc. Natl. Acad. Sci.* 102, 17231–17236. doi:10.1073/pnas.0508186102.

Ingouff, M., Jullien, P. E., and Berger, F. (2006). The female gametophyte and the endosperm control cell proliferation and differentiation of the seed coat in Arabidopsis. *Plant Cell* 18, 3491–501. doi:10.1105/tpc.106.047266.

Johnston, A. J., Matveeva, E., Kirioukhova, O., Grossniklaus, U., and Gruissem, W. (2008). A Dynamic Reciprocal RBR-PRC2 Regulatory Circuit Controls Arabidopsis Gametophyte Development. *Curr. Biol.* 18, 1680–1686. doi:10.1016/j.cub.2008.09.026.

Klepikova, A. V., Kasianov, A. S., Gerasimov, E. S., Logacheva, M. D., and Penin, A. A. (2016). A high resolution map of the *Arabidopsis thaliana* developmental transcriptome based on RNA-seq profiling. *Plant J.*

88, 1058–1070. doi:10.1111/tpj.13312.

Latrasse, D., Benhamed, M., Henry, Y., Domenichini, S., Kim, W., Zhou, D. X., et al. (2008). The MYST histone acetyltransferases are essential for gametophyte development in arabidopsis. *BMC Plant Biol.* 8, 1–16. doi:10.1186/1471-2229-8-121.

Levy, Y. Y. (2002). Multiple Roles of Arabidopsis VRN1 in Vernalization and Flowering Time Control. *Science (80-.).* 297, 243–246. doi:10.1126/science.1072147.

Li, N., Yuan, L., Liu, N., Shi, D., Li, X., Tang, Z., et al. (2009). SLOW WALKER2, a NOC1/MAK21 Homologue, Is Essential for Coordinated Cell Cycle Progression during Female Gametophyte Development in Arabidopsis. *PLANT Physiol.* 151, 1486–1497. doi:10.1104/pp.109.142414.

Liu, J., Zhang, Y., Qin, G., Tsuge, T., Sakaguchi, N., Luo, G., et al. (2008). Targeted Degradation of the Cyclin-Dependent Kinase Inhibitor ICK4/KRP6 by RING-Type E3 Ligases Is Essential for Mitotic Cell Cycle Progression during Arabidopsis Gametogenesis. *Plant Cell Online* 20, 1538–1554. doi:10.1105/tpc.108.059741.

Mansfield, S. G., and Briarty, L. G. (1991). Early embryogenesis in *Arabidopsis thaliana* . II. The developing embryo. *Can. J. Bot.* 69, 461–476. doi:10.1139/b91-063.

Mantegazza, O., Gregis, V., Mendes, M. A., Morandini, P., Alves-Ferreira, M., Patreze, C. M., et al. (2014). Analysis of the arabidopsis REM gene family predicts functions during flower development. *Ann. Bot.* 114, 1507–1515. doi:10.1093/aob/mcu124.

Matias-Hernandez, L., Battaglia, R., Galbiati, F., Rubes, M., Eichenberger, C., Grossniklaus, U., 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, 1702–15. doi:10.1105/tpc.109.068627.

McGraw, S., Morin, G., Vigneault, C., Leclerc, P., and Sirard, M.-A. (2007). Investigation of MYST4 histone acetyltransferase and its involvement in mammalian gametogenesis. *BMC Dev. Biol.* 7, 123. doi:10.1186/1471-213X-7-123.

Melnyk C. W.,Molnar A. and Baulcombea D. C. (2011) Intercellular and systemic movement of RNA silencing signals. EMBO J. 2011; 30(17): 3553–3563. doi: 10.1038/emboj.2011.274

Mendes, M. A., Guerra, R. F., Castelnovo, B., Silva-Velazquez, Y., Morandini, P., Manrique, S., et al. (2016). Live and let die: a REM complex promotes fertilization through synergid cell death in Arabidopsis. *Development* 143.

Miki, D., Itoh, R., and Shimamoto, K. (2005). RNA silencing of single and multiple members in a gene family of rice. *Plant Physiol.* 138, 1903–13. doi:10.1104/pp.105.063933.

Mlotshwa S, Voinnet O, Mette MF, Matzke M, Vaucheret H, Ding SW, Pruss G, Vance VB. (2002). RNA silencing and the mobile silencing signal. Plant Cell.;14 Suppl:S289-301. DOI:10.1105/tpc.001677

Muralla, R., Lloyd, J., and Meinke, D. (2011). Molecular Foundations of Reproductive Lethality in Arabidopsis thaliana. *PLoS One* 6, e28398. doi:10.1371/journal.pone.0028398.

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 (80-.).* 324.

Panoli, A., Martin, M. V., Alandete-Saez, M., Simon, M., Neff, C., Swarup, R., et al. (2015). Auxin Import and Local Auxin Biosynthesis Are Required for Mitotic Divisions, Cell Expansion and Cell Specification during Female Gametophyte Development in Arabidopsis thaliana. *PLoS One* 10, e0126164. doi:10.1371/journal.pone.0126164.

Park, S. K., Howden, R., and Twell, D. (1998). The Arabidopsis thaliana gametophytic mutation gemini pollen1 disrupts microspore polarity, division asymmetry and pollen cell fate. *Development* 125, 3789–3799.

Peterson, R., Slovin, J., and Cheng Changbin (2010). A simplified method for differential staining of aborted and non-aborted pollen grains. *Int. J. Plant Biol.* 1, e13. Available at: https://www.pagepress.org/journals/index.php/pb/article/view/pb.2010.e13/2309 [Accessed March 17, 2019].

Pillus, L. (2008). MYSTs mark chromatin for chromosomal functions. *Curr. Opin. Cell Biol.* 20, 326–33. doi:10.1016/j.ceb.2008.04.009.

Resentini, F., Cyprys, P., Steffen, J. G., Alter, S., Morandini, P., Mizzotti, C., et al. (2017). SUPPRESSOR OF FRIGIDA (SUF4) Supports Gamete Fusion via Regulating Arabidopsis EC1 Gene Expression. *Plant Physiol.* 173, 155–166. doi:10.1104/pp.16.01024.

Richter, R., Kinoshita, A., Vincent, C., Martinez-Gallegos, R., Gao, H., van Driel, A. D., et al. (2019). Floral regulators FLC and SOC1 directly regulate expression of the B3-type transcription factor TARGET OF FLC AND SVP 1 at the Arabidopsis shoot apex via antagonistic chromatin modifications. *PLoS Genet.* 15, e1008065. doi:10.1371/journal.pgen.1008065.

Romanel, E. A. C., Schrago, C. G., Couñago, 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. doi:10.1371/journal.pone.0005791.

Sapountzi, V., and Côté, J. (2011). MYST-family histone acetyltransferases: beyond chromatin. *Cell. Mol. Life Sci.* 68, 1147–1156. doi:10.1007/s00018-010-0599-9.

Skopelitis D.S., Hill K., Klesen S., Marco C. F., von Born P, Chitwood D. H. and. Timmermans M.C.P. (2018) Gating of miRNA movement at defined cell-cell interfaces governs their impact as positional signals. Nat Commun. 2018; 9: 3107. doi: 10.1038/s41467-018-05571-0

Shi, D.-Q., Liu, J., Xiang, Y.-H., Ye, D., Sundaresan, V., and Yang, W.-C. (2005). SLOW WALKER1, essential for gametogenesis in Arabidopsis, encodes a WD40 protein involved in 18S ribosomal RNA biogenesis. *Plant Cell* 17, 2340–54. doi:10.1105/tpc.105.033563.

Smyth, D. R., Bowman, J. L., and Meyerowitz, E. M. (1990). Early flower development in Arabidopsis. *Plant Cell* 2, 755–767. doi:10.1105/tpc.2.8.755.

Sung, S., and Amasino, R. M. (2004). Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature* 427, 159–164. doi:10.1038/nature02195.

Swaminathan, K., Peterson, K., and Jack, T. (2008). The plant B3 superfamily. *Trends Plant Sci.* 13, 647–655. doi:10.1016/j.tplants.2008.09.006.

Takeuchi, H., and Higashiyama, T. (2016). Tip-localized receptors control pollen tube growth and LURE sensing in Arabidopsis. *Nature* 531, 245–248. doi:10.1038/nature17413.

Wang, S.-Q., Shi, D.-Q., Long, Y.-P., Liu, J., and Yang, W.-C. (2012). GAMETOPHYTE DEFECTIVE 1, a Putative Subunit of RNases P/MRP, Is Essential for Female Gametogenesis and Male Competence in Arabidopsis. *PLoS One* 7, e33595. doi:10.1371/journal.pone.0033595.

Xiao, J., Zhang, H., Xing, L., Xu, S., Liu, H., Chong, K., et al. (2013). Requirement of histone acetyltransferases HAM1 and HAM2 for epigenetic modification of FLC in regulating flowering in Arabidopsis. *J. Plant Physiol.* 170, 444–451. doi:10.1016/J.JPLPH.2012.11.007.

Yao, X., Tian, L., Yang, J., Zhao, Y. N., Zhu, Y. X., Dai, X., et al. (2018). Auxin production in diploid microsporocytes is necessary and sufficient for early stages of pollen development. *PLoS Genet.* 14, 1–19. doi:10.1371/journal.pgen.1007397.

Zhao, X., Bramsiepe, J., Van Durme, M., Komaki, S., Prusicki, M. A., Maruyama, D., et al. (2017). RETINOBLASTOMA RELATED1 mediates germline entry in Arabidopsis. *Science* 356, eaaf6532. doi:10.1126/science.aaf6532.

Figures legends

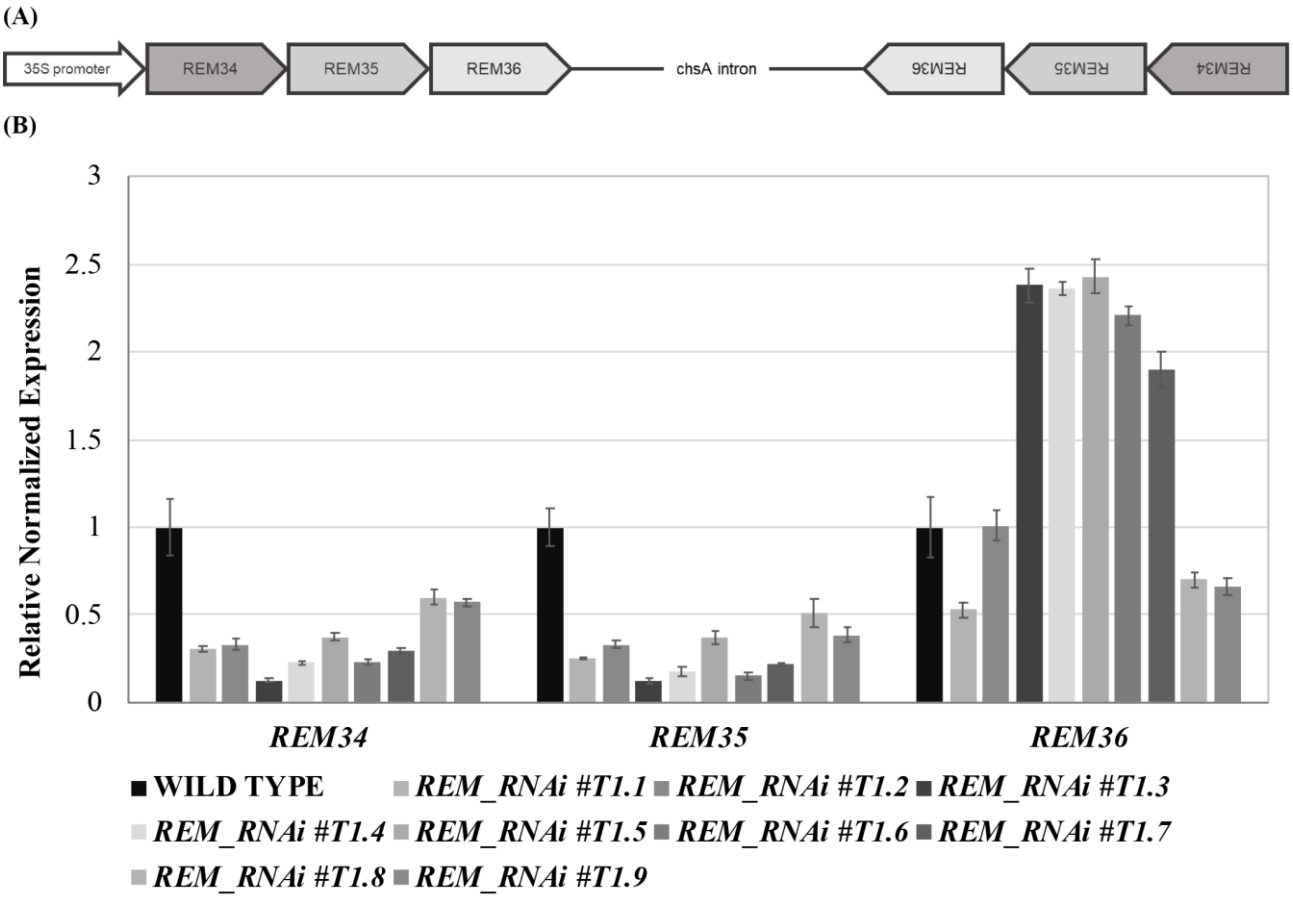


Figure 1

Multiple RNA interference lines

(A) Schematic representation of the RNAi construct. The REM34-REM35-REM36 sense and antisense fragments are separated by the chsA intron, to allow the hairpin structure formation. (B) qRT-PCR on 9 different *REM_RNAi* T1 inflorescences, showing a strong downregulation of *REM34* and *REM35* and different levels of *REM36* expression.

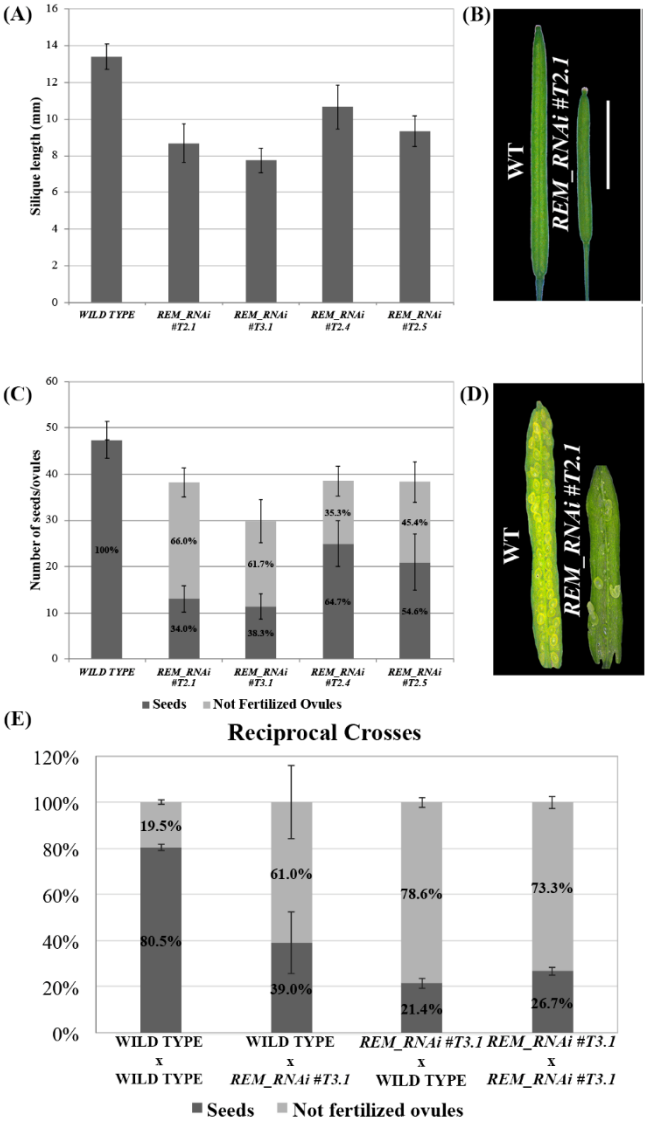


Figure 2

REM_RNAi lines have shorter siliques and a reduced seed set compared to the wild-type

(A) Graph showing the mean length of 10 wild-type and 10 *REM_RNAi #T2.1, #T3.1, #T2.4* and *#T2.5* siliques. A wild-type silique measures on average 13.4 mm, the siliques from the different *REM_RNAi* lines were found to measure on average between 7.8 and 10.7 mm. ($p<0.01$ for all comparison with the wild-type, ANOVA and Post Hoc Tukey HSD test were used). (B) Example of

wild-type and *REM_RNAi #T2.1* silique (Bar=5mm). (C) Graph showing the mean number of ovules/silique in the wild-type and *REM_RNAi #T2.1, #T3.1, #T2.4* and *#T2.5* plants, divided in seeds and not fertilized ovules. Compared to the wild-type situation, in which each silique contains on average 47.4 ovules, the *REM_RNAi* siliques have on average 29.8 to 38.5 ovules ($p<0.01$ for all comparison with the wild-type, ANOVA and Post Hoc Tukey HSD test were used). On average between 35.3% and 66.0% of ovules, depending from the analyzed line, failed to be fertilized, while no aborted ovules were detected in the wild type situation ($p<0.01$ for all comparison with the wild type, ANOVA and Post Hoc Tukey HSD test were used). (D) Example of wild-type and *REM_RNAi #T2.1* seed set (Bar=5mm).(E) Reciprocal crosses analysis between wild-type and *REM_RNAi #T3.1* plants. As a control both wild-type x wild-type and *REM_RNAi #T3.1* x *REM_RNAi #T3.1* crosses were performed. Crosses are indicated female x male. ($p<0.01$ for all comparison with the wild type of the non-fertilized ovules number, ANOVA and Post Hoc Tukey HSD test were used).

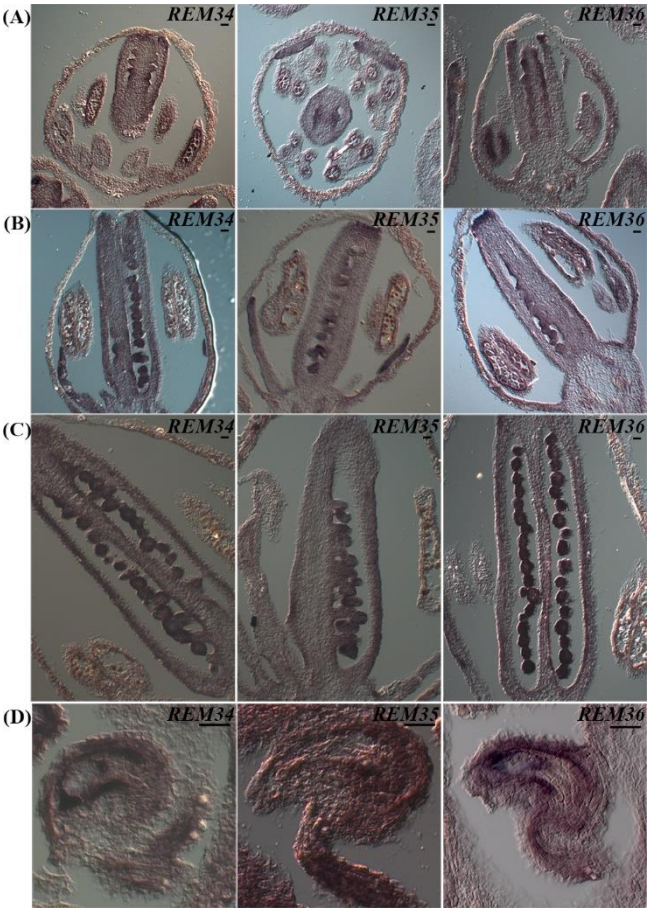


Figure 3

REM34, *REM35* and *REM36* expression pattern

(A) In flowers at ST8-9, the signal in the carpel is restricted in the tissue of the placenta and ovules primordia. At the same stage a clear signal is also visible in the anthers. (B) At ST10-11, the signal is present in the ovules, which are completing megagametogenesis, and in the anthers, where the pollen grains are undergoing the first mitotic division. (C) At ST12, when pollen reaches maturity, the signal is no longer visible in the anthers. (D) In flowers at anthesis, the target genes are expressed in the mature female gametophytes, in particular in the funiculus, inner integuments and central cell. Flower stages are described accordingly to Smyth et al., 1990. (Bar=20 µm).

