

**Molecular and Genetic Analysis of
the effects of SUMOylation on the
regulation of floral transition in
*Arabidopsis***

Inaugural-Dissertation

zur

Erlangung des Doktorgrades

der Mathematisch-Naturwissenschaftlichen Fakultät

der Universität zu Köln

vorgelegt von

Mitzi Villajuana Bonequi

aus Mexiko Stadt

Köln, November 2011



Max Planck Institute for
Plant Breeding Research

Die vorliegende Arbeit wurde am Max-Planck-Institut für Züchtungsforschung in Köln, in der Abteilung für Entwicklungsbiologie der Pflanzen (Direktor Prof. Dr. George Coupland) angefertigt.

Berichterstatter: Prof. Dr. George Coupland
Prof. Dr. Ute Höcker
Prüfungsvorsitzender: Prof. Dr. Martin Hülskamp

Tag der Disputation: 8 November 2011

I ching 52. Keeping still

First line

Keeping the toes still means halting before one has even begun to move. The beginning is the time of few mistakes. At that time one is still in harmony with primal innocence. Not yet influenced by obscuring interests and desires, one sees things intuitively as they really are. A man who halts at the beginning, so long as he has not yet abandoned the truth, finds the right way. But persisting firmness is needed to keep one from drifting irresolutely

Abstract

SUMOylation, the post-translational attachment of SUMO (Small Ubiquitin-like Modifier) to a substrate protein, regulates the activity of several proteins involved in critical cellular processes like cell division and transcriptional regulation. SUMO is subsequently removed from substrates by SUMO-specific proteases, making this modification reversible. In plants, SUMOylation has been implicated in several physiological responses and flowering time control. ESD4 (Early in Short Days 4) encodes a SUMO-specific protease that prevents the accumulation of SUMO-conjugates in Arabidopsis. The *esd4-1* mutant shows a very early flowering phenotype as well as several shoot developmental distortions suggesting an important role of SUMOylation in the regulation of plant development. To investigate the role of SUMOylation in flowering time control a suppressor screen of *esd4* was performed. 120 independent suppressors of *esd4* (*sed*) were isolated and 15 of them further characterized. The SUMO-conjugate levels of these *sed*s are more similar to those of *esd4-1* than to the wild type. Rough map positions for five of these *sed* mutants were established using classical genetic methods, and combined with Next Generation Sequencing *sed111-1* was fine-mapped to a region of chromosome I that contains only six candidate genes. In a different study, SUMOylation of SHORT VEGETATIVE PHASE (SVP) was assessed and SUMO attachment lysines were determined using *E. coli* strains that recapitulate the SUMO conjugation pathway. SVP interacts with FLOWERING LOCUS C (FLC) to form a strong floral repressor complex. To study the role of SUMOylation in SVP function, an *svp*-null mutant (*svp-41*) was transformed with constructs aiming to hyperSUMOylate (translational fusions with SUMO or AtSCE) or hypoSUMOylate (mutations in the putative SUMO-attachment sites) the SVP protein in transgenic plants. Mutant phenotypes caused by these constructs are discussed.

Zusammenfassung

SUMOylierung ist eine post translationale Proteinmodifikation, in der an das Substrat kovalent ein oder mehrere SUMO (Small Ubiquitin-like Modifier) gebunden werden. Sie kontrolliert die Aktivität verschiedener Proteine, die in kritische zelluläre Prozesse, wie Zellteilung und Transkriptionsregulation, involviert sind. Die kovalente Bindung von SUMO an das Substrat kann von SUMO-spezifischen Proteasen gespalten werden, wodurch diese Modifikation reversibel ist. In Pflanzen wurde gezeigt, dass die SUMOylierung in verschiedenen physiologischen Antworten und in der Kontrolle des Blühzeitpunkts involviert ist. Das Gen *ESD4* (Early in Short Days 4) kodiert eine SUMO-spezifische Protease, die die Akkumulation von SUMO-Konjugaten in Arabidopsis verhindert. Die *esd4-1* Mutante zeigt einen sehr früh blühenden Phänotypus, sowie Störungen in der Sprossentwicklung, was auf eine wichtige Rolle der SUMOylierung in der Pflanzenentwicklung schließen lässt. Um die Rolle der SUMOylierung in der Kontrolle des Blühzeitpunkts zu untersuchen wurde ein suppressor screen von *esd4* durchgeführt. 120 unabhängige Suppressoren von *esd4* (*sed*) konnten isoliert werden, wovon 15 weiter charakterisiert wurden. Die SUMO-Konjugat-Konzentrationen dieser *sed*s waren denen von *esd4-1* ähnlicher als denen des Wildtyps. 5 dieser *sed*-Mutanten wurden mit klassischen genetischen Methoden grob lokalisiert. Außerdem wurde die Mutation von *sedIII-1* mit Next Generation Sequencing Methoden auf eine Region im Chromosom 1 eingeschränkt, die nur sechs Kandidatengene enthält. In einer weiteren Studie wurde die SUMOylierung von SHORT VEGETATIVE PHASE (SVP) gezeigt und mittels *E. coli* Stämmen, die den SUMO-Konjugations-Weg rekapitulieren, Lysine in SUMO-Bindungsstellen von SVP identifiziert. SVP bildet durch die Interaktion mit FLOWERING LOCUS C (FLC) einen starken Blührepressorkomplex. Um die Rolle von SUMOylierung in der SVP-Funktion zu untersuchen, wurde eine *svp*-null Mutante (*svp-41*) mit genetischen Konstrukten transformiert: Durch translationale Fusion mit SUMO bzw. AtSCE wurde zum einen hyperSUMOylierung des SVP-Proteins in den transgenen Pflanzen bewirkt. Während zum anderen die Mutation der putativen SUMO-Bindestellen in einer

hypoSUMOylierung des SVP-Proteins resultierte. Die Phänotypen dieser genetischen Konstrukte werden in dieser Arbeit diskutiert.

Contents

Abstract	I
Zusammenfassung	II
Contents	IV

1. General Introduction

1.1. Posttranslational modifications	1
1.1.1 Molecular mechanism of Sumoylation	2
1.1.1.1 SUMO Activating Enzyme (E1)	3
1.1.1.2 SUMO Conjugating Enzyme (E2)	4
1.1.1.3 SUMO Ligases (E3)	6
1.1.1.4 SUMO Proteases	7
1.1.1.5 SUMO isoforms	9
1.1.1.6 SUMOylation and other Postranslational Modifications crosstalk	10
1.1.1.7 Molecular effects of SUMOylation	14
1.1.1.8 SUMOylation and Transcriptional Factors regulation	15
1.2. SUMOylation in <i>Arabidopsis thaliana</i>	16
1.2.1 ESD4, a SUMO specific protease	19
1.2.2 SUMOylation and floral transition	21
1.2.3 SVP, a strong floral repressor	24
1.3. Aim of the Study	26

2. Isolation and characterization of *suppressors of esd4* (*sed*s)

2.1. Introduction	27
2.2. EMS mutagenesis experiment	29
2.3. Isolation of <i>suppressors of early in short days 4</i> (<i>sed</i>)	30
2.4. Phenotypic characterization of selected <i>sed</i> mutants	31
2.4.1 Flowering time of <i>sed</i> mutants under LD and SD conditions	37
2.4.2 Flower, siliques and leaf phenotypes	37
2.4.3 Pollen: viability and relative amount	41

2.5. Molecular characterization of selected <i>sed</i> mutants	43
2.5.1 Expression of SUMO pathway components	43
2.5.2 Protein SUMOylation levels	46
2.6. Discussion	48
3. Molecular Identification of <i>suppressors of esd4</i>	
3.1. Introduction	53
3.2. Generation of Mapping Populations	55
3.3. Map based cloning of <i>sed</i> mutants	55
3.4. <i>sed111-1</i> and Next Generation Sequencing strategy	65
3.4.1 Rough Map position established by bulk segregation analysis	65
3.4.2 Background cleaning	66
3.4.3 DNA extraction, sequencing and data analysis	67
3.4.4 Generation of de novo dCAPS markers, genetic interval, gene candidates	68
3.5. <i>sed5-1</i> backcross mapping strategy	74
3.6. Discussion	78
4. Analysis of a potential role of SUMOylation in the activity of the transcriptional and floral repressor protein Short Vegetative Phase	
4.1. Introduction	83
4.2. SVP SUMOylation	85
4.2.1 <i>In silico</i> analysis predicted seven SUMO-attachment sites in SVP	85
4.2.2 <i>In vitro</i> analysis	92
4.3. Investigations towards a potential role of SUMOylation in SVP function- <i>In vivo</i> analysis	98
4.3.1 Translational fusions	100
4.3.2 The effect of mutations of SUMO-site lysines on SVP function	105
4.3.3 Morphological abnormalities associated with SVP overexpression	106
4.4. Discussion	118

5. General Conclusions and Perspectives	129
6. Material and Methods	
6.1. Material	133
6.1.1. Chemicals and antibiotics	133
6.1.2. Enzymes	133
6.1.3. Kits	134
6.1.4. Vectors	134
6.1.5. Bacterial Strains	134
6.1.6. Oligonucleotides	135
6.1.7. Plant Materials	135
6.2. Methods	135
6.2.1. Plant Work	135
6.2.2. Molecular Biology	137
6.2.2.1 General molecular biological techniques	137
6.2.2.2 RNA isolation	139
6.2.2.3 Reverse transcription	139
6.2.2.4 semi-quantitative RT-PCR	139
6.2.2.5 Gene Cloning and PCR-directed mutagenesis	140
6.2.3. Biochemical methods	141
6.2.3.1 Protein Extraction and purification	141
6.2.3.2 SDS-PAGE and Western blot analysis	142
6.2.4 Software and Websites	142
7. References	143
8. Appendix	VII
A.1. Figures	
A.2. Tables	
A.3. Abbreviations	
A.4. SVP homologs GI codes	
A.5. SVP CDS / Transgenic plants generated in this work	
Acknowledgements	XVI
Erklärung	XVII
Lebenslauf	XVIII

1.1 Post-translational modifications

Plants are sessile organisms that must promptly respond to challenging external stimuli in order to survive. These stimuli are integrated by intricate hierarchical genetic networks that are controlled by a limited number of genes (Carrera et al. 2009). Despite the limited number of these genes the protein diversity they encode is larger, sometimes by orders of magnitude (Walsh et al. 2005). This diversity is generated by mechanisms that act at different molecular levels, including post-transcriptional mechanisms such as alternative transcriptional initiation and mRNA splicing; translational mechanisms, like alternative translational initiation which affect the identity of the protein; and finally, by post-translational regulation or post-translational modifications (PTM), which are an efficient way to modulate the function of a protein, its stability, its protein partners or its cellular location (Peck 2005; Walsh et al. 2005).

PTM consist of the chemical modifications to substrate proteins and include the addition of smaller chemical groups, carbohydrates, lipids and other small polypeptides (Peck 2005). PTM can affect the activity, stability and localization of the target protein and have a direct impact on protein function and potential protein-protein interactions (Johnson 2004; Walsh et al. 2005). PTMs are also an effective and relatively energy efficient way to alter protein function and therefore are commonly used by independent cellular processes to provide a fine and fast modulation of protein function (Matafora et al. 2009).

The amino acid lysine (Lys) is target for many different post-translational modifications like methylation, acetylation, hydroxylation, ubiquitination and the addition of several ubiquitin-like proteins including SUMO (Small Ubiquitin-related MOdifier)(Seo and Lee 2004; Walsh et al. 2005; Jadhav and Wooten 2009). Methylation involves the addition of one or multiple methyl groups by the action of Methyltransferases, the addition of this group increases the hydrophobicity of the bearing Lysine (and protein) (Walsh et al. 2005). Acetylation changes the charge distribution on Lys side chains and has been implicated in blocking the availability of the Lys to other PTM, like Ubiquitination or SUMOylation, having in this way a direct impact on the proteins half-life and/or function (Walsh et al. 2005; Wu and

Chiang 2009). Hydroxylation is a type of oxidation which is mediated by enzymes, this is the only non-reversible modification (Walsh et al. 2005; Jadhav and Wooten 2009). In contrast to the previous modifications, where a chemical group is added, Ubiquitination and SUMOylation consist of the addition of small proteins (Ubiquitin or SUMO, respectively) to the substrate protein (Hershko and Ciechanover 1992; Matunis et al. 1996). Ubiquitin is a protein of around 76 amino acids whose most conspicuous effect is the turnover of proteins by the proteasome, however it is also involved in targeting proteins to subcellular regions, modulating enzymatic activity and protein-protein interactions (Ikeda and Dikic 2008). SUMO, on the other hand, is a protein of around 100 amino acids and has been more related with the establishment of protein-protein interactions and targeting of the substrate protein to the nucleus (Johnson 2004).

1.1.1 Molecular mechanism of SUMOylation

SUMOylation is a reversible post-translational modification process that involves the covalent attachment of SUMO to a target protein (Matunis et al. 1996; Mahajan et al. 1997). SUMO presents a very similar three-dimensional structure to Ubiquitin although they share only 18% amino acid sequence identity (Bayer et al. 1998). In contrast to Ubiquitin, SUMO presents a highly flexible protruding amino terminal tail consisting of around 15 residues (Bayer et al. 1998). This N-terminal tail might be involved in the formation of oligomeric SUMO chains (Tatham et al. 2001).

SUMO conjugation involves three enzymatic reactions very similar to those performed during the Ubiquitylation process (Johnson 2004). They are referred to as activation, conjugation and ligation or E1, E2 and E3, respectively. During SUMO-Activation a high energy thiolester bond is formed between a cysteine, located in the active site of the Activating enzyme (SAE), and the carboxyl group of the terminal glycine of SUMO. This high energy bond is necessary for the interaction between SUMO and the Conjugating enzyme (SCE) that, by a transesterification reaction, forms a thiolester bond between the active SUMO and a cysteine residue in the SCE. Afterwards, the conjugating enzyme and a SUMO-ligase transfer SUMO to a lysine residue in the target protein (Johnson 2004). This lysine residue is usually located

within a highly conserved motif called the SUMO-attachment motif ($\Psi\text{KxE/D}$; where Ψ is a hydrophobic amino acid, K is a lysine, x any amino acid and E or D are acidic amino acids) (Rodriguez et al. 2001) (Figure 1).

Like Ubiquitin, SUMO is synthesized as a precursor that requires a maturation step to expose a di-glycine motif necessary for the conjugation. The maturation as well as the removal of SUMO from conjugates is performed by SUMO-specific proteases also known as Ubiquitin Like specific Proteases (ULP) (Li and Hochstrasser 1999).

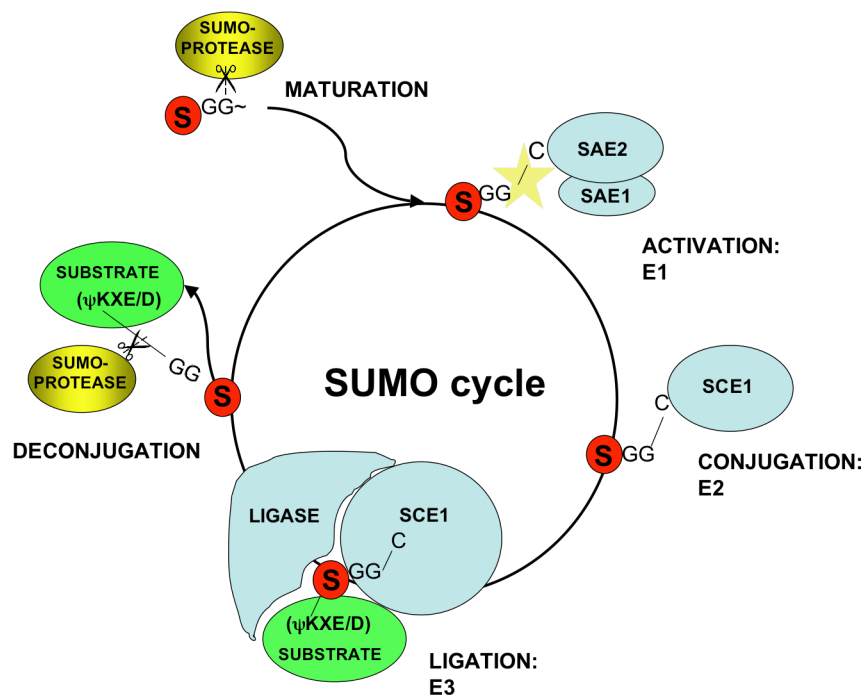


Figure 1. SUMO cycle. SUMO is encoded as a precursor that requires the removal of a carboxy terminal extension (processing or maturation) to expose the mature Gly-Gly C terminus motif required for the interaction with the downstream enzymes. The SUMO cycle comprises of three enzymatic steps consisting of the Activation (E1), Conjugation (E2) and Ligation (E3) reactions. SUMO is attached to Lysines that form part of what is known as the SUMO-attachment motif within the protein target. The maturation and deconjugation steps are performed by SUMO-specific proteases (see text for details).

1.1.1.1 SUMO Activating Enzyme (E1)

The SUMO-activating enzyme (SAE) or E1 is a heterodimeric enzyme composed of a large subunit (SAE2) and a small subunit (SAE1) (Dohmen et al. 1995; Desterro et al. 1999). The SAE heterodimer is very similar to Ubiquitin E1 monomer, the SAE1

resembles the N-terminal part of the Ubiquitin E1 while the SAE2 resembles the C-terminal part where the catalytic cysteine is found (Dohmen et al. 1995; Desterro et al. 1999). SAE1 contains only a single domain, which seems to participate in the adenylation of SUMO. The SAE2 subunit contains three domains, the adenylation domain, the catalytic Cys domain and the Ubiquitin Fold Domain (UFD) localized at the C-terminal region where also a Nuclear Localization Signal (NLS) can be found (Dohmen et al. 1995; Lois and Lima 2005). The adenylation domain is necessary for the recognition of the SUMO C-terminal part and the subsequent formation of an adenylate in an ATP dependent process. The catalytic cysteine domain attacks the adenylated SUMO C-terminus forming a high energy thiolester bond necessary for the subsequent transfer of SUMO to the SUMO Conjugating Enzyme (Lois and Lima 2005). The UFD serves as a surface interaction with the SUMO Conjugating Enzyme (SCE)(Wang et al. 2010).

1.1.1.2 SUMO Conjugating Enzyme (E2)

The SUMO Conjugating Enzyme (SCE) or E2 in mammals is a single domain structure composed of four-stranded antiparallel beta-sheet forming the core, surrounded by four alpha helices at both ends (Tong et al. 1997). The SCE presents a Cysteine in its active site or catalytic domain which interacts with the SUMO C-terminal Glycine residue to form an intermediate covalent thioester bond (Reverter and Lima 2005). The SCE active site is surrounded by a negative patch that interacts with the tetrapeptide motif (Ψ KxE/D) in the substrate protein (Bernier-Villamor et al. 2002; Tatham et al. 2003), however, variations of this consensus sequence have been reported indicating that this is not a required interacting surface. The SCE enzyme interacts also non-covalently with the SUMO and the SAE. The non-covalent SCE-SUMO interaction takes place on the opposite side of the catalytic domain of the SCE, while the interaction SCE-SAE, involves interacting interfaces located around the Cys-active sites from both enzymes which are important to transfer the activated SUMO (Knipscheer et al. 2007; Capili and Lima 2007; Duda et al. 2007; Wang et al. 2007). Interestingly, there is only one encoded SCE in all the organisms analyzed, including yeast, invertebrates, mammals and plants (Johnson and Blobel 1997; Johnson 2004; Kurepa et al. 2003).

SUMO consensus motifs and variations

SUMO consensus motifs are a set of amino acids that surround the SUMO-attachment lysine. These motifs provide an interacting surface for the E2 improving the protein substrate-E2 interaction as well as SUMO transfer (Bernier-Villamor et al. 2002). SUMO consensus motifs consist of the target lysine residue, a hydrophobic amino acid Ψ , any amino acid (x) and an acidic amino acid E or D ($\Psi K x E / D$) (Rodriguez et al. 2001). Different studies have identified variations of this SUMO consensus motif including the inverted SUMOylation consensus motif (ED) x K(VILFP), hydrophobic cluster motif ($\Psi\Psi\Psi K x E$), phosphorylation-dependent SUMOylation motif (PDSM, $\Psi K x E x x S P$) and negatively charged amino acid-dependent SUMO motif (NDSM, $\Psi K x E x x E E E E$) (Matic et al. 2010; Yang et al. 2006; Hietakangas et al. 2006). Apparently acidic residues enhance the interaction between the substrate and the E2, while phosphorylation sites can functionally substitute for acidic residues (Yang et al. 2006).

SUMO-interacting motifs

SUMO Interacting Motifs or SIMs are short-stretches of hydrophobic amino acids ((V/I/L)-X-(V/I/L)-(V/I/L)) that allow non-covalent interactions between SUMO and SIM-containing proteins (Song et al. 2004; Kerscher 2007). The hydrophobic core of the SIM forms an extended β -strand configuration which becomes inserted into a hydrophobic groove between the $\alpha 1$ -helix and a $\beta 2$ -strand in SUMO, this SIM β -strand can be bound in a parallel or antiparallel manner with respect to the $\beta 2$ -strand in SUMO (Kerscher 2007). SIMs are sometimes followed or preceded by a stretch of acidic amino acids that interact with a basic patch in the E2 surface and which probably determine the polarity of SIM-SUMO interaction (Gareau and Lima 2010). SIMs exist in the majority of the proteins involved in SUMO-dependent processes, including SUMO cycle enzymes, SUMO-substrates, SUMO-binding proteins and SUMO-targeted ubiquitin ligases (Song et al. 2004; Gareau and Lima 2010). Some SIM are followed by a patch of Serine residues that can be phosphorylated, these SIMs are called phospho-SIMs. When these Serines are phosphorylated they can mimic acidic amino acids enhancing the interactions with the E2 by extending the interaction surface (Gareau and Lima 2010). Remarkably, SIMs are commonly

observed in SUMO E3 ligases, and the deletion of the SIM may result in a decreased activity of the protein (Reverter and Lima 2005; Gareau and Lima 2010)

1.1.1.3 SUMO Ligases (E3)

SUMO Ligases or E3 are proteins that interact with the SUMO conjugating enzyme E2, and facilitate SUMO transfer to the protein substrate enhancing SUMOylation (Johnson 2004). E3 facilitates SUMOylation by one of two mechanisms, they can act as scaffold proteins allowing the interaction of the E2-SUMO thioester with the substrate or they can act as enhancers that help to position the E2-SUMO thioester and the target protein in a way that promotes SUMO transfer (Wilkinson and Henley 2010; Gareau and Lima 2010). It has been proposed that the E3 might be required to SUMOylate substrates that do not contain a consensus motif or to participate in the control of target specificity. At least three types of SUMO E3 ligases have been identified: PIAS family (protein inhibitor of activated STAT, with divergent RING-like motif), RanBP2 (nuclear pore complex (NPC)-associated protein with zinc finger domains) and Pc2 (polycomb group protein) (Johnson 2004).

PIAS family

Protein Inhibitor of Activated STAT or PIAS are proteins that share a conserved N-terminal domain of around 400 amino acids (Johnson 2004). This N-terminal domain contains four conserved structural motifs: a SAP (acronym derived from the scaffold attachment factor SAFA/B, ACINUS and PIAS proteins) motif involved in the interaction with DNA and chromatin in nuclear microdomains; a PINIT (for Pro-Ile-Asn-Ile-Thr) motif which in conjunction with the RING finger of the SP-RING zinc finger motif forms a functional module with E3 ligase activity; an SP-RING zinc finger motif similar to those found in ubiquitin E3 ligases; and a SIM, SXS, motif which interacts non-covalently with SUMO (Johnson 2004; Sharrocks 2006; Cheong et al. 2010). PIAS work as transcriptional co-regulators with activating or repressive effects, depending on the target gene and/or interacting transcriptional regulator (Sharrocks 2006). PIAS proteins are highly conserved and are present in yeast (Siz1, Siz2), humans (PIAS1, PIAS3, PIASXa, PIASXb and PIASy) and plants (AtSIZ1) (Johnson 2004; Sharrocks 2006; Cheong et al. 2010). PIAS, as an E3 SUMO ligase, enhances the interactions between the E2-SCE and the substrate protein. PIAS

proteins seem to also perform functions that are independent of their SUMO ligase activity (Johnson 2004; Sharrocks 2006; Cheong et al. 2010).

RanBP2

RanBP2 is the major component of the cytoplasmic filaments of the nuclear pore complex and a SUMO E3 ligase protein (Johnson 2004). The SUMO E3 ligase domain of RanBP2 consists of two repeats of around 50 amino acids each called internal repeat (IR) domain(s), which have not been found in any other protein (Johnson 2004). The RanBP2 IR domains can interact with the SUMOylated version of RanGAP1. RanGAP1 is the GTPase-activating protein for Ran, a nuclear Ras-like GTPase involved in nucleocytoplasmic transport of proteins and ribonucleoproteins. RanBP2, the SUMO1- RanGAP1 and Ubc9/E2 form a ternary complex that plays an important role in nuclear trafficking (Matunis et al. 1996).

Pc2

Some Polycomb group (PcG) proteins are evolutionarily conserved in animals and plants. PcG form large multimeric complexes involved in the remodeling and maintenance of the chromatin state by the methylation of histones, inducing an epigenetic gene silencing effect. Pc2 is a component of one of the two known PcG repressor complexes, the PRC1. Pc2 establishes an association between the transcriptional corepressor CtBP (C-terminal Binding Protein) and the PcG bodies. Pc2 can also interact with SUMO and the E2 functioning as a scaffold facilitating CtBP SUMOylation (Wilkinson & Henley 2010).

1.1.1.4 SUMO Proteases

SUMO proteases are enzymes that participate in the regulation of the SUMO-conjugated protein levels in the cell by performing two main functions, maturation of the SUMO precursor by a C-terminal hydrolase activity (processing) and SUMO deconjugation from the substrates by an isopeptidase activity (deconjugation) (Li and Hochstrasser 2003). SUMO maturation consists of the removal of a stretch of carboxyl terminal amino acids in the SUMO-precursor to expose a di-glycine motif necessary for the interaction with the downstream enzymes required for SUMOylation (Johnson 2004). SUMO deconjugation consists of the removal of SUMO from the

protein substrate by hydrolysis of the isopeptide bond between SUMO and the protein substrate (Johnson 2004).

SUMO proteases present a variable N-terminal domain and a conserved C-terminal domain of around 200 amino acids. The N-terminal domain has been shown to be a determinant for cellular localization and substrate specificity (Li and Hochstrasser 2003). The C-terminal domain is highly conserved among the SUMO proteases, it contains the catalytic triad His-Asp-Cys which classifies them as the family of Cys48 cysteine proteases of the clan CE (Drag and Salvesen 2008). CE Cysteine proteases active site forms a cleft that binds to the C-terminal domain of SUMO and cuts after a di-glycine sequence, making these proteases one of the most specialized ones (Drag and Salvesen 2008).

SUMO proteases have a strong impact on many functions within a cell, and therefore its expression and activity are tightly regulated (Mukhopadhyay and Dasso 2007; Kim and Baek 2009; Xu et al. 2009). Some of these mechanisms occur at the transcriptional level and appear to be crucial for normal cell function as the alteration of expression of SENP is observed in several carcinomas (Bawa-Khalfe and Yeh 2010). Other mechanisms of regulation act at the protein level like the subcellular localization directed by intrinsic protein sequences localized at the N-terminal region (Li and Hochstrasser 2003), or protein modifications. For example, oxidative stress inactivates the active site of the SUMO protease (Xu et al. 2009). On the other hand, SUMO itself can regulate the action of SUMO-proteases in mammals and plants. In this case, the amino acids that are located before and after the di-glycine motif in SUMO, which vary among the different SUMO paralogs, affect the affinity of the SUMO proteases for SUMO and therefore influence the maturation efficiency (Colby et al. 2006; Xu et al. 2009).

Yeast present two SUMO proteases or Ubiquitin-like proteases, Ulp1 and Ulp2. Ulp1 is localized at the Nuclear Pore Complex (NPC) and performs important roles in nuclear-cytoplasmic trafficking (Li and Hochstrasser 1999; Li and Hochstrasser 2003). Ulp2 is localized at the nucleoplasm and seems to be more important for deconjugation and SUMO-chain disassembly (Li and Hochstrasser 1999; Li and

Hochstrasser 2000; Li and Hochstrasser 2003). In mammals, the SUMO protease family is comprised of six members (SEN1-3, SEN5-7)(Kim and Baek 2009). SEN1 localizes at the nucleoplasm, SEN2 to the nuclear side of the NPC, SEN3 localizes at the nucleolus and SEN6 localizes predominantly to the cytoplasm (Kim and Baek 2009). Arabidopsis is predicted to present several potential SUMO proteases, until now four have been functionally characterized ESD4-LIKE SUMO PROTEASE 1 (ELS1, alternatively called AtULP1a), Early in Short Days 4 (ESD4), Overly Tolerant to Salt 1 (OTS1 or AtULP1d) and Overly Tolerant to Salt 2 (OTS2 or AtULP1c) (Murtas et al. 2003; Conti et al. 2008; Lois 2010; Hermkes et al. 2011). ELS1 is localized in the cytoplasm, ESD4 is localized at the periphery of the nuclear envelope, while OTS1 and OTS2 are localized in the nucleoplasm (Murtas et al. 2003; Conti et al. 2008; Lois 2010; Hermkes et al. 2011).

Analysis of the evolutionary relationships of the active site-protease domain, also referred to as Ulp-Domain (UD) of the SUMO proteases classifies them into two big evolutionarily related groups, the ULP1-like and Ulp2-like, these two groups share similar active site organization and cellular distributions (Li and Hochstrasser 2000; Mukhopadhyay and Dasso 2007). The mammal SUMO-proteases grouped into the Ulp1-like branch consisted of SEN1-3 and SEN5 and the Ulp2 branch comprises SEN6 and SEN7 (Mukhopadhyay and Dasso 2007). In Arabidopsis all the characterized SUMO proteases, namely ELS1, ESD4, OTS1 and OTS2, are members of the ULP1 type group (Murtas et al. 2003; Conti et al. 2008; Lois 2010; Hermkes, Fu, et al. 2011).

1.1.1.5 SUMO isoforms

SUMO is a small protein that forms part of the superfamily of ubiquitin-like modifiers. Like Ubiquitin, SUMO is conjugated in three enzymatic steps: activation (E1), Conjugation (E2) and Ligation (E3) (Johnson and Hochstrasser 1997). SUMO shares only 18% of sequence identity with Ubiquitin, but presents a highly similar three-dimensional structure characterized by a $\beta\beta\alpha\beta\beta\alpha\beta$ fold, the so called Ubiquitin-fold (Bayer et al. 1998). The ubiquitin fold, also known as β -GRASP motif (β -Golgi reassembly stacking protein motif), is a characteristic globular and compact three-

dimensional structure that can be found among functionally unrelated proteins (Orengo et al. 1994).

There are two main features that distinguish SUMO from ubiquitin: the presence of an amino-terminal tail in SUMO and a distinct surface charge distribution (Bayer et al. 1998). The amino-terminal tail is a highly flexible appendix containing around 20 amino acids (Bayer et al. 1998). This tail extends out from the globular core of the protein and has been related to the formation of poly-SUMO chains and might also provide an interacting surface to guide SUMO-specific protein-protein interactions (Bayer et al. 1998; Tatham et al. 2001).

The surface charge distribution on SUMO is mostly conserved among all human SUMO proteins and SMT3 in yeast (Bayer et al. 1998; Huang et al. 2004). These zones provide the surfaces for the interaction with the SUMO modifying enzymes (Liu et al. 1999; Huang et al. 2004). SUMO presents different charge surfaces, a mostly negative surface formed by four negatively charged amino acids and a positive surface provided by a conserved lysine, and a negative pocket at the junction of the tail and the globular core body of the protein (Bayer et al. 1998; Huang et al. 2004). The positive and negative surfaces are opposite to each other and have been shown to interact with the SUMO conjugating enzyme Ubc9 (Bayer et al. 1998; Liu et al. 1999; Tatham et al. 2003; Huang et al. 2004). Finally SUMO presents a small flexible C-terminal region where the glycines necessary for conjugation are located. This region is recognized by SUMO proteases as the crystal structure of yeast Ulp1-SMT3 revealed (Mossessova and Lima 2000). SUMO dynamics depend on several factors like: rates of SUMO conjugation and deconjugation, substrate function and biochemical properties and the used SUMO-isoform (Ayaydin and Dasso 2004).

Yeast contains a single SUMO, Smt3p (budding yeast) and pmt3 (fission yeast). Mammals present four SUMO isoforms, SUMO1-4 (Ayaydin and Dasso 2004). Plants have nine SUMO genes but only four are expressed (Kurepa et al. 2003; Saracco et al. 2007; van den Burg et al. 2010).

In mammals, SUMO1 is the most commonly expressed and conjugated isoform. It localizes at the nuclear envelope and to the nucleolus and within cytosolic dots

(Ayaydin and Dasso 2004). SUMO 2/3 are 96% identical, both are distributed throughout the nucleoplasm and can form polymeric chains (Ayaydin and Dasso 2004). SUMO4 is expressed only in the kidneys, spleen and lymph nodes (Wilkinson and Henley 2010). Mammalian SUMO paralogues are functionally distinct and specifically regulated in vivo (Ayaydin and Dasso 2004). Interestingly, the SUMO dynamics vary among substrates, SUMO2/3 are more rapidly deconjugated than SUMO1, which also coincides with a relatively larger free pool of SUMO2/3 (Ayaydin and Dasso 2004). SUMO2/3 can be readily conjugated to target proteins in reaction to stress stimuli (Johnson 2004).

In Arabidopsis nine SUMO genes have been described (SUMO 1 to SUMO 9), where SUMO 9 is more likely to be a pseudogene (Kurepa et al. 2003). Until now only SUM1-3 and SUM5 have been shown to encode for a functional mRNA (Kurepa et al. 2003; Saracco et al. 2007; van den Burg et al. 2010). All SUMOs present a very similar two intron/ three exon structure, with the exception of SUMO 5 (which only presents the second intron) suggesting a common-ancestor origin and a recent tandem duplication (Kurepa et al. 2003). In comparison to SUMO1, SUMO2 share 89% protein sequence identity, while SUMO3 shares a 48% and SUMO5 35% (Colby et al. 2006; Budhiraja, Hermkes, et al. 2009).

SUMO-chains

SUMO, like Ubiquitin, can be attached to its protein substrate one or multiple times as a monomer, or multiple times as a polymer (Ulrich 2008). The SUMO chain formation is proposed to be mediated by the E2 (Capili and Lima 2007). The E2/Ubc9 can interact covalently and non-covalently with SUMO, the SUMO-Ubc9 non-covalent interaction is via a SIM motif which is localized at the opposite side of the active site (Capili and Lima 2007). It is likely that an Ubc9 homodimer forms where one protein would perform the SUMO activation while the other would position the SUMO as a target to form the SUMO chain (Capili and Lima 2007).

The SUMO-attachment sites probed until now are localized within the unstructured N-terminal tail. Interestingly, these SUMO-attachment sites vary in number and localization among the different SUMO paralogs (Ulrich 2008). In mammals, only

the SUMO2/3 paralogs can form poly-SUMO chains while the SUMO1 can not (Matic et al. 2008; Wilkinson and Henley 2010). A similar scenario has been reported for Arabidopsis where SUMO1/2 form polymeric chains but neither SUMO3 nor SUMO5 (van den Burg et al. 2010; Lois 2010). The SUMO paralogs that are unable or inefficient at forming chains may act as terminators of SUMO chains (Denuc and Marfany 2010).

Substrate modification by poly-SUMO chains have different consequences than mono-SUMO modification. In Yeast, the attachment of poly-SUMO chains serves as a signal for protein degradation (Uzunova et al. 2007). Another level of complexity remains to be explored, the possibility of the formation of polymeric chains based on different SUMO-attachment lysines, this remains to be explored. This phenomenon has been reported for Ubiquitin where identity of the linkage formation of Ubiquitin chains has a direct impact on the targeted protein, for instance if the polyubiquitin chain is linked via the K48 the protein is targeted for degradation by the 26S proteasome, while the consequences of the formation of polymeric chains based on the other six lysines in the Ubiquitin promote different effects (Ikeda and Dikic 2008).

SUMO-targeted ubiquitin ligases (STUbLs)

SUMO-targeted ubiquitin ligases (STUbLs) are ubiquitin-ligases that present several SUMO-interaction motifs in their sequence, these SIMs interact with the poly-SUMO chains in their targets (Uzunova et al. 2007; Miteva et al. 2010). STUbLs target a SUMO substrate for degradation establishing in this way a link between the ubiquitylation and SUMOylation processes, and also demonstrating that SUMOylation can function as a signal for degradation by the proteasome (Uzunova et al. 2007).

SUMO-like domains

SUMO-like domains are three-dimensional β -GRASP fold structures similar to SUMO despite a very low sequence identity (Novatchkova et al. 2005). These domains are proposed to act as molecular mimics of SUMO and share its non-covalent interactions with a subset of SUMO pathway enzymes that contain SIMs (Prudden et al. 2009). These SUMO-like domains are part of multidomain proteins

involved in different processes (Novatchkova et al. 2005).

1.1.1.6 SUMOylation and other Post-translational Modifications crosstalk

SUMOylation can be induced or inhibited by other PTM that occur on the same target protein. These PTM include phosphorylation, acetylation and Ubiquitination.

Phosphorylation can either inhibit or induce SUMOylation. For example, phosphorylation induces SUMOylation on phosphorylation-dependent SUMOylation motifs or PDSM (Hietakangas et al. 2006; Mohideen et al. 2009). In PDSM the conventional consensus motif is followed by a serine and a proline residues which are target of phosphorylation (Hietakangas et al. 2006). The phosphorylation serine and proline provide a negative charge that enhances the interaction with some positive or basic patches in the SCE, increasing the levels of SUMO conjugation of the substrate (Hietakangas et al. 2006; Mohideen et al. 2009). This motif is present in proteins such as heat-shock factors (HSFs) and myocyte enhancer factor 2 (MEF2) (Hietakangas et al. 2006; Mohideen et al. 2009). In addition, phosphorylation of a substrate may lead to its re-localization or conformational changes that expose the SUMO attachment sites (Gareau and Lima 2010).

Acetylation establishes a complex crosstalk with SUMOylation during the control of gene expression. One example of this mechanism is the control of p53, a tumor suppressor and transcriptional regulator in mammals. p53 presents three lysine residues at its C-terminal end that are targets of acetylation and SUMOylation. Acetylation of these lysines enhances p53 binding to consensus DNA sites that are otherwise scarcely bound by unmodified p53. On the other hand, SUMOylation can activate or repress the transcriptional activity of p53 depending on the promoter context or cellular type. Interestingly, a recent report showed that SUMO-1 itself can be acetylated in a region that, according to the authors, resembles the acetylated site in the C-terminal region of p53, and that this SUMO-1 acetylation form can modify the p53 transcriptional program favouring apoptosis (Cheema et al. 2010).

Ubiquitination can compete for the same lysine residues with SUMOylation or act as a second PTM that targets the substrate to degradation (Wilkinson and Henley 2010). The crosstalk between SUMO and ubiquitin pathways was uncovered by the discovery of a conserved family of RING finger ubiquitin ligases, afterwards named SUMO-targeted ubiquitin ligases (STUbLs). These ligases exhibit several SUMO interaction motifs or SIMs that recognize their SUMOylated targets and target them for degradation (Miteva et al. 2010).

SUMO is also a target for PTM like phosphorylation or acetylation (Matic et al. 2008; Cheema et al. 2010). The functional consequences of these modifications remain to be further explored but open a new and intriguing level of regulation in the SUMOylation process.

1.1.1.7 Molecular effects of SUMOylation

The effects that the attachment of SUMO has on a substrate protein are very variable. SUMOylation may affect protein stability, localization or activity (Johnson 2004). The first reported effect of SUMOylation on a protein substrate function was the localization in distinct subcellular domains of RanGAP1 (Matunis et al. 1996). Other effects were later identified by characterizing the effects of SUMOylation on diverse proteins.

SUMO can act as a molecular “glue” by providing an extra interacting surface that allows recruitment of new interacting proteins. An example of this mechanism is provided by PML (Promyelocytic leukemia protein), a tumor suppressor protein in humans (Johnson 2004). PML is modified by SUMO at three distinct Lysines, PML SUMO modification is necessary for the formation of morphologically normal nuclear bodies (NB) (Johnson 2004). NB are dynamic molecular reservoirs that control the availability of certain transcription factors to active chromatin domains (Johnson 2004).

SUMO may also act as a steric hindrance modification or shield for protein interaction motifs or surfaces. For example, the SUMOylation of the Ubiquitin Conjugating enzyme E2-25k impedes its interaction with the E1-Ubiquitin enzyme

(Wilkinson and Henley 2010).

SUMOylation can lead to a conformational change in the target substrate. A clear example of this mechanism is presented by TDG (thymine-DNA glycosylase). TDG is a protein involved in the recognition and removal of mismatched bases in DNA. Once TDG removes the mismatched base it remains attached to the DNA with a high affinity, SUMOylation of TDG helps to decrease this TDG-DNA affinity by inducing a conformational change in TDG, which is mediated by the interaction of the SUMO and a SIM in the same TDG protein. The removal of SUMO from TDG by SUMO proteases reconstitute TDG to its functional state for a new round of activity (Johnson 2004).

1.1.1.8 SUMOylation and Transcription Factor regulation

The majority of the reported SUMO substrates are Transcription Factors (TF). In addition, a large number of other transcriptional regulators are also targets of SUMOylation, suggesting a link between SUMO and the regulation of gene expression (Gill 2005). The effects of SUMOylation on the TF function are mostly related to inhibition of TF activity or the promotion of TF repressor function. SUMOylation can regulate the activity of a transcription factor by different mechanisms, for example, competing with other PTM for the same lysine residue, impeding the interaction with other proteins or the DNA or allowing new interactions that are established based on the interaction surface of SUMO (Gill 2005). SUMO can establish protein-protein interactions with co-repressors like histone deacetylases (HDACs), which remove acetyl groups from the DNA correlating with transcriptional repression, or with PcG proteins, which regulate gene expression through chromatin remodeling (Garcia-Dominguez and Reyes 2009). Interestingly, the SUMO-attachment site was previously identified as a repressor domain in some transcription factors (Gill 2005).

SUMOylation takes place on lysines that can also be targets of other PTM like acetylation. HDACs can remove acetyl groups from lysines that may also be SUMO attachment sites (Gill 2005). Interestingly, HDAC4 has been reported to function as a E3-SUMO ligase (Gill 2005; Garcia-Dominguez and Reyes 2009).

MEF2

Myocyte enhancer factor-2 (MEF2) is a family of transcription factors (MEF2A-D) that regulate the expression of genes involved in the myogenesis and synaptogenesis in mammals (Santelli and Richmond 2000; Mohideen et al. 2009). MEF2 presents three domains, a MADS-domain of around 58 amino acids that mediates DNA binding and dimerization, a MEF2 domain of 28-residues that restricts the formation of heterodimers within the family excluding their formation with other MADS-proteins, and a C-terminal domain which has been related to transactivation activities (Santelli & Richmond 2000). Some MEF2 proteins are targets of SUMOylation mediated by the class II HDACs (Garcia-Dominguez & Reyes 2009). The characterization of the complex HDAC9/MEF2/DNA revealed that HDAC9 binds to the hydrophobic groove of the MEF2 dimer (Han et al. 2005). MEF2A presents a phosphorylation-dependent SUMO consensus motif (PSDM), and its SUMOylation converts this transcription factor from activator to repressor, which promotes synapse maturation in neurons (Mohideen et al. 2009). Remarkably, the SUMO-attachment lysine in the PSDM of MEF2A is also target of acetylation, this modification promotes the function of MEF2A as a transcriptional activator which results in synapse disassembly (Gareau & Lima 2010). MEF2 is a good example of the level of complexity that PTM can reach, how the different modifications are integrated and how the sequential order of such modifications has a direct impact on the output-function of the protein.

1.2. SUMOylation in *Arabidopsis thaliana*

SUMOylation in *Arabidopsis* combines diversity and specificity. Diversity is given by different SUMO isoforms and the distinct regulatory domains present in SUMO-proteases, whilst specificity is given by a combination of the selective function of SUMO ligases and SUMO-proteases as well as cellular location of components and expression patterns (Chosed et al. 2006).

SUMO

As mentioned before, only four *SUMO* genes (*SUM1*, 2, 3 and 5), out of the eight predicted, have been shown to be expressed (Kurepa et al. 2003; Saracco et al. 2007; van den Burg et al. 2010). SUMO 1 and 2 share 93.5% amino acid sequence identity

and its correspondant genes are the most transcriptionally active (Kurepa et al. 2003; Saracco et al. 2007; van den Burg et al. 2010). *SUM1* and *SUM2* present a similar broad spatial expression pattern and a related response to stress stimuli, while *SUMO3* and *SUMO5* are expressed at much lower levels and only in certain organs suggesting that they could perform a specific role (Saracco et al. 2007; van den Burg et al. 2010). Mutant analysis showed that single homozygous mutants *sum1* or *sum2* are viable, while the double homozygous mutant *sum1/sum2* is embryo lethal, suggesting that *SUMO1* and *SUMO2* are functionally redundant (Saracco et al. 2007). The *sum3* mutant is late flowering while characterization of a *sum5* mutant is lacking (Saracco et al. 2007; van den Burg et al. 2010). *In vitro* reconstitution of the Arabidopsis SUMOylation cascade showed that SUMO1 and SUMO2 are able to form polymeric chains while SUMO3 is not (Colby et al. 2006). The formation of homodimer SUMO1 chains has been confirmed in vivo by mass spectrometry analysis (Miller et al. 2010). The surfaces of plant SUMO paralogues that interact with E1 and E2 and SIMs are well conserved in SUMO1/2, in a lesser extent in SUMO3 and being most divergent in SUMO5 (Castaño-Miquel et al. 2011). SUMO3 and SUMO5 are deficient in establishing E2-non covalent interactions. Consistent with this observation these SUMOs have low conjugation rates (Castaño-Miquel et al. 2011). Until now there are no efficient specific SUMO3 and SUMO5 antibodies and thus the conjugation patterns are still unknown. Therefore there is a gap of information in the way those SUMOs act.

Interestingly, in Arabidopsis, SUMO1 forms poly-SUMO chains and is a target of ubiquitination, this last modification occurs after heat-shock stress (Miller et al. 2010). This result suggests that SUMO can work as a signal for protein degradation as in yeast.

SAE and SCE

As previously mentioned the SUMO-activating enzyme (SAE) is a heterodimeric enzyme. In Arabidopsis the large subunit (SAE2) is encoded by only one gene (*SAE2*). The small subunit (SAE1) is encoded by two genes (*SAE1a* and *SAE1b*), which seem to be functionally redundant. The SUMO-conjugating enzyme is encoded by only one gene, *SCE* (Kurepa et al. 2003; Murtas, Reeves, et al. 2003). The *sae2* and *sce* mutants are embryo lethal (Saracco et al. 2007).

SUMO-ligases

In *A.thaliana* multiple SUMO-ligases are predicted, but only two, SIZ1 (a protein that contains a SAP and a zinc-finger MIZ1 domain) and HIGH PLOIDY 2 (HPY2), have been characterized (Miura et al. 2005; Ishida et al. 2009; Miura and Hasegawa 2010).

SIZ1 is the SUMO machinery component most extensively characterized, it is involved in several physiological responses like freezing tolerance, phosphate starvation, basal thermotolerance, innate immunity, abscisic acid (ABA) responses, nitrogen assimilation and flowering time regulation (Miura et al. 2005; Yoo et al. 2006; Miura et al. 2007; Jin et al. 2008; Park et al. 2011). The transcript profile of *SIZ1* shows that it is present in all the analyzed tissues (Saracco et al. 2007). Protein analysis showed that SIZ1 compartmentalizes to nuclear speckles and that it is mainly involved in stress-responsive SUMO1/2 conjugation (Miura *et al*, 2005; Miura *et al*, 2007). Although SUMOylation through SIZ1 controls critical processes in plants it is not essential for survival. This could be due to the presence of other SUMO-ligases not yet identified that might perform similar functions (Miura et al. 2005; Ishida et al. 2009; Miura and Hasegawa 2010).

Arabidopsis High Ploidy 2 (AtHPY2) also called Methyl methanesulfonate-sensitivity protein 21 (AtMMS21) is an E3 SUMO ligase with two characteristic domains, a SP-RING domain exclusively found in SUMO E3s required for the SUMOylation activity and a binding motif that presents five conserved cysteine/histidine residues as Zn²⁺-coordinating ligands (Huang et al. 2009). AtMMS21 is localized in the nucleus and cytoplasm and is specifically involved in cytokinin signalling necessary for root development (Huang et al. 2009). AtMMS21 is expressed in almost all plant tissues and has been shown to directly interact with AtSCE1a in onion epidermal cells. *mms21-1* presents reduced meristem activity and size indicating that AtMMS21 is actively involved in cell-cycle regulation during root meristem development (Huang et al. 2009). Cytokinin signalling stimulates the G1/S transition of the cell cycle. *mms21-1* presents a short root phenotype, downregulated levels of several cytokinin primary response regulators and a decreased sensitivity to exogenous cytokinin treatment suggesting that the root phenotype in *mms21-1* is due to an alteration in cytokinin signalling (Huang et al. 2009).

Although the effects of SUMOylation in plants are diverse, only a few examples suggest the role of SUMOylation in the function of the modified proteins (Miura et al. 2007; Miura and Hasegawa 2010). AtSIZ1, for example, is needed for the freezing tolerance response in Arabidopsis. Low temperatures induce the transcription of the transcription factor CBF3/DREB1A required to activate the expression of downstream genes involved in the low temperature tolerance response in the plant. CBF3/DREB1A is controlled in part by ICE1, a MYC transcription factor that is a target of ubiquitylation and SUMOylation. SUMOylation stabilizes ICE1 and inhibits the ubiquitin-mediate degradation, allowing the activation of CBF/DREB1, which subsequently activates the expression of cold responsive genes (Miura, et al. 2007).

1.2.1 ESD4, a SUMO specific protease

ESD4 is composed of nine exons and eight introns, and is located on chromosome IV of the Arabidopsis genome. *ESD4* is ubiquitously expressed in roots, flowers, siliques, as well as in rosette and cauline leaves (Murtas et al. 2003). *ESD4* encodes a SUMO protease of 489 amino acids that is localized predominantly at the periphery of the nucleus and regulates the abundance of SUMO conjugates in Arabidopsis (Murtas et al. 2003). The *esd4-1* mutant presents an increased level of SUMO-conjugates and reduced levels of free SUMO; the mutant plant presents an extremely early-flowering phenotype besides several developmental alterations like distorted phyllotaxy, expansion of the silique and a decrease in stature (Reeves et al. 2002; Murtas et al. 2003).

According to *in vitro* and *in vivo* studies, ESD4 acts as a major isopeptidase that recycles SUMO from conjugates in Arabidopsis (Murtas et al. 2003; Chosed et al. 2006). As mentioned before Arabidopsis contains several SUMO-specific proteases of which at least four have been functionally characterized. All these SUMO peptidases present sequence similarity within the catalytic core but several differences within the amino-terminal regions. The N-terminal region seems responsible for the specificity of SUMO protein conjugate recognition and for modulation of enzymatic activity (Chosed et al. 2006; Miura et al. 2007). In yeast the distinct domains of Ulp1 are required for localization and substrate specificity (Li and Hochstrasser 2003). The functionally characterized SUMO proteases in Arabidopsis are classified as ULP1

type and include ELS1, ESD4, OTS1 and OTS2 (Murtas et al. 2003; Mukhopadhyay and Dasso 2007; Conti et al. 2008; Hermkes et al. 2011). ELS1 is the closest homologue known of ESD4 (Hermkes, Fu, et al. 2011). It shares 65% amino acid sequence identity and presents several functional similarities like the requirement of the N-terminal regulatory domain for peptidase activity *in vitro* and the ability to remove SUMO1 and SUMO2 from conjugates (Hermkes et al. 2011). ELS1 also removes SUMO3 as well as other non plant SUMOs from conjugates, making this protease more promiscuous than other ULPs (Chosed et al. 2006). Based on the strong mutant phenotype of *esd4-1*, it is clear that neither ELS1 nor other isopeptidases present in the plant complement ESD4 function.

ESD4 localizes at the Nuclear Pore Complex

The Nuclear Pore Complex (NPC) is a macromolecular structure formed by several proteins called nucleoporins (Nups). Nups localize at the nuclear envelope and selectively connect the cytoplasm with the nucleoplasm (Xu et al. 2007; Palancade and Doye 2008). Besides the regulation of the nucleocytoplasmic traffic, the NPC participates in cell division, DNA repair, DNA replication and mRNA quality control (Palancade and Doye 2008). The NPC can be structurally divided into three elements: a nuclear basket, a central pore, and cytoplasmic fibrils (Xu et al. 2007). Analysis of the ESD4:GFP recombinant protein indicated that ESD4 localizes at the NPC (Murtas et al. 2003; Xu et al. 2007; Palancade and Doye 2008; Xu et al. 2009). Interestingly, mutants of a protein that forms part of the nuclear basket, Nuclear Pore Anchor (NUA), phenocopy the *esd4* mutant in several aspects including an increase in SUMO conjugates, reduction of free SUMO, accumulation of nuclear poly(A)⁺RNA and alteration of the expression of flowering time genes (Xu et al. 2007). NUA and ESD4 proteins interacted in a yeast two-hybrid screen, however the localization of ESD4 was not affected in *nua* mutants indicating that other proteins are involved in tethering ESD4 to the nuclear envelope (Xu et al. 2007). However the characterization of the NPC components in plants is still in its infancy (Zhao and Meier 2011) and therefore ESD4 interacting partners have not been identified. Nevertheless, until now characterized loss of function Nups mutants (namely Nup62, Nup133, Nup96, Nup160, Nup88 and Tpr) have shown a dwarf and early flowering phenotype (Zhao and Meier 2011), very similar to the one of the *esd4* mutant, suggesting the existence

of a common underlying basis to these phenotypes. Remarkably, the *esd4* and *siz1* mutants accumulate nuclear mRNA indicating that alterations in SUMO homeostasis impair the mRNA export pathway (Muthuswamy and Meier 2011). Whether Nups are targets of SUMOylation or function as anchors of the SUMO machinery components remains to be explored. At least the human SENP2 and the budding yeast Ulp1 interact with Nups at the NPC (Li and Hochstrasser 2003; Palancade and Doye 2008).

