Higher order complex formation in the establishment of floral organ identity

Chairgroup of Molecular Biology:
MSc. Thesis - Molecular Life Sciences

Stuart Y. Jansma

Sep 2014 – April 2015

Abstract

MADS proteins were found to function as tetramers or even higher order complexes during flower development. However, it is unclear whether all floral MADS target genes are regulated by a higher order complex, or if a MADS dimer could also be sufficient to regulate target gene transcription. In order to address this question, we need to obtain a robust *sep1sep2sep3* triple mutant, to provide a background for complementation by a *SEP3ΔC-GR* construct. In addition to MADS-box, non-MADS proteins (e.g. ARF2 and SPL8) were identified to be in complex with plant MADS proteins involved in floral organ identity. According to the currently proposed mechanism, MADS domain proteins form a quaternary complex and bind two CArG boxes in close proximity, resulting in a DNA loop. Then, transcriptional cofactors and chromatin remodelling proteins are recruited, and gene expression is altered. Interaction of non-MADS-box transcription factors that are thought to form a higher order complex with MADS-box proteins was investigated using yeast-hybrid assays.

Unfortunately, it was not possible to generate the *sep1sep2sep3* triple mutant using T-DNA insertion lines. An insertion in the *sep2* allele is not available, and this study revealed that the two available *sep3* alleles do not contain a T-DNA insert in *SEP3*. Therefore, we explored alternative methods to generate stable mutants. Based on the results reported in literature, the CRISPR/Cas9 system meets all the requirements of an efficient tool for targeted mutagenesis. However, our results indicate that the efficiency in cloning the CRISPR construct into the pYB196 vector is very low, and requires alternative cloning methods (e.g. Golden Gate cloning).

Our results of the yeast-hybrid assays were inconclusive: during the first assay several combinations of MADS-box proteins with ARF2 and SPL8 showed weak growth on selective media, but these results could not be reproduced during a second assay. Without positive controls to verify the functionality of the BD-fused constructs of ARF2 and SPL8, we can only speculate why interaction wasn't observed. Additional experiments are necessary to clarify whether ARF2 and SPL8 can interact with MADS proteins or not.

Table of Contents

1. Introduction

1.1 Floral organ identity - the ABC(DE) model

A typical dicotyledonous flower comprises four concentric whorls of floral organs. The first whorl consists of green sepals that look similar to leaves. The second whorl contains showy petals, whose colours and scent have evolved to lure pollinators. The floral organs of the inner whorls function in reproduction. Stamens in the third whorl produce pollen, bearing the male gametes, and carpels in the fourth whorl comprise ovules that contain the female gametes. When a male gamete reaches a female gamete, fertilization takes place and a seed starts to develop from the ovule (Meyer, 1966). Genetic studies led to the identification of floral homeotic mutants (Figure 1), which showed a change in floral organ identity establishment, resulting in perturbed flower morphology. (Komaki et al., 1988).
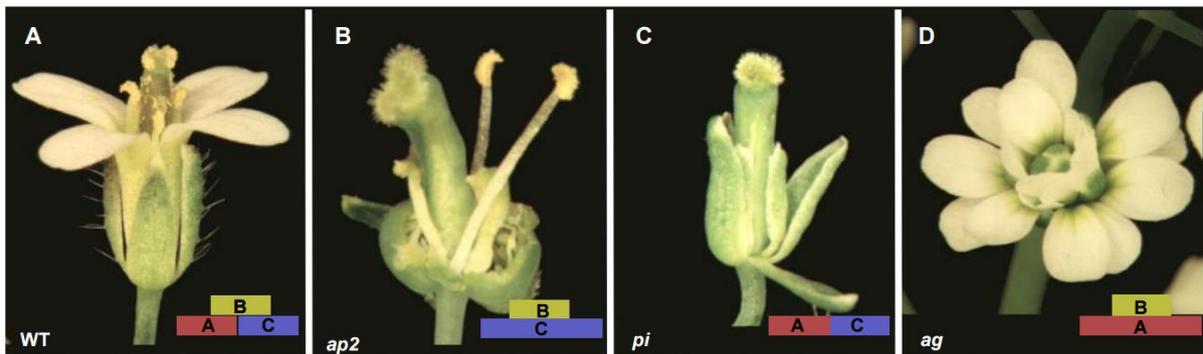

Figure 1: A) Wild type. B) Class A mutant with carpeloid organs in the first and stamenoid organs in the second whorl. C) Class B mutant with sepals in the second and carpels in the third whorl. D) Class C mutant that has petals in the third, and sepals in the fourth whorl. Adapted from Alvarez-Buylla et al. (2010).

Based on the changes induced by floral homeotic mutations, the affected genes were clustered in three functional classes (Bowman et al., 1989), and the **'ABC model of flower development**' was constructed (Figure 2) (Bowman et al., 1991; Coen and Meyerowitz, 1991; Meyerowitz et al., 1991). According to this model, combinatorial action of genes from these three classes specify the identities of the four floral organs.
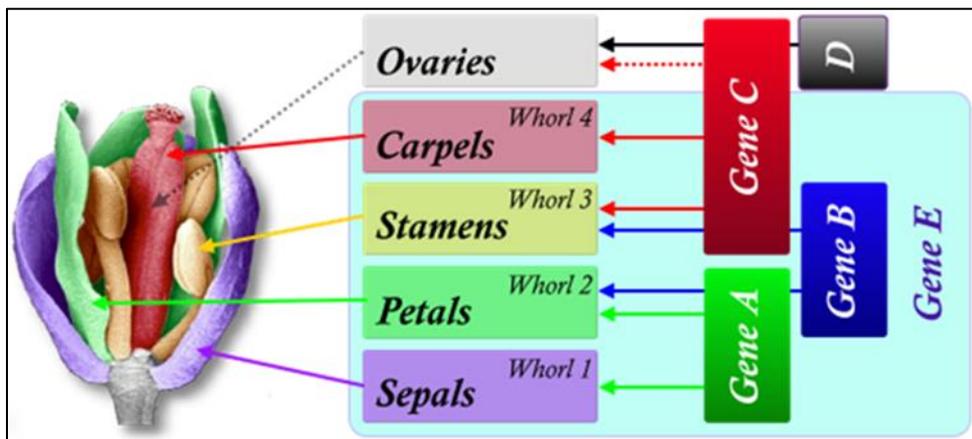

Figure 2: The ABC model of flower development hypothesizes that combinatorial action of genes expressed in specific whorls (classified as A, B, and C genes) results in establishment of floral organ identity. Retrieved from *http://www.adonline.id.au/flowers/floral-identity/*

Ectopic expression of these genes together did not generate flowers (Mizukami and Ma, 1992; Krizek and Meyerowitz, 1996), and it was therefore concluded that the three classes of genes of the ABC model were not sufficient for formation of floral organs. Additional research resulted in the identification of another class of floral homeotic genes that act together with the ABC genes. (Pelaz et al., 2000; Honma and Goto, 2001). These are the *SEPALLATA* genes, characterized by a *sep1sep2sep3* triple knockout mutant phenotype with only sepals (Figure 3b) (Pelaz et al., 2000). Additional knockout of *SEP4* converts those sepals into leaves, connecting *SEP4* function to floral organ identity as well (Figure 3c,d) (Ditta et al., 2004). 60% of *SEP3* overexpression mutants result in a severe dwarf phenotype with curled leaves, early flowering, and terminal flowers (Figure 3f) (Honma and Goto, 2001). Simultaneous overexpression of *PI*, *AP3*, and *SEP3* has been shown to convert true leaves into petaloid organs (Figure 3g) (Honma and Goto, 2001). Likewise, constitutive expression of *AP1, AP3,* and *PI*, together with *SEP3* (and *SEP2)* converts rosette leaves into petals (Figure 3h,i) (Pelaz et al., 2001).
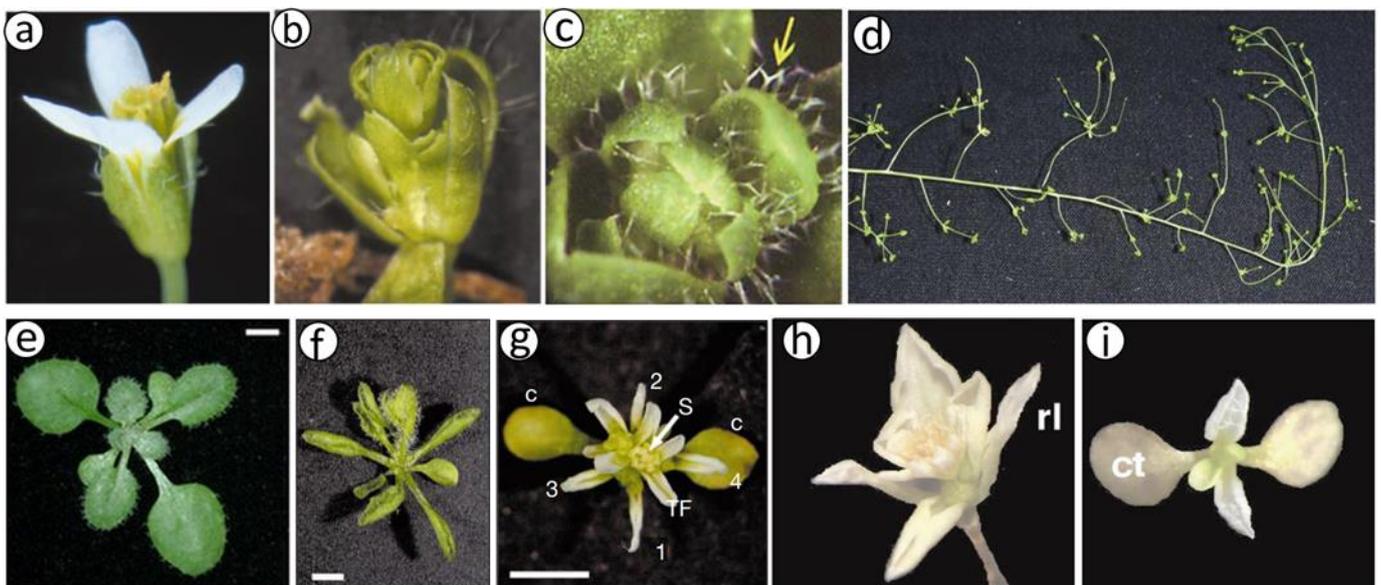

Figure 3: a) wildtype flower (adapted from (Pelaz et al., 2000)). b) *Sep1sep2sep3* triple knockout mutant phenotype with only sepals (adapted from (Pelaz et al., 2000)). c) *Sep1sep2sep3sep4* quadruple knockout mutant: true leaves with trichomes (arrow) (adapted from (Bowman et al., 1989; Ditta et al., 2004)). d) Whole plant of *sep1sep2sep3sep4* mutant showing loss of floral meristem identity (adapted from (Kunst et al., 1989; Ditta et al., 2004)). e) wildtype plant, 14 days after germination (adapted from (Yanofsky et al., 1990; Casanova-Sáez et al., 2014)). f) *35S::SEP3* showing severe phenotype (adapted from (Irish and Sussex, 1990; Honma and Goto, 2001)). g) *35S::AP3; 35S::PI; 35S::SEP3* showing conversion of true leaves into petaloid organs (adapted from (Alejandra Mandel et al., 1992; Honma and Goto, 2001)). h) *35S::AP1; 35S::AP3; 35S::PI; 35S::SEP3* and i) *35S::AP1; 35S::AP3; 35S::P; 35S::SEP2I; 35S::SEP3* show conversion of rosette leaves into petaloid organs (rl: rosette leaf; ct: cotyledon; adapted from (Pelaz et al., 2001))

These findings demonstrate that the *SEP* genes, together with the ABC genes, are sufficient to induce flower identity. The *SEPALLATA* genes represent class E floral homeotic genes in the revised ABC model of flower development. Another class of floral homeotic genes, the D class, was found to specify ovule identity (Colombo et al., 1995; Angenent and Colombo, 1996). The ABC model was adapted to include class D and E genes, as visualized in Figure 2 (Theißen, 2001). Table 1 displays the genes that belong to each class of the ABCDE model. Except for *AP2*, each of these genes encode transcription factors that are members of the MADS-box family, which is further discussed in the next section (Section 1.2).

Table 1: Genes in *Arabidopsis thaliana* of the ABCDE model of flower development that have been identified so far.

Class	Genes
A	*APETALA1* (*AP1*)(Irish and Sussex, 1990); *APETALA2* (*AP2*)(Komaki et al., 1988)
B	*APETALA3* (*AP3*); *PISTILLATA* (*PI*)(Bowman et al., 1989)
C	*AGAMOUS* (*AG*)(Bowman et al., 1989; Yanofsky et al., 1990)
D	*SHATTERPROOF1* (*SHP1*); *SHATTERPROOF2* (*SHP2*); *SEEDSTICK* (*STK*) (Colombo et al., 1995; Angenent and Colombo, 1996)
E	*SEPALLATA1* (*SEP1*); *SEPALLATA2* (*SEP2*); *SEPALLATA3* (*SEP3*)(Pelaz et al., 2000; Honma and Goto, 2001)

1.2 MADS-box proteins

MADS-box proteins are transcription factors, characterized by a conserved 180 bp DNA sequence motif, encoding a DNA-binding domain. Aside from plants, genes with striking similarity in the N-terminal region were also found in other eukaryotic species, such as yeast (*MINI CHROMOSOME MAINTENANCE 1*, *MCM1* (Passmore et al., 1988)) and humans (*SERUM RESPONSE FACTOR*, *SRF* (Norman et al., 1988)), indicating that this class of genes originates from a common ancestor (Ma et al., 1991; Winter et al., 2002). The term MADS is derived from the earliest identified members of the family: **M**CM1, **A**G, **D**EFICIENS, and **S**RF (Schwarz-Sommer et al., 1990).

Phylogenetic reconstructions allowed subdivision of the MADS-box family into several gene clades. A major clade of plant MADS-box genes has a conserved modular domain architecture, the **MIKC**-type domain structure (Ma et al., 1991). In angiosperms, a large proportion of the MADS-box genes encode transcription factors involved in flower development, and all MADS-box proteins involved in flower development have the MIKC structure (Pellegrini et al., 1995; Purugganan et al., 1995; Theißen et al., 1996). The *Arabidopsis thaliana* genome contains 107 members of MADS-box family, 46 of which are of this MIKC type (Riechmann and Ratcliffe, 2000; The_Arabidopsis_Genome_Initiative, 2000; De Bodt et al., 2003; Parenicova et al., 2003).

The following domains are characteristic for the MIKC-type structure (~260 aa):

M) (~56 aa)   The highly conserved **M**ADS domain is responsible for nuclear localization and DNA-binding to a consensus motif (Hayes et al., 1988; Schwarz-Sommer et al., 1990; Nurrish and Treisman, 1995; Pellegrini et al., 1995).

I) (27~42 aa)   The **I**ntervening domain is involved in selective formation of DNA-binding dimers and is relatively weakly conserved in sequence and length (Krizek and Meyerowitz, 1996; Riechmann et al., 1996b).

K) (~80 aa)   The **K**eratin-like domain, a domain present only in plant MADS-box proteins, encodes conserved regularly interspaced hydrophobic amino acid residues. The secondary structure comprises two amphipathic α-helices that separately mediate selective hetero-dimerization and tetramerization (Ma et al., 1991; Davies et al., 1996; Fan et al., 1997; Yang et al., 2003; Yang and Jack, 2004; Puranik et al., 2014).

C)   The most variable region, both in sequence and in length, is the **C**-terminal region, which is involved in transcriptional activation or repression, or in the formation of multimeric transcription factor complexes (Pellegrini et al., 1995; Riechmann et al., 1996b; Cho et al., 1999; Egea‑Cortines et al., 1999; Kaufmann et al., 2005).

1.3 Floral quartet model

A new insight in the combinatorial action of MADS-box genes was provided through publication of the crystal structure of a MADS-box dimer bound to DNA (Figure 4a) (Pellegrini et al., 1995). MADS-box proteins form homo- or heterodimers, which recognize and bind a 10 bp DNA sequence motif called CArG box, with the conserved consensus sequence $CC(A/T)_6GG$ (Riechmann et al., 1996a) (Muino et al., 2014). *In vitro* binding-site selection experiments, with chimeric MADS-box proteins, revealed that binding specificity is a sole function of the MADS-box N-terminal basic region and its flanking sequences (Nurrish and Treisman, 1995). Slight variation in the consensus sequence depends on which MADS-box proteins are present in the dimer, thus fitting the suggestion of the ABC model that different combination of MADS-box proteins activate different groups of target genes in each whorl of floral organ identity (Folter and Angenent, 2006).

Heterodimerization, however, is not sufficient to explain the different floral organ identities (Riechmann et al., 1996b), pointing out the need for a new model. This came into existence when Theißen and Saedler (2001) postulated **the floral quartet model** (Figure 4b), by combining the results of the discovery of a multimeric MADS complex in *Antirrhinum* consisting of *DEF, GLOBOSA* and *SQUAMOSA* (Egea- Cortines et al., 1999), with the discovery of *SEPALLATA* genes as the missing E class genes in *Arabidopsis* (Honma and Goto, 2001; Pelaz et al., 2001). This model assumes a combination of floral transcription factors that act as a tetramer to bind two CArG boxes in close proximity, thereby bending the DNA in a loop (West and Sharrocks, 1999; Theißen and Saedler, 2001; Melzer and Theissen, 2009; Melzer et al., 2009). The floral quartets determine the identity of each floral whorl. *SEP* proteins seem to play a pivotal role in these floral quartets (Immink et al., 2009), in combination with class B (*AP3* and *PI*) and either class A (*AP1*) or class C (*AG*), to specify petal or stamen identity, respectively (Honma and Goto, 2001).
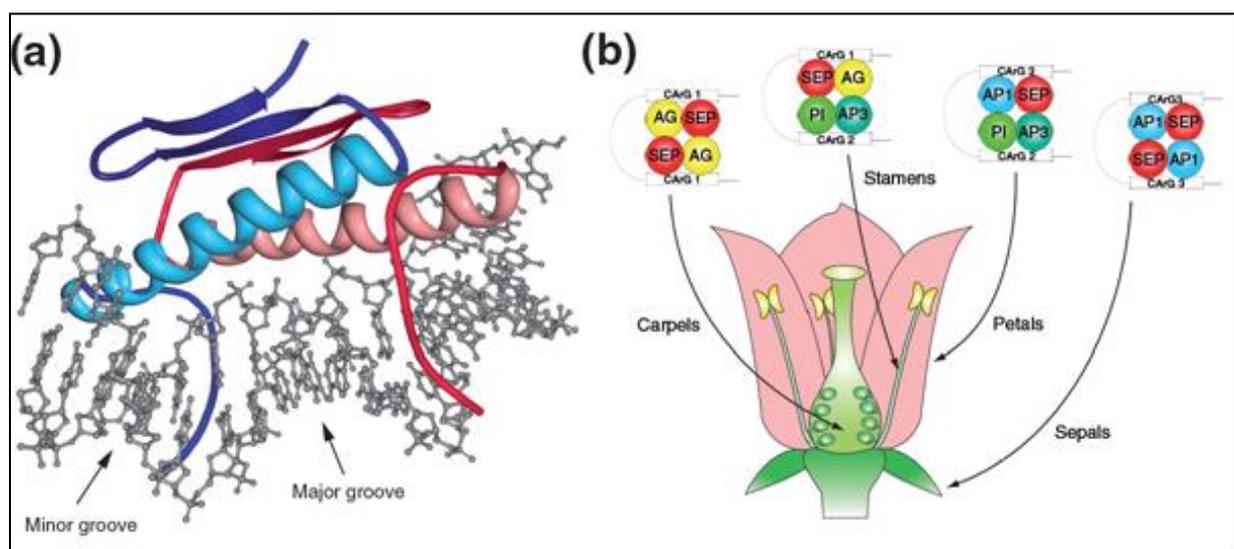

Figure 4: (a) crystal structure of SRF dimer bound to DNA (Pellegrini et al., 1995) (b) The floral quartet model (Theißen and Saedler, 2001).

Various evidence supports the floral quartet model, summarized as follows:

- Formation of a tetramer enhances the binding affinity to CArG-box repeats *in vitro* (Egea- Cortines et al., 1999).