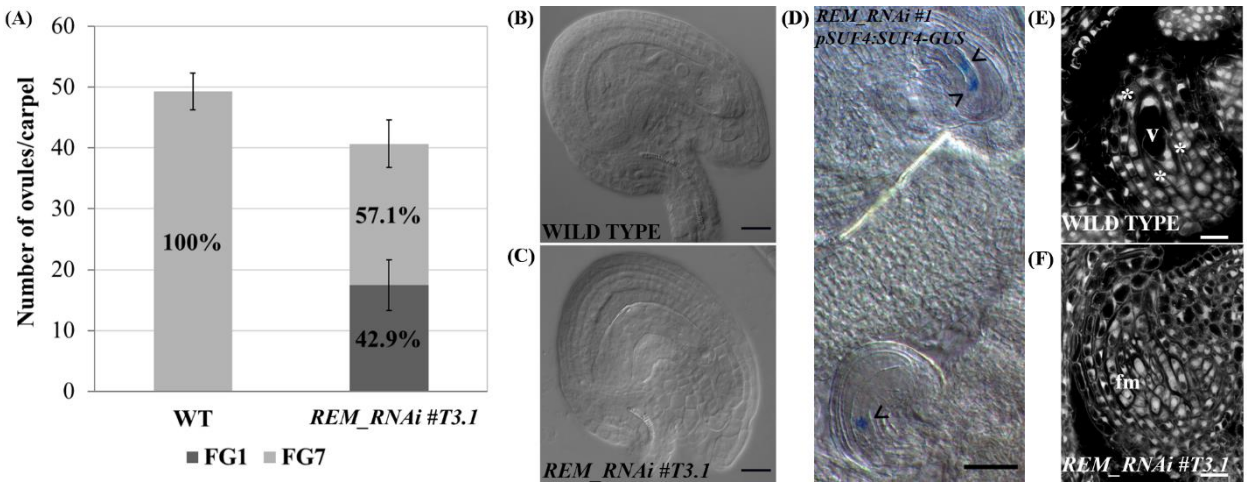


Figure 4

REM_RNAi #T3.1 female gametophyte characterization

(A) Analysis of cleared mature carpels of both wild-type (n=11) and *REM_RNAi #T3.1* (n=13). In wild-type mature carpels all the ovules reach the FG7 stage (542/542 ovules) while in the *REM_RNAi* line 227/529 ovules are arrested at the FG1 stage.(B)-(C) Cleared ovules collected from both wild-type (B) and *REM_RNAi #T3.1* (C) mature carpels. In the wild-type situation 100% of the embryo sac reach the FG7 stage, while in the RNAi line almost 60% of embryo sacs show an arrest in the FG1 stage (Bar=20µm).(D) *pSUF4:SUF4-GUS* in the *REM_RNAi* line. In the uppermost ovule 2 nuclei are stained, indicating the progression of gametogenesis till FG4 stage. In the lowest ovule 1 nucleus is stained indicating an arrest in FG1 stage. The arrowheads marked nuclei (Bar=50µm). (E)-(F) CLSM analysis of *REM_RNAi #T3.1* ovules. In the same carpel it was possible to observe ovules progressing in their development (E) and ovules arrested at the FG1 stage (F), asterisks indicate three out of the four nuclei. v = vacuole, fm = functional megaspore (Bar=10µm).

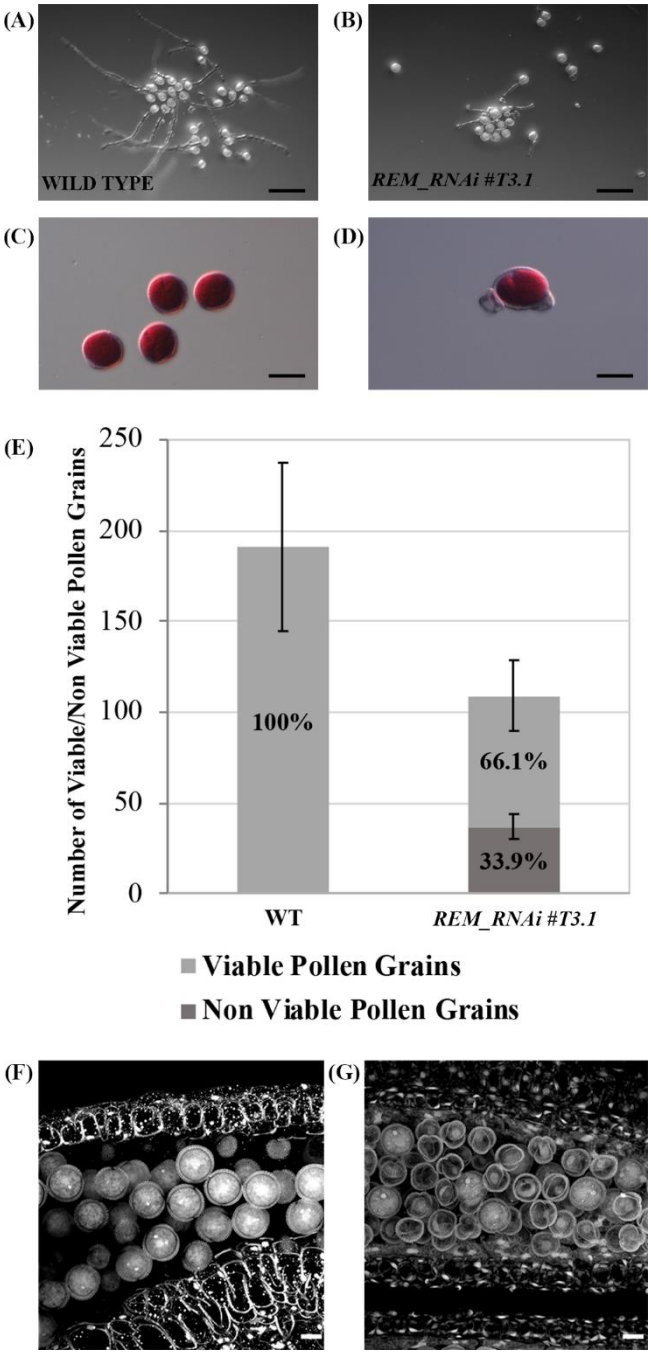


Figure 5

REM_RNAi #T3.1 male gametophyte characterization

(A)-(B) *In vitro* germination test of wild-type (A) and *REM_RNAi #T3.1* (B) pollen grains, plates were imaged 24h after plating (Bar=100µm).(C)-(D) Pollen grains, collected from mature anthers of both wild-type (C) and *REM_RNAi #T3.*1 anthers (D), were stained with Alexander's staining to check pollen grain viability. While all the wild type grains were viable, some *REM_RNAi#T3.*1 pollen grains appeared shrunken and unable to be stained in red (Bar=20µm). (E) Mature anthers from wild type and *REM_RNAi#T3.1* flowers were dissected, the released pollen was collected and treated

with Alexander's staining to discriminate between viable and non-viable pollen grains. In the wild type 100% of the pollen grains resulted vital while 33.9% of *REM_RNAi#T3.*1 pollen was found to be non-vital. (wt n= 1337, *REM_RNAi#T3.*1 *=874;* p<0.01 for all comparison with the wild-type, ANOVA and Post Hoc Tukey HSD test were used). (F)-(G) CLSM analysis of wild-type (F) and *REM_RNAi#T3.*1 (G) mature anthers. All the wild-type

grains are round and contain the vegetative nucleus and the two sperm cell nuclei, in the *REM_RNAi#T3.1* anthers it is possible to visualize both pollen grains at two nuclei stage, as well as degenerate pollen grains, without any visible nucleus (Bar=10µm).

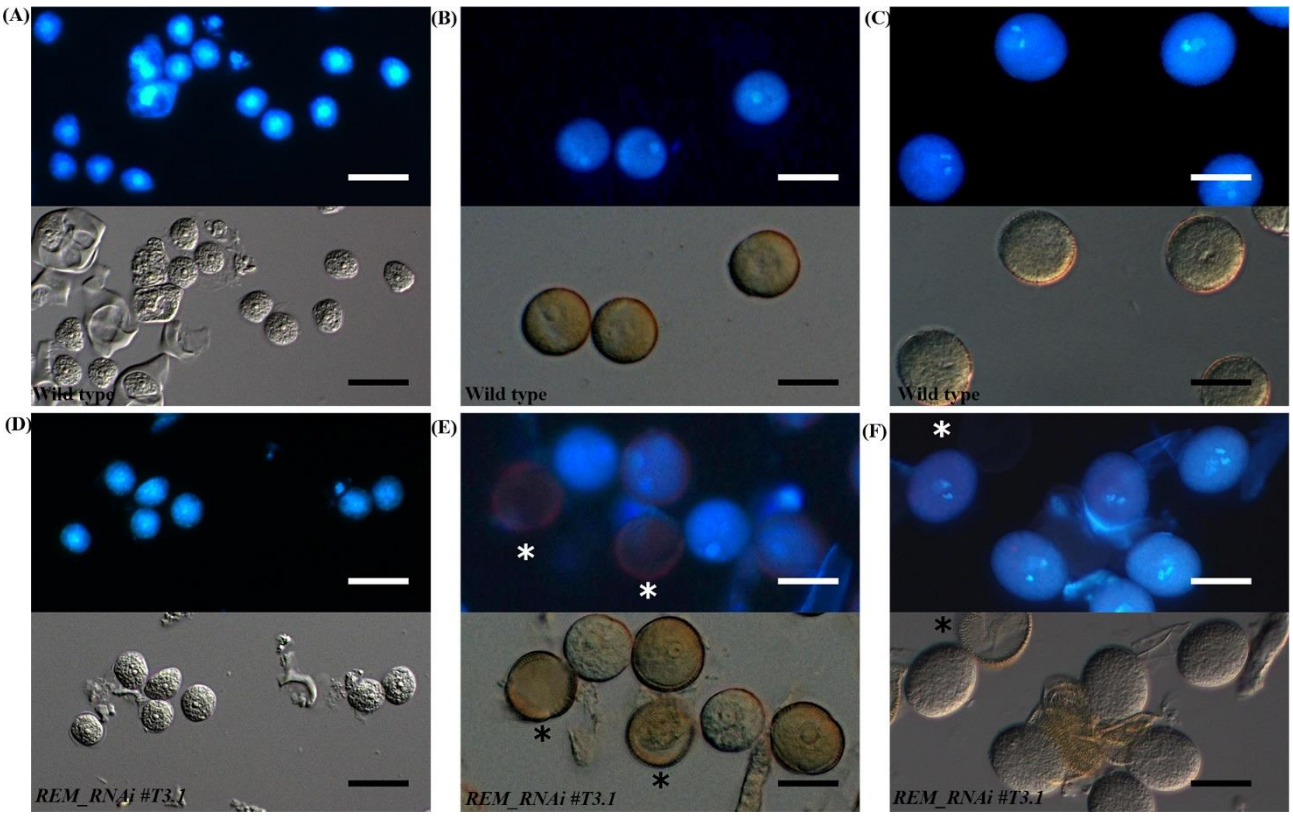


Figure 6

Wild type and *REM_RNAi* pollen development

DAPI staining of wild-type and *REM_RNAi#T3.*1 pollen grains at different developmental stages. At the microspore stage (A and D), all the grains contain a well-defined central nucleus. At the bicellular stage in all wild-type grains (B) the spermatic and vegetative nuclei are distinguishable, while in the *REM_RNAi#T3.*1 lines (E) some grains, marked with an asterisk, do not display any nucleus. Wild-type mature pollen grains (C), characterized by the presence of two sperm cells and

one vegetative nucleus. *REM_RNAi#T3.*1 mature pollen (F), the asterisk marks a mutant pollen grain without nucleus (Bar=10µm).

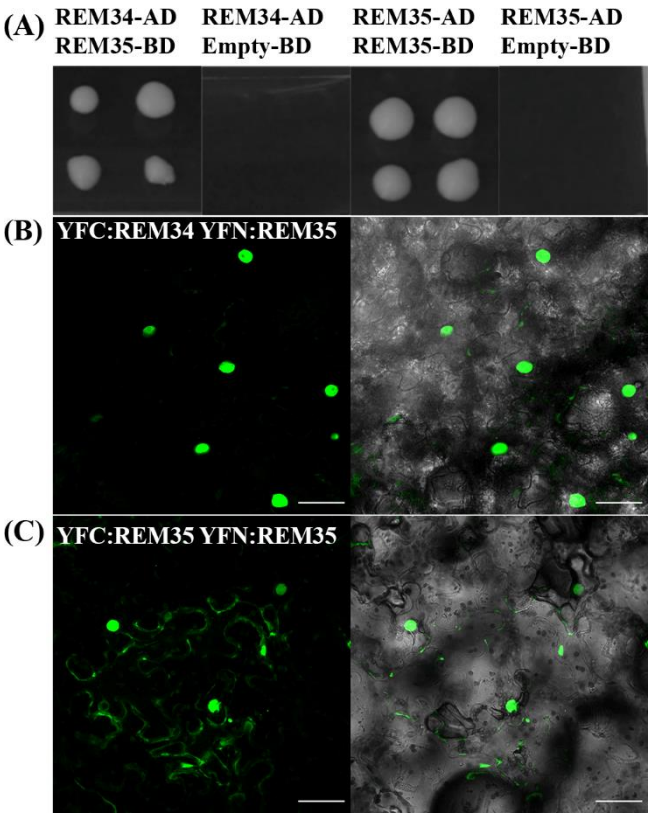


Figure 7

REM34 and REM35 interaction

(A) Yeast two hybrid assay showing the interaction between REM34 and REM35 and REM35 and REM35, on –L-W-H + 2.5 3-AT selective media. Empty pDEST32 vector was employed as a negative control. (B)-(C) BiFC experiments in tobacco leaf cells showing the reconstitute YFP fluorescence (green) between (B) REM34 and REM35 fusions to the C- and N-terminal fragments of YFP respectively (C) REM35 fusions to the C- and N-terminal fragments of YFP (Bar=50 µm).

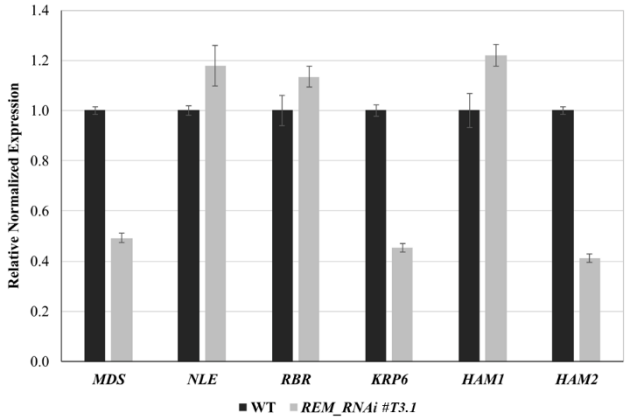


Figure 8

Expression analysis of genes involved in gametophyte development by qRT-PCR

Selected genes expression in inflorescence of wild-type and *REM_RNAi #T3.1*. The expression of selected genes was normalized to that of UBI and the expression level in Col was set to 1.

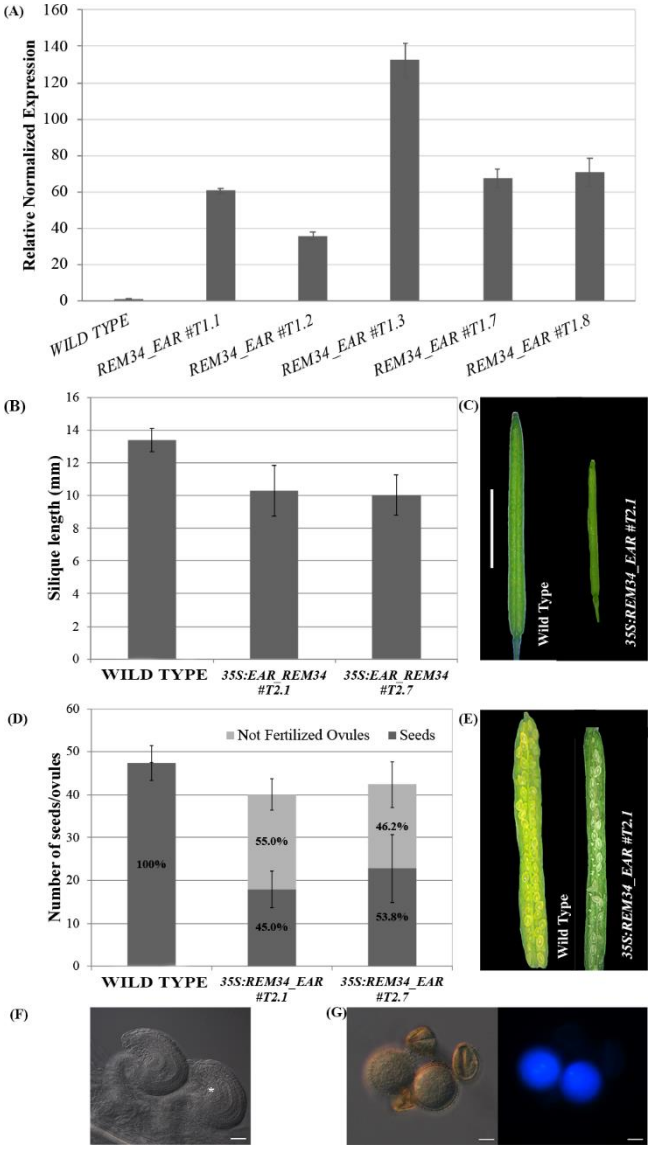


Figure 9

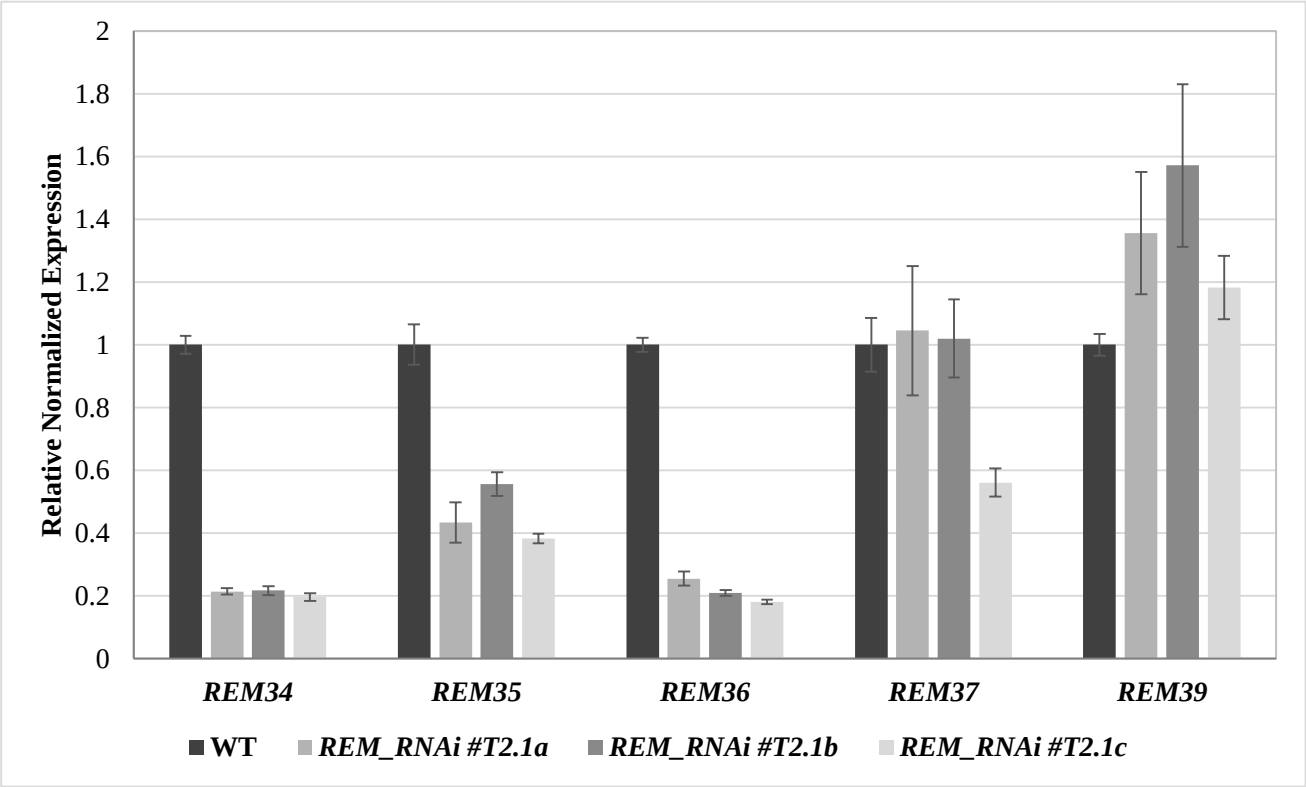
Analysis of the *35S:REM34_EAR* lines

(A) qRT-PCR for the evaluation of the *REM34_EAR* overexpression in 5 T1 transgenic lines. (B) Graph showing the mean length of 10 *REM34_EAR #T2.1 #T2.7* siliques, compared to the wild-type the two lines have a reduction in the silique length ($p<0.01$ for all comparison with the wild type, ANOVA and Post Hoc Tukey HSD test were used).(C) Example of wild-type and *REM34_EAR #T2.1* silique (Bar=5mm).(D) Graph showing the mean number of ovules/silique in the wild-type and *REM34_EAR #T2.1 #T2.7* plants. Both lines were characterized by a reduction in the total seed set of around 10% compared to the wild-type ($p<0.01$ for all comparison with the wild type, ANOVA and Post Hoc Tukey HSD test were used). On average between 55.0% and 46.2% of ovules, depending from the analyzed line, failed to be fertilized, while no aborted ovules were detected

in the wild type situation ($p<0.01$ for all comparison with the wild type, ANOVA and Post Hoc Tukey HSD test were used). (E) Example of wild-type and *REM34_EAR #T2.1* seed set (Bar=5mm).(F) Cleared ovules sampled from mature *REM34_EAR #T2.1* carpels, the asterisk marks the one blocked at the FG1 stage (Bar=20µm).

(G) DAPI stained pollen grains, sampled from mature *REM34_EAR #T2.1* anthers, some grains were able to reach the tricellular stage and showed fluorescent nuclei while other appeared degenerated and with no visible nucleus (Bar=20µm).