1.2.2 SUMOylation and floral transition

Several SUMO machinery mutants present an early-flowering phenotype including *siz1*, *ost1*, *ost2* and *esd4* indicating that the imbalance of SUMO-levels in the plant has a direct impact on this important developmental transition (Reeves et al. 2002; Jin et al. 2008; Conti et al. 2008; Jin and Hasegawa 2008; Lois 2010). As mentioned in the previous section, one of the causes of this phenotype might be the deregulation of the mRNA export pathway, which apparently has a direct effect on the regulation of the time to flower.

In Arabidopsis six genetically defined pathways control the flowering time, these include the photoperiodic, vernalization and ambient temperature pathways, which respond to environmental cues; and the gibberellin, autonomous and aging pathways, which involve the integration of endogenous cues (Fornara et al. 2010).

The photoperiodic pathway integrates and promotes flowering in response to changes in day length. The core of the photoperiodic pathway is composed by three genes: *GIGANTEA (GI)*, *CONSTANS (CO)* and *FLOWERING LOCUS T (FT)* (Turck et al. 2008), all of which act in the leaves. *GI* encodes a protein that does not show significant homology to other proteins. *CO* encodes a zinc finger transcription factor that is essential for the photoperiod-dependent induction of flowering (Putterill et al. 1995). *CO* is tightly regulated at the transcriptional and translational levels by light. Its mRNA is increased at the end of a long day and its protein is stabilized in light (Turck et al. 2008). *FT* encodes a protein similar to the RAF-kinase inhibitor-like protein (RKIP) and phosphatidylethanolamine binding protein (PEBP), and is a

potent floral promotor (Ahn et al. 2006). FT is part of the long distance signal that induces flowering by being transported from the leaves to the shoot apical meristem (Turck et al. 2008). Another important floral promotor is *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 (SOC1)* which not only integrates the signals from the photoperiodic pathway but also inputs from the gibberellin and vernalization pathways (Lee et al. 2000; Onouchi et al. 2000; Samach et al. 2000; Moon et al. 2003; Lee and Lee 2010).

The vernalization pathway accelerates flowering in response to prolonged periods of cold. The low temperatures induce the expression of *VERNALIZATION INSENSITIVE 3 (VIN3)* which interacts with components of the Polycomb-group Repressive Complex 2 (PRC2). VIN3 and PRC2 repress the expression of *FLOWERING LOCUS C (FLC)*, which encodes a MADS-box transcription factor that is a strong floral repressor (Boss et al. 2004; Srikanth and Schmid 2011).

The gibberellin pathway refers to the plant hormone gibberellin which induces the floral transition under non-floral inductive photoperiods. Mutations in the *GAREQUIRING 1 (GAI)* gene initially defined this effect. *GAI* encodes an ent-copalyl diphosphate synthetase enzyme that performs the first step in the biochemical synthesis of gibberellins. *gai* mutants flower later than wild type plants under LD but never flower under SD unless gibberellin is exogenously applied, indicating that this hormone plays an important role in the promotion of flowering under non-floral inductive photoperiods (Wilson et al. 1992; Reeves and Coupland, 2001).

The autonomous pathway refers to the genes whose proteins regulate the abundance of the *FLC* mRNA by various mechanisms (Boss et al. 2004). This pathway includes seven genes, loss-of-function mutants for four of them (*fca*, *fve*, *fy* and *fpa*) flower late independently of the photoperiod and respond to vernalization treatment (Koornneef et al. 1991; Boss et al. 2004). These genes form the core of the autonomous pathway, and are involved in different aspects of RNA metabolism or chromatin regulation (Amasino and Michaels, 2010).

More recently a plant-age flowering dependent pathway has been proposed, this pathway regulates developmental phase transitions by keeping a fine regulation of SQUAMOSA PROMOTER BINDING-LIKE (SPL) transcription factor expression mediated by a balance between two miRNA, *miR-156* and *miR-172* (Srikanth and Schmid 2011).

The *esd4* mutant was recovered in a genetic screen looking for plants that flowered early under short day conditions (Reeves et al. 2002). The *esd4* early flowering phenotype can be partly explained by the transcript levels of *FLC* being reduced in the mutant while the expression of the flowering-time promoter genes *SOC1* and *FT* are upregulated (Reeves et al. 2002).

FLC is a central floral repressor for the determination of flowering time in *Arabidopsis* (Sheldon et al. 1999; Desterro et al. 1999; Sheldon et al. 2000). *FLC* expression is tightly regulated by different chromatin states, some of which depend on the developmental state of the plant and are included in the autonomous pathway while others are influenced by environmental signals particularly vernalization (He 2009). Interestingly, *FLC* has recently been found to interact with both *AtSCE* and *ESD4* in the yeast two-hybrid assay and is SUMOylated *in vitro* (Elrouby, pers. Comm.), suggesting that *FLC* is a target of SUMOylation, however the regulation of this modification *in vivo* and its functional consequences remain to be explored

Several genes are involved in the regulation of flowering, among them *FWA* and *FLOWERING LOCUS D (FLD)* are discussed below in more detail as they have been related to the SUMOylation process (Reeves et al. 2002; Jin et al. 2008).

fwa is a semidominant mutation related to the long-day pathway (Koornneef et al. 1991; Koornneef et al. 1998). In previous work a mutant allele of *fwa-1*, which promotes an intermediate delay in flowering under LD conditions, was crossed to *esd4-1* to test for epistasis (Koornneef et al. 1991; Reeves et al. 2002). The double mutant *esd4-1 fwa-1* showed a clear recovery in several of the pleiotropic aspects of *esd4-1* including the time to flower (measured as the total leaf production at the time

of bolting) and phyllotaxy (Reeves et al. 2002). Interestingly, the number of leaves produced by the *esd4-1 fwa-1* double mutant was exactly the same as produced by the Wt under LD and SD conditions indicating that the *esd4* early flowering effects are strongly suppressed by *fwa-1* (Reeves et al. 2002). In Wt plants under SD conditions the *fwa-1* mutation has no effect on the time to flower indicating that in *esd4-1 fwa-1* has more effect in SD than in a Wt (Reeves et al. 2002). The *fwa* mutants are epialleles since they do not present changes in the nucleotide sequence of the *FWA* locus itself but instead a strong reduction in cytosine methylation across a region that includes the *FWA* promoter and the first two exons (Soppe et al. 2000). This cytosine methylation reduction allows the ectopic expression of *FWA*, the ectopically expressed FWA protein is able to interact with FT impeding its proper function in the Shoot Apical Meristem (SAM), therefore the *fwa* mutant is in reality a gain-of-function mutant (Soppe et al. 2000; Ikeda et al. 2007).

FLD encodes the plant ortholog of the mammalian KIAA0601/Lysine-Specific Demethylase 1 (LSD1), a component of the histone deacetylase 1,2 (HDAC1/2) co-repressor complexes, and is involved in H3K4 demethylation and deacetylation of *FLC* chromatin (He et al. 2003). A recent report suggests that FLD is target of SUMOylation mediated by SIZ1, and that the SUMOylated FLD protein is less efficient in repressing *FLC* expression by deacetylating the H4 in *FLC* chromatin (Jin et al. 2008).

1.2.3 SVP, a strong floral repressor

SHORT VEGETATIVE PHASE (SVP) is a member of the MADS-box family of transcription factors, MADS-box genes are considered master regulators of plant development. SVP is involved in the regulation of phase transitions in Arabidopsis in a dosage dependent manner (Hartmann et al. 2000). These developmental transitions include the flowering time transition and the maintenance of the identity of the floral meristem before the establishment of the identity of the floral whorls (Hartmann et al. 2000; Gregis et al. 2006; Liu et al. 2007; Gregis et al. 2008).

SVP interacts with FLC to form a floral repressor complex which binds the promoter regions of *FT* and *SOC1* inhibiting their expression (Michaels and Amasino 1999; Searle et al. 2006; Lee et al. 2007; Li et al. 2008). SVP is regulated by the autonomous, thermosensory and gibberellin pathways (Li et al. 2008), while FLC expression is controlled mainly by the vernalization and the autonomous pathways (Michaels and Amasino, 1999; Sheldon et al. 1999). If SVP and FLC interact SVP therefore acts as a repressor of flowering and its function is closely related to that of FLC.

As previously mentioned, SVP is not only involved in flowering time transition but it also prevents the precocious expression of floral organ identity genes mainly by repressing the expression of *SEPALLATA3* (Liu et al. 2009). SVP interacts with TERMINAL FLOWER 2/LIKE HETEROCHROMATIN PROTEIN 1 (TFL2/LHP1), a chromatin factor able to modulate trimethylation of histone H3 lysine 27 (H3K27me3), and apparently this interaction promotes a nonpermissive chromatin environment for *SEP3* expression (Liu et al. 2009).

The SVP capacity to bind to the regulatory sequences of *SOC1* and *FT* is partially modulated in the nucleus by *DNAJ HOMOLOG 3* (*J3*), a chaperone protein that mediates the integration of floral promoting signals (Shen et al. 2011). Apparently, *J3* perceives flowering signals from the photoperiodic, vernalization and gibberellin pathways and is able to regulate the expression of *SOC1* and *FT* (Shen et al. 2011).

How the interactions of SVP with proteins like FLC, *J3* or TFL2 are being regulated and which are the mechanisms that promote or inhibit such interactions are questions that remain to be answered.

1.3. Aim of the Study

The major goal of the thesis is a better understanding of the effects of SUMOylation on the regulation of floral transition in Arabidopsis. For this purpose two main approaches were taken, one is a genetic approach focused on the isolation of suppressors of *esd4*. The second approach is molecular based focused on the study of the role of SUMOylation in SVP function.

Identification of suppressors of *esd4*

In Arabidopsis, the alteration of SUMO-conjugate levels have a strong impact on plant development. Particularly, abolition of the expression of *ESD4*, the gene that encodes the major SUMO protease in Arabidopsis, causes severe phenotypic alterations and extreme early flowering. The aim of this project was to isolate and characterize mutants that suppress the early-flowering phenotype of *esd4-1* using a classical forward genetic approach combined with Next-Generation Sequencing. The isolation of suppressors of *esd4* will help to better understand the molecular mechanisms by which SUMOylation participates in the control of flowering time in *Arabidopsis thaliana* and how these are deregulated in *esd4*.

The role of SUMOylation in SVP function

In most cases, SUMOylation has a negative effect on transcription mainly by recruiting or stabilizing repressor complexes. SVP interacts with FLC to form a floral repressor complex that binds the promoter regions of *FT* and *SOC1* inhibiting their expression. Current research in our lab demonstrates that FLC interacts with both AtSCE and ESD4 in the yeast two-hybrid assay and is SUMOylated *in vitro* (Elrouby, pers. Comm.). The aim of this project is to determine if SVP is also SUMOylated and analyze the effects of this modification on SVP function by the analysis of transgenic plants expressing constructs aiming to hyperSUMOylate (translational fusions with SUMO or AtSCE) or hypoSUMOylate (mutations in the putative SUMO-attachment sites) SVP protein. The molecular characterization of the effects of SUMO on SVP will help to understand whether this modification influences the activity of this floral repressor to regulate the transition to flowering in Arabidopsis.

2. Isolation and characterization of *suppressors of esd4 (sed)*

2.1. Introduction

Post-translational modification of proteins by SUMOylation is involved in the regulation of several important cellular processes (Johnson 2004). The attachment of SUMO to a substrate protein is a reversible and dynamic process that is controlled by the action of SUMO conjugation enzymes (E1, E2, and E3) and SUMO proteases (Mukhopadhyay and Dasso 2007; Kim and Baek 2009). SUMO proteases control SUMOylation rates in two different ways; by processing SUMO precursor proteins to remove a carboxyl-terminal extension (processing or maturation), and by removing SUMO from the modified conjugates (deconjugation) (Mukhopadhyay and Dasso 2007; Kim and Baek 2009). The maturation step is important to expose a di-glycine motif required for conjugation to substrate proteins (Johnson 2004). By performing these maturation and deconjugation functions, SUMO proteases maintain balanced levels of SUMO conjugates in the cell, which is necessary for proper cellular function and development of organisms (Li and Hochstrasser 2003; Murtas et al. 2003). The importance of SUMO proteases in the regulation of the SUMO cycle and its impact on cell function makes them important candidates for study of the regulation of the SUMOylation process and as tools for the isolation of SUMO substrates.

Early in Short Day 4 (ESD4) encodes a SUMO-specific protease that regulates the levels of SUMO conjugates in Arabidopsis (Murtas et al. 2003). The *esd4-1* mutant was isolated as part of a screen after gamma irradiation of seeds to identify plants that flowered early under short-day (SD) conditions. The *esd4-1* mutation is caused by a deletion of 762 bp that removed a part of the 5' untranslated region and the first 30 bp of the coding region, leading to a null mutant (Murtas et al. 2003). The *esd4-1* mutant plants exhibit an extreme early flowering phenotype and a suite of strong developmental alterations, including a reduction in size, general reduction in number and size of all types of leaves, fewer flowers, short club-shaped siliques, and an alteration of phyllotaxy (Reeves et al. 2002). Interestingly, the strong phenotypic alterations presented by *esd4* mutant have not been observed in other Arabidopsis SUMO-proteases mutants (Conti et al. 2008; Hermkes et al. 2011), where only specific phenotypes have been observed. Additionally, *esd4* mutant plants

accumulate SUMO conjugates to high levels, and overexpression of free SUMO in the mutant does not restore any of its developmental alterations suggesting that the accumulation of SUMO conjugates (due to impaired ESD4 function) is the cause of the *esd4* phenotypes (Murtas et al. 2003). Together, the phenotypic and molecular evidence suggests that ESD4 acts as the major SUMO protease in Arabidopsis (Murtas et al. 2003)

Previous studies established the genetic interactions between *ESD4* and some of the genes that promote flowering in response to photoperiodic signals (Reeves et al. 2002), including *FLOWERING LOCUS T (FT)*, *CONSTANS (CO)* and *FLOWERING WAGENINGEN A (FWA)* (Koornneef et al. 1991; Turck et al. 2008). These studies established that the *esd4* mutant early flowering phenotype and some of its pleiotropic effects can be partially suppressed by *ft-1*, *co-2* and *fwa-1* so the mutations delay flowering of *esd4* mutant under short day photoperiods but were not fully epistatic to it (Reeves et al. 2002). Nevertheless the photoperiodic pathway mutants *ft-1* and *co-2* do not normally have a phenotype under SDs, so the hypersumoylated protein pool condition somehow makes the plant less photoperiodic responsive at least in part by causing the *CO* and *FT* genes to promote flowering under SDs.

In Arabidopsis, the misregulation of the SUMO cycle (by impairing the function of ESD4) causes early flowering. To investigate the role of SUMOylation in flowering time control, we sought to isolate and identify second-site mutations that would suppress the early flowering phenotype of *esd4*. These suppressors of *esd4* (or *sed*) may help identify genes involved in flowering time control and/or whose protein products are either SUMO substrates or regulate the SUMO cycle. In this chapter, we describe the isolation of 120 *sed* mutants that restored different aspects of the *esd4* mutant phenotype. Of these, 15 *sed* mutants were characterized further. These mutants had the strongest and most consistent phenotypes. We also report the morphological characterization of these mutants as well as analysis of the expression of SUMO-machinery genes and SUMO-conjugate levels.

2.2. EMS mutagenesis experiment

Mutant screens are particularly useful to assign functions to genes whose functions are poorly understood or correlate already described genes with new processes (Page and Grossniklaus 2002). These mutant screens help to better understand complex biological processes like developmental regulation and gene expression cascades (Kim et al. 2006).

A diverse variety of mutagens with different efficiencies and specificities exist, allowing a suitable mutagen to be selected for different experimental designs and purposes (Kim et al. 2006). Ethyl Methanesulfonate (EMS) is an extremely potent and specific chemical mutagen. EMS is an alkylating agent that induces single base-pair changes by chemical modification of nucleotides. More than 95 % of these modifications provoke guanine (G) to adenine (A) transition mutations. This is due to Guanine alkylation to O⁶_ethylguanine, an unusual base that preferentially pairs with thymine (T). During the DNA repair process the O⁶_ethylguanine is replaced with adenine (A) by the DNA repair machinery, giving rise to stable A:T base pairs (Kim et al. 2006).

EMS has several advantages so that it is often used in Arabidopsis mutagenesis experiments. First, mutations are randomly distributed. Second, chromosome break levels are low and therefore chromosomal rearrangements are very rare. Third, isolation of weak-allele mutants is high (around 65%). Such weak alleles are very useful in suppressor screens where the first mutation gives rise to a very sick phenotype (Greene et al. 2003; Kim et al. 2006).

In this study, 20,000 *esd4-1* seeds were mutagenized with 0.1% EMS to identify suppressors of *esd4-1* (or *sed*). M1 plants were grown under long day conditions and pooled into 15-20 plants per pool. M2 seeds from individual M1 pools were screened under short day conditions for plants that flowered later and/or looked healthier than the parental *esd4-1* plants. The heritability of the *sed* mutations was confirmed by growing M3 seeds under short day conditions. Suppressor mutant lines were assigned numbers that corresponded to the number of the pool batch.

2.3. Isolation of suppressors of *early in short days 4 (sed)*

M2 seeds from individual M1 pools (section 2.2) were sown on soil and screened under short day conditions for plants that flowered later and/or looked healthier than the parental *esd4-1* plants. From this screen a total of 120 independent suppressors were isolated. To confirm the inheritance and viability of these selected suppressor lines, around ten M3 plants from each of the 120 *sed* mutants were grown under long day (LD, 16-hours light/ 8-hours dark) and short day (SD, 16-hours light/ 8-hours dark) conditions, and their flowering times measured as the total number of leaves (rosette and cauline leaves) were compared with that of wild type and *esd4-1*. In general, none of the *sed* mutants flowered as late as wild type plants under either SD or LD conditions suggesting that none of the suppressor mutations completely suppressed the early flowering phenotype of *esd4-1*. Interestingly, more *sed* mutants flowered later than *esd4-1* under short day conditions (fifty *sed* mutants, Figure 2.1B) than under long day conditions (thirty *sed* mutants, Figure 2.1A). This is consistent with the fact that the early flowering phenotype of *esd4-1* is more severe in short days than in long days.

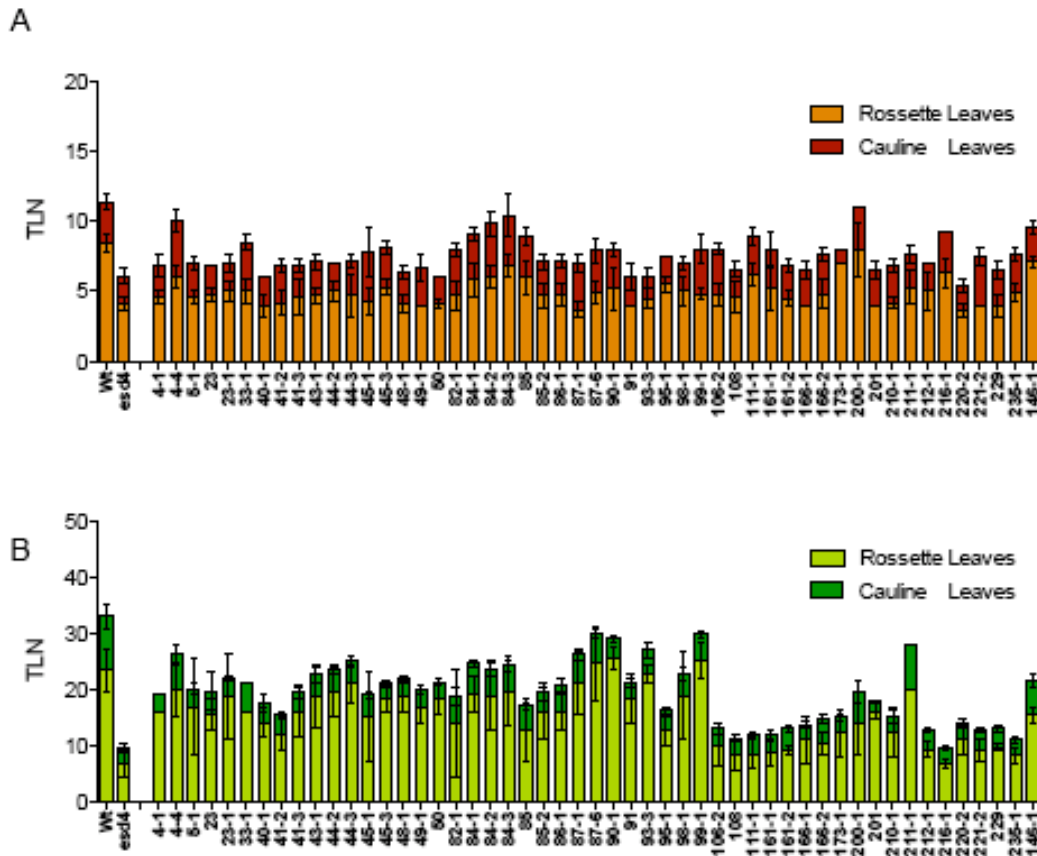


Figure 2.1. The *suppressors of esd4-1* exhibit a late flowering phenotype under Short Day (SD) and Long Day (LD) conditions. Flowering time, reported as the Total Leaf Number (TLN, Y axis) was registered for independent *sed* lines under LD (A) or SD (B) conditions. X axis show the genotype or identification number for each one of the different lines analyzed. Error bars illustrates the means \pm SE of $n \geq 8$ individuals.

2.4. Phenotypic characterization of selected *sed* mutants

The *esd4-1* mutant exhibits a pleiotropic phenotype which includes an extreme early flowering phenotype, a small stature, and a reduction in size of different organs like flowers, siliques, and cauline and rosette leaves (Reeves et al. 2002; Murtas et al. 2003)(Figure 2.2). This pleiotropic phenotype is likely to be caused by a general deregulation of the function of many proteins that accumulate in the form of SUMO-conjugates. In support of that, ESD4 was found to interact with hundreds of proteins involved in a wide variety of processes (Elrouby and Coupland 2010). It is likely that some of these proteins, and others not yet identified, may influence the different traits affected in *esd4*, including flowering time. Some of the isolated suppressors may

affect the same genes or pathways. Therefore morphological characterization of the *sed* mutants was carried out to determine whether they could be placed in subgroups that share similar phenotypes. These subgroups can afterwards be used for genetic tests including the establishment of allelism.

The fifty clear suppressor mutants that flowered later than *esd4-1* under SD conditions (Figure 2.1B) were further screened for suppression of some other aspects of the parental *esd4-1* phenotypes. For example, *sed* mutants with bigger rosette and cauline leaves, increase in stature and overall robustness were selected. As a result, fifteen independent *sed* mutants were chosen to be further characterized (Figure 2.3 and 2.4). In the following sections I will describe some of the phenotypes observed in these fifteen *sed* mutants.

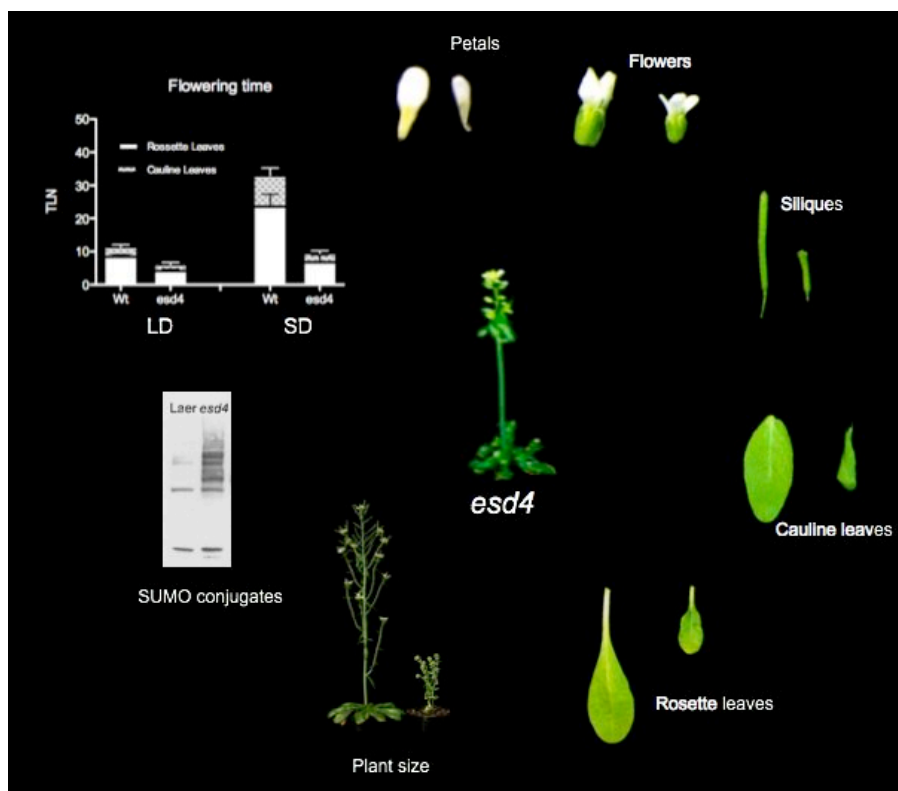


Figure 2.2. Phenotype of the *esd4-1* mutant. *esd4* mutant exhibits a pleiotropic phenotype that includes an early flowering phenotype, reduced floral organ size, small and club-shaped siliques, reduction of leaves and plant size, and an accumulation of SUMO conjugates. All phenotypes show the wild type (Wt) reference to the left and the *esd4* mutant to the right.



L er *esd4-1* *sed4-1*



L er *esd4-1* *sed5-1*



L er *esd4-1* *sed23-1*



L er *esd4-1* *sed33-1*



L er *esd4-1* *sed43-1*



L er *esd4-1* *sed44-3*



L er *esd4-1* *sed45-1*



L er *esd4-1* *sed48-1*

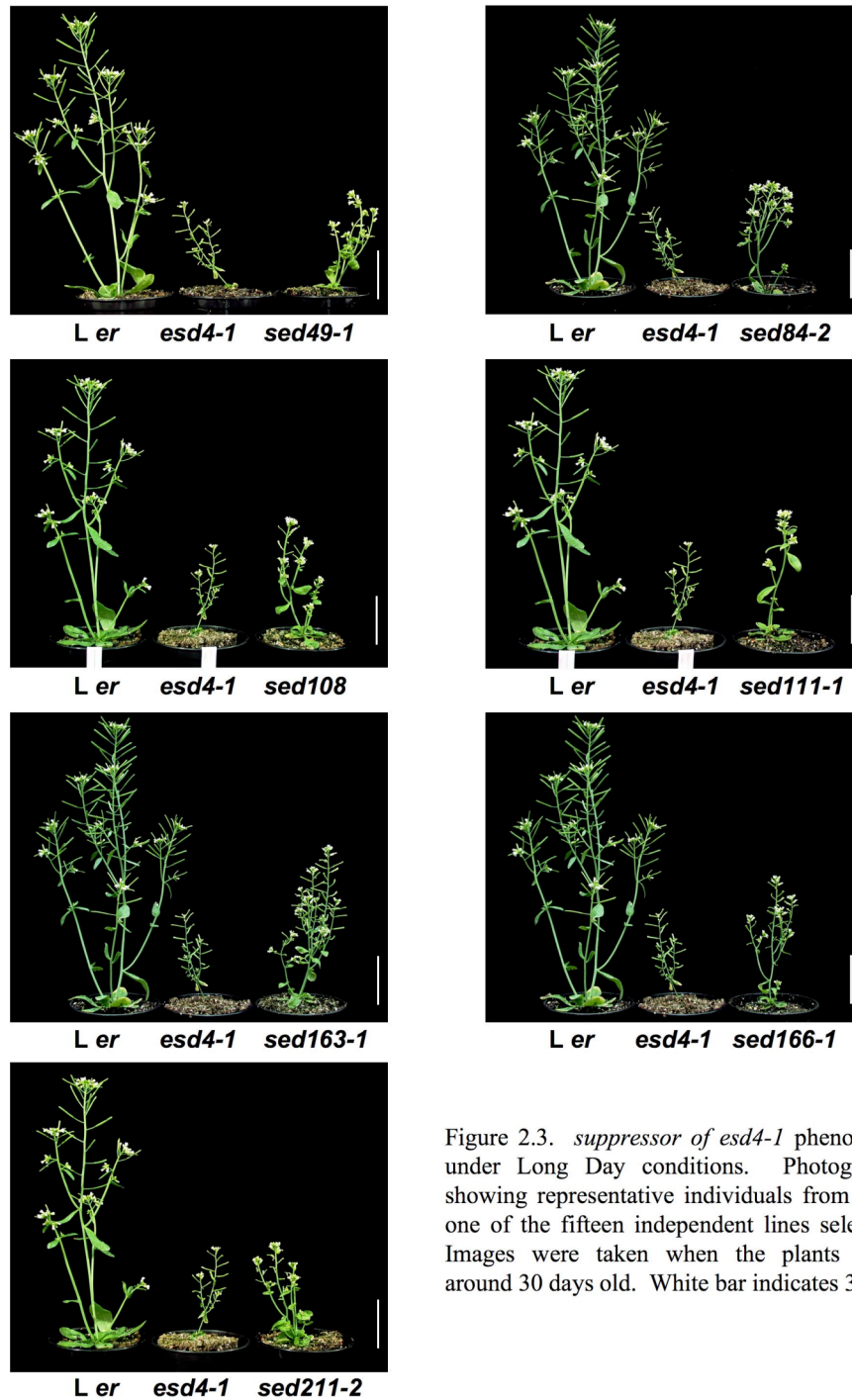


Figure 2.3. *suppressor of esd4-1* phenotypes under Long Day conditions. Photographs showing representative individuals from each one of the fifteen independent lines selected. Images were taken when the plants were around 30 days old. White bar indicates 3 cm.



L er esd4-1 sed4-1



L er esd4-1 sed5-1



L er esd4-1 sed23-1



L er esd4-1 sed33-1



L er esd4-1 sed43-1



L er esd4-1 sed44-3



L er esd4-1 sed45-1



L er esd4-1 sed48-1

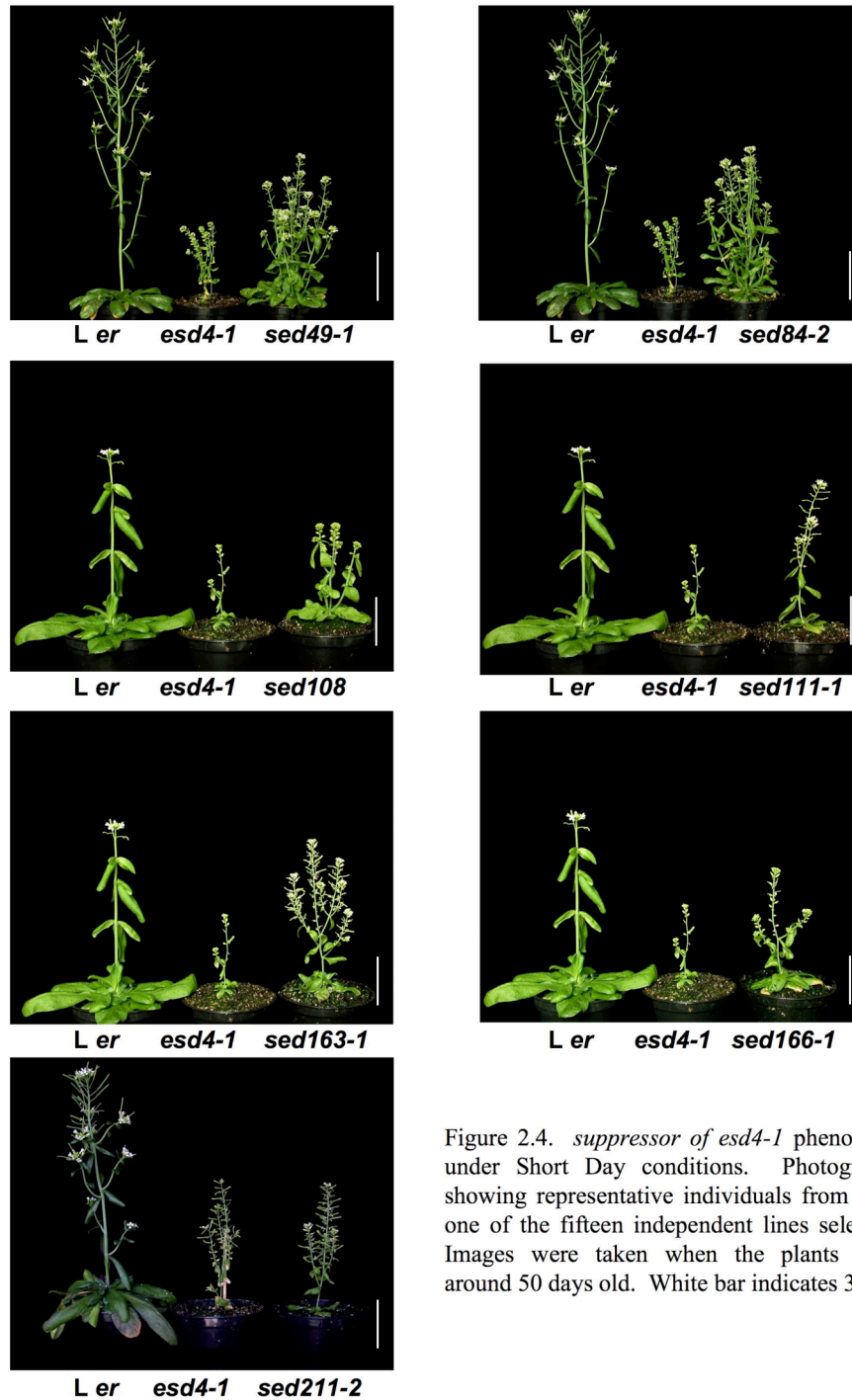


Figure 2.4. *suppressor of esd4-1* phenotypes under Short Day conditions. Photographs showing representative individuals from each one of the fifteen independent lines selected. Images were taken when the plants were around 50 days old. White bar indicates 3 cm.

2.4.1 Flowering time of *sed* mutants under LD and SD conditions

To confirm the flowering time phenotype of the fifteen independent *sed* lines a new experiment was performed under the same conditions (LD and SD). Consistent with the previous results, all 15 *sed* mutants flowered later than *esd4-1* under SD conditions while only a few of them (*sed111-1* and *sed84-2* and *sed33-1*) also flowered also later under LD conditions (Figure 2.5).

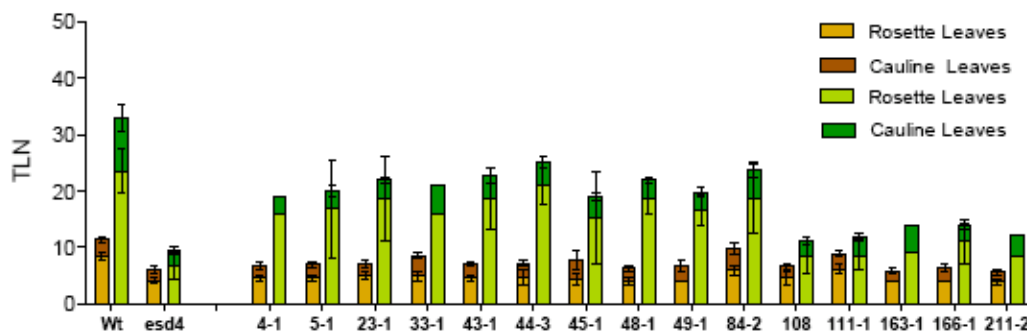


Figure 2.5. Flowering time of the fifteen selected *suppressors of esd4-1*. Flowering time, reported as the Total Leaf Number (TLN, Y axis) was registered for the 15 independent *sed* mutants under SD (green tones) and LD (orange tones) conditions. X axis show the identification number for each one of the *sed* mutants analyzed. Error bars illustrates the means \pm SE of $n \geq 10$ individuals.

2.4.2 Flower, siliques and leaf phenotypes

To determine if the *sed* mutants are able to suppress other phenotypic aspects of *esd4*, besides the early flowering time, I compared the general morphology and organ size of different structures like flowers, leaves and siliques among *sed* mutants, *esd4* and wild type plants.

Flowers

Besides a reduced size, *esd4* flowers are morphologically indistinguishable from those of wild type plants (Figures 2.2 and 2.6). To determine if *sed* mutants exhibited small flowers as *esd4*, twenty flowers (in recent anthesis) obtained from ten independent plants per genotype were collected and examined. This analysis revealed that *sed* flowers exhibit a regular number of floral organs when compared to the wild type and *esd4*. On the other hand, all *sed* flowers were as large or just slightly larger (*sed4-1* and *sed43-1*) than the *esd4-1* flowers, with the only exception of *sed44-3*, which

produce flowers as large as wild type (Figure 2.6).

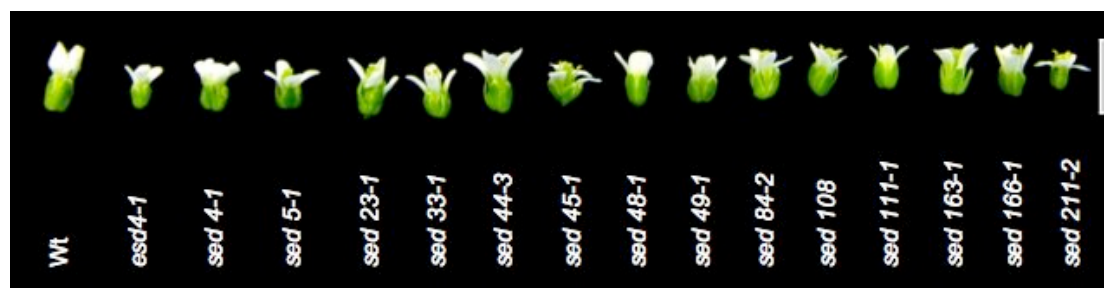


Figure 2.6. *suppressors of esd4-1* present normal developed flowers that are larger than the flowers of *esd4*. A representative flower from each *sed* mutant collected from plants grown under LD was photographed, genotypes are indicated below each flower. White bar to the left is 5 mm long.

Siliques

The siliques of *esd4-1* mutant plants are short and club-shaped (Reeves et al. 2002). To determine if these silique traits were conserved in the suppressors a comparison of six-day old post anthesis mature siliques (not abortive) was performed. This comparison showed that *seds* siliques were longer than those of *esd4* but still smaller than those of wild type. The only exception to this is *sed43-1*, which presented siliques that were even smaller than those of *esd4*. *sed43-1* siliques were mostly abortive (Figure 2.7).

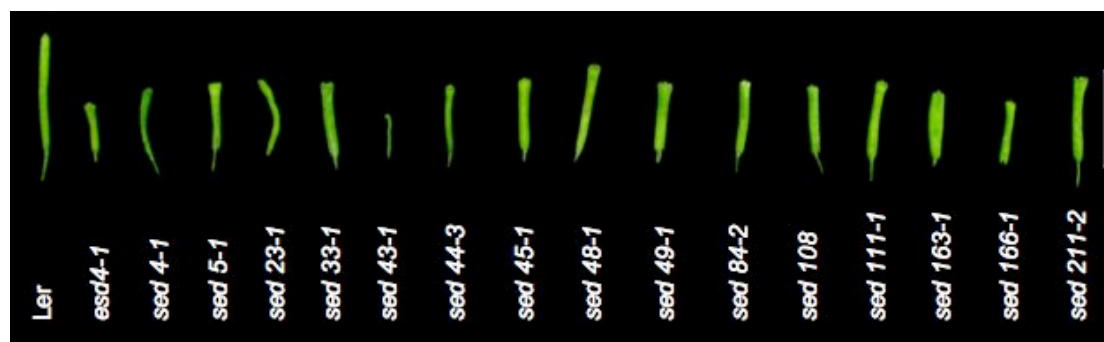


Figure 2.7. Silique phenotypes of the fifteen selected *suppressors of esd4-1* under LD conditions. White bar to the right represents 1 cm length.

The club-shaped silique of *esd4-1* was also observed in the *sed* mutants, but with a slight variation in the broadness of the silique tip (Figure 2.7). Siliques of *sed43-1* and *sed44-3* are narrow towards the basal part, a phenotype that correlates with lower seed production. In these mutants (*sed44-3* and *sed43-1*) a considerable number of basal siliques were aborted. This is particularly apparent in *sed43-1* where frequently we could not obtain any seeds. Interestingly, *sed43-1* is one of the strongest

suppressor mutants when other aspects of the *esd4* phenotype are considered e.g. it is late flowering (Figure 2.5) and exhibits larger leaves and is overall enhanced in vegetative growth (see below, Figures 2.8 and 2.9).

***suppressors of esd4-1* mutants show larger leaves**

In general, *sed* mutants showed larger leaves than *esd4-1* under LD (Figure 2.8A) and SD (Figure 2.8B) conditions. Rosette leaf morphology varied among the *sed* mutants, some of them presented a wavy shape, like *sed*s mutants *sed45-1*, *sed49-1*, *sed163-1* and *sed211-2*, or a downward curling of the leaf, like *sed111-1*. In some cases like *sed4-1*, *sed43-1*, *sed44-3* and *sed111-1* leaves tend to look paler than those of wild type or *esd4* (Figure 2.9). In the case of the cauline leaves most of them looked broader than wild type cauline leaves with twists across the long axis under SD (Figure 2.9 and 2.4, Data not shown).

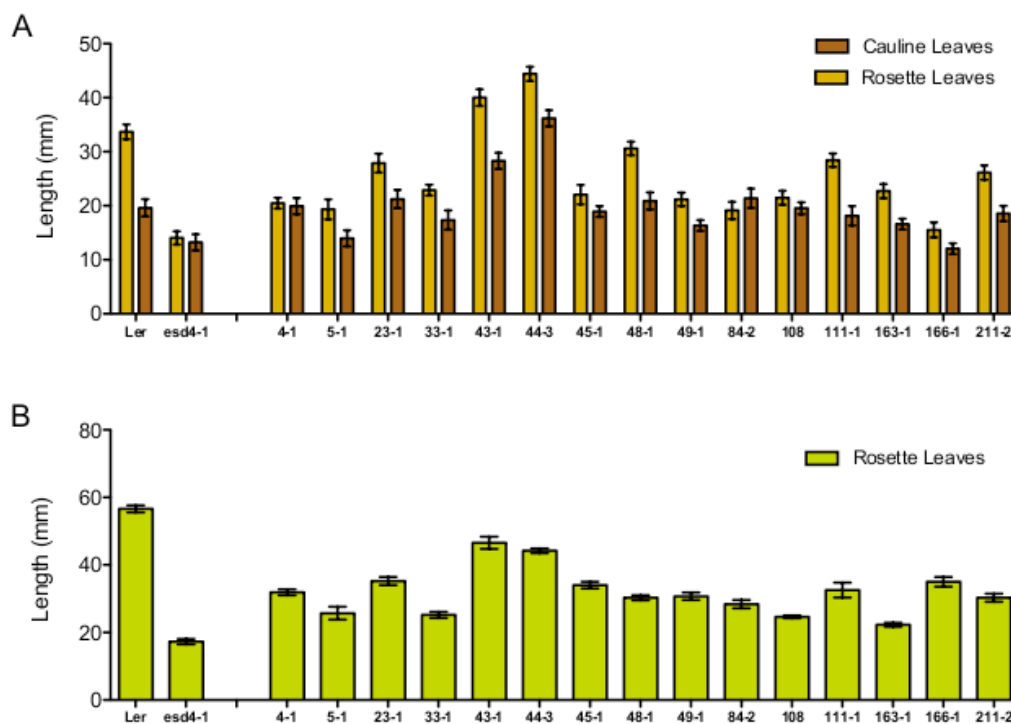


Figure 2.8. *suppressors of esd4-1* produce larger leaves than *esd4*. Leaf lengths of the fifteen selected *suppressors of esd4-1* under LD (A) and SD (B) conditions. The longest cauline and rosette leaves from ten independent individuals from wild type (*Ler*), *esd4-1* and each one of the suppressor mutants were measured after the first anthesis. Genotypes are indicated to the bottom of each column, for the *sed* mutants the identification number is indicated. Error bars represent the means \pm SE of $n \geq 10$.

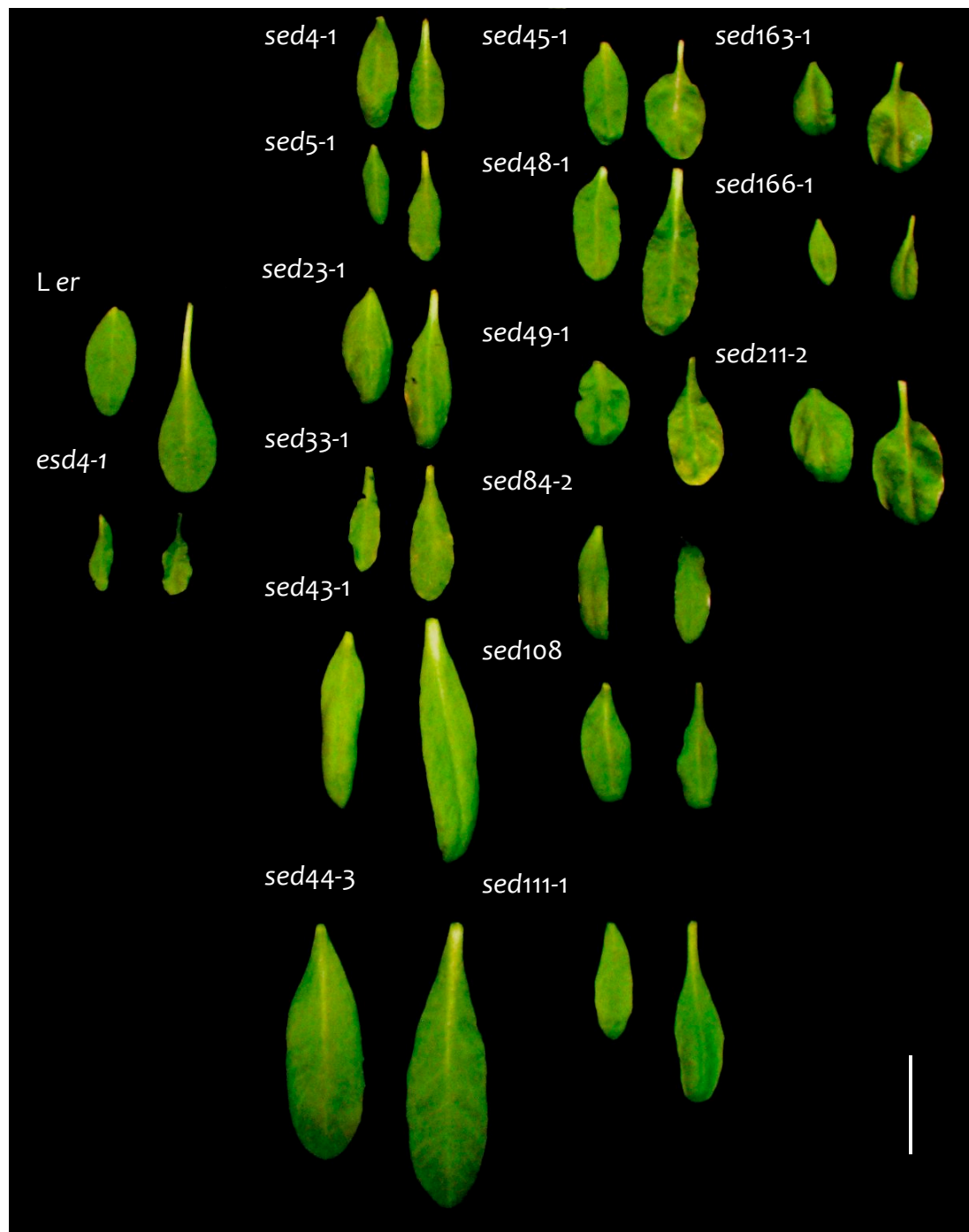


Figure 2.9. Photograph illustrating the *suppressors of esd4-1* leaf phenotypes under LD conditions. Representative Cauline (Left) and Rosette (Right) leaves from Wild type (*L er*), *esd4-1* and each one of the suppressor lines are presented. Genotypes are indicated to the upper left of each pair for clarity. The white bar represents 2 cm length as reference.

2.4.3 Pollen: viability and relative amount

Several factors can have an impact on plant seed production, including environmental conditions like temperature and water availability (Blum 2011; Bartrina et al. 2011). On the other hand, environmental stress conditions like high temperature and desiccation triggers responses that have a negative impact on seed yield (Mittler and Blumwald 2010; Blum 2011). SUMO-conjugate levels increase in plants subjected to different stresses like high temperatures and oxidative stress (Kurepa et al. 2003). Cells of *esd4* plants accumulate a pool of constitutively sumoylated proteins that could be interpreted as a “constitutively stressed” phenotype, maybe as a consequence *esd4* exhibits a low seed production among other phenotypes. *sed* mutants are larger and present a more robust aspect than *esd4* (see section 2.4.2), however, they seem more susceptible to environmental changes or stress. *sed43-1* and *sed44-3* mutants show a negative correlation of increased vegetative development with low seed yield, sometimes to the point of producing no seeds. As a consequence, the work with some *sed* lines has been hampered by low seed yield both in self and cross fertilizations.

Low *sed* mutant seed yield may also be caused by developmental defects in the formation of either the male or female gametes, or both. To determine if there was a difference in the pollen viability and/or pollen production between *sed* mutants and *esd4* or the wild type, a pollen viability test was performed. In brief, mature pollen from recently dehiscent anthers was extracted, distributed on a microscope slide and mounted in a solution of MTT (2,5-diphenyl tetrazolium bromide or thiazolyl blue), a chemical that is reduced to a purple compound (formazan) by metabolically active mitochondria (Khatun and Flowers 1995). Viable pollen, identified by the presence of purple strips, was counted under the microscope. In general, more than 80% of the total pollen produced by the *sed* mutants was viable. This value was also true for wild type plants (Figure 2.10A), indicating that pollen viability is not the cause of the low seed yield.

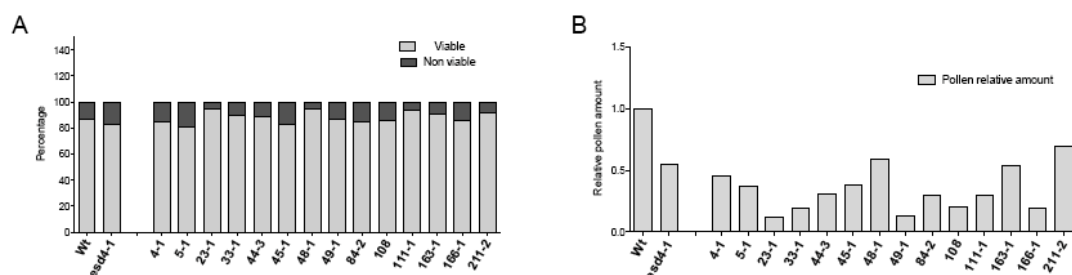


Figure 2.10. *suppressors of esd4-1* produce viable pollen but in lower amount than *esd4* under SD. (A) Pollen viability. Pollen of four flowers per genotype was mounted in a solution of MTT and counted. The bars indicate the total of viable (gray) or inviable (black) pollen per line as a percentage of the total pollen obtained in the same line. (B) Relative pollen amount. The total pollen of four flowers per genotype was counted. The percentage indicates total pollen in reference to the total pollen presented by the Wt. Y axis indicates the percentage. X axis indicates the genotypes analyzed.

Another possibility for low seed yield could be reduced pollen production, to assess this, the total amount of pollen produced by the anthers of four flowers per *sed* mutant was counted and compared to that produced by the wild type. In general *esd4* produces half the amount of pollen produced by the wild type plants, while the *sed* mutants produce less pollen than *esd4-1*, with the exception of *sed48-1*, *sed163-1* and *sed211-2* which produce equal or slightly more pollen than *esd4* (Figure 2.10B; Table 2.1).

Table 2.1. Pollen viability and relative amount in *sed* mutants under SD conditions

Genotype	Pollen			
	Viable ^a	Non viable	Total	Relative Amount ^b
Wt	1552	236	1788	1,00
<i>esd4-1</i>	816	168	984	0,55
<i>sed4-1</i>	700	128	828	0,46
<i>sed5-1</i>	532	128	660	0,37
<i>sed23-1</i>	202	12	214	0,12
<i>sed33-1</i>	310	35	345	0,19
<i>sed44-3</i>	494	62	556	0,31
<i>sed45-1</i>	556	115	671	0,38
<i>sed48-1</i>	996	55	1051	0,59
<i>sed49-1</i>	202	30	232	0,13
<i>sed84-2</i>	464	81	545	0,30
<i>sed108</i>	315	51	366	0,20
<i>sed111-1</i>	505	34	539	0,30
<i>sed163-1</i>	874	92	966	0,54
<i>sed166-1</i>	296	50	346	0,19
<i>sed211-2</i>	1154	104	1258	0,70

^aViable pollen indicates the total number of pollen grains per genotype that presented purple strips on its surface after MTT treatment.

^bRelative Amount indicates the rate of Total pollen produced by the *sed* in reference to the Total amount of pollen produced by the Wt, which was set as the maximum or 1.

2.5. Molecular characterization of selected *sed* mutants

2.5.1 Expression of SUMO pathway components

To test whether any of the *sed* mutants might be due to a mutation within one of the previously reported core SUMO-machinery genes (Kurepa et al. 2003) the expression of the full-length mRNA from each one of these genes (Table 2.2) was analyzed in the background of the fifteen isolated *sed* mutants by semi-quantitative reverse-transcription polymerase-chain reaction (RT-PCR). For this purpose the aerial parts of 15-day old seedlings grown in soil was collected for the analysis since all the SUMO machinery genes have been reported to be ubiquitously expressed (Saracco et al. 2007).

Protein	Gene	Accession number
SUMO	AtSUM1	At4g26840
	AtSUM2	At5g55160
	AtSUM3	At5g55170
	AtSUM5	At2g32760
E1	AtSAE1a	At4g24940
	AtSAE1b	At5g50580
	AtSAE2	At2g21470
E2	AtSCE	At3g57870
E3	AtSIZ1	At5g60410

SUMO

In Arabidopsis eight SUMO genes have been identified, but only four are consistently expressed (Kurepa et al. 2003; van den Burg et al. 2010; Castaño-Miquel et al. 2011). Previous reports have shown that SUMO isoforms are expressed in most of the tissues, with SUMO1 being the most transcriptionally active followed by SUMO2, whereas SUMO3 and SUMO5 are weakly expressed (Kurepa et al. 2003; Budhiraja et

al. 2009; van den Burg et al. 2010). Expression analysis of the SUMO isoforms in the fifteen *sed* mutants indicated that all SUMO genes are expressed, suggesting that none of the *sed* phenotypes are due to loss of the mRNA expression of these genes (Figure 2.11). Interestingly, while *esd4* contains the same SUMO mRNA levels as the wild type, the majority of the *sed*s showed a slight reduction in SUMO3 expression (Figure 2.11C). It is worth noting though that this gene was difficult to amplify and more PCR cycles were needed to detect it. On the other hand, a clear and significant increase in SUMO5 expression was detected in *sed43-1* and *sed44-3* (Figure 2.11D). These expression differences require further testing, maybe under different conditions, of how the expression of the different SUMO isoforms is induced in the *sed* mutant background. On the other hand it is also important to sequence these genes to exclude any possible point mutations that might alter gene expression or the encoded protein.