- EMSA/gel shift experiments indicated that a complex larger than a MADS-dimer could bind CArG boxes (Egea- Cortines et al., 1999).

- *In planta* experiments showed that simultaneous expression of *AP3, PI, SEP3,* and *AG* converts cauline leaves into stamen-like organs (Honma and Goto, 2001; Pelaz et al., 2001).

- The findings of large-scale yeast-2-hybrid and yeast-3-hybrid experiments, followed by fluorescence resonance energy transfer-fluorescence lifetime imaging (FRET-FLIM), attribute a pivotal role to SEP3 in mediating formation of a multimeric complex with AP3 and PI (Immink et al., 2002; Immink et al., 2009).

- Immuno-precipitation of GFP-fused MADS-box proteins confirmed interaction as predicted by the quartet model (Smaczniak et al., 2012b).

- MADS tetramers can bind DNA at two CArG boxes in close proximity of each other, resulting in bending of the DNA into a loop, mediated by SEP3 (West and Sharrocks, 1999; Melzer et al., 2009; Smaczniak et al., 2012c).

- Yeast-3-hybrid and yeast-4-hybrid experiments confirm complex formation of AP3, PI, AG and SEP3 (Smaczniak et al., 2012c).

1.4 Higher order complex

In addition to MADS-box tetramers, immuno-precipitation and EMSA experiments led to the identification of non-MADS proteins as interaction partners of plant MADS proteins involved in floral organ identity. For example, SEP3 was found to be part of a large protein complex of 670 kDa, significantly larger than the size of a MADS tetramer (Smaczniak et al., 2012c). *AUXIN RESPONSE FACTOR 2 (ARF2)* and *SQUAMOSA PROMOTOR BINDING PROTEIN-LIKE 8 (SPL8)*, both known to function in early flower development (Okushima et al., 2005; Xing et al., 2013), were revealed to be involved in a complex with *AG*, and *AP1* (Smaczniak et al., 2012b). The current hypothesized mechanism of regulatory action by MADS-box tetramers is described as follows: MADS domain proteins form a quaternary complex and bind two CArG boxes in close proximity, resulting in a DNA loop. Then, transcriptional cofactors and chromatin remodelling proteins are recruited, and gene expression is altered (Figure 5) (Smaczniak et al., 2012c).
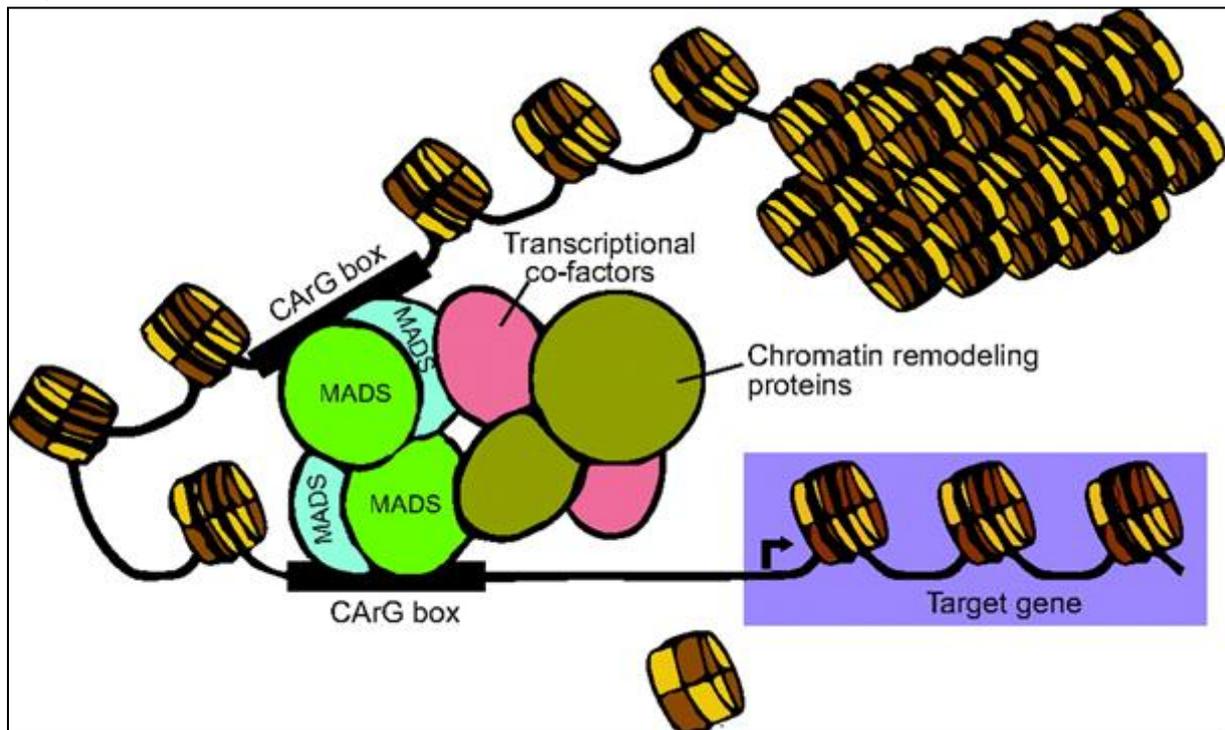

Figure 5: This schematic representation shows the hypothesized mechanism of regulatory action: MADS domain proteins form a quaternary complex and bind two CArG boxes in close proximity, resulting in a DNA loop. Then, transcriptional cofactors and chromatin remodelling proteins are recruited, and gene expression is altered *(Smaczniak et al., 2012c)*.

1.5 Research aim, question & hypothesis

MADS dimer targets

As described in the floral quartet model (Section 1.3), MADS proteins were found to function as tetramers or even higher order complexes during flower development. However, it is unclear whether all floral MADS target genes are regulated by a higher order complex, or if a MADS dimer could also be sufficient to regulate target gene transcription. To address this question, we will use the novel tool for targeted mutagenesis *in planta*, CRISPR/Cas9 (Jinek et al., 2012; Hyun et al., 2015). We intend to create a robust *sep1sep2sep3* triple knockout mutant, followed by introducing a truncated version of *SEP3* that lacks the tetramerization domain, connected to a receptor to enable glucocorticoid-inducible gene expression (Aoyama and Chua, 1997), *SEP3ΔC-GR*. RNA-seq will be used to compare differential expression upon *SEP3ΔC-GR* induction with differential expression after *SEP3-GR* induction to gain information on the function of the tetramer, as opposed to a dimer. A more detailed description of the CRISPR system is presented in Section 2.2.

Interaction of *ARF2* & *SPL8* with MADS

Second, the recruitment of non-MADS-box transcription factors that are thought to form a higher order complex will be investigated. Recently, Smaczniak *et al.* (2012) performed Immuno-Precipitation (IP) experiments that identified enrichment of *AUXIN RESPONSE FACTOR 2 (ARF2)* and *SQUAMOSA PROMOTOR BINDING PROTEIN LIKE 8 (SPL8)* within the samples targeting AP1, AG and SEP3 (Smaczniak et al., 2012c). ChIP-SEQ data of AP1 and SEP3 revealed enrichment of ARF and SPL8 binding motifs (Kaufmann et al., 2009). This suggests that these transcription factors assemble into complexes that bind to nearby sites in the same genomic region, but interaction has to be confirmed by another method. Yeast-2-Hybrid (Y2H) and Yeast-3-Hybrid (Y3H) screening methods will be applied to determine if ARF2 and SPL8 interact directly with MADS-box transcription factors and known MADS-box dimers (de Folter et al., 2005), respectively.

2. Methods

2.1 Generation of a *sep1sep2sep3* triple mutant line

In order to investigate the different actions of MADS tetramers and MADS dimers, a knock-out mutant must be created to allow observation of the different effects of rescue constructs containing *SEP3-GR* or *SEP3ΔC-GR,* respectively. Due to redundancy of *SEP3* with *SEP1* and *SEP2,* a homozygous *sep1sep2sep3* triple knockout mutant is required.

Two *Arabidopsis thaliana* lines with mutations in *SEPALLATA* genes were already present at the start of this thesis. The *sep3-1* mutant line S1056 has an insertion of the *En-1* transposon element in *SEP3,* 1050 bases after the start codon, in the second exon, with the 5' end towards the start of the gene. The mutant line S3843 combines *sep1-1*, isolated from the DuPont collection (Feldmann, 1991), with *sep2-1*, that has an *En-1* insertion 1889 bases after the start codon, in the seventh intron; the 5' end of the transposon element is towards the 5' end of *SEP2*. Based on similar mutations in other MADS-box genes, *sep2-1* is believed to represent an intermediate allele (Pelaz et al., 2000). Furthermore, this line contains a *pSEP3:SEP3-GFP* construct/transgene. The *En-1* insertion alleles of *SEP2* and *SEP3* were identified by Pelaz et al. (2000).

Additional mutant lines were ordered from the SALK collection. These SALK lines contain a T-DNA insert in the gene of interest, with a selection gene providing plants with resistance to kanamycin (*KAN*). An overview of the available *sepallata* mutant lines is presented in Table 2.

Table 2: Mutant lines. The SALK seeds are segregated T3 lines.

Mutant line designation		Mutated gene	Genotype
S3843		*sep1* (DuPont); *sep2* (*En-1*); *SEP3-GFP*	Homozygous for *sep1-1* and *sep2-1*; Heterozygous for *SEP3-GFP*
S1056		*sep3* (*En-1*)	Heterozygous (?)
S6435	SALK_121233	*sep1* (SALK)	Homozygous
S6436	SALK_065340	*sep3* (SALK)	Heterozygous
S6556	SALK_065223	*sep3* (SALK)	Heterozygous

Due to the unstable nature of the *En-1* transposable element, occasional restoration of the wild-type allele may occur, making the transposon lines less favoured to work with. The seeds ordered from the SALK collection are T3 from a mixed population of self-pollinated T2. Table 2 includes the genotype of the received lines. PCR was performed on all plant lines, in order to confirm their genotypes. The detailed description of all materials and protocols used throughout this project are presented in Supplementary Data IV.

The SALK collection lacks mutant lines for *SEP2*, therefore another way is required to obtain a stable *sep1sep2sep3* triple knockout mutant line. A new tool for targeted mutagenesis, CRISPR/Cas9, has been proposed (Jinek et al., 2012). Application of the CRISPR/Cas9 system has recently been shown to also be successful *in planta*, in addition to other

eukaryotes (Feng et al., 2013; Jiang et al., 2013; Li et al., 2013; Mao et al., 2013; Fauser et al., 2014; Feng et al., 2014). CRISPR/Cas9 is therefore selected to create a mutation in *SEP2*. The principle of CRISPR/Cas systems is explained in the next section (Section 2.2), followed by a description of the approach to target *SEP2* for mutagenesis employed in this project (Section 2.3).

2.2 CRISPR: Principle & Application

Clustered Regularly Interspaced Short Palindromic Repeats, or CRISPRs, were found to be part of a prokaryotic RNA-mediated adaptive defence mechanism, CRISPR/Cas (from **C**RISPR-**as**sociated), which protects organisms against invading plasmids and viruses (Barrangou et al., 2007; Bhaya et al., 2011; Gasiunas et al., 2012). During the adaptive phase, short DNA sequences (22bp), originating from these potentially dangerous intruders, are integrated into CRISPR arrays as so-called spacers, interspersed with identical repeats. After this first step of CRISPR/Cas-mediated immunity, during the expression phase, the repeat-spacer element is transcribed into precursor CRISPR RNA, pre-crRNA. Next, Cas proteins cleave the pre-crRNA into short spacer crRNAs, which target the proto-spacer sequences that have a 3' proto-spacer adjacent motif (PAM) of previously encountered viruses or plasmids. Target recognition results in silencing of these foreign sequences, through generation of double stranded breaks (DSB) in the target DNA, by Cas proteins in complex with the crRNAs (Barrangou et al., 2007; Sternberg et al., 2014).

Of the three types of CRISPR/Cas systems known, type II shows promising potential for application in research, as a tool for targeted mutagenesis in higher eukaryotes (Jinek et al., 2012; Chang et al., 2013; Hwang et al., 2013; Wang et al., 2013), including higher plants (Feng et al., 2013; Jiang et al., 2013; Li et al., 2013; Mao et al., 2013; Fauser et al., 2014; Feng et al., 2014). Mutations created by CRISPR type II seem to be integrated stably and are heritable in *Arabidopsis thaliana* (Feng et al., 2014). In this system, trans-activating crRNA (tracrRNA), that is complementary to the repeats in the pre-crRNA, triggers processing by RNaseIII in complex with Cas9 (Gasiunas et al., 2012; Jinek et al., 2012). Cas9, in complex with the processed repeat-spacer element and tracrRNA, binds a target sequence complementary to the spacer, and generates a double-stranded break within the target sequence. Endogenous repair mechanisms such as non-homologous end joining (NHEJ) attempt to reconstruct the DNA. When this occurs successfully, the target sequence is reassembled and the process is repeated until NHEJ results in alteration of the sequence through insertion, deletion or nucleotide exchange at the location of the DSB (Lieber, 2010). For a schematic representation of the CRISPR/Cas type II system, see Figure 6.
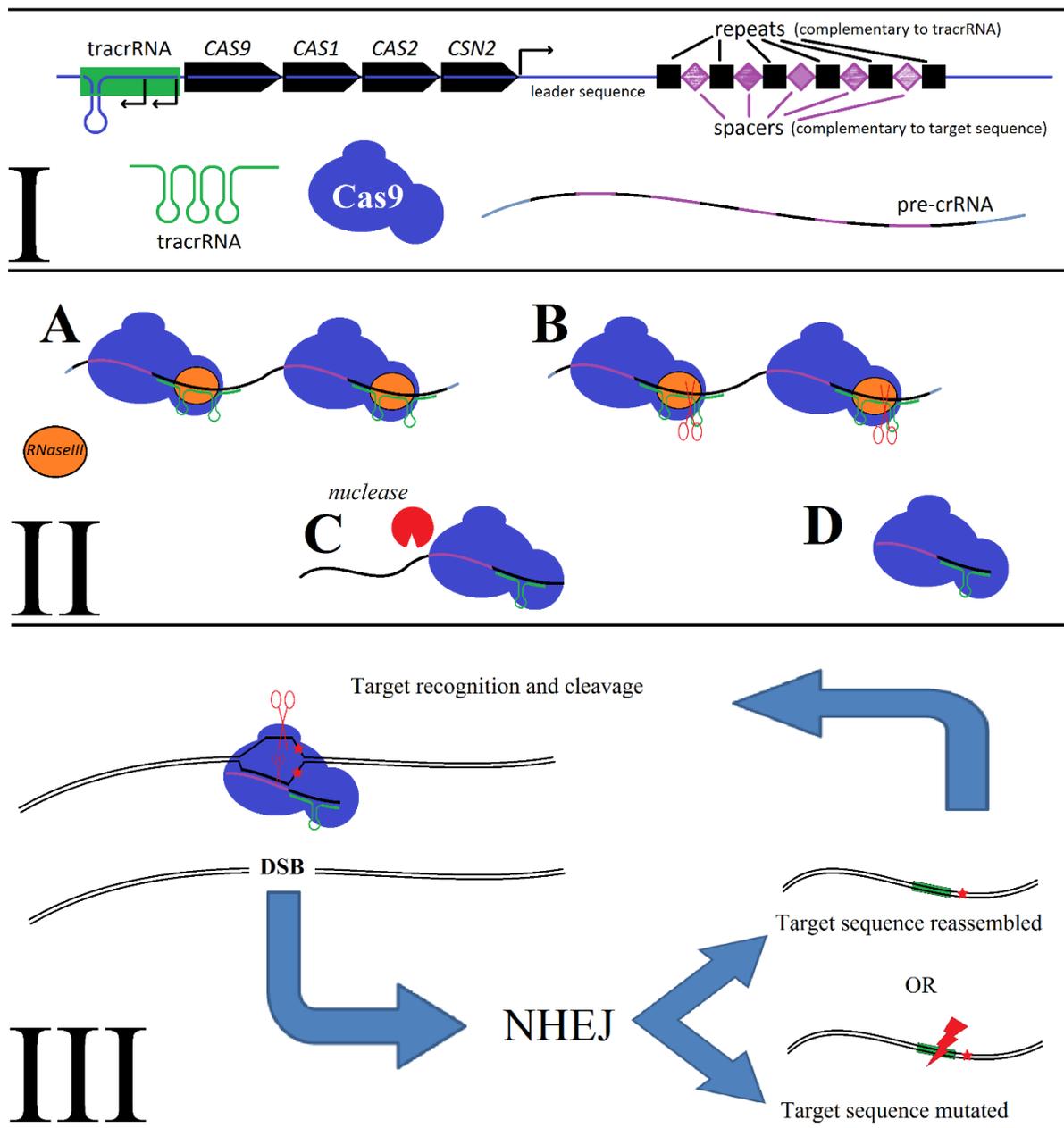
Figure 6: Phase I: the tracrRNA and pre-crRNA are transcribed; the *CAS* genes are translated and Cas9 assembles. Phase II: (A) Pairing of the tracrRNA with the complementary repeats in the pre-crRNA triggers cleavage (B) by the housekeeping endoribonuclease ⬤ RNase III in presence of Cas9. (C) unknown nucleases further process the tracrRNA and intermediate crRNA, resulting in (D) mature crRNA in complex with tracrRNA and Cas9. Phase III: the Cas9-crRNA complex scans the DNA for proto-spacer sequences with the required protospacer adjacent ⭐ motif (PAM). Upon target recognition, the DNA helix is opened at the PAM, and two catalytic domains in the Cas9 protein each cleave one strand of the DNA at a specific site relative to the PAM, within the target sequence. Non Homologous End Joining (NHEJ) often successfully repairs the double stranded break, leading to reassembly of the target sequence, which is then again cleaved by the CRISPR/Cas9 complex. Only when NHEJ introduces a mutation, the target sequence is no longer recognized (adapted from Jinek et al. (2012)).

Jinek *et al.* have constructed a fusion of tracrRNA and crRNA in a single transcriptional unit, which has been shown to efficiently guide Cas9 to any site in the genome that is homologous to the guide RNA followed by a NGG motif (Jinek et al., 2012). This chimeric sgRNA allows a relatively simple method to create CRISPR constructs that target genes of interest for mutagenesis (Jinek et al., 2012; Hyun et al., 2015). The method for creating such a CRISPR construct is displayed in Figure 7.
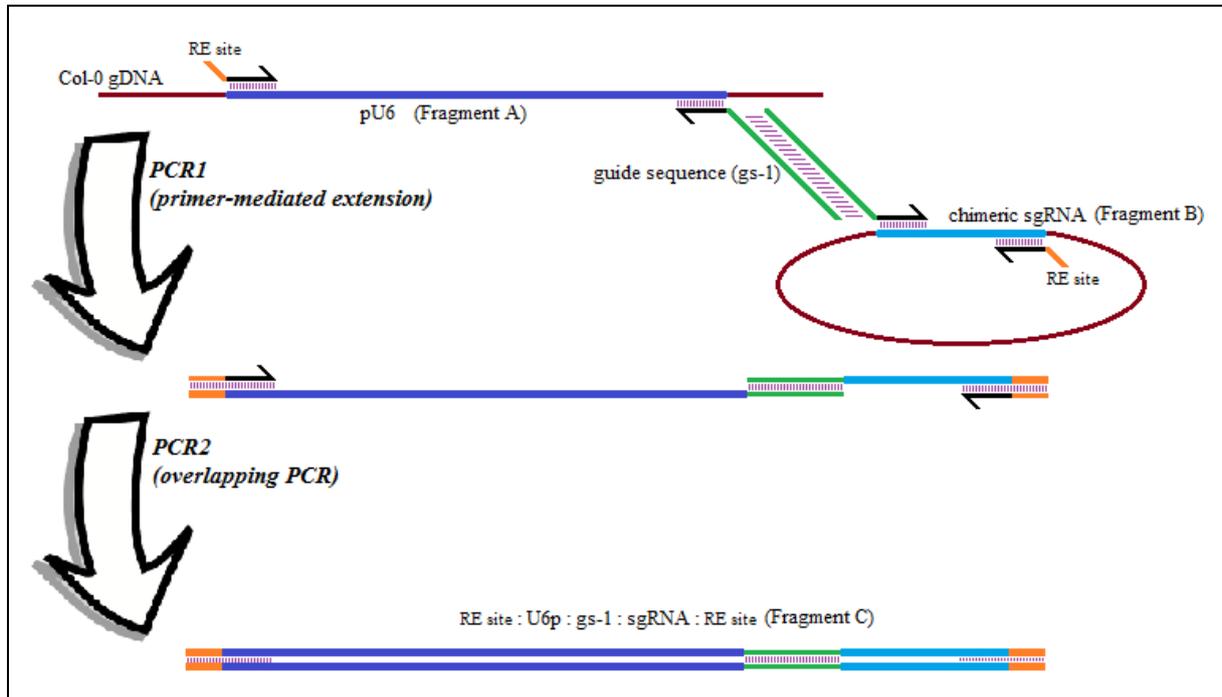

Figure 7: Approach for the generation of a CRISPR construct with chimeric sgRNA. During the first step, PCR with primer-mediated extension is used to attach a restriction enzyme recognition site and the guide sequence to both pU6 (from *Arabidopsis thaliana Col-0* genomic DNA) and the chimeric sgRNA (from vector pRG_ext_CCR5) in separate PCRs. Next, overlapping PCR is performed on the two fragments with overlapping guide sequences in one reaction, resulting in *U6p*: guide sequence: *sgRNA*, flanked by RE sites.