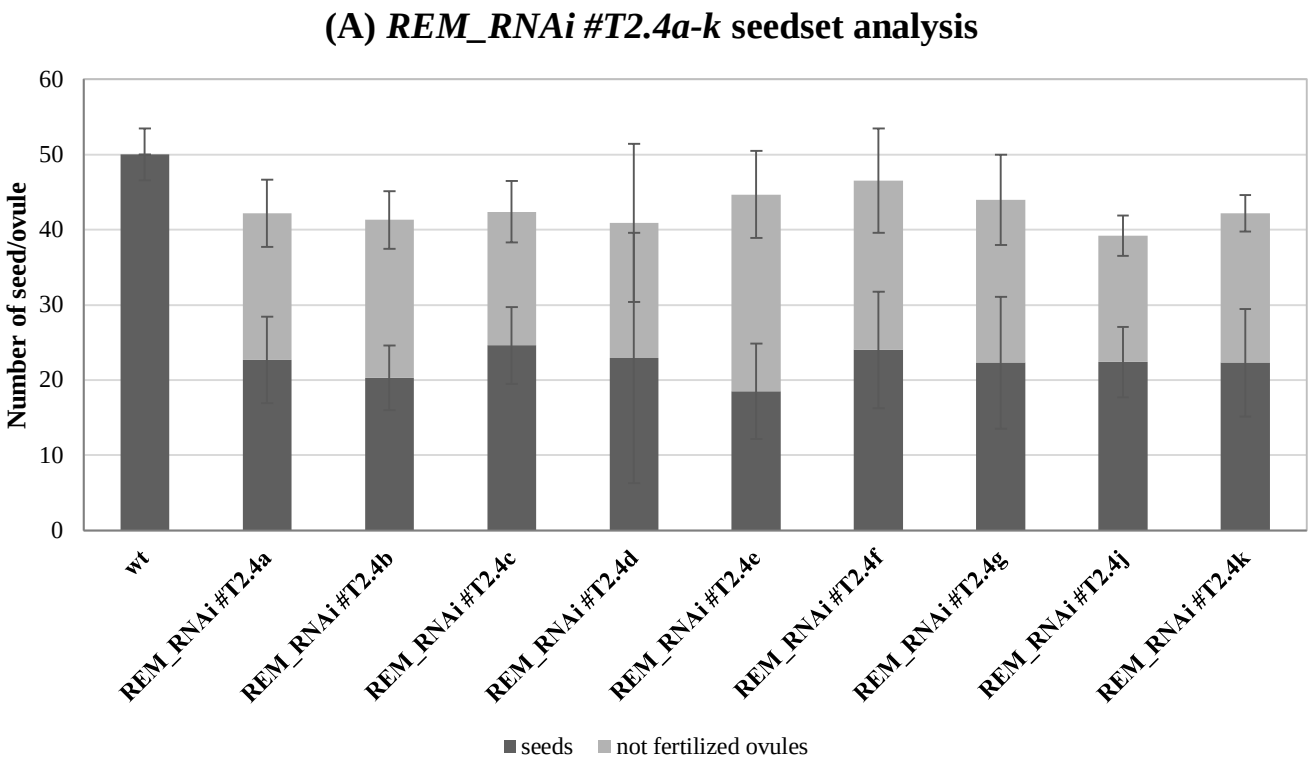
Supplementary Material

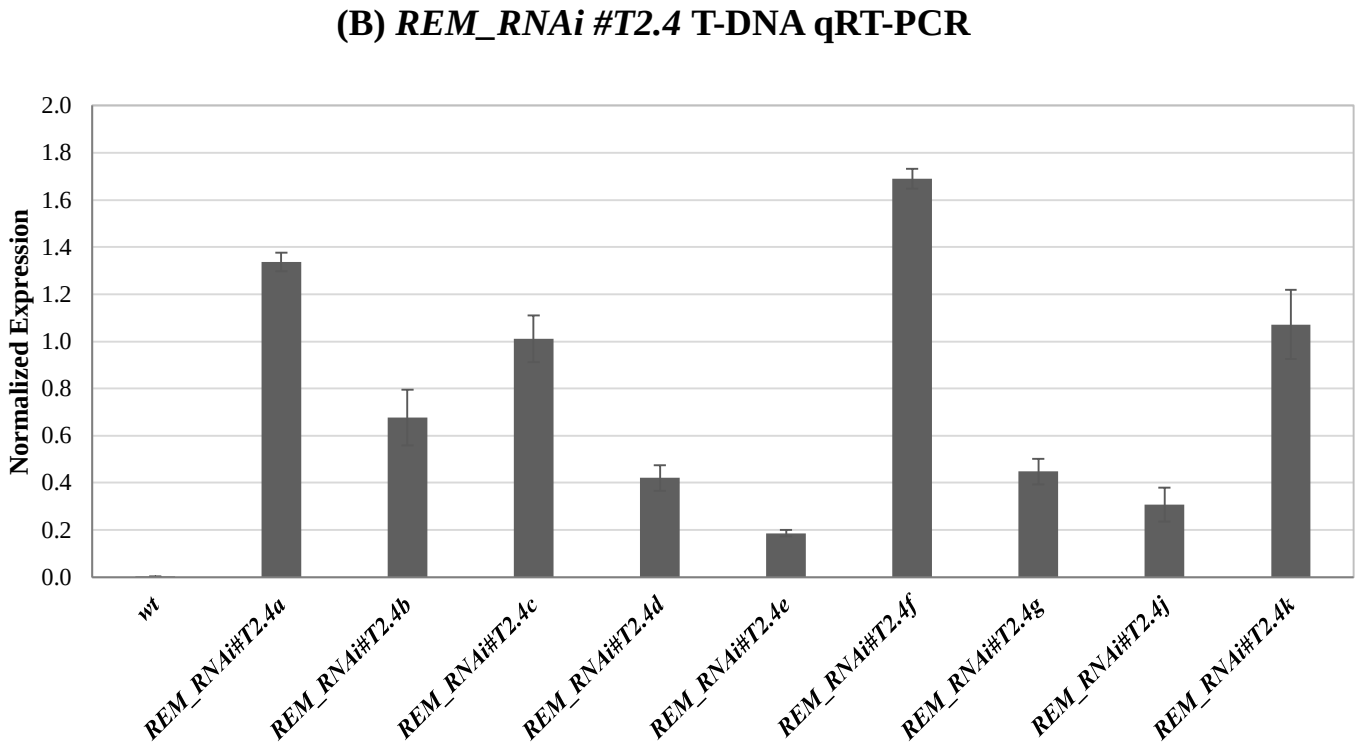


Supplementary Figure 1

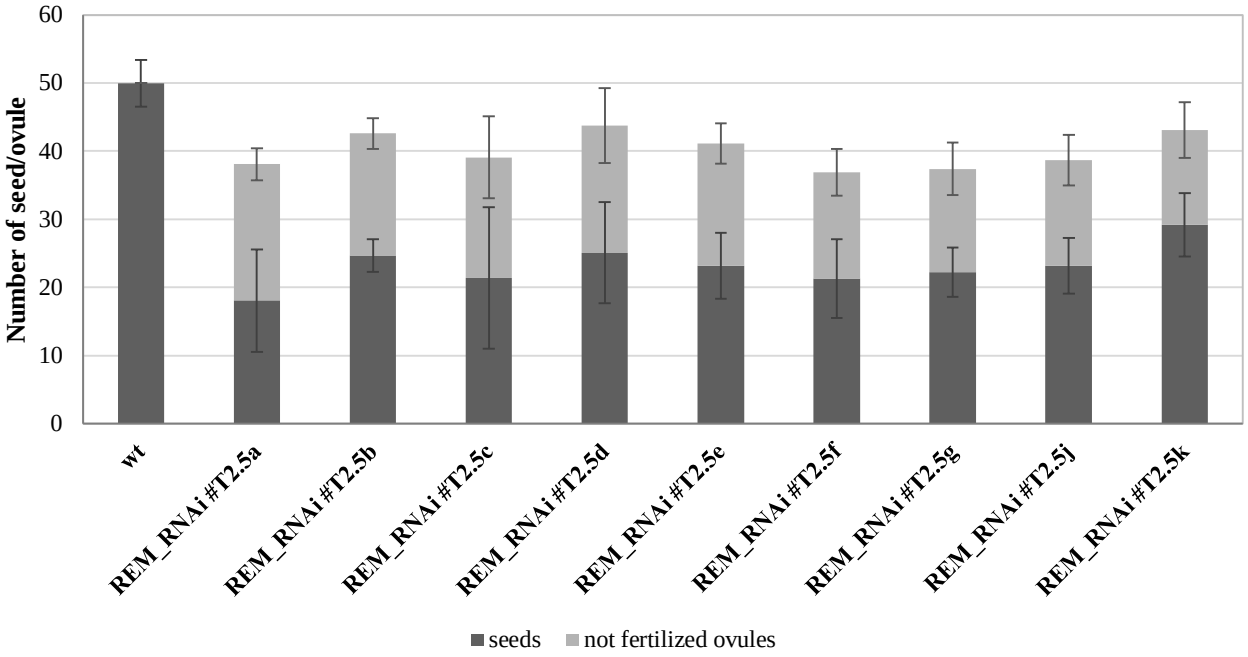
qRT-PCRs, performed on three different *REM_RNAi #T2.1* plants, to evaluate the specificity of the three fragments chosen for the RNA interference construct. Two genes were selected: *REM37*, which was the possible off target of one of the fragments, and *REM39*, that was chosen due to the fact that is the *REM* gene with the higher expression in the tissues of interest belonging to the same cluster as *REM34, REM35* and *REM36* (Mantegazza et al., 2014). Downregulation of neither *REM37* nor *REM39* was measured. The expression levels of *REM34, RERM35* and *REM36* were also

measured, to make sure that the silencing of the target genes was maintained in the T2 generation.

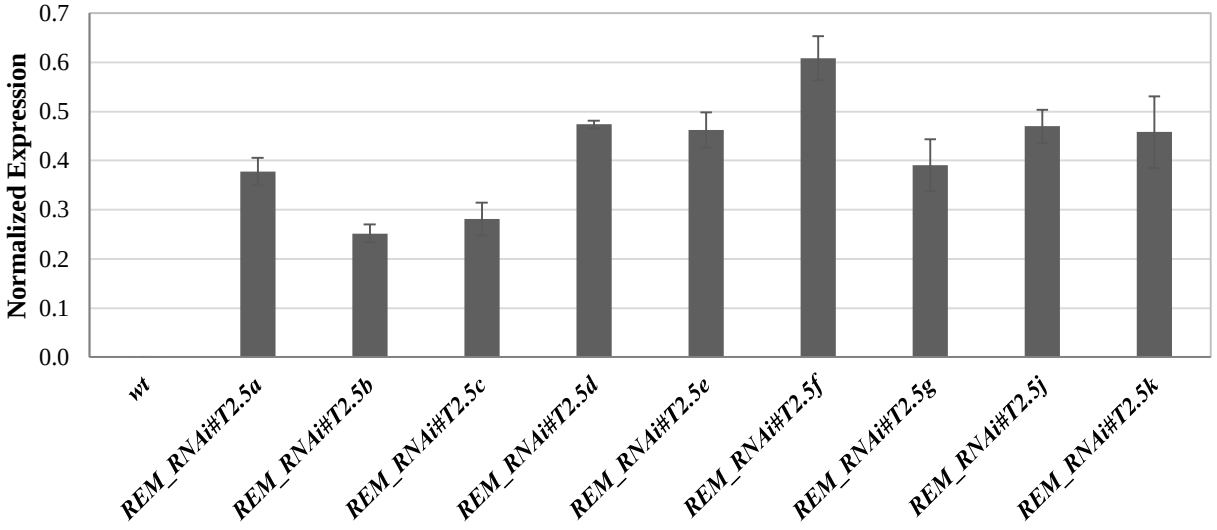




(C) *REM_RNAi #T2.5* seedset analysis



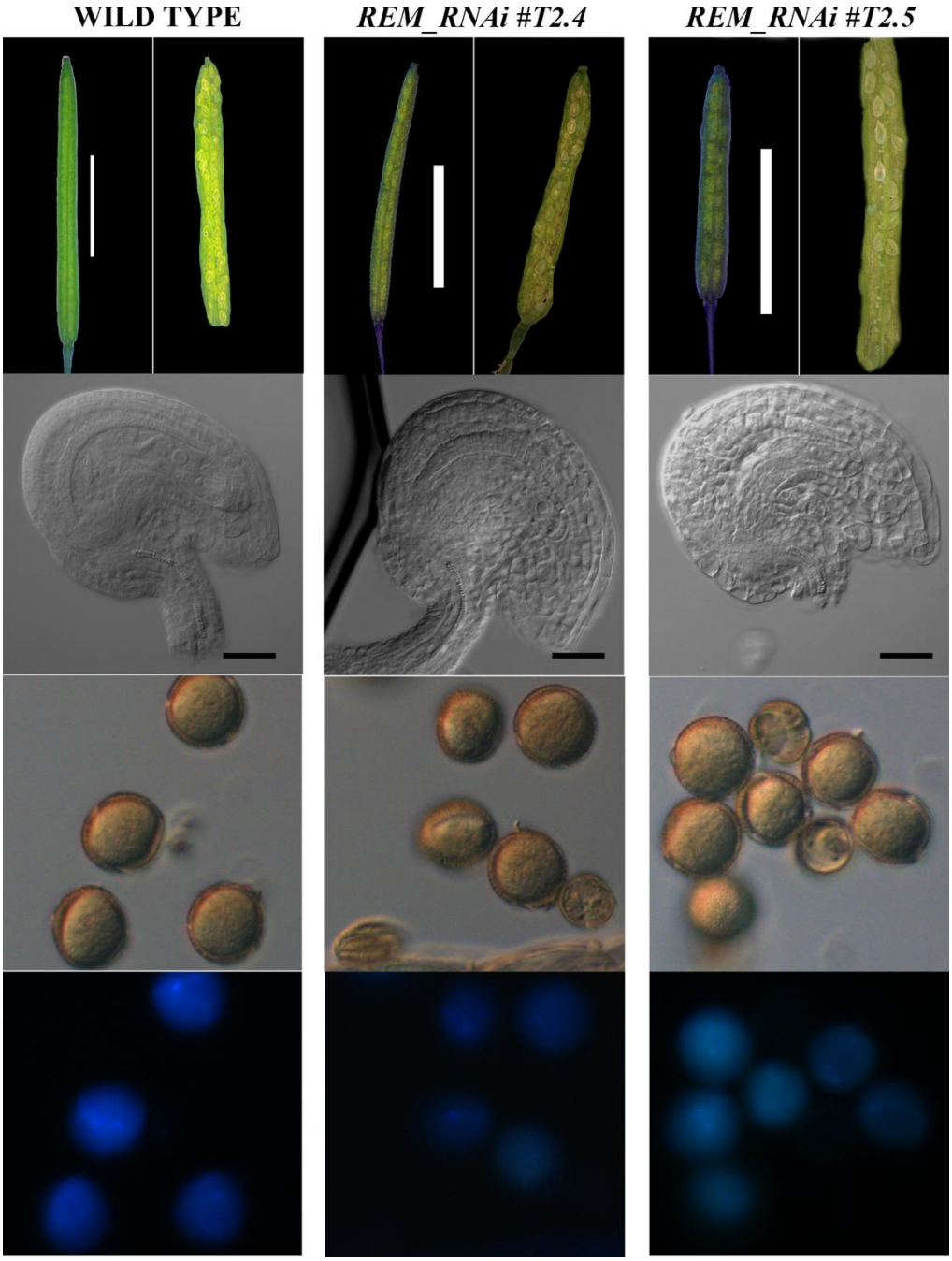
(D) *REM_RNAi #T2.4* T-DNA qRT-PCR



Supplementary Figure 2

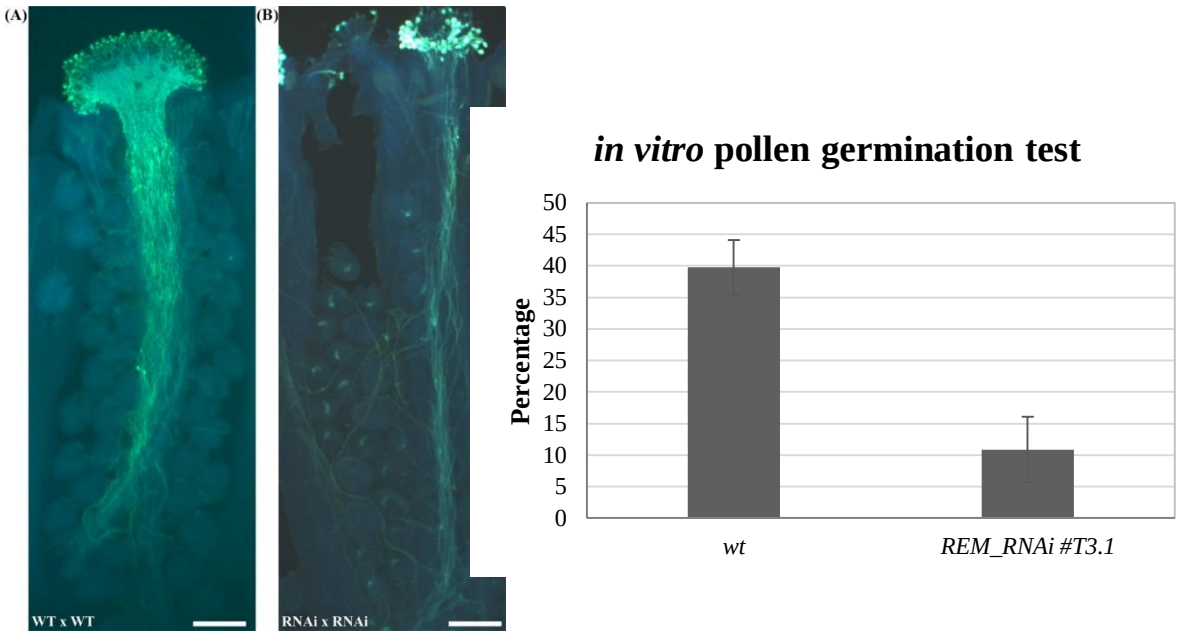
REM_RNAi T2 lines have a reduced seed set compared to the wild-type

(A) and (C) Graphs showing the mean number of ovules/silique in the wild-type and *REM_RNAi #1T2.4a* to *REM_RNAi #1T2.4j* and *REM_RNAi #1T2.5a* to *REM_RNAi #1T2.5j* plants resistant to the herbicide selection, divided in seeds and not fertilized ovules. Compared to the wild-type situation, in which each silique contains on average 50 ovules, the *REM_RNAi* siliques have on average 36.9 to 46.5 ovules. On average between 32.25% and 58.61% of ovules, depending from the analyzed line, failed to be fertilized, while no aborted ovules were detected in the wild type situation. Bars indicate the Standard deviation. (B) and (D) For each line an evaluation of T-DNA abundancy in each of the 9 plants analyzed, is presented. The RT-PCR analyses shows a various amount of T-DNA amplicon which is clearly unrelated to the ovule abortions and the overall seed set in both *REM_RNAi #1T2.4* and *REM_RNAi #1T2.5* lines. The primers used are on ACTIN7 used as normalizer and the herbicide resistance gene BAR used to estimate the copy of T-DNA.