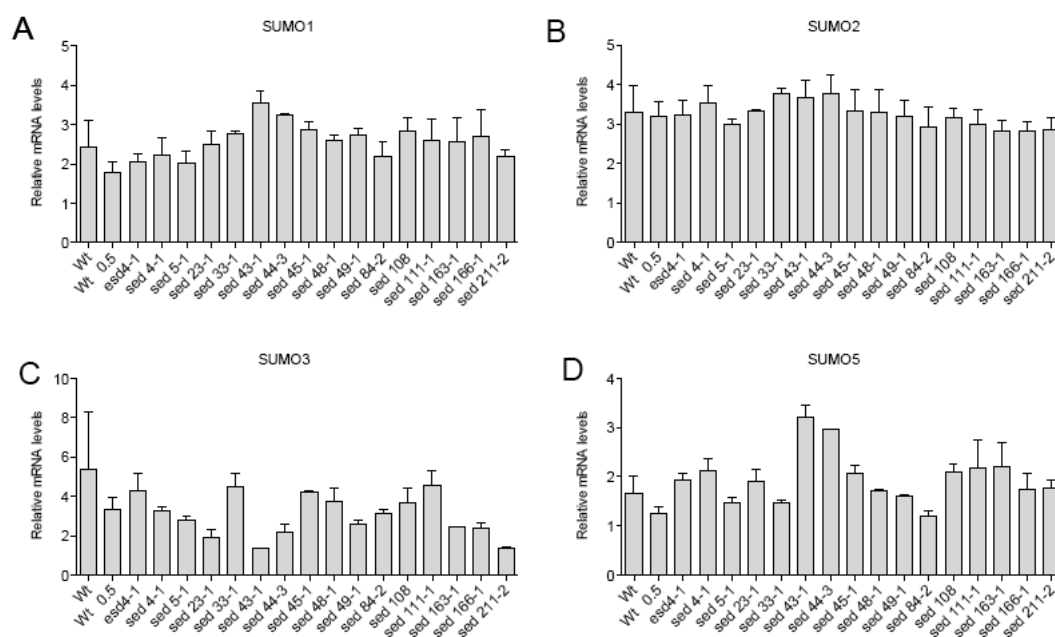


Figure 2.11. SUMO expression levels in *sed* mutants grown under LD conditions. The expression of SUMO1 (A), SUMO2 (B), SUMO3 (C) and SUMO5 (D) genes was analyzed in each *sed* line by RT-PCR. Primers used were specific for each SUMO and include the complete cDNA and small parts of the 5' and 3' UTR regions (see Appendix). Genotypes are indicated at the bottom of each column. Wt 0.5 indicates a 0.5 dilution of the Wt cDNA as fold control. RNA samples were isolated from 15-day-old seedlings grown in soil under long days (LD) and collected 12h after dawn. Actin was amplified and used as control to normalise the cDNA amounts. Data are based on one biological experiment. Results are presented as the mean of two repetitions \pm SE.

SUMO Activating Enzyme (SAE)

No difference in expression pattern of all subunits of the SAE (or E1) were detected when comparing *sed* mutants and wild type. As in wild type, and as previously reported the subunit SAE2 is less expressed in comparison to the SAE1a and SAE1b subunits (Kurepa et al. 2003; Saracco et al. 2007; Figure 2.12).

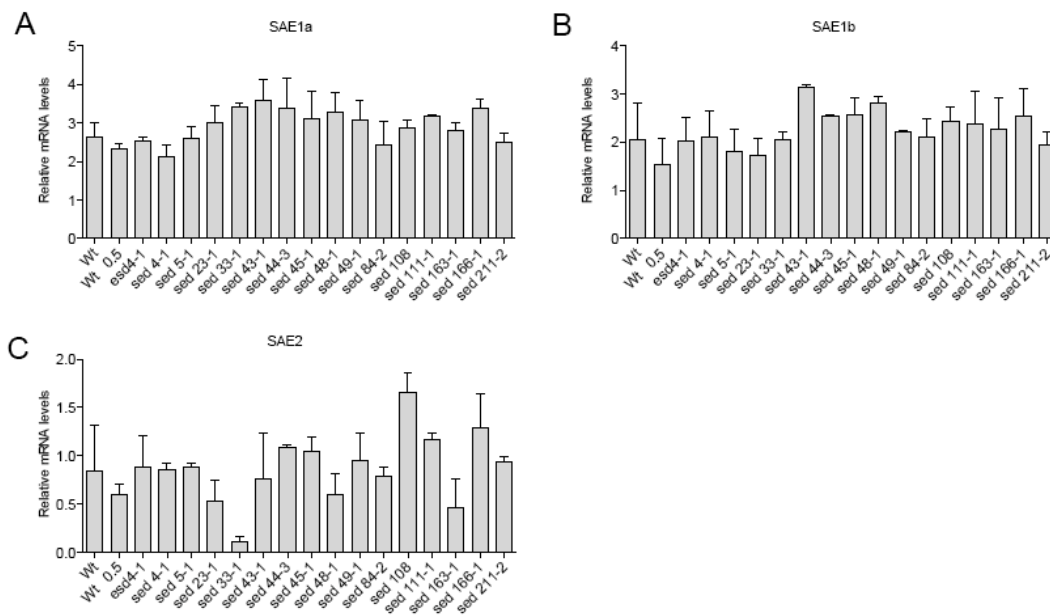


Figure 2.12. SUMO Activating Enzyme (SAE) subunits expression levels in *sed* mutants grown under LD conditions. The expression of SAE1a (A), SAE1b (B) and SAE2 (C) genes was analyzed in each *sed* line by RT-PCR. Primers used were specific for each SAE subunit and include the complete cDNA and small fragments of the 5' and 3' UTR regions (see Appendix). Genotypes are indicated to the bottom of each column. Wt 0.5 indicates a 0.5 dilution of the Wt cDNA as fold control. RNA samples were isolated from 15-day-old seedlings grown in soil under long days (LD) and collected 12h after dawn. Actin was amplified and used as control to normalise the cDNA amounts. Data are based on one biological experiment. Results are presented as mean of two repetitions after standardization with respect to Actin levels \pm SE.

SUMO Conjugating Enzyme (SCE) and SUMO E3 ligase SAP and MIZ1 (SIZ1)

The Arabidopsis genome contains only one SCE (or E2)-encoding gene that, when mutated, leads to a lethal phenotype (Saracco et al. 2007). On the other hand, it is predicted that the Arabidopsis genome contains several E3 SUMO ligases, while until now SIZ1 is the best characterized (Miura et al. 2009; Miura and Hasegawa 2010). Both SCE and SIZ1 are expressed in all *sed* mutants (Figure 2.13). *esd4* exhibits the

same expression levels of the SCE in comparison to the wild type while it is significantly more expressed in *sed44-3* (Figure 2.13A). On the other hand, SIZ1 expression levels show a tendency to be more expressed in *esd4* and all *sed* mutants in comparison to the Wt (Figure 2.13B and 2.13C).

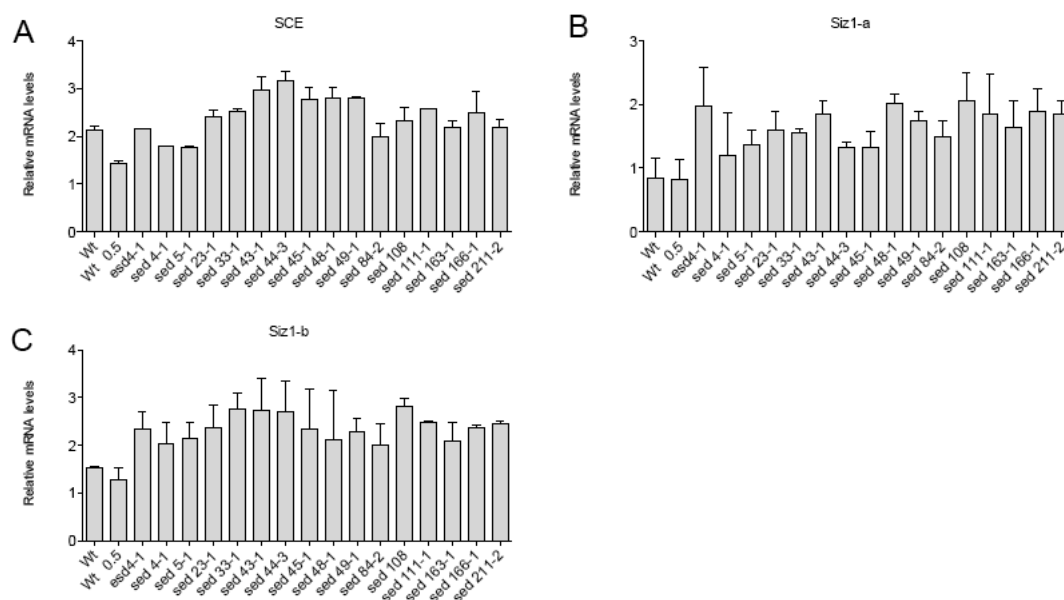


Figure 2.13. SUMO Conjugating Enzyme (SCE) and E3 Ligase Siz1 expression levels in *sed* mutants grown under LD conditions. The expression of SCE (A), SIZ1-a (B) and SIZ1-b (C) genes was analyzed in each *sed* line by RT-PCR. SIZ1-a and SIZ1-b represent only the amplification of different regions of the gene. Primers used were specific for each gene including the complete cDNA and small parts of the 5' and 3' UTR regions (see Appendix). Genotypes are indicated to the bottom of each column. Wt 0.5 indicates a 0.5 dilution of the Wt cDNA as fold control. RNA samples were isolated from 15-day-old seedlings grown in soil under long days (LD) and collected 12h after dawn. Actin was amplified and used as control to normalise the cDNA amounts. Data are based on one biological experiment. Results are presented as mean of two repetitions after standardization with respect to Actin levels \pm SE.

2.5.2 Protein SUMOylation levels

As described earlier, none of the *sed* mutations restores the *esd4-1* parental phenotype to wild type. Additionally, the different SUMO machinery genes are expressed in all fifteen *sed* mutants analyzed (Section 2.5.1). Therefore, the mutations

in the *sed* mutants were thought unlikely to cause a major change in the SUMO-conjugate pattern observed in *esd4-1*. To determine if the levels of SUMO-conjugates of the *sed* mutants change in comparison to the *esd4* mutant, crude protein extracts from 15-day old seedlings grown under LD from each one of the 15 *sed* mutants (plus *esd4-1* and wild type as controls) were extracted and analyzed with an anti-SUMO1-antibody. In general the levels of SUMO-conjugates in the *sed*s are more similar to those of *esd4-1* than to the wild type. This is in agreement with the morphological characterization of these mutants where none of them recovered to the wild type phenotype (Figures 2.3 and 2.4). However, some small differences can be appreciated, for example, *esd4* presents a clear strong band of approximately 55 kDa which is shared by the majority of the suppressors except for *sed4-1*, *sed5-1*, *sed33-1*, *sed84-2* and *sed166-1* where is much reduced. On the other hand *sed23-1* and *sed84-2* present extra bands at around 45 kDa and 60 kDa, respectively. These bands are more abundant in *sed23-1* and *sed84-2* than in either wild type or *esd4* plants (Figure 2.14).

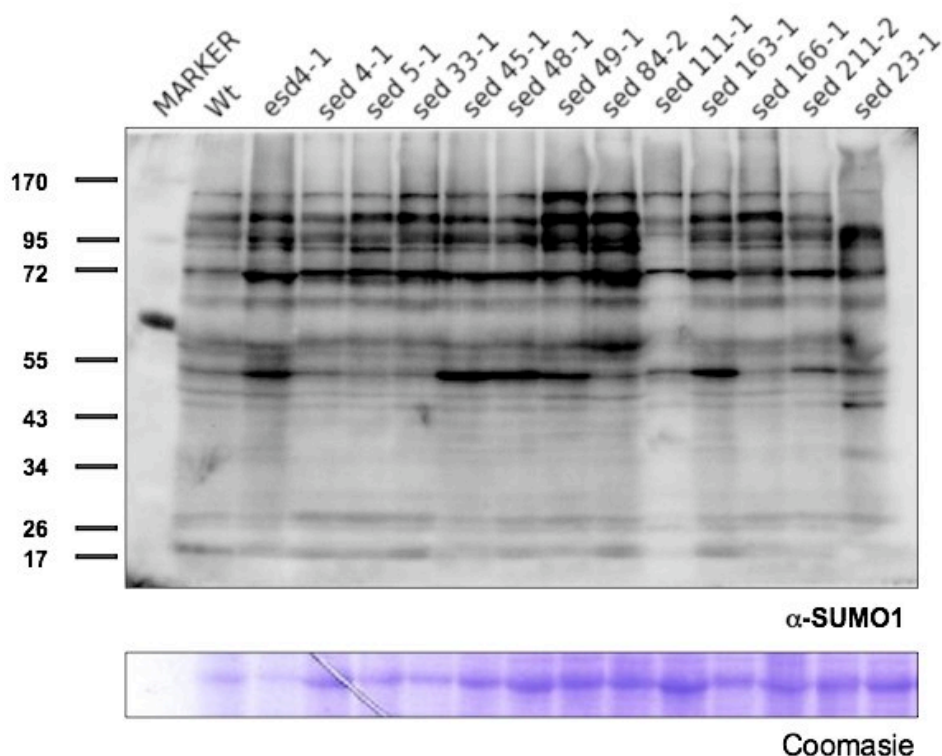


Figure 2.14. Analysis of SUMO-conjugate protein levels in the *sed* mutants under LD. SUMO conjugate levels were determined by Western-blot analysis with anti-SUMO1. Genotypes are indicated to the top of each column. Numbers to the left indicate Molecular Weight in kDa.

2.6. Discussion

ESD4 is the major SUMO protease in Arabidopsis. The *esd4* mutant is pleiotropic and flowers extremely early under both SD and LD conditions (Murtas et al 2002). Here I describe the first suppressor screen for *esd4*, focusing on the isolation of suppressors that reduce the extreme early flowering phenotype of this mutant. For this purpose *esd4* seeds were mutagenized with EMS and M2 populations visually screened for suppressors that flower later and look healthier than their *esd4* parents. In this way, 120 suppressors were isolated, these suppressors showed a wide range of phenotypes suggesting that they may be independent mutations at different loci that alleviate to varying degrees the *esd4* mutant phenotype.

Of the 120 suppressors, fifty partially suppressed the early flowering phenotype of *esd4* under SD conditions, and thirty out of these fifty also flowered late under LD conditions (Figure 2.1). This result suggests that in some cases the mechanisms that can bypass the *esd4* effects may be dependent on daylength. Alternatively, SD conditions may cause a degree of stress to Arabidopsis plants that normally grow in northern latitudes (with longer days), and these suppressors may alleviate some of the stress caused by both the first-site mutation (*esd4*) and SD conditions.

The *esd4* mutant offered several advantages and disadvantages for a second-site suppressor screen. Among the advantages are the strong developmental phenotype, namely the extreme early flowering time which allowed a relatively easy visual screen for later flowering plants. Additionally, the *esd4* mutation is a deletion of 762 bp, which eliminates the risk of isolating suppressor mutants caused by reversion at the *esd4* locus. Finally, the early flowering phenotype also allows for a rapid generation time. On the other hand, the disadvantages include a strong pleiotropic phenotype which does not allow for one suppressor mutation to completely restore the *esd4* parental phenotype to wild type. In addition, the *esd4* mutation causes a certain degree of stress in the plant, and hence, it has low tolerance to the manipulation and stress caused by crosses required for gene mapping- in many cases this results in reduced fertility.

Of the fifty late flowering suppressor of *esd4*, fifteen were selected based on their late flowering phenotypes (Figure 2.5), taking also into account a general partial recovery in overall health and viability. The morphological characterization of these suppressors indicated that they were able to alleviate other aspects of the pleiotropic *esd4* mutant phenotype (see Figure 2.2), in particular a general increase in size (Figures 2.3 and 2.4). In general *esd4* flowers, siliques and leaves are approximately half of the size of those of wild type plants. The suppressor mutants exhibit organs with intermediate sizes (Figures 2.6 to 2.9). This partial recovery towards a wild type phenotype in all the suppressors suggests that these mutations might act in several pathways in which *esd4* may function. Interestingly, the club-shaped silique phenotype was retained in the majority of the suppressors (Figure 2.7). The retention of the club-shaped silique phenotype has been previously observed in double mutants (like *esd4 co-2*, *esd4 ft-1* and *esd4 fwa-1*) which alleviate the early flowering phenotype of *esd4* under LD conditions (Reeves et al. 2002), indicating that the mechanism regulating this phenotype is independent of that regulating the time to flower. Two *sed* mutants (*sed43-1* and *sed44-3*), on the other hand, are remarkably the healthiest of all fifteen *sed* mutants as far as vegetative growth is involved. However, these mutants exhibit low fertility, small siliques and very low seed production.

All fifteen suppressor mutants analyzed express the full-length mRNA of each one of the previously reported core SUMO-machinery genes (Section 2.5.1), it remains to sequence these core SUMO-machinery genes to exclude the possibility of point mutations that might alter the encoded protein. Interestingly, null mutants of the SUMO machinery pathway genes like SUMO1 in combination with SUMO2, the SUMO-activating enzyme subunit SAE2 and the SCE, are embryo lethal; while mutants of the SUMO E3 ligases HIGH PLOIDY (HYP2), SAP and MIZ1 (SIZ1) show a dwarf phenotype as in *esd4* (Murtas et al. 2003; Miura et al. 2005; Ishida et al. 2009; Huang et al. 2009; van den Burg et al. 2010), indicating that at least for these genes the imbalance of SUMO levels affects viability or plant size.

In general, all the *sed* mutants exhibit wild type expression levels for the different SUMO machinery genes analyzed, with the exception of the SUMO ligase SIZ1, which appeared to be slightly increased in both the *sed* lines as well as their

progenitor *esd4-1* (Section 2.5.1 and Figure 2.13). This indicates that the partial recovery of the early flowering phenotype of *esd4-1* observed in *sed* mutants is unlikely due to a specific increase in SIZ mRNA expression. SIZ1 is a flowering time repressor that acts through two mechanisms: reduction of salicylic acid (SA) levels and SUMO modification of FLOWERING LOCUS D (FLD), a repressor of the strong floral repressor FLOWERING LOCUS C (FLC) (Jin et al. 2008).

Interestingly, *sed43-1* and *sed44-3* showed high levels of SUMO5 and low levels of SUMO3 mRNAs (Figure 2.11). SUMO5 was reported to be highly expressed in shoot tips and flowers, specifically carpels (Saracco et al. 2007). It is possible that this gene has a specific function in these tissue types which could affect flowering time. On the other hand, previous reports have shown that *sum3* mutants display a late flowering phenotype under SD conditions, and that attempts to overexpress SUMO3 were hampered by low expression levels of the transgene (van den Burg et al. 2010). It is possible that the low levels of SUMO3 observed in *sed43-1* and *sed44-3* are caused by the mutations in these lines, with the implication that the genes affected might act to regulate the expression levels of this gene (SUMO3), and possibly SUMO5.

sed33-1 has severely reduced levels of SAE2 mRNA (Figure 2.12). *sed33-1* also exhibits a low relative pollen production (Figure 2.10). Previous studies suggest that null mutants of SAE2 are embryo lethal suggesting that it is essential for growth and development (Saracco et al. 2007). Our results suggest that the mutation in *sed33-1* reduces expression levels of SAE2, which in turn may lead to reduced pollen production. Alternatively, the mutation in *sed33-1* may directly affect genes involved in pollen development.

The analysis of the SUMO-conjugate levels in the *sed* mutants showed that they are more similar to those of *esd4* rather than those of wild type. Interestingly *esd4* presents a clear strong band of approximately 55 kDa which is shared by the majority of the suppressors except for *sed4-1*, *sed5-1*, *sed33-1*, *sed84-2* and *sed166-1* where is much reduced. The identity of this band and why its level is reduced in these mutants is unclear. On the other hand *sed23-1* and *sed84-2* present extra bands at around 45 kDa and 60 kDa, respectively. These bands are more abundant in *sed23-1* and *sed84-*

2 than in either wild type or *esd4* plants. Despite these minor differences, it is noteworthy that none of the *sed* mutants characterized in this study completely restores SUMO-conjugate levels to those of wild type plants. This is in agreement with the morphological characterization of these mutants where none of them recovered to the wild type phenotype.

In this chapter I describe a suppressor screen to identify genes that are likely to be regulated by SUMO and may control floral transition in Arabidopsis. The morphological and molecular characterization of *sed* mutants show that in these mutants several phenotypes of the parental *esd4* were partially restored, including the early flowering phenotype and body and organ size. A major change in SUMO-conjugate levels was not observed though in the *sed* mutants. This result, together with the fact that all the SUMO-machinery components are expressed at the mRNA level, suggest that the *sed* mutations are not in genes that are part of the SUMO machinery. Although the identity of these genes is still to be determined, it is tempting to speculate that some of them may encode SUMO substrates or proteins that regulate their expression or activity.

3. Molecular Identification of *suppressors of esd4*

3.1. Introduction

Posttranslational modification of proteins is a common cellular mechanism to control the activity, stability and localization of target proteins, and therefore has a direct impact on protein function as well as potential protein-protein interactions (Johnson 2004; Walsh et al. 2005). SUMOylation is a reversible posttranslational modification involved in the control of several important cellular activities (Johnson 2004). SUMOylation occurs on a lysine residue usually located within a highly conserved motif called SUMO-attachment motif (Ψ KxE/D) where Ψ is a hydrophobic amino acid, K is a lysine, x any amino acid and E or D are acidic amino acids (Rodriguez et al. 2001). SUMO attachment to substrate proteins involves three enzymatic reactions referred to as activation, conjugation and ligation or E1, E2 and E3, respectively. SUMO is produced as a precursor protein, and its maturation (removal of a C-terminal extension) as well as the removal of SUMO from conjugates is performed by SUMO-specific proteases also known as Ubiquitin Like specific Proteases (ULP) (Li and Hochstrasser 1999). SUMO can also interact non-covalently with the proteins through SUMO-Interacting Motifs or SIMs, which are short-stretches of hydrophobic amino acids (Kerscher 2007). This non-covalent interaction provides another layer of complexity in the effects of SUMO modification on the proteins.

In plants, SUMOylation is important for proper development since mutants of the SUMO activating enzyme large subunit (E1, SAE2), SUMO conjugating enzyme (SCE or E2), or the double mutant *sumo1/sumo2* are embryonic lethal (Saracco et al. 2007; van den Burg et al. 2010; Lois 2010). Additionally, mutations in the SUMO-specific protease *ESD4* or the SUMO ligase *SIZ1* cause gross developmental and physiological defects (Murtas et al. 2003; Miura et al. 2005; Miura et al. 2007; Lee et al. 2007). Despite its importance, few substrates have been identified and characterized using traditional genetic approaches. These include PHR1, ICE1, FLD and ABI5 (Miura et al. 2007; Miura et al. 2007; Miura et al. 2009). Consequently, large-scale proteomic approaches have been taken to identify potential SUMO substrates including bioinformatic analysis, yeast two hybrid screens, biochemical purification and *in plant* screens based on 2-DE analysis (Miller et al. 2010; Elrouby

and Coupland 2010; Park et al. 2011). These screens have shown that SUMO substrates are involved in chromatin and genome stability, transcription, nucleic acid metabolism, protein translation, folding, stability, trafficking, metabolism, and stress response –mainly heat shock. Interestingly, among the hundreds of different proteins identified by these approaches, only a very small set is shared between them indicating that the experimental design affects the identity of SUMO substrates found. This notion supports the idea that SUMO is involved in the regulation of a variety of cellular processes and also that it can act on a specific pool of substrates according to the stimulus provided- or the mechanism under study.

Genetic screens help uncover parallel pathways, novel genetic interactions and molecular regulatory components, with the advantage of exploring these interactions and pathways directly *in planta*, avoiding the manipulation that *in vitro* studies require. In the previous chapter I described a suppressor screen to identify genes that are likely to be regulated by SUMO and may control floral transition in Arabidopsis. The morphological and molecular characterization of these *suppressors of esd4* (*sed*) mutations showed that they partially restored several phenotypes of *esd4*, including early flowering and body and organ size. Interestingly, no general change in the SUMO-conjugate levels nor in the expression of the major SUMO-machinery components at the mRNA level was observed in these lines suggesting that the *sed* mutations are not likely to be in genes that are part of the SUMO machinery. In this chapter I aimed to uncover the molecular basis for some of these *sed* mutants. Using classical map-based cloning I established rough map positions for four independent *sed* mutations. To speed up the cloning process I used Next Generation Sequencing to fine map one of these mutations, *sed111-1*. I could delimit the mutation to a region of chromosome that contains six candidate genes. Finally, T-DNA mutant lines for four potential candidates were crossed to the *esd4* mutant allele to test whether they confer the suppressor phenotype.

3.2. Generation of Mapping Populations

To identify the causative mutations of the different *sed* mutants (in *esd4-1*, accession Landsberg *erecta* – L *er*), mapping populations were established by crossing each independent *sed* mutant line with a heterozygous *esd4-2/+* (accession Columbia – Col; Figure 3.1A). *esd4-2* is a SALK line carrying a T-DNA insertion within the fourth intron of *ESD4* and has been previously reported (Murtas et al. 2003; Figure 3.1B). All the F1 plants were genotyped by PCR for the *esd4* allele, plants that presented an *esd4-1/esd4-2* genotype were kept and seeds from each plant harvested independently.

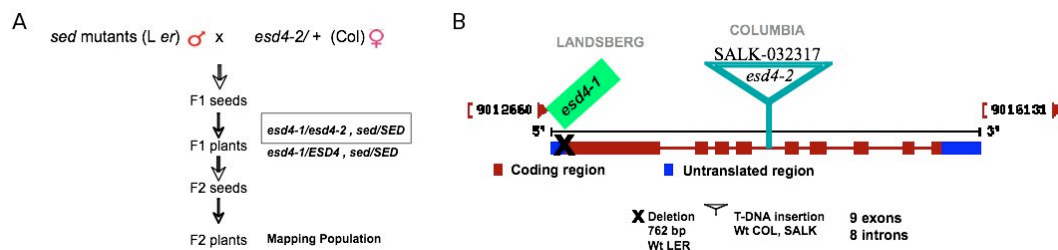


Figure 3.1. Generation of mapping populations of *sed* mutants. (A) Mapping populations were generated by crossing each one of the independent *sed* mutant lines, genotype *sed/sed; esd4-1/esd4-1*, which are in the L *er* accession, to heterozygous plants of the *esd4-2* mutant, which is an allele in the Col accession. *sed* mutants were used as pollen donors. Only F1 plants that presented an *esd4-1/esd4-2* genotype, denoted by a rectangle, were selected to be further analyzed. In this way it is assured that they were derived from a real cross. (B) *ESD4* genomic sequence representation. *esd4-1* is a gamma radiation mutant allele in the L *er* accession, this allele presents a deletion of 762 bp which includes 30 bp of the coding region starting from the ATG (Murtas et al. 2003). The *esd4-2* allele is in a Columbia background and is a SALK line carrying a T-DNA insertion within the fourth intron of *ESD4*.

3.3. Map based cloning of *sed* mutants

To assess the nature of segregation in the *sed* F2 mapping populations a first exploratory analysis was performed. For this purpose 86 F2 seeds from each one of the independent *sed* mapping populations generated were sown on soil and grown under SD conditions. This analysis revealed that the plants segregating from the mapping populations showed a considerable variation in phenotype. Flowering time looked more as a continuum instead of a discrete trait. This phenotypic variation could be due to the combination of alleles from the Col and L *er* accessions in the hybrid population. This form of heterosis has been reported before (Koornneef et al. 2004). To distinguish among the plants that present a suppressor phenotype from the ones that are only due to heterosis we generated a control cross of *esd4-1* (L *er*) and

esd4-2 (Col) mutants and determined the effect on flowering time and plant size in the F1 and F2 hybrid populations. Interestingly, the F1 hybrids showed a recovery of different aspects of the *esd4* mutant phenotype such as recovery of stature, leaf and silique size (larger and more similar to the wild type) (Figure 3.2A-B). A subtle delay in flowering in the F1 hybrid plants was also observed, however, these F1 hybrids were still early flowering in comparison to the wild type (Figure 3.2C). All these phenotypic aspects were very homogeneous among the independent F1 hybrids, indicating that somehow the combination of *esd4-1* (*L er*) and *esd4-2* (Col) genotypes has a certain compensatory effect over the *esd4* mutation or combining background alleles from *L er* and Col can partially suppress the effect of *esd4*. However, the phenotypic recovery in the F1 hybrids was surprising since the *esd4* mutation causes a very strong dwarf phenotype in both Col and *L er* accessions, which is even more dramatic under SD conditions under which the mapping populations were grown.

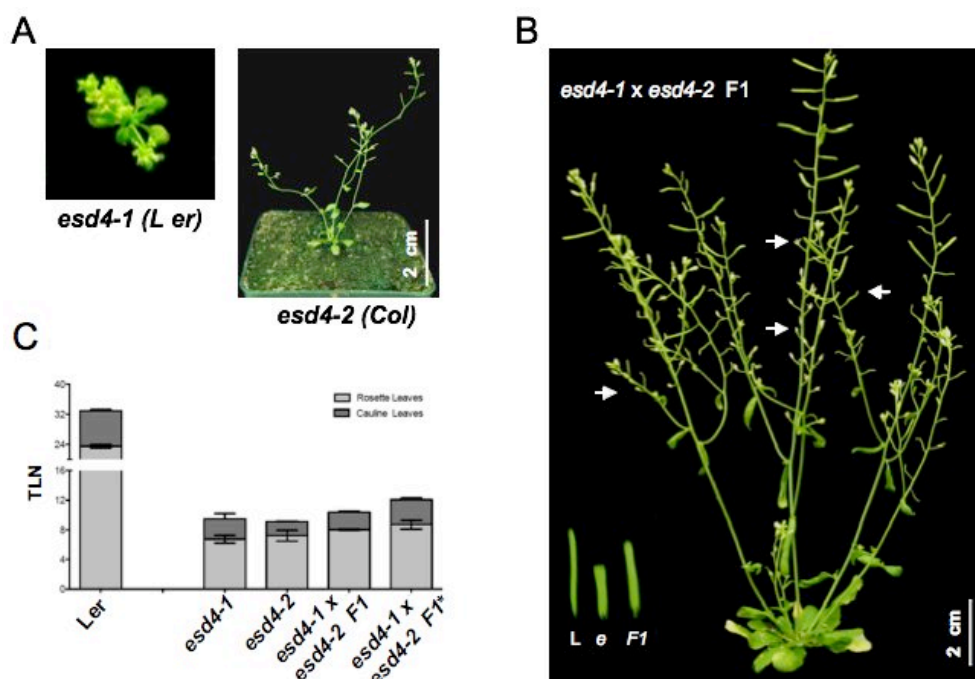


Figure 3.2. Phenotype of the F1 hybrid of the *esd4-1* (*L er*), *esd4-2* (Col) mutants. (A) *esd4* mutants exhibit a dwarf phenotype in *L er* and Col backgrounds. (B) F1 *esd4-1* (*L er*), *esd4-2* (Col) hybrid grown under SD conditions. Arrows denote several aborted flowers and siliques along the main and secondary inflorescence stems. Letters below the siliques to the bottom left indicate L=Landsberg *erecta*, e= *esd4-1* and F1=F1 hybrid of *esd4-1* (*L er*) x *esd4-2* (Col). (C) Flowering time for the *esd4* mutants as well as the F1 hybrid registered as the total leaf number (TLN) produced under SD conditions. The asterisk denotes a set of F1 hybrids that were grown under SD conditions in the glasshouse. Error bars illustrate the means \pm SE of $n \geq 12$ individuals.

Since the F1 hybrid population presented a weaker *esd4* homogeneous phenotype, most likely due to heterosis, we wondered if in the F2 generation strong and weak *esd4* phenotypes would segregate. To determine this, 86 *esd4-1* (*L er*) x *esd4-2* (*Col*) F2 mutant plants were grown under SD conditions and flowering time was scored. This experiment revealed a bell shaped frequency distribution of flowering time, which ranged from 10 to 25 total leaf number (TLN) in the F2 population, much later than the *esd4-1* (*L er*) x *esd4-2* (*Col*) F1 hybrids which flowered from 6 to 12 TLN or the parental single mutant lines, *esd4-1* (*L er*) or *esd4-2* (*Col*), which flowered within a range of 6 to 14 TLN (Figure 3.3). These results indicated that the *esd4-1* (*L er*) x *esd4-2* (*Col*) F2 population present a transgressive segregation and suggest that the combination of background alleles from *L er* and *Col* can partially suppress the effect of *esd4*. Interestingly, besides the variation in flowering phenotype other morphological phenotypes like variation in leaf size and shape were observed (Figure 3.4).

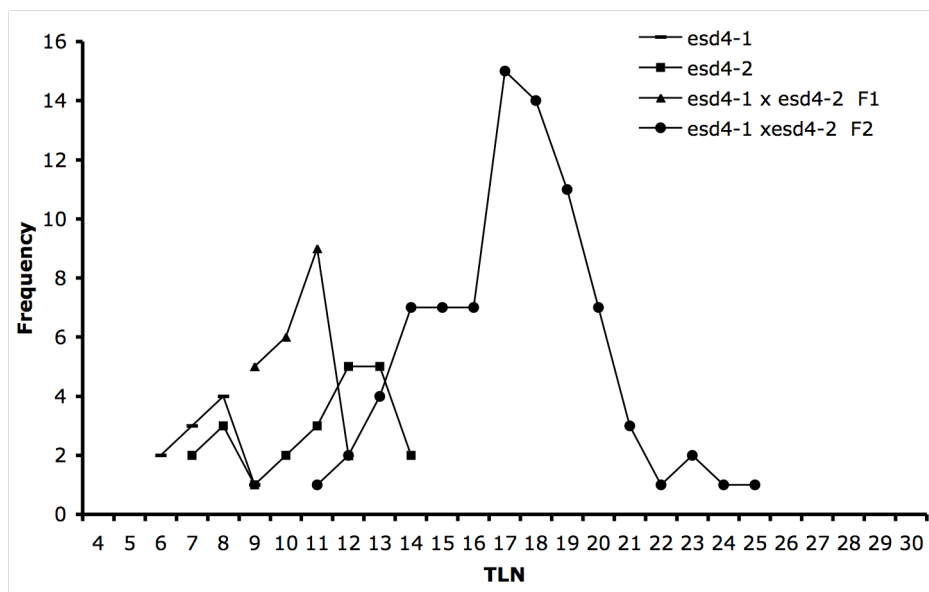


Figure 3.3. Frequency distribution of the flowering time of *esd4-1* (*L er*) x *esd4-2* (*Col*) F2 population under SD conditions. X axis indicates the TLN and Y axis the number of individuals. F2 population includes 86 individuals.



Figure 3.4. The *esd4-1* (*L er*) x *esd4-2* (*Col*) F2 population presents a broad phenotypic variation. This photograph illustrates the phenotypes of recombinant plants growing under SD conditions.

Since the *esd4-1* (*L er*) x *esd4-2* (*Col*) F2 population, from now on called control cross, presented a large variation in flowering time I wondered if it would be possible to discriminate late-flowering individuals in the *sed* (*L er*) x *esd4-2* (*Col*) F2 mapping populations. For this purpose the F2 populations from *sed4-1* x *esd4-2*, *sed5-1* x *esd4-2*, *sed43-1* x *esd4-2*, *sed44-3* x *esd4-2*, *sed45-1* x *esd4-2*, *sed111-1* x *esd4-2* as well as the control cross were grown under SD conditions and the flowering time compared. This experiment revealed that late-flowering individuals were distinguishable in the mapping F2 populations of *sed4-1*, *sed43-1*, *sed44-3* and *sed111-1* indicating that these populations are useful for mapping (Figure 3.5). On the other hand, the flowering time frequency distribution of the *sed5-1* and *sed45-1* populations fully overlapped with the one of the control cross, meaning that the effect of the suppressor was not strong enough to overcome the hybrid genotype background variation and therefore these populations were not analyzed further (Figure 3.6).

Interestingly, the late flowering individuals in the F2 segregating populations of *sed4-1*, *sed43-1*, *sed44-3* and *sed111-1* also presented large rosette leaves in comparison to the segregating plants of the control cross; this trait became useful for selecting the individuals to be genotyped for establishing rough map positions (next subsection).

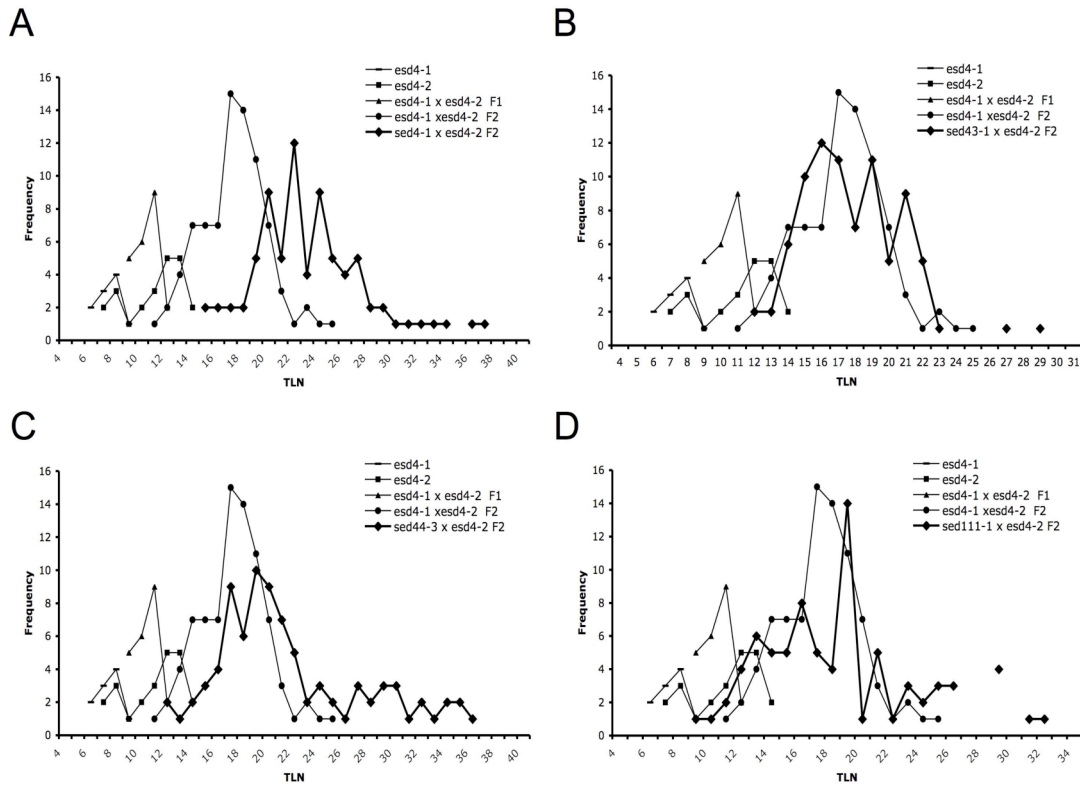


Figure 3.5. Flowering time frequency distribution of four independent *sed* (*L er*) x *esd4-2* (*Col*) F2 segregating populations. Flowering time frequency distribution of the F2 populations of *sed4-1* x *esd4-2* (A), *sed43-1* x *esd4-2* (B), *sed44-3* x *esd4-2* (C) and *sed111-1* x *esd4-2* (D). Plants were grown under SD conditions and the flowering time registered as the Total Leaf Number produced (TLN). X axis indicate the number of leaves and Y axis the number of individuals.

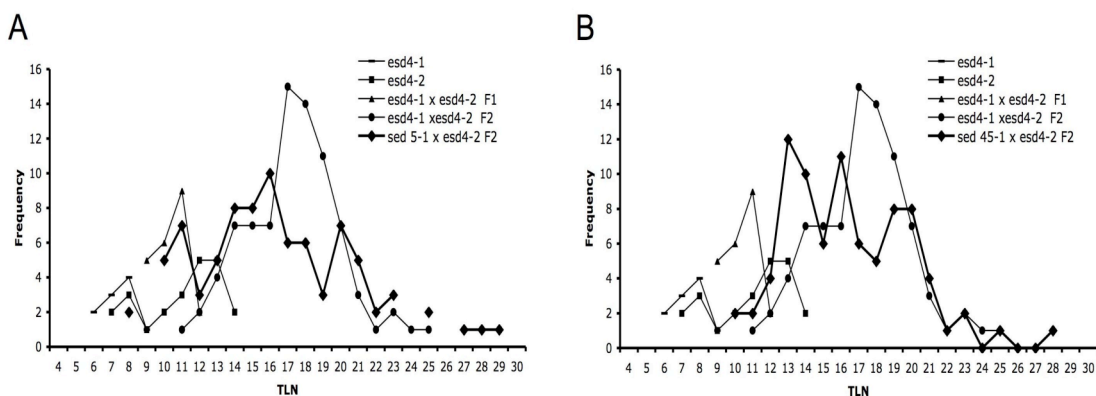


Figure 3.6. Flowering time frequency distribution of the *sed5-1* x *esd4-2* and *sed45-1* x *esd4-2* F2 segregating populations. Flowering time frequency distribution of the *sed5-1* x *esd4-2* (A) and *sed45-1* x *esd4-2* (B) populations. Plants were grown under SD conditions and the flowering time registered as the Total Leaf Number produced (TLN). X axis indicate the number of leaves and Y axis the number of individuals.

500 F2 plants from each F2 segregating population of *sed4-1* x *esd4-2*, *sed43-1* x *esd4-2* and *sed44-3* x *esd4-2* were grown under SD conditions and 100 plants, from each population, that presented a late flowering and big rosette area were selected to be genotyped (Figure 3.7 A-C, next section). To assure that the selected plants contained the suppressor mutation individuals that presented a late flowering phenotype and a bigger rosette area than the hybrid genetic control cross were selected since the presence of both characters was never found in the *esd4-1* x *esd4-2* F2 population (Figure 3.7D). In all cases the F2 *esd4-1* x *esd4-2* hybrid populations were grown alongside to compare as a phenotypic control. For the *sed111-1* a different strategy was followed and therefore it will be addressed in a later section (3.4.1).

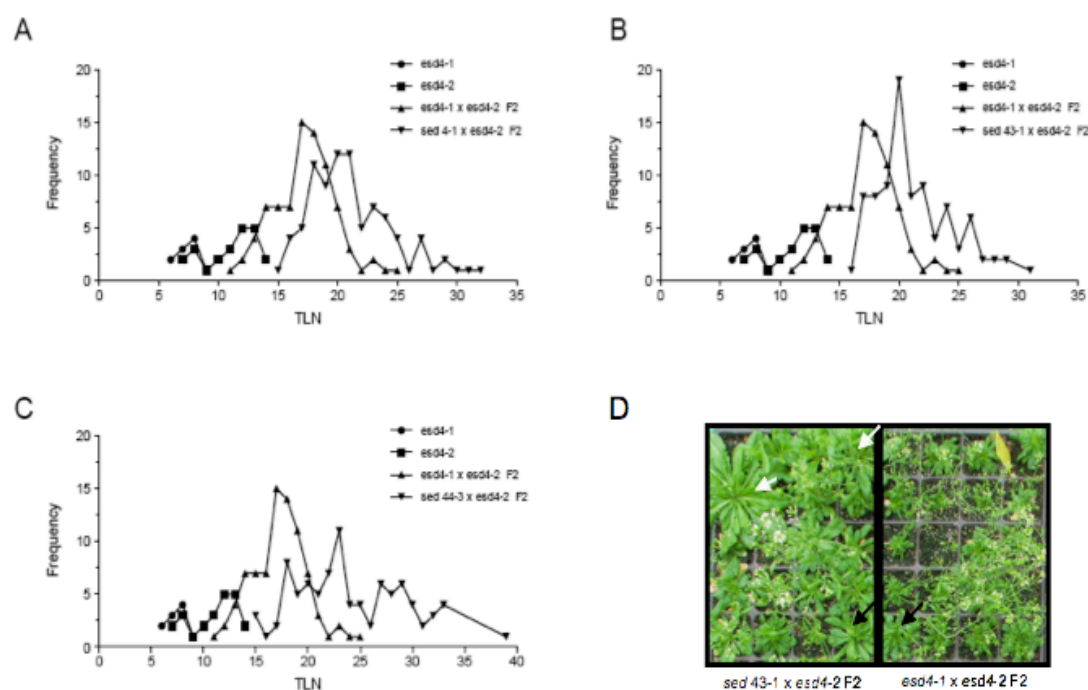


Figure 3.7. Flowering time of the F2 mapping populations of the *sed4-1*, *sed43-1* and *sed44-3* suppressor lines. Flowering time frequency distribution of the F2 mapping populations from the *sed4-1* x *esd4-2* (A), *sed43-1* x *esd4-2* (B) and *sed44-3* x *esd4-2* (C). Photograph of the F2 segregating plants from the *sed43-1* x *esd4-2* and the *esd4-1* x *esd4-2* populations. White arrows indicate representative plants showing a late flowering phenotype and a big leaf area, only present in the *sed43-1* x *esd4-2* population. Black arrows indicate plants morphologically similar between the *sed43-1* x *esd4-2* and the *esd4-1* x *esd4-2* populations, in the last case this plant represent one of the latest and biggest plants found in the control population. All plants were growth under SD conditions.

Suppressor Line	<i>sed</i> x <i>esd4-2</i> Mapping population		<i>sed</i> x <i>L er</i>		<i>sed</i> x <i>esd4-1</i> Backcross	
	G	A	G	A	G	A
<i>sed43-1</i>	F2	RM	No	-	-	aborted
<i>sed44-3</i>	F2	RM	F2	NP	-	aborted
<i>sed4-1</i>	F2	RM	No	-	F1	ND
<i>sed111-1</i>	F2	RM	F2	NP	F4	Yes
<i>sed5-1</i>	F2	NA	F2	P	F2	NA
<i>sed45-1</i>	F2	NA	F2	NP	-	aborted
<i>sed23-1</i>	F2	ND	No	-	No	-
<i>sed33-1</i>	F2	ND	No	-	No	-
<i>sed48-1</i>	F2	ND	F2	ND	No	-
<i>sed49-1</i>	F2	ND	F2	ND	No	-
<i>sed84-2</i>	F2	ND	No	-	No	-
<i>sed108</i>	F2	ND	No	-	No	-
<i>sed163-1</i>	F2	ND	No	-	No	-
<i>sed166-1</i>	F2	ND	F2	NP	No	-
<i>sed211-2</i>	F2	ND	F2	NP	No	-

Three different groups were established according to the strength of the phenotype, from top to bottom in High, Medium and Low.
G= generation; A= analysis; No= not cross performed
F2 indicates the generation; RM= Rough Map position established;
NA=Not analyzable; ND= not data, population has not been analyzed yet
NP= No visible phenotype detected; P=phenotype detected

Beside the mapping populations *sed* lines were also crossed to *L er* in order to look for a possible phenotype independent of the *esd4-1* mutation. In this case only *sed5-1* segregated a clear and distinctive phenotype, this will be addressed in another section (3.5).

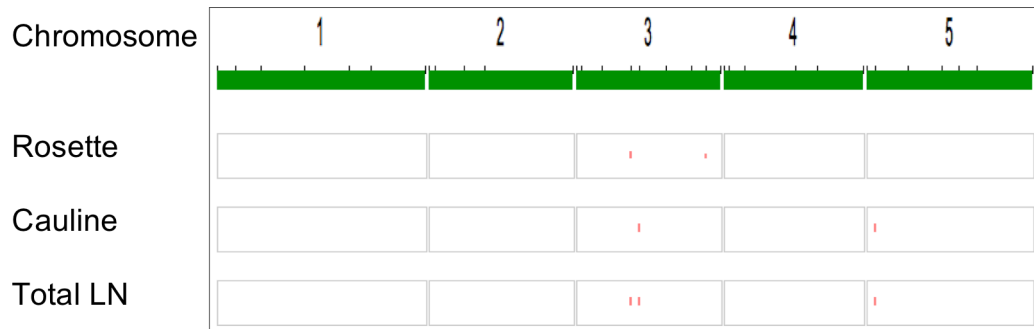
Mapping

DNA from the 100 F2 late flowering plants potentially *sed/sed*, *esd4/esd4* lines from each mapping population of the *sed4-1*, *sed43-1* and *sed44-3* suppressor lines were analyzed using a set of simple sequence length polymorphism markers (SSLP) to establish rough map positions. This analysis revealed several genomic regions enriched for the homozygous *L er* genotype in the mapping populations of *sed4-1* and *sed44-3*, but not *sed43-1* line (Data not shown). Remarkably, no single marker or region was absolutely linked to *sed4-1* and *sed44-3*. Therefore in order to distinguish the markers that are linked to the flowering time phenotype and the *sed* mutation, a basic marker-trait association analysis was performed using the GGT program (van Berloo 2008). This program calculates the correlation between marker data and trait values and the associated probabilities of the correlation values (van Berloo 2008). This analysis revealed that for the *sed4-1* mapping population, markers CIW11 and MUO22 (9.8 - 11.40 Mbp) on the upper arm of Chromosome III present a positive correlation with the flowering time (Figure 3.8A), and for the *sed43-1* mapping population markers F26K10 and T9A14 (13.0 - 17.0 Mbp) on the lower arm of Chromosome IV, and F4P9 (14.20 Mbp) in the lower arm of Chromosome II present a positive correlation with flowering time (Figure 3.8B). Finally the *sed44-3* mapping population showed a strong linkage to the upper arm of Chromosome V, between T32M21 and NGA139 (1.4 – 8.4 Mbp), but also the upper arm of Chromosome II and lower arm of Chromosome III might have a positive or additive effect (Figure 3.8C). Interestingly, these three *sed* mutations presented a positive marker-flowering time correlation in different marker intervals suggesting that the genetic basis might be different. To corroborate that these marker-trait associations were not related to the hybrid genetic background, namely *esd4-1* (*L er*) x *esd4-2* (*Col*) background, we also analyzed the same set of molecular markers in this population. This analysis revealed a relatively homogeneous heterozygous background where no single correlation was found (Figure 3.9).

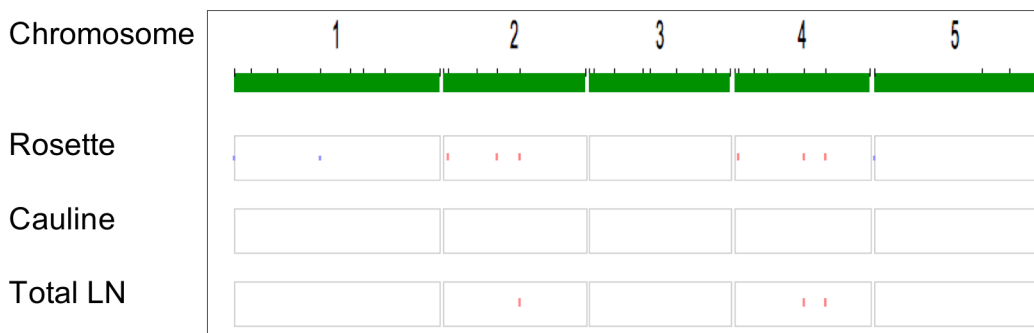
As previously shown (Figures 3.2, 3.3 and 3.4), it seems that the hybrid genetic background is very variable and might affect the mutant phenotype. The selected

plants that present big leaves and late flowering phenotype are enriched in the *L. er* background at several positions. A possible explanation for this is that the mutations may interact with alleles present in the Landsberg *erecta* background, and therefore the selected plants are enriched on this genotype at several positions.

A



B



C

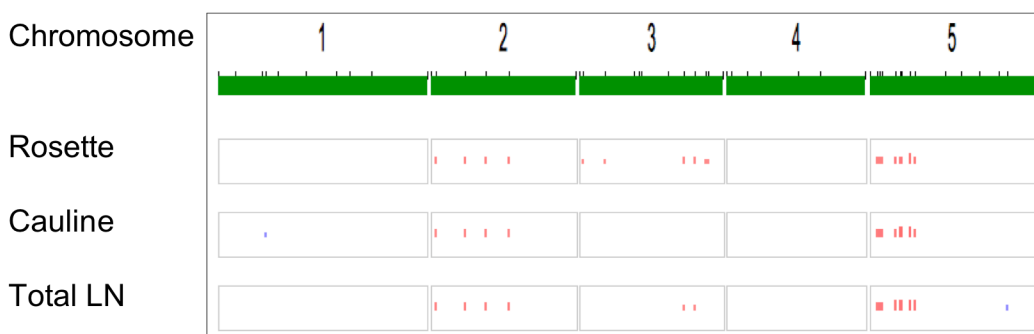


Figure 3.8. Marker-trait association plot of the *sed4-1*, *sed43-1* and *sed44-3* mapping populations. Marker-trait association plot of the *sed4-1* (A), *sed43-1* (B) and *sed44-3*(C). Each horizontal bar represents a chromosome, the vertical ticks on top represent each one of the analyzed markers. Each horizontal block represent a specific trait, in this case number of rosette leaves, cauline leaves and total leaf number. Red ticks signal the markers that present a positive association with the trait according to a correlation analysis. Correlation significances are corrected by a false discovery rate which establishes a threshold for the p-values of individual association tests (van Berloo 2008).