2.3 CRISPR approach: targeting the *SEP2* gene for mutagenesis

In order to create a *sep2* mutant for this research project, the approach described in Figure 7 was selected as a tool. The target sequence in the DNA needs to be accompanied by a Proto-spacer Adjacent Motif (PAM), which is 'NGG' for the CRISPR/Cas type II system. All NGG motifs in the sequence of the *SEP2* gene were identified, after which a target sequence inside an exon was selected. The guide sequence (complementary to the target sequence) was then aligned to the genome sequence of *Arabidopsis thaliana* using NCBI BLAST, to ensure that the sequence is unique within the genome. Additionally, the Cas OFFinder tool was used to screen the genome for potential mismatches (Table S3). The results indicated that a minimum of three mismatches need to occur for the target sequence to bind another site in the DNA that has the adjacent PAM.

After selecting the guide sequence, primers were designed to create a construct containing the guide sequence that targets *SEP2*. The sequence of this construct is shown in Figure 8. During the first round of PCR, the promoter region of the *U6-26* gene was amplified from *Arabidopsis thaliana* genomic DNA (ecotype: Col-0), and the chimeric sgRNA was amplified from the vector plasmid pRG_ext_CCR5 (Cho et al., 2013), which was kindly provided by Youbong Hyun (Max Planck Institute for Plant Breeding Research). Using primer-mediated

extension, restriction sites with short linkers were attached directly upstream of the *U6-26* promotor region sequence and downstream of the sgRNA sequence. The chosen guide sequence was attached directly downstream of *pU6* and directly upstream of the sgRNA using primer-mediated extension, creating fragment A and B, respectively. The second round of PCR makes use of the overlap (guide sequence), attached by primer-mediated extension, to allow annealing of fragment A and B at the overlap. The resulting sequence in displayed in Figure 8.
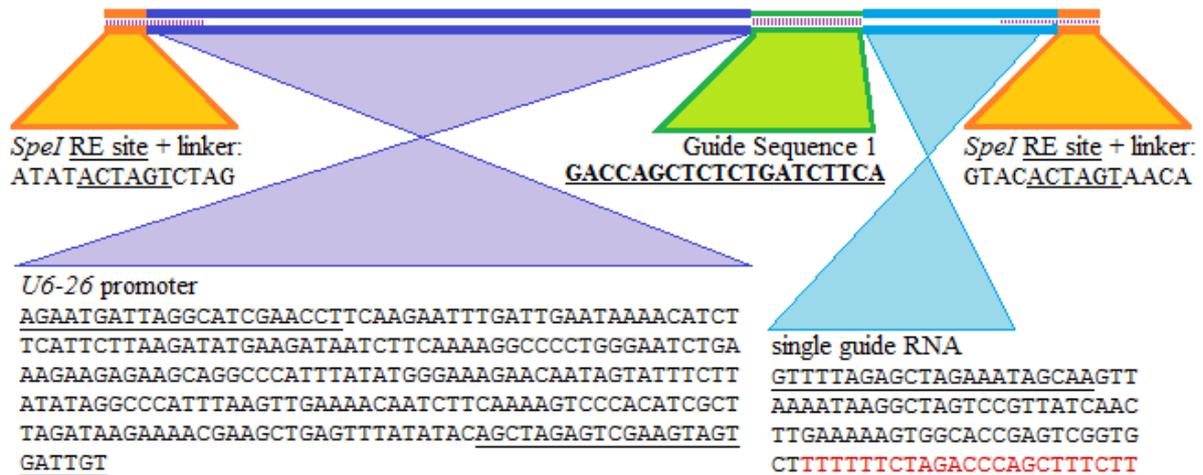

Figure 8: Complete sequence of the CRISPR construct targeting *SEP2* for mutagenesis, using guide sequence 1 and SpeI recognition sites

In order to facilitate expression of both *CAS9* and the p*U6::gs-1::sgRNA* cassette in *Arabidopsis thaliana*, the binary vector pYB196 (~14.5Kbp) (Hyun et al., 2015) was used. This plasmid was constructed by (Hyun et al.) in a pGREEN0229 background, through insertion of *CAS9* under the promoter of *INCURVATA2 (ICU2)*, which is highly active in proliferating cells (Hyun et al., 2013). Moreover, pYB196 contains two adjacent unique restriction sites (BamHI and SpeI), allowing two cassettes to be inserted, thus two different guide sequences can be expressed simultaneously. The sequence of the pYB196 vector can be found under NCBI GenBank accession number KJ816368.

After fragment C was created, both the construct and the binary vector pYB196 were digested using the endonuclease SpeI. Additionally, phosphatase was added to the plasmid digestion sample (Antarctic Phosphatase, AP, or Calf Intestinal Phosphatase, CIP) which catalyzes dephosphorylation of 5' (and 3') ends of DNA. This will prevent re-ligation of the empty vector. Next, fragment C was cloned into the pYB196 vector and subsequently transformed into *E.coli*. Several colonies that had grown on selective media were selected for Colony PCR. Colonies that gave rise to a band of the expected length were sequenced to confirm correct integration into the binary vector. A more elaborate description of these protocols is presented in the supplementary data, including an overview of primers (Table S4 and S5) and part of the sequence of pYB196, in which the SpeI site and primers for Colony PCR are indicated (Figure S3).

2.4 Protein-protein interaction: YnH

Smaczniak et al. (2012c) have identified ARF2 and SPL8 as enriched peptides in Immuno Precipitation (IP) experiments targeting AP1, AG, and SEP3. As *ARF2* and *SPL8* are known to function in early flower development (Okushima et al., 2005; Xing et al., 2013), and DNA-binding motifs for ARF2 and SPL8 were found to be enriched in AP1 and SEP3 ChIP-SEQ peaks (Kaufmann et al., 2009), these non-MADS-box transcription factors may act in complex with MADS-box transcription factors (Smaczniak et al., 2012c). Co-occurrence of certain proteins in an immuno-precipitation sample provides an indication of interaction. In this project, we attempt to further unravel this interaction, by examining direct protein-protein interaction of ARF2 and SPL8 with MADS-box transcription factors, using Yeast-2-Hybrid and Yeast-3-Hybrid assays.

Yeast-n-Hybrid assays provide a standardized means to determine protein-DNA and protein-protein interactions (Ma and Ptashne, 1988; Fields and Song, 1989). Interaction of proteins fused to the separated DNA-binding domain (BD) and activation domain (AD) of the *GAL4* transcription factor results in expression of a reporter gene that, in this case, enables synthesis of histidine, thus allowing growth on selective medium lacking addition of this amino acid (Figure 9). The method used is based on the protocol described by de Folter and Immink, (2011): Yeast protein-protein interaction assays and screens.
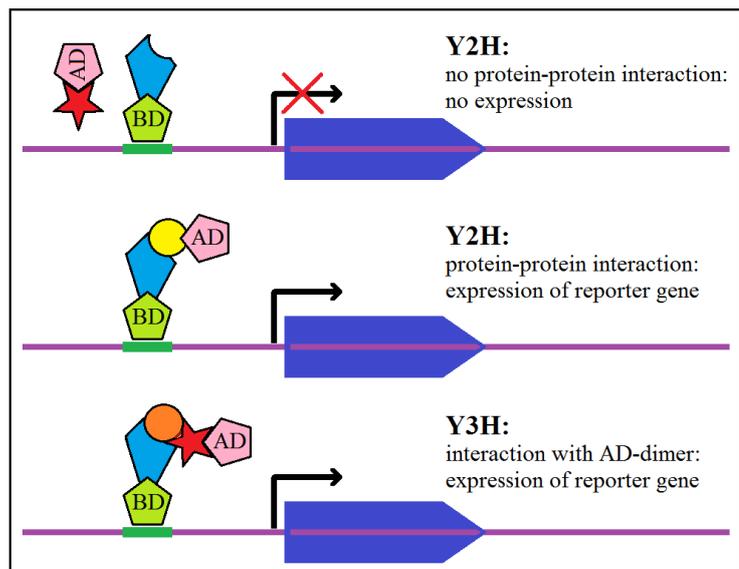
Figure 9: The binding domain (BD) and activation domain (AD) of the GAL4 transcription factor will induce expression of an adjacent gene, if present in the right orientation. Fusing proteins to these domains allows to distinguish whether interaction between these proteins takes place.

2.5 Yeast-n-Hybrid: Approach

A previously constructed prey library of MADS box transcription factors in the *GAL4-AD* vector pDEST22 (Trp), and MADS box dimers in pDEST22 (Trp) and pTFT1/pARC352 (*Ade2*), transformed into the yeast strain *PJ69-4A*, was screened for protein-protein interaction with *SPL8* and *ARF2*. The coding sequences of *ARF2* and *SPL8* were reverse transcribed from mRNA isolated from cauline leaf of *Arabidopsis thaliana Col-0*. The Gateway Cloning system was used to create *BD* vectors (pDEST32, Leu) suitable for Yeast-n-Hybrid experiments. After sequencing confirmed that the genes were in frame with the BD sequence, the constructs were transformed into *S. cerevisiae PJ69-4α*. Yeast colonies that showed growth on SD medium lacking leucine were evaluated by amplifying the insertion region of the destination vector, using colony PCR. The supplementary data includes an overview of primers (Table S6 and S7) and more elaborate protocols. A glycerol stock of the previously made FUL-BD in *PJ69-4α* (de Folter et al., 2005) was retrieved from storage, and used as positive control.

It is possible that the bait protein contains an intrinsic activation domain, which would allow growth on selective medium in the absence of protein-protein interaction. Alternatively, a reversion towards a functional gene could occur at the mutant *his3* locus (Scherer et al., 1982). Therefore, an auto-activation test was performed on yeast containing the BD construct: 3-Amino-1,2,4-triazole (3-AT) is a competitive inhibitor of the *HIS3* gene product, and can be applied to repress auto-activation.

The selection markers in these vectors allow the yeast strain to synthesize specific amino acids, and therefore a strain containing a certain vector is able to grow on selective medium lacking the corresponding amino acid. After selection for yeast cells that contained the pDEST22 or pDEST32 vector on SD medium supplemented with dropout complete that lacks tryptophan or leucine, respectively, the two vectors were combined in one diploid cell, by allowing the *PJ69-4A* strains to mate with the *PJ69-4α* strains. Selection occurred on SD medium supplemented with dropout solution -Leu/-Trp. The positive colonies were plated out on new selective media, containing SD-Leu/-Trp/-His (optionally supplemented with 3-AT to repress auto-activation) and SD-Leu/-Trp/-Ade. Possible positive interactions were selected for verification of the activity of the reporter genes. The approach is summarized in Figure 10.
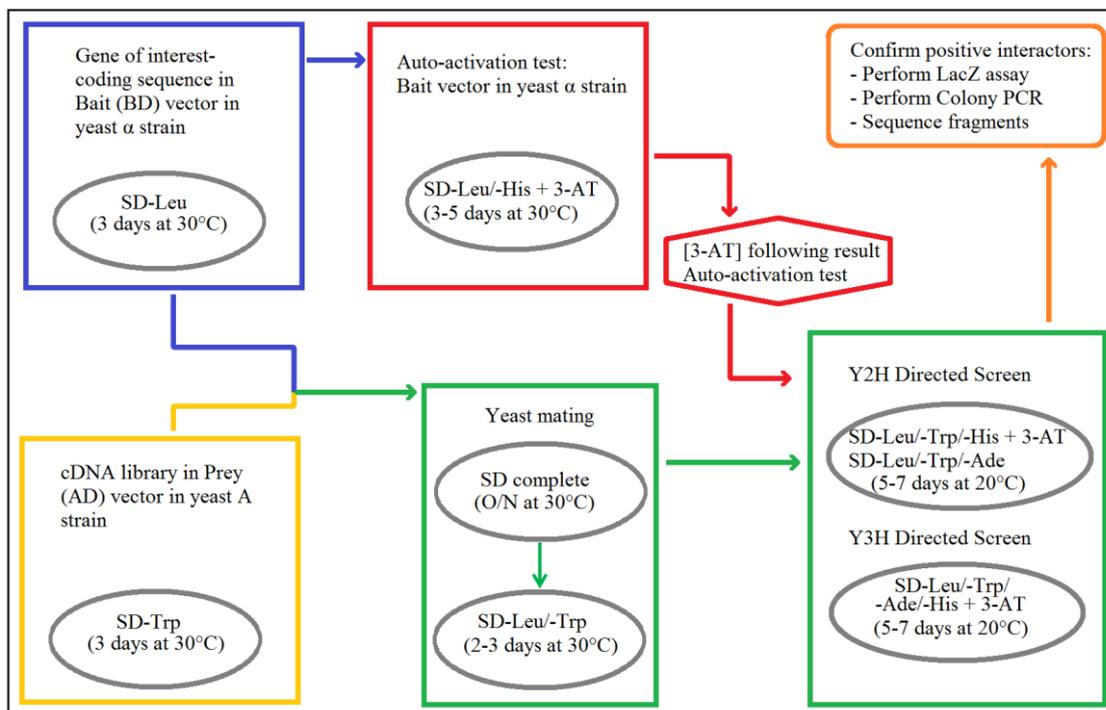
Figure 10: The BD vectors in *PJ69-4α* (blue) were allowed to mate with each AD-vector *PJ69-4A* strain (yellow) in order to combine the vectors (green). Additionally, the *PJ69-4α* strains with a BD-vector were subjected to an auto-activation test (red), leading to the concentration of inhibitor (3-AT) to use during the directed library screen (green). Positive interactors are confirmed by LacZ assays, Colony PCR, and sequencing of the fragments inserted in the YnH vectors

3. Results

The first goal of this project was to obtain a robust *sep1sep2sep3* triple mutant, to provide a background for the complementation constructs. To achieve this, the genotype of several mutant lines was determined.

3.1 Genotyping (& Phenotyping) *SEP* mutants

Transposon mutant lines.

Two *Arabidopsis thaliana* lines with mutations in *SEPALLATA* genes were already present at the start of this thesis. The mutant line S1056 contains *sep3-1,* annotated by Pelaz et al. (2000), which has an insertion of the *En-1* transposon element in *SEP3,* 1050 bases after the start codon, in the second exon, with the 5' end towards the start of the gene.

The mutant line S3843 combines *sep1-1*, isolated from the DuPont collection, with *sep2-1* that has an *En-1* insertion 1889 bases after the start codon; the 5'end of the transposon element is towards the 5'end of *SEP2*. Additionally, this line includes *SEP3-GFP* under its native promoter. PCR was performed on several plants, in order to confirm this genotype. During gel electrophoresis of the PCR samples, a 1Kb marker (M) was included to determine the fragment lengths afterwards (Figure 11b).

   a)   S3843 sep1-1/sep1-1 sep2-1/sep2-1 SEP3-GFP

PCR was performed on genomic DNA extracted from young rosette leaves of an *Arabidopsis thaliana* S3843 plant, using primers that amplify the wild-type or mutant allele of the *SEP1* and *SEP2* genes. A control with wild-type genomic DNA shows the expected bands of 500 bp and 797 bp for *SEP1* and *SEP2*, respectively. In combination with their respective mutant allele primers, mutant alleles should result in amplification of a fragment of 1600 bp for *sep1* and 400 bp for *sep2*. The results are presented in Figure 11a.
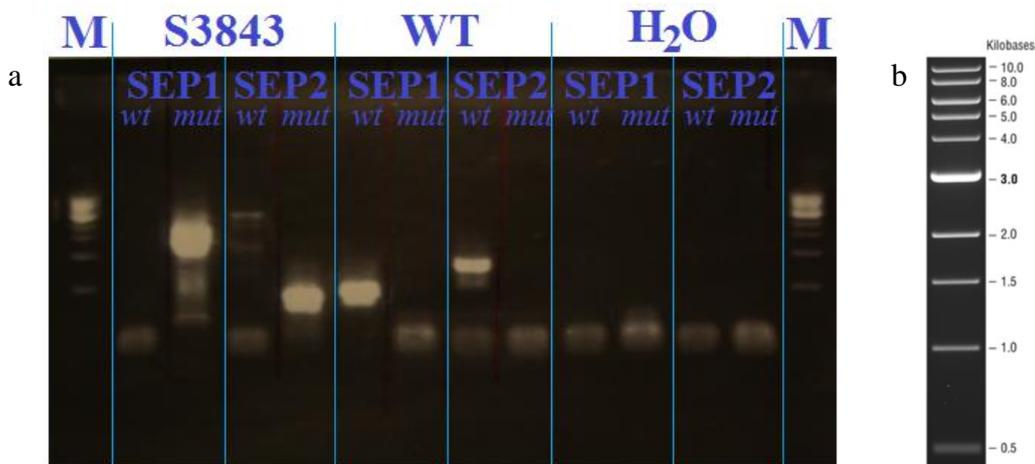
Figure 11: (a) Results of PCR on the S3843 plant.   (b) 1Kb DNA ladder
(www.neb.com)

As can be seen in Figure 11a, the expected mutant bands appear in the mutant lanes of S3843, hereby confirming the homozygous mutant phenotype of the plant. Several vague bands were observed in the *SEP2* wt lane of S3843. This is due to the unstable nature of transposable element: excision occurs in some cells, resulting in wild-type product (797bp) in all lanes (Feschotte et al., 2002). The other bands in this lane likely resulted from a-specific annealing of the primers. After the homozygous *sep1sep2* double knock-out mutant genotype was confirmed, seeds were isolated after self-pollination. Several offspring plants were grown on rockwool.

