

**TECHNISCHE UNIVERSITÄT MÜNCHEN**  
Fachgebiet für Entwicklungsbiologie der Pflanzen

**Molecular and Functional Analysis of the LRRV/SRF  
Family of Putative Leucine-Rich Repeat Receptor-Like  
Kinases in *Arabidopsis thaliana***

**Banu Eyüboğlu**

Vollständiger Abdruck der von der Fakultät Wissenschaftszentrum  
Weihenstephan für Ernährung, Landwirtschaft und Umwelt der  
Technischen Universität München zur Erlangung des akademischen  
Grades eines

**Doktors der Naturwissenschaften (Dr. rer. nat.)**

genehmigten Dissertation.

Vorsitzender: Univ.-Prof. Dr. A. Gierl  
Prüfer der Dissertation: 1. Univ.-Prof. Dr. K. H. Schneitz  
2. Priv.-Doz. Dr. E. J. Glawischnig

Die Dissertation wurde am 11.03.2008 bei der Technischen Universität  
München eingereicht und durch die Fakultät Wissenschaftszentrum  
Weihenstephan für Ernährung, Landnutzung und Umwelt am 08.04.2008  
angenommen.

*To my father and mother*

# Table of Contents

<b>Table of Contents</b> .....	i
<b>Summary</b> .....	vi
<b>Zusammenfassung</b> .....	viii

## Chapter 1. The plant receptor-like kinases

1.1 Evolution of the receptor-like kinases.....	1
1.2 Classification of plant receptor-like kinases.....	2
1.3 Leucine-rich repeat receptor-like kinases.....	3
1.3.1 Functions of the plant receptor-like kinases.....	3
1.3.2 Two examples of the plant LRR-RLKs having dual function.....	4
1.3.2.1 BAK1.....	4
1.3.2.2 ERECTA.....	6
1.4 Ligand specificity of the receptor-like kinases.....	7
1.5 <i>STRUBBELIG RECEPTOR FAMILY (SRF)</i> gene family.....	9
1.5.1 <i>STRUBBELIG (SUB)</i> , a leucine-rich repeat receptor-like kinase involved in plant development.....	9
1.5.2 <i>STRUBBELIG RECEPTOR FAMILY (SRF)</i> members.....	9
1.6 The scope of the study.....	10
1.6.1 In this study.....	10

## Chapter 2. Materials and Methods

2.1 Materials.....	11
2.1.1 Chemicals, reagents, and media.....	11
2.1.1.1 Chemicals.....	11
2.1.1.2 Restriction endonucleases.....	11
2.1.1.3 Polymerases.....	11
2.1.1.4 Bacteria and yeast growth media.....	11
2.1.1.5 Plant growth media.....	11
2.1.2 Microorganisms and plant materials used in the study.....	12
2.1.2.1 <i>Escherichia coli</i> strains and vectors.....	12
2.1.2.2 <i>Agrobacterium</i> strains and vectors.....	12
2.1.2.3 Yeast strains and vectors.....	12
2.1.2.4 Plant lines.....	12
2.2 Methods.....	13
2.2.1 Plant work.....	13
2.2.1.1 Plant growth conditions.....	13
2.2.1.2 Plant transformation.....	13
2.2.1.3 Seed sterilization.....	13
2.2.2 Nucleic acid purification.....	13
2.2.2.1 Bacterial plasmid isolation.....	13
2.2.2.2 Yeast plasmid isolation.....	13
2.2.2.3 <i>Arabidopsis</i> genomic DNA isolation.....	14
2.2.2.4 <i>Arabidopsis</i> RNA isolation.....	14
2.2.2.5 DNA purification from agarose gel.....	14
2.2.3 Polymerase chain reaction (PCR) method.....	15

2.2.3.1	Oligonucleotides.....	15
2.2.3.2	Rapid Amplification of cDNA Ends (RACE) .....	15
2.2.3.3	Generation of the full-length <i>SRF</i> cDNAs.....	17
2.2.3.3.1	Restriction digestion and ligation.....	17
2.2.3.3.2	Overlapping PCR.....	17
2.2.3.3.3	Asymmetric PCR.....	18
2.2.3.3.4	End-to-end PCR.....	19
2.2.3.3.5	Cloning of <i>SRF1A/B</i> gene from <i>Ler</i> cDNA.....	19
2.2.3.3.6	Site-directed Mutagenesis (SDM).....	20
2.2.3.3.7	Alternative splicing analysis of the <i>SRF</i> genes by RT-PCR.....	21
2.2.3.3.8	Semi-quantitative RT-PCR.....	22
2.2.4	Examination of the expressed sequence tags (EST).....	22
2.2.5	Generation of expression and yeast two-hybrid constructs.....	23
2.2.5.1	Overexpression constructs.....	23
2.2.5.2	Generation of <i>35S::SRF4:EGFP</i> construct.....	24
2.2.5.3	Yeast two-hybrid constructs.....	26
2.2.6	PCR screening strategy for T-DNA lines.....	27
2.2.6.1	T-DNA insertion lines analysis.....	27
2.2.6.1.1	Investigation of the <i>srf4</i> T-DNA insertion lines.....	27
2.2.6.1.2	Identification of the T-DNA insertion of <i>daeumling</i> ( <i>At1g79870</i> ).....	28
2.2.7	Nucleotide sequencing and bioinformatics analysis.....	28
2.2.7.1	Nucleotide sequencing.....	28
2.2.7.2	Sequence analysis.....	28
2.2.7.3	Polymorphism analysis.....	29
2.2.7.4	Phylogenetic analysis.....	29
2.2.8	Phenotypic analysis.....	30
2.2.8.1	Phenotypic analysis of the overexpressing transgenic lines.....	30
2.2.8.2	Phenotypic analysis of <i>SRF4</i> and <i>DAEUMLING</i> ( <i>DLG</i> ).....	31
2.2.8.2.1	Leaf size assessments.....	31
2.2.8.2.2	Stem size measurements.....	32
2.2.8.2.3	Hypocotyl measurements.....	32
2.2.8.2.4	Root measurements.....	32
2.2.9	Microscopy and art work.....	32
2.2.9.1	Cell size analysis.....	32
2.2.9.2	Analysis of the <i>35S::SRF4:EGFP</i> transgenic lines.....	33
2.2.9.2.1	Screening of EGFP transgenic lines.....	33
2.2.9.2.2	Live imaging of <i>35S::SRF4:EGFP</i> transgenic lines using confocal microscopy.....	33
2.2.9.2.3	Plasmolysis.....	33

### **Chapter 3. Cloning and structural analysis of the *STRUBBELIG RECEPTOR FAMILY* (SRF) members**

3.1	Introduction.....	34
3.1.1	Extracellular organization of LRR-RLKs.....	34
3.1.1.1	Leucine-rich repeats (LRRs).....	34
3.1.1.2	Proline rich region (PRR).....	34