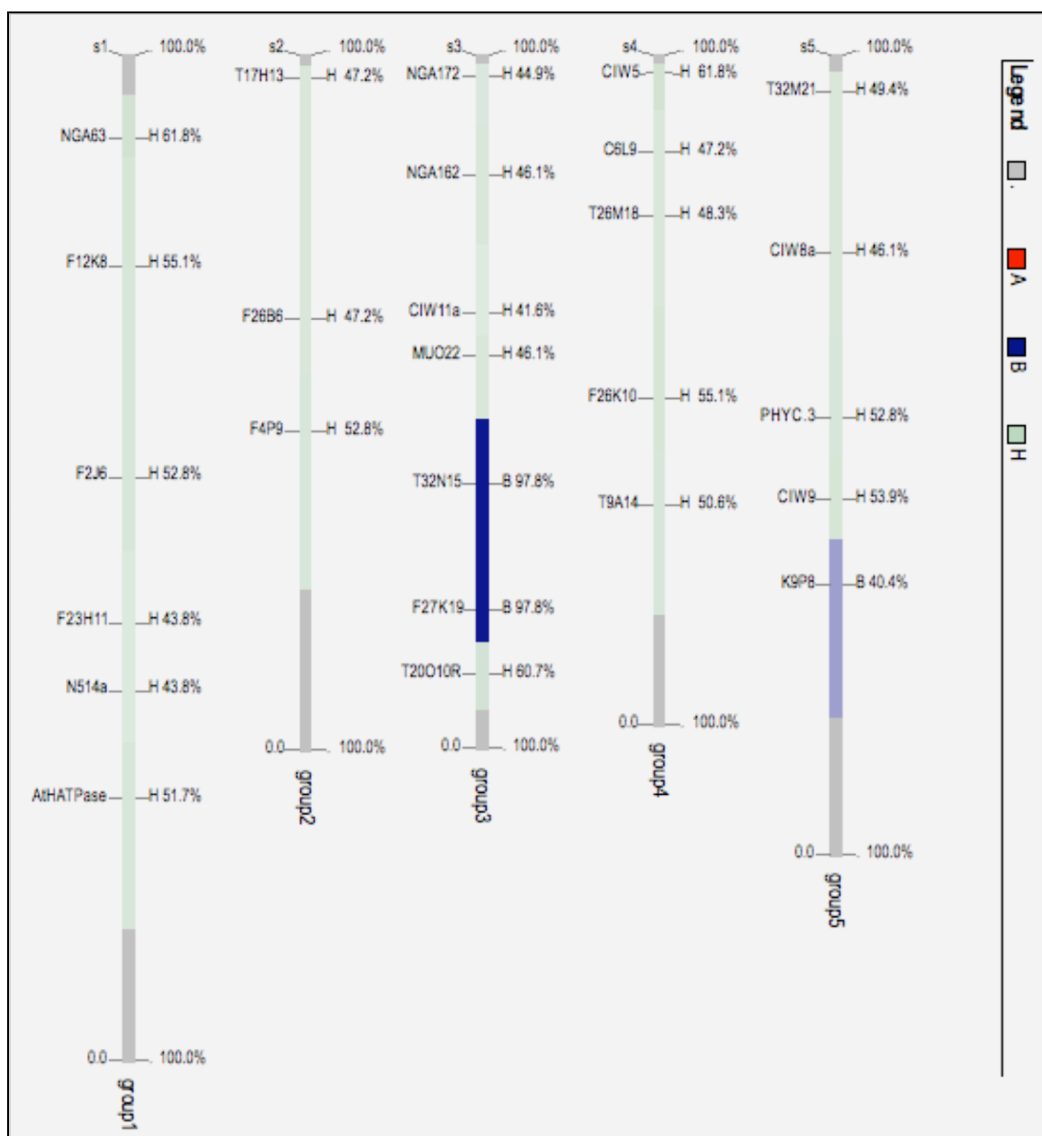


Figure 3.9. Consensus genotype map of F2 *esd4-1* x *esd4-2* hybrid population. The map is defined by 26 SSLP markers. The L *er* genotype is represented in red, Columbia in blue, and heterozygous regions in green. Gray regions are non determined. Each bar represents one chromosome from I-V from left to right. Marker name is depicted to the left of each chromosome and the percentage of individuals that present the most common genotype is written to the right.

3.4. *sed111-1* and Next Generation Sequencing strategy

3.4.1 Rough Map position established by bulk segregation analysis

The *sed111-1* mapping population was genotyped by bulked segregant analysis (BSA) (Michelmore, Paran, and Kesseli 1991) to determine a rough map position. The original BSA method consists of making two bulks of DNA obtained from a population that is segregating for a particular trait; one bulk would be constituted from individuals that share a trait, and another bulk would be individuals that do not present such trait. These two bulked DNA samples are afterwards analyzed with a set of molecular markers to establish linkage or a rough map position. Segregant populations derived from three independent F1 crosses of M3 *sed111-1* to the *esd4-2* mutants were used instead of a population segregating from a single cross. DNA from 94 F2 late flowering plants was pooled; two independent controls were used, one consisted of a 50-50 percent of a mix of Columbia and Landsberg *erecta* wild type DNA and the other was a 50-50 percent of a mix of *esd4-1* (*L er*) and *esd4-2* (Col) mutants. This BSA analysis showed that the *sed111-1* was linked to the bottom of Chromosome I (Figure 3.10), because *sed111-1* parental polymorphism was strongly overrepresented in this region.

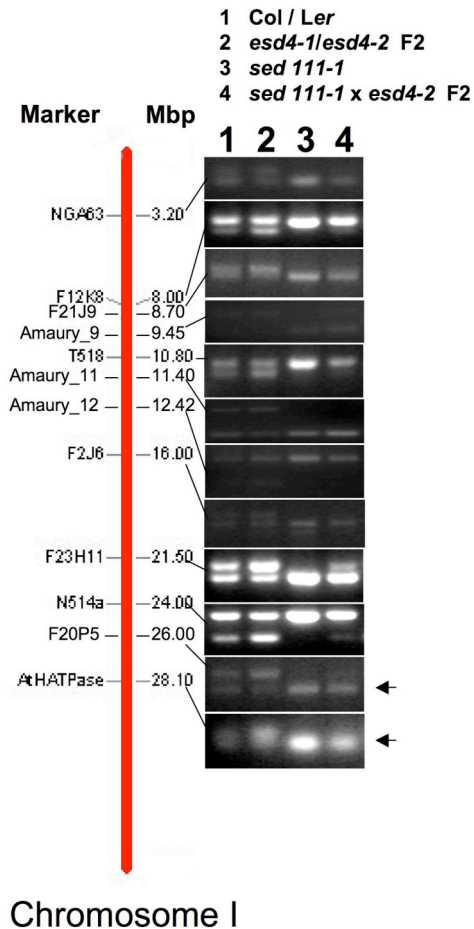


Figure 3.10. *sed111-1* rough map position by bulk segregant analysis. Chromosome I is represented by the red line, molecular markers name is depicted to the left of the chromosome while the position in mega base pairs (Mbp) is depicted to the right. The gel electrophoresis of PCR products for each molecular marker is shown, numbers on the top of these photographs refer to the genotype analyzed, Col / L *er* (1) consist of a mix of 50% of each accession DNA as control, *esd4-1 / esd4-2* F2 (2) consist of a mix of the DNA from 90 F2 individuals from this cross as a background control, *sed111-1* (3) works as the original parental control, *sed111-1 x esd4-2* F2 (4) consist of a mix of the DNA from 93 F2 individuals from this cross that constitute the mapping population. Black arrows indicate the markers that defect preferentially L *er*.

3.4.2 Background cleaning

The *sed111-1* suppressor mutant line was backcrossed with the parental *esd4-1* mutant to reduce the number of unlinked extra mutations induced by the EMS mutagenesis. The F1 BC1 seeds were sown on soil and individual F1 plants selfed. The resultant F2 seeds were grown under SD conditions to analyze *sed111-1* segregation. Individual F2 plants were predicted to be homozygous *esd4-1/esd4-1*; *sed111-1/sed111-1*, heterozygous *esd4-1/esd4-1*; *sed111-1/SED* or only *esd4-1/esd4-1*; SED/SED according to the phenotype observed, namely, late flowering phenotype and bigger leaves for the homozygous *sed111-1* segregating plants and early flowering and dwarf for the *esd4-1/esd4-1*; *sed111-1/SED* and *esd4-1/esd4-1*; SED/SED plants. To confirm the predicted genotypes F3 seeds obtained from individual F2 plants were

sown and grown under SD conditions, as expected the progeny of the predicted homozygous *sed111-1* plants presented a late flowering and big leaves phenotypes. The phenotypes were once more confirmed in the F4 generation (Figure 3.11).

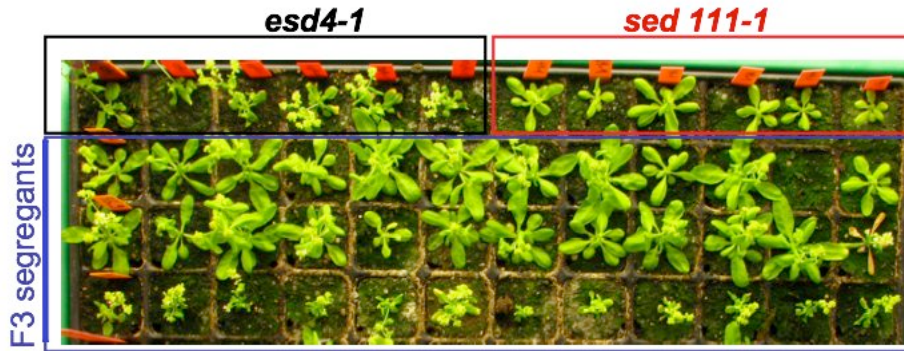


Figure 3.11. *sed111-1* suppressor mutant background cleaning, F3 BC1 segregants. The *sed111-1* line was backcrossed with the parental *esd4-1* mutant and the resultant progeny segregated in the F2 population. F2 plants predicted to be homozygous for the *sed111-1* mutation gave place to homogeneous late flowering plants with big leaves in the F3 generation (second and third rows) in comparison to the plants predicted to be *esd4-1/esd4-1*; SED/SED (third row). Plants were grown under SD conditions.

3.4.3 DNA extraction, sequencing and data analysis

Genomic DNA was extracted from *sed111-1* F3 BC1 (backcrossed to *esd4-1*) seedlings obtained from a single F2 BC1 family that showed a consistent segregation and homogeneous phenotype (section 3.4.2). Genomic DNA was also extracted from the parental *esd4-1* mutant to use as control. Whole-genome sequencing was performed using Illumina/Solexa Genome Analyzer II. A total of 2 x 40 million 95 bp long paired-end reads were generated, this is equivalent to an average of 61-fold sequence coverage across the entire genome. The obtained reads were mapped to the TAIR9 version of the Arabidopsis thaliana reference genome (Columbia ecotype) for variation discovery (Single Nucleotide Polymorphisms-SNPs, indels and Copy Number Variations-CNV) using the aligner program Burrows-Wheeler Aligner (BWA). BWA is an aligner that achieves a good balance between performance and accuracy for SNP calling (Li and Durbin 2009). Since the assembled reference genome is in the Columbia background while the *sed111-1* as well as their progenitor *esd4-1* are in the

Landsberg *erecta* (L *er*) ecotype, all the SNPs reported for L *er* in the TAIR database were also included in the analysis to determine the unique SNPs for *sed111-1* and *esd4-1*. A score matrix was generated from the BWA analysis and used to call homozygous SNPs in *sed111-1* and homozygous as well as heterozygous SNPs in *esd4-1* mutant. The *esd4-1* mutant heterozygous SNPs were necessary to get the maximum variation since this is the background genotype. The *esd4-1* mutant homozygous and heterozygous SNPs list was compared to the homozygous SNPs list of *sed111-1* to identify the SNPs that were homozygous and unique for *sed111-1*. The *sed111-1* SNPs were required to be covered with reads in *esd4-1* to exclude regions that are not covered by the sequencing or are difficult to align because either they contain many variations or large insertions or deletions.

3.4.4 Generation of *de novo* dCAPS markers, genetic interval, gene candidates

Generation of *de novo* dCAPS markers

CAPS (Cleaved Amplified Polymorphic Sequences) markers are amplified DNA fragments that contain one or more polymorphisms differing between two accessions. When these PCR fragments are digested with a restriction endonuclease they display a Restriction Fragment Length Polymorphism (RFLP) that differentiates the genotype of origin (Konieczny and Ausubel 1993). These markers are PCR-based and are co-dominant, meaning that both chromosomes/alleles may be genotyped (Lukowitz et al. 2000). dCAPS markers are derived CAPS markers where a synthetic restriction site is created by a combination of a mismatched PCR primer, which introduces a new-synthetic polymorphism, with the naturally occurring polymorphism in one of the accessions (Neff et al. 1998). *de novo* dCAPS markers are simply dCAPS markers that are created based on the SNPs produced by the EMS mutagenesis. In our case, these SNPs were detected by next generation sequencing in a segregant mutagenized population that presented the phenotype of interest (Figure 3.12). In particular for this work, several *de novo* dCAPS were designed to fine-map the causative mutation of the *sed111-1* suppressor phenotype.

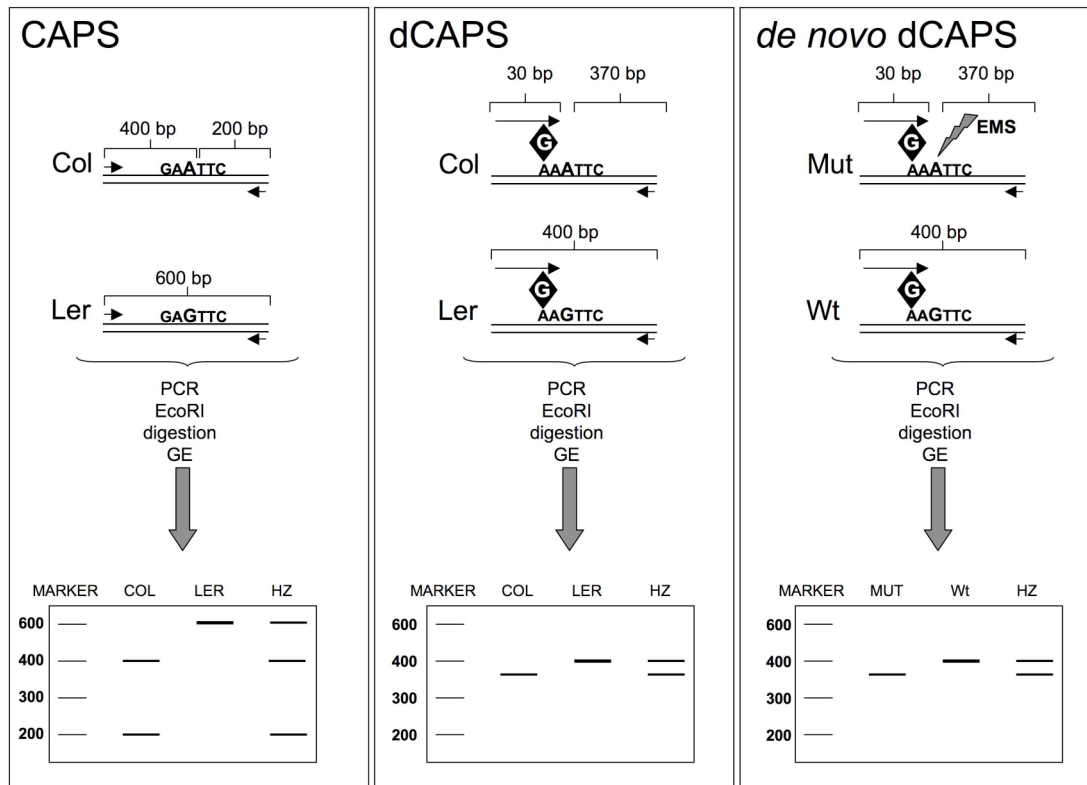


Figure 3.12. *de novo* dCAPS markers. CAPS markers consist of PCR amplicons that contain one or more polymorphisms between two different accessions, in this case Col and *L er*, and can be digested enzymatically giving place to different restriction fragment patterns. In this example the enzyme used is *EcoRI* and only one site is predicted to exist for Col. dCAPS is a variation of CAPS where a base pair change is introduced by a mismatched primer in order to create a synthetic restriction site when combined with the SNP already present in the accession. In this case the Columbia presents an adenine (A) whereas the *L er* presents a guanine (G), the synthetic introgression of G instead of A in both ecotypes generates an *EcoRI* site only in the Col accession. *de novo* dCAPS are the same as dCAPS with the only variation that the SNP variation is the product of the mutagenesis, in this case EMS depicted as a small lightning sign. This dCAPS marker distinguishes the genotype and at the same time corroborates the presence of the SNP in the mutant. GE is an abbreviation for Gel Electrophoresis. Adapted and modified from (Lukowitz et al. 2000).

Analysis of *sed111-1* x *esd4-1* F2 Backcross segregants by *de novo* dCAPS

In the previous section I explained how the *de novo* dCAPS are designed. In this section these molecular markers were used to narrow down the genetic interval that encompasses *sed111-1* to a few candidate genes. *de novo* dCAPS markers work as the normal dCAPS markers with the advantage that here the analysis is performed in the same background where the mutation was generated, avoiding complications in

scoring phenotypes due to heterosis in the hybrid populations as previously shown (Section 3.3, Figures 3.2, 3.3 and 3.4). Another advantage is that the marker itself corroborates the existence of the mutation. The bulk segregant analysis of the *sed111-1* F₂ mapping population established linkage to the bottom arm of Chromosome I, below 24.0 Mbp (Section 3.4.1). To corroborate this genetic interval several *de novo* dCAPS were designed based on different high score-quality SNPs that were distributed along chromosome I and tested on 30 F₂ plants showing the *sed111-1* phenotype. This analysis confirmed the linkage to the bottom of chromosome I (Figure 3.13), the analysis of another 166 F₂ plants narrowed down the genetic interval to 26.49 -28.07 Mbp. In this interval 13 SNPs were present. Five of these SNPs lead to non-synonymous mutations, five give rise to synonymous mutations, two are located in intronic/non coding regions and one is located at the 5'UTR (Table 3.2). Currently the SNPs located in the outer limits of the interval, At1g70320 and At1g74710, are being used to look for recombinants to narrow down the genetic interval containing the *sed111-1*.

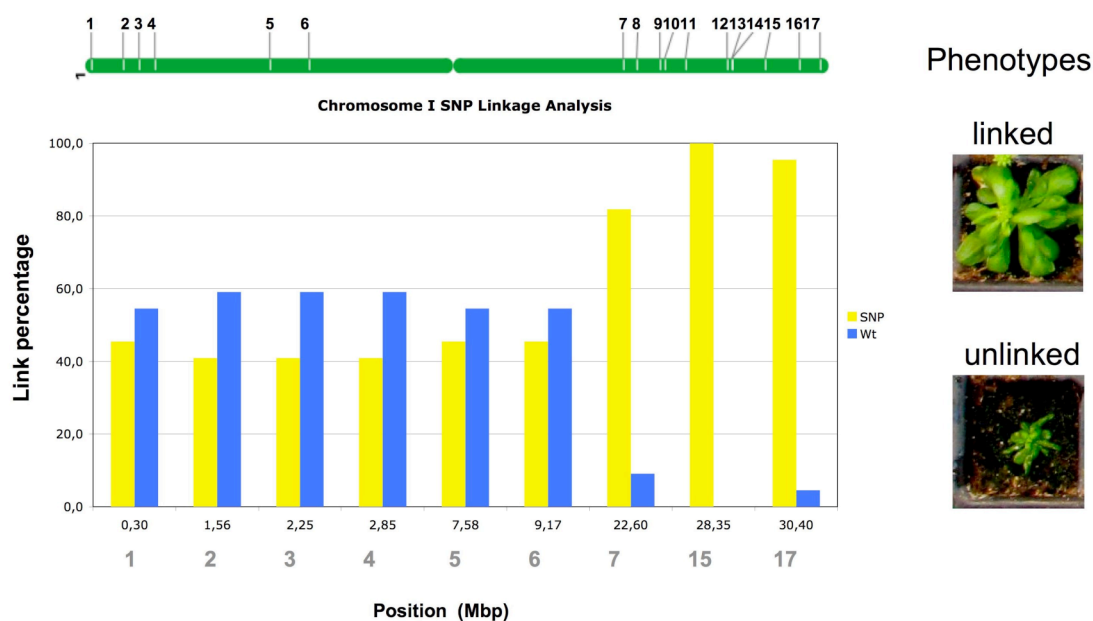


Figure 3.13. Linkage analysis of some of the SNPs found in Chromosome I in *sed111-1* plants. In order to assess some of the candidates in the linked area *de novo* dCAPS markers were generated based on the sequencing data and analyzed in the F₂ segregants from a backcross to *esd4-1* (Figure 3.11). This analysis confirmed the previous mapping data (Figure 3.10). Numbers at the bottom of the graph and on top of the green bar are references to localize the physical position of each marker. Green bar represents Chromosome I and each white tick on it corresponds to one SNP-Marker. Plant photographs present the phenotypes for a representative plant that showed linkage or not to the marker depicted as number 15 in the graph.

AGI code	gene	Mutation nature		Protein function	
AT1G70320		Non synonymous	R	Stop	E3 ubiquitin-protein ligase UPL2
AT1G70460		five_prime_UTR	-	-	Similar to protein kinase
AT1G70940		Synonymous	T	T	Auxin efflux carrier component 3
AT1G71240		Intronic/non coding	-	Stop	Unknown protein
AT1G71340		Intronic/non coding	-	-	Glycerolphosphodiesterase 4
AT1G72100		Synonymous	A	A	LEA domain containing protein
AT1G72410		Non synonymous	P	L	similar to COP1-interacting protein-related
AT1G73130		Synonymous	L	L	Unknown protein
AT1G73500		Non synonymous	T	P	AtMKK9
AT1G74170		Non synonymous	S	F	leucine-rich repeat family protein AtRLP13
AT1G74450		Synonymous	E	E	Unknown
AT1G74540		Synonymous	R	R	similar to CYP98A9 (cytochrome P450, family 98, subfamily A, polypeptide 9)
AT1G74710		Non synonymous	G	R	Isochorismate synthase 1

Analysis of *L er* x *sed111-1* F2 segregants

Eight F2 segregants were selected from a 90-individual population of *L er* x *sed111-1*. These individuals were analyzed with *de novo* dCAPS markers outside and within the genetic interval containing the *sed111-1* to confirm that they inherited all or some of the mutations in the genetic interval. All eight F2 segregants presented a heterozygous genotype for the complete or part of the linked interval but were mostly wild type outside this genetic interval, they were also heterozygous for the *esd4-1* mutation. These individuals will serve for three purposes: first to increase the number of potential recombinants to narrow down the genetic interval containing the *sed111-1* since only few seeds are produced by the *sed111-1* x *esd4-1* cross, second to isolate the mutation in the same *L er* genotype in case the phenotype is only visible in this accession, and third, the F3 generation will also allow to test more directed and intensively the candidates than was done before (Table 3.1) to determine whether they have a phenotype in the absence of the *esd4-1* mutant background.

***sed111-1* candidate genes**

The suppressor screen of the null mutant *esd4-1* is expected to allow the recovery of new mutations that overcome the effect of *esd4-1* on plant development. I focused on flowering time and overall growth. The combination of classical forward genetics, map based cloning and whole genome sequencing allowed me to reduce the number of genes linked to *sed111-1* to 13 candidates from which only six caused a non-synonymous mutation (Table 3.2). This reduced the number of candidate genes dramatically, because the original interval contained 469 loci (26.49 -28.07 Mbp).

Since *esd4-1* is a null mutant, recovery of mutations in genes that directly interact with ESD4 was not expected. In agreement with this idea, none of the six candidates of the *sed111-1* line are found in the database of SUMOylated proteins that interact with ESD4 (Elrouby and Coupland 2010), nor with the SUMOylated proteins reported in other studies (Miller et al. 2010; Park et al. 2011).

Four out of the six genes that induce a non-synonymous mutation have been reported in the literature. This information is valuable to determine what these genes do and how much they might contribute to the observed phenotype. Here I present a general description of these four genes.

At1g70320 E3 Ubiquitin-protein ligase UPL2 (Stop). UPL2 is a member of the HECT-E3 family in Arabidopsis, which consists of seven members that are grouped into four subfamilies according to the presence of domains upstream of the HECT domain (Downes et al. 2003). HECT domain Ubiquitin-protein ligases promote the transfer of Ubiquitin to appropriate targets by directly interacting with Ubiquitin via a conserved Cysteine in the HECT domain to which a high energy Ub-E3 thiol-ester intermediate is formed (Bates and Vierstra 1999). UPL2 encodes a protein of 405 kDa that is 94% identical to UPL1, indicating that it is a recently duplicated gene (Bates and Vierstra 1999). Both UPL1 and UPL2 are located on Chromosome I, 26 cM apart, on the upper and lower chromosome arms, respectively, and are products of a recent duplication (Bates and Vierstra 1999). The functionality of the HECT domain of UPL1 was tested by Ubiquitin conjugation assays *in vitro* (Bates and Vierstra 1999).

At1g73500 Mitogen-activated protein kinase 9 AtMKK9. MAPK cascades are important means of signal transduction; they are involved in pathways that regulate plant growth, development, and responses to various stress stimuli. Studies of transgenic plants that express an active version of MKK9 showed that this kinase-kinase activates MPK3 and MPK6 proteins, the activation of these proteins leads to an induction of the synthesis of ethylene (important phytohormone induced by biotic and abiotic stresses) and camalexin, an indole phytoalexin (sulfur-containing indole alkaloid) employed in defense against infection by microorganisms (Glawischnig et al. 2004; Xu et al. 2008).

At1g74170 leucine-rich repeat family protein AtRLP13. Receptor like proteins (RLP) are cell surface receptors that consist of an extracellular leucine-rich repeat domain, a transmembrane domain and a short cytoplasmic tail (Wang et al. 2008). These receptors have been related to disease resistance, plant growth and development, and sensitivity to various stress responses and abscisic acid (Wang et al. 2008). Until now 57 RLP have been identified in the Arabidopsis genome, and all are expressed. Some of the interesting representatives are CLAVATA 2, an important gene in the regulation of meristem maintenance, and TOO MANY MOUTHS (TMM), a gene involved in the regulation of stomatal distribution (Wang et al. 2008).

At1g74710 Isochorismate synthase 1 (ICS1). ICS1 is an essential enzyme for Salicylic Acid (SA) synthesis following pathogen recognition (Wildermuth et al. 2001). SA deficient plants are late flowering under SD conditions. This phenotype seems to involve the photoperiod pathway but without the requirement of CO, FCA or FLC (Martínez et al. 2004). Interestingly, ICS1 mutants called *sid2-1*, suppress the growth defect of leaf curling in 35S::SUM2 plants but not the dwarfness nor leaf elongation (van den Burg et al. 2010). Remarkably, a previous study showed that *sid-2* does not suppress the *esd4-2* mutant growth phenotype (Hermkes et al. 2011) under LD conditions where presumably the *sid-2* mutant has minimal or no effects.

***esd4* complementation by crossing with the candidate genes**

As a second strategy the sequence-indexed insertion line database of the Salk Institute Genomic Analysis Laboratory was screened for T-DNA insertions in the six genes containing non-synonymous mutations and that were identified in the previous section. These include similar to COP1-interacting protein related (At1g72410; SAIL_182_A11) with an insertion in the middle of the ninth exon, AtMKK9, (At1g73500; SAIL_60_H06) with an insertion in the middle of the gene (encoded by only one exon), AtRLP13 (At1g74170; GABI_574F03) in the middle of the eighth and last exon, (SALK_122999.30.40 and SALK_020984.49.10) both in the fifth exon. I also identified a T-DNA insertion for Isochorismate synthase I (At1g74170; SALK_042603.55.70) at the end of the fourth exon. These lines were ordered, screened to identify homozygotes and crossed to the *esd4-2* (Col) allele (since all these lines are in the Col background) to test whether any of them recapitulate the *sed111-1* suppressor phenotype.

3.5. *sed5-1* backcross mapping strategy

In a previous section (3.3) I mentioned that in addition to the mapping populations that were generated by crosses to *esd4-2* (Col) mutant, *sed* lines were also crossed to *L er* in order to segregate the *sed* mutation away from the *esd4-1* first-site mutation. The purpose of this experiment was to both assess the phenotype independent of the *esd4-1* mutant phenotype, and to eventually generate mapping populations by crossing to Wt Col. In this case the phenotypic characterization would be easier since only one mutation would have an effect on the plant. From the different *sed* lines backcrossed to *L er* (Table 3.1), only the *sed5-1* line segregated a clear phenotype independent of the *esd4-1* mutation in the F2 generation. From 80 F2 *sed5-1* x *L er* plants three presented the same mutant phenotype. To confirm the inheritance of these mutant phenotypes seeds from each one of the three plants were sown on soil and the phenotypes analyzed. All plants presented the same mutant phenotype indicating that they were homozygous. These mutants were called A6-C6 (in reference to the coordinates where they were growing on the tray). A6-C6 was backcrossed to *L er* to follow segregation of the phenotype in the F2 plants. A6-C6

plants presented diamond-shaped cotyledons, narrower and pale green rosette leaves, an apical inflorescence which dies without elongation, several flowers and siliques that do not mature or are aborted, and very often ball-shaped structures instead of trichomes are formed at the abaxial side of the leaves. Interestingly all these phenotypes are stronger under SD (Figure 3.14).

A rough mapping analysis by BSA from the *sed5-1/ sed5-1*; ESD4/ESD4 x Col showed a clear linkage to the molecular marker K9P8 (20,0 Mbp) at the bottom of Chromosome V. Analysis of a larger population allowed narrowing down the linked genetic interval with two recombinants; one at K9P8 (20,0 Mbp) and the other at M322 (dCAPS marker at 20,75 Mbp). This region contains 295 loci. Analysis of a much larger mapping population is required to narrow down this linked region (Figure 3.15).

Once a rough map position was established for A6-C6 mutant, the putative *sed5-1* line, a cross to the *esd4-1* mutant was performed to confirm that the isolated mutant is indeed the suppressor *sed5-1*. All F1 plants presented a L *er* phenotype indicating the recessiveness of both mutants, namely *esd4-1* and A6-C6. The F2 population is yet to be analyzed.

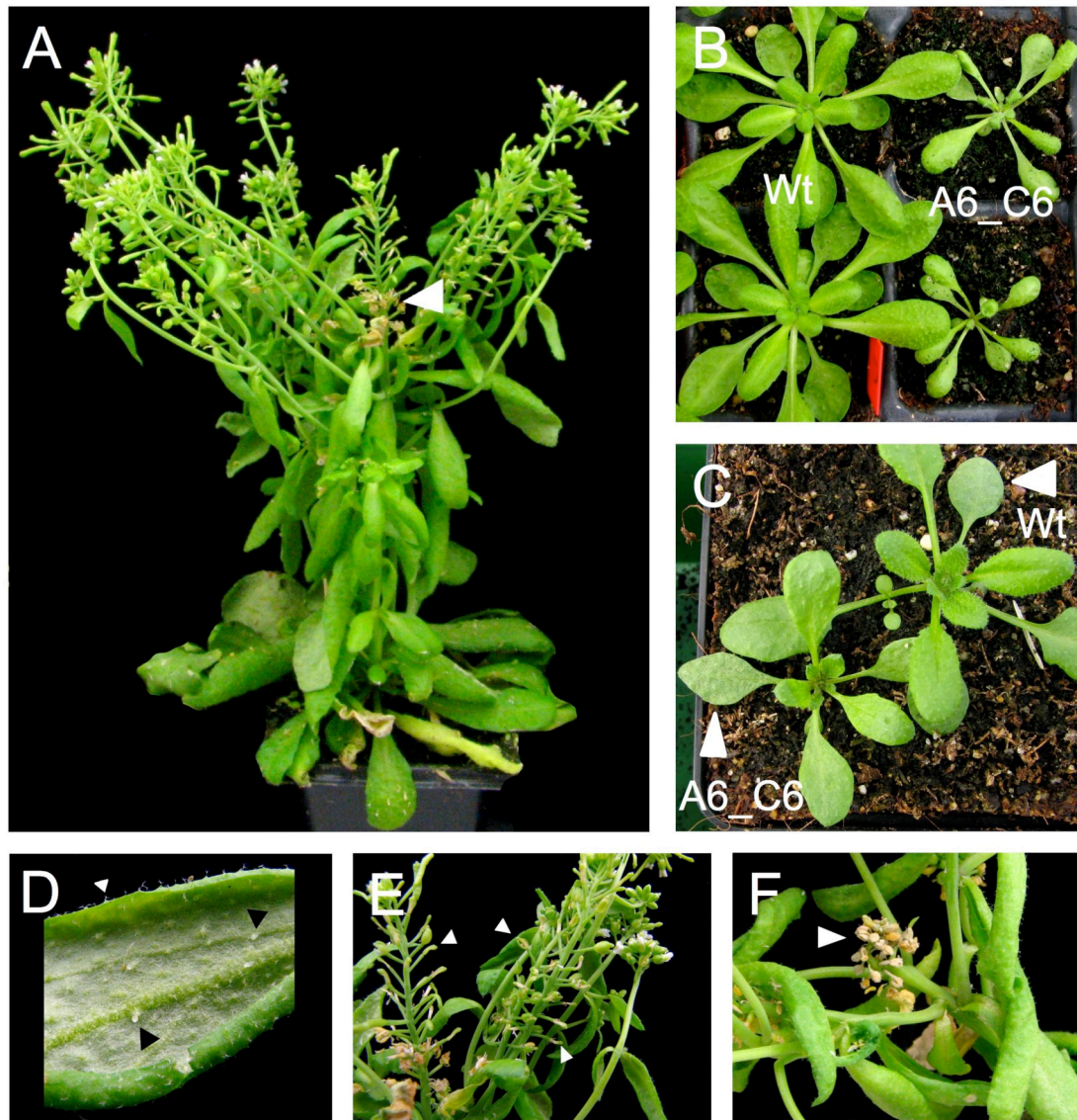


Figure 3.14. A6-C6 plants phenotypes. (A) A6-C6 plants present several aborted siliques and frequently the main inflorescence dries out. (B) A6-C6 rosette leaves (right) show longer petioles, small, pale and inward growth in comparison to the Wt. (C) A6-6C cotyledons show a diamond shape in contrast to the more circular shape of the Wt cotyledons. (D) Ball-shaped structures on the abaxial part of the rosette leaves. (E) Secondary inflorescences frequently present aborted siliques. (F) Primary inflorescence detail showing all the senesced flowers. Arrow heads serve to indicate the structures. All photographs, except C, are from plants growing under SD conditions.

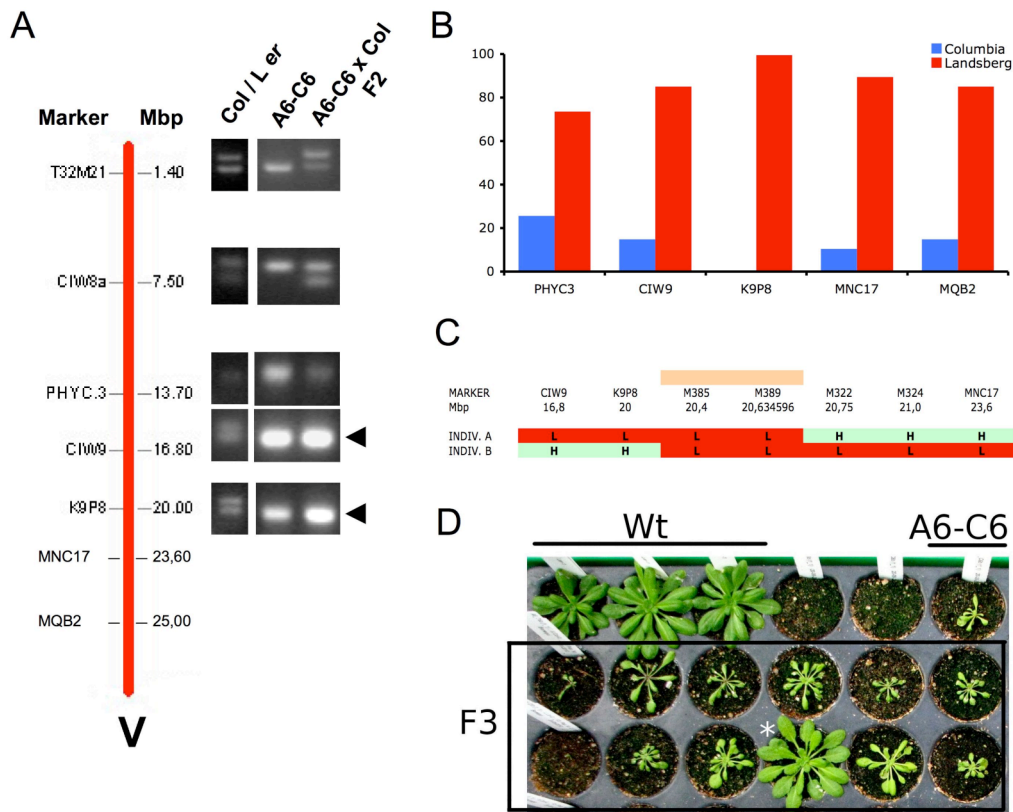


Figure 3.15. Bulk Segregant Analysis and rough map position of A6-C6. (A) Chromosome V is depicted as a red line, molecular markers name and their position in mega base pairs (Mbp) are indicated. The GE of PCR products for each molecular marker is shown. Col/L *er* indicates a 50-50 percent mix of Col and L *er* DNA, A6-C6 mutant DNA, and A6-C6 x Col F2 is a mix of 20 segregants presenting the mutant phenotype from a mapping population. Arrows indicate the linked markers. (B) Recombinant analysis from 138 individuals from an F2 mapping population showing a 100 percent linkage to the K9P8. (C) Genetic interval linked to the A6-C6 mutation showing the only two recombinant individuals found. (D) F3 plants from two independent F2 segregant plants from the mapping population. Upper row shows the progeny of a plant homozygous L *er* for the linked region. Lower row shows the progeny of an heterozygous plant. The white asterisk indicates a plant that presented a Wt phenotype, this plant was heterozygous for the linked region. Plants grown under SD conditions.

3.6. Discussion

SUMOylation, the post-translational attachment of SUMO (Small Ubiquitin-like Modifier) to a substrate protein, regulates the activity of several proteins involved in critical cellular processes. SUMO is subsequently removed from substrates by SUMO-specific proteases, making this modification reversible. In plants, SUMOylation has been implicated in several physiological responses and flowering time control. ESD4 (Early in Short Days 4) encodes a SUMO-specific protease that prevents the accumulation of SUMO-conjugates in Arabidopsis. The *esd4-1* mutant shows a very early flowering phenotype as well as several shoot developmental distortions suggesting an important role of SUMOylation in the regulation of plant development. In order to identify potential SUMO-substrates involved in the regulation of flowering, an enhanced mutant screen was performed. In this way 120 independent suppressors of *esd4* (*sed*) were isolated and 15 of them further characterized (Chapter II). Rough map positions for five of these *sed* mutations using classical genetic methods were established (this Chapter). As an alternative approach we used Next Generation Sequencing technology to fine map one of these *sed* mutations to high resolution. Doing this, we were able to use the whole genome sequencing data of an EMS mutagenized plant to generate *de novo* dCAPs markers to fine map the causative mutation of the suppressor phenotype. This powerful and fast sequencing technology in combination with state of the art bioinformatic analysis reduced the 469 genes localized within the linked interval to only six clear candidate genes that are being currently tested in independent segregant populations and reverse genetics. T-DNA insertion lines were ordered for these genes, screened to identify homozygotes and crossed to the *esd4-2* (Col) allele (since all these lines are in the Col background) to recapitulate the *sed111-1* suppressor phenotype.

Mapping populations

*The control cross *esd4-1(L er)* x *esd4-2 (Col)* leads to partial suppression of the *esd4* phenotype*

In order to identify the causal mutations of the *sed* mutants a map-based cloning strategy was followed. Contrary to what was expected the genetic mix of the two

accession genomes, Columbia and Landsberg *erecta*, led to a considerable recovery of the *esd4* mutant phenotype in the F1, and a surprising phenotypic variation in the F2 generation (Figures 3.2, 3.3 and 3.4). Such phenotypes can be explained by phenomena like hybrid vigour, in the case of the F1 population, and a compensation result of an allelic combination in the F2 generation which can commonly occur when two accessions, like Col and L *er*, that are highly inbred are crossed (Koornneef et al. 2004). Despite the wide range of the flowering time frequency distribution (TLN range 11-26) in this *esd4-1* (L *er*) x *esd4-2* (Col) F2 control background cross, the population still represents an early flowering distribution if we compared to the flowering time of the Col or L *er* plants, which under SD normally flower with around 50 and 35 total Leaves, respectively. This indicates that the *esd4* mutation is still exerting a strong effect on the flowering time in the segregating population.

Only sed mutants that present a strong late flowering phenotype and big rosette leaves can be mapped in the esd4-1(L er) x esd4-2 (Col) mutant background

To reduce the phenotypic mis-scoring of the segregant plants containing the *sed* mutation (compared to the ones that were only caused by the allelic combination in the Col - L *er* background) two traits were taken into account, flowering time and rosette leaf area. These traits were chosen because only few individuals in the background cross (*esd4-1* (L *er*) x *esd4-2* (Col)) present both traits in comparison to some of the *sed* mapping populations where almost a quarter of the population present plants with this combination of phenotypes (Figure 3.7D). However, after establishing the selection method and analyzing different mapping populations I realized that not all the suppressors exert an effect strong enough to overcome the variability presented by the background, so that both the flowering time frequency distribution as well as the rosette leaf area were overlapping (Figure 3.6).

Therefore only *sed* mapping populations that presented segregating individuals with a late flowering time and large rosette leaf area distinguishable from the control cross background were selected (Figure 3.7). Since neither single region nor marker was absolutely linked to the *sed4-1*, *sed43-1* or *sed44-3*, candidate areas were established by analyzing the positive correlations of the flowering time and candidate markers. The identified positions remain to be confirmed by analyzing F3 lines that segregate

for these putative linked regions (Figure 3.8). The fact that several areas segregate with a *L er* constitution might indicate either that the trait is caused by more than one locus that can act in an additive fashion, in which case it would look more as a Quantitative Trait Loci (QTL), or that the effect observed in the suppressor is favoured in a *L er* background and the plants collected to generate the mapping populations were enriched in this genotype. In both cases, it would be necessary to dissect the components that have an effect on the trait by reducing the complexity of the mapping population. One way to do this could be by creating Recombinant Inbred Lines (RIL) where only one of the loci remains polymorphic and its effect on the trait can be tested separately.

Remarkably, the flowering time trait and marker association analysis in the background *esd4-1(L er) x esd4-2 (Col)* did not show any association (Section 3.3, Figure 3.9), if only the late flowering individuals were analyzed, indicating that the patterns observed in the *sed4-1* and *sed44-3* mapping populations are due to the suppressor mutations, and that in this control cross the phenotypic variation can not be attributed to one gene but it is likely the product of a complex multiallelic interaction.

Next-Generation sequencing as a method of fine mapping the *sed111-1* suppressor

After working with different mapping populations and establishing rough map positions for different *sed* mutations I realized that the advance to narrow down the linked areas would be hampered for different reasons like the variation in flowering time of the genetic background and the sensitivity of the mutant phenotype to the environmental conditions. These inherent characteristics of the suppressors and the *esd4* mutation therefore became a limiting factor to screen for larger populations to fine map the causative mutations. I decided to accelerate the mapping process by using Whole-genome sequencing to identify the causative mutation in the *sed111-1* line. This strategy provided information on all the SNPs exclusively present in the *sed111-1* line. The corroboration of the linked genetic interval in a cross of *sed111-1* back to the *esd4-1* mutant avoided the need for a wide cross between accessions. For this *de novo* dCAPS, designed to use the SNP as part of the restriction enzyme site,

were used and supported the validity of the established rough map position (Figure 3.10 and 3.13). This technology proved to be very useful since only 13 SNPs were found in the linked area even if the genetic interval established by the analysis of recombinants is relatively large (26.49 -28.07 Mbp). These 13 candidates represent only 3% of the total number of loci (469) found in this area. From these 13 mutant candidates five SNPs cause non-synonymous mutations and one is located at an intron/exon junction (At1G71240).

Whole-genome sequencing has also unravelled complex genetic interactions by identifying mutants whose phenotypes depend on more than one mutation (Srivatsan et al. 2008; Harper et al. 2011). In the extremely rare case that *sed111-1* phenotype would be due to a synergic effect of independent mutations it would be necessary to analyze and determine the contribution of each mutation to the phenotype by generating individuals with single or combinations of the mutations. A limitation for this procedure would be the proximity of the potential mutations which would not allow to generate double crosses of independent T-DNA lines in the candidate genes to generate the mutant combinations. In this case the backcross to *L er* results are useful because the isolation of recombinants that only retain the genetic interval with the candidate genes will recapitulate the suppressor phenotype.

Overall, our results demonstrate that the new sequencing technologies are well suited for the isolation of suppressors of a mutant with a pleiotropic phenotype as *esd4*. This approach helped to accelerate the identification of clear candidates for the *sed111-1* causative mutation and proved very useful in a situation where the phenotype is restrictive, as in our case, where the plants have reduced fertility, low seed yields, and the mutant phenotype is susceptible to environmental influences, making precise scoring very difficult.

4. Analysis of a potential role of SUMOylation in the activity of the transcriptional and floral repressor protein Short Vegetative Phase

4.1. Introduction

During *Arabidopsis* development, the transition from the vegetative phase to the reproductive phase is controlled by a complex genetic network that includes a set of transcription factors acting upstream of effector genes (Mouradov et al. 2002; Simpson and Dean 2002; Lee and Lee 2010). Some of these transcription factors belong to the MADS-box protein family and play important roles during this developmental transition (Theissen et al. 2000; Simpson and Dean 2002; Michaels et al. 2003; Alexandre and Hennig 2008). The MADS-box family is characterized by the presence of a highly conserved domain of about 60 amino acids, the MADS-box (Kaufmann et al. 2005). This domain interacts directly with a conserved sequence, the CArG box, in the DNA of its target genes to regulate their expression (Shore and Sharrocks 1995; Pellegrini et al. 1995). MADS-box genes are extensively involved in the regulation of different developmental processes in plants, remarkably in reproductive development (Kaufmann et al. 2005; Theissen et al. 2000; Hemming and Trevaskis 2011).

SHORT VEGETATIVE PHASE (SVP), a member of the MADS-box family, is involved in the regulation of phase transitions in *Arabidopsis* in a dosage dependent manner (Hartmann et al. 2000). These developmental transitions include the flowering time transition and the maintenance of the identity of the floral meristem before the establishment of the identity of the floral whorls (Hartmann et al. 2000; Gregis et al. 2006; Liu et al. 2007; Gregis et al. 2008). SVP interacts with *FLOWERING LOCUS C (FLC)*, another important negative regulator of flowering in *Arabidopsis*, to form a floral repressor complex which binds the promoter regions of *FLOWERING LOCUS T (FT)* and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS (SOC1)* inhibiting their expression (Michaels and Amasino 1999; Searle et al. 2006; Lee et al. 2007; Li et al. 2008). FT and SOC1 are two important floral integrators which promote flowering in response to internal and environmental signals (Turck et al. 2008; Lee and Lee 2010).

Regulation of transcription factor functions by SUMOylation has been reported (Gill 2005). In most cases, SUMOylation has a negative effect on transcription mainly by recruiting or stabilizing repressor complexes (Gill 2005). FLC interacts with both AtSCE and ESD4 in the yeast two-hybrid assay and is SUMOylated *in vitro* (Elrouby, pers. Comm.). In this study, I used the *E. coli* SUMOylation system developed in our lab (Elrouby and Coupland 2010) to show that SVP is also SUMOylated. I also mapped the SUMO-attachment sites in SVP by site-directed mutagenesis. In order to study the role of SUMOylation in SVP function, an *svp*-null mutant (*svp-41*) was transformed with constructs aiming to hyperSUMOylate (translational fusions with SUMO or AtSCE) or hypoSUMOylate (mutations in the putative SUMO-attachment sites) SVP protein.

4.2. SVP SUMOylation

4.2.1 *In silico* analysis predicted seven SUMO-attachment sites in SVP

SUMOylation takes place on lysine residues that usually reside in what is known as the SUMO-attachment site (Rodriguez et al. 2001; Xu et al. 2008). This motif consists of a defined sequence of amino acids (ψ KxE/D) and can be predicted by bioinformatic analysis (Xue et al. 2006; Xu et al. 2008). Using SUMOplotTM, a publicly available program (<http://www.abgent.com/tool/sumoplot>), seven putative SUMO-attachment lysines were identified with different score probabilities (Table 4.1).

Table 4.1. Predicted SUMO-attachment sites in SVP

Residue position ^a	SUMO-attachment context ^b	Score ^c
K78	QSKNLE <u>E</u> KLDQPSLE	0.50
K30	RRGLF <u>K</u> KAEELSVL	0.48
K31	RRGLF <u>K</u> KAEELSVL	0.48
K10	EKIQI <u>R</u> KIDNATAR	0.44
K155	EISEL <u>Q</u> KKGMQLMD	0.33
K156	EISEL <u>Q</u> KKGMQLMD	0.33
K53	IFSST <u>G</u> KLFEFCSS	0.32

Amino acids are indicated by single-letter abbreviations. ^aThe number to the right indicates the amino acid position in the primary sequence. ^bSUMO motifs are highlighted in bold letters and the putative SUMO-target lysine is also underlined. ^cScore indicates the probability for the SUMO consensus sequence to be used in SUMO attachment based on two criteria: 1) direct amino acid match to SUMO consensus sites, and 2) substitution of the consensus amino acid residues with amino acid residues exhibiting similar hydrophobicity.

MADS-MIKC proteins contain three conserved domains (**MADS**, **I**ntervening and **K**eratin-like domains) and a C-terminal region (Ma et al. 1991). The **MADS** domain guides the DNA-protein interaction and lies flat on the DNA minor groove (Pellegrini et al. 1995; Santelli and Richmond 2000). The **I** domain plays a role as a key molecular determinant for the selective formation of DNA binding dimers (Riechmann et al. 1996). The **K** domain provides specificity in the formation of protein-protein interactions (Zachgo et al. 1995; Yang et al. 2003; Yang and Jack

2004), and finally, the C-terminal region is likely to be involved in transcriptional activation or in the formation of multimeric transcription factor complexes (Riechmann and Meyerowitz 1997). Interestingly, the majority of the predicted attachment lysines in SVP were located in the MADS domain (K10, 30, 31, 53), with only one in the I domain (K78) and two in the K domain (K155, 156; Figure 4.1).

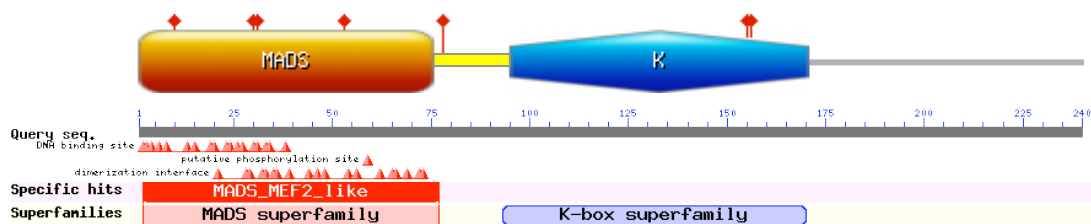


Figure 4.1. SVP protein conserved domains and potential SUMO-attachment sites. SVP presents the canonical three domains and a C-terminal extension of MADS-MIKC proteins. Four SUMO-attachment sites are located in the MADS domain (K10, 30, 31, 53), two are part of the K-box domain (K155, 156) and one (K78) is located in the I region. MADS-box is represented by the red rectangle, the I-domain is in yellow and the K-domain is represented by the blue polygon. Red diamonds on top of each domain represent the SUMO-target lysines. The small red triangles below the gray ruler represent amino acids that participate in the formation of DNA binding surface (upper), conserved phosphorylation site (middle) or the dimerization interface (bottom) (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>).

There are three parameters that can be taken into account to support the functionality of a protein motif: evolutionary conservation, surface accessibility and co-localization in the cell with the enzymes that carry out the posttranslational modification (PTM) (Diella et al. 2008; Diella et al. 2009). To explore the conservation of the potential SUMO-attachment motifs in different SVP homologues, a protein-protein BLAST analysis was performed using the SVP protein sequence as query (<http://www.ncbi.nlm.nih.gov/BLAST>). This analysis retrieved 100 homologous sequences that were then aligned using the Multiple Alignment tool COBALT (<http://www.ncbi.nlm.nih.gov/tools/cobalt>). As expected, the lysines located within the MADS domain (K30, 31 and 53) were highly conserved while lysines 10, 78, 155 and 156 were more variable (Figure 4.2).

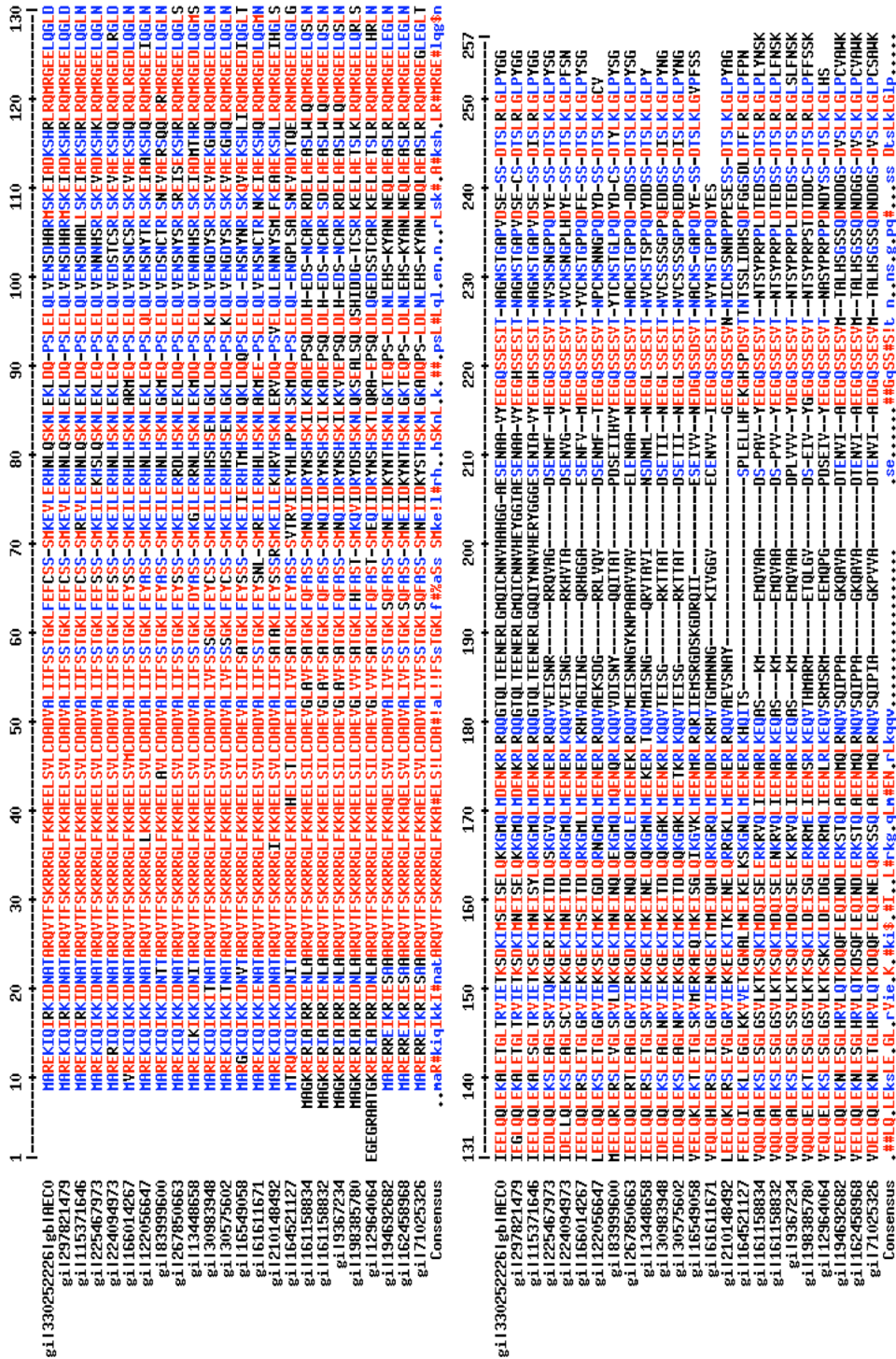


Figure 4.2. Multiple sequence clustal global alignment of potential SVP homologues in different plant species. Alignment of SVP homologues from 23 different species, AtSVP sequence corresponds to the first sequence denoted as gi 330252226. Amino acids are represented as letter code and the colors indicate highly conserved (red), medianly conserved (blue) and polymorphisms (black). Dashes indicate gaps.