For one of the offspring plants, #9, a mild phenotype was observed, with the following characteristics: a limited number of flowers showed a perturbed morphology, such as lack of petals and/or sepals, multiple carpels and short siliques with a reduced seed set. Some of these flowers are displayed in Figure 12.
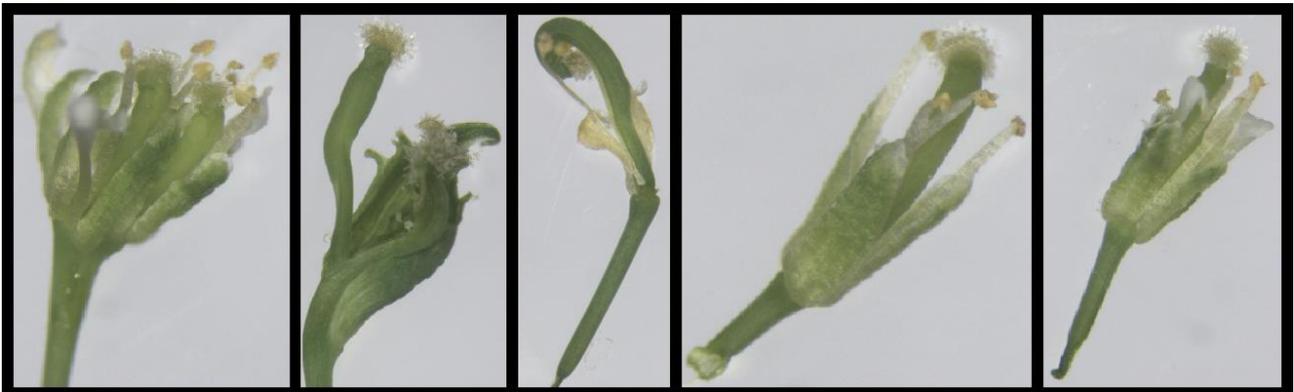
Figure 12: Mild phenotype of the S3843 offspring, plant number 9

DNA was extracted from young cauline leaves of the #9 offspring plant, this time including primers to confirm presence of *SEP3-GFP*. The results are presented in Figure 13, and confirm that the plant showing this phenotype is homozygous *for sep1sep2* mutant alleles, while the band amplified for *SEP3-GFP* resembles the band produced by the WT control.
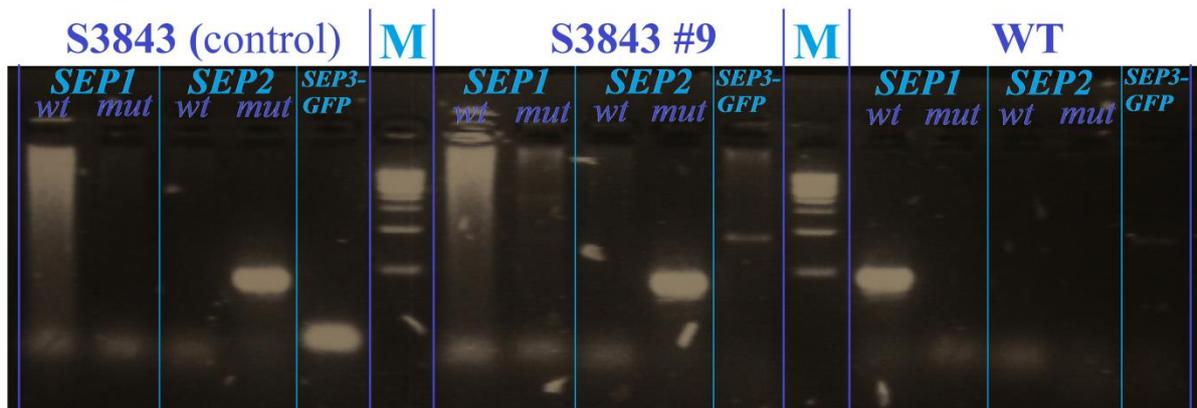
Figure 13: PCR results for S3843 offspring, plant number 9

b) S1056 sep3-1 (Pelaz et al., 2000)

A PCR was performed on genomic DNA extracted from an *Arabidopsis thaliana* S1056 plant using two primer sets. The "wt" set (8096+8097) amplifies a fragment of 734 bp if the wild-type allele of *SEP3* is present in the template DNA. The second set binds the *En-1* transposon and a region upstream of *SEP3*. A *sep3* transposon homozygous mutant should give rise to a fragment with a length of 418 bp when using the mutant primer set (8095+776). The results are presented in Figure 14.

16
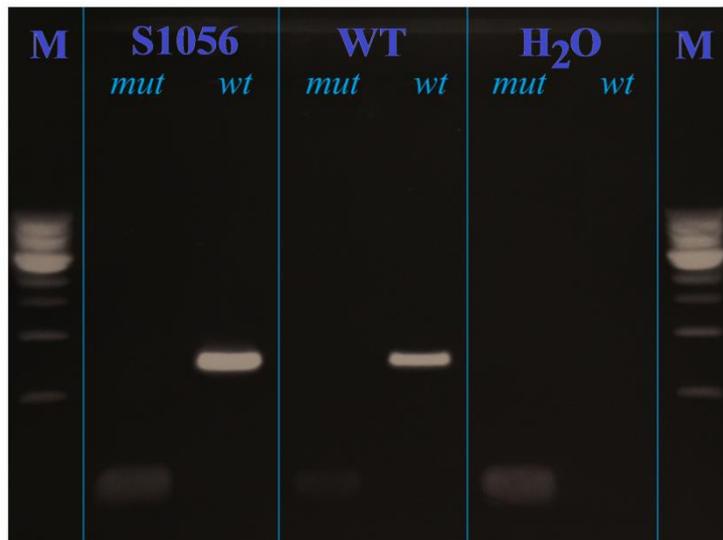
Figure 14: Results of PCR on the S1056 plant

The genomic DNA of S1056 contains at least one copy of the wild-type allele for *SEP3*, but the mutant allele could not be amplified. At the end of this project, it was discovered that the primers used to amplify the mutant allele of *sep3-1*, were designed to amplify the mutant allele of *sep3-2* (Pelaz et al., 2000), which has the *En-1* transposable element in the second exon, 1050 bases after the start codon of *SEP3*, in an orientation opposite to that of the *sep3-1* mutant. This explains the lack of amplification of a mutant alleles: both primers bind in the same orientation and can therefore not lead to amplification of a fragment of defined length.

SALK mutant lines.

Although the knock-out genotype of S3843 was confirmed, it appears that the *En-1* transposable element has been excised in part of the cells, restoring the wild-type allele for *SEP2*. This can be observed as a vague band of 797bp in Figure 11a. This may also have occurred in the S1056 line, in which the mutant allele of *SEP3* could not be confirmed either. As a result of the instable alleles, it has been very difficult to maintain the *sepallata* triple mutant in different labs. Therefore, new mutant lines were ordered from the SALK collection. This is a sequence indexed library of insertion mutations generated using the pROK2 T-DNA vector, which provides resistance to kanamycin as a selection marker.

c) S6435 sep1, (SALK_121233)

Seeds from a homozygous *sep1* mutant, SALK_121233, were ordered, and the line was named S6435. Very few seeds germinated (3 out of ±50 seeds), and DNA was extracted from two viable seedlings. PCR was performed using three primers in four combinations as displayed in Figure 15: PDS8093 and PDS8094 bind regions that flank the *SEP1* gene, while PDS5156 binds the T-DNA insert. Aside from the S6435 samples (S1 and S2), two wild-type controls were used (W1 and W2), and also mixtures of the S6435 genomic DNA with WT were included in the PCR: M1 (S1+W1) and M2 (S2+W1). The results are presented in Figure 15.
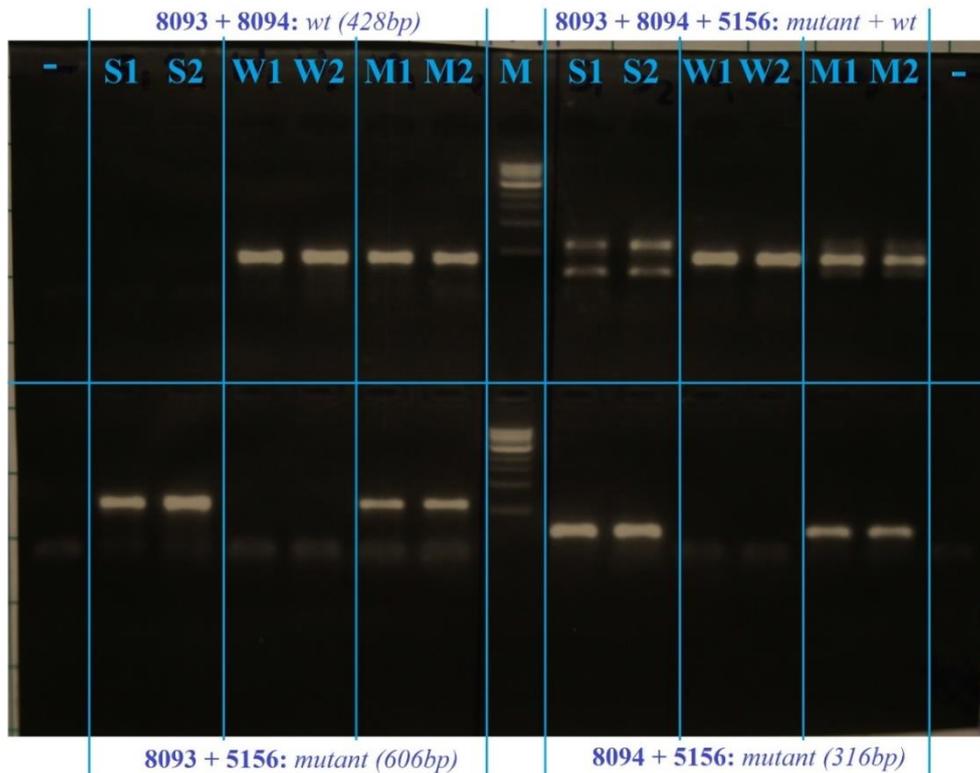
Figure 15: Results of PCR on the S6435 plant

Based on the results shown in Figure 15, the homozygous *sep1 T-DNA* genotype is confirmed. It seems that the T-DNA insert is present in tandem and in two opposite orientations facing away from each other, as both combinations of PDS5156 (the primer binding *T-DNA)* with either of the two primers flanking *SEP1,* result in amplification of a fragment.

d) S6436 (sep3, SALK_065340)

The *sep3 T-DNA* line SALK_065340 was named S6436, and delivered in the form of seeds derived from an T2 offspring population, thus containing a mix of WT, heterozygous and homozygous mutant plants. All seeds germinated, but showed sensitivity to kanamycin (30µg/L) in ½ MS growth medium. PCR was performed, using two primers, PDS8091 and PDS8092, flanking the *SEP3* gene to amplify the WT allele, or PDS5156 in combination with either one of the flanking primers, to amplify mutant allele 1 or 2, representing *sep3* with the T-DNA insertion in two possible orientations. The results are presented in Figure 16.
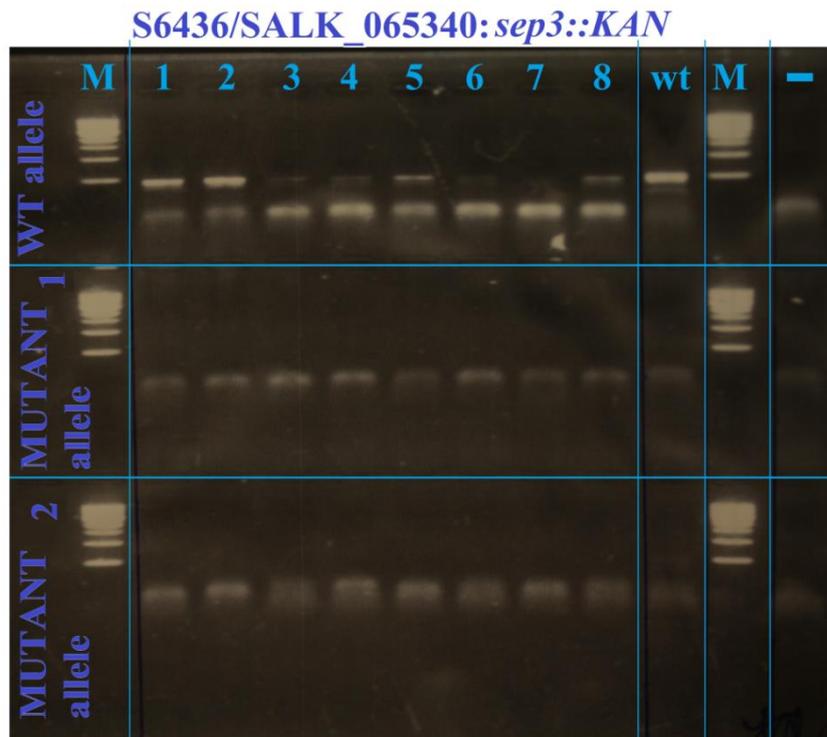
Figure 16: Results of PCR on the S6436 plant

Although every sample gives rise to (more or less) wild-type product, none of the samples seems to contain the T-DNA insert in the *SEP3* gene. Another line from the SALK collection was ordered, that should also contain the T-DNA insertion within the *SEP3* coding sequence, and was named S6556.

e) S6556 (sep3, SALK_065223)

A PCR was performed to amplify part of the T-DNA insert, using primers that amplify part of the *KAN* gene. The genomic DNA of a *ful-7* SALK mutant, containing the same T-DNA, is used as template in the positive control. This allowed a quick screening to exclude plants without T-DNA, thus homozygous for the wild-type allele of *SEP3*. The results are presented in Figure 17.
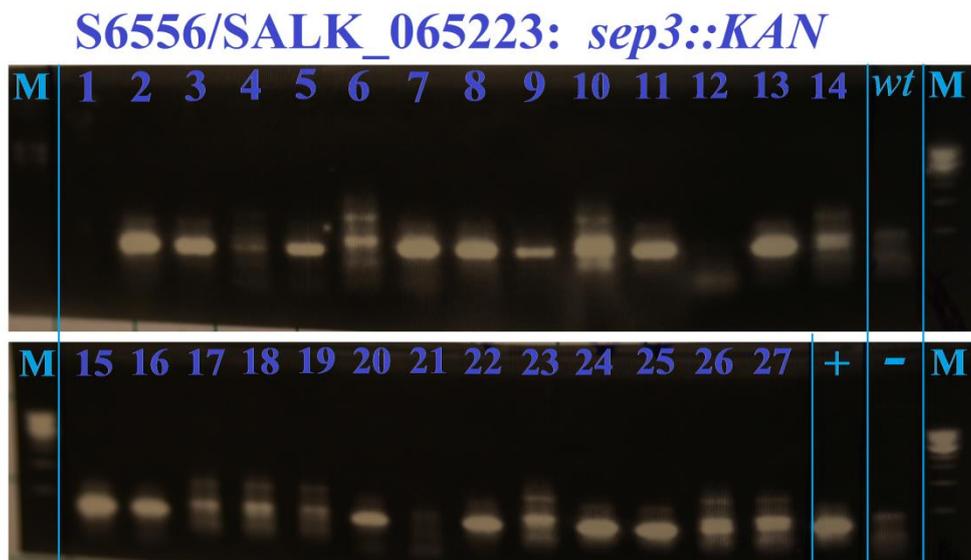
Figure 17: Results of PCR on the S6556 plant

Almost every plant sample leads to amplification of a fragment using primers that target the T-DNA. The positive control indicates which fragment length represents a T-DNA insertion. Three of the S6556 plants shown in Figure 17, whose genomes encode the *T-DNA* (#22, 24, & 25), were selected for PCR, using three combinations of primers: PDS8091 and PDS8092 flank the *SEP3* gene, while PDS5048 binds the T-DNA insert (Left Border of pROK2). The results are presented in Figure 18.

Figure 18: Results of PCR on the S6556 plants #22, 24, & 25

Although the S6556 samples show a band for the *SEP3* wild-type allele, no DNA fragment is amplified using the two mutant allele combinations. In combination with confirmed presence of the T-DNA insert, it was concluded that SALK_065223 is not a *sep3* mutant line. An explanation for these observations could be that the T-DNA is inserted in another gene and wrongly classified as a *sep3* mutant. Thus, neither for SALK_065340 (S6436), nor for SALK_065223 (S6556), a T-DNA insertion in the *sep3* allele could be confirmed, and true T-DNA insertion lines do not seem to be present for *SEP3*.

3.2 CRISPR

Due to the unstable nature of the *En-1* transposable element, occasional restoration of the wild-type allele may occur, making the transposon lines less favoured to work with. The SALK collection lacks mutant lines for *SEP2*, therefore another way is required to obtain a stable *sep1sep2sep3* mutant line. A novel tool for targeted mutagenesis *in planta*, CRISPR/Cas type II, was investigated as an alternative means to obtain the required triple knockout mutant, by targeting *SEP2* for mutagenesis in a *sep1sep3* double mutant background. The principle and approach of this method, as designed by Hyun et al. (2015), is more elaborately described in section 2.3 and in the supplementary data. The following section describes the generation of a CRISPR construct that targets *SEP2.*

a) Construct: fragment A + B = C

By means of the first round of PCR, fragment A (285 bp) and B (134 bp) were created, containing a short (20 bp) overlapping sequence: the chosen guide sequence (gs-1). During the first cycle of the second PCR, fragment A and B were connected through annealing of the overlapping guide sequence, resulting in fragment C (399 bp), which was then amplified during the following PCR cycles. The results are presented in Figure 19. As can be seen, the fragments created, A, B, and C, seem to be of the expected lengths.
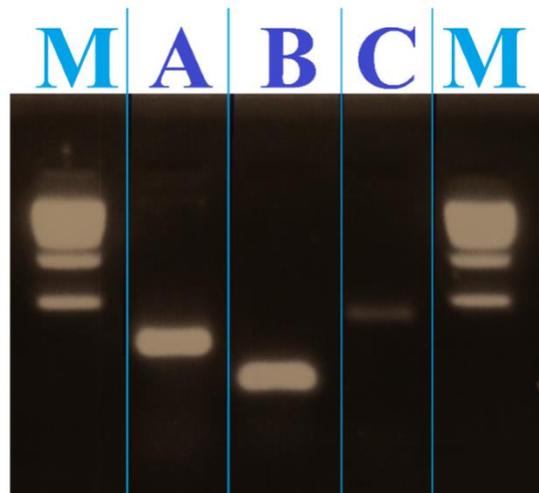
Figure 19: Results of PCR 1 and 2 for CRISPR gs-1

Purified PCR product of fragment C was sequenced, using the same primers as for the overlapping PCR, and aligned with the sequence of the desired construct. The results are presented in Figure 20. The guide sequence is indicated in yellow, with part of the surrounding sequence in green. The quality of the sequencing is lower nearing the end of the fragment, but this is complemented by sequencing in the opposite direction. It seems that the overlapping PCR was successful, as the sequence results have the exact guide sequence at the right location. The quality of the guide sequence and complete alignment are presented in supplementary data.

```
fragment_C    GCCCATTTAAGTTGAAAACAATCTTCAAAAGTCCCACATCGCTTAGATAAGAAAACGAAGCT  230
8066_RV_      GCCCATTTAAGTTGAAAACAATCTTCAAAAGTCCCACATCGCTTAGATAAGAAAACGAAGCT  235
8064          GCCCATTTAAGTCGAAAACAATCTTCAAAAGTCCCACATCGCTTAGATAAGAAAACGAAGCT  204
              ***********. ************************************************

fragment_C    GAGTTTATATACAGCTAGAGTCGAAGTAGTGATTGTGACCAGCTCTCTGATCTTCAGTTTTA  290
8066_RV_      GAGTTTATATACAGCTAGAGTCGAAGTAGTGATTGTGACCAGCTCTCTGATCTTCAGTTTTA  295
8064          GAGTTTATATACAGCTAGAGTCGAAGTAGTGATTGTGACCAGCTCTCTGATCTTCAGTTTTA  264
              ************************************************************

fragment_C    GAGCTAGAAATAGC-AAGTTAAAATAAGGCT-AGTCC-GTTATC-AAC-TTGAAAAAGTGGC  347
8066_RV_      GAGCTAGAAATAGCGAAGTTAAAATAAGGCTTAGTCCTGTTATCTAACCTTT----------  342
8064          GAGCTAGAAATAGC-AAGTTAAAATAAGGCT-AGTCC-GTTATC-AAC-TTGAAAAAGTGGC  321
              ************** **************** ***** ****** *** **.
```