3.1.1.3 PEST region.....	36
3.1.2 The intracellular kinase domain.....	36
3.1.3 Outlook on Chapter 3.....	37
3.2 Results.....	38
3.2.1 Generation of the full-length <i>SRF</i> members.....	38
3.2.1.1 Expressed sequenced tag analysis of <i>SRF</i> genes.....	38
3.2.1.2 Cloning and annotation of the full-length <i>SRF</i> cDNA sequences.....	38
3.2.2 Alternative splicing analysis of the <i>SRF</i> genes.....	40
3.2.3 Exon-intron organization and chromosomal distribution.....	42
3.2.4 Amino acid sequence analysis.....	45
3.2.5 Investigation of the extracellular domains of the <i>SRF</i> members.....	45
3.2.5.1 Leucine rich repeats (LRRs).....	47
3.2.5.2 Putative dimerization modules.....	47
3.2.5.3 PEST sequence.....	49
3.2.5.4 Proline-rich region (PRR).....	49
3.2.6 Examination of intracellular domains.....	49
3.2.7 Phylogenetic analysis.....	51
3.2.8 Assessment of substitution patterns.....	52
3.2.9 Polymorphism analysis in <i>SRF1</i> .....	55
3.2.10 Nucleotide substitutions in <i>SRF</i> gene family members.....	58
3.3 Discussion.....	59
3.3.1 Generation of full-length cDNA clones of all <i>SRF</i> members.....	59
3.3.1.1 EST.....	59
3.3.1.2 Generation of <i>SRF</i> full-length cDNAs by RACE approach.....	59
3.3.2 Sequence analysis of <i>SRF</i> genes.....	60
3.3.3 Extracellular domain analysis of <i>SRF</i> members.....	61
3.3.3.1 Leucine-rich repeats (LRRs).....	61
3.3.3.2 PEST motifs of <i>SRF</i> members.....	62
3.3.3.3 Proline-rich region (PRR).....	63
3.3.3.4 Paired cysteines.....	63
3.3.4 Transmembrane (TM) domain.....	64
3.3.5 Intracellular domain analysis of <i>SRF</i> members.....	65
3.3.5.1 Juxtamembrane domain.....	65
3.3.5.2 Kinase domain.....	65
3.3.5.2.1 Do <i>SRF</i> family members possess an inactive kinase domain?.....	67
3.3.5.2.2 Phylogenetic analysis of <i>SRFs</i> .....	69
3.3.5.3 C-terminal extension.....	69
3.3.6 Is <i>SRF1</i> under evolutionary selection?.....	70
3.3.7 The <i>SRF1</i> gene encodes two isoforms.....	72

## **Chapter 4. Functional analysis of the STRUBBELIG RECEPTOR FAMILY members**

4.1 Introduction.....	74
4.1.1 Anther development.....	74
4.1.1.1 Genes involved in anther development.....	75
4.1.1.2 LRR-RLKs play roles in anther and pollen development.....	75

4.1.2 Outlook on Chapter 4.....	77
4.2 Results.....	77
4.2.1 Overexpression analysis of the <i>SRF</i> family members.....	77
4.2.1.1 Male sterility phenotype.....	78
4.2.1.2 Seedling lethality phenotype.....	80
4.2.1.3 Examination of floral organs in <i>SRF</i> overexpression lines.....	81
4.2.1.4 The stem length of the respective <i>SRF</i> overexpression lines.....	82
4.2.1.5 Stem size, rosette leaves number.....	84
4.2.1.6 Examination of the hypocotyl length of the overexpression lines in Col-0 background.....	85
4.2.1.7 Analysis of the root length, hypocotyl length and rosette leaves number of the overexpression lines.....	88
4.2.1.8 Could <i>SRF</i> overexpression in <i>sub-1</i> background rescue the sub-1 phenotype?..	89
4.2.2 Identification of putative interaction partners of <i>SRF4</i> , <i>SRF5</i> , and <i>SRF6</i> .....	90
4.2.2.1 Examination of the putative interacting partners.....	92
4.3 Discussion.....	94
4.3.1 Ectopic expression of number of the <i>SRF</i> genes exhibit male-sterile phenotype.....	95
4.3.1.1 Expression level of <i>SRF4</i> , <i>SRF5</i> , and <i>SRF7</i> affect the strength of male sterility.....	96
4.3.2 Overexpression of <i>SRF1</i> and <i>SRF8</i> resulted in seedling phenotype.....	96
4.3.3 What might be the function of <i>SRF4</i> , <i>SRF5</i> , and <i>SRF6</i> ?.....	97
4.3.3.1 Yeast two-hybrid analysis of <i>SRF4</i> .....	98
4.3.3.1.1 <i>At1g79870</i> , an oxidoreductase family protein.....	98
4.3.3.1.2 <i>At2g19640</i> , SET-domain containing protein.....	99
4.3.3.1.3 <i>At1g54290</i> , Eukaryotic translation initiation factor <i>SUI1</i> .....	100
4.3.1.4.4 <i>At1g43170</i> , Arabidopsis ribosomal protein 1.....	101
4.3.3.2 Putative interacting candidates of <i>SRF5</i> .....	101
4.3.3.2.1 <i>CALLOSE SYNTHASE 1</i> .....	101
4.3.3.2.2 C2 domain containing protein.....	102
4.3.3.2.3 <i>SEC5A</i> .....	104
4.3.3.3 Putative interacting partners of <i>SRF6</i> .....	105
4.3.4 Functional redundancy of the <i>SRF</i> genes.....	106

## Chapter 5. *SRF4* is a positive regulator of leaf development

5.1 Introduction.....	109
5.1.1 Leaf organogenesis .....	109
5.1.1.1 Genes involved in cell expansion.....	110
5.1.1.2 Genes involved in cell proliferation.....	111
5.1.1.3 Compensatory phenomena.....	112
5.1.2 Leaf morphogenesis not only depend on the cell morphology.....	113
5.1.2.1 Control of leaf development via plant hormones.....	113
5.1.2.2 Endoreduplication.....	114

5.1.3 Outlook on Chapter 5.....	115
5.2 Results .....	115
5.2.1 Functional analysis of the <i>SRF4</i> gene.....	115
5.2.1.1 <i>SRF4</i> affects the rosette leaves size.....	115
5.2.1.2 Phenotypic characterization of <i>dlg-1</i> .....	118
5.2.1.3 <i>SRF4</i> and <i>DLG</i> affect different type of leaves.....	119
5.2.1.3.1 Analysis of cotyledons.....	120
5.2.1.3.2 Cauline leaves measurements of transgenic plants.....	120
5.2.1.4 Investigation of stem length.....	121
5.2.1.5 Hypocotyl measurements of <i>srf4</i> and <i>dlg-1</i> mutants and wild type.....	122
5.2.1.6 Root size investigation.....	122
5.2.2 Cell size measurement.....	123
5.2.3 The cellular localization of the <i>SRF4</i> protein.....	123
5.3 Discussion.....	124
5.3.1 <i>SRF4</i> is a direct positive regulator of leaf size.....	124
5.3.2 Does <i>SRF4</i> affect cell number or cell size of leaf cells?.....	126
<b>Chapter 6. Conclusion and future studies.....</b>	<b>129</b>
<b>Chapter 7. References.....</b>	<b>131</b>
APPENDIX A.....	144
APPENDIX B.....	146
APPENDIX C.....	147
APPENDIX D.....	148
APPENDIX E.....	149
APPENDIX F.....	150
APPENDIX G.....	151
APPENDIX H.....	152
<b>Acknowledgements.....</b>	<b>153</b>
<b>Lebenslauf.....</b>	<b>154</b>

## SUMMARY

The *Arabidopsis* genome encodes more than 600 receptor-like kinases (RLKs) corresponding to about 2.5% of the *Arabidopsis* protein coding genes. These RLKs play important roles in cell-cell communication during development, hormone perception, pathogen resistance and self-incompatibility.