Multiple sequence alignments indicated that three out of the seven potential SUMO-attachment lysines were conserved. To determine whether the SUMO-attachment motifs were also conserved, the retrieved 100 putative SVP homologue sequences were used to identify conserved stretches of 10 amino acids in length, that might include the predicted SUMO-attachment sites using MEME, a tool designed to discover motifs in a group of related protein sequences (<http://meme.sdsc.edu/meme>). This analysis revealed that the SUMO-attachment motifs were among the 15 most conserved motifs, with motifs spanning K53, K30/K31, K10, K155/K156, and K78 (ranking 3rd, 4th, 9th, 11th, 15th, respectively) (Figure 4.3).

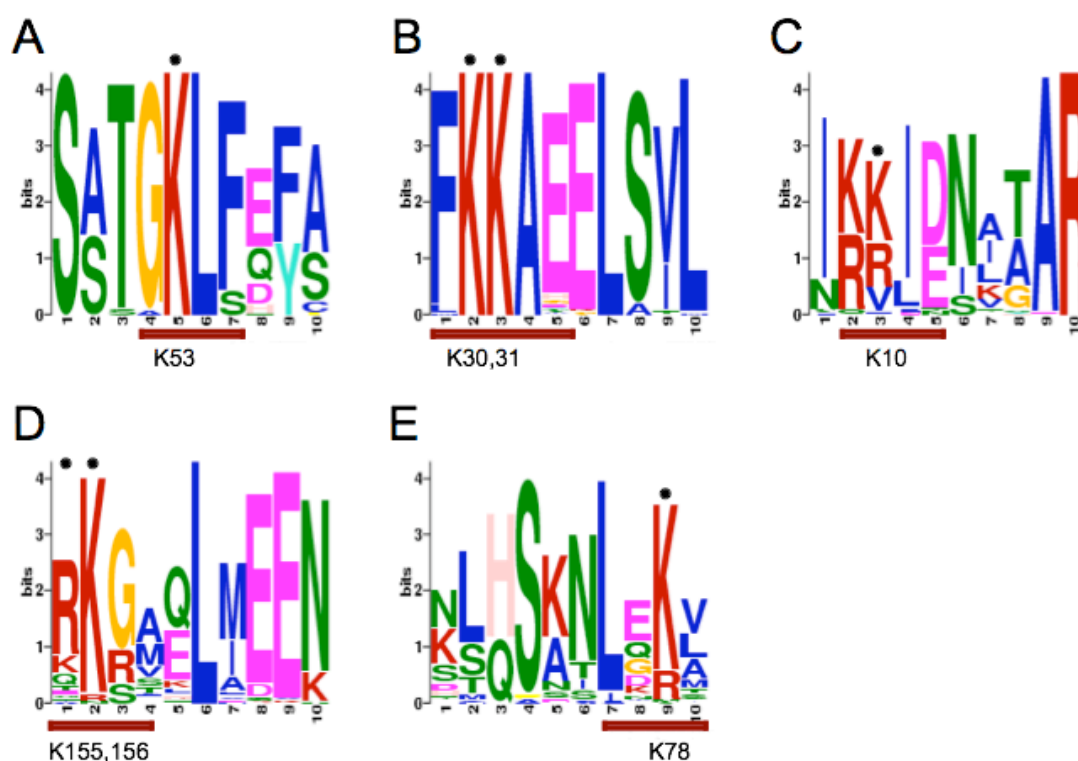


Figure 4.3. Residue conservation of SUMO-attachment motifs in SVP homologues. Each letter represents an amino acid, the height of each letter shows the percentage of sequences that presented such amino acid in that particular position. Since not all the sequences presented all the motifs, the probability is reported based on the number of sequences that did present such motif. In this case all the motifs were present in a minimum 88 out of the 90 sequences and can be considered the same. E-values, that report the statistical significance of the motif, are as follows: Motif A ($1.4e-618$, 90/90 sites), Motif B ($2.2e-576$, 88/90 sites), Motif C ($7.1e-447$, 89/90 sites), Motif D ($1.0e-381$, 89/90 sites), Motif E ($4.2e-325$, 89/90 sites). Numbers on the X-axis indicate the position on the 10 residues motif. Numbers in the Y-axis indicate the “information content” of that position in the motif in bits or position probability matrix. Letter colors indicate the biochemical properties of the various amino acids: Blue, most hydrophobic; Green, polar, non-charged, non-aliphatic; Magenta, acidic; Red, positively charged; Pink=H; Orange=G; Yellow=P; Turquoise=Y. For clarity SUMO motifs are underlined and SUMO attachment lysines are denoted with a dot.

Amino acid residues need to be accessible or exposed to PTM enzymes in order to be posttranslationally modified (Diella et al. 2008). In general, loops or coils are not “regular” protein secondary structures and most of the time are situated at the protein surface so that they are more likely to be targets of PTM (Bernier-Villamor et al. 2002). To determine which SUMO-attachment sites are putatively more accessible on the SVP surface I performed an *in silico* secondary structure prediction analysis using the sequence-based prediction of relative Solvent AccessiBiLitiEs (SABLE, <http://sable.cchmc.org>) for proteins of unknown structure and graphed the results with the protein structure visualization server POLYVIEW-2D (<http://polyview.cchmc.org>). In general, SVP exhibits the two characteristic regular secondary structures predicted for the MADS-MIKC proteins: the MADS domain consisted of two antiparallel alpha helices (Santelli and Richmond 2000) and the K domain which is predicted to form three amphipatic alpha-helices (Yang et al. 2003; Yang and Jack 2004). The SVP MADS-domain contains an alpha helix (residues 29-38) and two antiparallel beta sheets (residues 4-28 and 43-56) which appear interrupted by two short coil sequences (residues 11-17 and 49-53; Figure 4.4). The four SUMO-attachment motifs predicted within the MADS-domain (K10, K30, K31 and K53) are part of structured domains and not likely to be exposed (Figures 4.4 and 4.5). The I-domain, here predicted to form an alpha helix (residues 61-108), seems to be interrupted by a coiled structure (residues 73-83; Figure 4.4). K78 is localized in this I coil region which is also predicted to be disordered (Figure 4.5) and therefore suitable for interaction with the SUMOylating enzymes. The K-domain appears as a long non interrupted alpha helix (residues 121-180; Figure 4.4) in agreement with previous structure predictions for this domain (Yang et al. 2003). Here, lysines K155 and K156 seem to be moderately exposed. Finally, SVP shows a C-terminal region predicted as a long coiled structure, characteristic of all MIKC proteins (Figure 4.4). In this area there are no predicted SUMO-attachment sites.

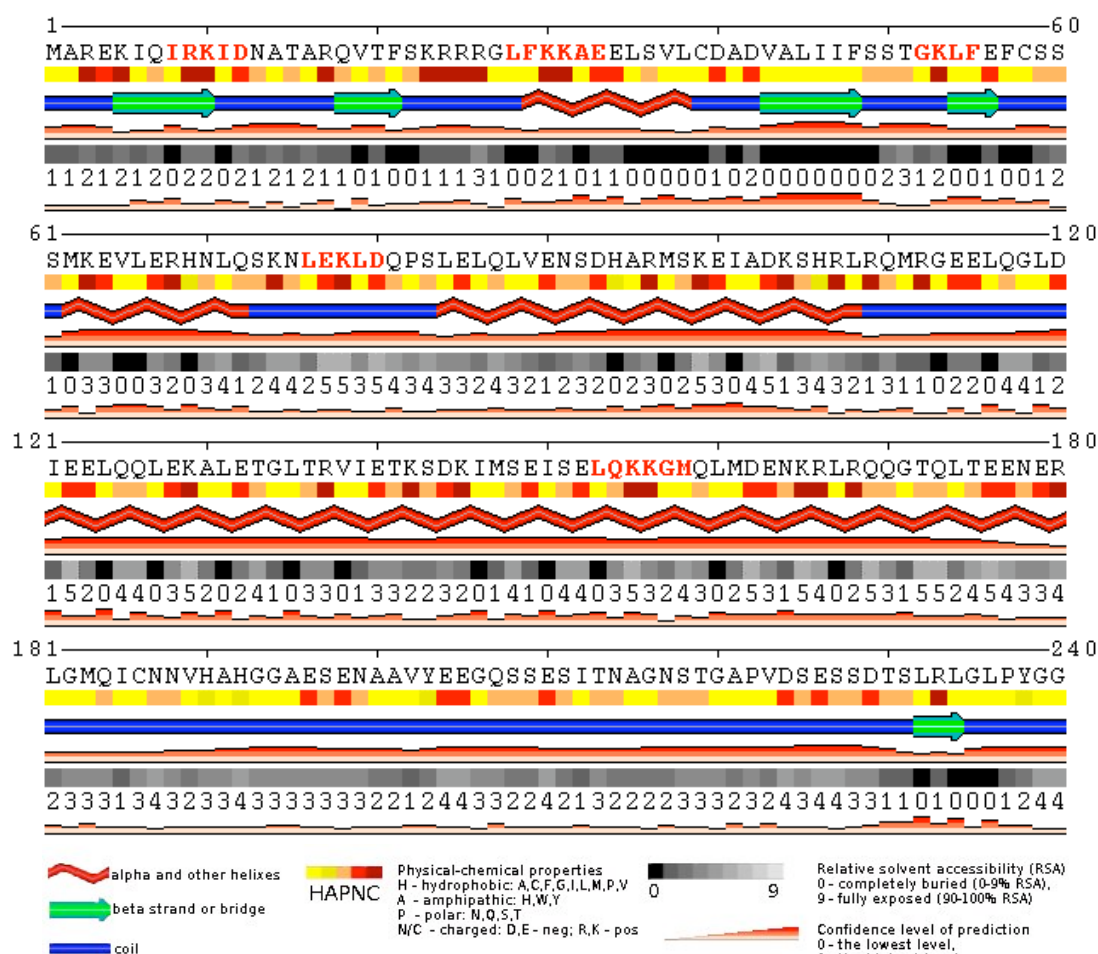


Figure 4.4. SVP secondary structure prediction. SUMO-attachment sites are highlighted in red bold letters. Numbers above the amino acid sequence indicate the amino acid position while the numbers below the gray gradient indicate the Relative solvent accessibility. POLYVIEW-2D: <http://polyview.cchmc.org/>

As mentioned before, surface accessibility is one of the parameters that can be taken into account to support the functionality of a protein motif (Diella et al. 2008). The protein motif surface accessibility is necessary for the temporal interaction with the modifying enzymes (Perkins et al. 2010). Some protein motifs occur outside the structured α -helices or β -strands domains, and are prone to be part of intrinsically disordered regions (Diella et al. 2008). To identify the potential disordered regions in SVP that might contain SUMO attachment sites I analyzed the SVP protein sequence with DisEMBLTM (<http://dis.embl.de/>), a bioinformatic tool that predicts disordered peptide segments based on three different definitions (Linding et al. 2003). This analysis revealed that K78 is the only SUMO-attachment lysine that was consistently found in a disordered region according to the loops and hot loops disorder definitions

and therefore is more likely to be modified. On the other hand, K10, K78, and to a lesser extent K30 and K31 are predicted to be disordered based on the hot loops or high mobility loops prediction definition (Figure 4.5).

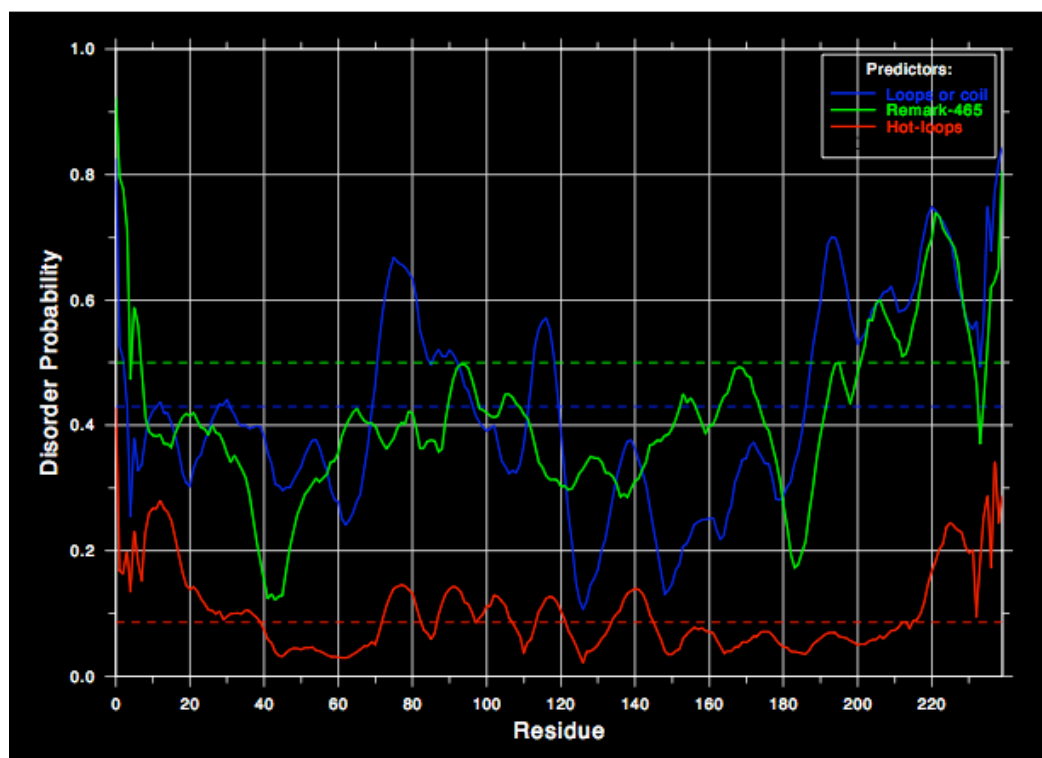


Figure 4.5. Disorder prediction for the SVP protein sequence. Prediction of disordered peptide segments based on three different definitions of disorder. Loops/coils (blue line) are all secondary structures that are not forming an α -helix or β -strand. Hot loops (red line) are loops with a high mobility and therefore predicted to be highly dynamic. REMARK465 (green line) represent the amino acids for which there are no X-Ray structure coordinates reported in the PDB entries and therefore most likely to be intrinsically disordered regions. The horizontal dashed lines represent the random “expectation level for each predictor” (<http://dis.embl.de/>).

In general, the disorder prediction power of a bioinformatic tool can be increased by cross comparing the results with other bioinformatic programs that perform similar analysis based on different methods (Diella et al. 2008). In this case the protein disorder prediction results of DisEMBLTM (Figure 4.5) were compared to the results of SPRITZ (<http://protein.cribi.unipd.it/spritz>), this second analysis supported the prediction that K78 is located in a disordered region (Data not shown). In summary, the bioinformatic analysis of the predicted SUMO-attachment sites in SVP showed that these motifs exhibit characteristics that support them, in different degree, as real motifs. K78 highest SUMOplot score and its location within an exposed and highly disordered region rank it as the most likely SUMO-attachment site (Table 4.2).

Table 4.2. SVP bioinformatic analysis and prediction synthesis					
Residue position	SUMOplot Score	MEME motif conservation E-value	POLYVIEW 2ry Structure	DisEMBL™ Disordered region	
				Coil ^a	Hot ^b
K78	0.50	4.2e-325	Coil	D	D
K30	0.48	2.2e-576	Alpha helix	O	D
K31	0.48	2.2e-576	Alpha helix	O	D
K10	0.44	7.1e-447	Beta sheet	O	D
K155	0.33	1.0e-381	Alpha helix	O	O
K156	0.33	1.0e-381	Alpha helix	O	O
K53	0.32	1.4e-618	Coil	O	O

Indicates the predicted results for Disordered (D) or Ordered (O) region based on the Coil^a or the Hot loop^b disorder definitions

4.2.2 *in vitro* analysis

SVP SUMOylation

SVP sequence analysis revealed seven putative SUMO conjugation sites; each with different properties that would support them as SUMO-attachment sites (Table 4.2). To determine if SVP is a SUMO target, wild type *SVP* cDNA was cloned into pET32b, a bacterial expression vector that allows expression of recombinant proteins as fusions with bacterial thioredoxin (Trx). Trx is used as a tag for immunodetection and to increase solubility of tagged proteins. pET32b-SVP was transformed into *E.coli* strains that contain the genes encoding the SUMO-activating and SUMO-conjugating enzymes as well as one of two SUMO-isoforms (SUMO1 or SUMO3) (Elrouby and Coupland 2010). After induction of gene expression, recombinant proteins were purified, separated by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE) and subjected to western blotting for immunodetection. In addition to the expected 55 KDa Trx-SVP fusion protein, several high molecular weight forms distributed between 72 to 95 KDa were detected (Figure 4.6A). These may correspond to SUMO-modified Trx-SVP. Two very large forms (170 and 130 KDa) were also observed and may represent Trx-SVP multimers, that are possibly SUMOylated as well (Figure 4.6A). The forms larger than 72 kDa did not appear if a mutant form of the E2 enzyme was used, providing support for them being SUMOylated forms of Trx-SVP. To test further if those high molecular weight

entities were indeed SUMOylated forms of SVP I performed immunoblot analysis again using anti-SUMO1 or anti-SUMO3 antibodies. This experiment confirmed that the high molecular entities were SUMOylated versions of SVP and also revealed that the SVP SUMO1 and SUMO3 SUMOylation patterns were different (Figure 4.6B).

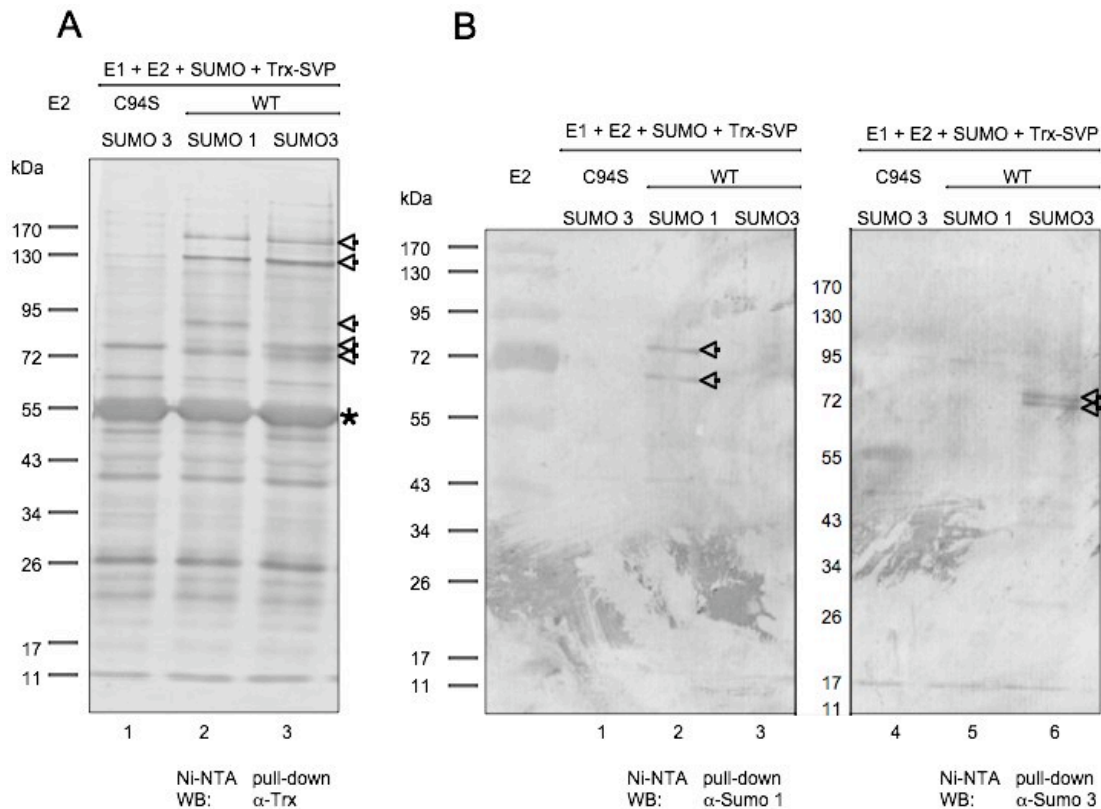


Figure 4.6. SVP is SUMOylated *in vitro* by SUMO1 and SUMO3. SVP exhibits a characteristic SUMOylation pattern in an *E.coli* SUMOylation system depending on the SUMO isoform provided. Total protein extracts of different cell lines were obtained, recombinant proteins were purified on Ni-NTA and fractionated by SDS-PAGE, blotted a membrane and probed using an anti-Trx (A) or anti-SUMO antibodies (B). Lane 1, Negative control where the SUMOylation strain is impaired by a point mutation in the SCE. Lane 2, SUMOylation strain using SUMO 1. Lane 3, SUMOylation strain using SUMO 3. Asterisk indicate the Trx-SVP fusion protein and arrows indicate SUMOylated SVP. Relative molecular mass standards are indicated with numbers to the left of each blot.

Determination of SUMOylation sites in SVP

From our previous experiments we concluded that SVP, at least in this *E.coli* SUMOylation system, is SUMOylated by both SUMO1 and SUMO3 and that the

SUMOylation pattern changes according to the SUMO isoform provided. In the next section, we attempt to utilize the *E. coli* SUMOylation system described above to determine which of the predicted SUMO-attachment lysines can actually be modified by SUMO.

SUMO1 sites

To assess which of the predicted SUMO-attachment sites were the targets of modification by SUMO1, I performed site-directed mutagenesis to replace each one of the seven predicted SUMO-attachment lysines by arginine. The *SVP* cDNA clones containing single point mutations (K10R, K30R, K31R, K53R, K78R, K155R and K156R) were subcloned into pET32b and expressed in the *E. coli* SUMOylation system as previously described for wild type *SVP*. After induction of gene expression, recombinant proteins were isolated by Ni-NTA purification under denaturing conditions, separated onto SDS-PAGE, immunoblotted and probed with anti-SUMO1 antiserum to test for modification of SVP by SUMO1. As shown in Figure 4.7, the K10R, K30R, K31R and K156R mutations seem to decrease the abundance of one of the SUMO-modified SVP forms but none of them abolished SVP SUMOylation completely.

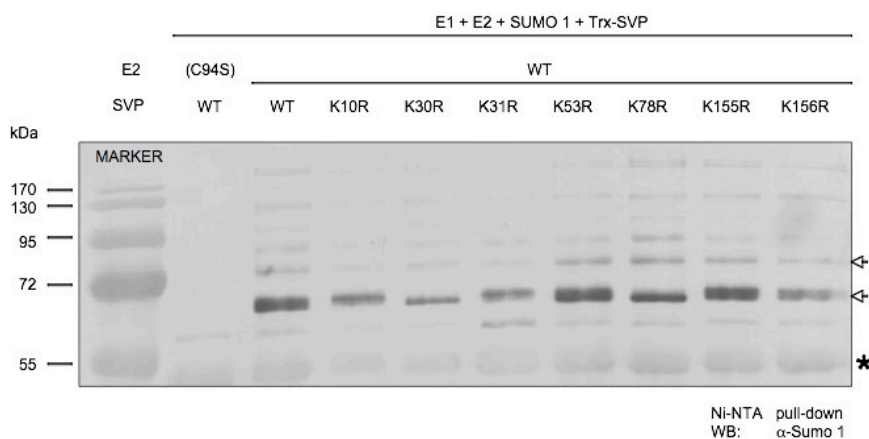


Figure 4.7. Effect of single point mutations on SVP SUMO1 SUMOylation. SVP presents a characteristic SUMO1 SUMOylation pattern of two bands (arrows) of around 72 KDa. Total protein extracts from independent cell cultures transformed with different mutant versions of SVP were fractionated by SDS-PAGE and immunoblotted using a specific anti-SUMO1 antibody. A Wild type (WT) or mutant version (C94S) of the E2 was provided. Arrows indicate SVP SUMOylated fractions. The asterisk indicates the expected Trx-SVP fusion protein size. Relative molecular mass standards are indicated with numbers to the left of the blot.

Since none of the single point mutations were able to abolish SUMO1 modification of SVP, we tested whether combining two mutations may abolish modification by SUMO1. All possible combinations of double mutants were established and analyzed in the *E coli* SUMOylation strains as before. This approach proved to be informative since three of the double mutant combinations (K30,53R; K30,78R and K31,53R) reduced significantly SUMO1 attachment (Figure 4.8B). Whereas these three mutant combinations caused the upper band (SUMO1-modified SVP, denoted by the upper white arrow) to disappear, several other mutant combinations only caused a slight reduction of the intensity of this band (Figure 4.8). Remarkably, none of the double mutant combinations affected the abundance of the lower band, suggesting that none of the combinations abolished SVP-SUMO1 modification completely (Figure 4.8).

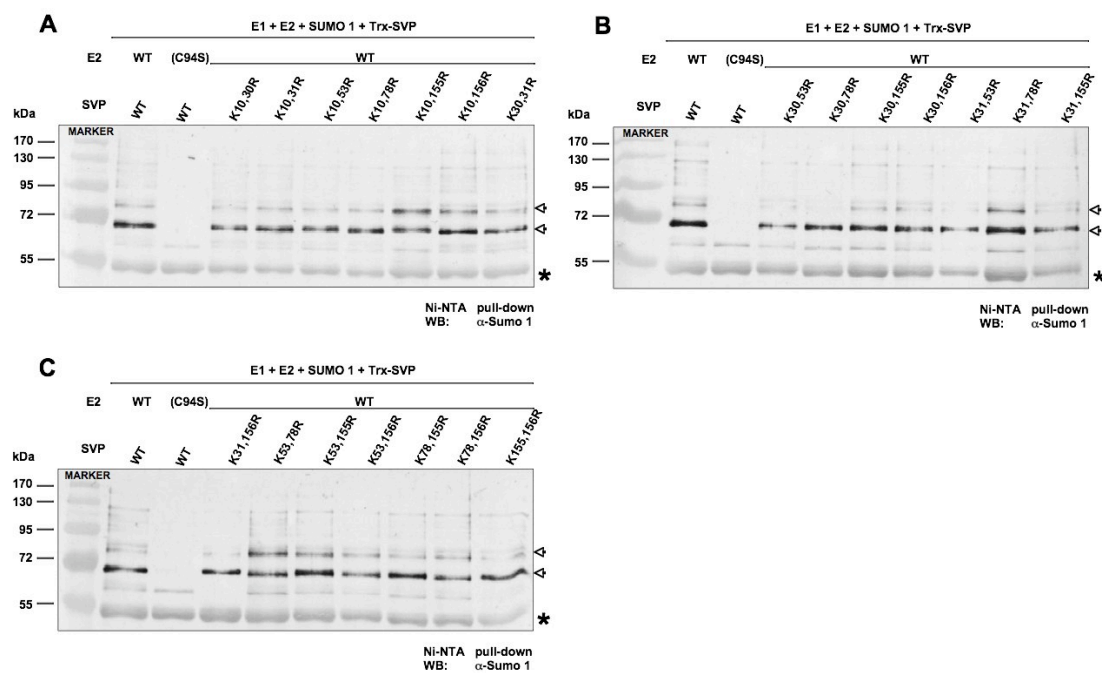


Figure 4.8. Effect of double point mutations on modification of SVP by SUMO1. SVP is modified by SUMO1 (two bands indicated by the arrows). Total protein extracts from independent cell cultures transformed with different double point mutant versions of SVP were fractionated by SDS-PAGE and immunoblotted using anti-SUMO1 antibody. A Wild type (WT) or mutant version (C94S) of the E2 were used as positive and negative controls, respectively. Arrows indicate SUMOylated SVP forms while the asterisk indicates the expected Trx-SVP fusion protein size. Relative molecular mass standards are indicated with numbers to the left of the blot.

As seen above, none of the single or double point mutants of SVP abolished SUMOylation by SUMO1 completely. Although the upper of two SUMO1-modified forms disappeared when K30,53R; K30,78R or K31,53R were used, the lower SUMO1-modified form persisted. We combined five point mutations (K30R, K31R, K53R, K78R, K156R) to generate a Penta-SUMO-site Mutant form of SVP (referred as PSM-SVP) and tested whether it can still be SUMOylated by SUMO1. Immunoblot analysis suggests that the PSM-SVP completely abolished SUMOylation of SVP by SUMO1 (Figure 4.9).

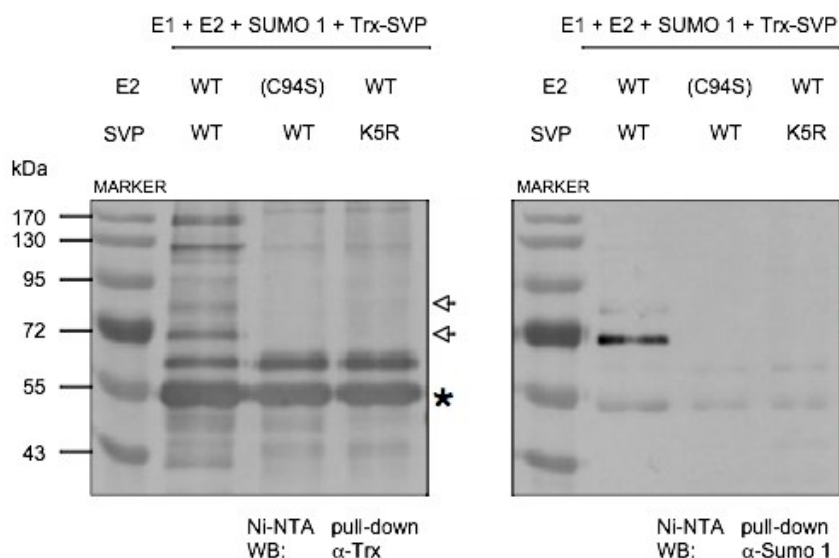


Figure 4.9. SUMOylation of SVP by SUMO1 is abolished in the K30, 31, 53, 78, 156R Penta-SUMO-site Mutant. SVP exhibits a characteristic SUMO1 SUMOylation pattern mostly consisted of two bands (arrows). Total protein extracts from independent cell cultures transformed with different mutant versions of SVP were purified by Ni-NTA fractionated by SDS-PAGE and immunoblotted using a specific anti-Thioredoxin or anti-SUMO1 antibody. As indicated a Wild type (WT) or mutant version (C94S) of the E2 was provided. Arrows indicate SVP SUMOylated fractions. The asterisk indicates the expected Trx-SVP fusion protein size. Relative molecular mass standards are indicated with numbers to the left of the blot.

SUMO3 sites

Previous results indicated that SVP is SUMOylated by SUMO1, in *E coli* strains expressing the SUMO conjugation machinery, at lysine residues located at positions 30, 31, 53, 78 and 156 (Figure 4.9). To assess whether the SVP SUMOylation pattern

may change with the SUMO isoform used, the same set of SVP mutants used above were tested in *E. coli* SUMOylation strains expressing SUMO3 instead of SUMO1. Similar to SUMOylation by SUMO1, none of the single point mutations was effective in impairing SVP SUMOylation by SUMO3 (Figure 4.10). In contrast, when the set of double mutants used previously with SUMO1 was used to test SUMOylation by SUMO3 (Figure 4.11), I found that the K30,78R double mutation (Di-SUMO-site Mutant, or DSM) abolishes SVP SUMOylation by SUMO3. This is intriguing since complete abolition of SVP SUMOylation by SUMO1 required the mutation of five lysine residues whereas mutation of only two of these lysines are sufficient for loss of its SUMOylation by SUMO3, underlying differences in SUMO isoform target site selection.

In summary, I mapped the SUMOylation sites in SVP to lysines K30, K31, K53, K78, K156 for SUMO1, and to lysines K30, K78 for SUMO3. Lysines K30 and K78 are common targets of SUMO1 and SUMO3 suggesting a major role for these lysines in SVP SUMOylation.

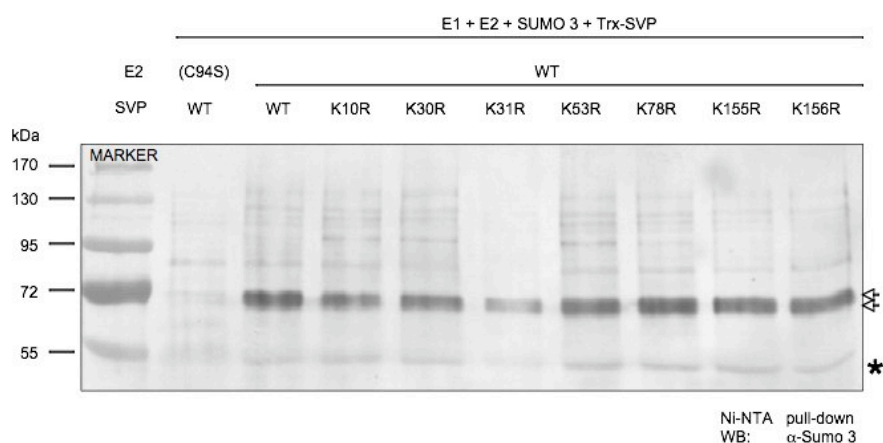


Figure 4.10. Effect of single point mutations on SUMOylation of SVP by SUMO3. SVP exhibits a characteristic SUMO3 SUMOylation pattern mostly consisting of two very close bands of around 72 kDa (arrows). Total protein extracts from independent cell cultures transformed with different mutant versions of SVP were fractionated by SDS-PAGE and immunoblotted using anti-SUMO3 antibody. A wild type (WT) or mutant version (C94S) of the E2 was used. The asterisk indicates the expected Trx-SVP fusion protein size. Relative molecular mass standards are indicated with numbers to the left of the blot.

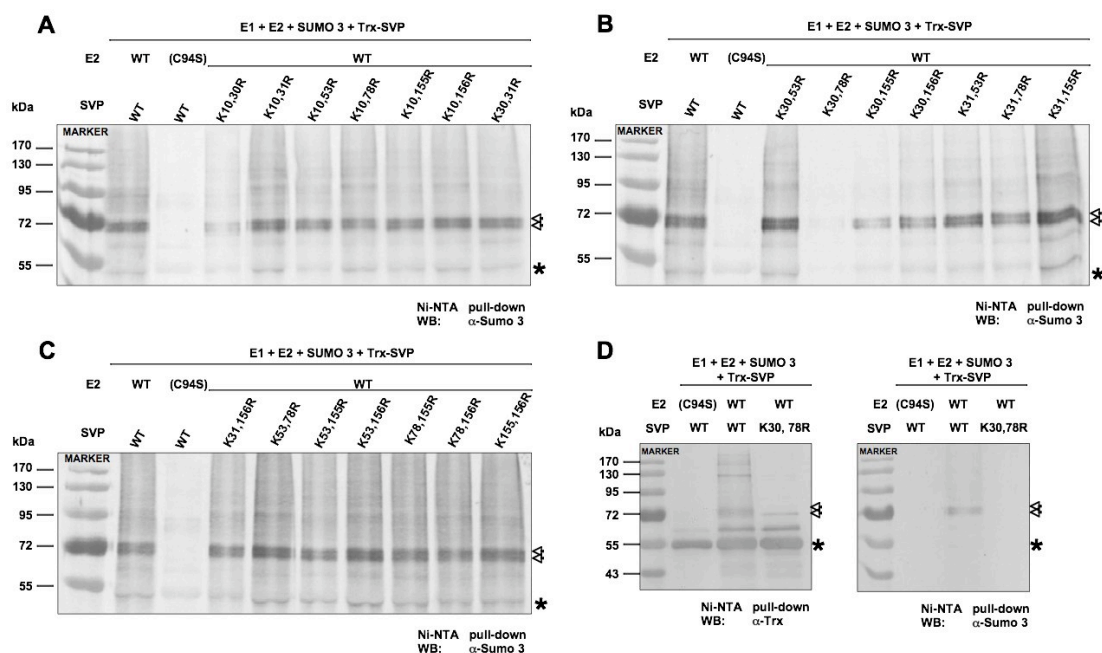


Figure 4.11. Effect of double point mutations on SUMOylation of SVP by SUMO3. SVP exhibits a characteristic SUMO3 SUMOylation pattern mostly consisting of two close bands (arrows). Total protein extracts from independent cell cultures transformed with different double point mutant versions of SVP were fractionated by SDS-PAGE and immunoblotted using a specific anti-SUMO3 antibody. A Wild type (WT) or mutant version (C94S, as negative control) of the E2 were provided. Arrows indicate SVP SUMOylated fractions. The asterisk indicates the expected Trx-SVP fusion protein size. Relative molecular mass standards are indicated with numbers to the left of the blot.

4.3. Investigations towards a potential role of SUMOylation in SVP function-*in vivo* analysis

Experiments using the *E. coli* strains described above suggest that SVP can be SUMOylated in that system by both SUMO1 and SUMO3, and that at least five lysines (at positions 30, 31, 53, 78, 156) may attach to SUMO1 whereas only two of these (positions 30, 78) are modified by SUMO3. To test whether these sites also attach SUMO *in vivo*, and to assess the significance of SVP SUMOylation *in planta*, I carried out a series of experiments using a null Arabidopsis mutant of SVP (*svp-41*). I transformed *svp-41* with different constructs expressing translational fusions of SVP with SUMO or the SUMO-conjugating enzyme (AtSCE), or mutant versions of SVP (where SUMO-attachment lysine were converted to arginine). These constructs are likely to represent states of SVP where it is “hyperSUMOylated” or “hypoSUMOylated” (Figure 4.12).

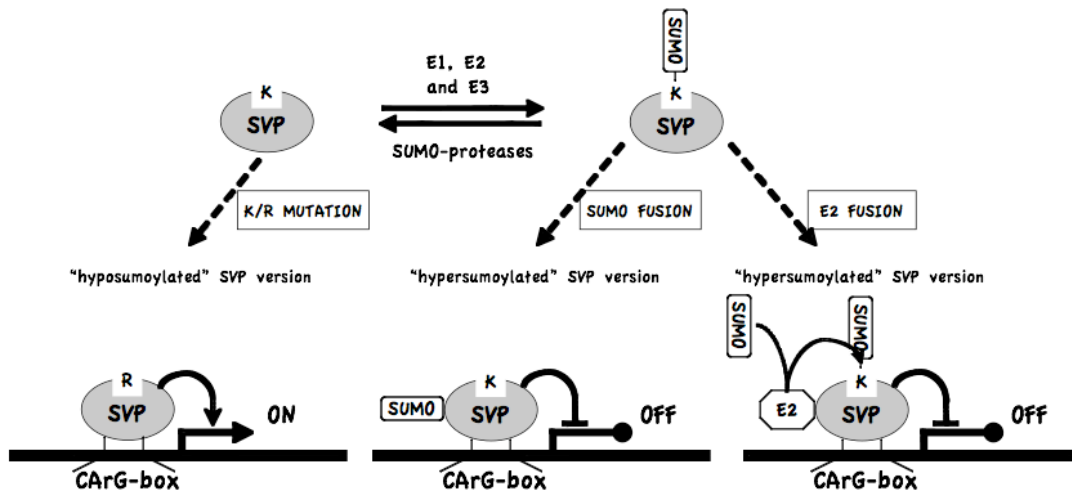


Figure 4.12. Explicative model for the creation of “Hyper” and “Hypo” SUMOylated versions of SVP. SUMOylation involves the action of E1, E2 and E3 enzymes. The SUMOylated substrate can be restored to the non-SUMOylated version by the action of SUMO-proteases. A hypoSUMOylated or SUMO-impaired version of SVP can be created by mutating the putative SUMO-attachment lysines to arginines (Left) while hyperSUMOylated versions can be obtained by generating translational fusions of SVP with either SUMO (Middle) or the SUMO Conjugating Enzyme (SCE or E2, Right). E2 translational fusions have been shown to enhance SUMOylation of the fused protein in a constant and specific manner (Jakobs et al. 2007). In this model it is assumed that SUMOylation of the transcription factor SVP provokes transcriptional repression of the target gene. Diagram adapted and modified from (Ouyang, Valin, and Gill 2009).

As mentioned before, SVP and FLC participate in a repressor complex that suppresses the expression of *FT* and *SOC1*, two genes required for the transition to flower in Arabidopsis (Li et al. 2008). As a trait, the time to flower can be determined in Arabidopsis by the number of leaves that the plant produces at the onset of flowering since the apical meristem makes a transition from a vegetative state (which produces leaves) to a reproductive state (which produce flowers) (Koornneef et al. 1991). The apical meristem is then referred to as the “inflorescence meristem”(Alvarez-Buylla 2010). Therefore the degree of genetic complementation of the flowering-time phenotype (registered as the total number of leaves at flowering) of the *svp-41* mutant will reflect the activity of SVP. The analysis of the “hyperSUMOylated” or “hypoSUMOylated” versions of SVP in this process will give some indication of the role of SUMO in SVP function.

4.3.1 Translational fusions

Previous studies have shown that MADS-box proteins are very sensitive to translational fusions, both N- and C-terminal (de Folter et al. 2007). This is most likely because important functions of MADS-box proteins are mediated by its N-terminal and C-terminal regions. For instance, the N-terminal domain of MADS-box proteins interacts directly with the DNA while the C-terminal region participates in protein-protein interactions (Theissen et al. 2000). However, not all of the MADS-box proteins respond in the same way to translational fusions, therefore it is necessary to assess each protein specifically (de Folter et al. 2007). To circumvent these potential complications we constructed both N-terminal and C-terminal fusions as explained in the following sections.

N-terminal fusions

We constructed a series of N-terminal fusions where SUMO1 or SUMO3 was fused to the coding region of SVP at its N-terminus. We also fused the SUMO-conjugating enzyme (AtSCE) or a mutant version of AtSCE (C94S) to the N-terminus of SVP (Figure 4.12). SUMO translational fusions may be able to mimic the effects of SUMO on the target protein and perform as the natively SUMO-conjugated protein in the cell (Ross et al. 2002; Ouyang et al. 2009). In some cases, SUMO fusion constructs even enhance the protein SUMO-dependent activities (Nayak et al. 2009; Kang et al. 2010; Smith et al. 2011). On the other hand, fusions with the SCE have been suggested to enhance SUMOylation of the associated protein in an efficient and selective way (Jakobs et al. 2007). The C94S mutation converts the active site cysteine to a serine and renders the AtSCE inactive so it may act as a negative control.

The results indicated that only two independent T2 lines of the 35S::FLAG-SUMO1-SVP construct (2_2 and 2_9) were able to complement the early flowering phenotype of *syp-41* mutant (Figure 4.13). On the other hand, the rest of the constructs, including the 35S::FLAG-SVP control, caused all plants to flower just slightly later than the *syp-41* mutant (around 10 to 11 leaves; Figure 4.13). This suggests that the N-terminal translational fusions may interfere with the proper function of SVP.

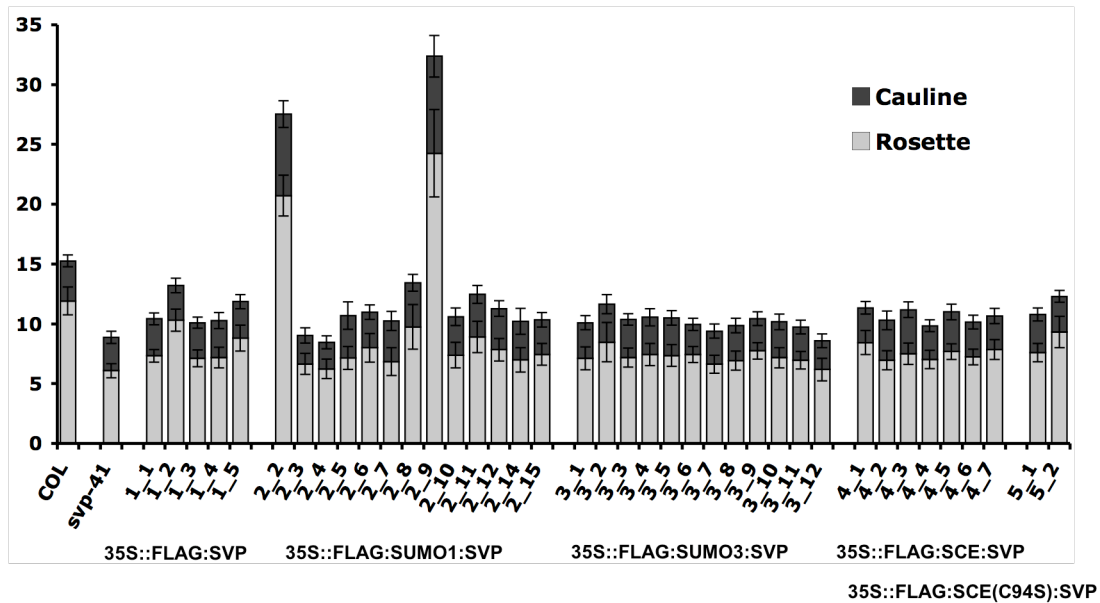


Figure 4.13. Flowering time of independent T2 lines overexpressing SVP amino-terminal fusions to SUMO or the SCE grown under long day conditions. Flowering time is presented as the Total Leaf Number (TLN). Y axis indicate the number of leaves while the X axis show the genotypes analyzed: 1=35S::FLAG:SVP, 2=35S::FLAG:SUMO1:SVP, 3=35S::FLAG:SUMO1:SVP, 4=35S::FLAG:SCE:SVP, 5=35S::FLAG:SCE(C94S):SVP. Error bars illustrate the means \pm SD of $n \geq 25$ plants.

Previous reports have shown that wild type plants transformed with a 35S::SVP construct flower late (Lee et al. 2007). Our 35S::FLAG-SVP construct in *svp-41* mutant background does not rescue the early flowering phenotype of *svp-41* (Figure 4.13). On the other hand, two independent 35S::FLAG:SUMO1:SVP lines rescued the *svp-41* mutant and even strongly delayed the flowering relative to wild type plants, indicating that SVP is functional in these plants. It is possible that the N-terminal FLAG-tag in the 35S::FLAG-SVP construct interferes with the proper function of SVP in this and the other constructs, whereas FLAG:SUMO1 is cleaved off from the FLAG:SUMO1:SVP protein by endogenous SUMO proteases releasing a free and functional SVP protein in those late flowering lines. To test this hypothesis we performed a whole protein extract from these independent FLAG:SUMO1:SVP lines and analyzed the levels of recombinant protein by western-blot using an anti-FLAG antibody. As suspected, a unique band of around 17 KDa corresponding to the expected size for the FLAG:SUMO1 fusion protein was identified (Figure 4.14C-D). In summary, these results allow us to conclude that the N-terminal fusions interfere with the proper function of SVP (Figure 4.13), and that the late flowering phenotype

observed in the FLAG:SUMO1:SVP lines is due to the removal of the FLAG:SUMO1 part by the action of SUMO-proteases allowing the normal function of SVP (Figure 4.14C-D).

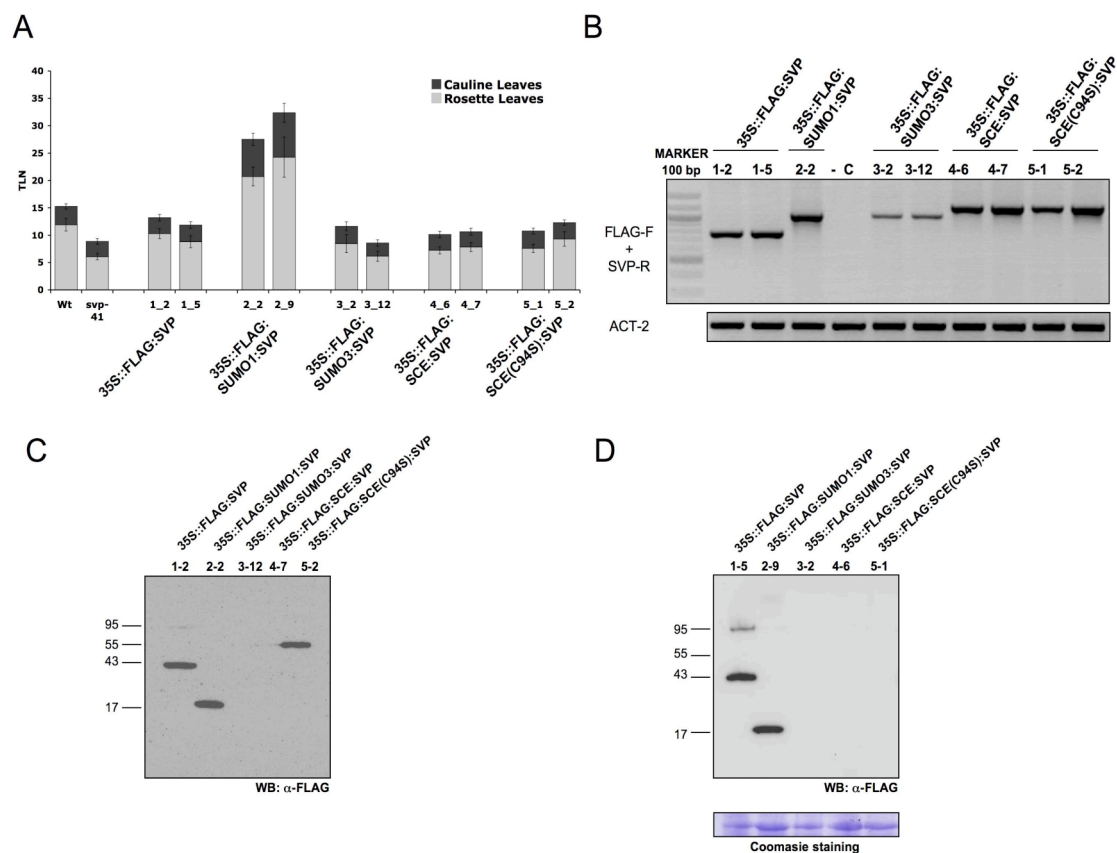


Figure 4.14. Molecular characterization of the FLAG N-terminal SVP fusion constructs. Two independent lines per construct were characterized in better detail, numbers in the top right corner of the pictures indicate the construct. (A) Flowering time of two independent T2 lines per construct under Long days. (B) RT-PCR expression analysis of the different N-terminal constructs. For this analysis the specific forward primer, indicated as FLAG-F in the figure, for each N-terminally attached gene was used in combination with an SVP specific reverse primer. Actin-2 gene (ACT-2) was amplified as a quantitative control. (C-D) Western blots protein expression analysis of independent T2 transgenic lines. Total protein extracts from 15 day-old seedlings were fractionated by SDS-PAGE and immunoblotted using an anti-FLAG antibody. Numbers to the left of each blot indicate the molecular mass in kilodaltons. In panel D a portion of the gel stained with Coomassie Brilliant Blue is shown as a loading control.

C-terminal fusions

The same set of fusion constructs were created but with SUMO1, SUMO3, SCE, or SCE(C94S) fused at the C-terminal end of SVP. In contrast to the N-terminal constructs where all except for the 35S::FLAG:SUMO1:SVP show the same flowering time as the *svp-41* mutant in the T1 generation (Figure 4.13), the C-terminal constructs caused a delay in the flowering time in comparison to *svp-41* already in the T1 generation, which was enhanced in the T2 generation (Figure 4.15). Although these lines flowered later than *svp-41*, they still flowered at values comparable to wild type. This flowering time phenotype suggests that the C-terminal fusions might interfere less with flowering time functions of SVP.