WILD TYPE
REM_RNAi #T2.4
REM_RNAi #T2.5

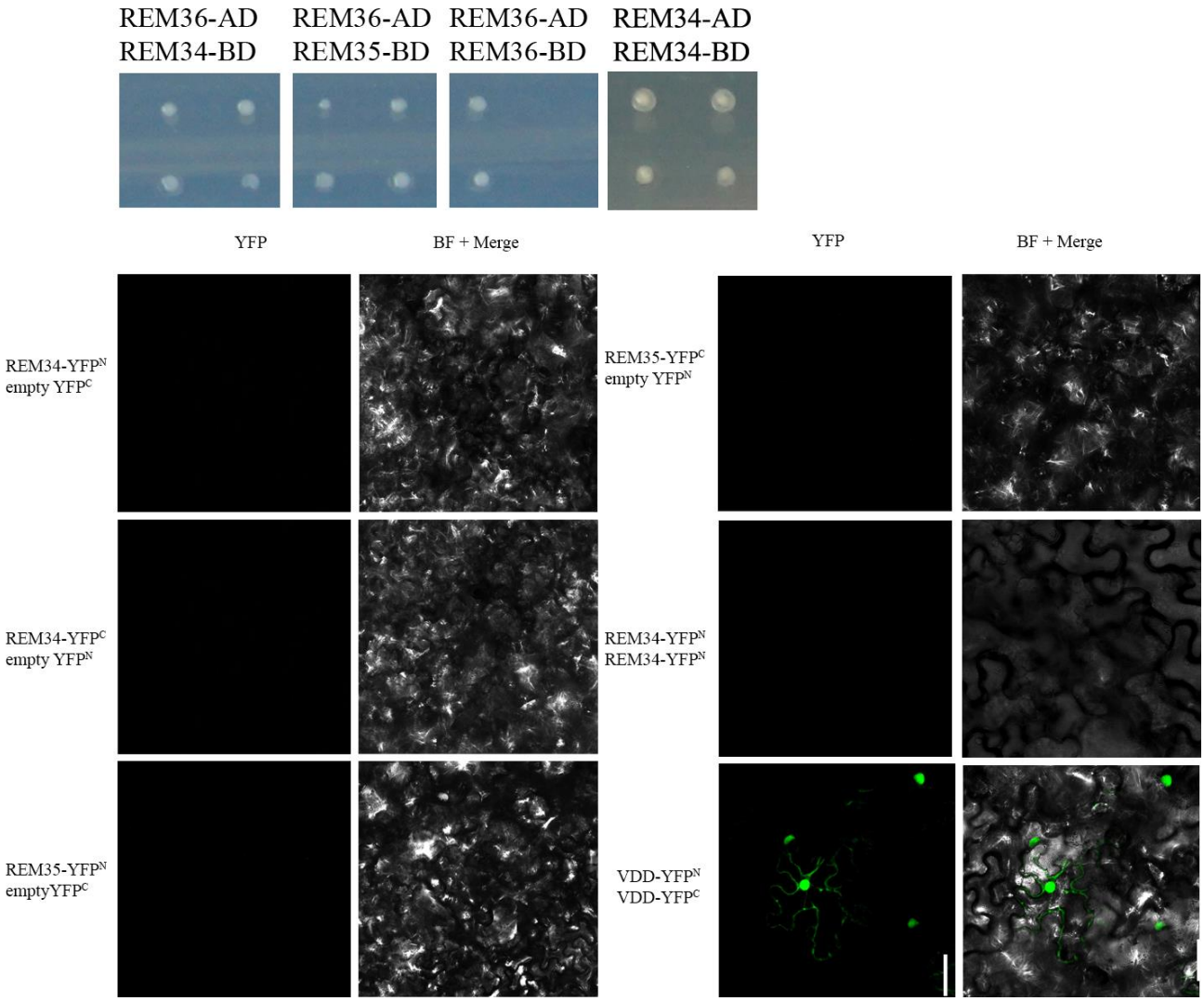
Supplementary Figure 3

Analysis of three different *REM_RNAi* T2 transgenic lines, all characterized by shorter siliques compared to the wild-type, a reduced seed set and ovule abortions. The majority of the embryo sacs of these lines showed a block in the earliest stages of megagametogenesis. Pollen collected from mature anthers is characterized by the presence of some degenerate grains, with no nucleus.(White bar=5mm, black bar=20µm)



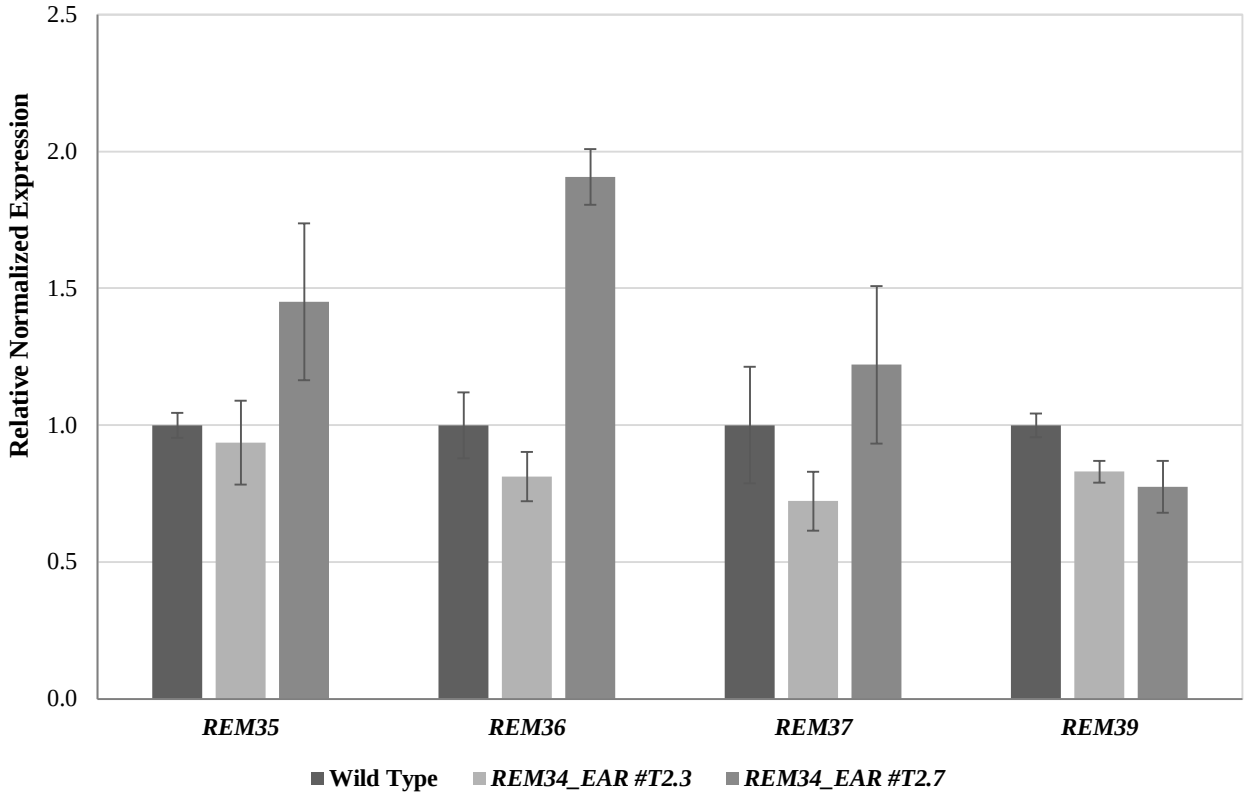
Supplementary Figure 4

*REM_RNAi #T3.*1 pollen germination test *in vitro,* showing a 30% decrease in the germination capability in the *REM_RNAi* line compared to the wild type. Aniline staining was employed to visualize pollen grains adhesion to the stigma and pollen tube germination and growth. Compared to the wild-type situation (A), a lower number of REM_RNAi #1 grains were able to adhere to the stigma and to germinate (B). In both cases the pollen tubes grow correctly until the end of the transmitting tract, In *REM_RNAi #1T3.1*x *REM_RNAi #1T3.1* cross, several ovules are not targeted by a pollen tube correlated to the fact that about 40% of the ovules are blocked in FG1 stage and so synergids are not properly formed (Bar=100µm).



Supplementary Figure 5

In the Y2H assays, no interactions were detected for REM36 and REM34 homodimers. Negative and positive controls for the BiFC experiment: all the constructs were cotranformed with the corresponding empty vector to test for false positive interaction. The REM34_REM34 interaction, which was found to be negative in the Y2H screening, was also employed as a negative control. Finally, as a positive control, the VDD_VDD interaction (M. Mendes et al, 2016) was tested.



Supplementary Figure 6

q-RTPCR on inflorescence on *REM_EAR #T2.1* and *#T2.7.* Compared to the wild type, none of the *REM* genes analyzed were found to be downregulated.

Supplementary Table 1: PRIMERS

GGTCTCACACCTGAAGTTTCCAAAGGAAAGG	REM34 FW golden gate cloning	
GGTCTCTATCTCTCTCCAACCTCTTC	REM34 REV golden gate cloning	
GGTCTCAAGATTCCAAGTCCAAGGACAAG	REM35 FW golden gate cloning	
GGTCTCGTGTCAACAATAATCTGTTTC	REM35 REV golden gate cloning	
GGTCTCAGACATCATCAAGTCTAGAAGGGAAG	REM36 FW golden gate cloning	
GGTCTCGCCTTAATCATCCCACAAGCACAC	REM36 REV golden gate cloning	
CACCATGGCGGATCCACCACATTTC	REM34 FW for REM34_EAR fusion	
CTAAGCAAATCCAAGTCTAAGTTCAAGATCAAGATCAAGAACCAGATTACTGCTGAGG	REM34 RV for REM34_EAR fusion	
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCGGATCCACCACATTTCTC	REM34 FW for CDS cloning GW	
GGGGACCACTTTGTACAAGAAAGCTGGGTTCAAACCAGATTACTGCTGAGG	REM34 REV for CDS cloning GW	
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGATGATCCAGCAATTTC	REM35 FW for CDS cloning GW	
GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATGGCGAATCATCCACTA	REM36 FW for CDS cloning GW	
GGGGACCACTTTGTACAAGAAAGCTGGGTCTTACTTGAGGATTTTGTTGATTTCCG	REM35/REM36 RV for CDS cloning GW	
CACTCAGGTTTCATCACAGCACG	REM34 FW in situ probe	From Mantegazza et a 2014
TAATACGACTCACTATAGGGTAGCCCATTAGCGCAGCAGAAG	REM34 REV +T7 in situ probe	From Mantegazza et a 2014
TCCTATGTAGCTTCTGGCGATGG	REM35 FW in situ probe	From Mantegazza et a 2014
TAATACGACTCACTATAGGGAGTCTCCCCTCCTTCATCAAATGG	REM35 REV +T7 in situ probe	From Mantegazza et a 2014
CTCACTGCTTCCAACCTACG	REM36 FW in situ probe	From Mantegazza et a 2014
TAATACGACTCACTATAGGGAGCGTCCACGGATAAAAGCCTG	REM36 REV +T7 in situ probe	From Mantegazza et a 2014
AGCTTGTGAGACTGCTCCAC	REM34 FW expression analysis	
CCTGATCGGAGACTGAGCAC	REM34 REV expression analysis	
CATTTGATGAAGGAGGGGAGAC	REM35 FW expression analysis	
CTTTCTAGCTCTGACCGAATCC	REM35 REV expression analysis	
TCACTTGCTGGACACACCTC	REM36 FW expression analysis	
TCGTCTCGAAGACAGTGTGC	REM36 REV expression analysis	
TGGCATAGAGTGGAAGTCGCATC	REM37 FW expression analysis	
GTCATTCGGGGTTTCCTATCC	REM37 REV expression analysis	
GGAGAAGTTTCTGCCGTGAG	REM39 FW expression analysis	
GGTCACTGGCCACTCTTCTC	REM39 REV expression analysis	
GCAAGCTCAGTGGTGACTAC	MDS FW expression analysis	
ACATCCACTTTCTGACATGC	MDS REV expression analysis	
GTTAACCGTTGCTCACAGAC	NLE FW expression analysis	

GCCTTTGCAAGTAAACAATG	NLE REV expression analysis	
AGTCGCCTGCTGCTAAGACAAA	RBR FW expression analysis	
ATGACAGTCCTGAGCCACTTGG	RBR REV expression analysis	
CTTCTGATTCATCACCGGACTC	KRP6 FW expression analysis	From Liu J et al. 2008
ACACCAAACGACGAACTGTTCT	KRP6 REV expression analysis	From Liu J et al. 2008
ATGGGATCGTCTGCGGATACA	HAM1 FW expression analysis	From Latrasse et al., 2
GAATTCGTGAGAGCGAGTATCGCA	HAM1 REV expression analysis	From Latrasse et al., 2
CCTTTAACTCCTGATCAAGCTAT	HAM2 FW expression analysis	From Latrasse et al., 2
CTACAGCGCACTCTACTGAATC	HAM2 REV expression analysis	From Latrasse et al., 2
ATCTCGGTGACGGGCAGGACC	BAR gene FW	
TCTACACCCACCTGCTGAAG	BAR gene REV	
CGTTTCGCTTTCCTTAGTGTTAGCT	Actin 7 FW	
AGCGAACGGATCTAGAGACTCACCTTG	Actin 7 REV	

Manuscript #3: Spatiotemporal restriction of *FUSCA3* expression by class I BPC promotes ovule development and coordinates embryo and endosperm growth

This manuscript explores the role of *FUSCA3* in endosperm growth. Here, I generated the *35S*:BPC1 that led to the constitutive expression of *BPC1* throughout plant development. On those lines, I performed several ChIP assays in distinct reproductive tissues to determine whether the BPC factor could bind the promoter of FUS3. My results show that FUS3 is a direct target of BPC1 in *vivo*.

Preprint on bioRxiv; in revision to Plant Cell

Spatiotemporal restriction of *FUSCA3* expression by class I BPC promotes ovule development and coordinates embryo and endosperm growth

Jian Wu [1,4], Rosanna Petrella [3], Sebastian Dowhanik[1], Veronica Gregis [3] and Sonia Gazzarrini[1,2,§]

[1] Department of Biological Sciences, University of Toronto Scarborough, 1265 Military Trail, Toronto, ON Canada, M1C 1A4

[2] Department of Cell and Systems Biology, University of Toronto, 25 Willcocks Street, Toronto, ON Canada, M5S 3B2

[3] Università degli Studi di Milano, Dipartimento di Bioscienze, via Celoria 26 20133 Milano, Italia

[4] Beijing Key Laboratory of Development and Quality Control of Ornamental Crops, Department of Ornamental Horticulture and Landscape Architecture, China Agricultural University, Beijing, China

[§] Corresponding author: e-mail gazzarrini@utsc.utoronto.ca

Running Title: *FUSCA3* repression by BPCs in stems, ovules, seeds

Abstract

Spatiotemporal regulation of gene expression plays an important role in developmental timing in plants and animals. The transcription factor *FUSCA3 (FUS3)* regulates developmental phase transitions by acting as a link between hormonal pathways in Arabidopsis. However, the mechanisms governing its spatiotemporal expression pattern are poorly understood. Here, we show that *FUS3* is expressed in stems, chalaza and funiculus of mature ovules and seeds, but is repressed in the embryo sac, integuments and endosperm. *FUS3* repression requires class I BASIC PENTACYSTEINE (BPC) proteins, which directly bind GA/CT *cis*-elements in *FUS3* and restrict its expression pattern. During vegetative and reproductive development, derepression of *FUS3* in *bpc1/2* or *pML1:FUS3* misexpression lines results in dwarf plants carrying defective flowers and aborted ovules. Post-fertilization, ectopic *FUS3* expression in *bpc1/2* or *pML1:FUS3* endosperm increases endosperm nuclei proliferation and seed size, leading to delayed or arrested embryo development. These phenotypes are rescued in *bpc1/2 fus3-3*. Lastly, class I BPCs interact with FIS-PRC2 (FERTILIZATION-INDEPENDENT SEED-Polycomb Repressive Complex 2), which represses *FUS3* in the endosperm. We propose that BPC1/2 promote the transition from reproductive to seed development by repressing *FUS3* in ovule integuments. After fertilization, BPC1/2 and FIS-PRC2 repress *FUS3* in the endosperm to coordinate endosperm and embryo growth.

Introduction

Throughout their lifetime, plants integrate endogenous and environmental signals to correctly time the expression of developmental genetic programs. Plants transition through three major phases of development during their life cycle: vegetative, reproductive and seed development. Developmental phase transitions are characterized by large changes in gene expression, which are under flexible epigenetic regulation and respond to developmental and environmental cues (Mozgova and Henning, 2015; Mozgova et al., 2015).

Reproductive development in seed plants starts with the production of female and male gametes, and is followed by fertilization and seed development. During ovule development, the maternal sporophytic integuments originate from the chalaza and enclose the female gametophyte (embryo sac), which contains two gametes: the haploid egg cell and the diploid central cell (Gasser and Skinner, 2019). After fertilization of the central cell, the triploid endosperm nuclei undergo multiple rounds of division, which are followed by cellularization. In most Angiosperms the function of the endosperm is to nourish the developing embryo. Fertilization of the egg cell generates the diploid zygote, which divides multiple times to form the embryo (Lafon-Placette and Kohler, 2014; Dresselhaus et al., 2016; Gasser and Skinner, 2019). Auxin is a major player in establishing apical-basal polarity and patterning of the embryo, as well as regulating integuments and endosperm development (Figueiredo et al., 2015; Figueiredo et al., 2016; Robert et al., 2018; Lau et al., 2012; de Vries and Weijers, 2017). In the absence of fertilization, seed development is repressed by the Polycomb-Repressive Complex2 (PRC2). In particular, the FIS-PRC2 complex represses autonomous endosperm development, while EMF-PRC2 and VRN-PRC2 prevent seed coat development prior to fertilization. Mutations in FIS-PRC2 or impairment of auxin synthesis and signaling leads to seed abortion (Roszak and Kohler, 2011; Figueiredo and Kohler, 2018; Robert, 2019).

Seed maturation is characterized by cell expansion and very little cell division. During this stage of development, the embryo accumulates seed storage compounds, acquires dormancy and establishes desiccation tolerance. These processes are largely controlled by the hormone abscisic acid (ABA) and the LAFL genes, which include the B3 domain transcription factors LEAFY COTYLEDON2 (LEC2), ABSCISIC ACID INSENSTIVE3 (ABI3) and FUSCA3 (FUS3), as well as the NF-YB subunits of the CCAAT-binding complex, LEC1 and LEC1-LIKE. LAFL genes are expressed during seed development, where they promote seed mturation while inhibiting germination, and some

of them do so through hormones (Sreenivasulu and Wobus, 2013; Jia et al., 2014; Fatihi et al., 2016; Carbonero et al., 2017; Lepiniec et al., 2018). The heterochronic gene *FUS3* promotes seed maturation by increasing ABA levels while inhibiting vegetative growth and flowering by repressing gibberellins (GA) synthesis. These hormones feed back by positively (ABA) and negatively (GA) regulating FUS3 levels (Keith et al., 1994; Curaba et al., 2004; Gazzarrini et al., 2004; Chiu et al., 2016). FUS3 also inhibits germination and vegetative phase change by repressing ethylene signaling (Lumba et al., 2012). Thus, *FUS3* regulates phase transitions by modulating hormones syntheses/signaling.

During germination, the seed maturation program is repressed by epigenetic mechanisms, which lead to dormancy break and the transition to the next phase of development; these include: CHROMODOMAIN HELICASE DNA BINDING3 (CHD3)/PICKLE (PKL)-dependent chromatin remodeling; Polycomb Repressive Complex2 (PRC2)-mediated histone 3 lysine 27 trimethylation (H3K27me3); H2AK121ub monoubiquitination by the PRC1 components RING-finger homologs AtBMI1A and AtBMI1B; and VIP1/ABI3/LEC (VAL) mediated recruitment of histone deacetylases (HDAC) and PRC complexes (Jia et al., 2014; Lepiniec et al., 2018). Mutations that affect these processes result in *LAFL* derepression, leading to expression of seed-specific traits and development of embryonic structures during vegetative growth. Accordingly, ectopic expression of *LAFL* genes post-embryonically results in similar phenotypes (Lotan et al., 1998; Stone et al., 2001; Gazzarrini et al., 2004; Braybrook et al., 2006). Epigenetic regulation of *LAFL* genes has also been observed during early embryonic development. For example, *FUS3* is ectopically expressed in the endosperm of the PRC mutant *medea* (*mea)* (Makarevich et al., 2006), but the mechanism and function of *FUS3* repression in this tissue is unknown.

Recently, we have shown that FUS3 plays a critical role also in reproductive development. The *fus3-3* loss-of-function mutant displays seed abortion, which is enhanced in plants grown at elevated temperature and dependent on FUS3 phosphorylation (Chan et al., 2017; Tsai and Gazzarrini, 2012). *pML1:FUS3-GFP* plants that mis-express *FUS3* during reproductive development also show aborted siliques, suggesting that spatiotemporal expression of *FUS3* must be tightly regulated at this stage of development (Gazzarrini et al., 2004). Here, we show that class I BASIC PENTACYSTEINE (BPC) proteins interact the FIS-PRC2 complex and directly bind to the *FUS3* chromatin. BPC1/2 repress *FUS3* in the stem, integuments of mature ovules, as well as in the endosperm of developing seeds. *FUS3* misexpression in the *bpc1-1 bpc2 (bpc1/2)* mutant and *pML1:FUS3-GFP* misexpression lines reduces plant height, impairs the development of flowers,

ovules and endosperm leading to seed abortion or arrested embryogenesis. The *bpc1/2* phenotypes can be partially rescued in the *bpc1/2 fus3-3* background, strongly indicating that they are partly caused by ectopic *FUS3* expression. We propose that during reproductive development BPC1/2- and PRC2-mediated repression of *FUS3* is necessary for ovule development, while after fertilization *FUS3* repression in the endosperm by BPC1/2 and FIS-PRC2 is required to coordinate endosperm and embryo growth. Hence, correct spatiotemporal expression of *FUS3* regulates the transition from plant reproduction to seed development and from embryo pattern formation to seed maturation.

Results

FUS3 localizes to reproductive organs before fertilization and is required for ovule development

The *fus3-3* loss-of-function mutant displays seed abortion, which is enhanced at elevated temperature (Chan et al., 2017). To investigate the role of *FUS3* in reproductive development, we first determined FUS3 localization pattern in flower buds using a *pFUS3:FUS3-GFP* translational reporter (Gazzarrini et al., 2004). However, no FUS3-GFP fluorescence was detected, likely due to the fast turnover rate of FUS3 (Lu et al., 2010). We then used a *pFUS3:FUS3ΔC-GFP* reporter, which lacks the PEST instability motif of FUS3 and allows detection of low FUS3 protein levels (Lu et al., 2010). This reporter is non-functional (it doesn't rescue *fus3-3*), but recapitulates *FUS3* expression patterns determined by qRT-PCR, *pFUS3:GUS* and *pFUS3:GFP* reporters (Lu et al., 2010). Using the *pFUS3:FUS3ΔC-GFP* reporter, the FUS3 protein was found to be localized to the pistil (septum, valves and funiculus) and ovules, in agreement with microarray data (Figure 1 A-F and Supplemental Figure S1A). In developing ovules FUS3ΔC-GFP was localized to the epidermis of the nucellus, the chalaza, and funiculus, while in mature ovules (FS12) it was localized to the chalaza and funiculus (Figure 1 C-F). After fertilization, (6-48 hours) FUS3ΔC-GFP was present in the funiculus, outer layer of the seed coat, chalaza and micropyle; it was also localized to the embryo at early stages of embryogenesis (Figure 1G-L and Supplemental Figure S1B).