In this study, the cloning, structural and functional analysis of the nine LRR-VI *STRUBBELIG RECEPTOR FAMILY* (SRF) members encoding leucine-rich repeat receptor-like kinases (LRR-RLKs) is reported. Sequence analysis showed that all SRF family members except SRF1B have RLK configuration. According to phylogenetic analyses SRF2 is the most basal and primitive SRF member and SRF1/SRF3, SRF4/SRF5, and SRF6/SRF7 seem to have originated from relatively recent gene duplication events.

Interestingly, among the *SRF* members *SRF1* undergoes alternative splicing and is predicted to encode two different isoforms. *SRF1A* would encode an LRR-RLK whereas *SRF1B* a receptor-like protein (LRR-RLP) lacking the most part of the intracellular domain. Moreover, we found a high number of polymorphisms for the *SRF1* gene between the Col and *Ler* background indicating a probable evolutionary selection.

To understand the function of the *SRF* family members, investigation of loss-of function and gain-of-function mutants of the *SRF* genes in Col-0 and *Ler* backgrounds was carried out. The loss-of-function and gain-of function analysis of *SRF4* suggested that *SRF4* is a positive regulator of leaf size. In addition to the rosette leaves size, the results indicate that *SRF4* affects hypocotyl and stem length as well as the length of the cauline leaves, sepals and petals. These results provided direct evidence that the *SRF4* gene is a regulatory component controlling specifically organ size. Moreover, microscopic analysis of *SRF4* mutant lines showed that *SRF4* probably affects cell size rather than cell proliferation.

In addition, investigation of ectopic expression of several *SRF* members, using the 35S promoter, revealed that *SRF2-5* and *SRF7* might be related to anther development. Furthermore, overexpression of *SRF1A* and *SRF8* resulted in seedling lethality.

In addition to these approaches, yeast-two hybrid analysis of SRF4, SRF5, and SRF6 was carried out to identify probable interacting partners. This analysis also provided hints about the functions of the respective *SRF* genes. Especially, one of the putative interacting proteins of SRF4, DAEUMLING (DLG), a predicted D-isomer specific hydroxyacid reductase, may play a role in the regulatory process during leaf development because the phenotypes caused by T-DNA insertions into the two genes are related.

In summary, this study provides a comprehensive structural overview of the SRF family members as well as valuable first insights in their functions. A solid basis for further functional analyses was created.

## ZUSAMMENFASSUNG

Das *Arabidopsis* Genom kodiert für mehr als 600 Rezeptor-ähnliche Kinasen (RLKs), was in etwa 2,5% der Protein-kodierenden Gene in *Arabidopsis* entspricht. Diese RLKs spielen eine wichtige Rolle in der Zellkommunikation während der Entwicklung sowie in der Hormonwahrnehmung, Pathogenresistenz und Selbstinkompatibilität.

In dieser Arbeit wird über die Klonierung und die strukturelle und funktionelle Analyse der 9 Mitglieder umfassenden LRR-V/"*STRUBBELIG RECEPTOR FAMILY*" (SRF), die "leucin-rich repeat" RLKs (LRR-RLKs) kodieren, berichtet. Sequenzanalyse zeigte, dass alle SRF Familienmitglieder mit Ausnahme von *SRF1B* eine RLK Konfiguration besitzen. Phylogenetischen Analysen zufolge ist *SRF2* das basalste und primitivste der SRF Mitglieder und *SRF1/SRF3*, *SRF4/SRF5* und *SRF6/SRF7* scheinen aus relativ jungen Genduplikationsereignissen hervorgegangen zu sein.

Interessanterweise ist unter den *SRF* Mitgliedern *SRF1* von alternativem Splicing betroffen und kodiert wahrscheinlich für zwei verschiedene Isoformen. Dabei würde *SRF1A* eine LRR-LRK und *SRF1B* ein Rezeptor-ähnliches Protein (LRR-RLP), bei dem der größte Teil der intrazellulären Domäne fehlt, kodieren. Darüber hinaus haben wir eine große Zahl an Polymorphismen für das *SRF1* Gen zwischen dem Col und dem Ler Background gefunden, was auf eine mögliche evolutionäre Selektion hindeutet.

Um die Funktion der *SRF* Familienmitglieder zu verstehen, wurden Untersuchungen zu den „loss of function“ und „gain of function“ Mutanten der *SRF* Gene im Col und Ler Background durchgeführt. Dabei deutete die Analyse von „loss of function“ und „gain of function“ bei *SRF4* darauf hin, dass *SRF4* ein positiver Regulator der Blattgröße ist. Die Ergebnisse besagen, dass *SRF4* zusätzlich zu der Größe der Rosettenblätter auch die Länge der Hypokotylen und Stengel sowie die Länge der Kaulinblätter, der Sepalen und der Petalen beeinflusst. Diese Ergebnisse erbrachten einen direkten Hinweis darauf, dass das *SRF4* Gen eine spezifisch die Organgröße kontrollierende regulatorische Komponente ist. Darüberhinaus zeigte die mikroskopische Analyse von *SRF4* Mutantenlinien, dass *SRF4* wahrscheinlich eher auf die Zellgröße als auf die Zellvermehrung einwirkt.

Desweiteren offenbarte die Untersuchung der ektopischen Expression von mehreren *SRF* Mitgliedern unter Verwendung des 35S Promotors, dass *SRF2-5* und *SRF7* an der Antherenentwicklung beteiligt sein könnten. Darüber hinaus wiesen Pflanzen mit überexprimierten *SRF1A* und *SRF8* Genen Sämlingssterblichkeit auf.

Zusätzlich zu diesen Ansätzen wurden „Yeast-two hybrid“ Analysen von *SRF4*, *SRF5* und *SRF6* durchgeführt, um mögliche Interaktionspartner zu identifizieren. Diese Analyse erbrachte auch Hinweise über die Funktionen der jeweiligen *SRF* Gene. Insbesondere einer der putativen Interaktionspartner von *SRF4*, eine wahrscheinliche D-Isomer spezifische Hydroxysäurereduktase, *DAEUMLING* (*DLG*), könnte eine Rolle in dem regulatorischen Prozess während der Blattentwicklung spielen, da die durch T-DNA Insertionen in die beiden Gene verursachten Phänotypen ähnlich sind.

Diese Arbeit bietet zusammengefasst einen umfassenden strukturellen Überblick über die *SRF* Familie sowie erste wertvolle Einblicke in ihre Funktionen. Damit wurde eine solide Basis für weitere funktionelle Analysen geschaffen.

Molekulare und funktionale Analyse der Familie der LRRV/*SRF* "leucine-rich repeat" Rezeptorkinasen in *Arabidopsis thaliana*

## Chapter 1. Receptor-like kinases (RLKs) in plants

All living organisms perceive and assimilate external and internal signals and stimuli to continue their life. In animals, one of the important mechanisms to perceive signals and stimuli is mediated via the cell surface receptors (Hubbard and Till, 2000). In plants, different kinds of cell-surface receptors, notably the receptor-like kinases (RLKs), are involved in the stimuli perception (Shiu and Bleecker, 2001a). Genes encoding plant RLKs constitute one of the largest gene family in plant genomes and RLKs are implicated in all aspects of plant biology from early embryogenesis to disease resistance. Since plants are sessile organisms, these receptors possess a special importance for a quick response and high flexibility for the adaptation of the plants to the environment.