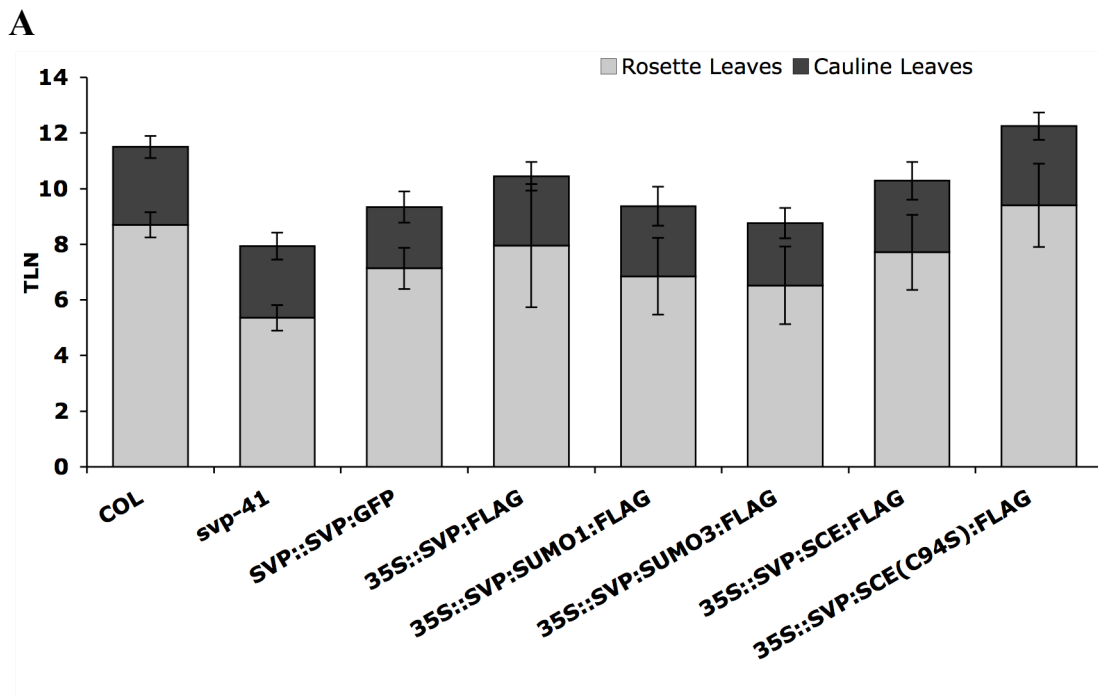
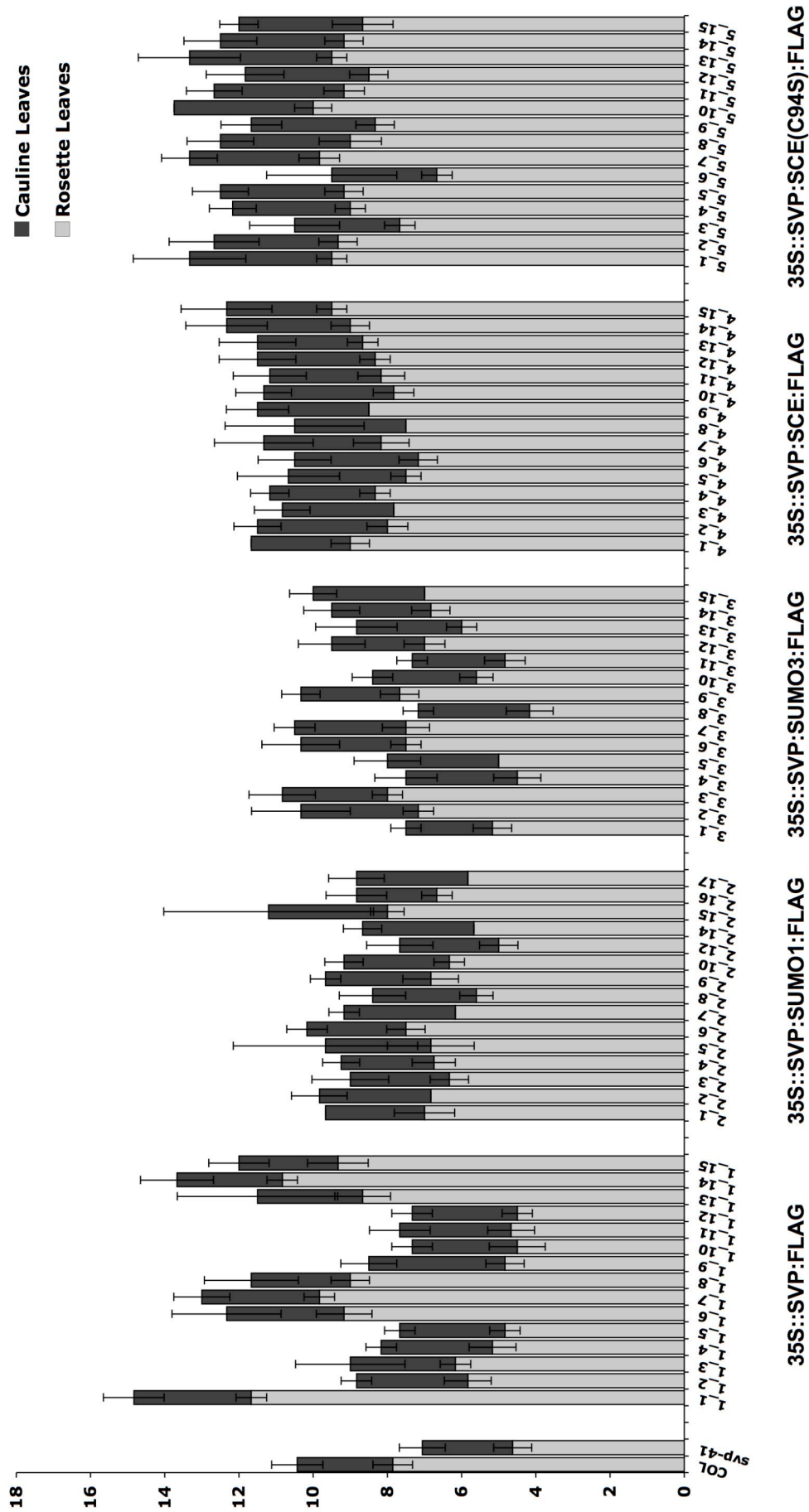
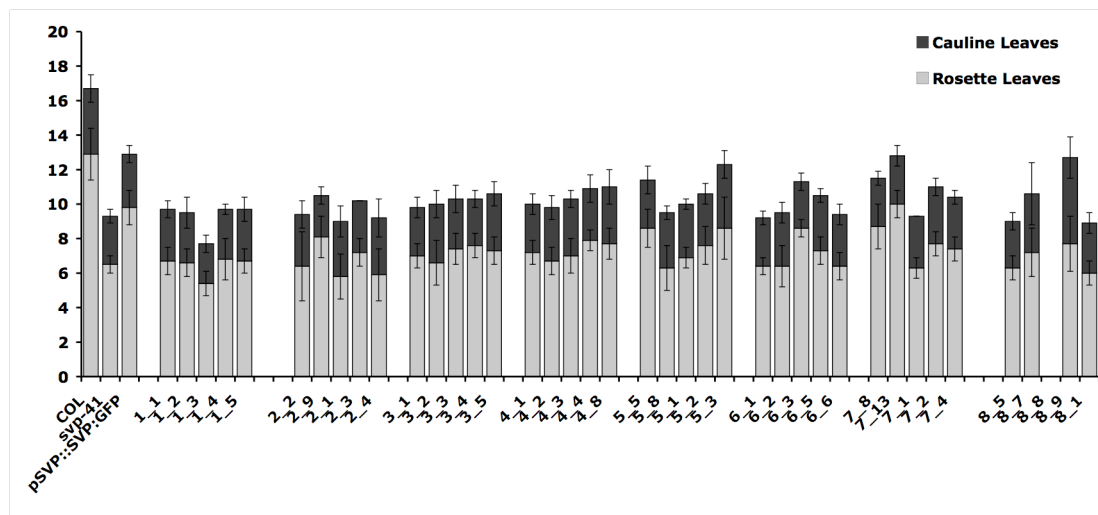


Figure 4.15. Flowering time of independent lines overexpressing SVP carboxyl-terminal fusions to SUMO or the SCE grown under long day conditions. Flowering time is presented as the Total Leaf Number (TLN) produced per genotype for T1 (A) or T2 (B, see next page) generations. Y axis indicate the number of leaves while the X axis show the genotypes analyzed. 1=35S::SVP:FLAG, 2=35S::SVP:SUMO1:FLAG, 3=35S::SVP:SUMO3:FLAG, 4=35S::SVP:SCE:FLAG, 5=35S::SVP:SCE(C94S):FLAG. Error bars illustrate the means \pm SD of: $n \geq 15$ plants for the controls, $n \geq 20$ plants for the T1 and $n \geq 6$ plants for the T2.



4.3.2 The effect of mutations of SUMO-site lysines on SVP function

The PSM and DSM mutations that abolished SUMOylation of SVP by SUMO1 and SUMO3 in *E. coli* were also used to transform the *svp-41* mutant. The purpose of this experiment was to assess whether these lysines are important for SVP function. A FLAG tag was included at the N-terminus to allow detection of the proteins on Western-blot. The results suggest that none of the constructs were able to complement the early flowering mutant phenotype of *svp-41* in the T1 generation (Data not shown), and only a subtle delay in flowering was observed in the T2 generation (Figure 4.16). Since the SUMO-site mutant constructs were made at the same time as the N- and C- terminal fusions described in the previous section we were not able to foresee the negative effects of the small FLAG tag on the function of SVP in the regulation of flowering time. Thus, it is difficult to conclude whether the inability of these constructs to delay flowering is caused by the mutations introduced in the recombinant SVP (PSM and DSM) or due to rendering the protein inactive by the N-terminal FLAG. However, I observed phenotypic alterations in the flowers of some of these SUMO-site mutant constructs. Such mutant phenotypes were never observed in the 35S::FLAG:SVP plants, which serve as a control, indicating that the induced mutations might have an impact on the function of SVP.



1. 35S::FLAG-SVP^{K30,31,53,78,156R}
2. 35S::FLAG-SVP^{k30, 31, 53R}
3. 35S::FLAG-SVP^{k30, 31R}
4. 35S::FLAG-SVP^{K53, 78, 156R}

5. 35S::FLAG-SVP^{K78,156R}
6. 35S::FLAG-SVP^{K30, 78R}
7. 35S::FLAG-SVP^{K53, 156R}
8. 35S::FLAG-SVP^{30, 31, 78R}

Figure 4.16. Flowering time of independent T2 lines overexpressing different versions of SVP where the putative SUMO-attachment sites are mutated. Flowering time is presented as the Total Leaf Number (TLN) produced per genotype under Long Day conditions. Y axis are number of leaves while the X axis indicate the construct. Error bars illustrates the means \pm SE of $n \geq 8$ plants.

4.3.3 Morphological abnormalities associated with SVP overexpression

Phenotypes associated with N-terminal SVP fusions.

In the previous section I showed that the FLAG:SUMO1:SVP was the only construct that restored the *svp-41* early flowering phenotype and that this effect was likely due to the cleavage of the FLAG:SUMO1 part from the SVP protein allowing restoration of a functional SVP. On the other hand, the rest of the constructs do not recover the early flowering *svp-41* phenotype. Strikingly, some of these constructs caused morphological alterations in the *svp-41* background which, except for an early flowering phenotype, is otherwise indistinguishable from wild type plants. In this section I briefly describe the molecular and phenotypic characterization of the observed alterations and speculate on a possible explanation.

To better characterize the N-terminal SVP constructs I decided to analyze in more detail two of the strongest lines per construct. In agreement with the previous results, this analysis revealed that only the FLAG:SUMO1:SVP construct caused a late flowering phenotype while the other constructs caused an intermediate phenotype between Wild type and *svp-41* (Figure 4.14A). Analysis of the mRNA expression in the different transgenic plants indicated that all transgenes were fully and strongly expressed, with the exception of the 35S::FLAG:SUMO3:SVP which contained very little mRNA (Figure 4.14B). To determine if the levels of mRNA correspond to the amount of protein synthesized I analyzed the protein levels by Western-blot using an anti-FLAG antibody. Plants expressing FLAG:SVP presented two bands, an expected band of around 43 kDa which was detected in high amount, and a second unexpected and faint higher molecular weight band of around 95 kDa (Figure 4.14C-D). In the case of the FLAG:SUMO1:SVP lines I only identified a very abundant band of around 17 kDa which corresponds to the expected size for the FLAG:SUMO1,

confirming the cleavage of this part by endogenous SUMO-proteases (Figure 4.14C-D).

An independent western-blot analysis using an anti-SUMO1 antibody confirmed the identity of this band (Data not shown). No protein was detected for FLAG:SUMO3:SVP in any of the two lines tested (Figure 4.14C-D). FLAG:SCE(C94S):SVP and FLAG:SCE:SVP presented a band of around 54 kDa corresponding to the expected size for the fusion protein, interestingly, the FLAG:SCE:SVP band was very faint although both constructs presented similar mRNA expression levels (Figure 4.14 B-C).

FLAG:SUMO1:SVP lines exhibited a dark green colour and grew slower than the rest of the lines (Figure 4.17A). On the other hand, FLAG:SUMO3:SVP and FLAG:SCE:SVP T2 lines presented plants that segregate with similar phenotypes, these plants were small with rosettes consisting of distorted and unarranged leaves (Figure 4.17A). Whether this phenotype is caused by SUMOylation of SVP by SUMO3 remains to be tested. Intriguingly, these effects are not shared by the 35S::FLAG:SVP nor the 35S::FLAG:SCE (C94S) lines. It would be interesting to characterize more independent homozygous lines to determine if there is a correlation of the protein levels and this phenotype.

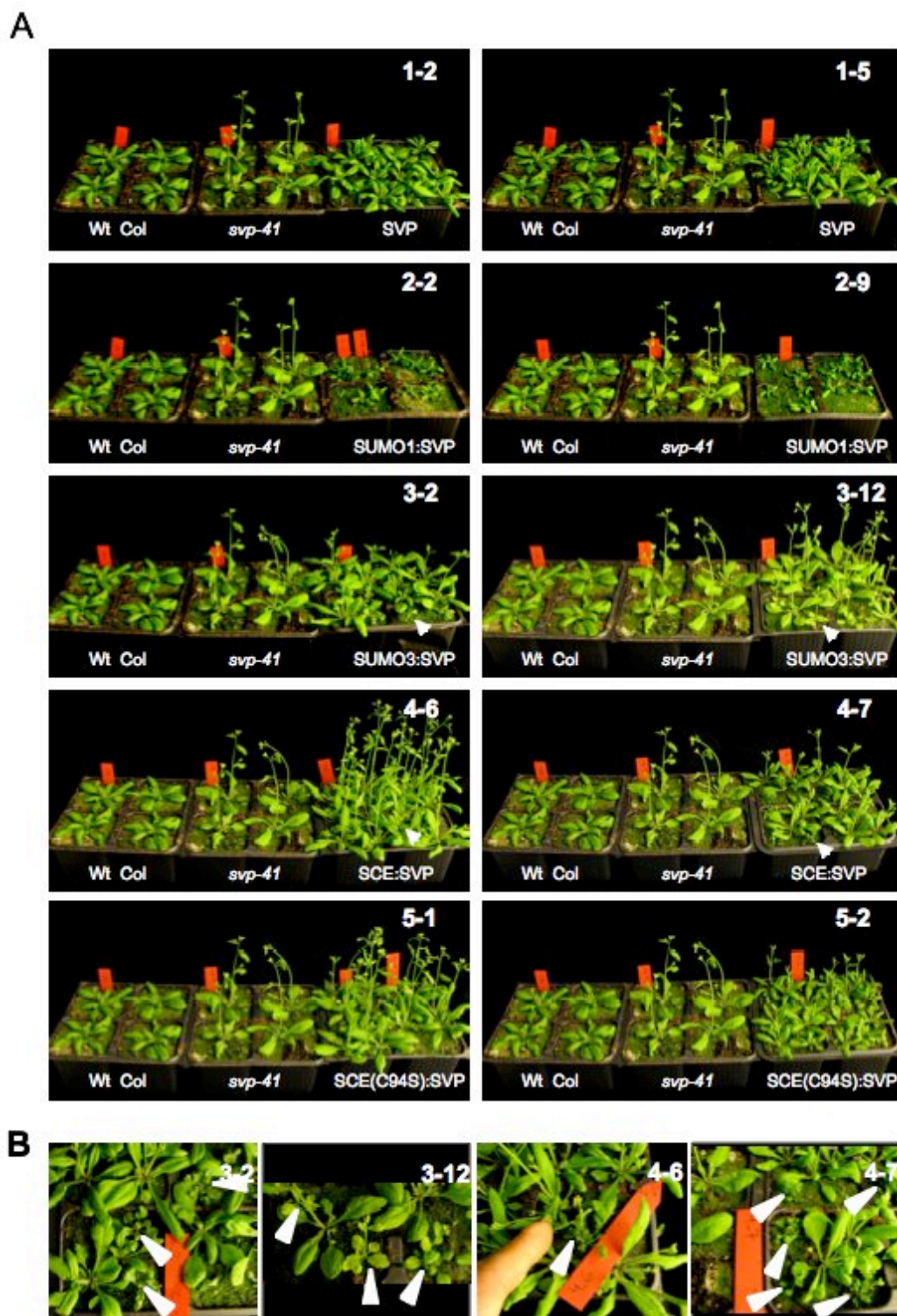


Figure 4.17. Phenotypical characterization of the FLAG N-terminal SVP fusion constructs under Long Days. Two independent lines per construct were characterized in better detail, numbers in the top right corner of the pictures indicate the construct. (A) Photographs of plants from independent T2 lines, the number to the upper left in each picture indicates the construct-line. (B) Plant phenotype details for the FLAG:SUMO3:SVP construct in lines 3-2 and 3-12 as well as FLAG:SCE:SVP construct in lines 4-6 and 4-7.

Further analysis of mature plants uncovered that FLAG:SUMO1:SVP plants developed flowers with six sepals and seven petals (Figure 4.18C). Other interesting phenotypes included the brief acquisition of a purple colour in the sepals, petals and siliques; and the bended form of the siliques at early developmental stages (Figure 4.18).



Figure 4.18. Phenotypes of 35S::FLAG:SUMO1:SVP, *svp-41* line under Long Day conditions. (A) FLAG:SUMO1:SVP lines present flowers where the carpels are bent, this hook like shape persisted in the elongating siliques and tended to disappear later when the siliques are fully developed. Sepals are retained until late stages of silique development in contrast to its almost immediate abscission in Wt plants. (B) Sepals, petals and siliques present a pink-purple coloration that disappear later on the organ development. Flowers present hypernumerary petals (C) and sepals (D).

The 35S::FLAG:SUMO3:SVP lines exhibited a pleiotropic phenotype that included the production of several secondary inflorescences and pale and slightly narrow leaves. Some plants also produced disordered rosettes and fewer and thinner inflorescences (Figure 4.19). On the other hand, some of the 35S::FLAG:SCE:SVP plants presented multiple secondary inflorescences, a slightly pale colour and flowers that did not fully develop (Figure 4.20). These plants produced no seeds.

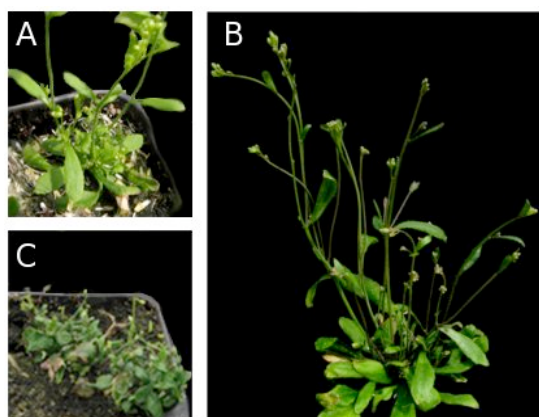


Figure 4.19. Phenotype of 35S::FLAG:SUMO3:SVP, *svp-41* line grown under Long Day conditions. FLAG:SUMO3:SVP lines present narrow pale leaves (A), several secondary inflorescences as well as multiple organs emerging from the internodes (B) or a dwarf phenotype which includes small rosettes composed of crumpled leaves (C).



Figure 4.20. Phenotype of 35S::FLAG:SCE:SVP, *svp-41* amino-terminal fusion construct under Long Day conditions. p35S::FLAG:SCE:SVP, *svp-41* lines present a dwarf phenotype including many secondary branches. Flowers are small and mostly do not fully develop.

Phenotypes associated with C-terminal SVP fusions.

As noted above, C-terminal fusions of SVP seem to be less disruptive of SVP function in floral transition. However, as with the N-terminal fusions, morphological alterations were also observed in plants expressing these C-terminal fusions. The 35S::SVP:SUMO1:FLAG lines presented individuals with fused stems, sometimes these fused stems split giving rise to chimerical internodes where several organs, including cauline leaves, inflorescence stems and flowers, grew out from the same axillary meristem (Figure 4.21A-B). Another common phenotype was the appearance of flowers with five sepals and six petals, very similar to what we observed for the FLAG:SUMO1:SVP lines. Occasionally, a small flower grew out from an existing flower, usually from one of the older floral whorls (e.g. petals), or flowers with a chimeric organ structure (half-sepal half-petal) appeared (Figure 4.21C-D). Some

plants expressing 35S::SVP:SCE:FLAG produced smaller pointy leaves as well as a main shoot whose internodes did not elongate and where the flowers started to open very close to the rosette (Figure 4.22A-B).



Figure 4.21. 35S::SVP:SUMO1:FLAG lines recover the early flowering phenotype of *svp-41* and present morphological alterations in the aerial shoot and flowers under Long Day conditions. (A) SVP:SUMO1:FLAG line presenting fused stems. (B) Multiple organs -like cauline leaves, inflorescences and flowers- coming from the same internode. (C) Flowers frequently present five sepals and six petals and no single stamen (upper arrow), rarely flowers also presented a second flower emerging from the same whorl as the petals (arrow down left). (D-C) Flower picture details showing a flower with leaf-like sepals (D) and a flower showing, greenish petals, no stamens and a chimerical “sepal-petal” organ where half of the organ is a white-thricome-less region while the other half shows a green-thricome region (E).

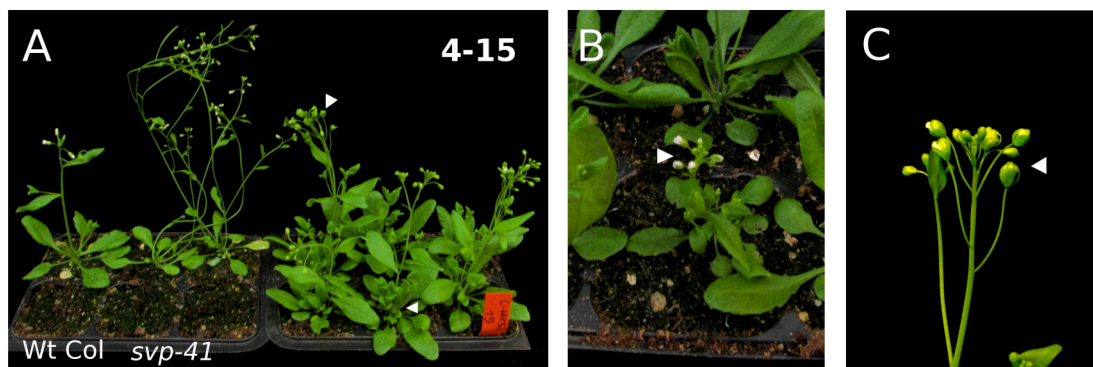


Figure 4.22. 35S::SVP:SCE:FLAG lines recover the early flowering phenotype of *svp-41* and flower later than the Wt under Long Day conditions. (A) Some lines present an increment in the production of secondary leaves in the rosette and multiple secondary inflorescences (lower arrow right), and round flowers (upper arrow left). (B) Independent lines segregate smaller plants that flower as late as the Wt. The flowers of these plants start to open without previous elongation of the shoot showing bunches of open flowers. (C) Photograph detail of the inflorescence of one of the SVP:SCE:FLAG lines showing the ball-like buds. These characteristic buds are formed by overgrown sepals that are bound to each other by the margins.

35S::SVP:SCE(C94S):FLAG lines showed a very consistent late flowering phenotype (Figure 4.23), in many independent T2 lines small plants segregated (Figure 4.23A). These small plants presented the same characteristics as the ones of the SVP:SCE:FLAG line (Figure 4.22B). These phenotypes included smaller rosettes, pointy leaves, stunted main shoot growth and late flowering (Figure 4.23A and 4.22B).

Finally, the 35S::SVP:FLAG lines did not present other phenotypes than the recovery of the early flowering phenotype of the *svp-41* mutant to a flowering time similar to the Wt and some ball-shaped like flowers that appear more frequently in most of the late flowering plants. No visible and/or a characteristic phenotype could be determined for the SVP:SUMO3:FLAG line (Data not shown).

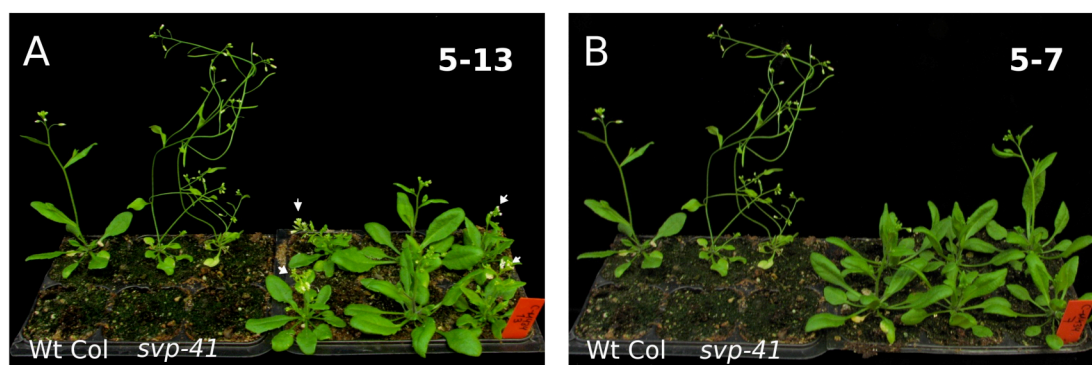


Figure 4.23. 35S::SVP:SCE(C94S):FLAG lines flower later than the Wt under Long Day conditions. (A) Independent lines segregate smaller plants that flower as late as the Wt, and very similar to the ones segregating in the SVP:SCE:FLAG line. The flowers of these plants present very short internodes and flower without previous elongation of the shoot (indicated by white arrows). (B) In general all the plants with the SVP:SCE(C94S):FLAG construct in independent lines flower later than the Wt.

In summary, the C-terminal translational fusions prove to be less disruptive for SVP protein function than the N-terminal fusions as suggested by a modest recovery of the early flowering phenotype of the *svp-41* mutant in plants expressing C-terminal fusion constructs. Remarkably, both N- and C- terminal fusions produce morphological alterations in the aerial parts of transgenic plants. These alterations may be directly caused by increased activity of a functional SVP. Alternatively, they may be caused indirectly through unregulated interference of recombinant SVP with complexes to which SVP normally contributes or to heterologous complexes.

Homeotic phenotypes caused by SUMO-site mutations of SVP

FLAG:SVP-PSM plants senesced faster than the *svp-41*, FLAG:SVP and the wild type plants (Figure 4.24). This phenotype was observed in independent experiments and was not seen for any of the other construct lines, however, formal experiments are required to confirm this phenotype.

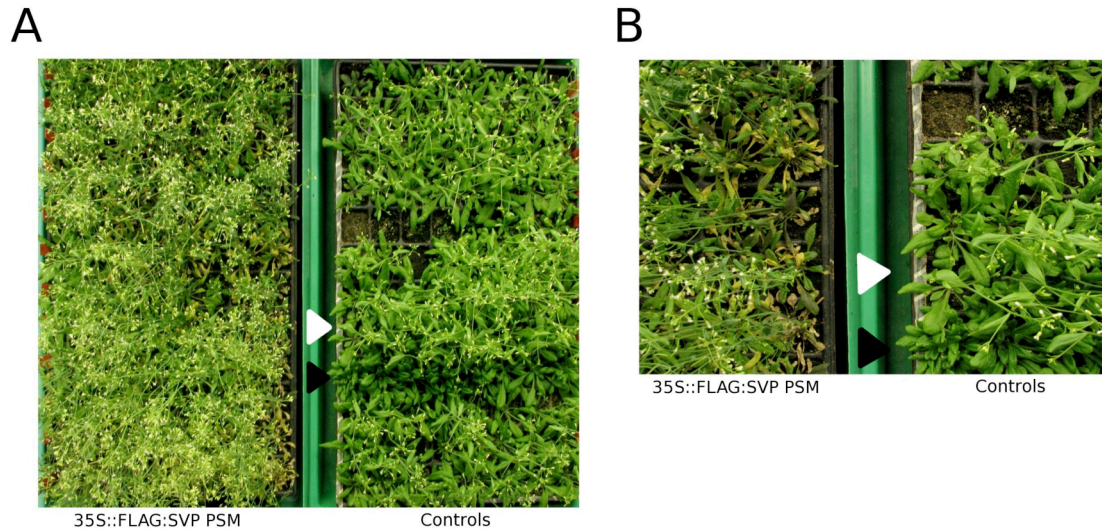


Figure 4.24. 35S::FLAG:SVP^{K30,31,53,78,156R} or FLAG:SVP PSM in *svp-41* show a faster senescence in Long day conditions. (A) Photograph showing a tray with 12 independent T2 FLAG:SVP PSM transformants (left) and another tray with Wt and *svp-41* plants (right). All the FLAG:SVP PSM plants have already flowered and started to senesce in comparison to the other plants. (B) Close up showing the yellowish-purple senescing leaves of the FLAG:SVP PSM plants (left) in comparison to the green leaves of the controls (right). Arrowheads indicate the genotype per row, black corresponds to Wt Columbia, white indicates *svp-41*, the rest of the plants in the tray correspond to 35S::FLAG:SVP.

FLAG:SVP^{K30,31,53R} plants produce flowers that occasionally contain secondary flowers emerging from the second whorl. These secondary flowers are normally smaller than the primary flower but exhibit a normal perianth and carpels. The primary flower also contains leaf-like sepals, four or sometimes only two petals, a variable number of stamens (always less than produced by wild type flowers) and apparently normal carpels (Figure 4.25). On the other hand FLAG:SVP^{K30,31,78R} lines

exhibit a very dramatic phenotype where inflorescence shoots are produced instead of flowers (Figure 4.26). $FLAG:SVP^{K53,78,156R}$ lines presented flowers with alterations in the first three whorls. In the strongest mutant flowers sepals are broad and resemble a leaf structure, and no petals or stamens are formed (Figure 4.27). A similar phenotype was observed for the plants expressing $FLAG:SVP^{K78,156R}$ which causes floral abnormalities that include the disappearance of the stamens and the production of chlorophyll-containing petals; in some cases there is also an increase in the number of organs in the perianth from four organs per whorl to five (Figure 4.28).



Figure 4.25 $35S::FLAG:SVP^{K30,31,53R}$ lines present secondary flowers emerging from the second flower whorl of the primary flower. Photographs were taken from different individuals from independent lines.

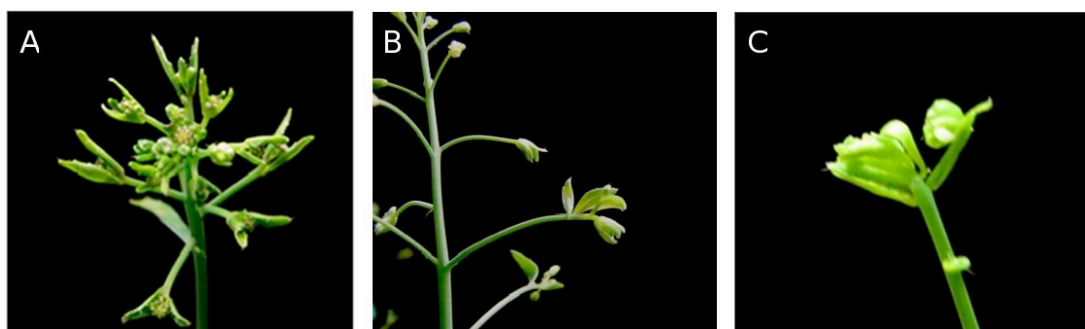


Figure 4.26. $35S::FLAG:SVP^{K30,31,78R}$ plants present inflorescence shoots instead of flowers. Inflorescence in an early (A) or almost completely elongated state (B) showing whorls of leaves instead of flowers. Detail of a “flower” consisted only of organs that resemble a mixture of sepal-leaves (C).



Figure 4.27 35S::FLAG:SVP^{K53,78,156R} lines present alterations in the first three whorls of the flower. Such alterations include flowers with broad sepals, greenish and small petals; and in extreme cases no stamens are formed (flower to the right).

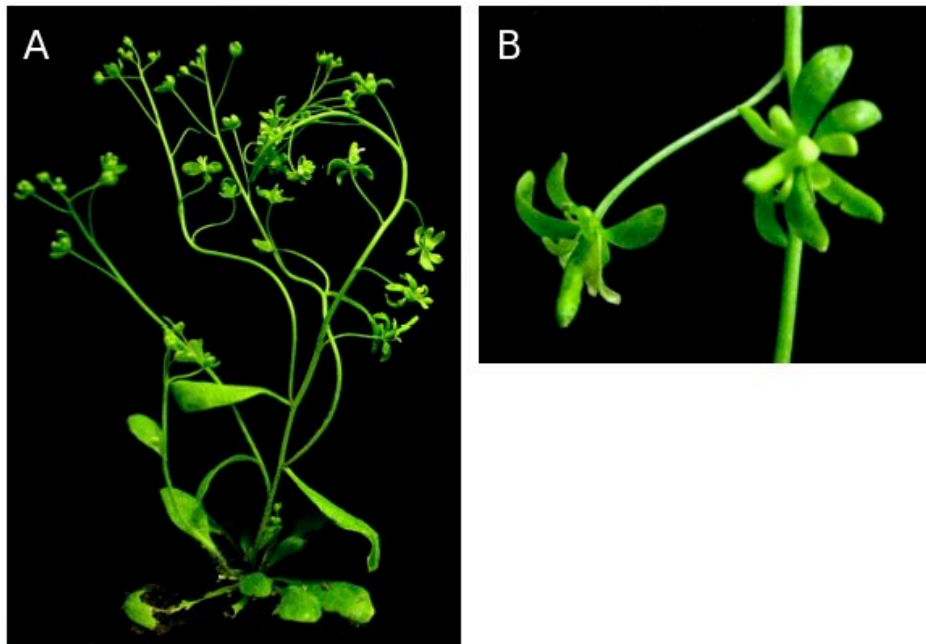


Figure 4.28. 35S::FLAG:SVP^{K78,156R} plants present flowers with alterations in the number of organs per floral whorl and conversions of petals into small green structures when grown under Long Day conditions. (A) p35S::FLAG:SVP plants are normal in size but present flowers with leaf-like sepals and green petals. (B) Flower detail showing two flowers that present leaf-like sepals, green petals and no stamens.

The levels of mRNA and protein in the plants that present the strongest phenotypes were analyzed to determine if the introduced constructs were expressed in the plant. This analysis showed that all the constructs were expressed at the level of mRNA but in different amounts, and that the levels of mRNA correlate with the amount of protein detected. Interestingly, the high molecular weight form of around 95 kDa that was seen before (Figure 4.14C-D) is present in the FLAG:SVP^{K30,31,53R} line (Figure 4.29).

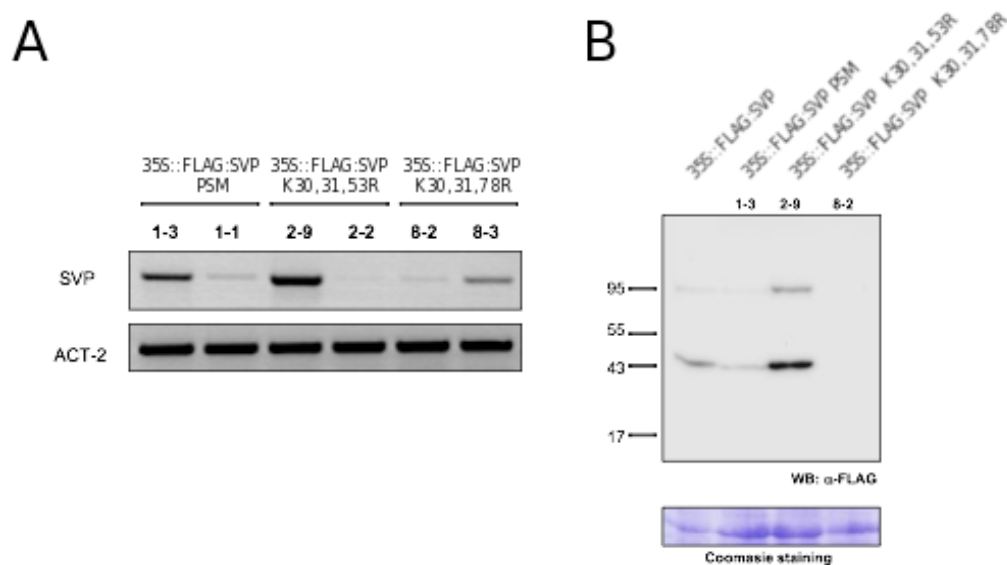
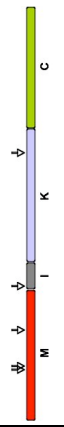
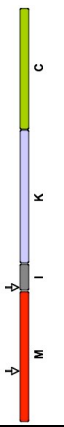
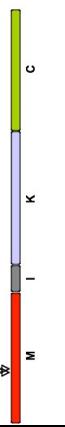


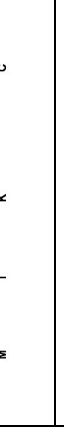
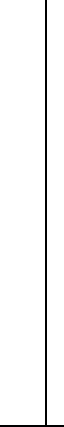



Figure 4.29. mRNA and protein expression analysis in the FLAG:SVP PSM and FLAG:SVP^{K30,31,53R} and FLAG:SVP^{K30,31,78R} construct lines. (A) RT-PCR expression analysis of different SUMO-site mutant constructs. For this analysis the SVP gateways primers, were used. Actin-2 cDNA (ACT-2) was amplified as a quantitative control. (B) Western blot protein expression analysis of independent T2 transgenic lines. Total protein extracts from 15 day-old seedlings were fractionated by SDS-PAGE and immunoblotted using an anti-FLAG antibody. Numbers to the left of the blot indicate the molecular mass in kilodaltons. A portion of the gel stained with Coomassie Brilliant Blue is shown as a loading control.

In summary, the introduction of point mutations in the SVP sequence promotes the formation of particular mutant phenotypes. The FLAG:SVP PSM construct causes premature senescence while no phenotype was observed for the FLAG:SVP DSM. Interestingly, the FLAG:SVP^{K30,31,78R} line, which only has an extra amino acid substitution (K31R) compared with FLAG:SVP DSM, presents a strong phenotype consisting of the conversion of flowers into shoot-like structures which produce only leaf-like organs. Finally, the FLAG:SVP^{K30,31,53R} line present extra secondary flowers that emerge from the second whorl of the primary flower. In general, all the above mentioned lines, which present a phenotype, expressed the transformed constructs at the level of mRNA and protein, with the exception of FLAG:SVP^{K30,31,78R} line for which no protein was detected due to low mRNA levels.

TABLE 4.3. SVP SUMO-site mutants and its effects on flower development

DOMAIN LOCALIZATION OF THE POINT MUTATIONS IN SVP	CONSTRUCT NUMBER	LYSINES MUTATED	PHENOTYPE	MUTANTS WITH SIMILAR PHENOTYPE
	1	K30,31,53,78,156R	Abolition of SVP SUMO1 SUMOylation in <i>E.coli</i> Plants senesce faster than <i>syp-41</i>	
	6	K30,78R	Abolition of SVP SUMO3 SUMOylation in <i>E.coli</i> Plants similar to <i>syp-41</i>	
	3	K30,31R	Plants similar to <i>syp-41</i>	
	2	K30,31,53R	Production of secondary flowers from the axils of the sepals or petals	<i>apl</i> , <i>apl-1 syp-41</i> , <i>agl24 syp apl</i> , <i>35S::AGL24 35S::SVP</i> , <i>35S::OsMADS47</i>
	8	K30,31,78R	Production of leaf-like shoots instead of flowers	<i>sep1 sep2 sep3 sep4</i> (Pelaz et al,2000; Ditta et al, 2004)
	4	K53,78,156R	Flowers with big and wide leaf like sepals, green petals, stamens sometimes absent, perianth with more than four organs per whorl	35S::AGL24 (Michaels et al, 2003), <i>lfy-2 sep3-2</i> (Liu et al, 2009)
	7	K53,156R	Plants similar to <i>syp-41</i>	
	5	K78,156R	Flowers with big and wide leaf like sepals, green petals, stamens sometimes absent, perianth with more than four organs per whorl	<i>lfy-2 sep3-3</i> (Liu et al, 2009)

4.4. Discussion

SVP-SUMOylation

SUMOylation is an important post-translational modification that modulates the function of several transcription factors in different species (Colby *et al*, 2006; Gill, 2005; Miura *et al*, 2007). SVP is a MADS-box transcription factor involved in the regulation of meristem phase transition in Arabidopsis (Hartmann *et al*. 2000). In this study I showed that SVP can be SUMOylated in an *E.coli* SUMOylation system, and I established that a different SUMOylation pattern is visible depending on the SUMO isoform provided (Figure 6). In the case of SUMOylation by SUMO1, two forms of SUMO1-modified SVP were detected and may correspond to the SVP protein modified by one or two SUMO1 molecules. SUMOylation by SUMO3 produced a doublet, that most likely corresponds to a mono-SUMOylated SVP.

SUMOylation occurs on lysine residues that most of the time reside within a conserved motif known as the SUMO-attachment site (Rodriguez, Dargemont, and Hay 2001), such motifs can be determined by bioinformatic analysis (Xue *et al*. 2006). However, SUMOylation can also take place on lysines that are not part of a canonical SUMO-attachment sites; therefore each one of the predicted sites needs to be confirmed experimentally (Matic *et al*. 2010). In this study I combined computational and experimental approaches to determine SUMO-attachment lysines in SVP. Each one of the predicted SUMO-attachment lysines was systematically mutated and analyzed. The results showed that lysines 30, 31, 53, 78, 156 are target for SUMOylation by SUMO1 (Figure 4.9) whereas lysines 30, 78 are target for SUMOylation by SUMO3 in *E.coli* (Figure 4.11B and 4.11D). Interestingly, lysines 30 and 78 are common targets for SUMO1 and SUMO3, providing additional support for these as SUMOylation sites of SVP.

Although the pattern of SVP SUMOylation by SUMO1 is consistent with only mono- and di- SUMOylation, five lysine residues had to be mutated to abolish SUMOylation in the *E.coli* system. A possible explanation is that SUMO1 uses alternative sites for SUMOylation once a major SUMOylation site is mutated. This scenario has been

proposed for other SUMO substrates (Okada et al. 2009). In the case of SUMOylation by SUMO3, where lysines 30 and 78 are involved (Figure 4.6 and Figure 4.11D), two modified forms of SVP (a doublet) are also observed, but in this case the molecular weight difference between them is too small to assume di-SUMOylation. A similar pattern was observed for the SUMOylation of FLC by SUMO3 (N Elrouby, personal communication), and is likely due to an aberrant migration pattern of SUMO3-modified SVP or FLC.

Amino-acid motifs that serve as targets for post-translational modifications most likely occur in exposed surfaces of the protein (or those forced to be exposed through interaction with other proteins) and usually consist of a few consecutive amino acids forming linear motifs (LMs) (Diella et al. 2008). These LM usually occur as highly disordered regions in folded proteins (Diella et al. 2009). Recent analysis shows that SUMOylation, like other PTM, frequently occurs in LM (Diella et al. 2009). The bioinformatical analysis revealed that K78 is the only SUMO-attachment site in SVP that seems to be located in a LM that is highly exposed and unstructured (Figure 4.4 and 4.5, Table 4.2). Interestingly, K78 exhibits the highest probability score for SUMOylation as well (Table 4.1), and as mentioned above, is a common target for SUMO1 and SUMO3. Taken together, these data support K78 as the most likely SVP SUMO-attachment site.

Phenotypes of plants transformed with potentially “hypo” and “hyper” SUMOylated versions of SVP

In general, SUMO facilitates transcriptional repression by transcription factors by either inhibiting the activation or stimulating the repression properties of these proteins (Gill 2005). The effects of SUMO on the function of a transcription factor are difficult to assess mainly because SUMO is readily deconjugated by the action of SUMO-proteases and normally only a small fraction of the total pool of a protein is SUMOylated at a given point of time (Johnson 2004). To overcome these difficulties two strategies were followed, first to create translational fusions aiming to hyperSUMOylate SVP (translational fusions with SUMO or AtSCE) , and second to create potentially “hypoSUMOylated” versions of SVP (using SVP mutations in the

putative SUMO-attachment sites), which cannot be SUMOylated. The hypoSUMOylated SVP versions included the PSM (K30R, K31R, K53R, K78R, K156R) and DSM (K30R, K78R) versions of SVP, which abolished SUMOylation by SUMO1 and SUMO3 in the *E.coli* system (Section 4.2.2. of this Chapter).

To study the role of SUMOylation in SVP function and determine if this modification has a physiological importance the potentially “hyperSUMOylated” and “hypoSUMOylated” SVP versions were fused to the 35S promoter and transformed into an *svp*-null mutant (*svp-41*). Since MADS-box proteins are very sensitive to translational fusions (de Folter et al. 2007) both N- and C- terminal fusions were constructed. Our results indicated that the C-terminal fusion lines better complement the *svp-41* early flowering phenotype in comparison to the N-terminal fusion lines (Figure 4.15 and Figure 4.13). This result indicates that the C-terminal fusions to SVP have less disruptive effects. The reason for this is probably that the DNA interacting MADS-domain is located at the N-terminal part of the protein and hence the addition of an extra protein tag at this end may have a negative impact on SVP DNA binding and function.

SVP participates in the control of the transition from vegetative to reproductive phase, as well as in the maintenance of floral meristem identity (Hartmann et al. 2000; Gregis et al. 2008). Interestingly, both N- and C- terminal fusions produced morphological alterations in the aerial part of the transgenic plants, this might be caused by the misexpression, whether temporally or spatially, of the recombinant protein since all the constructs were expressed under the control of the constitutive Cauliflower Mosaic Virus 35S promoter (CaMV 35S). The overexpression of SVP would promote the interaction of the protein with its natural interacting proteins but in times and tissues where these interactions normally do not take place. It may also force interactions with other proteins with which SVP does not normally interact.

For clarity we will discuss these and other effects of the N- and C- terminal fusion constructs separately.

N- terminal SUMO fusion constructs

SVP “hyperSUMOylated” versions

Two aspects have to be considered in generating a fusion-tagged protein for functional analysis, first, that the fusion protein has to be properly and stably expressed, and second that the recombinant protein has to be active (de Folter et al. 2007). In general, the transgenic lines showed that the mRNA levels correlate with the levels of protein detected for most of the constructs. Interestingly, in the case of the FLAG:SVP line, besides the expected band for the fused protein, a higher molecular weight band was also detected (Figure 4.14C and 4.14D). This form of the protein could correspond to a SUMOylated SVP fraction. Preliminary analysis to determine the identity of this form using an anti-SUMO antibody failed, most likely due to its low abundance. More experiments remain to be done to determine the identity of this band. In contrast, the amount of protein of the FLAG:SCE:SVP construct detected was extremely low despite the high amount of mRNA expressed (Figure 4.14B and 4.14C). Whether this is an effect caused by increased SUMOylation of SVP (mediated by the attached SCE) remains to be tested. No protein was detected for the FLAG:SUMO3:SVP construct in agreement with its low mRNA expression (Figure 4.14 B-D). This result is intriguing since attempts to overexpress SUMO3 (using a 35S::SUMO3 construct) in wild type plants failed to establish high levels the *SUMO3* mRNA and SUMO3 protein levels were undetected (van den Burg et al. 2010). This suggests that SUMO3 protein is unstable and an active mechanism is used to downregulate its expression. Interestingly, SUMO3 is expressed in mature pollen and in developing ovaries (van den Burg et al. 2010), and maybe more stable specifically in those tissues. Finally, FLAG:SUMO1:SVP construct was highly expressed at the mRNA level while immunodetection analysis detected a protein band of around 17 kDa indicating that the FLAG:SUMO1 part was cleaved off from SVP. This 17 kDa band was recognized by an anti-SUMO1 antibody (Data not shown). This suggests that the FLAG:SUMO1:SVP line exhibited delayed flowering because the N-terminal fusion part was cleaved (by

SUMO-proteases) releasing an active SVP protein containing no N-terminal tag.

FLAG:SUMO1:SVP lines presented interesting phenotypes besides the late flowering time. These included the formation of flowers with six sepals and seven petals, the formation of hook-like siliques, purple coloration of the petals and young siliques (which disappeared later in development), and the persistence of the sepals at the base of the mature silique (in wild type plants, abscission of these organs usually occurs when the siliques are mature) (Figure 4.18).

These phenotypes are not likely due to the FLAG:SUMO1:SVP fusion protein since its life time should be very short due to the action of SUMO-proteases but maybe due to overexpression of SVP from the 35S promoter (Figure 4.14C and 4.14D). Until now four plant SUMO-proteases have been characterized, three of them localize to the nucleus (ESD4, OST1 and OST2) (Murtas et al. 2003; Xianfeng Morgan Xu et al. 2007; Conti et al. 2008; Hermkes et al. 2011) and one (ELS1) to the cytoplasm (Hermkes et al. 2011). All four proteases have high activity towards SUMO1 (Colby et al. 2006; Chosed et al. 2006). The broad cellular distribution of SUMO proteases and their high proteolytic activity towards SUMO1 suggests FLAG:SUMO1:SVP would be efficiently processed. These phenotypes are also not likely to be due to free FLAG:SUMO1 because the overexpression of AtSUMO1/2 has no obvious effects in plant development under LD (Luisa Maria Lois, Lima, and Nam-Hai Chua 2003) and the phenotypes under SD, which include a compact rosette and early flowering (van den Burg et al. 2010), are not similar to the ones exhibited by the FLAG:SUMO1:SVP. Additionally, these phenotypes are not due to impairing SUMO1 conjugation or SUMO chain formation because the wild type copy of *SUMO1* in these plants is not affected, and the N-terminal FLAG tag in recombinant SUMO1 has been shown not to impede the conjugation of SUMO1 *in vivo* (Budhiraja et al. 2009). These results altogether indicate that the flower morphology and late flowering phenotypes are likely due to increased SVP (freed from its tag by SUMO proteases) function.

Previous studies have reported some floral phenotypes caused by 35S::SVP (Liu et al. 2007), but none is similar to those observed in the FLAG:SUMO1:SVP plants. Interestingly, Arabidopsis plants that express the rice SVP homologues OsMADS22 and OsMADS47 from the CaMV 35S promoter showed floral phenotypes similar to those observed for FLAG:SUMO1:SVP plants. These phenotypes included the purple petals and the continuous adhesion (lack of abscission) of the sepals at the base of the mature siliques (Fornara, Gregis, Pelucchi, et al. 2008). These rice SVP homologues did not restore the early flowering phenotype of *svp-41* but were able to interact with the same protein partners as SVP, indicating that the observed floral phenotypes are the product of improper interactions with other proteins (out of proper time and space) (Fornara, Gregis, Pelucchi, et al. 2008). It is conceivable that a similar scenario might be happening in the FLAG:SUMO1:SVP line.

AP1 can bind to the promoter of *SVP* and regulate its expression (Liu et al. 2007; Gregis et al. 2008). The downregulation of *SVP* expression is necessary for the establishment of floral meristem identity by AP1 (Liu et al. 2007; Gregis et al. 2008). Once floral identity has been established *SVP* expression is required again to maintain this identity until organ identity is established for each whorl (Gregis et al. 2008). The formation of the sepals and petals is partially determined by class A floral identity genes which include *AP1* and *AP2*. AP1 can also specify petal formation by regulating the expression domain of class B genes (Alvarez-Buylla 2010). AP1 also interacts with SVP protein (de Folter et al. 2005; Gregis et al. 2006). The ectopic expression of SVP allows stronger and spatially broader interactions of SVP with AP1, and most likely this enhanced interaction leads to the formation of extra sepals and petals in the FLAG:SUMO1:SVP line. It has been previously shown that *SEP3* is upregulated in *svp-41* mutants and downregulated in 35S::SVP. SVP binds to the promoter of *SEP3* and represses its expression (Liu et al. 2009). Since *SEP3* expression is initiated at the end of stage 2 during the formation of the floral meristem, exactly after *SVP* expression decreases, the correct timing of the downregulation of *SVP* expression (and upregulation of *SEP3* expression) is necessary for proper organ development. This process may also be misregulated in the overexpressor lines

described here.

FLAG:SUMO3:SVP plants are small with pale green leaves. This may be a consequence of the overexpression of SUMO3 since this phenotype was not seen in FLAG:SVP or FLAG:SUMO1:SVP lines. However, in the transgenic plants described here no FLAG:SUMO3:SVP could be detected by western blotting, suggesting that the protein is not very strongly expressed (Figure 4.14). SUMO3 is normally expressed at very low levels, and its expression pattern is restricted to the hydathodes, a pore normally found at the tip of a leaf serration, and leaf vasculature of mature leaves, as well as in the mature pollen and developing ovaries (van den Burg et al. 2010). Previous experiments have shown that the overexpression of SUMO3 has toxic effects on the plant (Budhiraja 2005) or, when it is weakly overexpressed causes an early flowering phenotype (van den Burg et al. 2010). As mentioned above, we do not see these phenotypes in lines expressing FLAG:SVP or FLAG:SUMO1:SVP (where FLAG:SUMO1 is cleaved off the fusion) ruling out that they may be caused by free SVP. In addition, SUMO-protease activity towards SUMO3 is low (Colby et al. 2006; Chosed et al. 2006; Budhiraja et al. 2009; Castaño-Miquel, Seguí, and L Maria Lois 2011) suggesting that in the FLAG:SUMO3:SVP lines, the fusion may be maintained. Although this may suggest that the phenotypes observed are caused by the FLAG:SUMO3:SVP fusion, at this point we can not rule out if they are due to free SUMO3 (FLAG:SUMO3) as well.

Lines expressing FLAG:SCE:SVP showed phenotypes similar to those of FLAG:SUMO3:SVP (Figure 4.17B) suggesting that they may be caused by a common mechanism; SUMOylation of SVP by SUMO3. However, more experiments need to be done to test this since the FLAG:SCE:SVP protein was only once very faintly detected by western blotting (Figure 4.14). The phenotypes observed must be correlated with the presence of the protein, and whether the FLAG:SCE:SVP protein is modified by SUMO3 should be tested. Previous reports have shown that the overexpression of AtSCE:(His)₆ or the AtSCE(C94S):(His)₆ causes no detectable differences in the levels of SUMO-conjugates in comparison to wild type (Luisa

Maria Lois, Lima, and Nam-Hai Chua 2003). While a previous thesis states that plants overexpressing the 35S::SCE do not exhibit any phenotype, the 35S::SCE(C94S) plants showed a stunted phenotype and an early flowering phenotype under LD and SD conditions (Budhiraja 2005). Only alterations in the phenotypes of FLAG:SCE:SVP plants were observed while all the FLAG:SCE(C94S):SVP lines looked very similar to the *svp-41* control background. These differences support the idea that the phenotype observed in the FLAG:SCE:SVP plants is more likely due to enhanced SUMOylation of SVP.