To further address the role of *FUS3* in reproduction, we monitored ovule development in *fus3-3* loss-of-function mutant and *pML1:FUS3-GFP* misexpression lines (Gazzarrini et al., 2004). *pML1:FUS3-GFP* was shown to rescue all *fus3-3* seed maturation defects, including desiccation intolerance, however misexpression during postembryonic development caused additional

phenotypes (Gazzarrini et al., 2004). Strong *pML1:FUS3-GFP* lines show delayed vegetative growth and flowering, reduced plant height and aborted siliques, as previously described (Figure 2A; Gazzarrini et al., 2004; Lu et al., 2010). In *pML1:FUS3-GFP* lines FUS3-GFP was mislocalized to the endothelium, outer and inner integuments of developed ovules, while in aborted ovules FUS3-GFP surrounded the aborted embryo sac (Figure 2B); after fertilization, FUS3-GFP was mislocalized to the endosperm (Figure 2B). Moreover, siliques of intermediate-to-strong *pML1:FUS3-GFP* lines contained aborted seeds or seeds with delayed development (Figure 2C, D).

Next, we determined if seed abortion in *fus3-3* and *pML1:FUS3-GFP* is the result of impaired ovule development. We found that the embryo sac of wild type ovules at FS12 stage contained the egg nucleus, the secondary endosperm nucleus, the synergids, and was surrounded by inner and outer integuments. However, at FS12 stage the embryo sac of some *fus3-3* and *pML1:FUS3-GFP* lines was delayed at various stages, from FG1 to FG6, arrested or not fully wrapped by the integuments (Figure2E). The arrest of female megagametogenesis resulted in seed abortion in *fus3-3* and more so in strong *pML1:FUS3-GFP* lines (Figure 2C, D).

Taken together, these results show that spatiotemporal localization of FUS3 is tightly regulated and that lack or misexpression of *FUS3* severely impairs embryo sac and integument development, indicating that spatiotemporal control of *FUS3* expression is required for proper ovule development.

Class I BPC transcription factors bind to $(GA/CT)_n$ motifs in *FUS3*.

To understand the mechanisms controlling the spatiotemporal patterns of *FUS3* expression, we identified upstream regulators of *FUS3* by yeast one-hybrid. To increase screening specificity, a short genomic region of 615bp upstream of the *FUS3* translation start (*pFUS3*) was screened against an Arabidopsis transcription factor library (Figure 3A; Mitsuda et al., 2010). Sequencing of the cDNA inserts revealed that all colonies contained BPC3. BPCs are a small family of plant specific transcription factors consisting of six genes and a pseudogene (BPC5) and divided into 3 classes based on sequence similarity: class I (BPC1/2/3), class II (BPC4/5/6) and class III (BPC7) (Meister et al., 2004). We retested individually all class I BPCs (BPC1-3) and also included class II BPC4, which is not present in the cDNA library but it is highly expressed in embryos and flowers (Berger et al., 2011). All three class I BPCs bound to *pFUS3* by yeast one-hybrid, but not class II BPC4 (Figure 3A).

BPCs were shown to bind to $(GA/CT)_n$ cis elements in several plant species, with a preference for different numbers of repeats (Berger and Dubreucq, 2012; Simonini and Kater, 2014). When all $(GA/CT)_n$ motifs of the *pFUS3* were mutated (*pFUS3*^{MUT}), none of the class I BPCs interacted with the *FUS3* sequence, confirming binding specificity (Figure 3B; Supplemental Figure S2). To identify the binding location of BPCs on *pFUS3*, we generated truncations of approximately 200bp fragments (F1 to F3); the first exon/intron region containing 2 $(GA/CT)_n$ repeats (F4) was also tested (Figure 3C). In Y1H, BPC1-3 bound to the 5'UTR (F3) and first exon/intron regions (F4), where $(GA/CT)_n$ motifs are enriched (Figure 3D). BPC1-4 did not bind the promoter region further upstream, corresponding to the F1 or F2 truncations, where there is only one $(GA)_5$ or no $(GA/CT)_n$ motif, respectively (Figure 3D). To determine if BPC1 also binds to the *FUS3* locus *in vivo* during reproductive development, we generated BPC1 overexpression lines and performed ChIP in inflorescences, which show that BPC1 directly binds to this region (Figure 3F).

Altogether, this indicates that class I BPCs bind to the 5'UTR and first intron/exon regions of *FUS3* in Y1H. Furthermore, BPC1 directly binds to *FUS3 in vivo* during reproductive development.

Class I BPCs repress *FUS3* during vegetative growth

In a genome-wide study, BPC1 was found to interact with and recruit the conserved PRC2-complex subunit FIERY (FIE) *in vivo* and trigger polycomb-mediated gene silencing in imbibed seeds (Xiao et al., 2017). We first analyzed ChIP-seq data from Xiao et al. (2017) and found that the first exon/intron and 5'UTR of *FUS3* was bound by BPC1 in seedlings (Figure 3E). Furthermore, this same region was bound by FIE and associated with H3K27me3, a repressive mark (Figure 3E). Lastly, BPC1/2 interact with EMBRYONIC FLOWER2 (EMF2), which belongs to the EMF-PRC2 complex involved in repressing the vegetative-to-reproductive and embryo-to-seedling phase transitions (Xiao et al., 2017; Mozgova et al., 2015). This suggests that *FUS3* may be repressed in germinating seeds by BPC1 recruitment of EMF-PRC2. To confirm this, we mutated all BPC binding sites $(GA/CT)_n$ in the *FUS3* sequence (*pFUS3*^{MUT}) and showed that *pFUS3*^{MUT}*:GUS/GFP* is indeed derepressed post-embryonically in leaf and root tips (Figure 3G,H). Together with previous data showing that *FUS3* was strongly upregulated in *swinger curly leaf* (*swn clf*) (Makarevich et al., 2006), these results strongly suggest that BPC1 binds to and represses *FUS3* during vegetative development by recruiting the EMF-PRC2 complex.

Class I BPCs repress FUS3 during reproductive and seed development

Previous ChIP assays showed that in closed flowers the *FUS3* locus is also associated with the FIS-PRC2 complex component MEA and H3K27me3 repressive marks, and that *FUS3* is upregulated in the endosperm of *mea/MEA* seeds at 3 days after flowering (DAF) (Makarevich et al., 2006). Given that BPC1 bind to the *FUS3* locus in closed flowers (Figure 3F), we hypothesized that during reproductive development *FUS3* may be repressed by BPCs through FIS-PRC2 recruitment. To test this hypothesis, we first monitored *pFUS3^{MUT}:GUS* staining and found that *FUS3* is upregulated in flower buds (Figure 3G). Next, we determined if class I BPCs interact *in planta* with the FIS-PRC2 complex, which acts during gametophyte and endosperm development. BPC1-3 interacted with the unique subunits of this PRC2 complex, FIS2 and MEA, and also with the PRC2-shared component, MSI1, in BiFC assays; all but BPC3 also interacted with FIE (Supplemental Figure S3). In agreement with previous Y2H results, class I BPCs also interacted with each other *in planta*, and BPC2 and 3 could also form homodimers (Supplemental Figure S4; Simonini et al., 2012). No class I BPC member or FIS-PRC2 subunit interacted with FUS3, suggesting that these BiFC interactions are specific (Supplemental Figure S5). Last, given that BPC6 recruits PRC2 by interacting with LIKE HETEROCHROMATIN PROTEIN1 (LHP1; Hecker et al., 2015), we also tested the interaction between class I BPCs and LHP1 *in planta*. However, no interaction was found, suggesting class I and class II BPCs recruitment of the PRC2 complex may differ (Supplemental Figure S6). We conclude that class I BPCs can form homo- and hetero-dimers and recruit the FIS-PRC2 complex *in planta*.

Class I BPCs were shown to be expressed in ovules (Monfared et al., 2011). To better understand their role in reproductive and seed development, we determined their expression patterns before (FS4-12) and after (1-11DAF) fertilization. BPC1-3 had largely overlapping expression patterns before fertilization and they were all highly expressed in almost all tissues of developing ovules, embryos, as well as the endosperm and seed coat (Supplemental Figure S7). BPC1 had a more restricted pattern before (chalaza and micropile) and after (chalaza, micropyle, seed coat)

fertilization. This suggests that class I BPCs may act redundantly during ovule and embryo development.

As previously shown, the FIS-PRC2 complex subunits FIS2 and MEA were only expressed in the central cell of developing ovules and in the endosperms at 2DAF (Supplemental Figure S7; Wang et al., 2006). Collectively, these data show that BPCs can interact with each other and with FIS-PRC2 to regulate gene expression. Given the specific localization of FIS and MEA to the central cell and endosperm, and *FUS3* derepression in the endosperm of *mea*/MEA, we conclude that aside from their role in silencing *FUS3* during vegetative growth through EMF-PRC2, class I BPCs repress *FUS3* during reproductive and seed development by recruiting FIS-PRC2 in the central cell and endosperm. Furthermore, BPCs may recruit sporophytic PRC2 (EMF/VRN PRC2) to repress FUS3 in the integuments and seed coat.

Reproductive defects of *bpc1/2* are partially rescued by *fus3-3*

Previously, *bpc* mutants were shown to display pleiotropic phenotypes during vegetative and reproductive development (Monfared et al., 2011). Higher order *bpc1/2 and bpc1/2/3* mutants are dwarf, have shorter or aborted siliques, display severe seed abortion and defects in embryo sac development, while most single *bpc* mutants resemble wild type, suggesting functional redundancy (Figure 5A-F; Supplemental Figure S8A-D; (Monfared et al., 2011). These phenotypes are remarkably similar to those shown by *pML1:FUS3* misexpression lines (Figure 2; Gazzarrini et al., 2004), suggesting that *bpc1/2* phenotypes may be caused by ectopic expression of *FUS3*. To address the genetic relationship between class I *BPCs* and *FUS3*, we crossed *bpc1/2* with *fus3-3*. The *bpc1/2 fus3-3* indeed showed partial rescue of these phenotypes, including plant height (Figure 4A,D), silique and seed abortion (Figure 4B,C,E,F; Supplemental Figure S10), as well as embryo sac development (Figure 4H), supporting our hypothesis that *FUS3* is misexpressed in *bpc1/2*.

After fertilization, the endosperm of some *bpc1/2* mutants appeared very dense and some ovules were not fertilized (Figure 4H; Supplemental Figure S8E). In fertilized seeds, most *bpc1/2* also display delayed or arrested embryo development (Figure 4E,F,H; Supplemental Figure S8A,B,F,G). Overall, reproductive defects in higher order *bpc* mutants result in severe reduction of seed yield (Figure 4G). The *bpc1/2 fus3-3* triple mutant partially rescue endosperm and embryo development

(Figure 4E,F,H). These data strongly suggest that BPCs repress *FUS3* during reproductive and seed development.

BPC1/2 repress *FUS3* to promote inflorescence stem elongation, ovule and endosperm development.

To confirm a repressive role of BPCs on *FUS3* function, we analyzed *FUS3* expression level and patterns in *bpc1/2* mutants. We show that *FUS3* transcript level is indeed increased in *bpc1/2* inflorescence stem (Figure 5A). Consistent with the transcript analysis, *pFUS3:GUS* activity is also increased in *bpc1/2* inflorescence stem and flower buds (Figure 5B). In WT, low *FUS3* expression in the inflorescence stem is shown by transcriptomic data and detected with the *pFUS3:FUS3ΔC-GFP* sensitive reporter (Supplemental Figure S1). Together with previous findings showing that plant height is reduced in *pML1:FUS3-GFP* misexpression plants (Gazzarrini et al., 2004), while increased in the *fus3-3* mutant (Figure 4D), these results indicate that BPC1/2 downregulates *FUS3* in the stem to promote stem elongation.

During reproductive development FUS3ΔC-GFP is mislocalized to the integuments at the micropylar region of developing *bpc1-1 and bpc1/2* ovules, while after fertilization ectopic *pFUS3:GUS* activity and FUS3ΔC-GFP localization were detected in *bpc1* and *bpc1/2* endosperms (Figure 5B,C,D). Combined with the above functional analysis, these results show that before fertilization BPCs restrict *FUS3* expression to the funiculus and chalaza to promote ovule development, while after fertilization *FUS3* is repressed by BPCs in most of the endosperm to coordinate embryo and endosperm growth.

To analyze the repressive role of class I BPCs, we also crossed *pFUS3:FUS3:GFP* translational reporter with *bpc1/2* mutant. However, we were only able to isolate *bpc1-1 pFUS3:FUS3:GFP* lines. As shown in Supplemental Figure S8 and previous research (Monfared et al., 2011) *bpc1-1* doesn't have any visible phenotype compared with wild type, nor does *pFUS3:FUS3:GFP,* which rescues the *fus3-3* mutant phenotypes (Gazzarrini et al., 2004; Chan et al., 2017). However, in *bpc1-1 pFUS3:FUS3:GFP* some flower buds were arrested and never opened, petals and anthers filaments did not elongate and anthers were aborted, similar to *bpc1/2* double mutant (Figure 6A). Seed abortion was increased and delayed embryogenesis was evident in *bpc1-1 pFUS3:FUS3:GFP* plants (Figure 6C; Supplemental Figure S8B). The *bpc1-1 pFUS3:FUS3:GFP* plants were shorter and resembled the *bpc1/2* double mutant (Figure 6D and Figure 4A,D). Thus, our inability to isolate *bpc1/2 pFUS3:FUS3:GFP* line may be due to the severe phenotype of such a mutant. The presence

of the *pFUS3:FUS3:GFP* transgene enhanced the *bpc1-1* phenotype likely due to higher or ectopic *FUS3* expression. Accordingly, we could detect strong GFP fluorescence in the integuments, seed coat and funiculus of *bpc1-1 pFUS3:FUS3:GFP*, while *pFUS3:FUS3:GFP* showed no fluorescence in WT (in contrast to the stable *pFUS3:FUS3ΔC-GFP*). Furthermore, FUS3:GFP was mis-localized in *bpc1-1 pFUS3:FUS3:GFP* endosperm after fertilization (Figure 6E), in agreement with FUS3ΔC-GFP mislocalization and *pFUS3:GUS* misexpression in *bpc1-1* and *bpc1/2* endosperm (Figure 5). These results further support a repressive role of BPCs on *FUS3* expression in different tissues during reproductive and seed development.

Last, *bpc1/2*, *bpc1-1 pFUS3:FUS3:GFP* and *pML1:FUS3-GFP* ovules that were successfully fertilized displayed an increased number of endosperm nuclei, which correlated with an increase in seed size (Figure 7A,B,C,D; Supplemental Figure S9), and some embryos were delayed or arrested at various stages of development (globular to early torpedo) (Figure 7E; Supplemental Figure S9). *bpc1/2* also showed aberrant cell division patterns in the embryo proper and suspensor, which resulted in defective embryos and were partially rescued by *fus3-3* (Figure 7E; Supplemental Figure S9). Collectively, these data show that repression of *FUS3* in the endosperm of developing seeds is required to coordinate endosperm and embryo growth.

Discussion

PRC2 play important roles in balancing cell proliferation with differentiation and regulating developmental phase transitions in plants and animals. Recently, genome wide studies have shown that the plant-specific, class I BPC transcription factors bind Polycomb response elements (PREs), recruit EMF-PRC2 and trigger gene silencing during germination (Xiao et al., 2017). Similar to GAGA factors in *Drosophila melanogaster*, BPCs recognize $(GA/CT)_n$ *cis* elements, despite the lack of sequence similarity between these transcription factors, suggesting convergent evolution (Berger and Dubreucq, 2012). BPCs play essential roles during vegetative and reproductive development, as shown by the dwarf stature and severe seed abortion displayed by higher order *bpc* mutants, however the molecular mechanisms are largely unknown (Kooiker et al., 2005; Monfared et al., 2011; Simonini et al., 2012; Simonini and Kater, 2014;). Here we show that BPC1/2 interact with FIS-PRC2 and bind to the *FUS3* chromatin to restrict *FUS3* expression to specific tissues during reproductive and seed development. BPC-mediated spatiotemporal regulation of *FUS3* expression is required to i) suppress stem elongation during vegetative-to-reproductive phase change, ii) promote ovule development before fertilization and iii) coordinate embryo and

endosperm development after fertilization (Figure 7F). Several lines of evidence support these conclusions. First, Y1H show that class I BPCs bind to $(GA/CT)_n$ around the *FUS3* transcription start, and ChIP assays in flower buds show that BPC1 binds *in vivo* to the *FUS3* chromatin. Mutations in these $(GA/CT)_n$ sites abolish BPCs binding and derepress *FUS3*. Furthermore, *FUS3* is upregulated in the inflorescence stem of *bpc1/2* dwarf plants, which is consistent with *fus3-3* tall plant and *ML1:FUS3-GFP* dwarf plant phenotypes, as well as FUS3 role as repressor of vegetative-to-reproductive phase change (Gazzarrini et al., 2004; Lumba et al., 2012). Second, class I BPCs interact with FIS-PRC2 complex *in planta,* and the *in vivo* BPC1-binding region on *FUS3* was shown to associate with MEA and H3K27me3 repressive marks (Makarevich et al., 2006), strongly suggesting BPC1 recruits FIS-PRC2 to repress *FUS3* during reproductive/seed development. Third, FUS3 is transiently localized to the integuments during early ovule development and later restricted to the funiculus and chalaza of mature wild type ovules. Ectopic and persistent expression of *FUS3* in the integuments of *bpc1/2* and *ML1:FUS3* mis-expression lines impairs integument and embryo sac development leading to seed abortion, which can be partially rescued in *bpc1/2 fus3-3*. Last, after fertilization FUS3 is localized to the funiculus, chalaza and outer integument, aside from its known localization to the embryo (Gazzarrini et al., 2004). Ectopic expression of *FUS3* in *bpc1/2* and *ML1:FUS3* endosperm leads to increased proliferation of the endosperm nuclei and delayed or arrested embryo development, which are rescued in *bpc1/2 fus3-3*. The latter phenotypes are also displayed by mutants in *FIS-PRC2* subunits (Kiyosue et al., 1999; Kohler and Grossniklaus, 2002). We conclude that BPCs recruit PRC2 to restrict spatiotemporal *FUS3* expression during reproductive and seed development; this is required to regulate tissue development locally and modulate developmental phase transitions in Arabidopsis. The genomic sequences of *FUS3* orthologs in other species show conservation of $(GA/CT)_n$ repeats (Supplemental Figure S11), suggesting that similar mechanisms may regulate the expression of *FUS3*-like transcription factors in other species.

Inflorescence stem elongation and flower development require repression of *FUS3* by Class I BPCs

During germination BPC1 interacts with PRC2 and directly binds to the genomic region of *FUS3* proximal to the transcription start, which is marked by H3K27me3 repressive marks and associates with FIE (Figure 3; Xiao et al., 2017). Furthermore, *FUS3* is strongly expressed in *swn clf* seedlings, which show embryonic traits (Makarevich et al., 2006), suggesting that during germination *FUS3*

is repressed through BPC1-recruitment of EMF/VRN-PRC2. Here we show that mutations of all BPC binding sites on the *FUS3* promoter derepress *FUS3* in vegetative and reproductive organs, and that lack of BPCs results in ectopic *FUS3* expression in leaves, inflorescence stem and flower buds. Furthermore, ectopic *FUS3* in *bpc1/2*, *bpc1 pFUS3:FUS3-GFP* or *pML1:FUS3-GFP* leads to similar phenotypes, including reduced internode elongation and defective flowers (arrested flower bud development, flowers with a protruding carpel and shorter floral organs), suggesting *FUS3* inhibits the elongation of the stem and floral organs during flowering. Recently, deletion of a small region in the *FUS3* promoter near the BPC binding sites and corresponding to the PRC2 recruitment region, also caused ectopic *FUS3* expression in vegetative and reproductive tissues (Roscoe et al., 2019). Thus, we propose that class I BPCs recruit sporophytic VRN/EMF-PRC2 to repress *FUS3* post-embryonically, in vegetative and reproductive organs (Figure 7F).

BPC-mediated restriction of *FUS3* expression in developing ovules and seeds is required to promote ovule development and to coordinate endosperm and embryo growth

During ovule development, the funiculus supplies nutrients and signaling molecules from the mother plant to the chalaza, initiates the integuments that grow around the nucellus and protect the developing female gametophyte (Schneitz et al., 1995). Our data show that during megagametogenesis FUS3 is initially localized to the nucellus epidermis and tissues surrounding the nucellus, including the integuments and chalaza. However, BPC1/2 later repress *FUS3* in the integuments of mature ovules, and ectopic *FUS3* expression in *bpc1/2* or *pML1:FUS3* inhibits integuments and embryo sac development, triggering ovule abortion. This is in agreement with previous findings showing that the integuments are required for female gametogenesis (Elliott et al., 1996; Klucher et al., 1996; Baker et al., 1997) and strongly suggests that spatiotemporal restriction of *FUS3* is required for integuments, embryo sac and ovule development (Figure 7F).