### 1.1 Evolution of RLKs

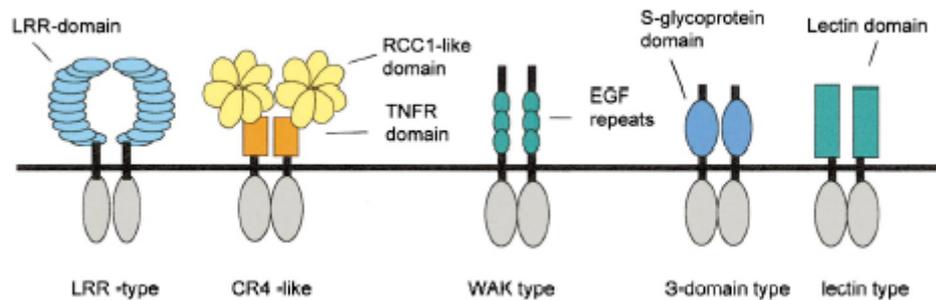
The comparison of plant and animal receptor kinases showed that animal genomes contain serine/threonine (ser/thr), histidine (his), and tyrosine (tyr) kinases while plant genomes consist of ser/thr and his kinases (Chevalier and Walker, 2005). In *Arabidopsis*, most of the RLKs have Ser/Thr. Although plant RLKs are structurally related to the tyrosine and serine/threonine receptor kinase families found in animals, these three classes of receptor kinases are grouped into a distinct monophyletic family (Shiu and Bleecker, 2001b; Cock *et al.*, 2002; Johnson and Ingram, 2005). Plant RLKs belong to the monophyletic interleukin-1 receptor-associated kinase (IRAK) or RLK/Pelle family whereas animal receptor tyrosine kinases (RTKs) and Raf kinases form another monophyletic group (Shiu and Bleecker, 2001b; Shiu *et al.*, 2004). Comparison of the number of the plant RLKs and *Drosophila* Pelle members revealed that the plant genome has the elevated number of RLKs (Shiu and Bleecker, 2001b). The RLKs represent one of the largest gene families of the *Arabidopsis* genome with more than 610 members corresponding to about 2.5% of *Arabidopsis* protein coding genes (Shiu and Bleecker, 2001a). Interestingly, 470 of the 610 RLK/Pelle family members are located within segmentally duplicated regions in *Arabidopsis* genome indicating that tandem duplications and segmental/whole

genome duplications are two major mechanisms for the expansion of the RLK/Pelle family (Shiu and Bleecker, 2003). The expansion of the plant RLKs could be related to the evolution of the plant lineages from the aquatic to the terrestrial environment which need for a quick response and high flexibility to adopt the rapidly changing environmental conditions (Shiu and Bleecker, 2001b).

These transmembrane receptors typically possess an amino-terminal extracellular domain, a transmembrane domain (TM), and a carboxyl-terminal intracellular kinase domain (Shiu and Bleecker, 2001a). Based on the presence and absence and identity of the extracellular domains and the kinase domain phylogeny, the RLK/Pelle family is subdivided into 46 subfamilies (Shiu and Bleecker, 2003). Most of the RLKs belong to the type I membrane proteins except for proline extensin receptor kinases (PERKs) from *Brassica* and thylakoid-associated kinase 1 (TAK1) from *Arabidopsis* (Shiu and Bleecker, 2001a; Snyders and Kohorn, 1999). In addition to receptor-like kinases, plants also utilize receptor like proteins (RLPs) that lack an intracellular domain and receptor-like cytoplasmic kinases (RLCKs) possessing an intracellular domain. Very few of the RLCKs have a known function, one exception is Arabidopsis PBS1, which is involved in disease resistance (Swiderski and Innes, 2001).

## **1.2 Classification of plant receptor-like kinases**

The plant RLKs are classified into subfamilies according to their domain organization, structure of the extracellular domains and the phylogenetic relationships between kinase domains of subfamily members (Torii and Clark, 2000; Shiu and Bleecker, 2001a). The S-domain class, tumor necrosis factor receptor (TNFR) class, epidermal growth factor (EGF) class, pathogenesis related protein (PR) class, lectin class, and leucine rich repeat (LRR) classes are the best known classes of the plant receptor-like kinases (Figure 1.1).



**Figure 1.1:** Structure and classification of plant RLKs. Major families of plant RLKs are classified according to their extracellular domains (McCarty and Chory, 2000).

### 1.3 Leucine-rich repeat receptor-like kinases

The largest RLK subfamily, including 239 members, consists of the leucine-rich repeat receptor-like kinases (LRR-RLKs) (Shiu and Bleecker, 2001a). LRR receptor-like kinases are further subclassified into 13 subfamilies according to the amino acid relationships between their kinase domains. Moreover, each LRR-RLK subfamily has characteristic repeats of LRRs in their extracellular domains.

#### 1.3.1 Functions of plant receptor-like kinases

LRR-RLKs play fundamental roles in cell-cell communication during development, hormone perception, abiotic and biotic stress responses (Torii, 2004).

Several LRR-RLKs have functions in plant growth and development. For instance, CLAVATA1 (CLV1), a LRR-RLK, together with CLAVATA2 (CLV2) and CLAVATA3 (CLV3), regulates the balance between cell proliferation and cell differentiation in the shoot meristem (Jeong *et al.*, 1999; Trotochaud *et al.*, 2000). CLV2 and CLV3 encode an LRR-receptor like protein and a small peptide, respectively. ERECTA plays a role in organ elongation (Torii *et al.*, 1996). BRASSINOSTEROIDS INSENSITIVE1 (BRI1) is involved in growth-promoting brassinosteroid signaling (Li and Chory, 1997; Li *et al.*, 2002; Nam and Li, 2002). Another LRR-RLKs, SOMATIC EMBRYOGENESIS RECEPTOR KINASES1 and 2 (SERK1 and SERK2), are crucial for anther development (Colcombet *et al.*, 2005). Plant LRR-RLKs also function as pattern recognition receptors in host innate immunity. FLAGELLIN SENSITIVE 2 (FLS2) and EF-Tu

receptor (EFR) play a role in bacterial elicitor perception (Gomez-Gomez and Boller, 2000; Zipfel *et al.*, 2006). Xa21 from *Oryza sativa* is involved in resistance to bacterial pathogens (Song *et al.*, 1995). As they play diverse biological roles it is important to understand them at a structural and functional level.

Interestingly, few of the RLKs are known to play roles in different biological functions such as ERECTA and SOMATIC EMBRYOGENESIS RECEPTOR KINASE 3 (SERK3)/BRI1 ASSOCIATED RECEPTOR KINASE1 (BAK1) RLKs (Torii *et al.*, 1996; Colcombet *et al.*, 2005). These different pathways will be described in the next section.