Figure 20: Alignment of sequences from fragment C to the theoretical sequence of the construct

b) RE digestion & ligation

Cohesive ends were created on fragment C through digestion of the fragment by the restriction enzyme SpeI (target sequence AGTACT). The plasmid pYB196 was linearized by SpeI, and the vector ends were de-phosphorylated using CIP. Visualized through agarose gel electrophoresis, fully digested pYB196 should appear as a sharp single band. The results are presented in Figure 21. One clear, sharp band is observed for the pYB196 digested sample, indicating that RE digestion by SpeI was successful and complete.
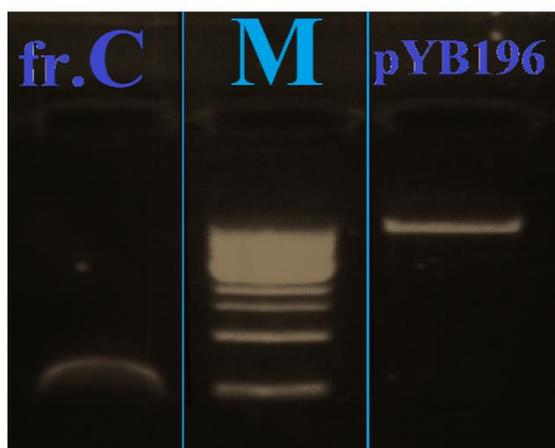
Figure 21: Digestion of fragment C and the pYB196 plasmid

c) colony PCR & sequencing results

After ligation and transformation into *E.coli DH5α*, several colonies were selected that showed resistance to kanamycin (100 mg/L). PCR was performed to amplify a region on pYB196 that includes the SpeI restriction site. The empty plasmid vector pYB196 should give a band of 307 bp and was used as positive control. PCR on colonies that contain a pYB196 vector with an insertion at the SpeI site will amplify a larger fragment of 686 bp, for insertion of the *pU6*:gs-1:*sgRNA* construct. The results of the PCR are presented in Figure 22.
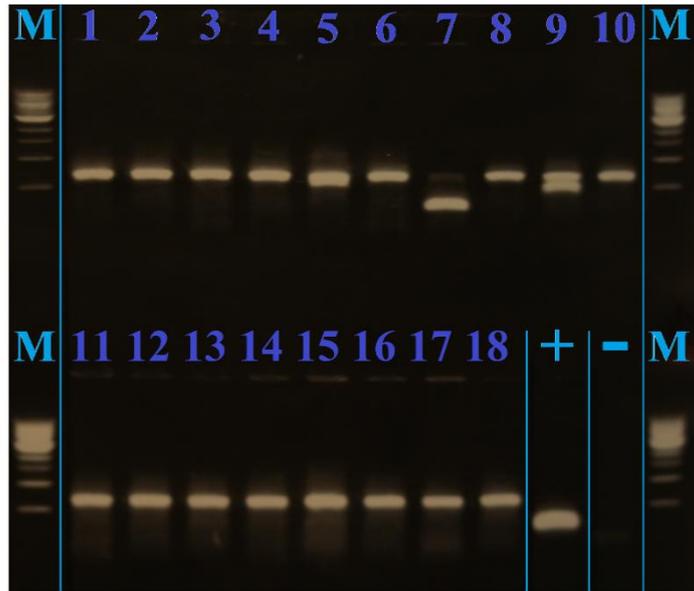
Figure 22: Results of colony PCR

As evident in Figure 22, many samples showed an increase in fragment length of approximately 400 bp in comparison to the empty vector (+). Four colonies (#1-4) were selected for outsourced sequencing by Macrogen. However, none of the Mini-prep samples resulted in trustworthy sequences. The samples were run on agarose gel to determine the purity and concentration of the plasmids. The results are shown in Figure 23.
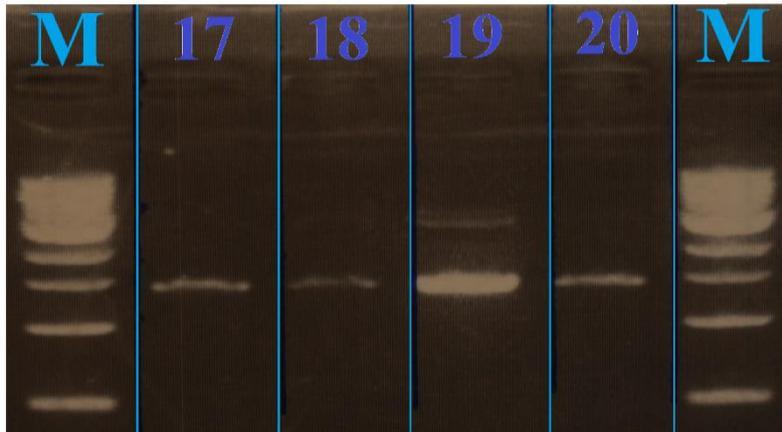
Figure 23: Mini-prep samples on agarose gel

As can be seen in Figure 23, the samples show a single band, indicating only one type of plasmid is present. However, instead of the expected ~14.5Kb of pYB196, the plasmids seem to be of a much smaller size. The digestion and ligation, transformation, and colony PCR were repeated, not leading to the desired sequence inserted at the SpeI site. Among the results was a partial sequence of pYB196. After several unsuccessful attempts, the CRISPR project within this thesis was terminated.

3.3 Yeast-n-Hybrid

Immuno-precipitation experiments identified two non-MADS-box transcription factors (ARF2 and SPL8) that may act in a higher order complex with AP1, AG, or SEP3. Yeast-2-hybrid and yeast-3-hybrid directed library screen assays were performed to investigate direct protein-protein interaction of ARF2 and SPL8 with a set of MADS-box monomers and dimers.

cDNA of the coding sequences of *ARF2* and *SPL8* was reverse transcribed from RNA extracted from cauline leaf, and cloned into bait vectors by means of Gateway Technology (Invitrogen). The screening method was described by de Folter and Immink (2011). A previously made FUL-BD construct was used as positive control (de Folter et al., 2005).

a) auto-activation test

An auto-activation test was performed on the three strains containing pDEST32 with FUL-BD, ARF2-BD or SPL8-BD, by determining growth on different selective media. The results are shown in Figure 24.
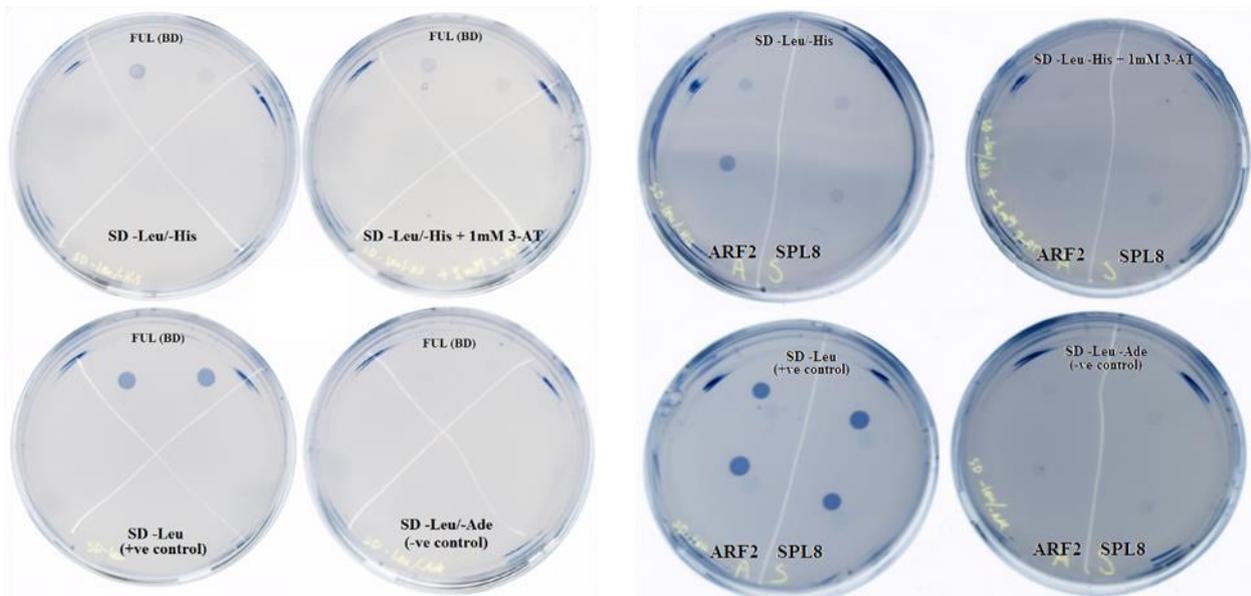
Figure 24:Results of auto-activation test. +ve control: SD-Leu; -ve control: SD-Leu/-Ade. Growth is decreased on SD-Leu/-His, and near absent on SD-Leu/-His + 1mM 3-AT for the three BD vectors in *PJ69-4α*

All strains show growth on the positive control SD-Leu, selecting for presence of the pDEST32 vector. Growth is decreased on plates lacking Leu and His, and is nearly absent when 1 mM 3-AT is added. Selection for protein-protein interaction during the library screening will therefore occur at an inhibitor concentration of 1 mM 3-AT.

b) library screen assay 1

The three strains were allowed to mate with every *PJ69-4A* strain with pDEST22 containing MADS-AD for Y2H, and pDEST22 + pARC352 expressing MADS dimers for Y3H. After incubation on media selective for diploid cells (-Leu/-Trp), the colonies were transferred to selective media for the library screen assay(-Leu/-Trp/-His + 3-AT and –Leu/-Trp/-Ade). The results of the library screen are shown in Figure 25.
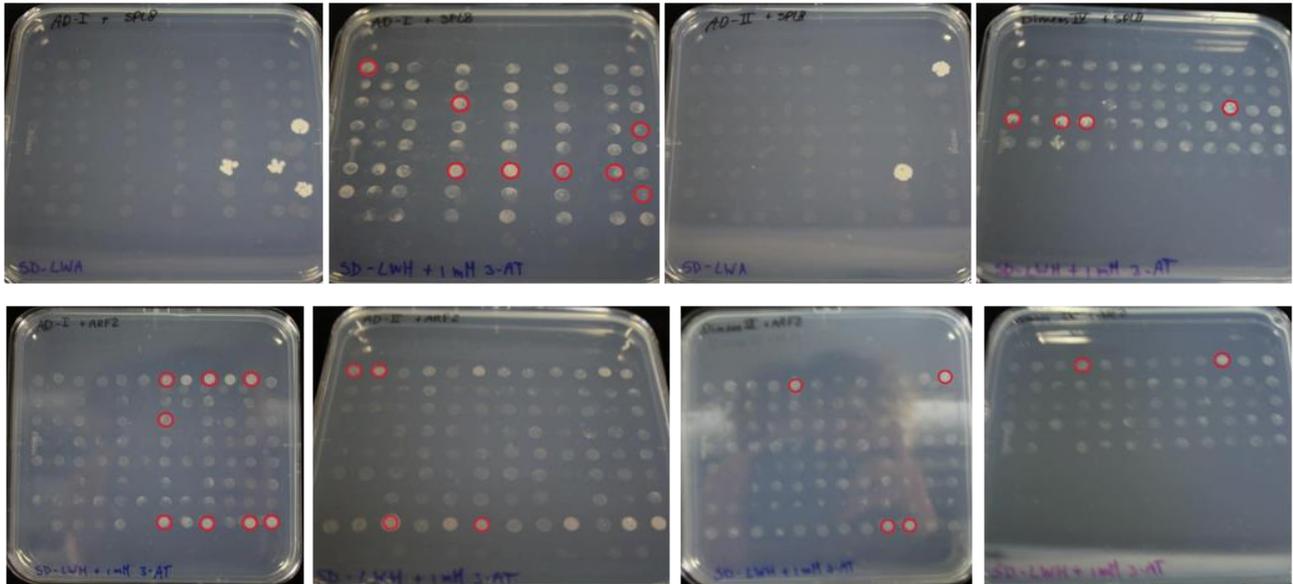
Figure 25:Results of the first library screening assay for SPL8 and ARF2.
Selected colonies are indicated by red circles.

After 7-9 days of incubation at 20°C, several colonies were selected that seemed to show some growth (indicated with red circles in Figure 25). These were inoculated in 2ml liquid SD medium (-Leu/-Trp) and grown overnight at 30°C and 300 rpm. By means of OD_{600} measurements, the selected colonies were plated out at equal cell concentrations (OD_{600} = 0.1 and 0.02 Au) on plates with selective media (SD-LW; -LWA; -LWAH; -LWH; -LWH + 1/3/10 mM 3-AT).

c) FUL-BD results to confirm successful YnH assay

The control screening using *FUL* as bait shows growth of many colonies during the initial screening against the MADS library. The putative positive interactors were compared to an earlier performed screening (de Folter et al., 2005), which showed significant resemblance, as discussed in Supplementary Data III. The potential *FUL* interactors were not confirmed further, but several control colonies were selected to be included with the second screening of *SPL8* and *ARF2* as controls. The controls used are shown in Table 3. These include:

- a *FUL*-BD control in combination with an AD-vector that also shows growth with *SPL8* or *ARF2* **(+/+)**.
- a *FUL*-BD control that grows in combination with an AD-vector, but this AD-vector doesn't show growth in combination with *SPL8* or *ARF2* **(+/-)**.

Table 3: Controls used during the second YnH library screening assay

	SPL8 controls	**ARF2 controls**
(+/+)	FUL ADI-D1	FUL dIII-H4
(+/-)	FUL ADI-B8	FUL ADI-F2
(-/-)	FUL ADI-A1	FUL ADI-A1
(-/+)	SPL8 ADI-B8	ARF2 ADI-B8

- a *FUL*-BD control that doesn't grow in combination with an AD-vector, nor does this AD-vector show growth in combination with *SPL8* or *ARF2* **(-/-)**.
- an *SPL8*-BD or *ARF2*-BD control that doesn't grow in combination with an AD-vector, while that AD-vector does show growth in combination with *FUL*-BD **(-/+)**.

d) library screen assay 2

The second assay showed growth for all colonies on SD-LW. Colonies containing plasmids for Y3H also show growth on SD-LWA, because the pTFT1 vector has adenine as a selection marker for interaction with the AD vector, and is thus already produced in the

25

pTFT1/pDEST22 dimer containing colonies. During the auto-activation test (Figure 24), it was observed that the SPL8-BD construct is a bit leaky in the synthesis of histidine. This explains the slight growth observed during the second screening, on selective medium lacking L/W/H and L/W/A/H for strains containing the SPL8-BD construct in Y2H and Y3H assays, respectively. No growth was observed that could be attributed to interaction of the bait- and prey vectors, thus interaction could not be confirmed for any of the colonies selected after the first screening. The positive FUL-controls show growth on every selective medium, confirming (strong) interaction and indicating that the method was applied successfully. Figure 26 and Figure 27 display the plates of the second assay of ARF2-BD and SPL8-BD, respectively. The results are summarized schematically in Table S10.

Figure 26: Growth of the selected combinations of ARF2-BD with MADS-AD (yellow) and MADS dimers (blue), positive (green), and negative (red) controls on several selective media. Pictures were taken after 3 days of incubation at 20°C.
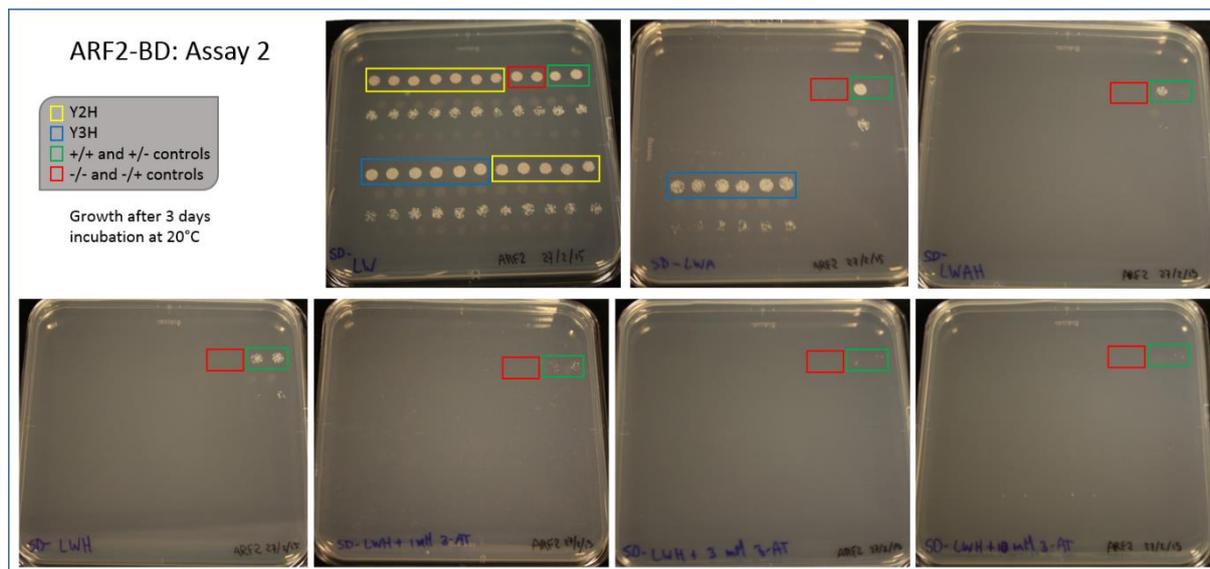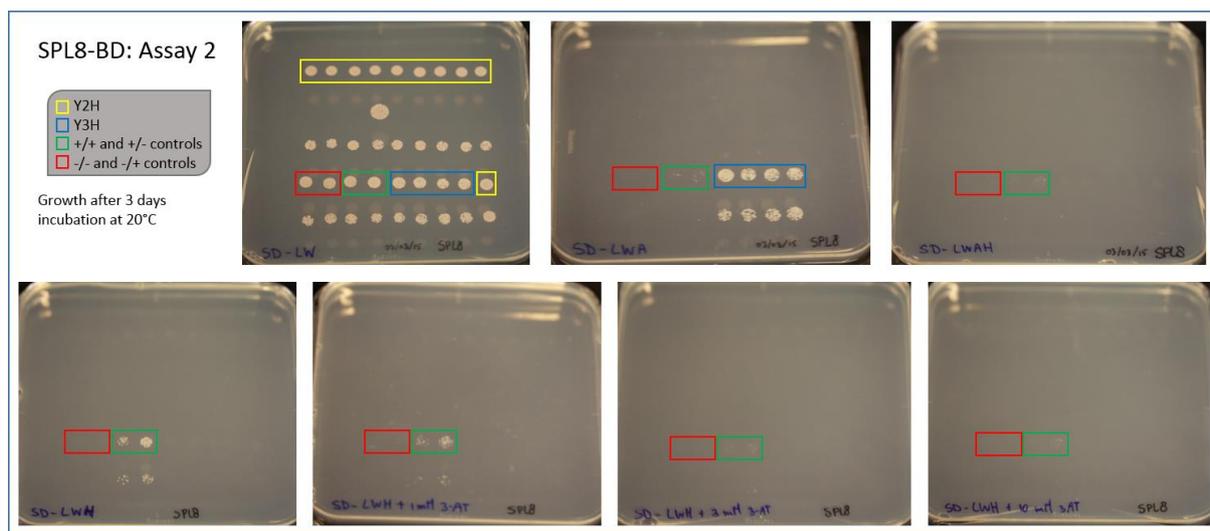

Figure 27: Growth of the selected combinations of SPL8-BD with MADS-AD (yellow) and MADS dimers (blue), positive (green), and negative (red) controls on several selective media. Pictures were taken after 3 days of incubation at 20°C.

4. Discussion, Conclusions, and Recommendations

4.1 Genotyping

In this project, we attempted to generate a new robust *sep1sep2sep3* triple mutant using stable T-DNA insertion alleles for *SEP1* and *SEP3*, and applying the new CRISPR technique to generate a mutant *sep2* allele. However, although genotyping confirmed the presence of a T-DNA insertion in the *sep1* line, this could not be confirmed for the two independent *sep3* T-DNA insertion lines. These lines seem to lack a T-DNA in the *SEP3* gene and are thus probably not correctly annotated. Unfortunately, also the generation of a *sep2* mutant using the CRISPR/Cas9 technique failed, as discussed in the next paragraph.

4.2 CRISPR/Cas9

After analysis of the results obtained during this project, it is clear that multiple problems were encountered during the generation of the CRISPR construct, such as inefficient digestion/ligation, and contamination by another plasmid, probably before or during electroporation. It may be possible that longer linkers are required around the SpeI restriction sites flanking the fragment C construct, in order to allow efficient digestion by the restriction enzyme. Although sequencing indicated that fragment C was not ligated into pYB196, the colony PCR results led us to believe that we selected colonies with an insert of the expected length. This resulted in a significant loss of time, and the project was aborted after several attempts. Outside the context of this project, our group managed to create a CRISPR construct with another guide sequence. However, the process was successful only after optimization of all steps, and still had a very low efficiency (data not shown, Bemer 2015, unpublished).