Following fertilization, the zygote together with the endosperm and the integuments develop in a coordinated manner to form the embryo and the seed coat of the mature seed. FUS3 was previously shown to localize to developing embryos from globular to cotyledon stages (Gazzarrini et al., 2004). Using the sensitive/stable FUS3ΔC-GFP reporter, we found that FUS3 localizes also to the funiculus, chalaza and outer seed coat of developing seeds, partially mirroring its expression

pattern in ovules. Ectopic FUS3 localization to the endosperm of *bpc1/2* or *pML1:FUS3-GFP* increases cell proliferation, resulting in enlarged endosperm and larger seed at the expense of embryo development (delayed or arrested). These phenotypes are reminiscent of some FIS-PRC2 mutant alleles of *mea* (Kiyosue et al., 1999). Given that *FUS3* is derepressed in *mea* endosperm and that MEA and H3K27me3 repressive marks associate in a repressive region of the *FUS3* locus where BPC1 also binds, we propose that BPC1/2 recruit FIS-PRC2 to repress *FUS3* in the endosperm (Makarevich et al., 2006); this is required to reduce the rate of endosperm nuclei proliferation, promoting endosperm differentiation and embryo growth (Figure 7F).

The FIS-PRC2 specific subunits, MEA and FIS2, are targeted solely to the central cell in the ovule and to the endosperm in the seed, and thus are likely to participate in *FUS3* repression in these tissues (Luo et al., 2000; Wang et al., 2006). The MEA homolog SWN, which belongs to the VRN-PRC2 and FIS-PRC2 complexes, has a broader localization pattern, but plays a partially redundant function with MEA in repressing central cell/endosperm nuclei proliferation in the absence of fertilization (Wang et al., 2006). Thus, SWN may also be involved in repressing *FUS3* in the central cell/endosperm. In contrast, autonomous seed coat development in the ovule is repressed by the sporophytic complexes VRN-PRC2 and EMF-PRC2, which may be involved in repressing *FUS3* in the integuments (Kohler and Grossniklaus, 2002; Roszak and Kohler, 2011). In accordance, *FUS3* and other seed-specific genes were derepressed and showed reduced H3K27me3 repressive marks in siliques of a weak *curly leaf* (*clf)* allele, although the tissue specific expression was not investigated (Liu et al., 2016).

Although BPCs can recruit EMF- and FIS-PRC2 complexes for transcriptional silencing, BPCs were also shown to positively regulate a close *FUS3* family member, *LEC2* (Berger et al., 2011). This is in accordance with the role of GAGA binding proteins in animals, which have dual function of activators and repressors (Berger and Dubreucq, 2012). Interestingly, *FUS3* is expressed in the embryo and in specific sporophytic tissues of the ovule and seed (chalaza, funiculus, seed coat), where all class I BPCs are expressed. Thus, it will be important to determine the mechanisms of BPCs activation and repression of *FUS3* and other *LAFL* genes during reproductive and seed development.

An important question is how does FUS3 regulate tissue development and phase transitions. FUS3 was shown to regulate embryonic-to-vegetative and vegetative-to-reproductive phase changes by controlling ABA/GA ratio and ethylene signaling (Gazzarrini et al., 2004; Lu et al., 2010; Chiu et al.,

2012). A positive feedback regulatory loop has also been established between auxin and FUS3 in the embryo, whereby auxin induces *FUS3* expression and FUS3 promotes auxin synthesis; however, the role of this feedback regulation is currently unclear (Gazzarrini et al., 2004). Given that auxin is required for the synchronized growth of the fruit, the different tissues within the seed (integuments, endosperm and embryo), and that FUS3 localization patterns in ovules, seeds and embryos largely mirror those of auxin, we propose that FUS3 may regulate auxin level/localization and that auxin may in turn regulate FUS3 expression/activity (Gazzarrini et al., 2004; Figueiredo et al., 2015; Figueiredo et al., 2016; Larsson et al., 2017; Robert et al., 2018). Reduced auxin accumulation in the chalaza and funiculus of *fus3-3* or increased auxin levels in the integuments and endosperm of *pML1:FUS3* or *bpc1/2* would impair ovule and seed development resulting in seed abortion and delayed embryo development, respectively, as shown by delayed endosperm cellularization and embryo growth arrest triggered by auxin overproduction in the endosperm (Figueiredo and Kohler, 2018; Batista et al., 2019; Robert, 2019).

In conclusion, mutations affecting FIS-PRC2 or PRE binding TF BPCs cause severe seed abortion, however the molecular mechanisms are not fully understood (Monfared et al., 2011; Wang and Kohler, 2017; Figueiredo and Kohler, 2018). Here we show that BPC1/2-mediated spatiotemporal restriction of *FUS3*, a target of the PRC2 complex, is required for the transition from reproductive to seed development, as well as from early embryogenesis to seed maturation in Arabidopsis.

Material and methods

Plant material

T-DNA insertion lines *bpc1-1* (SALK_072966C), *bpc2* (SALK_090810), *bpc1-1/bpc2* (*bpc1/2*; CS68700), and *bpc1-1/bpc2/bpc3-1* (CS68699), and an EMS mutant *bpc3-1* (CS68805) were previously described (Monfared et al., 2011). T-DNA insertion lines *bpc1_salk* (SALK_101466C), *bpc2_salk* (SALK_110830C), *bpc3_sail* (SAILseq_553_B09.0) were obtained from ABRC. All primers used for genotyping are listed in the Supplemental Table 1. The *pFIE:FIE:GFP*, *pMSI1:MSI1:GFP*, *pMEA:MEA:YFP* and *pFIS2:GUS* reporter lines were previously described (de Lucas et al., 2016). The *pML1:FUS3-GFP* construct previously described (Gazzarrini et al., 2004) was transform into *fus3-3* loss-of-function mutant (Keith et al., 1994). *pFUS3:FUS3ΔC-GFP* construct previously described was transformed into *Col-0* (Lu et al., 2010). *pFUS3:FUS3-GFP* construct was previously

described (Gazzarrini et al., 2004). *pFUS3(1.5kb)*:*GUS/GFP* and *pFUS3^{MUT}(1.5kb)*:*GUS/GFP* were generated as described in Supplementary Methods. Eight to ten transgenic lines per constructs were analyzed for GUS staining or GFP fluorescence. Sterilized Arabidopsis seeds were germinated on half-strength Murashige and Skoog (MS) medium, transferred to soil and grown under 16/8h light/darkness 22°C/18°C. Frequencies of seed phenotypes displayed by various genotypes were calculated with half dissected siliques (n=10); experiments were repeated three times with similar results and one is shown. Total seed yield per plant was calculated with 5 plants per pot, experiments were repeated three times.

Yeast one-hybrid screening

Yeast one-hybrid (Y1H) library screening and one-on-one retests were performed as described by Deplancke et al. (2006) with some modifications. See Supplementary Methods for bait construction and Y1H screen).

Bimolecular fluorescence complementation (BiFC) assay

The CDS of *BPCs*, *FIE* (AT3G20740), *MSI1* (AT5G58230), *MEA* (AT1G02580), *FIS2* (AT2G35670) and *LHP1* (AT5G17690) were cloned into BiFC vectors pB7WGYN2 (YNE) or pB7WGYC2 (YCE) (Tsuda et al., 2017) by Gateway. These recombined vectors were transformed into *Agrobacterium tumefaciens* strain GV2260 and infiltrated into *Nicotiana benthamiana* leaves as described previously (Duong et al., 2017). At least three biological replicates were performed.

Differential interference contrast (DIC) microscopy

Pistils at FS12 or siliques were dissected and immersed in fixing solution (9:1, ethanol:acetic acid, v/v) for 2h before washing them twice with 90% ethanol. The siliques were then cleared with clearing solution (2.5g/ml chloral hydrate and 30% glycerol) overnight. Images were taken with a Zess Axioplant 2 microscope equipped with DIC optics. The quantification of seed size and endosperm nuclei were performed by Image J.

Confocal microcopy

To observe the expression of GFP signal in transgenic Arabidopsis, fresh tissues was dissected and mounted on the slides with 10% glycerol. Visualization was done with a Zeiss LSM510 confocal microscope (488 nm excitation and a 515-535 nm band pass filter).

GUS staining

The *pBPC3:GUS* line was previously described (Monfared et al., 2011). The promoter regions of *BPC1/2* described in Monfared et al. (2011) were PCR-amplified and transformed into the pGWB3 vector to generate *pBPC1:GUS* and *pBPC2:GUS*. Several transformed homozygous lines were selected on kanamycin and hygromycin plates and analyzed and two lines were selected for further analysis. The GUS staining assays were performed as previously described (Wu et al., 2019) with some modifications. The concentration of ferri/ferrocyanide used for *pBPC3:GUS* was 2mM, while 5mM was used for *pBPC1:GUS* and *pBPC2:GUS*. To detect low expression of *FUS3* in inflorescences, leaves or flowers of *pFUS3(1.5kb):GUS* and *pFUS3^{MUT}(1.5kb):GUS* lines, ferri/ferrocyanide was not included in the buffer. Cleared tissues were imaged by DIC microscopy using Zeiss Axioplant 2.

Glutaraldehyde staining

To visualize ovule/seed structures, whole pistils/siliques at FS12 or 1-2DAF were fixed in 3% paraformaldehyde in PBS for 15min at room temperature and rinsed twice with PBS. The treated tissues were stained in 5% glutaraldehyde in PBS at 4°C overnight in the dark. Tissues were washed 3 times with PBS and cleared for about 1 to 2 weeks with ClearSee buffer (Kurihara et al., 2015). The images were photographed with a Zeiss LSM510 confocal microscope (530nm excitation and a 560nm long pass filter).

Gene expression assay

RNA was extracted using the RNeasy Plant Mini Kit (Qiagen). About 1µg of RNA was used for reverse transcription. Quantitive real-time PCR was performed using Step One Plus real-time PCR system (Applied Biosystems) with SYBR premix. *PP2AA3* was chosen as the internal reference gene. Primers used are listed in Supplemental Table 1. Three biological replicates were performed.

ChIP assay

To generate *35S:BPC1-RFP*, the *BPC1* coding sequence was first cloned into pDONR221 (Life Technologies) and subsequently transferred to pB7RWG2 (Flanders Interuniversity Institute for Biotechnology, Gent, Belgium). Arabidopsis plants were transformed with the *35S:BPC1-RFP* using the Agrobacterium tumefaciens-mediated floral dip method (Clough and Bent, 1998). Transformed plants were sown on soil and selected by BASTA; the presence of the construct was assessed by genotyping and analysis of RFP expression. Arabidopsis plants were directly sown on soil and kept under short-day conditions for 2 weeks (22°C, 8h light and 16h dark) and then moved to long-day conditions (22°C, 16h light and 8h dark). ChIP assays were performed as described by Gregis et al. (2009) using for BPC1-RFP an anti-RFP V_HH coupled to magnetic agarose beads RFP-trap_MA® (Chromotek). 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 in Supplemental Table 1. ChIP-qPCR experiments and relative enrichments were calculated as reported by Gregis et al. (2009).

Acknowledgements

We thank C.S. Gasser for *pBPC3:GUS* reporter; F. Parcy for *pFUS3:GUS*; SM Brady and M. De Lucas for *pFIE:FIE:GFP*, *pMSI1:MSI1:GFP*, *pMEA:MEA:YFP* and *pFIS2:GUS* reporter lines as well as *FIE* and *MSI1* vectors; C. Koehler and R. Yadegari for *MEA/pBluescript II KS* and *FIS2/pGBKT7* vectors. JW was supported by the National Natural Science Foundation projects (31701952) and China Postdoctoral Council scholarships. V.G. was supported by the Ministero dell'Istruzione, Università e Ricerca MIUR, SIR2014 MADSMEC (RBSI14BTZR). R.P. was supported by Doctorate School in Molecular and Cellular Biology, Università degli Studi di Milano Fellowship. This work was funded by a Natural Sciences and Engineering Research Council of Canada Discovery Grant to SG.

Authors contribution

JW and SG conceived the study and wrote the paper. JW conducted most of the experiments. SD contributed to the identification of higher order mutants. RP and VG conducted ChIP assays. All read and approved the manuscript.

REFERENCES

Baker SC, Robinson-Beers K, Villanueva JM, Gaiser JC, and Gasser CS. 1997. Interactions among genes regulating ovule development in Arabidopsis thaliana. *Genetics* 145: 1109-1124

Batista RA, Figueiredo DD, Santos-Gonzalez J, and Kohler C. 2019. Auxin regulates endosperm cellularization in Arabidopsis. *Genes Dev* 33: 466-476

Berger N, and Dubreucq B. 2012. Evolution goes GAGA: GAGA binding proteins across kingdoms. *Biochim Biophys Acta* 1819: 863-868

Berger N, Dubreucq B, Roudier F, Dubos C, and Lepiniec L. 2011. Transcriptional regulation of Arabidopsis LEAFY COTYLEDON2 involves RLE, a cis-element that regulates trimethylation of histone H3 at lysine-27. *Plant Cell* 23: 4065-4078

Braybrook SA, Stone SL, Park S, Bui AQ, Le BH, Fischer RL, Goldberg RB, and Harada JJ. 2006. Genes directly regulated by LEAFY COTYLEDON2 provide insight into the control of embryo maturation and somatic embryogenesis. *Proc Natl Acad Sci USA* 103: 3468-3473

Carbonero P, Iglesias-Fernandez R, and Vicente-Carbajosa J. 2017. The AFL subfamily of B3 transcription factors: evolution and function in angiosperm seeds. *J Exp Bot* 68: 871-880

Chan A, Carianopol C, Tsai AYL, Varatharajah K, Chiu RS, and Gazzarrini S. 2017. SnRK1 phosphorylation of FUSCA3 positively regulates embryogenesis, seed yield, and plant growth at high temperature in Arabidopsis *J Exp Bot* 68: 5981-5981

Chiu RS, Pan S, Zhao R, Gazzarrini S. 2016. ABA-dependent inhibition of the ubiquitin proteasome system during germination at high temperature in Arabidopsis. *Plant J* 88: 749–761

Curaba J, Moritz T, Blervaque R, Parcy F, Raz V, Herzog M, Vachon G. 2004. AtGA3ox2, a key gene responsible for bioactive gibberellin biosynthesis, is regulated during embryogenesis by LEAFY COTYLEDON2 and FUSCA3 in Arabidopsis. *Plant Physiol*. 136:3660-9.

Clough SJ, and Bent AF. 1998. Floral dip: a simplified method for Agrobacterium-mediated

transformation of Arabidopsis thaliana. *Plant J* 16: 735-743

de Lucas M, Pu L, Turco G, Gaudinier A, Morao AK, Harashima H, Kim D, Ron M, Sugimoto K, Roudier F, et al. 2016. Transcriptional Regulation of Arabidopsis Polycomb Repressive Complex 2 Coordinates Cell-Type Proliferation and Differentiation. *Plant Cell* 28: 2616-2631

de Vries SC, and Weijers D. 2017. Plant embryogenesis. *Curr Biol* 27: R870-R873

Dean G, Cao Y, Xiang D, Provart NJ, Ramsay L, Ahad A, White R, Selvaraj G, Datla R, and Haughn G. 2011. Analysis of gene expression patterns during seed coat development in Arabidopsis. *Mol Plant* 4: 1074-1091

Deplancke B, Vermeirssen V, Arda HE, Martinez NJ, and Walhout AJ. 2006. Gateway-compatible yeast one-hybrid screens. *CSH Protoc* 2006: doi: 10.1101/pdb.prot4590

Dresselhaus T, Sprunck S, and Wessel GM. 2016. Fertilization Mechanisms in Flowering Plants. *Curr Biol* 26: R125-139

Duong S, Vonapartis E, Li CY, Patel S, and Gazzarrini S. 2017. The E3 ligase ABI3-INTERACTING PROTEIN2 negatively regulates FUSCA3 and plays a role in cotyledon development in Arabidopsis thaliana. *J Exp Bot* 68: 1555-1567

Elliott RC, Betzner AS, Huttner E, Oakes MP, Tucker WQ, Gerentes D, Perez P, and Smyth DR. 1996. AINTEGUMENTA, an APETALA2-like gene of Arabidopsis with pleiotropic roles in ovule development and floral organ growth. *Plant Cell* 8: 155-168

Fatihi A, Boulard C, Bouyer D, Baud S, Dubreucq B, and Lepiniec L. 2016. Deciphering and modifying LAFL transcriptional regulatory network in seed for improving yield and quality of storage compounds. *Plant Sci* 250: 198-204

Figueiredo DD, and Kohler C. 2018. Auxin: a molecular trigger of seed development. *Genes Dev* 32: 479-490

Figueiredo DD, Batista RA, Roszak PJ, and Kohler C. 2015. Auxin production couples endosperm development to fertilization. *Nat Plants* 1: 15184

Figueiredo DD, Batista RA, Roszak PJ, Hennig L, and Kohler C. 2016. Auxin production in the endosperm drives seed coat development in Arabidopsis. *eLife* 5

Gasser CS, and Skinner DJ. 2019. Development and evolution of the unique ovules of flowering plants. *Curr Top Dev Biol* 131: 373-399

Gazzarrini S, Tsuchiya Y, Lumba S, Okamoto M, and McCourt P. 2004. The transcription factor FUSCA3 controls developmental timing in Arabidopsis through the hormones gibberellin and abscisic acid. *Devel Cell* 7: 373-385

Gietz RD, and Schiestl RH. 2007. Large-scale high-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. *Nat Protoc* **2:** 38-41

Gregis V, Sessa A, Dorca-Fornell C, and Kater MM. 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

Grossniklaus U, Vielle-Calzada JP, Hoeppner MA, and Gagliano WB. 1998. Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. *Science* **280:** 446-450

Hecker A, Brand LH, Peter S, Simoncello N, Kilian J, Harter K, Gaudin V, and Wanke D. 2015. The Arabidopsis GAGA-Binding Factor BASIC PENTACYSTEINE6 Recruits the POLYCOMB-REPRESSIVE COMPLEX1 Component LIKE HETEROCHROMATIN PROTEIN1 to GAGA DNA Motifs. *Plant Physiol* **168:** 130-141

Heckman KL, and Pease LR. 2007. Gene splicing and mutagenesis by PCR-driven overlap extension. *Nat Protoc* **2:** 924-932

Jia H, Suzuki M, and McCarty DR. 2014. Regulation of the seed to seedling developmental phase transition by the LAFL and VAL transcription factor networks. *Wiley Interdiscip Rev Dev Biol* **3:** 135-145

Keith K, Kraml M, Dengler NG, and McCourt P. 1994. fusca3: A Heterochronic Mutation Affecting Late Embryo Development in Arabidopsis. *Plant Cell* **6:** 589-600

Kiyosue T, Ohad N, Yadegari R, Hannon M, Dinneny J, Wells D, Katz A, Margossian L, Harada JJ, Goldberg RB, et al. 1999. Control of fertilization-independent endosperm development by the MEDEA polycomb gene in Arabidopsis. *Proc Natl Acad Sci USA* **96:** 4186-4191

Klucher KM, Chow H, Reiser L, and Fischer RL. 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

Kohler C, and Grossniklaus U. 2002. Epigenetic inheritance of expression states in plant development: the role of Polycomb group proteins. *Curr Opin Cell Biol* **14:** 773-779

Kooiker M, Airoldi CA, Losa A, Manzotti PS, Finzi L, Kater MM, and Colombo L. 2005. BASIC PENTACYSTEINE1, a GA binding protein that induces conformational changes in the regulatory region of the homeotic arabidopsis gene SEEDSTICK. *Plant Cell* **17:** 722-729

Kurihara D, Mizuta Y, Sato Y, and Higashiyama T. 2015. ClearSee: a rapid optical clearing reagent for whole-plant fluorescence imaging. *Devel* **142:** 4168-4179

Lafon-Placette C, and Kohler C. 2014. Embryo and endosperm, partners in seed development. *Curr*

Opin Plant Biol 17: 64-69

Larsson E, Vivian-Smith A, Offringa R, and Sundberg E. 2017. Auxin Homeostasis in Arabidopsis Ovules Is Anther-Dependent at Maturation and Changes Dynamically upon Fertilization. *Front Plant Sci* doi: 10.3389/fpls.2017.01735

Lau S, Slane D, Herud O, Kong J, and Jurgens G. 2012. Early embryogenesis in flowering plants: setting up the basic body pattern. *Annu Rev Plant Biol* 63: 483-506

Lepiniec L, Devic M, Roscoe TJ, Bouyer D, Zhou DX, Boulard C, Baud S, and Dubreucq B. 2018. Molecular and epigenetic regulations and functions of the LAFL transcriptional regulators that control seed development. *Plant Reprod* 31: 291-307

Liu J, Deng S, Wang H, Ye J, Wu HW, Sun HX, and Chua NH. 2016. CURLY LEAF Regulates Gene Sets Coordinating Seed Size and Lipid Biosynthesis. *Plant Physiol* 171: 424-436

Lotan T, Ohto M, Yee KM, West MAL, Lo R, Kwong RW, Yamagishi K, Fischer RL, Goldberg RB, and Harada JJ. 1998. Arabidopsis LEAFY COTYLEDON1 is sufficient to induce embryo development in vegetative cells. *Cell* 93: 1195-1205

Lu QS, dela Paz J, Pathmanathan A, Chiu RS, Tsai AYL, and Gazzarrini S. 2010. The C-terminal domain of FUSCA3 negatively regulates mRNA and protein levels, and mediates sensitivity to the hormones abscisic acid and gibberellic acid in Arabidopsis. *Plant J* 64: 100-113

Lumba S, Tsuchiya Y, Delmas F, Hezky J, Provart NJ, Shi Lu Q, McCourt P, and Gazzarrini S. 2012. The embryonic leaf identity gene FUSCA3 regulates vegetative phase transitions by negatively modulating ethylene-regulated gene expression in Arabidopsis. *BMC Biol* doi: 10.1186/1741-7007-10-8

Luo M, Bilodeau P, Dennis ES, Peacock WJ, and Chaudhury A. 2000. Expression and parent-of-origin effects for FIS2, MEA, and FIE in the endosperm and embryo of developing Arabidopsis seeds. *Proc Natl Acad Sci USA* 97: 10637-10642

Makarevich G, Leroy O, Akinci U, Schubert D, Clarenz O, Goodrich J, Grossniklaus U, and Kohler C. 2006. Different Polycomb group complexes regulate common target genes in Arabidopsis. *EMBO Rep* 7: 947-952

Meister RJ, Williams LA, Monfared MM, Gallagher TL, Kraft EA, Nelson CG, and Gasser CS. 2004. Definition and interactions of a positive regulatory element of the Arabidopsis INNER NO OUTER promoter. *Plant J* 37: 426-438

Mitsuda N, Ikeda M, Takada S, Takiguchi Y, Kondou Y, Yoshizumi T, Fujita M, Shinozaki K, Matsui M, and Ohme-Takagi M. 2010. Efficient Yeast One-/Two-Hybrid Screening Using a Library

Composed Only of Transcription Factors in Arabidopsis thaliana. *Plant Cell Physiol* 51: 2145-2151

Monfared MM, Simon MK, Meister RJ, Roig-Villanova I, Kooiker M, Colombo L, Fletcher JC, and Gasser CS. 2011. Overlapping and antagonistic activities of BASIC PENTACYSTEINE genes affect a range of developmental processes in Arabidopsis. *Plant J* 66: 1020-1031

Mozgova I, Köhler C, Hennig L. 2015. Keeping the gate closed: functions of the polycomb repressive complex PRC2 in development. *Plant J*. 83: 121-32.