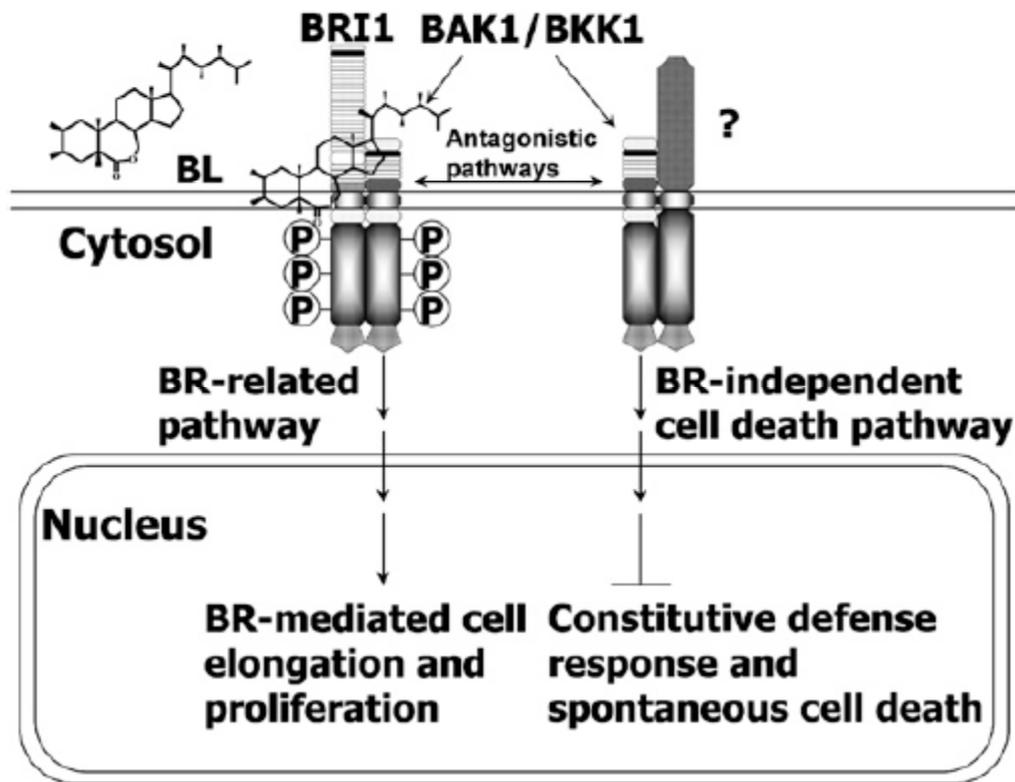
### **1.3.2 Two examples of the plant LRR-RLKs having dual function**

#### **1.3.2.1 BAK1**

Brassinosteroids (BRs) play crucial roles in plant growth and development. Plants defective in BR biosynthesis or BR perception exhibit typical phenotypes including dwarfed stature, darker green round leaves, male fertility, delayed flowering and senescence, and photomorphogenetic defects (Li *et al.*, 2002; Nam and Li, 2002). BRASSINOSTEROID-INSENSITIVE1 (BRI1) that encodes LRR-RLK binds BR in an extracellular domain (Li and Chory, 1997; Wang *et al.*, 2001). SERK3/BAK1 is involved in the BRI1 mediated BR signaling pathway forming a heterodimer with BRI1 on the cell membrane (Nam and Li, 2002). *Bak1* mutant plants displayed compact rosettes with round leaves and short petioles and reduced stem and hypocotyls lengths and are weaker than the *bri1* phenotype. BAK1 possesses 5 LRRs and belongs to the LRR type II subfamily which contains 14 members, five of which were named SOMATIC EMBRYOGENESIS RECEPTOR KINASE 1 (SERK1) to SERK5 (Shiu and Bleecker, 2001b; Nam and Li, 2002). The *Arabidopsis* genome encodes 13 BAK1 homologs. Most probably because of the redundancy between BAK1 and its homologs, *bak1* plants might display weaker phenotypes. BRI1 and BAK1, which exist as inactive monomers, can form ligand independent heterodimers on the cell surface with the interaction of extra- and intracellular domains (Nam and Li, 2002). Heterodimer formation results in transphosphorylation of specific Ser/Thr residues of the partners triggering the signaling cascade.

In addition to the BRI1 and BAK1 interaction, recent studies revealed that BAK1 can also form a heterodimer with FLAGELLIN SENSING2 (FLS2), an LRR-RLK (Chinchilla *et al.*, 2007; Heese *et al.*, 2007). Mutant plants possessing *bak1* mutations lacked an early and late flagellin triggered response indicating that BAK1 is a positive regulator of the flagellin signaling. The formation of a complex between BAK1 and FLS2 takes place in a ligand dependent manner. Interestingly, BAK1 is classified as RD class RLK since it contains an arginine and aspartic acid motif in the catalytic site located in the kinase domain. However, FLS2 is a non-RD class RLK. Therefore, BAK1 most probably has a role as a co-receptor to regulate the different receptors.

The *bak1* null mutant only shows a subtle phenotype suggesting that functionally redundant proteins might be present in the *Arabidopsis* genome (Kemmerling *et al.*, 2007; He *et al.*, 2007; Ingram, 2007). Studies revealed that SERK4/BAK1-LIKE1 (BKK1), which is the closest paralog of BAK1, functions redundantly with BAK1 because overexpressed SERK4/BKK1 partially rescues the *bri1-5* phenotype. However, the double mutants of *bak1 bkk1* do not display the expected *bri1*-like phenotype. The double mutant plants show unexpected post-embryonic seedling lethality due to spontaneous cell death. Transcriptome analysis of the double mutant exhibited increased expression of defense and senescence related genes. Interestingly, analyses revealed that BAK1 and BKK1 have dual roles in two antagonistic signaling pathways (He *et al.*, 2007) (Figure 1.2). These two proteins not only positively regulate a BR-dependent pathway, but also negatively regulate a BR-independent cell death pathway. It was shown that BAK1 and BKK1 negatively regulate the microbial-infection-induced cell death pathway via regulation of genes involved in defense response (Kemmerling *et al.*, 2007; He *et al.*, 2007).



**Figure 1.2:** A proposed model showing that BAK1 and BKK1 positively regulate a BR signaling pathway and negatively regulate a spontaneous cell-death pathway (He *et al.*, 2007).

### 1.3.2.2 ERECTA

The cells, during organ primordia formation, have to coordinate their growth to develop organs with correct size and shape. Therefore, cell-cell communication is crucial for the determination of cell size and shape. One of the most known gene having a function during organ development is ERECTA (ER). *Erecta* mutant plants display compact inflorescence, short internodes, round leaves with short petioles, and pedicels, short and blunt siliques, reduced plant height and smaller organ size (Torii *et al.*, 1996). *ERECTA* encodes a 976 amino acids LRR-RLK which comprises 20 LRRs. Studies revealed that ERECTA regulates the number of the cells within the ground tissues, which specifies the size and shape of the mature organs (Torii *et al.*, 1996; Shpak *et al.*, 2004). In the

*Arabidopsis* genome *ERECTA* belongs to the LRR-XIII family consisting of 7 members. *ERECTA* has two functional paralogs, *ERECTA-LIKE1* (*ERL1*) and *ERL2*, playing redundant roles in ER signaling pathway (Shpak *et al.*, 2004). Although *erl1* and *erl2* single mutant and double mutant plants displayed no obvious phenotype, each single mutant in *er* background enhanced the *erecta* phenotype. Moreover, loss of all *ERECTA* family members caused extreme dwarfism, small-misshaped cotyledons, diminished primary leaves and abnormal flower development. *ER* is highly expressed in shoot meristems, developing organ primordia, flowers and siliques and young rosette leaves whereas expression is low in mature shoot organs, such as stem, rosette leaves and expression in roots was barely observed (Torii *et al.*, 1996; Yokoyama *et al.*, 1998). At the early developmental stages all three genes are expressed in an overlapping manner, while at the later stages their expression become restricted to different proliferating tissues. Although *ERL1* and *ERL2* could substitute *ER* function when they are expressed under the *ER* promoter, expression analysis revealed that *ERL1* and *ERL2* display overlapping and unique expression patterns. Therefore, the roles of *ERL1* and *ERL2* are parallel to the ER signaling pathway which regulates the cell proliferation along the longitudinal dimension of organ growth (Shpak *et al.*, 2004). In addition to these functions Shpak and coworkers (2005) showed that ER-family members have a function during stomatal patterning. During this function there is a complex interaction between TOO MANY MOUTHS (TMM), a receptor-like protein, and ER-family members. TMM negatively regulates the specific ER-family members during stomatal differentiation.