Although our results indicate low efficiency for the creation of the CRISPR construct, which led us to abandon the technique, literature suggests that after the creation of the construct and floral infiltration, targeted mutagenesis occurs successfully and at high efficiency (Feng et al., 2014). Over the past few years, various research groups have successfully applied CRISPR/Cas9 as a tool for targeted mutagenesis in higher eukaryotes (Jinek et al., 2012; Chang et al., 2013; Hwang et al., 2013; Wang et al., 2013), including plants (Feng et al., 2013; Jiang et al., 2013; Li et al., 2013; Mao et al., 2013; Fauser et al., 2014; Feng et al., 2014; Hyun et al., 2015). Moreover, the mutations generated by CRISPR/Cas9 were shown to be stable and inheritable (Feng et al., 2014; Jiang et al., 2014). Besides that, the option to generate multiple mutations at once provides a huge advantage over classical crossing methods (Belhaj et al., 2015; Hyun et al., 2015). Finally, there seems to be a preference for integration of T-DNA at a DSB during NHEJ (Chilton and Que, 2003), making CRISPR/Cas9 a promising tool for creating insertion mutants as well (Fauser et al., 2014; Belhaj et al., 2015).

In conclusion, it seems that CRISPR is an easy to design, quick, and efficient tool for targeted mutagenesis (Feng et al., 2014; Belhaj et al., 2015). However, the efficiency to create the CRISPR construct and to clone into pYB196 needs to be improved. A new system to create CRISPR constructs, using Golden Gate cloning, is planned to be implemented in the near future (van der Wal, 2015, unpublished data).

4.3 YnH

While some growth was observed in the first assay of SPL8-BD and ARF2-BD with the MADS-AD and MADS-dimer libraries (Figure 25), none of these putative interacting combinations showed growth in a second assay, where the putative positive combinations were spotted at equal cell densities (Figure 26 and Figure 27). Possible explanations for the lack of growth are discussed below.

The results from the *FUL*-BD construct, elaborately discussed in Supplementary Data III, show significant resemblance to those obtained in an earlier screening performed by de Folter et al. (2005). This indicates that the *FUL* controls used in the second assay of this screening are trustworthy. In turn, growth of the control-colonies showed that the second assay was performed correctly (Table S10). However, no control for the actual interaction of the *SPL8*-BD and *ARF2*-BD constructs was included. For example, a combination of *ARF2*-BD with *ARF1*-AD, *Aux/IAA*-AD, or *TOPLESS (TPL)*-AD should indicate whether the *ARF2*-BD construct is functional (Causier et al., 2012; Piya et al., 2014).

If these controls would show interaction with *ARF2* and *SPL8*, it could be concluded that the constructs are functional. In this case the lack of growth could be attributed to a lack of interaction of *ARF2* and *SPL8* with the MADS-box proteins and MADS-box dimers in the libraries. Complete absence of a physical relationship with the higher order complex is considered to be unlikely, as the earlier described IP experiments performed by (Smaczniak et al.) indicated a significant occurrence of these proteins in pull-down samples targeting *AG*, *SEP3,* or *AP1*. Another explanation could be that no protein-protein interaction was determined in these yeast-hybrid assays, due to a lack of other factors. For example, another protein may be required to serve as a bridge between the proteins tested. It is possible that a conformational change, established for example by a post-translational modification (PTM), is required for physical interaction with MADS-box proteins. (Birkenbihl et al.) suggested SBP-domain proteins such as *SPL8* are regulated by PTMs. *SPL8* has two, and *ARF2* has three alternative splicing variants. Furthermore, it has been hypothesized that MADS-box transcription factors bind to the DNA first, after which other factors may be recruited, thereby facilitating transcriptional regulation (Smaczniak et al., 2012a; Pajoro et al., 2014). It may be possible that *SPL8* or *ARF2* are not interacting with MADS-box proteins, prior to formation of a tetramer and/or DNA-loop.

If positive controls like *TPL* would not show growth when co-expressed with *ARF2* and *SPL8*, the lack of reporter gene expression should be attributed to the constructs. If the constructs are not functional because fusion to the binding domain causes a conformational change of the protein, it might be an option to fuse the BD to the other terminal end of the protein, or to perform a reciprocal transfer of proteins (switch AD and BD constructs). *ARF2* was found to be able to function as a negative regulator of transcription (Okushima et al., 2005), whereas *SPL8* was found to be a negative allosteric modulator (Zhang et al., 2007). Both proteins may have an intrinsic suppression domain that could overrule the fused activation domain, resulting in repression of the reporter genes during the YnH assays, thus inhibiting growth on selective medium. Using similar reasoning, bait proteins that have an intrinsic activation domain may result in activation of the reporter gene, also in absence of protein-protein interaction. This will result in the occurrence of false positives. Alternative yeast-hybrid methods of determining physical protein-protein interactions, which do not rely on a transcriptional readout, may be more suitable to investigate interactions of transcription factors. Such methods can include Split-Protein sensors (Johnsson and Varshavsky, 1994),

Intracistronic complementation of β-galactosidase mutants (Rossi et al., 1997), or the Sos Recruitment System (Aronheim et al., 1997).

4.4 In conclusion

In conclusion, the present study highlights the need for a molecular tool to establish stable and inheritable mutations, at selectable locations. Compared to collections of randomly integrated elements (e.g. T-DNA and transposons), and based on the results reported in the literature, the CRISPR/Cas9 system meets all the requirements for an efficient tool for targeted mutagenesis. However, our results indicate that the efficiency in cloning the CRISPR construct into the pYB196 vector is very low, and requires alternative cloning methods (e.g. Golden Gate cloning).

Regarding the yeast-hybrid experiments, our results are inconclusive: during the first assay several combinations of MADS-box proteins with ARF2 and SPL8 showed some growth on selective media, but these results could not be reproduced by the second assay. Without positive controls to verify the functionality of the BD-constructs of ARF2 and SPL8, we can only speculate why interaction wasn't observed. As a recommendation, alternative methods to determine physical protein-protein interactions, which do not rely on a transcriptional readout, may be more suitable to investigate interactions of transcription factors. Overall, the methods employed in the present study were unsuccessful in answering our initial research questions.

REFERENCES

Alejandra Mandel, M., Gustafson-Brown, C., Savidge, B., and Yanofsky, M.F. (1992). Molecular characterization of the Arabidopsis floral homeotic gene APETALA1 **360,** 273-277.

Alvarez-Buylla, E.R., Benítez, M., Corvera-Poiré, A., Chaos Cador, Á., de Folter, S., Gamboa de Buen, A., Garay-Arroyo, A., García-Ponce, B., Jaimes-Miranda, F., Pérez-Ruiz, R.V., Piñeyro-Nelson, A., and Sánchez-Corrales, Y.E. (2010). Flower Development. The Arabidopsis Book**,** e0127.

Angenent, G.C., and Colombo, L. (1996). Molecular control of ovule development. Trends in Plant Science **1,** 228-232.

Aoyama, T., and Chua, N.H. (1997). A glucocorticoid-mediated transcriptional induction system in transgenic plants. The Plant journal : for cell and molecular biology **11,** 605-612.

Aronheim, A. (1997). Improved efficiency sos recruitment system: expression of the mammalian GAP reduces isolation of Ras GTPase false positives. Nucleic Acids Res **25,** 3373-3374.

Aronheim, A., Zandi, E., Hennemann, H., Elledge, S.J., and Karin, M. (1997). Isolation of an AP-1 repressor by a novel method for detecting protein-protein interactions. Mol Cell Biol **17,** 3094-3102.

Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., Romero, D.A., and Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science **315,** 1709-1712.

Belhaj, K., Chaparro-Garcia, A., Kamoun, S., Patron, N.J., and Nekrasov, V. (2015). Editing plant genomes with CRISPR/Cas9. Current Opinion in Biotechnology **32,** 76-84.

Bhaya, D., Davison, M., and Barrangou, R. (2011). CRISPR-Cas systems in bacteria and archaea: versatile small RNAs for adaptive defense and regulation. Annual review of genetics **45,** 273-297.

Birkenbihl, R.P., Jach, G., Saedler, H., and Huijser, P. (2005). Functional dissection of the plant-specific SBP-domain: overlap of the DNA-binding and nuclear localization domains. Journal of molecular biology **352,** 585-596.

Bowman, J.L., Smyth, D.R., and Meyerowitz, E.M. (1989). Genes directing flower development in Arabidopsis. The Plant Cell Online **1,** 37-52.

Bowman, J.L., Smyth, D.R., and Meyerowitz, E.M. (1991). Genetic interactions among floral homeotic genes of Arabidopsis. Development **112,** 1-20.

Casanova-Sáez, R., Candela, H., and Micol, J.L. (2014). Combined haploinsufficiency and purifying selection drive retention of RPL36a paralogs in Arabidopsis **4**.

Causier, B., Ashworth, M., Guo, W., and Davies, B. (2012). The TOPLESS interactome: a framework for gene repression in Arabidopsis. Plant Physiol **158,** 423-438.

Chang, N., Sun, C., Gao, L., Zhu, D., Xu, X., Zhu, X., Xiong, J.W., and Xi, J.J. (2013). Genome editing with RNA-guided Cas9 nuclease in zebrafish embryos. Cell Res **23,** 465-472.

Chilton, M.-D.M., and Que, Q. (2003). Targeted Integration of T-DNA into the Tobacco Genome at Double-Stranded Breaks: New Insights on the Mechanism of T-DNA Integration. Plant Physiology **133,** 956-965.

Cho, S., Jang, S., Chae, S., Chung, K., Moon, Y.-H., An, G., and Jang, S. (1999). Analysis of the C-terminal region of Arabidopsis thaliana APETALA1 as a transcription activation domain. Plant Molecular Biology **40,** 419-429.

Cho, S.W., Kim, S., Kim, J.M., and Kim, J.-S. (2013). Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease **31,** 230-232.

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

Colombo, L., Franken, J., Koetje, E., Van Went, J., Dons, H.J., Angenent, G.C., and Van Tunen, A.J. (1995). The petunia MADS box gene FBP11 determines ovule identity. The Plant cell **7,** 1859-1868.

Davies, B., Egea-Cortines, M., de Andrade Silva, E., Saedler, H., and Sommer, H. (1996). Multiple interactions amongst floral homeotic MADS box proteins. The EMBO Journal **15,** 4330-4343.

De Bodt, S., Raes, J., Van de Peer, Y., and Theißen, G. (2003). And then there were many: MADS goes genomic. Trends in Plant Science **8,** 475-483.

de Folter, S., and Immink, R.G. (2011). Yeast protein-protein interaction assays and screens. Methods in molecular biology (Clifton, N.J.) **754,** 145-165.

de Folter, S., Immink, R.G.H., Kieffer, M., Pařenicová, L., Henz, S.R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M.M., Davies, B., and Angenent, G.C. (2005). Comprehensive Interaction Map of the Arabidopsis MADS Box Transcription Factors. The Plant Cell Online **17,** 1424-1433.

Ditta, G., Pinyopich, A., Robles, P., Pelaz, S., and Yanofsky, M.F. (2004). The SEP4 Gene of Arabidopsis thaliana Functions in Floral Organ and Meristem Identity. Current Biology **14,** 1935-1940.

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.

Fan, H.Y., Hu, Y., Tudor, M., and Ma, H. (1997). Specific interactions between the K domains of AG and AGLs, members of the MADS domain family of DNA binding proteins. The Plant journal : for cell and molecular biology **12,** 999-1010.

Fauser, F., Schiml, S., and Puchta, H. (2014). Both CRISPR/Cas-based nucleases and nickases can be used efficiently for genome engineering in Arabidopsis thaliana. Plant Journal **79,** 348-359.

Feldmann, K.A. (1991). T-DNA insertion mutagenesis in Arabidopsis: mutational spectrum. The Plant Journal **1,** 71-82.

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.L., Wei, P., Cao, F., Zhu, S., Zhang, F., Mao, Y., and Zhu, J.K. (2013). Efficient genome editing in plants using a CRISPR/Cas system. Cell Research **23,** 1229-1232.

Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D.L., Wang, Z., Zhang, Z., Zheng, R., Yang, L., Zeng, L., Liu, X., and Zhu, J.K. (2014). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. Proceedings of the National Academy of Sciences of the United States of America **111,** 4632-4637.

Feschotte, C., Jiang, N., and Wessler, S.R. (2002). Plant transposable elements: where genetics meets genomics **3,** 329-341.

Fields, S., and Song, O. (1989). A novel genetic system to detect protein-protein interactions. Nature **340,** 245-246.

Folter, S.d., and Angenent, G.C. (2006). trans meets cis in MADS science. Trends in Plant Science **11,** 224-231.

Gasiunas, G., Barrangou, R., Horvath, P., and Siksnys, V. (2012). Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. Proc Natl Acad Sci U S A **109,** E2579-2586.

Hayes, T.E., Sengupta, P., and Cochran, B.H. (1988). The human c-fos serum response factor and the yeast factors GRM/PRTF have related DNA-binding specificities. Genes Dev **2,** 1713-1722.

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

Hwang, W.Y., Fu, Y., Reyon, D., Maeder, M.L., Kaini, P., Sander, J.D., Joung, J.K., Peterson, R.T., and Yeh, J.R. (2013). Heritable and precise zebrafish genome editing using a CRISPR-Cas system. PloS one **8,** e68708.

Hyun, Y., Kim, J., Cho, S.W., Choi, Y., Kim, J.S., and Coupland, G. (2015). Site-directed mutagenesis in Arabidopsis thaliana using dividing tissue-targeted RGEN of the CRISPR/Cas system to generate heritable null alleles. Planta **241,** 271-284.

Hyun, Y., Yun, H., Park, K., Ohr, H., Lee, O., Kim, D.-H., Sung, S., and Choi, Y. (2013). The catalytic subunit of Arabidopsis DNA polymerase α ensures stable maintenance of histone modification. Development **140,** 156-166.

Immink, R., Tonaco, I., de Folter, S., Shchennikova, A., van Dijk, A., Busscher-Lange, J., Borst, J., and Angenent, G. (2009). SEPALLATA3: the 'glue' for MADS box transcription factor complex formation. Genome Biology **10,** R24.

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. Proceedings of the National Academy of Sciences of the United States of America **99,** 2416-2421.

Irish, V.F., and Sussex, I.M. (1990). Function of the apetala-1 gene during Arabidopsis floral development. The Plant Cell Online **2,** 741-753.

Jiang, W., Yang, B., and Weeks, D.P. (2014). Efficient CRISPR/Cas9-Mediated Gene Editing in <italic>Arabidopsis thaliana</italic> and Inheritance of Modified Genes in the T2 and T3 Generations. PloS one **9,** e99225.

Jiang, W., Zhou, H., Bi, H., Fromm, M., Yang, B., and Weeks, D.P. (2013). Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucleic Acids Research **41.**

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science **337,** 816-821.

Johnsson, N., and Varshavsky, A. (1994). Split ubiquitin as a sensor of protein interactions in vivo. Proceedings of the National Academy of Sciences of the United States of America **91,** 10340-10344.

Kaufmann, K., Melzer, R., and Theissen, G. (2005). MIKC-type MADS-domain proteins: structural modularity, protein interactions and network evolution in land plants. Gene **347,** 183-198.

Kaufmann, K., Muiño, J.M., Jauregui, R., Airoldi, C.A., Smaczniak, C., Krajewski, P., and Angenent, G.C. (2009). Target Genes of the MADS Transcription Factor SEPALLATA3: Integration of Developmental and Hormonal Pathways in the <italic>Arabidopsis</italic> Flower. PLoS Biol **7,** e1000090.

Komaki, M.K., Okada, K., Nishino, E., and Shimura, Y. (1988). ISOLATION AND CHARACTERIZATION OF NOVEL MUTANTS OF ARABIDOPSIS-THALIANA DEFECTIVE IN FLOWER DEVELOPMENT. Development **104,** 195-&.

Krizek, B.A., and Meyerowitz, E.M. (1996). The Arabidopsis homeotic genes APETALA3 and PISTILLATA are sufficient to provide the B class organ identity function. Development **122,** 11-22.

Kunst, L., Klenz, J.E., Martinez-Zapater, J., and Haughn, G.W. (1989). AP2 Gene Determines the Identity of Perianth Organs in Flowers of Arabidopsis thaliana. The Plant Cell Online **1,** 1195-1208.

Li, J.F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M., and Sheen, J. (2013). Multiplex and homologous recombination-mediated genome editing in Arabidopsis and Nicotiana benthamiana using guide RNA and Cas9. Nature Biotechnology **31,** 688-691.

Lieber, M.R. (2010). The Mechanism of Double-Strand DNA Break Repair by the Nonhomologous DNA End Joining Pathway. Annual review of biochemistry **79,** 181-211.

Ma, H., Yanofsky, M.F., and Meyerowitz, E.M. (1991). AGL1-AGL6, an Arabidopsis gene family with similarity to floral homeotic and transcription factor genes. Genes & Development **5,** 484-495.

Ma, J., and Ptashne, M. (1988). Converting a eukaryotic transcriptional inhibitor into an activator. Cell **55,** 443-446.

Mao, Y., Zhang, H., Xu, N., Zhang, B., Gou, F., and Zhu, J.K. (2013). Application of the CRISPR-Cas system for efficient genome engineering in plants. Molecular Plant **6,** 2008-2011.

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

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

Meyer, V. (1966). Flower abnormalities. The Botanical Review **32,** 165-218.

Meyerowitz, E.M., Bowman, J.L., Brockman, L.L., Drews, G.N., Jack, T., Sieburth, L.E., and Weigel, D. (1991). A genetic and molecular model for flower development in Arabidopsis thaliana. Development **112,** 157-167.

Mizukami, Y., and Ma, H. (1992). Ectopic expression of the floral homeotic gene AGAMOUS in transgenic Arabidopsis plants alters floral organ identity. Cell **71,** 119-131.

Muino, J.M., Smaczniak, C., Angenent, G.C., Kaufmann, K., and van Dijk, A.D. (2014). Structural determinants of DNA recognition by plant MADS-domain transcription factors. Nucleic Acids Res **42,** 2138-2146.

Norman, C., Runswick, M., Pollock, R., and Treisman, R. (1988). Isolation and properties of cDNA clones encoding SRF, a transcription factor that binds to the c-fos serum response element. Cell **55,** 989-1003.

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

Okushima, Y., Mitina, I., Quach, H.L., and Theologis, A. (2005). AUXIN RESPONSE FACTOR 2 (ARF2): a pleiotropic developmental regulator. The Plant journal : for cell and molecular biology **43,** 29-46.

Pajoro, A., Madrigal, P., Muino, J.M., Matus, J.T., Jin, J., Mecchia, M.A., Debernardi, J.M., Palatnik, J.F., Balazadeh, S., Arif, M., O'Maoileidigh, D.S., Wellmer, F., Krajewski, P., Riechmann, J.L., Angenent, G.C., and Kaufmann, K. (2014). Dynamics of chromatin accessibility and gene regulation by MADS-domain transcription factors in flower development. Genome Biol **15,** R41.

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

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

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

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

Pellegrini, L., Tan, S., and Richmond, T.J. (1995). Structure of serum response factor core bound to DNA. Nature **376,** 490-498.

Piya, S., Shrestha, S.K., Binder, B., Stewart, C.N., Jr., and Hewezi, T. (2014). Protein-protein interaction and gene co-expression maps of ARFs and Aux/IAAs in Arabidopsis. Frontiers in plant science **5,** 744.

Puranik, S., Acajjaoui, S., Conn, S., Costa, L., Conn, V., Vial, A., Marcellin, R., Melzer, R., Brown, E., Hart, D., Theißen, G., Silva, C.S., Parcy, F., Dumas, R., Nanao, M., and Zubieta, C. (2014). Structural Basis for the Oligomerization of the MADS Domain Transcription Factor SEPALLATA3 in Arabidopsis. The Plant Cell Online **26,** 3603-3615.

Purugganan, M.D., Rounsley, S.D., Schmidt, R.J., and Yanofsky, M.F. (1995). Molecular evolution of flower development: diversification of the plant MADS-box regulatory gene family. Genetics **140,** 345-356.

Riechmann, J.L., and Ratcliffe, O.J. (2000). A genomic perspective on plant transcription factors. Current Opinion in Plant Biology **3,** 423-434.

Riechmann, J.L., Wang, M., and Meyerowitz, E.M. (1996a). DNA-Binding Properties of Arabidopsis MADS Domain Homeotic Proteins APETALA1, APETALA3, PISTILLATA and AGAMOUS. Nucleic Acids Research **24,** 3134-3141.

Riechmann, J.L., Krizek, B.A., and Meyerowitz, E.M. (1996b). Dimerization specificity of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. Proceedings of the National Academy of Sciences **93,** 4793-4798.