Mozgova I, Hennig L. 2015. The polycomb group protein regulatory network. *Annu Rev Plant Biol*. 66: 269-96.

Robert HS. 2019. Molecular Communication for Coordinated Seed and Fruit Development: What Can We Learn from Auxin and Sugars? *Int J Mol Sci* doi: 10.3390/ijms20040936

Robert HS, Park C, Gutierrez CL, Wojcikowska B, Pencik A, Novak O, Chen JY, Grunewald W, Dresselhaus T, Friml J, et al. 2018. Maternal auxin supply contributes to early embryo patterning in Arabidopsis. *Nat Plants* 4: 548-553

Roscoe TJ, Vaissayre V, Paszkiewicz G, Clavijo F, Kelemen Z, Michaud C, Lepiniec L, Dubreucq B, Zhou DX, and Devic M. 2019. Regulation of FUSCA3 expression during seed development in Arabidopsis. *Plant Cell Physiol* 60: 476-487

Roszak P, and Kohler C. 2011. Polycomb group proteins are required to couple seed coat initiation to fertilization. *Proc Natl Acad Sci USA* 108: 20826-20831

Schmid M, Davison TS, Henz SR, Pape UJ, Demar M, Vingron M, Scholkopf B, Weigel D, and Lohmann JU. 2005. A gene expression map of Arabidopsis thaliana development. *Nat Genet* 37: 501-506

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

Simonini S, and Kater MM. 2014. Class I BASIC PENTACYSTEINE factors regulate HOMEOBOX genes involved in meristem size maintenance. *J Exp Bot* 65: 1455-1465

Simonini S, Roig-Villanova I, Gregis V, Colombo B, Colombo L, and Kater MM. 2012. Basic pentacysteine proteins mediate MADS domain complex binding to the DNA for tissue-specific expression of target genes in Arabidopsis. *Plant Cell* 24: 4163-4172

Sreenivasulu N, and Wobus U. 2013. Seed-Development Programs: A Systems Biology-Based Comparison Between Dicots and Monocots. *Annu Rev Plant Biol* 64: 189-217

Stone SL, Kwong LW, Yee KM, Pelletier J, Lepiniec L, Fischer RL, Goldberg RB, and Harada JJ. 2001.

LEAFY COTYLEDON2 encodes a B3 domain transcription factor that induces embryo development. *Proc Natl Acad Sci USA* 98: 11806-11811

Suh MC, Samuels AL, Jetter R, Kunst L, Pollard M, Ohlrogge J, and Beisson F. 2005. Cuticular lipid composition, surface structure, and gene expression in Arabidopsis stem epidermis. *Plant Physiol* 139: 1649-1665

Swanson R, Clark T, and Preuss DJSPR. 2005. Expression profiling of Arabidopsis stigma tissue identifies stigma-specific genes. *Sex Plant Reprod* 18: 163-171

Tsai AY, Gazzarrini S. 2012. AKIN10 and FUSCA3 interact to control lateral organ development and phase transitions in Arabidopsis. *Plant J* 69: 809–21.

Tsuchiya Y, Nambara E, Naito S, and McCourt P. 2004. The FUS3 transcription factor functions through the epidermal regulator TTG1 during embryogenesis in Arabidopsis. *Plant J* 37: 73-81

Tsuda K, Abraham-Juarez MJ, Maeno A, Dong Z, Aromdee D, Meeley R, Shiroishi T, Nonomura KI, and Hake S. 2017. KNOTTED1 Cofactors, BLH12 and BLH14, Regulate Internode Patterning and Vein Anastomosis in Maize. *Plant Cell* 29: 1105-1118

Wang D, Tyson MD, Jackson SS, and Yadegari R. 2006. Partially redundant functions of two SET-domain polycomb-group proteins in controlling initiation of seed development in Arabidopsis. *Proc Natl Acad Sci USA* 103: 13244-13249

Wang G, and Kohler C. 2017. Epigenetic processes in flowering plant reproduction. *J Exp Bot* 68: 797-807

Wu J, Wu W, Liang J, Jin Y, Gazzarrini S, He J, and Yi M. 2019. GhTCP19 Transcription Factor Regulates Corm Dormancy Release by Repressing GhNCED Expression in Gladiolus. *Plant Cell Physiol* 60: 52-62

Wu J, Jin Y, Liu C, Vonapartis E, Liang J, Wu W, Gazzarrini S, He J, and Yi M. 2018. GhNAC83 inhibits corm dormancy release by regulating ABA signaling and cytokinin biosynthesis in Gladiolus hybridus. *J Exp Bot* 70: 1221-1237

Xiao J, Jin R, Yu X, Shen M, Wagner JD, Pai A, Song C, Zhuang M, Klasfeld S, He C, et al. 2017. Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in Arabidopsis. *Nat Genet* 49: 1546

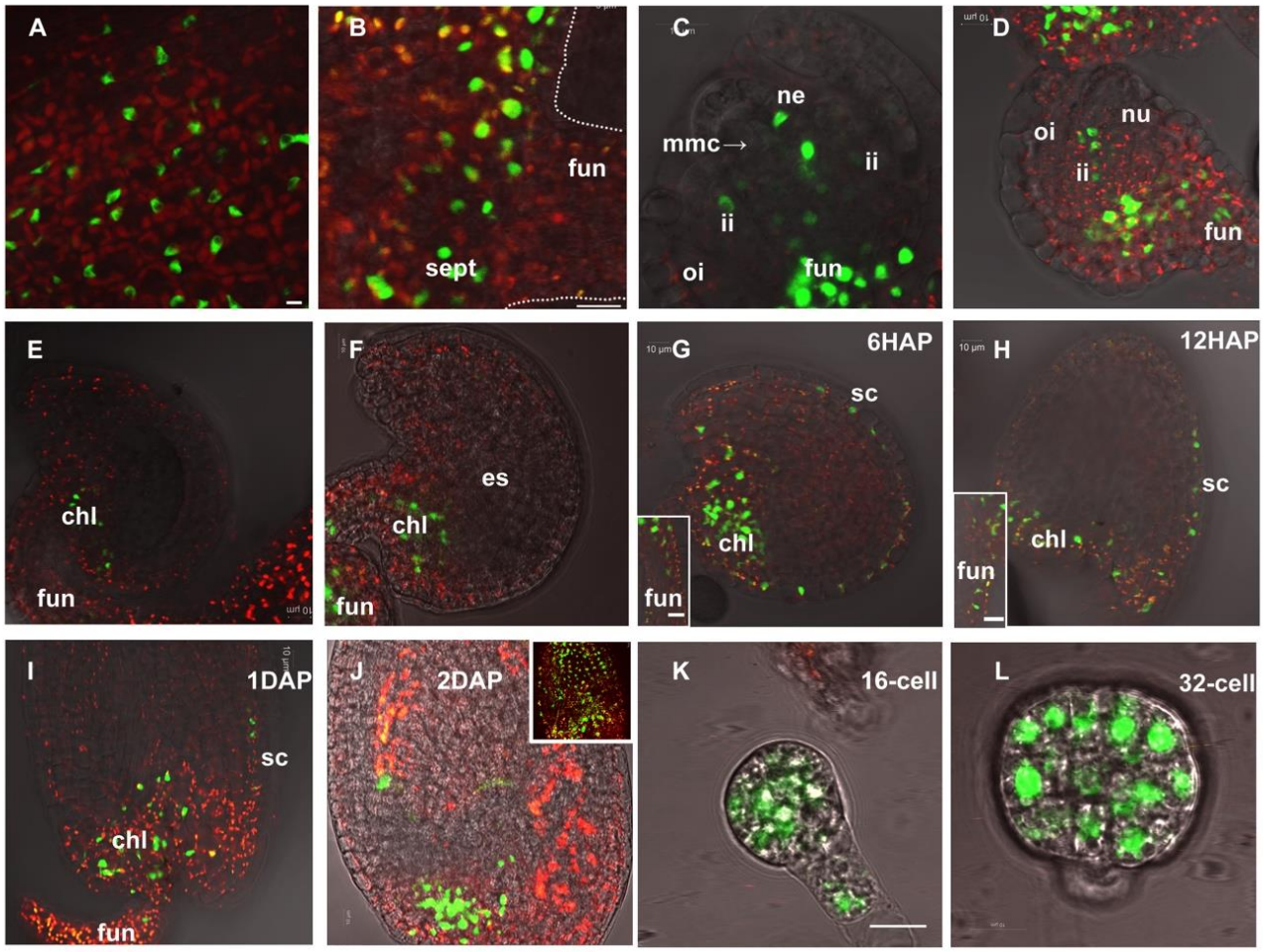


Figure 1. FUS3 localization in developing ovules and early stages of seed development.

Confocal images showing pFUS3:FUS3ΔC-GFP localization in Arabidopsis. (A) Valve and (B) septum of the pistil. (C-F) Ovules during female megasporogenesis (C) and megagametogenesis at FG1-FG7 (D-F); nucellar epidermis (C), inner and outer integuments (C,D), funiculus and chalazal (C,F). (G-J) Seeds at 6 hours (6HAP) to 2 days (2DAP) after pollination. Fluorescence was localized to the seed coat, chalaza and funiculus (G-J). (K) Suspensor and 16-cell stage embryo proper. (L) 32-cell stage embryo proper. chl: chalaza; es, embryo sac; fun, funiculus; ii: inner integument; megaspore mother cell; ne, nucellar epidermis; nu: nucellus; oi: outer integument; sept, septum. Red, autofluorescence. Purple dashed lines represent the outline of embryo sac. Scale bars, 10µm.

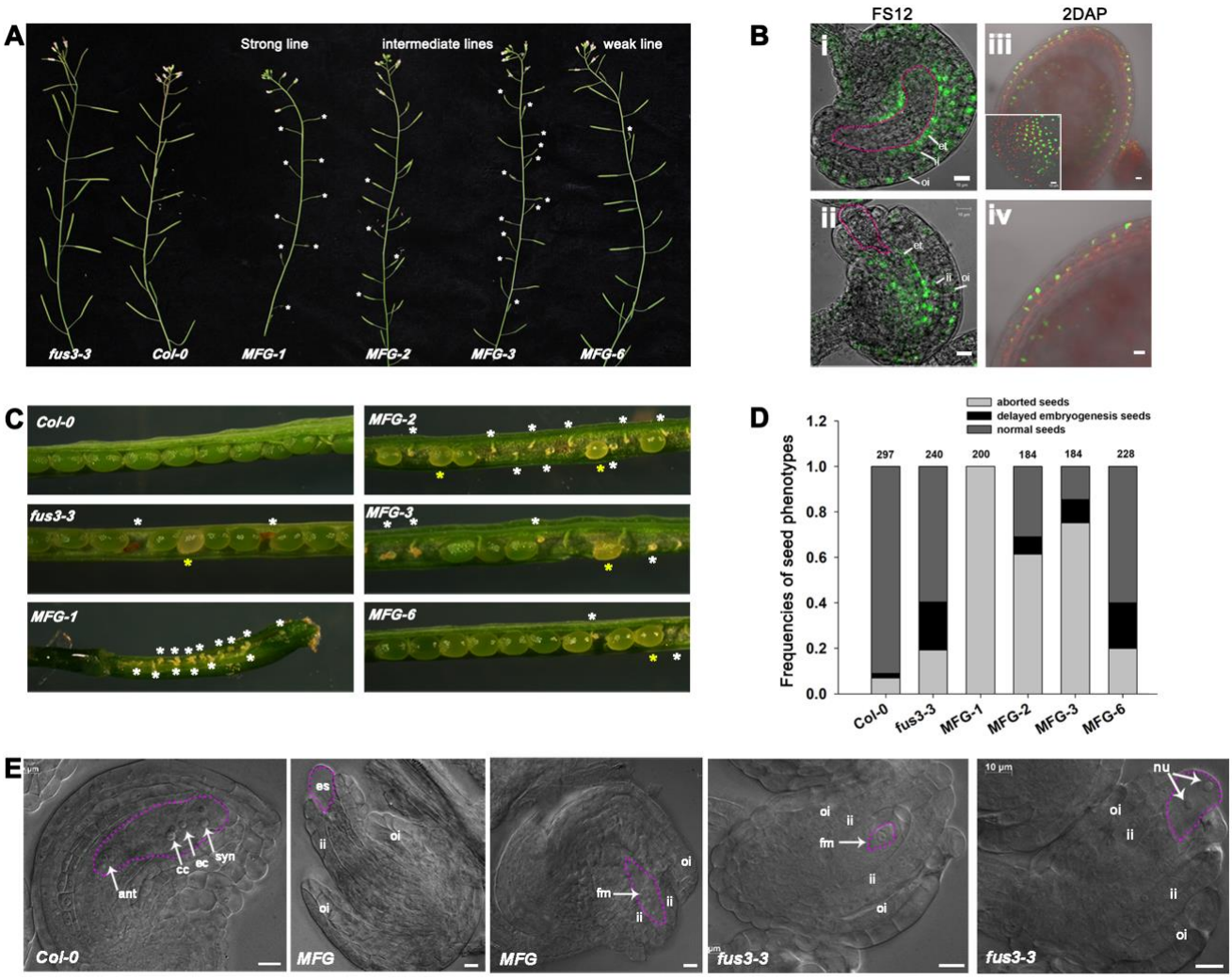


Figure 2. *FUS3* is required for ovule development.

A, Aborted silique (asterisks) in *fus3-3* and *fus3-3 pML1:FUS3-GFP* (*MFG*). B, pML1:FUS3-GFP localization to the integuments and endothelium of FS12 ovules and outer layer of the seed coat and endosperm (inset) of 2DAP seeds. (i) developed ovule; (ii) aborted embryo sac; (iii, iv). outer layer of the seed coat and the endosperm (inset) in 2DAP seeds; bar, 10µM. C, Aborted seeds (white asterisk) and delayed embryogenesis (yellow asterisk) in *MFG* and *fus3-3* siliques. D, Distribution of seed phenotypes in peeled, half side siliques of WT, *MFG* and *fus3-3* (n= ten siliques/genotype). E, DIC images of WT, *MFG* and *fus3-3* FS12 ovules. Pink dashed lines outline the embryo sac. Ant: anti: antipodals; ec: egg cell; es: embryo sac; et: endothelium; fm: functional megaspore; ii, inner integument; nu: nuclei; oi, outer integument; syn: synergid cell nuclei. Bars, 10µm.

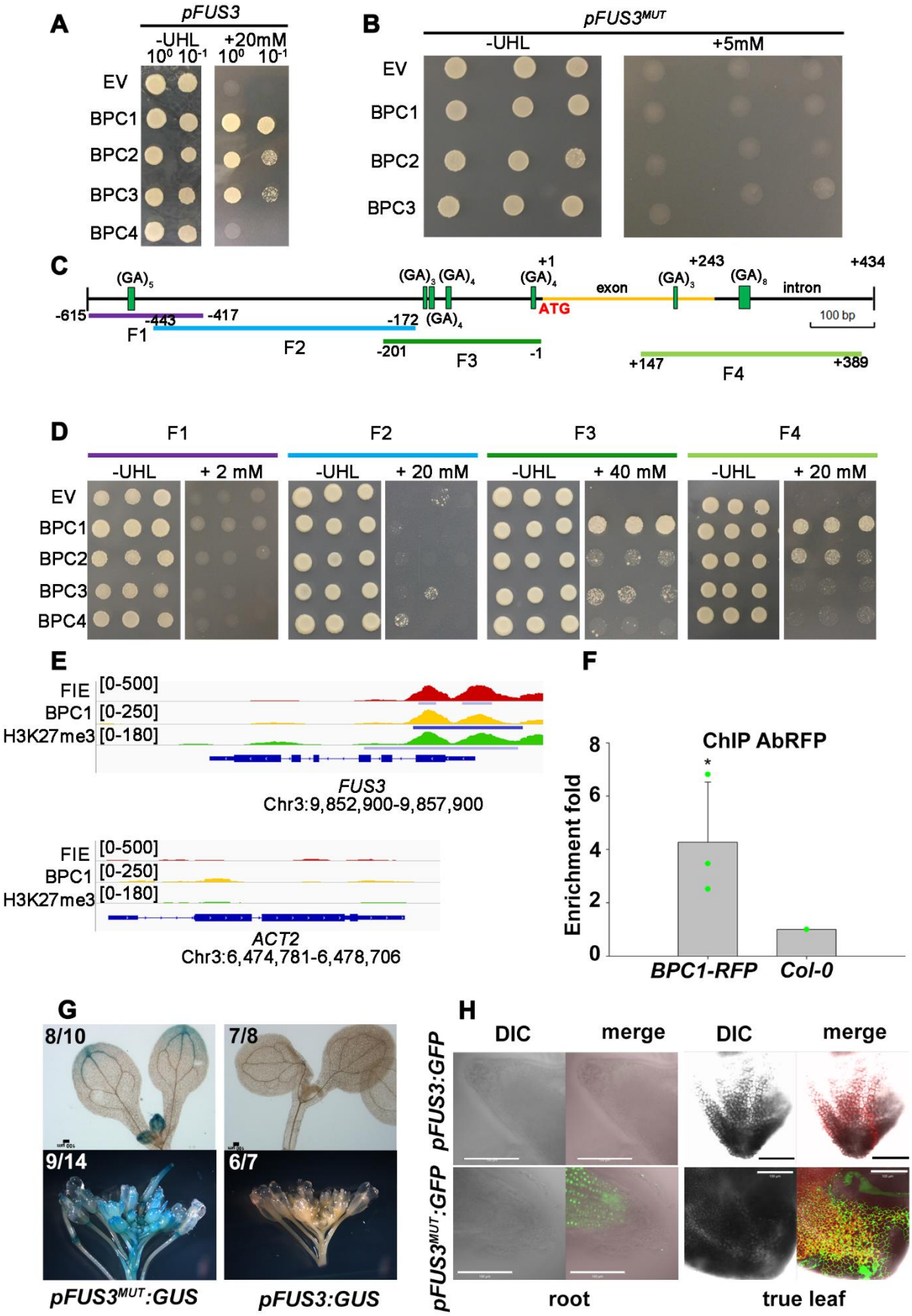


Figure 3. Class I BPCs bind to the *FUS3* genomic region proximal to the transcription start site.