Llorente and coworkers (2005) showed that *ERECTA* is required for resistance to the necrotrophic fungus *Plectosphaerella cucumerina*. Results indicated that especially LRRs and kinase domains of ER possess roles.

*BAK1* and *ERECTA* are two examples of the LRR-RLKs having dual roles in plant development and immunity. Interestingly, *ERECTA* and *BAK1* are also the examples of the LRR-RLKs acting redundantly.

#### **1.4 Ligand specificity of the receptor-like kinases**

The identification of the ligands of LRR-RLKs helps to better understand the signal transduction of the LRR-RLKs (Torii, 2004). However, until so far, only a

few ligands of the RLKs have been identified. Interestingly, identified ligands belong to different protein families.

CLV3, which is a member of the CLE (CLV3/ESR) family of putative polypeptide signaling molecules, is a ligand candidate of the CLV1/CLV2 heterodimer, (Fletcher *et al.*, 1999; Cock *et al.*, 2002; Torii, 2004). Recently, Ogawa and coworkers (2008) have shown that CLV3 can bind to the extracellular domain of CLV1. The maize ESR (Embryo Surrounding Region) proteins play roles as a signal between endosperm and embryo during early maize development. However, their receptors have not been determined. The *Arabidopsis* genome consists of 25 CLE genes expressed in various tissues as probable putative ligand candidates for LRR-RLKs. (Torii, 2004).

Another example for the ligand molecules is Phytosulfokine (PSK) family members one of which was cloned from a rice cell culture cDNA library (Yang *et al.*, 1999). *Arabidopsis* genome has four PSK-like genes. PSK binds to the PSK receptor, PSKR1, a leucine-rich repeat receptor-like kinase (Matsubayashi *et al.*, 2002). Another example of tyrosine-sulfated peptide is PSY1, a 5-amino acids tyrosine sulfated glycopeptide (Amano *et al.*, 2007). Interestingly, PSY1 is a putative ligand of At1g72300, an LRR-RLK, which promotes cellular proliferation.

Cysteine-rich extracellular proteins are other interesting examples of the LRR-RLK ligands. LAT52, a cys-rich protein, is a ligand of the *Petunia* LRR-RLK LePRK2 but not for LePRK3, a close homolog of LePRK2, indicating a ligand specificity for each LRR-RLKs (Muschietti *et al.*, 1994). S-locus cys-rich protein (SCR/SP11) is a functional ligand for S-LOCUS RECEPTOR KINASE (SRK) conferring a self-incompatible phenotype to pollen (Schopfer *et al.*, 1999). AtGRP-3, a glycine-rich extracellular protein, interacts with WALL ASSOCIATED KINASE (WAK) and acts as a ligand for WAK1 (Park *et al.*, 2001).

Hormones can also be a ligand of LRR-RLKs. For instance, brassinolide (brassinosteroids) signal through BRI1 and BAK1 (Wang *et al.*, 2001).

There are also some proteins secreted by fungus which function as a ligand. *C. fulvum* Avr9 and Avr4 proteins are small cys-rich secreted proteins detected by tomato disease-resistance proteins Cf9 and Cf4 (Jones and Jones, 1997). The bacterial flagellin and elongation factor EF-Tu, signaling in the

immunity system, are ligands for FLS2 and EFR, respectively (Gomez-Gomez *et al.*, 2001; Zipfel *et al.*, 2006).

## **1.5 STRUBBELIG RECEPTOR FAMILY (SRF) gene family**

### **1.5.1 STRUBBELIG (SUB), a leucine-rich repeat receptor-like kinase involved in plant development**

The *Arabidopsis* gene *STRUBBELIG* (*SUB*) was first identified based on female sterility during an EMS mutant screen (Schneitz *et al.*, 1997). The *sub* phenotype is due to an altered proliferation of the ovule outer integument responsible for the formation of severely defected ovules. The outer integument phenotype is visualized in the late stage 2-III/stage2-IV of ovule development. In addition to the ovule phenotype, *sub* plants display a reduced plant height compared to wild-type plants, and show twisted stems, carpel and petal pedicels (Chevalier *et al.*, 2005). At the cellular level, the L2 layer of stage-3 floral meristem displays an elevated frequency of periclinal cell divisions. In addition, 30-day old *sub* stems show reduced numbers of epidermal, cortex, and pith cells (Chevalier *et al.*, 2005). Furthermore, Kwak and coworkers (2005) showed that SUB/SCRAMBLED (SCM) plays a role in the specification of root epidermis cells. It has been also shown that *SUB/SCM* negatively regulates the expression of *WEREWOLF* (*WER*) (Kwak and Schiefelbein, 2007). Taken together, the *sub* phenotype suggests that *SUB* has functions in the orientation of the cell division plane and in the control of cell number, cell size and cell shape.

### **1.5.2 STRUBBELIG RECEPTOR FAMILY (SRF) members**

*SUB* belongs to the *LRR-V* gene family encoding putative LRR-RLKs (Shiu and Bleecker, 2001b; Shiu and Bleecker, 2003). This monophyletic family is represented by *SUB* and eight additional members. These eight homologous genes were named as *STRUBBELIG RECEPTOR FAMILY* (*SRF1-8*). However, the name of *SUB* remained same since *STRUBBELIG* is the first identified gene among the family members (Eyüboğlu *et al.*, 2007). Orthologs of the SRF family are also known. Three *LEUCINE RICH REPEAT TRANSMEMBRANE PROTEIN KINASE1-3* (*LTK1-3*) from *Zea mays* and six genes called *STRUBBELIG RELATED SEQUENCE1-6* (*SRS1-6*) from *Oryza sativa* also belong to the *LRR-V* subclass (Li and Wurtzel, 1998; Matsumoto *et al.*, 2001;

Goff *et al*, 2002). Although the function of LTK family members is unknown, LTK1 is ubiquitously expressed whereas LTK2 and LTK3 are specifically expressed in endosperm (Li and Wurtzel, 1998).

All members of SRF family consist of an N-terminal signal peptide (SP), an extracellular SUB domain, six leucine-rich repeats (LRRs), a transmembrane domain (TM), an intracellular juxtamembrane domain (JM), and an intracellular kinase domain (KD). In addition to these domains, certain SRF members possess a proline-rich region, PEST sequence, and a C-terminal extension.

## **1.6 The scope of the study**

The aim of the study is to understand the structure and function of the SRF family members.

### **1.6.1 In this study**

Although RLKs represent one of the largest gene families in the *Arabidopsis* genome, as less as 2% of the RLKs functions have been identified and among these RLKs much less is known about their signaling components. Thus, it is an important task to gain information about the function of the remaining RLKs.

In this study, we are interested in the *LRR-V/STRUBBELIG RECEPTOR FAMILY (SRF)* gene family encoding LRR-RLKs. To gain information about the function of the SRF members, full-length sequence information of the genes is necessary. Therefore, we performed cloning and structural analysis of the complete *SRF* gene family (Chapter 3). We outlined the sequence similarities and differences among family members and also between the RLKs with known functions. This approach allowed us to gain more hints about the probable function of the SRF family members. In addition to this approach, we analyzed T-DNA insertion lines and *35S::SRF* lines of the *SRF* members in order to understand the function of the genes (Chapter 4). Moreover, yeast-two hybrid approaches were used to identify the additional components of the SRF signaling pathways (Chapter 4).