Constructs expressing SVP versions with mutations in SUMO-attachment lysines

I generated SVP versions that are mutated in the lysines that are SUMOylated in the *E.coli* system. The purpose of this experiment was to generate “hypoSUMOylated” SVP forms. These constructs were transformed into the *svp-41* mutant to test for the biological activity of these modified forms of SVP. Ideally, if these forms of SVP were not functional (due to a requirement of the lysines at these positions for SUMOylation or otherwise), they were expected not to rescue the early flowering phenotype of *svp-41*. Although these mutant forms of SVP did not rescue the early flowering phenotype of *svp-41*, it is difficult to conclude that they are required for proper SVP function in flowering time control since the wild type copy of SVP did not delay flowering either, likely due to the presence of the FLAG tag as mentioned earlier. However, unlike the wild type version of SVP, these mutant forms caused developmental abnormalities. These included interesting and unique phenotypes for particular constructs indicating that certain amino acids are key for the function of SVP in flower development (Section 4.3.3). FLAG:SVP^{K30,31,53R} plants produced flowers that contained secondary flowers emerging from the second whorl (Figure 4.25). FLAG:SVP^{K30,31,78R} lines exhibited a very dramatic phenotype where inflorescence shoots are produced instead of flowers (Figure 4.26). The FLAG:SVP^{K53,78,156R} and FLAG:SVP^{K78,156R} lines produced flowers with alterations in the first three whorls (Figure 4.27 and 4.28). Also, FLAG:SVP PSM plants senesced faster than the *svp-41*, 35S::FLAG:SVP and the wild type (Figure 4.24). These phenotypes may be due to dominant negative effects caused by sequestering interaction partners in complexes that consequently become unfunctional. This

suggests that the amino acids involved are crucial for proper SVP function.

Interestingly, all the lines that showed novel developmental phenotypes included mutations at K53 or K78, pointing out that these lysines are key determinants of SVP function. In addition, the most dramatic phenotypes involved also K30 and K31, which seem to have an additive effect that is dependent on the K53 or K78 mutation. This hypothesis is based on the fact that in lines where only K30 and K31 were mutated, no phenotype was observed at all, indicating that these lysines by themselves can not cause the above mentioned phenotypes but somehow enhance them. Although I have not yet shown the importance of these residues in the SUMOylation of SVP *in vivo*, the computational analyses we performed suggest that they are functionally important. For example, K78 has the highest SUMOplot score among the seven predicted motifs and is exposed in a disordered region, while K53 has a low score but is both found in a coil region (Table 4.2) and is one of the most conserved motifs among the 100 SVP homologues analyzed (Figure 4.3). This suggests that both K78 and K53 are important for SVP function.

SVP regulates different developmental transitions in a dose-dependent manner (Hartmann et al. 2000). These developmental transitions include the floral transition and the maintenance of the identity of the floral meristem before the establishment of the identity of the floral whorls (Hartmann et al. 2000; Gregis et al. 2006; Liu et al. 2007; Gregis et al. 2008). In general, SVP functions by repressing these developmental switches until all proteins and signals required for the transition are in place. In addition to SVP, two other proteins, SOC1 and AGL24, are also required during flower development to repress *SEP3* expression and hence prevent ectopic floral organ formation (Liu et al. 2009). It is conceivable that in the *svp-41* mutant, either SOC1 or AGL24 compensate for the absence of SVP in complexes required to regulate *SEP3* expression and hence maintain floral organ identity (Liu et al. 2009). However, in plants expressing FLAG:SVP^{K30,31,53R}, this mutant form of SVP naturally participates in these complexes but renders them non-functional. In the performed experiments, plants expressing 35S::FLAG:SVP^{K30,31,53R} produce flowers where a secondary flower emerges from the second whorl of the primary flower (Figure 4.25).

Such a phenotype has been observed in the *apl-12*, *apl ag* double and *sep1 sep2 sep3* triple mutants where SVP is ectopically expressed (Gregis et al. 2008). This suggests that the floral organ abnormality we observe in FLAG:SVP^{K30,31,53R} lines may be due to dominant negative effects of SVP^{K30,31,53R}.

General discussion of the expression of SVP from the CaMV 35S promoter

The use of a constitutive promoter like CaMV 35S to express transcription factors has several drawbacks such as unexpected pleiotropic effects (de Folter et al. 2007) and unequal tissue and developmental expression patterns (Benfey, Ren, and N H Chua 1990; Sunilkumar et al. 2002). Additionally, overexpression of MADS-box proteins from the constitutive 35S CaMV promoter can be difficult. This has been shown to be due to co-suppression and to dominant negative effects (de Folter et al. 2007). On the other hand, genes encoding MADS-box proteins contain important regulatory sequences within their intron regions, which are important for correct spatiotemporal expression (Busch et al. 1999; Sheldon et al. 2002; Hong et al. 2003). Therefore, the ideal scenario to work with MADS-box genes is to generate genomic constructs expressed under the control of their endogenous promoters (de Folter et al. 2007). However, this strategy is impractical when testing many constructs such as the study presented here. The creation of many point mutations in large genomic constructs is time-consuming and an initial exploratory analysis is required. Also, whereas it genomic constructs under the native promoter will recapitulate in a better way the endogenous condition, they will also present other levels of regulation that might complicate the interpretation. For example, *FLC* expression is partially controlled by the sense-non-coding RNA *COLD ASSISTED INTRONIC NONCODING RNA (COLDAIR)*, which is a small RNA encoded in the first intron of *FLC* and interacts with the repressive complex Polycomb Repressive Complex 2 (PRC2) maintaining *FLC* repression established after vernalization (Heo and Sung 2011). I am interested in the protein behavior and potential effects of SUMOylation on SVP protein function and extra levels of transcriptional and post-transcriptional regulation would be drawbacks for the interpretation of this PTM on SVP function.

5. General Conclusions and Perspectives

5.1. Isolation and characterization of *suppressors of esd4 (sed)*

In the last few years the importance of SUMOylation in the regulation of different cellular mechanisms in eukaryotes has been clearly demonstrated. In Arabidopsis, imbalances of the SUMO-conjugate levels have a strong impact on plant development. *ESD4* encodes a SUMO-specific protease that participates in the regulation of the levels of SUMO conjugates in Arabidopsis. The *esd4-1* mutant plants present several developmental alterations as well as an extreme early flowering phenotype under both SD and LD conditions. Despite the importance of ESD4 in plant development and the regulation of floral transition little is known about the molecular mechanisms by which ESD4 acts or which are its protein targets.

In this work, a suppressor screen of *esd4* focused on lines that reverse the early flowering phenotype was performed to investigate the role of SUMOylation in flowering time control. In this screen 120 *suppressors of esd4 (sed)* were isolated, these suppressors showed a wide range of phenotypes indicating that they are independent mutations that alleviate to varying degrees the *esd4* mutant phenotype. Interestingly no single suppressor restored the *esd4* parental phenotype to wild type, in agreement with ESD4 having multiple targets and the pleiotropic *esd4* mutant phenotype having a complex basis that cannot be restored to wild type. More suppressors restored the early flowering phenotype of *esd4* under SD conditions than under LD conditions suggesting that in some cases the mechanisms that can bypass the *esd4* effects may be dependent on day length. The fifteen characterized *sed* mutants restored different aspects of the *esd4* mutant phenotype including stature, leaf size and flowering time. These phenotypic traits were always intermediate between *esd4* and the Wt. The *sed* suppressor phenotypes are not caused by major changes in the levels of SUMO-conjugates because *sed*s SUMO-conjugate levels are more similar to those of *esd4-1* than to the wild type, suggesting they might affect single substrates or affect other pathways that can partially bypass the effects of *esd4*.

Rough map positions localized five of these *sed* mutations to distinct genomic regions suggesting that the genetic basis of these suppressors is different. Next Generation

Sequencing technology was used to fine map one of these *sed* mutations, *sed111-1*. Only six non-synonymous mutations in candidate genes were found in the genetic interval containing the *sed111-1*. The *sed111-1* candidate genes are involved in process like ubiquitination, signaling cascades and salicylic acid synthesis. All these processes have been related to SUMOylation but very few of the molecular components involved have been identified. T-DNA mutant lines for four of these potential candidates were crossed to *esd4* to test whether any of them recapitulate the *sed111-1* suppressor phenotype, the double mutants are currently being analyzed. The comparison of the expression of all this candidate genes in the backgrounds of *esd4* and *sed111-1* is also in progress. This analysis demonstrates that Next Generation sequencing is a powerful tool to identify mutations that are recalcitrant to traditional map-based cloning approaches.

The cloning and functional characterization of the suppressors of *esd4* isolated in this work will help to better understand the molecular mechanisms by which ESD4 acts and may also help to identify regulators of the levels of SUMO-conjugates or biologically relevant substrates of ESD4.

5.2 Analysis of a potential role of SUMOylation in the activity of the transcriptional and floral repressor protein SHORT VEGETATIVE PHASE

In Arabidopsis two major floral repressors have been identified, FLOWERING LOCUS C (FLC) and SHORT VEGETATIVE PHASE (SVP). SVP interacts with FLC to form a floral repressor complex that binds to the promoter regions of *FLOWERING LOCUS T (FT)* and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS (SOC1)*, two strong floral promoters, inhibiting their expression. SVP SUMOylation may enhance or inhibit the interaction with other proteins like FLC. The molecular characterization of the effects of SUMO on SVP will help to understand how these floral repressors are acting to regulate the transition to flowering in Arabidopsis.

In this work I analyzed whether SVP is SUMOylated in an *E.coli* SUMOylation system. To study the role of SUMOylation in SVP function, an *svp*-null mutant (*svp-41*) was transformed with constructs aiming to hyperSUMOylate (translational

fusions with SUMO or AtSCE) or hypoSUMOylate (mutations in the putative SUMO-attachment sites) SVP protein.

SVP is a substrate for SUMOylation in an *E.coli* SUMOylation system and SVP SUMOylation patterns vary depending on the SUMO isoform provided. The SVP SUMO1 SUMOylation pattern suggests a di-SUMO1 SUMOylation, while the SUMOylation pattern with SUMO3 indicates a mono-SUMOylation. Since SUMO1 and SUMO3 isoforms have distinct molecular properties and are likely involved in distinct biological functions, they may have a different impact on the function of SVP. The SUMO1 SUMOylation sites in SVP map to lysines K30, K31, K53, K78, K156 and to lysines K30, K78 for SUMO3. K30 and K78 are common SUMOylation sites for both SUMO1 and SUMO3 suggesting a major role for these lysines in SVP SUMOylation. Transgenic plants transformed with C-terminal translational fusions prove to be less disruptive for SVP protein function than the N-terminal fusions. Remarkably, both N- and C-terminal fusions hyperSUMOylate and hypoSUMOylate SVP versions caused morphological alterations in the aerial parts of transgenic plants in the *svp-41* background. The results obtained in this Thesis also indicate that the introduction of point mutations in the SVP sequence promotes the formation of particular mutant phenotypes.

The biochemical analysis to determine if SVP is SUMOylated *in planta* and if SVP SUMOylation is abolished by the same set of point mutations that were determined in *E.coli* SUMOylation system is in progress. Biochemical analysis of these transgenic lines will reveal whether the phenotypes observed are related to the effects of SUMO on SVP function. Transgenic plants expressing hypoSUMOylated SVP versions with the FLAG-tag attached at the C-terminus will be important as these fusions should be less disruptive of SVP protein function.

Like any other PTM, SUMOylation effects on SVP function may be transient and important only at particular stages during development. Once the major SUMO-attachment sites in SVP *in planta* are determined it will be crucial to generate constructs under the endogenous promotor. It will also be informative to compare the expression of target genes that are under the control of this transcription factor in the different constructs generated. Finally, the co-immunoprecipitation of the SVP-FLC

complex compared to the different SVP mutants will also determine if SUMOylation has an effect on the formation of this repressor complex. Demonstration of the significance of SUMO for SVP/FLC function would place SUMO at a major point of integration of different flowering pathways, and provide an important function for this protein modifier in plant development.

6. Materials and Methods

6.1. Materials

6.1.1. Chemicals and antibiotics

Chemicals were purchased from the following companies: Amersham (Münich, Germany), Boehringer Ingelheim (Ingelheim am Rhein, Germany), Carl Roth (Karlsruhe, Germany), GIBCO (Darmstadt, Germany), Invitrogen (Karlsruhe, Germany), Merck (Darmstadt, Germany), Carl Roth (Karlsruhe, Germany) and Sigma Aldrich (Deisenhofen, Germany), unless otherwise stated. Standard solutions and common buffers were prepared according to the Laboratory Manual “Molecular Cloning” (Sambrook and Russell, 2001). Modified protocols are otherwise indicated.

Nickel nitriloacetate for purification of 6xHis-tagged proteins (Ni-NTA Agarose, QIAGEN, Düsseldorf, Germany).

PVDF membranes (Millipore, Schwalbach, Germany).

Anti-Trx (Sigma-Aldrich, Deisenhofen, Germany).

Secondary antibody (goat anti-rabbit conjugated to alkaline phosphatase; Sigma-Aldrich, Deisenhofen, Germany).

NBT/BCIP (Thermo Fisher Scientific Products, Pierce Protein, Bonn, Germany).

AtSUMO1 and AtSUMO3 antibodies (Murtas et al., 2003).

Table 6.1. Antibiotics

Antibiotics	Stock Conc. (mg/ml)	Solvent	Final Concentration for selection on LB or YEB medium (mg/l)	
			<i>E. coli</i>	<i>A. tumefaciens</i>
Ampicillin (Amp)	100	H ₂ O	100	-
Carbencillin (Carb)	50	Ethanol	-	50
Chloramphenicol (Cam)	50	Ethanol	25	-
Gentamicyn (Gent)	10	H ₂ O	10	10
Hygromycin (Hyg)	50	H ₂ O	50	50
Kanamycin (Kan)	50	H ₂ O	50	25
Rifampicin (Rif)	50	DMSO	-	50
Spectinomycin (Spec)	100	H ₂ O	100	100
Tetracyclin (Tet)	10	H ₂ O	10	10

6.1.2. Enzymes

Brown Taq (Home Made Taq Polymerase – Used for Genotyping)

Taq Polymerase (Invitrogen, Karlsruhe, Germany – semi-quantitative RT-PCR)

Expand High Fidelity Taq Polymerase (Roche, Mannheim, Germany - Cloning)

Restriction Enzymes (New England Biolabs, Frankfurt am Main)

Superscript Reverse Transcriptase (Invitrogen, Karlsruhe, Germany- cDNA synthesis)

RNAse inhibitor (Invitrogen, Karlsruhe, Germany –cDNA synthesis protocol)

6.1.3. Kits

DNA-free™ kit (Ambion, Cambridgeshire, UK)
RNeasy Plant™ Mini Kit (Qiagen, Hilden, Germany)
DNeasy™ Mini Kit (Qiagen, Hilden, Germany)
Protector RNase Inhibitor (Roche, Mannheim, Germany)
Biosprint kit for DNA extraction (Qiagen, Hilden, Germany)
PCR purification kit (Qiagen, Hilden, Germany)
Nucleospin Plasmid (Macherey-Nagel, Germany)
Nucleospin Extract II (Macherey-Nagel)
BP-Clonase (Invitrogen, Heidelberg, Germany)
LR-Clonase (Invitrogen, Heidelberg, Germany)
pENTR™ Directional TOPO Cloning Kit (Invitrogen, Karlsruhe, Germany)

6.1.4. Vectors

Gateway cloning vectors

pDONR[®] 201: Gateway System BP-Reaction, Kanamycin resistance
(Invitrogen, Karlsruhe, Germany)

Gateway bacterial expression vectors

pET32b cloning and high-level expression of peptide sequences fused with the 109aa Trx•Tag™ thioredoxin protein. Gateway Compatible Bacterial Expression Vector. Ampicillin resistant.

Vectors for plant expression

pLeela	Gateway compatible with 35S-promoter, Ampicillin resistant.
p2JB_35S::FLAG:GW	p2x35S::3xFLAG::GW, Ampicillin resistant (J. Stuttmann, MPIZ)
p2JB_35S::GW:FLAG	p2x35S::GW::3xFLAG, Ampicillin resistant (J. Stuttmann, MPIZ)

6.1.5. Bacterial Strains

E. coli strains used for cloning and expression

DH5a [®]	Cloning (Invitrogen, Karlsruhe, Germany)
BL21 Star [®] (DE3)	IPTG Induction-Expression (Novagen)
DB 3.1	Gateway System (Invitrogen, Karlsruhe, Germany)

***Agrobacterium tumefaciens* strains**

GV3101/pMP90RK Rifampicin, gentamicin and kanamycin resistant

6.1.6. Oligonucleotides

Oligonucleotides were purchased from Invitrogen (Karlsruhe, Germany). The full length oligonucleotids sequences are listed in the Appendix.

6.1.7. Plant Materials

Arabidopsis thaliana, accessions Columbia 0 (*Col-0*) and Landsberg *erecta* were from the stocks in our lab. The *esd4-1* mutation was generated in Landsberg *erecta* (*L er*) by gamma radiation which caused a deletion of 752 bp that eliminates a part of the proximal promoter and the first 10 codons of the *ESD4* coding sequence (Reeves et al. 2002; Murtas et al. 2003). A Columbia allele of *ESD4*, called *esd4-2*, was obtained from the Arabidopsis Biological Resource Center (ABRC, line SALK_032317) and contains a T-DNA insertion in the fourth intron, 1537 bp from the ATG. The *svp-41* mutant (*Col-0* background) is a 2-bp deletion that causes a frame shift (Hartmann et al. 2000).

Transgenic plants overexpressing an N-terminal or C-terminal triple FLAG-tagged version of SVP under the control of the cauliflower mosaic virus 35S promoter were all developed in an *svp-41* background.

6.2. Methods

6.2.1. Plant Work

Plant growth conditions and measurement of flowering time

Seeds were sowed directly on soil and stratified in darkness at 4°C for 2-3 days. Flowering time was measured as the total number of leaves (rosette and cauline leaves) under two different photoperiod conditions: Short Day condition (SD) consisted of 8 h light and 16 h dark and Long Day condition (LD) consisted of 16 h light and 8 h dark. Both conditions with a white illumination of 120 $\mu\text{mol m}^{-2} \text{sec}^{-1}$, 22°C and relative humidity 70% in Percival Growth chambers.

Plant flower, siliques and leaves measurements

All plant organ measurements were performed with a digital-caliper. Flower measurements were taken from flowers on recent anthesis and compared. Siliques were measured at six-day old post anthesis. Rosette and cauline leaves from plants grown under LD and SD conditions were measured after the plants bolted and presented the first flowers, the longest rosette and cauline leaves per individual was considered for the comparison.

Pollen viability assay

For the pollen viability test the total number of pollen grains provenient from four flowers per genotype was analyzed. Pollen from recently dehiscent anthers was extracted, distributed on a microscope slide and mounted in a water-based solution of 1% MTT (2,5-diphenyl tetrazolium bromide or thiazolyl blue) and 5% sucrose. The solution was always freshly maked to avoid MTT degradation due to its light sensitivity. The microscope slides containing the pollen in the MTT solution were incubated for 30 minutes at 37°C. The number of viable pollen grains, deep pink colored or with irregular black lines over the surface, were counted under the light microscope.

Plant crosses and generation of Mapping Populations

Floral buds from the receptor plants (mothers) were chosen and emasculate with help of fine grasp forceps. Pollination was performed by rubbing the stigma with a bunch of stamens from the desired donor plant (father). Finally all the non-crossed buds and flowers were removed from the inflorescence and the stem labelled with the corresponding parental cross information. Seeds were collected once the silique dried.

We normally worked with floral buds around developmental stages 11-12 (Smyth et al, 1990) because in these stages, when petals are not clearly visible, the gynoecium stigmatic papillae are already receptive for pollination, a requisit for a succesful cross. For the *sed* mutants crosses the overnight step together with a multi-pollen rubber was crucial for a succesful cross, this because often the stigma remained underdeveloped in the sed mutants and a high silique abortion rates in those mutants.

To generate mapping populations, *sed* mutant plants were crossed with *esd4-2* heterozygous plants to avoid homozygous plants reduced fertility. *sed* mutants were also backcrossed with *L er* to isolate the *sed* mutation from the *esd4-1* background.

***Agrobacterium tumefaciens*-mediated transformation of Arabidopsis plants**

The floral dipping transformation method (Clough and Bent, 1998) was used to perform all our transgenic plant experiments.

Transformation of *A. tumefaciens* was performed by electroporation. In general, 2 ul of plasmid miniprep from clone of interest were added to 50 ul of *A.tumefaciens* GV3101 strain glycerol and a brief electropulse Micropulser (Biorad/2.20kV) was given. Immediately after 500 ul of YEB (MgSO₄) were added and let it shake for 2 h at 28°C. 100 ul of the grown bacteria were plated on a solid YEB (MgSO₄) media plus the following antibiotics: Kanamycin, Rifampycin, Gentamycin and Carbincillin. The plates were incubated for two days at 28°C or until colonies were visible. Colonies were tested for the prescence of the insert by PCR with the corresponding primers. A single positive colony, per construct, was stroke in a new plate plus antibiotics and let it grow for two more days. Afterwards, a touch of the grown colony was inoculate in a tube containing 2 ml liquid YEB (MgSO₄ + antibiotics) and grown overnight or until a dense bacterial concentration was observed. Finally, each inoculum was poured into 200 ml YEB (MgSO₄ + antibiotics) and let grow until a OD 0.8-1.0 at 600 nm was reached (measured with Eppendorf BioPhotometer6). Cultures were centrifugated (Beckman Coulter –Avanti J-25-

Rotor JLA10,500) 40 minutes at 10,500 rpm / 4°C and the pellets resuspended in a 5% Sucrose – 0.03% Silwet solution. Each one of the plants to be transformed was soaked into the resuspended pellet solution during 3 minutes. To conclude each one of the soaked plants were covered with a plastic bag and let so overnight before to take them to the greenhouse, were plastic bags were removed and plants watered.

Selection of Plant transformants

Transgenic seeds provenient from the transformed plants were sown on soil until seedlings with full open cotyledons were visible, seedlings were sprayed with a BASTA solution of Glufosinat 200 mg/l (BASTA, Hoechst, Germany) to select those plants that contained the BASTA resistance gene.

EMS-mutagenesis of Arabidopsis seeds

EMS-mutagenesis was performed according to Weigel and Glazebrook protocol (2002).

6.2.2. Molecular Biology

If not indicated otherwise, the Molecular Biology methods were following standard protocols from the Laboratory manuals Sambrook and Russel (2001) and Ausubel (1994).

6.2.2.1 General molecular biological techniques

Isolation of genomic DNA for Next Generation Sequencing

The DNA for Next Generation Sequencing was extracted using the DNeasy™ Mini Kit (Qiagen, Hilden, Germany) according to the manufacturer's protocol. Four independent samples were collected for the *sedIII-1* line and four independent samples for the *esd4-1*, after DNA extraction the concentration was measured with a Nanodrop (PEQ-Lab/1000). A total of 5µg of high-molecular genomic DNA with the best O.D (260/280)>1.8, and concentration of 50ng/µl was used for sequencing and the rest kept for PCR analysis.

Isolation of genomic DNA

DNA was isolated according to Edwards *et al* (Edwards *et al*, 1991). A single *Arabidopsis* rosette leaf was harvested per plant/sample and collected in an eppendorf tube. Samples were frozen in liquid nitrogen and ground with an automatic drill with a metal pestle.

For genotyping, a semi-automated method for DNA extraction was performed using the BioSprint 96 robot (Qiagen). The samples, one rosette leave per individual plant, were collected in 96 collection tubes and frozen by placing them into liquid nitrogen. All collection tubes (boxes) were kept into -80°C freezer to avoid DNA degradation until use. Tissue disruption was performed by shaking for 30 seconds in a Tissue disruptor (Retsch MM301). 300 µl of Lysis Buffer (Qiagen) was added to each tube

and homogenized by shaking for 20 seconds. Afterwards samples were centrifuged for 15 minutes at 5,000 rpm and 200 µl from the cleared plant lysate was transferred into 96-well Plates to proceed according to manufacturer's protocol.

Plasmid DNA isolation from bacteria

Plasmid constructs from transgenic *E. coli* strains were purified with the kit Nucleospin Plasmid (Macherey-Nagel) according to the manufacturers' protocol. DNA was eluted with 50 µl sterile H₂O and stored at -20°C until use.

PCR conditions

esd4 mutant genotype confirmation

To confirm that all *sed* mutants contained the *esd4-1* mutation, two primers were used in Polymerase Chain Reaction (PCR). The two primers; NE5S_ESD4 5'-TGCAGTCTCATAGGCGTCT-3' and MV1A_ESD4 5'-CTAACGACAACGAATCCAGCATC-3', amplify a 1332-bp DNA fragment in ESD4 and 570-bp in *esd4-1* because *esd4-1* has a deletion of 752 bp (Murtas *et al*, 2003). To detect the T-DNA in *esd4-2* (SALK_032317, COL) PCR with primers LP 5'-TTC ATGGGATACAGAAGCCAG-3', RP 5'-CTTATGCAAAGTGCGGAGAAG-3' and Lbb1_BP 5'-GCGTGGACCGCTTGCTGCAACT-3' was performed. The bands correspond to a 1064-bp fragment in the absence of the insertion and a 482-bp fragment if the insertion is present. Crosses of *sed* mutants and *esd4-2* were analyzed by isolating DNA from individual F1 plants and PCR amplification in two independent reactions. The first reaction was performed using three primers; NE5S_ESD4, MV1A_ESD4 and MV2A_ESD4 5'-CTTGGTACTACTCTTCAGTC-3' which amplify a 570-bp fragment in *esd4-1* deletion mutant and 408 bp in Wt. The second PCR reaction was performed using the LP and Lbb1 primers that amplifies a 482 bp band. Each PCR cycle included incubations at 94°C for 30 s, 60°C for 30 s, and 72°C for 1.30 min.

PCR Genotyping conditions for SSLP and dCAPS

For SSLP each PCR cycle included incubations at 94°C for 15 s, 58°C for 15 s, and 72°C for 30 s, for a total of 35 cycles.

For dCAPS markers each PCR cycle included incubations at 94°C for 15 s, 58°C for 15 s, and 72°C for 30 s, for a total of 30 or 25 cycles. The PCR products were digested accordingly to the manufacturer's protocol or overnight for all the enzymes with incubation temperature of 37°C.

The PCR products from SSLP and dCAPS markers were resolved in 3 % agarose gels after electrophoresis in 100 V for 2 hours approximately.

Purification of PCR products

PCR products were separated on an agarose gel and the desired fragments were excised with a sterile scalpel, DNA was purified with the kit Nucleospin Extract II (Macherey-Nagel) according to the manufacturer's protocol.

General Sequencing

Sequencing was performed by the in-house facility Automatic DNA Isolation and Sequencing (ADIS) service unit of the MPIZ, Cologne. DNA sequences were determined on Applied Biosystems (Weiterstadt, Germany). Abi Prism 377, 3100 and 3730 sequencers using BigDye-terminator v3.1 chemistry. Premixed reagents were from Applied Biosystems.

Next Generation Sequencing

Next Generation Sequencing was performed at the Cologne Center for Genomics (CCG), University of Cologne with a Illumina Genome Analyzer IIX.

6.2.2.2 RNA isolation

Plant tissue was collected in sterile eppendorf tubes and immediately frozen in liquid Nitrogen. To avoid RNA degradation all samples were kept at -80°C until use. Before RNA extraction, the tissue was pulverized with help of a frozen mortar and pestle, always adding liquid Nitrogen to avoid melting of the sample. The pulverized tissue (100 mg) was processed using the RNeasy Plant™ Mini Kit (Qiagen) according to manufacturer instructions. To remove any possible contaminating DNA samples were treated with the DNA-free™ kit (Ambion) following manufacturers protocol.

6.2.2.3 Reverse transcription

For first strand DNA synthesis 5 ug of RNA (up to a final volume of 20 ul) were heated for 5 minutes at 65°C, centrifuged and cooled on ice. Afterwards 20 ul of a master mix (see below) were added and the samples incubated for 2 h at 42°C in a water bath. To inactivate the reaction an incubation for 5 minutes at 75°C was performed. Finally the samples were spun down, cooled on ice and 110 ul H₂O added. 2 ul of cDNA per each 20 ul PCR reaction were used.

Master Mix

Oligo dT21 (20 mM) (MV7-dT_21)	4 ul
dNTPs (20 mM) (Fermentas)	2 ul
5X RT-Buffer (Invitrogen)	8 ul
DTT (100 mM) (Invitrogen)	4 ul
Protector RNase Inhibitor (Roche)	1 ul
Superscript II Reverse Transcriptase	1 ul
(Invitrogen)	

6.2.2.4 semi-quantitative RT-PCR

For the semi-quantitative RT-PCR analysis 2 ul of cDNA per each 20 ul PCR reaction were used following the programm: 94°C for 4min, (94°C for 15 s, 57 °C for 15 s and 72°C for 45 s), PCRs of 20, 25 and 30 cycles were performed to establish the saturation cycle-levels for each gene, a final extension of 72°C 2min.

6.2.2.5 Gene Cloning and PCR-directed mutagenesis

SVP GW cloning

The different SVP constructs were constructed using the GATEWAY™ recombination system (Invitrogen). The pDONR201-SVP entry clone (Torti, Max-Planck Institute für Züchtungsforschung) was amplified with primers including the attB1 and attB2 sequences (M119-SVP_GW_F and M164-SVP_GW_R_NoStop for C-terminal fusions, or M76b-SVP_GW_R_Stop for N-terminal fusions, Appendix Oligos) with the Expand High Fidelity PCR system (Roche).

The resulting entry clone was used in a LR reaction with the destination binary vector p2x35S::FLAG:GW or p2x35S::GW:FLAG. This LR reaction generated a translational fusion between the 3xFLAG tag and the mature form of SVP at the N-or C-terminus of AtSVP and place the translational fusion downstream of the 35S promoter.

SVP PCR directed mutagenesis

To introduce point mutations into the SVP sequence we used the pDONR201-SVP plasmid as template. The procedure consist of the introduction of a single base pair change by PCR. For this purpose, specific oligonucleotides were designed in a way that the base to be changed was located in the middle of the primer sequence, both in a Forward and Reverse primers. The first PCR round consisted of two independent PCR reactions (A and B), each one amplifying towards opposite ends of the SVP sequence but having the same base change in a complementary manner (Table 6.2). A third independent PCR consist of the amplification of the complete SVP sequence but using a mix of the PCR products A and B as templates. This final PCR reaction produces a complete SVP fragment with the desired single base pair change.

We used 0.5 ul of DNA plasmid and 1 ul per PCR product accordingly.

DNA plasmid (each)	0.5 ul	DNA PCR	1.0 ul
Buffer	5.0 ul	Buffer	5.0 ul
dNTPs	1.0 ul	dNTPs	1.0 ul
Primer F	1.0 ul	Primer F	1.0 ul
Primer R	1.0 ul	Primer R	1.0 ul
Expand HiFi	0.5 ul	Expand HiFi	0.5 ul
ddH2O	41.0 ul	ddH2O	41.0 ul
Final	50.0 ul	Final	50.0 ul

PCR Conditions

94°C for 4 min; (94°C for 30 s, 58°C for 30 s, 72°C for 1 min) 15 cycles ; 72°C for 2 min (Plasmid).

94°C for 4 min; (94°C for 30 s, 56°C for 30 s, 72°C for 1 min) 20 cycles ; 72°C for 2 min (PCR).

Heat shock transformation of *E. coli*

An aliquot of competent cells was thawed on ice and incubated with 1 µl plasmid DNA (Miniprep) or 5 µl ligation (Gateway) for 10 – 30 min on ice. Afterwards cells were incubated for 2 min at 37°C and put directly on ice. The bacteria were incubated in 0.75 ml LB medium and incubated for 1 h at 37°C and 750 rpm. An aliquot was plated on LB plates containing antibiotics for selection of transformants.

Oligos / example	Schematic	Use
M119 (F) + M164 (R)	<p>attB1-SVP-GW-F → [GW SVP GW] ← attB1-SVP-GW-R</p>	SVP clone
SUMO1-GW-S (F) + M97a-SVP-SUM1-A / for A fragment M97b-SUM1-SVP-S + M164 (R) / for B fragment SUMO1-GW-S (F) + M164 (R) for full length final fragment	<p>SUMO-GW-F → A ← SVP-SUMO-A [GW SUMO SVP GW] SUMO-SVP-S → B ← SVP-GW-R-Stop</p>	To generate N-terminal translational fusions with SUMO or the SCE
SVP-GW-F (F) + M156-SUMO1-SVP-A / for A fragment M155-SVP_SUMO1-S + M157-SUMO1-CTERM GW for B fragment SVP-GW-F (F) + M157- SUMO1-CTERM GW for full length final fragment	<p>SVP-GW-F → A ← SUMO-SVP-A [GW SVP SUMO GW] SVP-SUMO-S → B ← SUMO-GW-R</p>	To generate C-terminal translational fusions with SUMO or the SCE
M119 (F) + M89b-SVP- K10R-R for A fragment M164 (R) + M89a-SVP-K10R for B fragment M119 (F) + M164 (R) using A and B as templates for full length final fragment	<p>attB1-SVP-GW-F → A ← SVP-K10R-R [GW SVP GW] SVP-K10R-F → B ← attB2-SVP-GW-R</p>	To generate point mutations in the SUMO-attachment Lysines

6.2.3. Biochemical methods

6.2.3.1 Protein Extraction and purification

E. coli recombinant proteins extraction

Ten milliliters of *E. coli* cultures were grown for 2h at 37°C and induced with 1 mM IPTG. After induction, cultures were incubated overnight at 28°C with a final 2 h incubation at 25°C. Afterwards cultures were centrifuged for 10 min at 4,000 rpm to recover cell pellets. Harvested cell pellets were thaw on ice for 15 min and resuspended in buffer B (100 mM NaH₂PO₄, 10 mM Tris.Cl, 6 M GuHCl, pH 8.0).

To improve cell lysis the resuspended cultures were shaken end-over-end for 1 h and sonicated twice at an interval of 0.5/0.5 for 15 min. with a Bioruptor™ Diagenode. Lysates were centrifuged at 10,000 x g for 30 min at room temperature and the supernatant recovered. 100 ul were kept for total protein extraction control. Recombinant proteins were purified on 50% Nickel Nitriloacetate resin in 8 M Urea (Ni-NTA, QIAGEN) following the Batch purification of 6xHis-tagged proteins under denaturing conditions of The QIAexpressionist™ handbook of QIAGEN.

***A.thaliana* recombinant proteins extraction**

Proteins were extracted from 10-15 days old seedlings following the protocol reported by Murtas et al. (Murtas et al., 2003).

Protein quantification

Protein quantification was performed using a standard Bradford method.

6.2.3.2 SDS-PAGE and Western blot analysis

Proteins were denatured by heat (10 minutes, 98°C) in SDS-protein loading buffer. Afterwards, proteins were separated on 10% SDS-PAGE gels. Samples were runned at 80 Volts for around 2-3 hours and afterwards blotted onto PVDF membranes (Millipore, Bedford, USA) using standard methods. Blots were blocked overnight in PBS with 5% nonfat milk, next day membranes were rinsed briefly with PBS with 5% nonfat milk and afterwards a PBS with 5% nonfat milk plus anti-Trx (Sigma-Aldrich) or anti-SUMO antisera solution buffer were added. After 2 h, the blots were washed twice in PBS-milk, and a secondary antibody (goat anti-rabbit conjugated to alkaline phosphatase; Sigma-Aldrich) was added. After washing, the blots were treated with BCIP and NBT to visualize the signals.

6.2.4. Software and websites

All bioinformatic analysis were performed with the default presets following the instructions from each correspondent web-site.

dCAPS design	http://helix.wustl.edu/dcaps/dcaps.html
Arabidopsis information resource	http://www.arabidopsis.org/
Bioinformatics tool	http://www.ncbi.nlm.nih.gov/
DNA sequences alignments	http://bioinfo.genopole-toulouse.prd.fr/multalin/
SUMOPlot™	http://www.abgent.com/tool/sumoplot
Conserved domains	http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi
BLAST	http://www.ncbi.nlm.nih.gov/BLAST
Alignment tool COBALT	http://www.ncbi.nlm.nih.gov/tools/cobalt
Motif discover	http://meme.sdsc.edu/meme
Secondary Structure Prediction	http://sable.cchmc.org
Secondary Structure visualization	http://polyview.cchmc.org
Disordered region analysis tools:	
DisEMBL™	http://dis.embl.de/
SPRITZ	http://protein.cribi.unipd.it/spritz

7. References

- Ahn, J.H., Miller, D., Winter, V.J., Banfield, M.J., Lee, J.H., Yoo, S.Y., Henz, S.R., Brady, R.L. and Weigel, D. 2006. A divergent external loop confers antagonistic activity on floral regulators FT and TFL1. *The EMBO Journal*, 8; 25(3): 605-14.
- Alexandre, C.M., and Lars H. 2008. FLC or not FLC: the other side of vernalization. *Journal of Experimental Botany*, 59 (6): 1127-1135.
- Alvarez-Buylla, E.R., Benítez, M., Corvera-Poiré, A., Cador, A.C., 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 January 2010*: Volume 8, e0127.
- Amasino, R.M. and Michaels, S.D. 2010. The timing of flowering. *Plant Physiology*, 154 (2): 516-20.
- Ausubel F. M. 1994. Current protocols in molecular biology. New York, John Wiley & Sons, Inc.
- Ayaydin, F. and Dasso, M., 2004. Distinct in vivo dynamics of vertebrate SUMO paralogues. *Molecular Biology of the Cell*, 15(12): 5208-5218.
- Bartrina, I., Otto, E., Strnad, M., Werner, T., and Schmülling, T., 2011. Cytokinin regulates the activity of reproductive meristems, flower organ size, ovule formation, and thus seed yield in *Arabidopsis thaliana*. *The Plant Cell*, 23 (1): 69-80.
- Bates, P.W., and Vierstra, R.D., 1999. UPL1 and 2, two 405 kDa ubiquitin-protein ligases from *Arabidopsis thaliana* related to the HECT-domain protein family. *The Plant Journal: For Cell and Molecular Biology*, 20 (2): 183-195.
- Bawa-Khalfe, T. and Yeh, E.T.H., 2010. SUMO Losing Balance: SUMO Proteases Disrupt SUMO Homeostasis to Facilitate Cancer Development and Progression. *Genes & Cancer*, 1 (7): 748-752.
- Bayer, P., Arndt, A., Metzger, S., Mahajan, R., Melchior, F., Jaenicke, R. and Becker, J., 1998. Structure determination of the small ubiquitin-related modifier SUMO-1. *Journal of Molecular Biology*, 280 (2): 275-286.
- Benfey, P.N., Ren, L. and Chua, N.H., 1990. Tissue-specific expression from CaMV 35S enhancer subdomains in early stages of plant development. *The EMBO Journal*, 9 (6): 1677-1684.
- Bernier-Villamor, V., Sampson, D.A., Matunis, M.J., and Lima, C.D., 2002. Structural basis for E2-mediated SUMO conjugation revealed by a complex between ubiquitin-conjugating enzyme Ubc9 and RanGAP1. *Cell*, 108 (3) : 345-356.

- Blum, A. 2011. *Plant Breeding for Water-Limited Environments*. New York, NY: Springer New York, 2011, XIV, 258p.
- Boss, P.K., Bastow, R.M., Mylne, J.S. and Dean, C. 2004. Multiple pathways in the decision to flower: enabling, promoting, and resetting. *The Plant Cell*, 16 Suppl:S18-31
- Budhiraja, R. 2005. Post-translational modification of proteins by SUMO in *Arabidopsis thaliana*. PhD Thesis, University of Cologne, Germany. 128 p.
- Budhiraja, R., Hermkes, R., Müller, S., Schmidt, J., Colby, T., Panigrahi, K., Coupland, G. and Bachmair, A., 2009. Substrates related to chromatin and to RNA-dependent processes are modified by *Arabidopsis* SUMO isoforms that differ in a conserved residue with influence on desumoylation. *Plant Physiology*, 149 (3): 1529-1540.
- Busch, M.A., Bomblies, K. and Weigel, D., 1999. Activation of a Floral Homeotic Gene in *Arabidopsis*. *Science*, 285 (5427) : 585 -587.
- Capili, A.D. and Lima, C.D., 2007. Structure and analysis of a complex between SUMO and Ubc9 illustrates features of a conserved E2-Ubl interaction. *Journal of Molecular Biology*, 369 (3): 608-618.
- Carrera, J., Rodrigo, G., Jaramillo A. and Elena S.F., 2009. Reverse-engineering the *Arabidopsis thaliana* transcriptional network under changing environmental conditions. *Genome Biology*, 10(9): p.R96.
- Castaño-Miquel, L., Seguí J., and Lois, L.M., 2011. Distinctive properties of *Arabidopsis* SUMO paralogues support the in vivo predominant role of AtSUMO1/2 isoforms. *The Biochemical Journal*, 436 (3): 581-590.
- Cheema, A., Knights, C.D., Rao, M., Catania, J., Perez, R., Simons, B., Dakshanamurthy, S., Kolukula, V.K., Tilli, M., Furth, P.A., Albanese, C. and Avantaggiati, M.L., 2010. Functional mimicry of the acetylated C-terminal tail of p53 by a SUMO-1 acetylated domain, SAD. *Journal of Cellular Physiology*, 225 (2): 371-384.
- Cheong, M.S., Park, H.C., Bohnert, H.J., Bressan, R.A. and Yun, D.J., 2010. Structural and functional studies of SIZ1, a PIAS-type SUMO E3 ligase from *Arabidopsis*. *Plant Signaling & Behavior*, 5 (5): 567-569.
- Chosed, R., Mukherjee, S., Lois, L.M. and Orth, K., 2006. Evolution of a signalling system that incorporates both redundancy and diversity: *Arabidopsis* SUMOylation. *The Biochemical Journal*, 398 (3): 521-529.
- Clough S.J. and Bent A.F. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal: For Cell and Molecular Biology*, 16(6): 735-43.

- Colby, T., Matthäi, A., Boeckelmann, A. and Stuible, H.P., 2006. SUMO-conjugating and SUMO-deconjugating enzymes from Arabidopsis. *Plant Physiology*, 142 (1) : 318-332.
- Conti, L., Price, G., O'Donnell, E., Schwessinger, B., Dominy, P. and Sadanandom, A., 2008. Small ubiquitin-like modifier proteases OVERLY TOLERANT TO SALT1 and -2 regulate salt stress responses in Arabidopsis. *The Plant Cell*, 20 (10) : 2894-2908.
- Denuc, A. and Marfany, G., 2010. SUMO and ubiquitin paths converge. *Biochemical Society Transactions*, 38: 34-39.
- Desterro, J.M., Rodriguez, M.S., Kemp, G.D. and Hay, R.T., 1999. Identification of the enzyme required for activation of the small ubiquitin-like protein SUMO-1. *The Journal of Biological Chemistry*, 274 (15): 10618-10624.
- Diella, F., Haslam, N., Chica, C., Budd, A., Michael, S., Brown, N. P., Trave, G. and Gibson, T.J. 2008. Understanding eukaryotic linear motifs and their role in cell signaling and regulation. *Frontiers in Bioscience*, 13: 6580-6603.
- de Folter, S., Immink, R.G.H., Kieffer, M., Parenicová, L., Henz, S.R., Weigel, D., Busscher, M., Kooiker, M., Colombo, L., Kater, M.M., Davies, B. and Angenent, G.C., 2005. Comprehensive interaction map of the Arabidopsis MADS Box transcription factors. *The Plant Cell*, 17 (5): 1424-1433.
- de Folter, S., Urbanus, S.L., van Zuijlen, L.G.C., Kaufmann, K., and Angenent, G.C., 2007. Tagging of MADS domain proteins for chromatin immunoprecipitation. *BMC Plant Biology*, 7 (47): 1471-2229.
- Dohmen, R.J., Stappen, R., McGrath, J.P., Forrová, H., Kolarov, J., Goffeau, A. and Varshavsky, A., 1995. An essential yeast gene encoding a homolog of ubiquitin-activating enzyme. *The Journal of Biological Chemistry*, 270 (30): 18099-18109.
- Drag, M. and Salvesen, G.S., 2008. DeSUMOylating enzymes-SENPs. *IUBMB Life*, 60 (11): 734-742.
- Downes, B. P., Stupar, R.M., Gingerich, D.J. and Vierstra, R.D., 2003. The HECT ubiquitin-protein ligase (UPL) family in Arabidopsis: UPL3 has a specific role in trichome development. *The Plant Journal: For Cell and Molecular Biology* 35 (6) : 729-742.
- Duda, D.M., van Waardenburg, R.C., Borg, L.A., McGarity, S., Nourse, A., Waddell, M.B., Bjornsti, M.A. and Schulman, B.A., 2007. Structure of a SUMO-binding-motif mimic bound to Smt3p-Ubc9p: conservation of a non-covalent ubiquitin-like protein-E2 complex as a platform for selective interactions within a SUMO pathway. *Journal of Molecular Biology*, 369 (3): 619-630.

- Edwards, K., Johnstone, C. and Thompson, C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acid Research*, 25;19 (6): 1349.
- Elrouby, N. and Coupland, G., 2010. Proteome-wide screens for small ubiquitin-like modifier (SUMO) substrates identify Arabidopsis proteins implicated in diverse biological processes. *Proceedings of the National Academy of Sciences of the United States of America*, 107 (40): 17415-17420.
- Fornara, F., Gregis, V., Pelucchi, N., Colombo, L. and Kater, M., 2008. The rice StMADS11-like genes OsMADS22 and OsMADS47 cause floral reversions in Arabidopsis without complementing the *svp* and *agl24* mutants. *Journal of Experimental Botany*, 59 (8): 2181-2190.
- Fornara, F., de Montaigu, A. and Coupland, G. 2010. SnapShot: Control of flowering in Arabidopsis. *Cell*, 30; 141 (3): 550.e1-2.
- Garcia-Dominguez, M. and Reyes, J.C., 2009. SUMO association with repressor complexes, emerging routes for transcriptional control. *Biochimica Et Biophysica Acta*, 1789 (6-8): 451-459.
- Gareau, J.R. and Lima, C. D., 2010. The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nature Reviews. Molecular Cell Biology*, 11(12): 861-871.
- Gill, G. 2005. Something about SUMO inhibits transcription. *Current Opinion in Genetics & Development*, 15 (5): 536-541.
- Glawischnig, E., Hansen, B.J., Olsen, C.E. and Halkier, B.A., 2004. Camalexin is synthesized from indole-3-acetaldoxime, a key branching point between primary and secondary metabolism in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, 101 (21): 8245-8250.
- Greene, E. A., Codomo, C.A., Taylor, N.E., Henikoff, J.G., Till, B.J., Reynolds, S.H., Enns, L.C., Enns, L.C., Burtner, C., Johnson, J.E., Odden, A.R., Comai, L., Henikoff, S., 2003. Spectrum of chemically induced mutations from a large-scale reverse-genetic screen in Arabidopsis. *Genetics*, 164 (2): 731-740.
- Gregis, V., Sessa, A., Colombo, L. and Kater, M.M., 2006. AGL24, SHORT VEGETATIVE PHASE, and APETALA1 redundantly control AGAMOUS during early stages of flower development in Arabidopsis. *The Plant Cell* 18 (6): 1373-1382.
- Gregis, V., Sessa, A., Colombo, L., and Kater, M.M., 2008. AGAMOUS-LIKE24 and SHORT VEGETATIVE PHASE determine floral meristem identity in Arabidopsis. *The Plant Journal: For Cell and Molecular Biology*, 56 (6): 891-902.

- Han, A., He, J., Wu, Y., Liu, J.O. and Chen, L., 2005. Mechanism of recruitment of class II histone deacetylases by myocyte enhancer factor-2. *Journal of Molecular Biology*, 345 (1): 91-102.
- Harper, M.A., Chen, Z., Toy, T., Machado, I.M., Nelson, S.F., Liao, J.C. and Lee, C.J., 2011. Phenotype sequencing: identifying the genes that cause a phenotype directly from pooled sequencing of independent mutants. *PLoS One*, 6 (2): e16517.
- Hartmann, U., Höhmann, S., Nettekheim, K., Wisman, E., Saedler, H. and Huijser P., 2000. Molecular cloning of SVP: a negative regulator of the floral transition in Arabidopsis. *The Plant Journal: For Cell and Molecular Biology*, 21 (4): 351-360.
- He, Y., 2009. Control of the transition to flowering by chromatin modifications. *Molecular Plant*, 2 (4): 554-564.
- He, Y., Michaels, S.D. and Amasino, R.M., 2003. Regulation of Flowering Time by Histone Acetylation in Arabidopsis. *Science*, 302(5651): 1751 -1754.
- Hemming, M.N., and Trevaskis, B., 2011. Make hay when the sun shines: the role of MADS-box genes in temperature-dependant seasonal flowering responses. *Plant Science: An International Journal of Experimental Plant Biology*, 180 (3): 447-453.
- Heo, J.B. and Sung, S., 2011. Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. *Science*, 331 (6013): 76-79.
- Hermkes, R., Yong-Fu, F., Nürrenberg, K., Budhiraja, R., Schmelzer, E., Elrouby, N., Dohmen, R.J., Bachmair, A. and Coupland, G. 2011. Distinct roles for Arabidopsis SUMO protease ESD4 and its closest homolog ELS1. *Planta* 233 (1): 63-73.
- Hershko, A. and Ciechanover, A., 1992. The ubiquitin system for protein degradation. *Annual Review of Biochemistry*, 61: 761-807.
- Hietakangas, V., Anckar, J., Blomster, H.A., Fujimoto, M., Palvimo, J.J., Nakai, A. and Sistonen, L., 2006. PDSM, a motif for phosphorylation-dependent SUMO modification. *Proceedings of the National Academy of Sciences of the United States of America*, 103(1): 45-50.
- Hong, R.L., Hamaguchi, L., Busch, M.A. and Weigel, D. 2003. Regulatory elements of the floral homeotic gene AGAMOUS identified by phylogenetic footprinting and shadowing. *The Plant Cell*, 15 (6) : 1296-1309.
- Huang, L., Yang, S., Zhang, S., Liu, M., Lai, J., Qi, Y., Shi, S., Wang, J., Wang, Y., Xie, Q., Yang, C., 2009. The Arabidopsis SUMO E3 ligase AtMMS21, a homologue of NSE2/MMS21, regulates cell proliferation in the root. *The Plant Journal: For Cell and Molecular Biology*, 60 (4): 666-678.

- Huang, W.C., Ko, T.P., Li, S.S. and Wang, A.H., 2004. Crystal structures of the human SUMO-2 protein at 1.6 Å and 1.2 Å resolution: implication on the functional differences of SUMO proteins. *European Journal of Biochemistry / FEBS*, 271 (20): 4114-4122.
- Ikeda, F. and Dikic, I., 2008. Atypical ubiquitin chains: new molecular signals. *EMBO Rep*, 9(6): 536-542.
- Ikeda, Y., Kobayashi, Y., Yamaguchi, A., Abe, M. and Araki, T., 2007. Molecular basis of late-flowering phenotype caused by dominant epi-alleles of the FWA locus in Arabidopsis. *Plant & Cell Physiology*, 48 (2): 205-220.
- Ishida, T., Fujiwara, S., Miura, K., Stacey, N., Yoshimura, M., Schneider, K., Adachi, S., Minamisawa, K., Umeda, M. and Sugimoto, K., 2009. SUMO E3 ligase HIGH PLOIDY2 regulates endocycle onset and meristem maintenance in Arabidopsis. *The Plant Cell* 21 (8): 2284-2297.
- Jadhav, T. and Wooten, M.W., 2009. Defining an Embedded Code for Protein Ubiquitination. *Journal of Proteomics & Bioinformatics*, 24(2): 316.
- Jakobs, A., Koehnke, J., Himstedt, F., Funk, M., Korn, B., Gaestel, M. and Niedenthal, R., 2007. Ubc9 fusion-directed SUMOylation (UFDS): a method to analyze function of protein SUMOylation. *Nature Methods*, 4 (3): 245-250.
- Jin, J.B. and Hasegawa, P.M., 2008. Flowering time regulation by the SUMO E3 ligase SIZ1. *Plant Signaling & Behavior*, 3 (10): 891-892.
- Jin, J.B., Jin, Y.H., Lee, J., Miura, K., Yoo, C.Y., Kim, W.Y., Van Oosten, M., Hyun Y, Somers, D.E., Lee, I., Yun, D.J., Bressan, R.A. and Hasegawa, P.M., 2008. The SUMO E3 ligase, AtSIZ1, regulates flowering by controlling a salicylic acid-mediated floral promotion pathway and through affects on FLC chromatin structure. *The Plant Journal: For Cell and Molecular Biology*, 53 (3): 530-540.
- Johnson, E.S. and Blobel, G., 1997. Ubc9p is the conjugating enzyme for the ubiquitin-like protein Smt3p. *The Journal of Biological Chemistry*, 272 (43): 26799-26802.
- Johnson, E.S. 2004. Protein modification by SUMO. *Annual Review of Biochemistry* 73: 355-382.
- Johnson, P.R. and Hochstrasser, M., 1997. SUMO-1: Ubiquitin gains weight. *Trends in Cell Biology*, 7 (10): 408-413.
- Kang, X., Qi, Y., Zuo, Y., Wang, Q., Zou, Y., Schwartz, R.J., Cheng, J., Yeh, E.T., 2010. SUMO-specific protease 2 is essential for suppression of polycomb group protein-mediated gene silencing during embryonic development. *Molecular Cell*, 38 (2): 191-201.