A, BPC1/2/3 bind to a *FUS3* genomic region [*pFUS3(0.6 kb)*; -615 to +1 base pairs] in Y1H. **B**, BPC1/2/3 do not bind the *FUS3* genomic sequence carrying mutations in all $(GA/CT)_n$ [$pFUS3^{MUT}$*(0.6 kb)*]. Colonies in **A** and **B** were selected on -ura-his-leu medium (-UHL) with or without 5 or 20mM 3-AT. **C**, Distribution of $(GA/CT)_n$ motifs in *FUS3* genomic sequence (-615 to +434). **D**. Binding of BPC1/2/3 to truncated *FUS3* genomic sequences shown in C (F1 to F4). **E**, Bowser view of chromatin occupancy of FIE, BPC1 and H3K27me3 at *FUS3* and *ACT2* (negative control) in 30-hours-old seedlings using ChIP-seq data from Xiao et al. (2017). Significant peaks ($Q<10^{-10}$) according to MACS2 are marked by horizontal bars. **F**. Real-time qPCR analysis of ChIP assay using chromatin from *35S:BPC1-RFP* and *Col-0* (negative control) inflorescences and primers for the F3 region of *pFUS3*. Antibodies against the RFP tag were used in the IP. Error bars represent the propagated error value using three biological replicates (*$p<0.05$; student t-test). **G**, *pFUS3(1.5kb):GUS* and $pFUS3^{MUT}$*(1.5kb):GUS* staining in 10-days-old seedlings and in flower buds; numbers refer to the number of transgenic lines displaying the same GUS stain pattern as shown in **G**. **H**, *pFUS3(1.5kb):GFP* and $pFUS3^{MUT}$*(1.5kb):GFP* fluorescence in the leaf and root tip of 15-days-old seedlings.

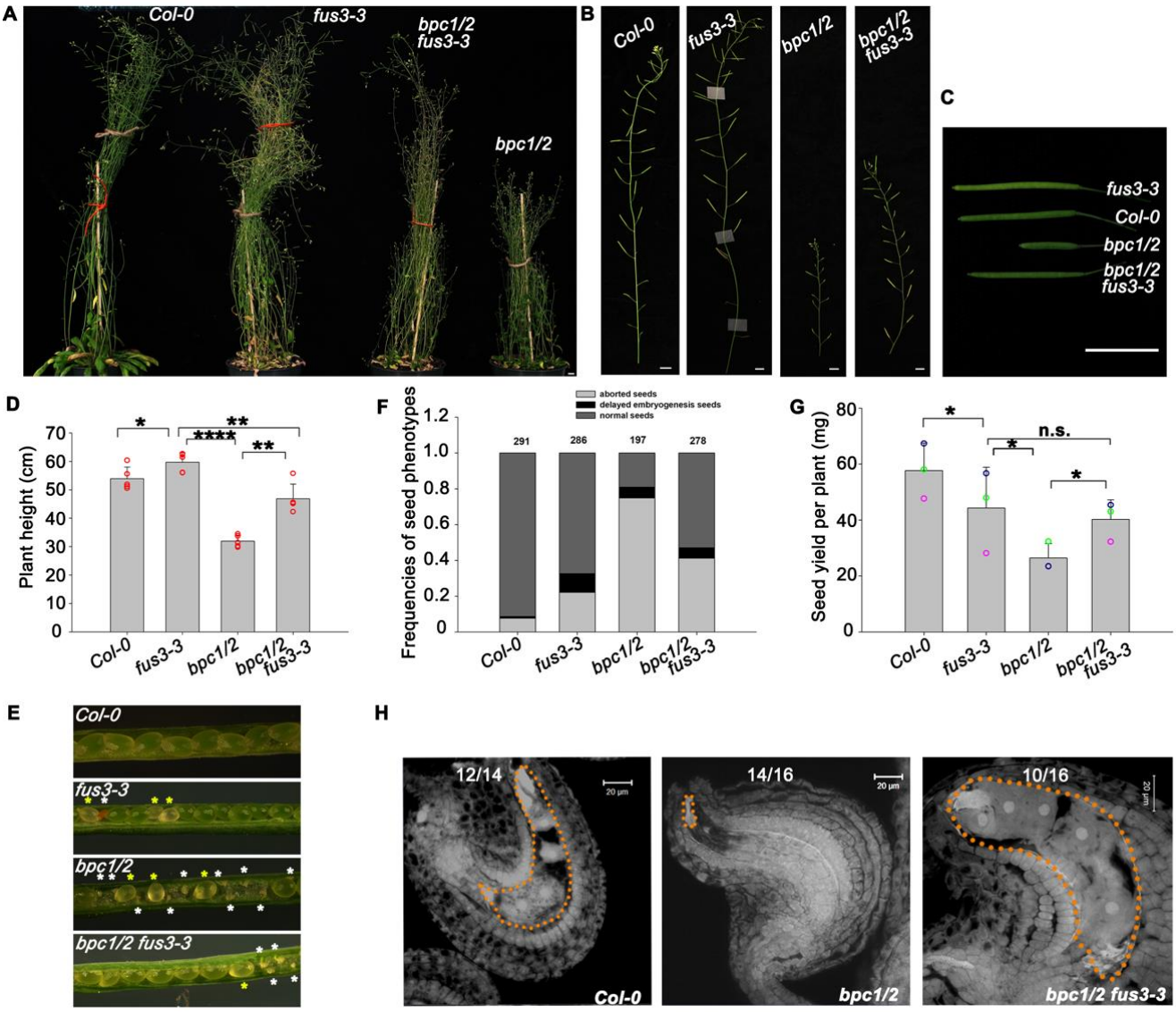


Figure 4. *P*artial rescue of *bpc1/2* stunted growth, aborted embryos and arrested seeds in *bpc1/2 fus3-3*.

A-C, Stunted growth and silique elongation of *bpc1/2* was partially rescued in *bpc1/2 fus3-3*. Scale bar in **C**, 1cm. **D**, Quantification of plant height. Five biological replicates were performed (n= five plants per genotype). **E, F**, *fus3-3* partially rescues *bpc1/2* severe seed abortion. White asterisks, aborted seeds; yellow asterisks, delayed embryogenesis seeds. **F,** Frequencies of seed phenotypes in *bpc1/2 fus3-3* mutants. Total number of seeds was calculated in 10 peeled, half side siliques. Three biological repeats were performed with similar results and one is shown (see also Supplemental Figure S10A). **G,** Seed yield; error bars represent the SD of three biological replicates (n=5). n.s.: no significant difference. (*p<0.05; **p<0.01; ****p<0.0001; student t-test). **H**, *fus3-3* partially rescues *bpc1/2* embryo sac defects. Images were taken at 1DAP; scale bars, 20µm. Numbers refer to the number of embryos displaying the phenotype shown.

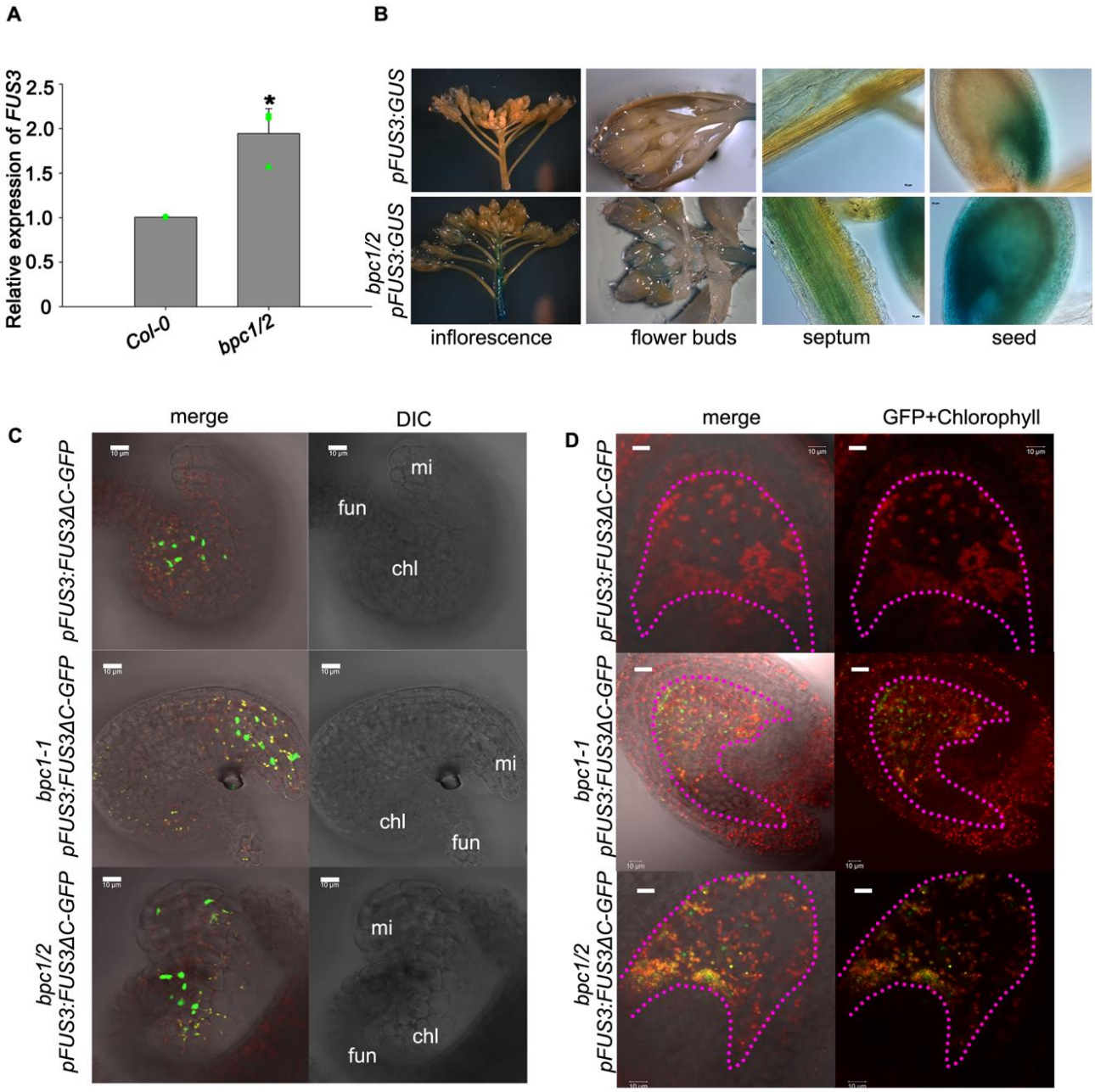


Figure 5. BPC1/2 negatively regulate *FUS3* expression in reproductive organs and seeds.

A, qRT-PCR showing increased *FUS3* transcript level in *bpc1/2* inflorescence stem. Error bars represent the SD of three biological replicates (*p<0.05; student t-test). B, GUS staining in the inflorescence stem, flower buds, septum and seed (2DAF) of *pFUS3:GUS* and *bpc1/2 pFUS3:GUS*. GUS staining was enhanced in the inflorescence stem and septum, while ectopically expressed in the endosperm of *bpc1/2*. C, D *pFUS3:FUS3ΔC-GFP* and *bpc1 pFUS3:FUS3ΔC-GFP* ovules were imaged before (C) and two days after (D) fertilization by confocal microscopy. FUS3ΔC-GFP was localized to the chalaza of developing WT ovules before fertilization, while ectopically localized to

the integuments at the micropilar region of *bpc1-1 and* of *bpc1/2* ovules (FS12) and endosperm of 2DAF *bpc1-1* seeds.

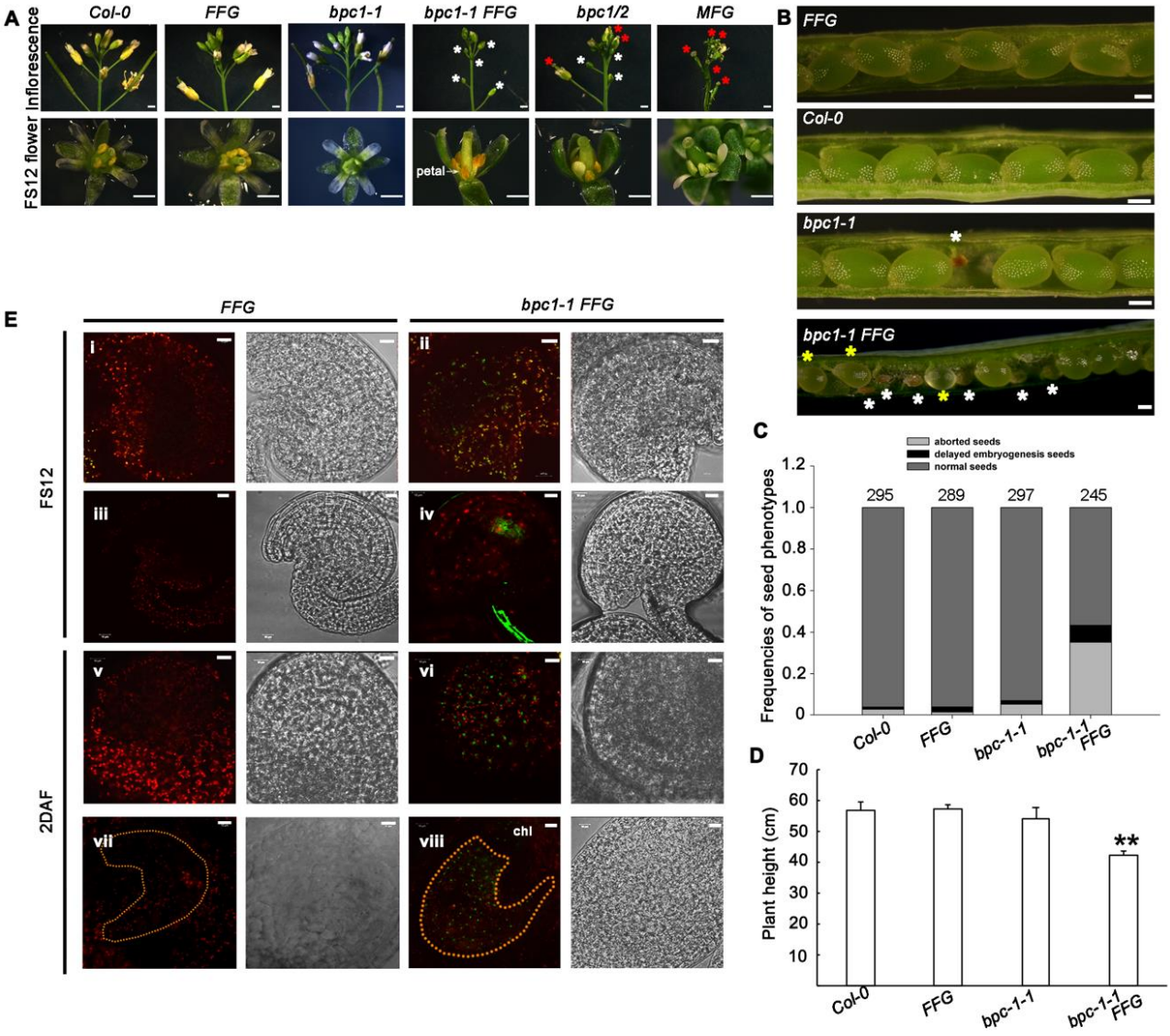


Figure 6. Ectopic *FUS3* expression negatively impacts reproductive organ development.

A, Introduction of a *pFUS3:FUS3-GFP (FFG)* transgene in *bpc1-1* results in arrested flower buds that never open (white asterisk), similar to *bpc1/2*. Arrested flower buds in *bpc1-1 FFG* have underdeveloped petals, non-elongated filaments and aborted anthers, similar to *bpc1/2*. *pML1:FUS3-GFP* (*MFE*) also show shorter filaments and underdeveloped anthers, but flower buds open prematurely. B, *bpc1-1 FFG* develops aborted seeds (white asterisk) and delayed embryogenesis (yellow asterisk). C, Frequencies of seed phenotypes. Total number of seeds was calculated in ten siliques (half side). Three biological repeats were performed and one representative is shown (see also Supplemental Figure S10B). D, *bpc1-1 FFG* plants display stunted

growth. The error bars represent SD of three biological replicates (n=5). (**: p<0.01; student t-test). **E**, FUS3-GFP is mis-localized to the integument (**ii**) and increased in the funiculus (**iv**) of *bpc1-1* ovules (FS12). Two days after fertilization (2DAF), FUS3-GFP is increased in *bpc1-1* seed coat (**vi**) and mis-expressed in the endosperm (**viii**).

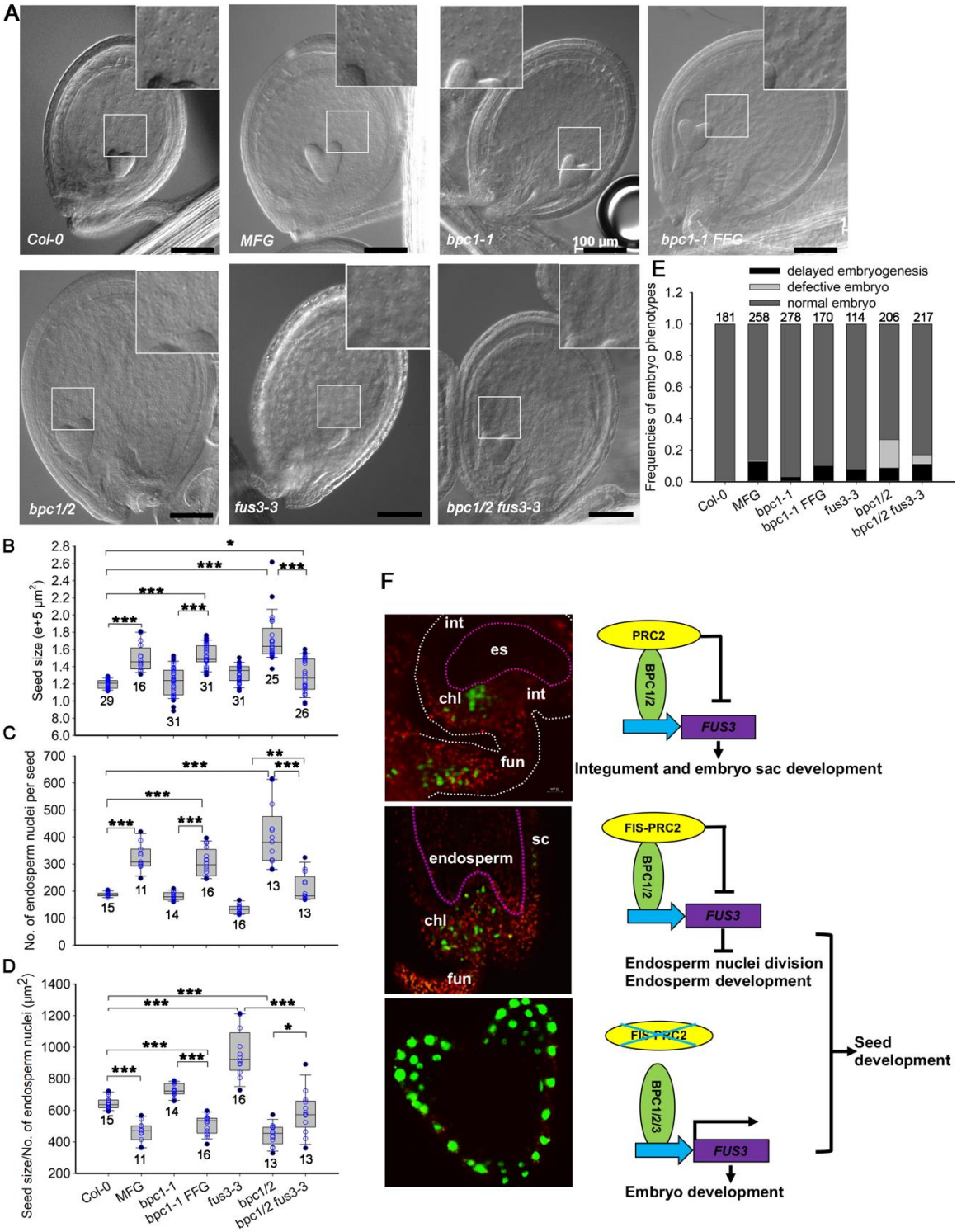


Figure 7. Negative regulation of endosperm nuclei proliferation by BPC1/2-mediated *FUS3* repression and model of spatiotemporal regulation of *FUS3* in reproductive and seed development.

A, Whole-mount clearing of six days after fertilization seeds. Scale bars, 100µm. **B-E** Ectopic expression of *FUS3* in *pML1:FUS3-GFP (MFG)*, *bpc1-1 pFUS3:FUS3-GFP (FFG)*, and *bpc1/2* leads to enlarged seed size **B)**, increased endosperm nuclei proliferation **C)** density **D)**, and delayed embryogenesis **E)**, which are partially rescued in *bpc1/2 fus3-3* (*p<0.05; **p<0.01; ****p<0.0001, in **B,C,D**; student t-test). **F**, Model depicting spatiotemporal regulation of *FUS3* by BPCs and PRC2 and its role in regulating reproductive and seed development. Before fertilization (top images), FUS3 becomes restricted to the funiculus and chalaza of mature ovules through BPC1/2-mediated *FUS3* repression in the integuments; this is required to promote integument and embryo sac development. After fertilization (middle and bottom images), FUS3 is localized to the seed coat, chalaza, funiculus and embryo, but is repressed in the endosperm by BPC1/2; this is required to decrease endosperm nuclei proliferation and promote embryo development. In integuments, BPC1/2-mediated *FUS3* repression may be orchestrated by sporophytic EMF/VRN-PRC2 (Liu et al., 2016; Xiao et al., 2017). After fertilization, FIS-PRC2 represses *FUS3* in the endosperm (Makarevich et al., 2006).

Supplementary informations can be found online at https://www.biorxiv.org/content/10.1101/612408v1.