## Chapter 2. Materials and Methods

### 2.1 Materials

#### 2.1.1 Chemicals, reagents, and media

##### 2.1.1.1 Chemicals

All chemicals used in this study were purchased from Sigma (USA), Fluka (USA), and Roth (Germany). Solutions were prepared according to Sambrook *et al.* (1989), unless otherwise described in the text.

##### 2.1.1.2 Restriction endonucleases

Restriction endonucleases were purchased from either New England Biolabs (NEB, USA) or Roche (Mannheim, Germany).

##### 2.1.1.3 Polymerases

For cloning purposes *Pfu* ultra high fidelity polymerase (Stratagene, USA) was used. Advantage cDNA polymerase mix enzyme was used for RACE-PCR (Clontech, USA). *Taq* DNA polymerase (MBI Fermentas, Germany) was used for genotyping.

##### 2.1.1.4 Bacteria and yeast growth media

Luria-Bertani (LB) broth or LB agar plates (1.5%) were used to culture the bacterial strains. Media were prepared according to Sambrook *et al.* (1989). For yeast growth Synthetic Complete (SC) media was prepared as described in Adams *et al.* (1997).

##### 2.1.1.5 Plant growth media

*Arabidopsis* seeds were germinated on Murashige and Skoog (MS) medium containing 1% sucrose, and 0.9% agar (Murashige and Skoog, 1962). MS agar with suitable antibiotics was used for the selection of the T1, T2, and T3 lines of the *35S::SRF1-8* and *35S::SRF4:EGFP* transgenic plants.

## 2.1.2 Microorganisms and plant materials used in the study

### 2.1.2.1 *Escherichia coli* strains and vectors

*E. coli* strain DH5 $\alpha$  (*supE44*  $\Delta$ *lacU169* (*F*  $\phi$ 80 *lacZ* $\Delta$ *M15*) *hsdR17* *recA1* *endA1* *gyrA96* *thi-1* *relA1*) was used for standard DNA manipulations. For cloning of the full-length *SRF1-8* cDNAs, the pCR II-TOPO vector (Stratagene, USA) and pGEM-T-Easy vector (Promega, USA) were used. Both vectors encode a kanamycin (50 $\mu$ g/ml) and an ampicillin (50 $\mu$ g/ml) resistance selectable marker, respectively.

### 2.1.2.2 *Agrobacterium* strains and vectors

*Agrobacterium tumefaciens* strain GV3101/pMP90 (Konez and Schell, 1986) was used for all plant transformation experiments. pMP90 encodes the selectable marker for gentamycin resistance (50 $\mu$ g/ml). *35S::SRF1-8* and *35S::SRF4:EGFP* constructs are derivatives of the modified pCAMBIA2300 binary vector (<http://www.cambia.org>), and pGREEN0029 vector (<http://www.pgreen.ac.uk>), respectively, which both encode a kanamycin resistance selectable marker (50 $\mu$ g/ml).

### 2.1.2.3 Yeast strains and vectors

For the yeast two-hybrid screen, the PJ69-4A (*MATa*, *trp1-901*, *leu2-3, 112*, *ura3-52*, *his3-200*, *gal4* $\Delta$ , *gal80* $\Delta$ , *LYS2::GALI-HIS3* *GAL2-ADE2* *met2::GAL7-lacZ*) yeast strain, maintained using standard conditions, was utilized as a host (Adams *et al.*, 1997; James *et al.*, 1996). pAS2-attr vector was used to prepare *SRF4*, *SRF5*, and *SRF6* intracellular construct to use as bait for yeast two-hybrid screening assay.

### 2.1.2.4 Plant lines

*Arabidopsis thaliana* (L.) Heynh var. Columbia (Col-0) and var. Landsberg (*erecta* mutant) (*Ler*) were used as wild type strains. The *sub-1* mutant was isolated in an ethyl methane sulfonate (EMS) mutagenesis in a *Ler* background (Schneitz *et al.*, 1997; Chevalier *et al.*, 2005).

## **2.2 Methods**

### **2.2.1 Plant work**

#### **2.2.1.1 Plant growth conditions**

Plants were grown in greenhouse under Philips SON-T Plus 400-W fluorescent bulbs in 16-h light/8-h dark cycles. Dry seeds treated with cold for four days were sown on soil (Patzer Einheitserde, extra-gesiebt, Typ T, Patzer GmbH & Co. KG, Sinntal-Jossa, Germany) situated above a layer of perlite and then placed in the greenhouse. In order to support equal germination and to provide humidity, all plants were kept under a lid for one week.

#### **2.2.1.2 Plant Transformation**

All *35S::SRF1-8* gene constructs based on the modified pCAMBIA2300 vector were transformed into the *Agrobacterium tumefaciens* GV3101 strain and then into the 4-week old Col wild-type, *Ler*, and *sub-1* background plants by means of the floral dip method (Clough and Bent, 1998). The same method was also used for the transformation of the *35S::SRF4:EGFP* construct.

#### **2.2.1.3 Seed sterilization**

For the seed surface sterilization, seeds were treated with 70% ethanol for 1 minute, then put for 10 minutes in 3% sodium hypochlorite (NaOCl) and finally rinsed with sterile distilled water for five times. Seeds were suspended in 0.1% sterile agarose.

### **2.2.2 Nucleic acid purification**

All molecular techniques, unless otherwise stated in the text, were carried out according to Sambrook *et al.* (1989).

#### **2.2.2.1 Bacterial plasmid isolation**

Plasmid DNA isolation from bacteria was performed by using the alkaline-lysis method (Sambrook *et al.*, 1989). High-purity DNA was isolated by using QIAprep spin miniprep kit according to the manufacturer's instruction (Qiagen, Germany).

### **2.2.2.2 Yeast plasmid isolation**

Yeast cells were grown in 10 ml Synthetic Complete (SC) media without leucine for overnight at 30°C. Cells centrifuged at 2,000 rpm were resuspended in 1 M sorbitol and 0.1 M Na<sub>2</sub>EDTA (pH 7.5). Cells were incubated with Zymolase 100T (2.5mg/ml) for 2 hours at 37°C. Then, lysed cells were centrifuged at 13,000 g for 1 min and resuspended in 50 mM Tris-HCl (pH 7.4) and 20 mM Na<sub>2</sub>EDTA. After addition of 10% SDS, suspension was incubated for 30 minutes at 65°C with agitation. Lysis was stopped with the treatment of 200 µl of 5 M potassium acetate for 1 hour on ice. DNA was precipitated by using isopropanol and 3 M sodium acetate in two steps. The pellet was resuspended in 20 µl of water and then, transformed to the *E. coli* DH5α electro-competent cells.

### **2.2.2.3 Arabidopsis genomic DNA extraction**

*Arabidopsis* genomic DNA was isolated from rosette leaves and stage 1-12 flowers (Smyth *et al.*, 1990) of 30-day old Col-0 or Ler plants according to Murray and Thompson (1980).

### **2.2.2.4 Arabidopsis RNA extraction**

For RACE, the Marathon cDNA amplification kit (Clontech, USA) and poly (A)+ RNA (Col) from flowers of stage 1-12 (Smyth *et al.*, 1990) (*SRF1A/B* to *SRF5*, *SRF7-8*) or rosette leaves (*SRF6*) were used and all steps were performed according to the manufacturer's recommendations.