Rossi, F., Charlton, C.A., and Blau, H.M. (1997). Monitoring protein–protein interactions in intact eukaryotic cells by β-galactosidase complementation. Proceedings of the National Academy of Sciences **94,** 8405-8410.

Scherer, S., Mann, C., and Davis, R.W. (1982). Reversion of a promoter deletion in yeast. Nature **298,** 815-819.

Schwarz-Sommer, Z., Huijser, P., Nacken, W., Saedler, H., and Sommer, H. (1990). Genetic Control of Flower Development by Homeotic Genes in Antirrhinum majus. Science **250,** 931-936.

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

Smaczniak, C., Li, N., Boeren, S., America, T., van Dongen, W., Goerdayal, S.S., de Vries, S., Angenent, G.C., and Kaufmann, K. (2012b). Proteomics-based identification of low-abundance signaling and regulatory protein complexes in native plant tissues. Nature protocols **7,** 2144-2158.

Smaczniak, C., Immink, R.G., Muino, J.M., Blanvillain, R., Busscher, M., Busscher-Lange, J., Dinh, Q.D., Liu, S., Westphal, A.H., Boeren, S., Parcy, F., Xu, L., Carles, C.C., Angenent, G.C., and Kaufmann, K. (2012c). Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development. Proc Natl Acad Sci U S A **109,** 1560-1565.

Sternberg, S.H., Redding, S., Jinek, M., Greene, E.C., and Doudna, J.A. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. Nature **507,** 62-67.

The_Arabidopsis_Genome_Initiative. (2000). Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature **408,** 796-815.

Theißen, G. (2001). Development of floral organ identity: stories from the MADS house. Current Opinion in Plant Biology **4,** 75-85.

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

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

Wang, H., Yang, H., Shivalila, C.S., Dawlaty, M.M., Cheng, A.W., Zhang, F., and Jaenisch, R. (2013). One-step generation of mice carrying mutations in multiple genes by CRISPR/Cas-mediated genome engineering. Cell **153,** 910-918.

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

Winter, K.-U., Weiser, C., Kaufmann, K., Bohne, A., Kirchner, C., Kanno, A., Saedler, H., and Theißen, G. (2002). Evolution of Class B Floral Homeotic Proteins: Obligate Heterodimerization Originated from Homodimerization. Molecular Biology and Evolution **19,** 587-596.

Xing, S., Salinas, M., Garcia-Molina, A., Hohmann, S., Berndtgen, R., and Huijser, P. (2013). SPL8 and miR156-targeted SPL genes redundantly regulate Arabidopsis gynoecium differential patterning. The Plant journal : for cell and molecular biology **75,** 566-577.

Yang, Y., and Jack, T. (2004). Defining subdomains of the K domain important for protein–protein interactions of plant MADS proteins. Plant Molecular Biology **55,** 45-59.

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

Yanofsky, M.F., Ma, H., Bowman, J.L., Drews, G.N., Feldmann, K.A., and Meyerowitz, E.M. (1990). The protein encoded by the Arabidopsis homeotic gene agamous resembles transcription factors **346,** 35-39.

Zhang, Y., Schwarz, S., Saedler, H., and Huijser, P. (2007). SPL8, a local regulator in a subset of gibberellin-mediated developmental processes in Arabidopsis. Plant Mol Biol **63,** 429-439.

SUPPLEMENTARY DATA

Table of Contents

Supplementary Data I - Genotyping

This chapter contains an overview of the primers used in the polymerase chain reactions (PCRs) performed to confirm the genotype of the *sepallata* mutant lines used in this project.

a) Overview of primers

Table S1: List of the primer combinations used for PCR during the genotyping of insertion mutant lines assessed in this thesis, with their respective target alleles and product lengths.

Target Line	Target Gene	Target Allele	FW primer	RV primer	Product
S1056	SEP3	wild type	PDS8096	PDS8097	734 bp
S1056	sep3-2	En-1 insertion	PDS8095	PRI776	418 bp
S3843	SEP1	wild type	PDS3909	PDS3980	500 bp
S3843	sep1-1	DuPont insert	PRI768	PRI775	1600 bp
S3843	SEP2	wild type	PRI772	PRI771	797 bp
S3843	sep2-1	En-1 insertion	PRI776	PRI771	400 bp
S6435	SEP1	wild type	PDS8093	PDS8094	428 bp
S6435	sep1	SALK insert, fw orientation	PDS8093	PDS5156	606 bp
S6435	sep1	SALK insert, rv orientation	PDS5156	PDS8094	316 bp
S6436/ S6556	SEP3	wild type	PDS8091	PDS8092	462 bp
S6436	sep3	SALK insert, fw orientation	PDS8091	PDS5156	644 bp
S6436	sep3	SALK insert, rv orientation	PDS5156	PDS8092	410 bp
S6556	sep3	SALK insert, fw orientation	PDS8091	PDS5048	562 bp
S6556	sep3	SALK insert, rv orientation	PDS5048	PDS8092	312 bp

Table S2: List of oligonucleotide sequences used for PCR during the genotyping of insertion mutant lines assessed in this thesis.

Primer	Forward/Reverse	Sequence
PRI768	fw	CACAACCTCCACACACTTCCAAACAC
PRI771	rv	TCCAGCCAGGGATGTAGCCGTTTCCTT
PRI772	fw	AGACTGAGACACATGCATGA
PRI775	rv	GATGCACTCGAATCAGCCAATTTTAGAC
PRI776	fw	GAGCGTCGGTCCCCACACTTCTATAC
PDS3909	fw	CCTGCCAAAGAACTTGAGGTGT
PDS3980	rv	GTACTGTGTCTGCAAACAATCATCAT
PDS5048	rv	AGGCGGTGAAGGGCAATCAGC
PDS5156	rv	ATTTTGCCGATTTCGGAAC
PDS8091	fw	GTTGTACCCAATTCTCTTCTCTTTC
PDS8092	rv	TGGTCAAGCATAAACTGTGTCTG
PDS8093	fw	CCTCCACACACTTCCAAACAC
PDS8094	rv	CACAGAGAACAGACAATTCATAAGC
PDS8095	fw	TGGAGGGTATATAGTTGAGTCTGA
PDS8096	fw	CACACTCTTACGAATCATACGA
PDS8097	rv	ACAAGATCAATAGGCAAGTGACG

Supplementary Data II - CRISPR

This chapter contains the supplementary data of the CRISPR project within this thesis. Part a) displays the potential off-targets identified by the Cas OFFinder tool. Part b) offers an overview of the primers used to generate the CRISPR construct and perform Colony PCR. Furthermore, a complete alignment of the sequencing results using fragment C as template with the theoretical sequence (c), and part of the pYB196 vector sequence surrounding the SpeI site (d), are provided.

a) Cas OFFinder results for guide sequence 1

Table S3: Overview of potential CRISPR mismatch targets of guide sequence 1 targeting *SEP2*, generated by the Cas OFFinder tool.

	target sequence	OFF-target chr	OFF-target position	OFF-target sequence	Strand	# of mismatches
SEP2; gs-1	GACCAGCTCTCTGATCTTCANNN	**chr3**	465251	GACCAGCTCTCTGATCTTCAAGG	Reverse (-)	0
SEP2; gs-1	GACCAGCTCTCTGATCTTCANNN	**chr3**	9998739	tctCAGCTCTCTGATCTTCACGG	Forward (+)	3
SEP2; gs-1	GACCAGCTCTCTGATCTTCANNN	**chr4**	505401	aACaAGCTCTaTGATCTTaAAGG	Reverse (-)	4
SEP2; gs-1	GACCAGCTCTCTGATCTTCANNN	**chr3**	6199759	GAtgcGCTCTCTGATaTTCATGG	Forward (+)	4
SEP2; gs-1	GACCAGCTCTCTGATCTTCANNN	**chr3**	8037395	agCCAGgTCTCTGATCTTCcGGG	Forward (+)	4
SEP2; gs-1	GACCAGCTCTCTGATCTTCANNN	**chr3**	8901504	GACCtGCTCagaGATCTTCACGG	Reverse (-)	4
SEP2; gs-1	GACCAGCTCTCTGATCTTCANNN	**chr2**	15653828	GAtaAGCTCTCTGAgCTgCAAGG	Forward (+)	4
SEP2; gs-1	GACCAGCTCTCTGATCTTCANNN	**chr3**	19829019	GACaAaCTCTCTGAgCTgCAAGG	Forward (+)	4
SEP2; gs-1	GACCAGCTCTCTGATCTTCANNN	**chr5**	22270510	GACtAGCTCTCTGAaCTgCcCGG	Reverse (-)	4
SEP2; gs-1	GACCAGCTCTCTGATCTTCANNN	**chr5**	25950392	GACCAGgTCTCTGtaCTTtAAGG	Reverse (-)	4
SEP2; gs-1	GACCAGCTCTCTGATCTTCANNN	**chr1**	26623763	tAaCtGCTCTgTGATCTTCAAGG	Forward (+)	4
SEP2; gs-1	GACCAGCTCTCTGATCTTCANNN	**chr1**	26910005	GAgCAtCTCTCTGgTCaTCACGG	Reverse (-)	4

b) Overview of primers

Table S4: Overview of the primers used to create the CRISPR construct, and for Colony PCR to determine the fragment length of inserts at the SpeI site of pYB196

FW primer	RV primer	Product	Template	PCR	Product
PDS8064	PDS8068	Fragment A + guide sequence 1	*U6p* (Col-0 gDNA)	primer-mediated extension	285 bp
PDS8070	PDS8066	Fragment B + guide sequence 1	sgRNA pRG_ext_CCR5	primer-mediated extension	134 bp
	PDS8320				
PDS8064	PDS8066	Fragment C + guide sequence 1	Fr. A + Fr. B	overlapping PCR	399 bp
	PDS8320				
PDS8116	PDS8117	insertion at SpeI	pYB196	Colony PCR	686 bp

Table S5: List of oligonucleotide sequences used for generation and confirmation of the CRISPR construct

Primer #	Primer name	FW/ RV	Sequence
PDS8064	SpeI:U6p	forward	ATATACTAGTCTAGAGAATGATTAGGCATCGAACCT
PDS8068	U6p:SEP2 gs-1	reverse	TGAAGATCAGAGAGCTGGTCACAATCACTACTTCGACTCTAGCT
PDS8070	SEP2 gs-1:sgRNA	forward	GACCAGCTCTCTGATCTTCAGTTTTAGAGCTAGAAATAGCAA
PDS8066	sgRNA:SpeI (pICH)*	reverse	GTACAAGAAAGCTGGGTCACTAGTTGGT
PDS8320	sgRNA:SpeI (pRG)	reverse	ATGCAGGAAGACAACTAGTCAA
PDS8116	Colony PCR (SpeI)	forward	GCAAGGCGATTAAGTTGGGTAACG
PDS8117		reverse	TTGTTCCCGCTCCAACCGTCAT

*Initially, the plasmid pICH86966, containing the chimeric sgRNA backbone, was used as a template to generate fragment B with primers PDS8070 and PDS8066. After several failed attempts to obtain sequenced confirmation of fragment C in pYB196, the plasmid pRG_ext_CCR5 was provided by Youbong Hyun, containing the same chimeric sgRNA. Another two attempts using this plasmid as a source for sgRNA, using a newly designed primer (PDS8320), did not result in successful generation of a CRISPR construct either.

c) Complete alignment of fragment C

Figure S1 below displays the complete alignment of sequencing result on fragment C. Sequencing of both the 5' and 3' terminal regions of the fragment lead to less significant results, therefore these regions are marked in red text. More thorough observation of the individual absorbance peaks clarifies many of the differences found between sequencing results and the expected sequence. The deviations, which can be attributed to limitations of the automated sequence determination software, are highlighted in grey. The sequence quality of the region of interest, guide sequence 1, highlighted in yellow, is high enough to be reliable. Figure S2 displays the absorbance peaks of this region when sequenced using PDS8064.

Figure S1: Complete alignment of sequencing results with theoretical fragment C sequence. Highlighted sequences: Light green: confirmed frag C sequence; Dark green: SpeI sites; Yellow: guide sequence; Red: unconfirmed nucleotides. The nucleotides marked with red font represent peaks of lower quality, and the ones highlighted in grey represent deviations observed due to the limitations of the automated sequence determination software, as explained above.

Figure S2: Sufficient sequencing quality at the location of the guide sequence (underlined) indicates that the sequence generated by Macrogen is reliable.

d) pYB196 SpeI insert sequence

…GCCTCTTCGCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGGCGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATACGACTCACTATAGGGCGAATTGGGTACCGGATCCACTAGTCAACGCTGCCAGATTCGATGTAAGTTACTGGAAACTGAACCGAAGGGAACTCTCCAGGAAACAGAGATGGAAACTGAAACTAACTCTGTTTCCAAACGTTAGGGTTCCTTGAAATTGCCTACCAATTGTTCCCGCTCCAACCGTCAT…

Figure S3: Sequence of the empty pYB196 vector with primer binding sites (indicated with cyan and magenta) and the SpeI RE site (indicated with yellow)

Supplementary Data III - YnH

In order to create ARF2-BD and SPL8-BD constructs, Gateway Cloning technology was applied. PCR1 was performed on cDNA, reverse-transcribed from RNA isolates from cauline leaf. By means of primer-mediated extension, part of AttB sites were attached to the terminal ends of the coding sequences of *ARF2* and *SPL8*. The remaining parts of the AttB sites were attached by another primer-mediated extension during PCR2. After purification of the PCR product by PEG precipitation, a BP and LR reaction were performed to clone the construct into pDONR211 and pDEST32, respectively, and each was followed by transformation into *E.coli* DH5α for amplification of the vector containing the BD-construct. Colony PCR was performed to select for an insertion of the right length. The constructs in pDEST32 were sequenced to confirm that the *ARF2* and *SPL8* sequences are in frame with the binding domain.

The protocols for the PCRs and Gateway BP and LR reactions are described in section IV(c) (Gateway Cloning). For the experimental procedures of this project that employ yeast as a host, the media and protocols used were applied (without deviations) as described by de Folter and Immink (2011). This chapter contains an overview of the primers used for PCR, and the interaction matrix containing the FUL-BD construct, as published by de Folter et al. (2005). In the last part, the results of the FUL-BD, used as positive control during these yeast-hybrid assays, are discussed.

a) Overview of primers

Table S6: Overview of the primers used to create the ARF2-BD and SPL8-BD constructs, and for Colony PCR to confirm insertion of the correct fragment length in pDONR211 and pDEST32

FW primer	RV primer	Product	Template	PCR	Product
PDS8308	PDS8309	*ARF2* CDS + AttB sites	Col-0 cDNA (cauline leaf)	PCR1	2586 bp
PDS 8314	PDS 8315	*SPL8* CDS + AttB sites	Col-0 cDNA (cauline leaf)	PCR1	1001 bp
PDS7387	PDS7388	BD-construct	pDEST32	PCR2; Colony PCR	CDS+58 bp

Table S7: List of oligonucleotide sequences used for generation and confirmation of the CRISPR construct

Primer #	Primer name	FW/RV	Sequence
PDS8314	SPL8 fw	forward	*AAAAAGCAGGCTCAATGTTGGACTACGAATGGGA*
PDS8315	SPL8 rv	reverse	*AGAAAGCTGGGTCTATCCGCTGGAGAAAAACA*
PDS8308	ARF2 fw	forward	AAAAAGCAGGCTTGGCGAGTTCGGAGGTTTC
PDS8309	ARF2 rv	reverse	AGAAAGCTGGGTtgtttgtTTAAGAGTTCCC
PDS7387	AttB forward primer	forward	GGGGACAAGTTTGTACAAAAAAGCAGGCT
PDS7388	AttB reverse primer	reverse	GGGGACCACTTTGTACAAGAAAGCTGGGT

b) Interaction matrix

Figure S4 below displays an interactome map of the *Arabidopsis* MADS box transcription factors, as published by de Folter et al. (2005). Proteins that did not interact in their screen were omitted from this figure. Protein-protein interactions with FUL-BD, according to their experiment, occurs with the following proteins: AG; AGL6; AGL14; AGL21; AGL24; SEP1; SEP3; SEP4-II; and SOC1.
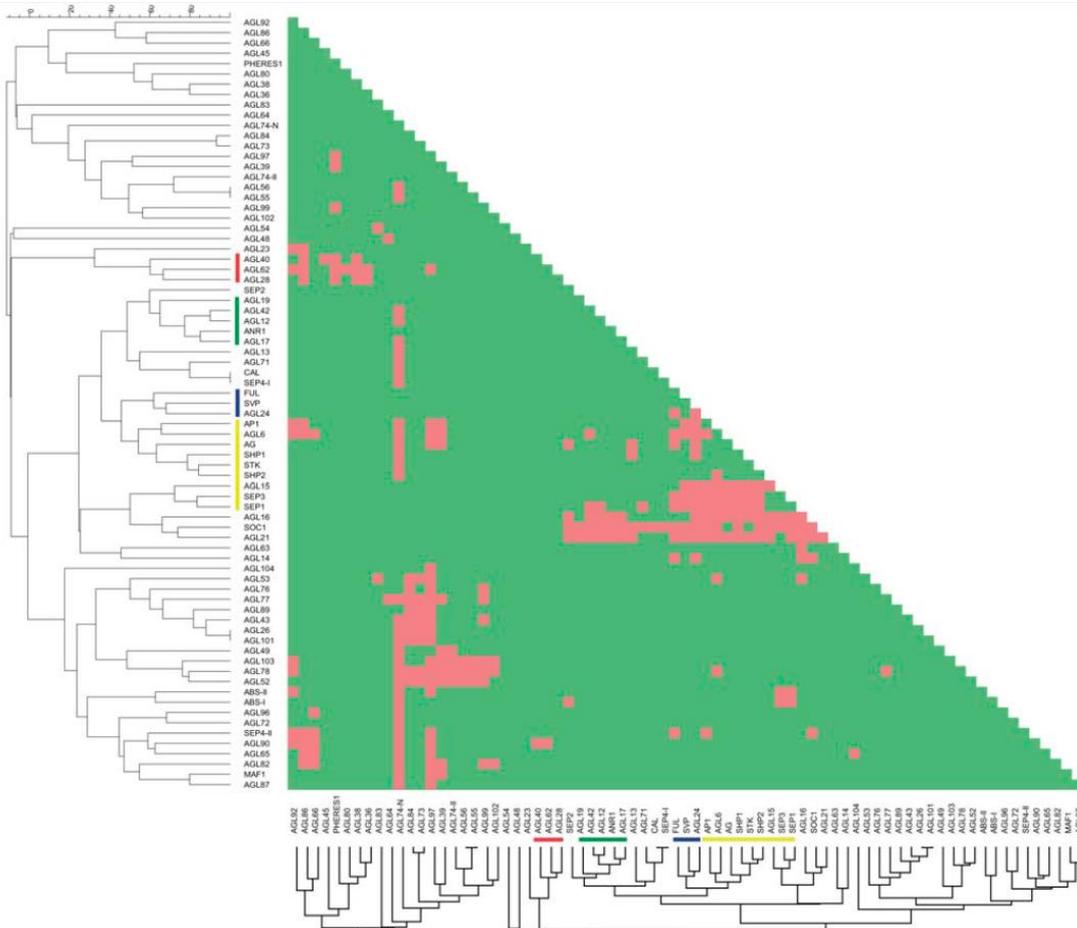
Figure S4: Interactome Map of the *Arabidopsis* MADS Box Transcription Factor Family. Proteins are organized based on hierarchical clustering of their protein-protein interaction patterns. Proteins that do not interact were omitted from this figure. Protein-protein interactions are indicated with red blocks, and no interactions with green blocks. Presence of clustered proteins with a putative similar function is indicated with a coloured bar on the left and bottom side of the figure: red for embryo; green for root; blue for flowering; and yellow for floral organs (from de Folter et al., 2005).

c) *FUL*-BD

The putative protein-protein interactions of the *Arabidopsis* MADS-box transcription factor family with FUL-BD, according to the library screening performed in this thesis, include the following proteins: AGL6; AGL21; AGL24; SEP1; SEP3; and SOC1. Slight growth is observed for combinations of FUL-BD with AGL14, and with SEP4-II. All proteins highlighted by de Folter et al. (2005) as putative interactors of FUL were found as positive interactors in our screening, except for AG. Based on the parity with the assay performed by de Folter et al. (2005), it was concluded that the present assay was performed correctly.