- 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 (2): 183-198.
- Kerscher, O. 2007. SUMO junction-what's your function? New insights through SUMO-interacting motifs. *EMBO Reports* 8 (6): 550-555.
- Khatun, S. and Flowers, T.J., 1995. The estimation of pollen viability in rice. *Journal of Experimental Botany*, 46 (1): 151 -154.
- Kim, J.H. and Baek, S.H., 2009. Emerging roles of desumoylating enzymes. *Biochimica Et Biophysica Acta*, 1792 (3): 155-162.
- Kim, Y.S., Schumaker, K.S. and Zhu, J.K., 2006. EMS mutagenesis of Arabidopsis. *Methods in Molecular Biology*, 323: 101-103.
- Knipscheer P., van Dijk, W.J., Olsen, J.V., Mann, M. and Sixma, T.K., 2007. Noncovalent interaction between Ubc9 and SUMO promotes SUMO chain formation. *The EMBO Journal*, 26 (11): 2797-2807.
- Konieczny, A. and Ausubel, F.M., 1993. A procedure for mapping Arabidopsis mutations using co-dominant ecotype-specific PCR-based markers. *The Plant Journal: For Cell and Molecular Biology*, 4 (2): 403-410.
- Koornneef, M., Alonso-Blanco, C., Blankestijn-de Vries, H., Hanhart, C.J. and Peeters, A.J., 1998. Genetic interactions among late-flowering mutants of Arabidopsis. *Genetics*, 148 (2): 885-892.
- Koornneef, M., Hanhart, C.J. and van der Veen, J.H., 1991. A genetic and physiological analysis of late flowering mutants in Arabidopsis thaliana. *Molecular & General Genetics: MGG*, 229 (1): 57-66.
- Koornneef, M., Alonso-Blanco, C. and Vreugdenhil, D., 2004. Naturally occurring genetic variation in Arabidopsis thaliana. *Annual Review of Plant Biology* 55: 141-172.
- Kurepa, J., Walker, J.M., Smalle, J., Gosink, M.M., Davis, S.J., Durham, T.L., Sung, D.Y. and Vierstra, R.D., 2003. The small ubiquitin-like modifier (SUMO) protein modification system in Arabidopsis. Accumulation of SUMO1 and -2 conjugates is increased by stress. *The Journal of Biological Chemistry*, 278 (9): 6862-6872.
- Lee, J., Miura, K., Bressan, R.A., Hasegawa, P.M. and Yun, D.J., 2007. Regulation of Plant Innate Immunity by SUMO E3 Ligase. *Plant Signaling & Behavior*, 2 (4): 253-254.
- Lee, H., Suh, S.S., Park, E., Cho, E., Ahn, J.H., Kim, S.G., Lee, J.S., Kwon, Y.M. and Lee, I. 2000. The AGAMOUS-LIKE 20 MADS domain protein integrates floral inductive pathways in Arabidopsis. *Genes and Development*, 15; 14 (18):2366-76.

- Lee, J.H., Yoo, S.J., Park, S.H., Hwang, I., Lee, J.S., Ahn, J.H., 2007. Role of SVP in the control of flowering time by ambient temperature in Arabidopsis. *Genes & Development*, 21 (4): 397-402.
- Lee, J. and Lee, I., 2010. Regulation and function of SOC1, a flowering pathway integrator. *Journal of Experimental Botany*, 61 (9): 2247-2254.
- Li, D., Liu, C., Shen, L., Wu, Y., Chen, H., Robertson, M., Helliwell, C.A., Ito, T., Meyerowitz, E. and Yu, H., 2008. A repressor complex governs the integration of flowering signals in Arabidopsis. *Developmental Cell*, 15 (1): 110-120.
- Li, H., and Durbin, R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics*, 25 (14): 1754-1760.
- Li, S.J. and Hochstrasser, M., 1999. A new protease required for cell-cycle progression in yeast. *Nature*, 398 (6724): 246-251.
- Li, S.J. and Hochstrasser, M., 2000. The yeast ULP2 (SMT4) gene encodes a novel protease specific for the ubiquitin-like Smt3 protein. *Molecular and Cellular Biology*, 20(7): 2367-2377.
- Li, S.J. and Hochstrasser, M., 2003. The Ulp1 SUMO isopeptidase: distinct domains required for viability, nuclear envelope localization, and substrate specificity. *The Journal of Cell Biology*, 160 (7): 1069-1081.
- Linding, R., Jensen, L.J., Diella, F., Bork, P., Gibson, T.J. and Russell, R.B., 2003. Protein disorder prediction: implications for structural proteomics. *Structure*, 11 (11): 1453-1459.
- Liu, C., Xi, W., Shen, L., Tan, C. and Yu, H., 2009. Regulation of floral patterning by flowering time genes. *Developmental Cell*, 16 (5): 711-722.
- Liu, C., Zhou, J., Bracha-Drori, K., Yalovsky, S., Ito, T. and Yu, H., 2007. Specification of Arabidopsis floral meristem identity by repression of flowering time genes. *Development*, 134 (10): 1901-1910.
- Liu, Q., Jin, C., Liao, X., Shen, Z., Chen, D.J. and Chen, Y., 1999. The binding interface between an E2 (UBC9) and a ubiquitin homologue (UBL1). *The Journal of Biological Chemistry*, 274(24): 16979-16987.
- Lois, L.M. 2010. Diversity of the SUMOylation machinery in plants. *Biochemical Society Transactions*, 38: 60-64.
- Lois, L.M. and Lima, C.D., 2005. Structures of the SUMO E1 provide mechanistic insights into SUMO activation and E2 recruitment to E1. *The EMBO Journal*, 24 (3): 439-451.

- Lois, L.M., Lima, C.D. and Chua, N.H., 2003. Small ubiquitin-like modifier modulates abscisic acid signaling in Arabidopsis. *The Plant Cell*, 15 (6): 1347-1359.
- Lukowitz, W., Gillmor, C.S. and Scheible, W.R., 2000. Positional cloning in Arabidopsis. Why it feels good to have a genome initiative working for you. *Plant Physiology*, 123 (3): 795-805.
- 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 (3): 484-495.
- Mahajan, R., Delphin, C., Guan, T., Gerace, L. and Melchior, F., 1997. A small ubiquitin-related polypeptide involved in targeting RanGAP1 to nuclear pore complex protein RanBP2. *Cell*, 88 (1): 97-107.
- Martínez, C., Pons, E., Prats, G. and León, J., 2004. Salicylic acid regulates flowering time and links defence responses and reproductive development. *The Plant Journal: For Cell and Molecular Biology*, 37 (2): 209-217.
- Matafora, V., D'Amato, A., Mori, S., Blasi, F. and Bachi, A., 2009. Proteomics analysis of nucleolar SUMO-1 target proteins upon proteasome inhibition. *Molecular & Cellular Proteomics: MCP*, 8 (10): 2243-2255.
- Matic, I., Macek, B., Hilger, M., Walther, T.C. and Mann, M., 2008. Phosphorylation of SUMO-1 occurs in vivo and is conserved through evolution. *Journal of Proteome Research*, 7(9): 4050-4057.
- Matic, I., Schimmel, J., Hendriks, I.A., van Santen, M.A., van de Rijke, F., van Dam H., Gnad, F., Mann, M. and Vertegaal, A.C., 2010. Site-specific identification of SUMO-2 targets in cells reveals an inverted SUMOylation motif and a hydrophobic cluster SUMOylation motif. *Molecular Cell*, 39 (4): 641-652.
- Matunis, M.J., Coutavas, E. and Blobel, G., 1996. A novel ubiquitin-like modification modulates the partitioning of the Ran-GTPase-activating protein RanGAP1 between the cytosol and the nuclear pore complex. *The Journal of Cell Biology*, 135 (6): 1457-1470.
- Michelmore, R.W., Paran, I., Kesseli, R.V., 1991. Identification of markers linked to disease-resistance genes by bulked segregant analysis: a rapid method to detect markers in specific genomic regions by using segregating populations. *Proceedings of the National Academy of Sciences of the United States of America*, 88 (21): 9828-9832.
- Michaels, S.D. and Amasino, R.M., 1999. FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *The Plant Cell*, 11 (5): 949-956.
- Michaels, S.D., Ditta, G., Gustafson-Brown, C., Pelaz, S., Yanofsky, M. and Amasino, R.M., 2003. AGL24 acts as a promoter of flowering in Arabidopsis

- and is positively regulated by vernalization. *The Plant Journal: For Cell and Molecular Biology*, 33 (5): 867-874.
- Miller, M.J., Barrett-Wilt, G.A., Hua, Z. and Vierstra RD., 2010. Proteomic analyses identify a diverse array of nuclear processes affected by small ubiquitin-like modifier conjugation in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, 107 (38): 16512-16517.
- Miteva, M., Keusekotten, K., Hofmann, K., Praefcke, G.J. and Dohmen, R.J., 2010. Sumoylation as a signal for polyubiquitylation and proteasomal degradation. *Sub-Cellular Biochemistry*, 54: 195-214.
- Mittler, R. and Blumwald, E., 2010. Genetic Engineering for Modern Agriculture: Challenges and Perspectives. *Annual Review of Plant Biology* 61 (1): 443-462.
- Miura, K. and Hasegawa, P.M., 2010. Sumoylation and other ubiquitin-like post-translational modifications in plants. *Trends in Cell Biology* 20 (4): 223-232.
- Miura, K., Jin, J.B. and Hasegawa, P.M., 2007. Sumoylation, a post-translational regulatory process in plants. *Current Opinion in Plant Biology*, 10 (5): 495-502.
- Miura, K., Jin, J.B., Lee, J., Yoo, C.Y., Stirm, V., Miura, T., Ashworth, E.N., Bressan, R.A., Yun, D.J. and Hasegawa, P.M., 2007. SIZ1-mediated sumoylation of ICE1 controls CBF3/DREB1A expression and freezing tolerance in Arabidopsis. *The Plant Cell*, 19 (4): 1403-1414.
- Miura, K., Lee, J., Jin, J.B., Yoo, C.Y., Miura, T. and Hasegawa, P.M., 2009. Sumoylation of ABI5 by the Arabidopsis SUMO E3 ligase SIZ1 negatively regulates abscisic acid signaling. *Proceedings of the National Academy of Sciences of the United States of America*, 106 (13): 5418-5423.
- Miura, K., Rus, A., Sharkhuu, A., Yokoi, S., Karthikeyan, A.S., Raghothama, K.G., Baek, D., Koo, Y.D., Jin, J.B., Bressan, R.A., Yun, D.J., Hasegawa, P.M., 2005. The Arabidopsis SUMO E3 ligase SIZ1 controls phosphate deficiency responses. *Proceedings of the National Academy of Sciences of the United States of America*, 102 (21): 7760-7765.
- Mohideen, F., Capili, A.D., Bilimoria, P.M., Yamada, T., Bonni, A. and Lima CD., 2009. A molecular basis for phosphorylation-dependent SUMO conjugation by the E2 UBC9. *Nature Structural & Molecular Biology*, 16 (9): 945-952.
- Moon, J., Suh, S.S., Lee, H., Choi, K.R., Hong, C.B., Paek, N.C., Kim, S.G. and Lee, I. 2003. The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. 2003. *The Plant Journal*. 35 (5):615-23.
- Mossessova, E. and Lima, C.D., 2000. Ulp1-SUMO crystal structure and genetic analysis reveal conserved interactions and a regulatory element essential for cell growth in yeast. *Molecular Cell*, 5(5): 865-876.

- Mouradov, A., Cremer, F. and Coupland, G., 2002. Control of flowering time: interacting pathways as a basis for diversity. *The Plant Cell*, 14 Suppl: S111-130.
- Mukhopadhyay, D. and Dasso, M., 2007. Modification in reverse: the SUMO proteases. *Trends in Biochemical Sciences*, 32 (6): 286-295.
- Murtas, G., Reeves, P.H., Fu, Y.F., Bancroft, I., Dean, C. and Coupland, G., 2003. A nuclear protease required for flowering-time regulation in Arabidopsis reduces the abundance of SMALL UBIQUITIN-RELATED MODIFIER conjugates. *The Plant Cell*, 15 (10): 2308-2319.
- Muthuswamy, S. and Meier, I., 2011. Genetic and environmental changes in SUMO homeostasis lead to nuclear mRNA retention in plants. *Planta*, 233(1), pp.201-208.
- Nayak, A., Glöckner-Pagel, J., Vaeth, M., Schumann, J.E., Buttman, M., Bopp, T., Schmitt, E., Serfling, E., Berberich-Siebelt, F., 2009. Sumoylation of the Transcription Factor NFATc1 Leads to Its Subnuclear Relocalization and Interleukin-2 Repression by Histone Deacetylase. *The Journal of Biological Chemistry* 284 (16): 10935-10946.
- Neff, M.M., Neff, J.D., Chory, J. and Pepper, A.E., 1998. dCAPS, a simple technique for the genetic analysis of single nucleotide polymorphisms: experimental applications in Arabidopsis thaliana genetics. *The Plant Journal: For Cell and Molecular Biology* 14 (3): 387-392.
- Novatchkova, M., Bachmair, A., Eisenhaber, B. and Eisenhaber, F., 2005. Proteins with two SUMO-like domains in chromatin-associated complexes: The RENi (Rad60-Esc2-NIP45) family. *BMC Bioinformatics*, 6(1): 22.
- Okada, S., Nagabuchi, M., Takamura, Y., Nakagawa, T., Shinmyozu, K., Nakayama, J. and Tanaka, K., 2009. Reconstitution of Arabidopsis thaliana SUMO pathways in E. coli: functional evaluation of SUMO machinery proteins and mapping of SUMOylation sites by mass spectrometry. *Plant & Cell Physiology*, 50 (6): 1049-1061.
- Onouchi, H., Igeño, M.I., Périlleux, C., Graves, K. and Coupland, G. 2000. Mutagenesis of plants overexpressing CONSTANS demonstrates novel interactions among Arabidopsis flowering-time genes. *The Plant Cell*, 12 (6): 885-900.
- Orengo, C.A., Jones, D.T. and Thornton, J.M., 1994. Protein superfamilies and domain superfolds. *Nature*, 372 (6507): 631-634.
- Ouyang, J., Valin, A. and Gill, G., 2009. Regulation of transcription factor activity by SUMO modification. *Methods in Molecular Biology*, 497: 141-152.
- Page, D.R. and Grossniklaus, U., 2002. The art and design of genetic screens: Arabidopsis thaliana. *Nature Reviews. Genetics*, 3 (2): 124-136.

- Palancade, B. and Doye, V., 2008. Sumoylating and desumoylating enzymes at nuclear pores: underpinning their unexpected duties? *Trends in Cell Biology*, 18(4): 174-183.
- Park, H.C., Choi, W., Park, H.J., Cheong, M.S., Koo, Y.D., Shin, G., Chung, W.S., Kim, W.Y., Kim, M.G., Bressan, R.A., Bohnert, H.J., Lee, S.Y., Yun, D.J., 2011. Identification and molecular properties of SUMO-binding proteins in Arabidopsis. *Molecules and Cells*, 32 (2): 143-51.
- Park, B.S., Song, J.T. and Seo, H.S., 2011. Arabidopsis nitrate reductase activity is stimulated by the E3 SUMO ligase AtSIZ1. *Nature Communications*, 2:400.
- Peck, S.C., 2005. Update on proteomics in Arabidopsis. Where do we go from here? *Plant Physiology*, 138 (2): 591-599.
- Pellegrini, L., Tan, S., and Richmond, T.J., 1995. Structure of serum response factor core bound to DNA. *Nature* 376 (6540): 490-498.
- Perkins, J.R., Diboun, I., Dessailly, B.H., Lees, J.G. and Orengo, C., 2010. Transient protein-protein interactions: structural, functional, and network properties. *Structure*, 18 (10): 1233-1243.
- Prudden, J., Perry, J.J., Arvai, A.S., Tainer, J.A. and Boddy, M.N., 2009. Molecular Mimicry of SUMO Promotes DNA Repair. *Nature structural & molecular biology*, 16 (5): 509-516.
- Putterill, J., Robson, F., Lee, K., Simon, R. and Coupland, G., 1995. The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell*, 24; 80 (6): 847-857.
- Reeves, P.H. and Coupland, G. 2001. Analysis of flowering time control in Arabidopsis by comparison of double and triple mutants. *Plant Physiology*, 126 (3):1085-91.
- Reeves, P.H., Murtas, G., Dash, S. and Coupland, G., 2002. early in short days 4, a mutation in Arabidopsis that causes early flowering and reduces the mRNA abundance of the floral repressor FLC. *Development*, 129 (23): 5349-5361.
- Reverter, D. and Lima, C.D., 2005. Insights into E3 ligase activity revealed by a SUMO-RanGAP1-Ubc9-Nup358 complex. *Nature*, 435 (7042): 687-692.
- Riechmann, J.L., Krizek, B.A. and Meyerowitz, E.M., 1996. Dimerization specificity of Arabidopsis MADS domain homeotic proteins APETALA1, APETALA3, PISTILLATA, and AGAMOUS. *Proceedings of the National Academy of Sciences of the United States of America*, 93 (10): 4793-4798.
- Riechmann, J.L., and Meyerowitz, E.M., 1997. MADS domain proteins in plant development. *Biological Chemistry*, 378 (10): 1079-1101.

- Rodriguez, M.S., Dargemont, C. and Hay, R.T., 2001. SUMO-1 conjugation in vivo requires both a consensus modification motif and nuclear targeting. *The Journal of Biological Chemistry*, 276 (16): 12654-12659.
- Ross, S., Best, J.L., Zon, L.I. and Gill, G., 2002. SUMO-1 modification represses Sp3 transcriptional activation and modulates its subnuclear localization. *Molecular Cell*, 10 (4): 831-842.
- Samach, A., Onouchi, H., Gold, S., Ditta, G., Schwarz-Sommer, Z., Yanofsky, M., and Coupland, G. (2000). Distinct roles of CONSTANS target genes in reproductive development of Arabidopsis. *Science*, 288:1613–1616.
- Sambrook, J. and Russel, D.W. 2001. Molecular cloning: A laboratory manual. Cold Spring Harbour Laboratories Press, New York, 3rd Edition.
- Santelli, E. and Richmond, T.J., 2000. Crystal structure of MEF2A core bound to DNA at 1.5 Å resolution. *Journal of Molecular Biology*, 297 (2): 437-449.
- Saracco, S.A., Miller, M.J., Kurepa, J. and Vierstra, R.D., 2007. Genetic analysis of SUMOylation in Arabidopsis: conjugation of SUMO1 and SUMO2 to nuclear proteins is essential. *Plant Physiology*, 145 (1): 119-134.
- Searle, I., He, Y., Turck, F., Vincent, C., Fornara, F., Kröber, S., Amasino, R.A., Coupland, G., 2006. The transcription factor FLC confers a flowering response to vernalization by repressing meristem competence and systemic signaling in Arabidopsis. *Genes & Development* 20 (7): 898-912.
- Seo, J. and Lee, K.J., 2004. Post-translational modifications and their biological functions: proteomic analysis and systematic approaches. *Journal of Biochemistry and Molecular Biology*, 37 (1):35-44.
- Sharrocks, A.D., 2006. PIAS proteins and transcriptional regulation—more than just SUMO E3 ligases? *Genes & Development*, 20(7): 754 -758.
- Sheldon, C.C., Burn, J.E., Perez, P.P., Metzger, J., Edwards, J.A., Peacock, W.J. and Dennis, E.S. 1999. The FLF MADS box gene: a repressor of flowering in Arabidopsis regulated by vernalization and methylation. *The Plant Cell*, 11 (3):445-458.
- Sheldon, C.C., Rouse, D.T., Finnegan, E.J., Peacock, W.J. and Dennis ES., 2000. The molecular basis of vernalization: the central role of FLOWERING LOCUS C (FLC). *Proceedings of the National Academy of Sciences of the United States of America*, 97 (7):3753-3758.
- Sheldon, C.C., Conn, A.B., Dennis, E.S. and Peacock, W.J., 2002. Different Regulatory Regions Are Required for the Vernalization-Induced Repression of FLOWERING LOCUS C and for the Epigenetic Maintenance of Repression. *The Plant Cell*, 14 (10): 2527 -2537.

- Shen, L., Kang, Y.G.G., Liu, L. and Yu, H. 2011. The J-Domain protein J3 mediates the integration of flowering signals in Arabidopsis. *The Plant Cell*, 23: 499-514.
- Shore, P. and Sharrocks, A.D., 1995. The MADS-box family of transcription factors. *European Journal of Biochemistry / FEBS* 229 (1): 1-13.
- Simpson, G.G. and Dean, C., 2002. Arabidopsis, the Rosetta stone of flowering time? *Science*, 296 (5566): 285-289.
- Smith, M., Mallin, D.R., Simon, J.A. and Courey, A.J., 2011. Small ubiquitin-like modifier (SUMO) conjugation impedes transcriptional silencing by the polycomb group repressor Sex Comb on Midleg. *The Journal of Biological Chemistry*, 286 (13): 11391-11400.
- Song, J., Durrin, L.K., Wilkinson, T.A., Krontiris, T.G. and Chen, Y., 2004. Identification of a SUMO-binding motif that recognizes SUMO-modified proteins. *Proceedings of the National Academy of Sciences of the United States of America*, 101 (40):14373-14378.
- Soppe, W.J., Jacobsen, S.E., Alonso-Blanco, C., Jackson, J.P., Kakutani, T., Koornneef, M. and Peeters, A.J., 2000. The Late Flowering Phenotype of *fwa* Mutants Is Caused by Gain-of-Function Epigenetic Alleles of a Homeodomain Gene. *Molecular Cell*, 6 (4):791-802.
- Srikanth, A. and Schmid, M., 2011. Regulation of flowering time: all roads lead to Rome. *Cellular and Molecular Life Sciences: CMLS*, 68 (12):2013-2037.
- Srivatsan, A., Han, Y., Peng, J., Tehranchi, A.K., Gibbs, R., Wang, J.D. and Chen, R., 2008. High-precision, whole-genome sequencing of laboratory strains facilitates genetic studies. *PLoS Genetics* 4 (8): e1000139.
- Sunilkumar, G., Mohr, L., Lopata-Finch, E., Emani, C. and Rathore, K.S., 2002. Developmental and tissue-specific expression of CaMV 35S promoter in cotton as revealed by GFP. *Plant Molecular Biology* 50 (3): 463-474.
- Tatham, M.H., Jaffray, E., Vaughan, O.A., Desterro, J.M., Botting, C.H., Naismith, J.H. and Hay, R.T., 2001. Polymeric chains of SUMO-2 and SUMO-3 are conjugated to protein substrates by SAE1/SAE2 and Ubc9. *The Journal of Biological Chemistry*, 276 (38): 35368-35374.
- Tatham, M.H., Chen, Y. and Hay, R.T., 2003. Role of two residues proximal to the active site of Ubc9 in substrate recognition by the Ubc9.SUMO-1 thiolester complex. *Biochemistry*, 42 (11): 3168-3179.
- Tatham, M.H., Kim, S., Yu, B., Jaffray, E., Song, J., Zheng, J., Rodriguez, M.S., Hay R.T. and Chen Y., 2003. Role of an N-terminal site of Ubc9 in SUMO-1, -2, and -3 binding and conjugation. *Biochemistry*, 42(33): 9959-9969.

- Theissen, G., Becker, A., Di Rosa, A., Kanno, A., Kim, J.T., Münster, T., Winter, K.U. and Saedler, H., 2000. A short history of MADS-box genes in plants. *Plant Molecular Biology*, 42 (1): 115-149.
- Tong, H., Hateboer, G., Perrakis, A., Bernards, R. and Sixma, T.K., 1997. Crystal structure of murine/human Ubc9 provides insight into the variability of the ubiquitin-conjugating system. *The Journal of Biological Chemistry*, 272(34): 21381-21387.
- Turck, F., Fornara, F. and Coupland, G., 2008. Regulation and identity of florigen: FLOWERING LOCUS T moves center stage. *Annual Review of Plant Biology*, 59: 573-594.
- Ulrich, H.D., 2008. The fast-growing business of SUMO chains. *Molecular Cell*, 32(3): 301-305.
- Uzunova, K., Götsche, K., Miteva, M., Weisshaar, S.R., Glanemann, C., Schnellhardt, M., Niessen, M., Scheel, H., Hofmann, K., Johnson, E.S., Praefcke, G.J. and Dohmen, R.J., 2007. Ubiquitin-dependent proteolytic control of SUMO conjugates. *The Journal of Biological Chemistry*, 282 (47): 34167-34175.
- van Berloo, R. 2008. GGT 2.0: versatile software for visualization and analysis of genetic data. *The Journal of Heredity*, 99 (2): 232-236.
- van den Burg, H.A., Kini, R.K., Schuurink, R.C. and Takken, F.L., 2010. Arabidopsis small ubiquitin-like modifier paralogs have distinct functions in development and defense. *The Plant Cell* 22 (6): 1998-2016.
- Walsh, C.T., Garneau-Tsodikova, S. and Gatto, G.J.Jr. 2005. Protein posttranslational modifications: the chemistry of proteome diversifications. *Angewandte Chemie*, 44 (45): 7342-7372.
- Wang, G., Ellendorff, U., Kemp, B., Mansfield, J.W., Forsyth, A., Mitchell, K., Bastas, K., Liu, C.M., Woods-Tör, A., Zipfel, C., de Wit, P.J., Jones, J.D., Tör M. and Thomma BP. 2008. A genome-wide functional investigation into the roles of receptor-like proteins in Arabidopsis. *Plant Physiology*, 147 (2): 503-517.
- Wang, J., Hu, W., Cai, S., Lee, B., Song, J. and Chen, Y., 2007. The intrinsic affinity between E2 and the Cys domain of E1 in ubiquitin-like modifications. *Molecular Cell*, 27(2): 228-237.
- Wang, J., Taherbhoy, A.M., Hunt, H.W., Seyedin, S.N., Miller, D.W., Miller, D.J., Huang, D.T. and Schulman, B.A., 2010. Crystal structure of UBA2(ufd)-Ubc9: insights into E1-E2 interactions in Sumo pathways. *PloS One*, 5(12), p.e15805.

- Weigel, D. and Glazebrook, J. *Arabidopsis: a laboratory manual*. New York, Cold Spring Harbour, 2002. 354 pp.
- Wildermuth, M.C., Dewdney, J., Wu, G. and Ausubel, F.M., 2001. Isochorismate synthase is required to synthesize salicylic acid for plant defence. *Nature*, 414 (6863): 562-565.
- Wilkinson, K.A. and Henley, J.M., 2010. Mechanisms, regulation and consequences of protein SUMOylation. *The Biochemical Journal*, 428 (2): 133-145.
- Wilson, R., Heckman, J. and Somerville, C. 1992. Gibberellin is required for flowering in *Arabidopsis thaliana* under short days. *Plant Physiology*, 100:403-408
- Wu, S.Y. and Chiang, C.M., 2009. Crosstalk between sumoylation and acetylation regulates p53-dependent chromatin transcription and DNA binding. *The EMBO Journal*, 28 (9): 1246-1259.
- Xu, J., He, Y., Qiang, B., Yuan, J., Peng, X. and Pan, X.M., 2008. A novel method for high accuracy sumoylation site prediction from protein sequences. *BMC Bioinformatics* 9: 8.
- Xu, J., Li, Y., Wang, Y., Liu, H., Lei, L., Yang, H., Liu, G. and Ren, D., 2008. Activation of MAPK kinase 9 induces ethylene and camalexin biosynthesis and enhances sensitivity to salt stress in *Arabidopsis*. *The Journal of Biological Chemistry*, 283 (40): 26996-27006.
- Xu, X.M., Rose, A., Muthuswamy, S., Jeong, S.Y., Venkatakrishnan, S., Zhao, Q. and Meier, I., 2007. NUCLEAR PORE ANCHOR, the *Arabidopsis* homolog of Tpr/Mlp1/Mlp2/megator, is involved in mRNA export and SUMO homeostasis and affects diverse aspects of plant development. *The Plant Cell* 19 (5): 1537-1548.
- Xu, Z., Chan, H.Y., Lam, W.L., Lam, K.H., Lam, L.S., Ng, T.B. and Au, S.W., 2009. SUMO proteases: redox regulation and biological consequences. *Antioxidants & Redox Signaling*, 11 (6):1453-1484.
- Xue, Y., Zhou, F., Fu, C., Xu, Y. and Yao, X., 2006. SUMOsp: a web server for sumoylation site prediction. *Nucleic Acids Research*, 1;34: W254-257.
- Yang, S.H., Galanis, A., Witty, J. and Sharrocks, A.D., 2006. An extended consensus motif enhances the specificity of substrate modification by SUMO. *The EMBO Journal*, 25 (21): 5083-5093.
- 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: For Cell and Molecular Biology*, 33 (1): 47-59.

- Yoo, C.Y., Miura, K., Jin, J.B., Lee, J., Park, H.C., Salt, D.E., Yun, D.J., Bressan, R.A. and Hasegawa, P.M., 2006. SIZ1 small ubiquitin-like modifier E3 ligase facilitates basal thermotolerance in Arabidopsis independent of salicylic acid. *Plant Physiology*, 142 (4): 1548-1558.
- Zachgo, S., Silva Ede, A., Motte, P., Tröbner, W., Saedler, H. and Schwarz-Sommer Z.. 1995. Functional analysis of the *Antirrhinum* floral homeotic DEFICIENS gene in vivo and in vitro by using a temperature-sensitive mutant. *Development*, 121 (9): 2861-2875.
- Zhao, Q. and Meier, I., 2011. Identification and characterization of the Arabidopsis FG-repeat nucleoporin Nup62. , 6(3), pp.330-334.

A.1. FIGURES

Figure 1.1	SUMO-cycle	3
Figure 2.1	The <i>suppressors of esd4-1</i> exhibit a late flowering phenotype under Short Day (SD) and Long Day (LD) conditions	31
Figure 2.2	Phenotype of the <i>esd4-1</i> mutant	32
Figure 2.3	<i>suppressors of esd4-1</i> phenotypes under Long Day conditions	33
Figure 2.4	<i>suppressors of esd4-1</i> phenotypes under Short Day conditions	35
Figure 2.5	Flowering time of the fifteen selected <i>suppressors of esd4-1</i>	37
Figure 2.6	<i>suppressors of esd4-1</i> present normal developed flowers that are larger than the flowers of <i>esd4</i>	38
Figure 2.7	Silique phenotypes of the fifteen selected <i>suppressors of esd4-1</i> under LD conditions	38
Figure 2.8	<i>suppressors of esd4-1</i> produce larger leaves than <i>esd4</i>	39
Figure 2.9	Photograph illustrating the <i>suppressors of esd4-1</i> leaf phenotypes under LD conditions	40
Figure 2.10	<i>suppressors of esd4-1</i> produce viable pollen but in lower amount than <i>esd4</i> under SD	42
Figure 2.11	SUMO expression levels in <i>sed</i> mutants grown under LD conditions	44
Figure 2.12	SUMO Activating Enzyme (SAE) subunits expression levels in <i>sed</i> mutants grown under LD conditions	45
Figure 2.13	SUMO Conjugating Enzyme (SCE) and E3 Ligase Siz1 expression levels in <i>sed</i> mutants grown under LD conditions	46
Figure 2.14	Analysis of SUMO-conjugate protein levels in the <i>sed</i> mutants under LD	47
Figure 3.1	Generation of mapping populations of <i>sed</i> mutants	55

Figure 3.2	Phenotype of the F1 hybrid of the <i>esd4-1</i> (L <i>er</i>), <i>esd4-2</i> (Col) mutants	56
Figure 3.3	Frequency distribution of the flowering time of <i>esd4-1</i> (L <i>er</i>) x <i>esd4-2</i> (Col) F2 population under SD conditions	57
Figure 3.4	The <i>esd4-1</i> (L <i>er</i>) x <i>esd4-2</i> (Col) F2 population presents a broad phenotypic variation	58
Figure 3.5	Flowering time frequency distribution of four independent <i>sed</i> (L <i>er</i>) x <i>esd4-2</i> (Col) F2 segregating populations	59
Figure 3.6	Flowering time frequency distribution of the <i>sed5-1</i> x <i>esd4-2</i> and <i>sed45-1</i> x <i>esd4-2</i> F2 segregating populations	59
Figure 3.7	Flowering time of the F2 mapping populations of the <i>sed4-1</i> , <i>sed43-1</i> and <i>sed44-3</i> suppressor lines	60
Figure 3.8	Marker-trait association plot of the <i>sed4-1</i> , <i>sed43-1</i> and <i>sed44-3</i> mapping populations	63
Figure 3.9	Consensus genotype map of F2 <i>esd4-1</i> x <i>esd4-2</i> hybrid population	64
Figure 3.10	<i>sed111-1</i> rough map position by bulk segregant analysis	66
Figure 3.11	<i>sed111-1</i> suppressor mutant background cleaning F3 BC1 segregants	67
Figure 3.12	<i>de novo</i> dCAPS markers	69
Figure 3.13	Linkage analysis of some of the SNPs found in Chromosome I in <i>sed111-1</i> plants	70
Figure 3.14	A6-C6 plants phenotypes	76
Figure 3.15	Bulk Segregant Analysis and rough map position of A6-C6	77
Figure 4.1	SVP protein conserved domains and potential SUMO-attachment sites	86
Figure 4.2	Multiple sequence clustal global alignment of potential SVP homologues in different plant species	87

Figure 4.3	Residue conservation of SUMO-attachment motifs in SVP homologues	88
Figure 4.4	SVP secondary structure prediction	90
Figure 4.5	Disorder prediction for the SVP protein sequence	91
Figure 4.6	SVP is SUMOylated <i>in vitro</i> by SUMO1 and SUMO3	93
Figure 4.7	Effect of single point mutations on SVP SUMO1 SUMOylation	94
Figure 4.8	Effect of double point mutations on modification of SVP by SUMO1	95
Figure 4.9	SUMOylation of SVP by SUMO1 is abolished in the K30, 31, 53, 78, 156R Penta-SUMO-site Mutant	96
Figure 4.10	Effect of single point mutations on SUMOylation of SVP by SUMO3	97
Figure 4.11	Effect of double point mutations on SUMOylation of SVP by SUMO3	98
Figure 4.12	Explicative model for the creation of “Hyper” and “Hypo” SUMOylated versions of SVP	99
Figure 4.13	Flowering time of independent T2 lines overexpressing SVP amino-terminal fusions to SUMO or the SCE grown under long day conditions	101
Figure 4.14	Molecular characterization of the FLAG N-terminal SVP fusion constructs.	102
Figure 4.15	Flowering time of independent lines overexpressing SVP carboxyl-terminal fusions to SUMO or the SCE grown under long day conditions	103
Figure 4.16	Flowering time of independent T2 lines overexpressing different versions of SVP where the putative SUMO-attachment sites are mutated	105

Figure 4.17	Phenotypical characterization of the FLAG N-terminal SVP fusion constructs under Long Days	108
Figure 4.18	Phenotypes of 35S::FLAG:SUMO1:SVP, <i>svp-41</i> line under Long Day conditions	109
Figure 4.19	Phenotype of 35S::FLAG:SUMO3:SVP, <i>svp-41</i> line grown under Long Day conditions	110
Figure 4.20	Phenotype of 35S::FLAG:SCE:SVP, <i>svp-41</i> amino-terminal fusion construct under Long Day conditions	110
Figure 4.21	35S::SVP:SUMO1:FLAG lines recover the early flowering phenotype of <i>svp-41</i> and present morphological alterations in the aerial shoot and flowers under Long Day conditions	111
Figure 4.22	35S::SVP:SCE:FLAG lines recover the early flowering phenotype of <i>svp-41</i> and flower later than the Wt under Long Day conditions	111
Figure 4.23	35S::SVP:SCE(C94S):FLAG lines flower later than the Wt under Long Day conditions	112
Figure 4.24	35S::FLAG:SVP ^{K30,31,53,78,156R} or FLAG:SVP PSM in <i>svp-41</i> show a faster senescence in Long day conditions	113
Figure 4.25	35S::FLAG:SVP ^{K30,31,53R} lines present secondary flowers emerging from the second flower whorl of the primary flower	114
Figure 4.26	35S::FLAG:SVP ^{K30,31,78R} plants present inflorescence shoots instead of flowers	114
Figure 4.27	35S::FLAG:SVP ^{K53,78,156R} lines present alterations in the first three whorls of the flower	115
Figure 4.28	35S::FLAG:SVP ^{K78,156R} plants present flowers with alterations in the number of organs per floral whorl and conversions of petals into small green structures when grown under Long Day conditions	115
Figure 4.29	mRNA and protein expression analysis in the FLAG:SVP PSM and FLAG:SVP ^{K30,31,53R} and FLAG:SVP ^{k30,31,78R} construct lines	116

A.2. TABLES

Table 2.1	Pollen viability and relative amount in <i>sed</i> mutants under SD conditions	42
Table 2.2	SUMO pathway components in Arabidopsis	43
Table 3.1	Suppressor mapping crosses and backcrosses synthesis	61
Table 3.2	Mutated genes within the linked genetic interval containing the <i>sed111-1</i>	71
Table 4.1	Predicted SUMO-attachment sites in SVP	85
Table 4.2	SVP bioinformatic analysis and prediction synthesis	92
Table 4.3	SVP SUMO-site mutants and its effects on flower development	117
Table 6.1	Antibiotics	133
Table 6.2	Schematic representation for hyper- and hypo- SUMOylation constructs generation	141

A.3. Abbreviations

General abbreviations

A	Adenine
aa	amino acid
AGI	Arabidopsis Genome Initiative
ABRC	Arabidopsis Biological Resource Center
AP	alkaline phosphatase
Arabidopsis	<i>Arabidopsis thaliana</i>
<i>A. tumefaciens</i>	<i>Agrobacterium tumefaciens</i>
At	<i>Arabidopsis thaliana</i>
attR1, attR2	Attachment sites for site-specific recombination in the Gateway cloning system
BC1	Back Cross 1 (in the context of genetic crosses)
BLAST	Basic Local Alignment Search Tool
BSA	Bulked Segregant Analysis
bp	base pair
BWA	Burrows-Wheeler Aligner
C	Cytosine
C-	carboxy-terminal
CaMV	Cauliflower Mosaic Virus
CAPS	Cleaved Amplified Polymorphic Sequences
cDNA	complementary DNA
CDS	Coding sequence
cm	centimeter
CNV	Copy Number Variations
Col	<i>Arabidopsis thaliana</i> ecotype Columbia-0
dCAPS	derived Cleaved Amplified Polymorphic Sequences
DNA	deoxyribonucleic acid
dNTP	deoxyribonucleic triphosphate
DSM	Di-SUMO-site Mutant
<i>E. coli</i>	<i>Escherichia coli</i>
<i>EcoRI</i>	<i>Escherichia coli</i> RY13 I, restriction enzyme
e.g.	<i>Exempli gratia</i> [Lat.] for example
et al.	<i>et alii / et aliae</i> [Lat.] and others
EMS	Ethyl Methanesulfonate
F1, F2, F3	first, second, third... filial generation after a cross
G	Guanine
g	gram
GA	gibberellic acid
GE	Gel Electrophoresis

h	hour
HRP	horse radish peroxidase
kb	kilobase pair
kDa	kilo Dalton
LD	long-day
<i>Ler</i>	Landsberg <i>erecta</i>
LM	Linear Motif
M	molar (mol/l)
M1, M2, M3	first, second, third generation after mutagenesis
Mbp	Mega base pairs
mm	milimeter
min	minute
mol	mole
mRNA	messenger RNA
MTT	2,5-diphenyl tetrazolium bromide
N-	amino-terminal
Ni-NTA	Nickel-nitrilotriacetic acid
nt	nucleotide
Os	<i>Oryzum sativa</i> (Rice)
PcG	Polycomb group genes
PCR	polymerase chain reaction
PDB	Protein Data Bank
PEBP	phosphatidyl ethanolamine binding domain protein
pH	negative logarithm of proton concentration
PRC	Polycomb repressive complex
PSM	Penta-SUMO-site Mutant
PTM	Posttranslational Modification
QTL	Quantitative Trait Loci
RIL	Recombinant Inbreeding Line
RFLP	Restriction Fragment Length Polymorphism
RNA	ribonucleic acid
RNase	ribonuclease
RT-PCR	reverse transcription PCR
SA	salicylic acid
SAM	shoot apical meristem
SD	short-day
SDS_PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SE	standard error
SIM	SUMO interacting motif
SNP	Single Nucleotide Polymorphism
SSLP	Single Sequence Length Polymorphism
StUbls	SUMO-targeted Ubiquitin ligases
T	Thymine

TAIR	
T1, T2, T3	first, second, third transgenic generation after plant transformation
T-DNA	transferred DNA
TLN	Total Leaf Number
Trx	Thioredoxin
ULP	Ubiquitin Like specific Protease
UTR	untranslated region
wt	wild type
x	crossed to (crosses are always indicated in the order: female x male)
:	fused to (in the context of gene fusion constructs)
::	under the control of (in the context of promoter-gene constructs)
-	minus, not present
%	percentage
°C	degrees Celsius
3'	three prime end of a DNA fragment
35S	promoter of the Cauliflower Mosaic virus
2-DE	two dimensional gel electrophoresis
5'	five prime end of a DNA fragment
α-	in the context of a Western blot represent antibody against
μ	micro
ψ	hydrophobic amino acid (in the context of a protein motif)

Gene and protein names

The nomenclature for plant genes follows the Arabidopsis standard: *GENES* are written in upper case italics. Mutant *genes* are indicated in lower case italics. *PROTEINS* appear in upper case regular letters. Mutant proteins in lower case regular letters.

ABI5	ABA INSENSITIVE 5
ACT	ACTIN
AGL24	AGAMOUS-LIKE 24
AP1	APETALA 1
CLV2	CLAVATA 2
CO	CONSTANS
<i>COLDAIR</i>	<i>COLD ASSISTED INTRONIC NONCODING RNA</i>
COP1	CONSTITUTIVE PHOTOMORPHOGENESIS 1
CYP98A9	CYTOCHROME P450, FAMILY 98, SUBFAMILY A, POLYPEPTIDE 9
EL1	EARLY IN SHORT DAYS 4- LIKE 1 (homolog of ESD4)
ESD4	EARLY IN SHORT DAYS 4
FCA	* This gene has not been given a full name

FD -	* This gene has not been given a full name
FLC	FLOWERING LOCUS C
FRI	FRIGIDA
FT	FLOWERING LOCUS T
FWA	FLOWERING WAGENINGEN A
ICE1	INDUCER OF CBP EXPRESSION 1
ICS1	ISOCHORISMATE SYNTHASE
LFY	LEAFY
MPK3	MITOGEN-ACTIVATED PROTEIN KINASE 3
MKK9	MITOGEN-ACTIVATED PROTEIN KINASE 9
PHR1	PHOSPHATE STARVATION RESPONSE 1
RLP13	RECEPTOR LIKE PROTEIN 13
SAE	SUMO ACTIVATING ENZYME / E1
SCE	SUMO CONJUGATING ENZYME / E2
<i>sed</i>	<i>suppressor of esd4</i>
SEP1,2,3,4	SEPALLATA1,2,3,4
SIZ1	SAP and MIZ1
SOC1	SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1
SUMO	SMALL UBIQUITIN RELATED MODIFIER PROTEIN
SVP	SHORT VEGETATIVE PHASE
TMM	TOO MANY MOUTHS
Ub	UBIQUITIN
UPL2	UBIQUITIN-PROTEIN LIGASE 2

Amino acids

Alanine	Ala	A	Methionine	Met	M
Cysteine	Cys	C	Asparagine	Asn	N
Aspartic acid	Asp	D	Proline	Pro	P
Glutamic acid	Glu	E	Glutamine	Gln	Q
Phenylalanine	Phe	F	Arginine	Arg	R
Glycine	Gly	G	Serine	Ser	S
Histidine	His	H	Threonine	Thr	T
Isoleucine	Ile	I	Valine	Val	V
Lysine	Lys	K	Tryptophane	Trp	W
Leucine	Leu	L	Tyrosine	Tyr	Y

A.4. SVP homologs

gi|225459099|ref|XP_002285687.1|
gi|147769366|emb|CAN68106.1|
gi|147744373|gb|ABQ51099.1|
gi|261393597|emb|CAX51278.1|
gi|161158834|emb|CAM59075.1|
gi|161158832|emb|CAM59074.1|
gi|147744439|gb|ABQ51132.1|
gi|95981864|gb|ABF57917.1|
gi|71025328|gb|AAZ17550.1|
gi|9367234|emb|CAB97350.1|
gi|13384052|gb|AAK21250.1|AF335237
gi|194692682|gb|ACF80425.1|
gi|9367232|emb|CAB97349.1|
gi|261391552|emb|CAX11663.1|
gi|162457969|ref|NP_001105148.1|
gi|120407344|gb|ABM21529.1|
gi|164521127|gb|ABY60423.1|
gi|326521456|dbj|BAK00304.1|
gi|71025326|gb|AAZ17549.1|
gi|223943985|gb|ACN26076.1|
gi|29372750|emb|CAD23409.1|
gi|115448827|ref|NP_001048193.1|
gi|161158818|emb|CAM59067.1|
gi|118767201|gb|ABL11476.1|
gi|261393654|emb|CAX51307.1|
gi|261393633|emb|CAX51296.1|
gi|55792842|gb|AAV65503.1|
gi|261393656|emb|CAX51308.1|
gi|261393562|emb|CAX51259.1|
gi|261393631|emb|CAX51295.1|
gi|261393599|emb|CAX51279.1|
gi|261393564|emb|CAX51260.1|
gi|261393579|emb|CAX51268.1|
gi|147744437|gb|ABQ51131.1|
gi|261393566|emb|CAX51261.1|
gi|261393568|emb|CAX51262.1|
gi|55792840|gb|AAV65502.1|
gi|261393650|emb|CAX51305.1|
gi|261391554|emb|CAX11664.1|
gi|55792846|gb|AAV65505.1|
gi|261393615|emb|CAX51287.1|
gi|261393641|emb|CAX51300.1|
gi|147744369|gb|ABQ51097.1|
gi|55792837|gb|AAV65501.1|
gi|2735764|gb|AAB94005.1|
gi|55792844|gb|AAV65504.1|
gi|316890774|gb|ADU56833.1|
gi|224094973|ref|XP_002310310.1|
gi|224063317|ref|XP_002301093.1|
gi|13448660|gb|AAK27151.1|
gi|225467973|ref|XP_002269295.1|
gi|267850663|gb|ACY82403.1|
gi|13448658|gb|AAK27150.1|
gi|122056647|gb|ABD66219.2|
gi|113207065|emb|CAL36572.1|
gi|83999600|emb|CAG27846.1|
gi|218192223|gb|EEC74650.1|
gi|30681743|ref|NP_179840.2|
gi|40806814|gb|AAR92206.1|
gi|297821479|ref|XP_002878622.1|
gi|30983948|gb|AAP40641.1|
gi|58201617|gb|AAW66885.1|
gi|30575602|gb|AAP33087.1|
gi|17433048|sp|Q9FUY6.1|
gi|6652756|gb|AAF22455.1|
gi|12964064|emb|CAC29335.1|
gi|33621117|gb|AAQ23142.1|
gi|115371646|gb|ABI96182.1|
gi|297600444|ref|NP_001049203.2|
gi|91207152|sp|Q5K4R0.2|
gi|194693938|gb|ACF81053.1|
gi|166014267|gb|ABY78023.1|
gi|198385780|gb|ACH86229.1|
gi|108706565|gb|ABF94360.1|
gi|194698260|gb|ACF83214.1|
gi|195625994|gb|ACG34827.1|
gi|224081933|ref|XP_002306534.1|
gi|95981862|gb|ABF57916.1|
gi|326415788|gb|ADZ72841.1|
gi|16549058|dbj|BAB70736.1|
gi|125490315|dbj|BAF46766.1|
gi|210148492|gb|ACJ09169.1|
gi|7672991|gb|AAF66690.1|AF144623_
gi|61611671|gb|AAX47170.1|
gi|326415786|gb|ADZ72840.1|
gi|297799552|ref|XP_002867660.1|
gi|23304690|emb|CAD48304.1|
gi|15233857|ref|NP_194185.1|
gi|255635649|gb|ACU18174.1|
gi|224095810|ref|XP_002310488.1|

A.5. SVP CDS sequence

SVP CDS sequence plus Gateway oligonucleotides* (* indicated in CAPITALS)

```

1   GGGGACAAGT TGTACAAAA AAGCAGGCTT Catggcgaga gaaaagattc      51
51  agatcaggaa gatcgacaac gcaacggcga gacaagtgac gttttcgaaa      101
101 cgaagaagag ggcttttcaa gaaagctgaa gaactctccg ttctctgcga      151
151 cgccgatgtc gctctcatca tcttctcttc caccggaaaa ctgttcgagt      201
201 tctgtagctc cagcatgaag gaagtcctag agaggcataa cttgcagtca      251
251 aagaacttgg agaagcttga tcagccatct cttgagttac agctggttga      301
301 gaacagtgat cacgcccgaa tgagtaaaga aattgcggac aagagccacc      351
351 gactaaggca aatgagagga gaggaacttc aaggacttga cattgaagag      401
401 cttcagcagc tagagaaggc ccttgaaact ggtttgacgc gtgtgattga      451
451 aacaaagagt gacaagatta tgagtgaagt cagcgaactt cagaaaaagg      501
501 gaatgcaatt gatggatgag aacaagcggg tgaggcagca aggaacgcaa      551
551 ctaacggaag agaacgagcg acttggcatg caaatatgta acaatgtgca      601
601 tgcacacggt ggtgctgaat cggagaacgc tgctgtgtac gaggaaggac      651
651 agtcgtcgga gtctattact aacgccggaa actctaccgg agcgcctggt      701
701 gactccgaga gctccgacac ttcccttagg ctcggcttac cgtatggtgg      751
751 ttagGACCCA GCTTCTTGT ACAAAGTGGT CCCC

```

Transgenic designation	Transformed into	Selection
HyperSUMOylation N-term		
p35S::3xFLAG:SVP	<i>svp-41</i> , Col	BASTA resistant
p35S::3xFLAG:SUMO1:SVP	<i>svp-41</i> , Col	BASTA resistant
p35S::3xFLAG:SUMO3:SVP	<i>svp-41</i> , Col	BASTA resistant
p35S::3xFLAG:SCE:SVP	<i>svp-41</i> , Col	BASTA resistant
p35S::3xFLAG:SCE(C94S):SVP	<i>svp-41</i> , Col	BASTA resistant
HypoSUMOylation		
p35S::3xFLAG:SVP ^{K30,31,53,78,156R}	<i>svp-41</i> , Col	BASTA resistant
p35S::3xFLAG:SVP ^{K30,31,53R}	<i>svp-41</i> , Col	BASTA resistant
p35S::3xFLAG:SVP ^{K30,31R}	<i>svp-41</i> , Col	BASTA resistant
p35S::3xFLAG:SVP ^{K53,78,156R}	<i>svp-41</i> , Col	BASTA resistant
p35S::3xFLAG:SVP ^{K78,156R}	<i>svp-41</i> , Col	BASTA resistant
p35S::3xFLAG:SVP ^{K30,78R}	<i>svp-41</i> , Col	BASTA resistant
p35S::3xFLAG:SVP ^{K53,156R}	<i>svp-41</i> , Col	BASTA resistant
p35S::3xFLAG:SVP ^{K30,31,78R}	<i>svp-41</i> , Col	BASTA resistant
HyperSUMOylation C-term		
p35S::SVP:3xFLAG	<i>svp-41</i> , Col	BASTA resistant
p35S::SVP:SUMO1:3xFLAG	<i>svp-41</i> , Col	BASTA resistant
p35S::SVP:SUMO3:3xFLAG	<i>svp-41</i> , Col	BASTA resistant
p35S::SVP:SCE:3xFLAG	<i>svp-41</i> , Col	BASTA resistant
p35S::SVP:SCE(C94S):3xFLAG	<i>svp-41</i> , Col	BASTA resistant

Acknowledgements

I would like to thank my supervisor Professor Dr. George Coupland for his constant support, guidance, key suggestions, corrections and vivid discussions along the PhD journey. Thanks George for the opportunity to work in this challenging project and allow me to develop as a scientist.

I would also like to thank Professor Dr. Ute Höcker for being chair of the disputation and to Professor Dr. Martin Hülskamp for being the second examiner. Thank you for your kindness and patience.

Thanks to Dr. Wim Soppe, my second supervisor, for his kindness and helpful comments.

I want to thank very specially to Dr. Nabil Elrouby for unconditional support, discussion and help during the realization of this project. I want to thank him also for the critical reading, comments and corrections that helped me substantially in the edition of this thesis. Thanks Nabil for encourage me to give my very best and do not give up.

I want also to thank Dr. Karl Nördstrom for the analysis of the Next Generation Sequencing data and help.

Furthermore, I would like to thank all the members of the group of George Coupland at Max-Planck Institute for Plant Breeding Research for a nice atmosphere and encouraging phrases during hard moments, especially Dr. Amaury de Montaigne and Dr. Mark Rühl for interesting discussions and orientation; and Elisa de Ansorena Pablos for technical help.

I would also like to thank Max Planck Gesellschaft and SFB635 for the financial support of this project.

I want to especially thank and say *Gracias* to my parents, Fred and Maria, and my beloved brother, Sinuhé, and grandmother, Emma, for their unconditional help, love and support and for teaching me with their example that perseverance, discipline, and work would always help to surmount almost any obstacle. You inspire me I love you all.

I would like to give my deepest thanks to Asis Hallab for his help with some bioinformatic programmes and for his support full of love, tenderness and fun, I also want to thank with all my heart to all Hallabs Family members, Dr. Mohammed Hallab, Christa, Dr. Amina Hallab (:) and Leila, for being supportive, loving and caring.

I want to thank all the rest of my family and all my friends in *Deutschland* and back in *México* for nice chats and unforgettable experiences from which I have learned a lot.

Erklärung

Ich versichere, dass ich die von mir vorgelegte Dissertation selbständig angefertigt, die benutzten Quellen und Hilfsmittel vollständig angegeben und die Stellen der Arbeit - einschließlich Tabellen, Karten und Abbildungen -, die anderen Werken im Wortlaut oder dem Sinn nach entnommen sind, in jedem Einzelfall als Entlehnung kenntlich gemacht habe; dass diese Dissertation noch keiner anderen Fakultät oder Universität zur Prüfung vorgelegen hat; dass sie - abgesehen von unten angegebenen Teilpublikationen - noch nicht veröffentlicht worden ist sowie, dass ich eine solche Veröffentlichung vor Abschluss des Promotionsverfahrens nicht vornehmen werde. Die Bestimmungen dieser Promotionsordnung sind mir bekannt. Die von mir vorgelegte Dissertation ist von Prof. Dr. George Coupland betreut worden.

Köln, den 8. November 2011

Mitzi Villajuana Bonequi

Teilpublikationen

Es liegen keine Teilpublikationen vor.

Lebenslauf

Persönliche Informationen

Name	Villajuana Bonequi
Vorname	Mitzi
Geburtsort	Mexiko
Geburtstag	11. November 1980
Nationalität	Mexikanisch
Familienstand	Ledig

Ausbildung

1999	Abitur, Staatslyceum 6 “Antonio Caso”, Mexiko Stadt.
1999-2004	Studium Biologie, Staatsuniversität von Mexiko, Mexiko Stadt.
2004-2006	Masters “Functional characterization of two MADS-box genes <i>SOCI</i> like, (<i>AGL19</i> and <i>AGL14</i>) in the flowering time transition network”. Staatsuniversität von Mexiko – Institut von Ökologie. Betreuerin: Prof. Dr. Elena Álvarez-Buylla Roces.
Januar 2007	Beginn der Doktorarbeit am Max-Planck-Institut für Züchtungsforschung, Köln in Der Arbeitsgruppe von Prof. Dr. George Coupland.

Köln, den 8 November 2011