Some variants of positive interactors that did not show growth during the screening by de Folter et al. (2005), did show growth in our assay: SEP3ΔC; SEP4-I; and SOC1β. Additionally, several other putative interactors were identified, that showed strong growth (AGL15; AGL26; AP1; FUL; SPH1; SVP1; and SVP2) or slight growth (AGL13; AGL17; AGL28; FLM; and SEP2). A comparison of the results is displayed in Table S8.

Table S8: An overview of the putative interactors of FUL-BD, as determined by de Folter et al. (2005) (left column), and during this screening (middle and right column). An estimation of the degree of growth on SD-LWA and/or SD-LWH+1mM 3-AT is presented as +/++/+++. Proteins in parentheses showed (strong) growth on only one of the selective media.

FUL-BD Y2H putative protein-protein interactions				
Results produced by de Folter et al. (2005)	This YnH project			
	Corresponding &		Additional results	
AG			AGL15	++
AGL6	AGL6	+++	AGL26	+++
AGL14	(AGL14)	++	(AGL13)	+
AGL21	AGL21	++	(AGL17)	++
AGL24	AGL24	+++	(AGL28)	++
SEP1	SEP1	+++	AP1	+++
SEP3	SEP3	+++	(FLM)	++
	SEP3ΔC	+++	FUL	+++
SEP4-II	(SEP4-II)	++	(SEP2)	++
	SEP4-I	++	SPH1	+++
SOC1	SOC1	+++	SVP1	+++
	SOC1β	+++	SVP2	++

Among the combinations of proteins in the Y3H assay that show growth with FUL-BD, most of the AD vectors encode a transcription factor that interacts directly with FUL-BD, as was observed in the Y2H screens. Presence of another gene encoding a MADS-box transcription factor in the TFT1 vector may not be of influence in these assays. To confirm these results, a reciprocal transfer of proteins should be performed, in order to determine whether growth occurs when the second protein is connected to the activation domain while the first protein is encoded by the TFT1 vector (Table S9).

Table S9: Dimer library strains that showed growth in combination with FUL-BD. Proteins highlighted in purple also interact directly with FUL-BD in both Y2H assays. Orange highlights proteins that showed direct interaction with FUL-BD during the Y2H assay of this thesis, while these were not identified by de Folter et al. (2005). The blue highlight indicates AGAMOUS, which previously has been shown to interact with FUL-BD, although not in our assay.

FUL-BD Y3H putative protein-protein interactions									
	Protein1 (AD)	Protein2 (TFT1)			Protein1 (AD)	Protein2 (TFT1)			Protein1 (AD)	Protein2 (TFT1)
#1	AGL15	AGL16		#22	SOC1	SOC1		#43	AGL97	AGL89
#2	AGL14	AGL16		#23	AGL21	AGL24		#44	SEP3	SHP2
#3	SOC1	AGL19		#24	AGL21	SOC1		#45	ABS-II	SEP3
#4	SOC1	AGL17		#25	AGL21	AGL19		#46	SEP3	ABS-I
#5	SOC1	AGL16		#26	AGL21	AGL13		#47	SEP3	AGL24
#6	SOC1	AGL14		#27	AGL21	STK		#48	SEP3	SVP
#7	SOC1	AGL13		#28	AGL21	SHP1		#49	SEP3	SOC1
#8	SOC1	AGL12		#29	AGL24	SHP1		#50	SEP3	AGL16
#9	SOC1	CAL		#30	AGL21	AGL6		#51	SEP3	STK
#10	SOC1	SHP1		#31	SEP1	SOC1		#52	SEP3	SHP1
#11	SOC1	FUL		#32	SEP1	STK		#53	AG	FUL
#12	SOC1	AGL71		#33	AGL24	AGL16		#54	SEP3	AG
#13	SOC1	AP1		#34	SEP1	AG		#55	SEP3	FUL
#14	SOC1	AGL6		#35	SEP1	FUL		#56	SEP3	AP1
#15	SOC1	SHP2		#36	SEP1	AGL71		#57	SEP3	AGL6
#16	SOC1	ANR1		#37	SEP1	AP1		#58	ABS-II	SEP3ΔC2
#17	SOC1	AGL42		#38	SEP1	AGL6		#59	AGL24	SEP3ΔC2
#18	SOC1	SEP2		#39	SEP1	SHP2		#60	SOC1	SEP3ΔC2
#19	SOC1	SEP4-II		#40	AGL90	AGL40		#61	STK	SEP3ΔC2
#20	SOC1	SEP4-I		#41	FUL	SEP4-II		#62	SHP2	SEP3ΔC2
#21	SOC1	SVP		#42	AGL97	SEP4-II				

After the first assay, colonies containing the putative interacting combinations of ARF2 and SPL8 with MADS box transcription factor were subjected to a second assay, in which the cell density was equalized and growth on several selective media was examined. Although none of the colonies containing an ARF2-BD or SPL8-BD construct showed growth, the FUL-BD constructs used as controls do show growth on several media, indicating that the method was applied successfully. The results are summarized schematically in Table S10. (Note that the results are grouped, e.g. 'ARF2' includes all combinations of ARF2-BD with the MADS-AD vectors)

Table S10: Schematic representation of the growth observed in the second YnH screening. Green indicates strong growth; light green represents slight growth; red shows which colonies did not grow. (L=Leucine; W=Tryptophan; A=Adenine; H=Histidine; 3-AT= **3-A**mino-1,2,4-**T**riazole)

ASSAY	BD-vector	SD-LW	SD-LWA	SD-LWH	SD-LWAH	SD-LWH+ 3-AT
Y2H -BD (Leu) -AD (Trp)	*ARF2*	green	red	red	red	red
	SPL8	green	red	light green	red	red
	FUL **(+ve)**	green	light green	green	light green	green
	FUL **(-ve)**	green	red	red	red	red
Y3H -BD (Leu) -AD (Trp) -TFT1 (Ade)	*ARF2*	green	green	red	red	red
	SPL8	green	green	light green	light green	red
	FUL **(+ve)**	green	green	green	green	green

Supplementary Data IV – Protocols

This chapter contains several protocols used during this thesis.

a) Genotyping protocols

1. DNA extraction

For extraction of genomic DNA from plant tissue (leaves)

1. Fill eppendorf tubes with 5 tungsten beads each.
2. Put one or two small leave(s) in every tube.
3. Place in liquid nitrogen.
4. Grind the tissue in Capmix$^®$ for 10 seconds.
5. Add 400 µl extraction buffer at 60°C to every sample.
6. Invert the tubes a couple of times.
7. Centrifuge for 10' at 14.000 rpm
8. Transfer 200 µl of the supernatant to a new tube.
9. Add 200 µl cold isopropanol.
10. Invert the tubes a couple of times.
11. Centrifuge for 20' at 14.000 rpm, 4°C
12. Wash with 70% EtOH
13. Dry 5' in the vacuum desiccator.
14. Dissolve in 400 µl dH$_2$O

Extraction buffer (100 ml, prewarm at 60°C)

- 20 ml 1M Tris, pH 7.5
- 5 ml 5M NaCl
- 0.5ml 0.5M EDTA, pH 8.0
- 5 ml 10% SDS

2. Seed sterilization

1. Transfer an aliquot of the seeds to a new tube.
2. Add 0.5 ml 2% bleach.
3. Incubate for 10'
4. Remove bleach using a syringe.
5. Add 0.5 ml sterile H2O
6. Mix, incubate for 1'
7. Remove H2O using a syringe, repeat step 5-7
8. Add 0.5 ml 0.1% Agar
9. Store until use (4°C)

3. ½MS growth medium

- ½ MS: 2.2 g/l (0.44 g) (including vitamins)
- MES: 0.5 g/l (0.1 g)
- Set pH to 5.8 using 1M KOH
- Daishin Agar: 9 g/l (1.8 g)
- Add dH$_2$O to 200 ml
- Autoclave 20' at 121°C
- Cool down to ~60°C; pour 4 plates (4x25ml)
- Add 30 µl Kanamycin; pour 4 plates (4x25ml)

The plates were stored at 4°C

b) CRISPR construction protocol

In this section, the CRISPR construction protocol is displayed, as performed to generate constructs targeting the *SEP2* gene.

1. Miniprep (plasmid isolation)

> ➢ Perform Miniprep according to protocol (Nucleospin® plasmid Quickpure Kit by Macherey_Nagel, REF740615.250)

Vector pYB196 (**ICU2p:Cas9**:NLS:HA:NosT)
Template vector pICH86966 (chimeric **sgRNA** backbone)

> ➢ Measure nucleic acid concentration using NanoDrop®

2. PCR (primer-mediated extension with guide sequence)

Fragment A.   *Promoter U6* (PCR on *Arabidopsis thaliana Col-0* gDNA)

a) Guide 1:   SpeI For U6p
                SEP2 Guide1 Rev
b) Guide 2:   SacI For U6p
                SEP2 Guide2 Rev

<div align="center">

Fragment A = *REs:U6p:guide*

</div>

Fragment B.   *sgRNA* (PCR on pICH86966) (10.000x diluted)

a) Guide 1:   sgRNA SEP2 Guide1 For
                SpeI sgRNA pICH Rev
b) Guide 2:   sgRNA SEP2 Guide2 For
                SacI sgRNA pICH Rev

<div align="center">

Fragment B = *guide:sgRNA;REs*

</div>

V (µl)	PCR mix	time	temp. (°C)		
2	template DNA	1'00"	98		
10	5x Q5 buffer	0'10"	98	}	5 cycles
1	10mM dNTPs	0'20"	58		
1.5	FW primer	0'30"	72		
1.5	RV primer	0'10"	98	}	30 cycles
33.5	dH2O	0'40"	72		
0.5	Q5 polymerase	1'00"	72		
50		∞	12		

> ➢ **5 µl**: Perform gel electrophoresis on a 2% Agarose gel

3. Overlapping PCR (U6p::sgRNA generation)

➢ Dilute both samples 100x.

V (µl)	PCR mix	time	temp. (°C)		
1	U6p:guide1 (100x diluted)				
1	sgRNA:guide1 (100x diluted)	1'00"	98		
10	5x Q5 buffer	0'10"	98	}	5 cycles
1	10mM dNTPs	0'20"	55		
1.5	FW primer (SpeI For U6p)	0'30"	72		
1.5	RV primer (SpeI sgRNA pICH Rev)	0'10"	98	}	15 cycles
33.5	dH2O	1'00"	72		
0.5	Q5 polymerase	2'00"	72		
50		∞	12		

➢ Perform gel electrophoresis on a 1.5% Agarose gel using 5 µl of the sample

Fragment C = *REs:U6p:guide:sgRNA:REs*

4. RE digestion

a) Fragment C (Overlapped PCR fragments for Guide1)

➢ Purify the remaining 45 µl on a column, elute in 50 µl
➢ Digest the total sample:
  ○ 43 µl purified sample
  ○ 5 µl 10x SpeI buffer
  ○ 2 µl SpeI

b) pYB196 vector

  ○ ~6 µl pYB196 vector**(~1 µg)**
  ○ 10 µl 10x SpeI buffer
  ○ 4 µl SpeI
  ○ 2 µl CIP
  ○ 78 µl dH$_2$O

➢ Incubate at 37°C for 3 hours
➢ Heat-inactivate enzymes at 80°C for 20'00"
➢ Purify digestion of (only!) **Fragment C**
➢ Confirm digestion on a 1% agarose gel (use 5 µl of each sample)

5. Ligation

➢ Ligate Fragment C into pYB196:
  ○ 1 µl 10x T4 buffer
  ○ 5 µl digested pYB196 vector (linearized & dephosporylated)
  ○ 2 µl digested Fragment C (Guide1 overlap product)
  ○ 1.5 µl dH$_2$O
  ○ 0.5 µl T4 DNA ligase
➢ Incubate O/N at 16°C

6. Transformation into *E.coli* DH5α (electro-competent)

➢ Add 1 µl DNA to 50 µl electro-competent cells
➢ Transfer to clean cuvette (keep on ice)
➢ Press "Raise" and "Lower" buttons simultaneously
➢ Adjust to 1.7 KV using "Raise" and "Lower" buttons

- ➢ Insert cuvette; slide back to close
- ➢ Press pulse (~2 seconds)
- ➢ Add 1 ml LB medium
- ➢ Incubate for 1 hour at 37°C
- ➢ Plate out on LB agar (+ 100 mg/L Kanamycin) plates (100 µl; 300 µl; 600 µl)
- ➢ Incubate O/N at 37°C

7. Colony PCR on *E.coli* DH5α + pYB196

- ➢ Add 10 µl H_2O to a pcr tube
- ➢ Pick a colony from the plate using a toothpick and suspend in the pcr tube.
- ➢ Use 1 µl as template for PCR
- ➢ Add 100 µl LB (+ Kan) to the remainder and incubate O/N at 37°C

V (µl)	PCR mix	time	temp. (°C)		
1	template DNA	3'00"	95		
3	10x Taq buffer	0'30"	95		
0.6	10mM dNTPs	0'30"	58	}	35 cycles
0.6	FW primer	1'00"	72		
0.6	RV primer	3'00"	72		
24	dH2O	∞	12		
0.2	MM Taq polymerase				
30					

- ➢ Perform gel electrophoresis on a 1% Agarose gel
- ➢ Add 50 µl of the colonies with pYB196 + CRISPR construct to 2 ml LB (+ Kan) in a culture tube and incubate O/N at 37°C

8. Miniprep (plasmid isolation)

Vector pYB196 + U6p:gs-1:sgRNA

- ➢ Perform Miniprep according to protocol (Nucleospin® plasmid Quickpure Kit by Macherey_Nagel, REF740615.250)
- ➢ Measure nucleic acid concentration using NanoDrop®
- ➢ Prepare mixture for Sequence reaction:
  - ○ 5 µl plasmid (100 ng/µl)
  - ○ 5 µl primer (5 pmol) (**OR** *2.5µl * primer(10µM) + 2.5µl H_2O*)

c) Gateway cloning protocol using pDONR & pDEST vectors

1. PCR1: Amplify fragment from gDNA or cDNA and attach part of the AttB sites

V (µl)	PCR mix	time	temp. (°C)		
1	template DNA (~20ng)	1'00"	98		
6	5x Q5 buffer	0'10"	98		
0.6	10mM dNTPs	0'15"	60	}	5 cycles
0.6	FW primer	1'00"*	72		
0.6	RV primer	0'10"	98		
20.9	dH2O	0'15"	65	}	30 cycles
0.3	Q5 polymerase	1'00"*	72		
30		5'00"	72		
	depending on fragment length	∞	12		

1. Use 10 µl for agarose gel electrophoresis.
   If a single product of the expected length is created, proceed with PCR2. If multiple products are visible, repeat PCR1 with increased annealing temperatures or perform a gel extraction of the right band.

2. PCR2: Add remaining part of the AttB sites
2. Dilute the PCR1 product 50x and use 1µl for PCR2

V (µl)	PCR mix	time	temp. (°C)		
1	template DNA (~20ng)	1'00"	98		
10	5x Q5 buffer	0'10"	98		
1	10mM dNTPs	0'15"	48	}	5 cycles
1	FW primer PDS 7387	1'00"*	72		
1	RV primer PDS 7388	0'10"	98		
35.5	dH2O	0'15"	58	}	15 cycles
0.5	Q5 polymerase	1'00"*	72		
50		5'00"	72		
	depending on fragment length	∞	12		

3. PEG precipitation
Use 25 µl of the PCR2 product to perform a PEG precipitation:

- 25 µl PCR product
- 75 µl TE
- 50 µl PEG/MgCl$_2$ solution*
  in the Gateway freezer box; cut the end of a yellow tip for easy pipetting

3. Vortex vigorously
4. Centrifuge 15' at max speed, RT
5. Immediately pipet off the supernatant carefully
6. Dissolve in 15 µl TE
7. Use 5 µl of the PCR2 product and 5 µl of the PEG product for agarose gel electrophoresis.
   If a clear, specific fragment is created, proceed with the BP reaction.
8. Estimate the DNA concentration of the PEG product visually on gel.
9. Use approximately 50 ng of product in the BP reaction

4. Gateway Cloning BP reaction

- X µl PEG purified product (~50 ng)
- 1 µl pDONR vector (100 ng/µl)
- X µl TE (total reaction volume should be 5 µl)
- 1 µl BP enzyme mix *(Gateway box)*

10. Incubate > 3 hours at RT
11. Add 0.5 µl proteinase K to stop the reaction *(Gateway box)*
12. Incubate 10' at 37°C
13. Keep on ice until transformation to *E.coli* DH5α

5. Transformation to E.coli DH5α

14. Add 1 µl DNA to 50 µl electro-competent cells
15. Transfer to clean cuvette (keep on ice)
16. Press "Raise" and "Lower" buttons simultaneously
17. Adjust to 1.7 KV using "Raise" and "Lower" buttons
18. Insert cuvette; slide back to close
19. Press pulse (~2 seconds)
20. Add 1 ml LB medium
21. Incubate for 1 hour at 37°C
22. Plate out on LB agar *(+ antibiotic*)*
23. Incubate O/N at 37°C
    *pDONR221 has KanR; pDONR207 has GentaR

6. Colony PCR on E.coli DH5α + pDONR

24. Add 10 µl H2O to a pcr tube
25. Pick a colony from the plate using a pipet or toothpick and suspend in the pcr tube.
26. Use 1 µl as template for PCR

V (µl)	PCR mix	time	temp. (°C)		
1	template DNA	3'00"	95		
3	10x Taq buffer	0'30"	95		
0.6	10mM dNTPs	0'30"	58	}	35 cycles
0.6	FW primer	1'00"*	72		
0.6	RV primer	3'00"	72		
24	dH2O	∞	12		
0.2	MM Taq polymerase				
30		** depending on fragment length*			

27. Perform gel electrophoresis on a 1% agarose gel
28. Add 50 µl of the colonies with the correct construct to 2 ml LB *(+ antibiotic)* in a culture tube and incubate O/N at 37°C

7. Miniprep

29. Perform Miniprep according to protocol (Nucleospin® plasmid Quickpure Kit by Macherey_Nagel, REF740615.250)
30. Measure nucleic acid concentration using NanoDrop®

8. Gateway Cloning LR reaction

- X µl pDONR + fragment (~50 ng)
- 1 µl pDEST vector (100 ng/µl)
- X µl TE (total reaction volume should be 5 µl)
- 1 µl LR enzyme mix *(Gateway box)*

31. Incubate > 3 hours at RT
32. Add 0.5 µl proteinase K to stop the reaction *(Gateway box)*
33. Incubate 10' at 37°C
34. Keep on ice until transformation to *E.coli* DH5α

5. Transformation to E.coli DH5α

9. Prepare sequence reaction mix

35. Culture *E.coli DH5α* with pDEST + fragment in LB *(+ antibiotic*)* O/N at 37°C
    *pDEST32 has GentaR
36. Perform Miniprep according to protocol (Nucleospin® plasmid Quickpure Kit by Macherey_Nagel, REF740615.250)
37. Measure nucleic acid concentration using NanoDrop®
38. Prepare mixture for Sequence reaction:
    - 5 µl plasmid (100 ng/µl)
    - 5 µl primer (5 pmol) *(**OR** 2.5µl * primer(10µM) + 2.5µl H_2O)*

Upon confirmation of the correct sequence, in frame with the BD sequence encoded on pDEST32, the vector is transformed to the yeast strain *S. cerevisiae PJ69-4α*. This part, and the following auto-activation test and library screen assays, were previously described by de Folter and Immink (2011): "Yeast Protein-Protein Interaction Assays and Screens."

REFERENCES

de Folter, S., and Immink, R.G. (2011). Yeast protein-protein interaction assays and screens. Methods in molecular biology (Clifton, N.J.) **754,** 145-165.

de Folter, S., Immink, R.G.H., Kieffer, M., Pařenicová, L., Henz, S.R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M.M., Davies, B., and Angenent, G.C. (2005). Comprehensive Interaction Map of the Arabidopsis MADS Box Transcription Factors. The Plant Cell Online **17,** 1424-1433.