Total RNA was extracted from stage 1-12 flowers (stage according to Smyth *et al.*, 1990) and rosette leaves, harvested from 30-day old Col plants and extracted by the hot phenol RNA isolation method (Verwoerd, 1989) and the same tissue types were used to create the cDNA template of the RACE-PCR.

### **2.2.2.5 DNA purification from agarose gel**

DNA fragments were cut from gel and purified by using QIAquick gel extraction kit according to the manufacturer's instruction (Qiagen, Germany).

## 2.2.3 Polymerase chain reaction (PCR) methods

### 2.2.3.1 Oligonucleotides

All primers used for RACE-PCR, RT-PCR, genotyping, PCR amplification, sequencing were designed using the Vector NTI suite 7.1 computer program (Invitrogen, USA). All primers used in the study were represented in relevant Methods sections.

### 2.2.3.2 Rapid Amplification Of cDNA Ends (RACE)

Rapid amplification of cDNA ends is a technique to obtain 5' and 3' ends of the cDNA by PCR (5' RACE-PCR and 3' RACE-PCR, respectively). A specific cDNA is amplified by RACE-PCR using a gene specific primer (GSP) that anneals to the region of known sequence and an adaptor primer (AP) that targets to Marathon cDNA adaptor located at the end of the cDNA.

To clone *SRF2-6* and *SRF8* two steps of RACE-PCR were made. Used RACE-PCR primers are shown in Table 2.1. All genes specific primers were designed approximately 100 to 200 basepairs (bp) before and after the transmembrane domain of the *SRF* genes. Using GSP1 having a  $T_m$  more than 65°C and AP1 (5'-CCATCCTAATACGACTCACTATAGGGC-3'), RACE-PCR reactions were performed. During the first-round thermal cycling, the GSP is extended to the end of the adaptor where it creates the AP1 binding site at the 5' or 3' end of the cDNA. PCR conditions were: 94°C for 30 sec, 5 cycles of 94°C for 5 sec and 72°C 3 min, 5 cycles of 94°C 5 sec and 70°C 3 min, followed by 25 cycles of 94°C 5 sec and 68°C 3 min. To increase the specificity of the RACE product a second-round of PCR was performed with GSP2 and AP2 (5'-ACTCACTATAGGGCTCGAGCGGC-3') primers and 1:50 diluted first-round amplification product was used as a template.

For *SRF1A*, *SRF1B* and *SRF7*, only 5' RACE-PCR was carried out. A certain part of the *SRF1A* cDNA including 3' UTR generated by Dr. David Chevalier was available. For *SRF7* 3' UTR sequence information was obtained from sequenced RAFL-09-89-K12 EST clone to generate the full-length *SRF7* cDNA clone.

To find out the correct length of the cDNA, the inserts were released with proper restriction endonucleases from the clones isolated with the alkaline lysis method (Sambrook *et al.*, 1989). For this purpose, at least 60 clones for 5'

RACE and 30 clones for 3' RACE were examined. Selected inserts for 3' and 5' RACE-PCR cDNA clones of each gene were sequenced on both strands.

**Table 2.1: Primers used for RACE-PCR of *SRF* genes**

<b>Gene name</b>	<b>Primer name</b>	<b>Primer sequence</b>
<i>SRF1A-B</i>	SRS15'RACE R3	5'-CCCGATGTCTTTCAAGACCTTGGCC-3'
	SRS15'RACE R1	5'-TGGAGGAAGTATCCCATCCACGTTGC C-3'
<i>SRF2</i>	SRS2 3'race F3	5'-ACTGGCTATCCAACAACAGAGTCAAGTG CC-3'
	SRS2 3'race F4	5'-GAAGAAAGGCATAGGAGCAGGAAGTAC C-3'
	SRS25'raceR1	5'-TGGAGGAAGTATCCCATCCACGTTGC C-3'
	SRS25'raceR2	5'-CTCAGCTGCAGAGCTTCAACTGGCAA C-3'
<i>SRF3</i>	SRS33'RACE F1	5'-ACCTCGACTGCACCATCATTGTCTC-3'
	SRS33'RACE F2	5'-GGGAAAGTTGCTGATGGACCTTCTG-3'
	SRS35'RACE R1	5'-CCTCCTCCTTTGGATGAGAAGGTCA-3'
	SRS35'RACE R2	5'-TGCCACCACCTGGTCGATCTGAAAA-3'
<i>SRF4</i>	SRS43'RACE F1	5'-AAACTGGAGGAAACAAGTGGTCAAGC GG-3'
	SRS4_3'RACEF2	5'-AGGAGGCGGTGGAGGAAGCAGCAAG-3'
	SRS4_5'RACER1	5'-TCGCCGAATCGCCTTCTTGGTGAAG-3'
	SRS4_5'RACER2	5'-CTGACCGTTCTCCACCTCCCAAGA-3'
	SRS4_5'RACER3	5'-TGCCATATTCCGGTGTCTCTCATGAACG A-3'
	SRS4_5'RACER4	5'-CTCAAAGGCTGGTCTTCAAGCGGAGG-3'
	SRS53'RACE F1	5'-CGTCGAAGACAACCAGTTTGAAGGATG G-3'
	SRS53'RACE F2	5'-CTTCGGGTTTCGAAAGACGGAGGAGG-3'
<i>SRF5</i>	SRS55'RACE R1	5'-TCCGACGGACGTACTIONTTGGCGGTTA-3'
	SRS55'RACE R2	5'-TCGCGTGTGATGTCTTTCACTGATACC G-3'
	SRS6 3'-race-F1	5'-TCCACCACCACCTGGTACACCTCCTA-3'
	SRS6 3'-race-F2	5'-TCCAAGAAATCAGGAATCGGAGCGGG-3'
<i>SRF6</i>	SRS65'RACE R1	5'-TTCGTCTGCTCTCCACATGGCATG-3'
	SRS65' RACE R2	5'-TGCTGTCAAGAAATCCACAGTGGTGG-3'
<i>SRF7</i>	SRS7RACER1	5'-CGAAATCCGGGAACCGTGGAAACC-3'
	SRS7RACER2	5'-ACAATCGGTTACAGGCTGGATCCC-3'
<i>SRF8</i>	SRS8 3'-race-F1	5'-ACACCTGAGGTGCAGGAGCAGAGGG T-3'
	SRS8 3'-race-F2	5'-CCTCTCTGCAAGTTGCAACAAACAGC-3'
	SRS8 5'-race-R1	5'-CATGCTTGAGCTCTTAAGTGGTCGGA-3'
	SRS8 5'-race-R2	5'-CATTTGGATACAGCGCTCCAGAGTTTG C-3'
	SRS85'RACER3	5'-TGGCTCTTGGAACTGCCAAGGCT-3'
	SRS85'RACER4	5'-TGTTCCCTTGGCTGGATACTGCACCG-3'

RACE-PCR fragments for all *SRF* genes were cloned into the PCR II-TOPO vector according to the manufacturer's instructions (Invitrogen, USA). Then 2  $\mu$ l of the reaction mixture were transformed into the *E. coli* TOP10F' chemical competent cells according to the manufacturer's recommendations (Invitrogen, USA).

### **2.2.3.3 Generation of the full-length *SRF* cDNAs**

The generation of the full-length *SRF* cDNAs was obtained by the assembly of the different fragments. For this purpose, four different methods were used.

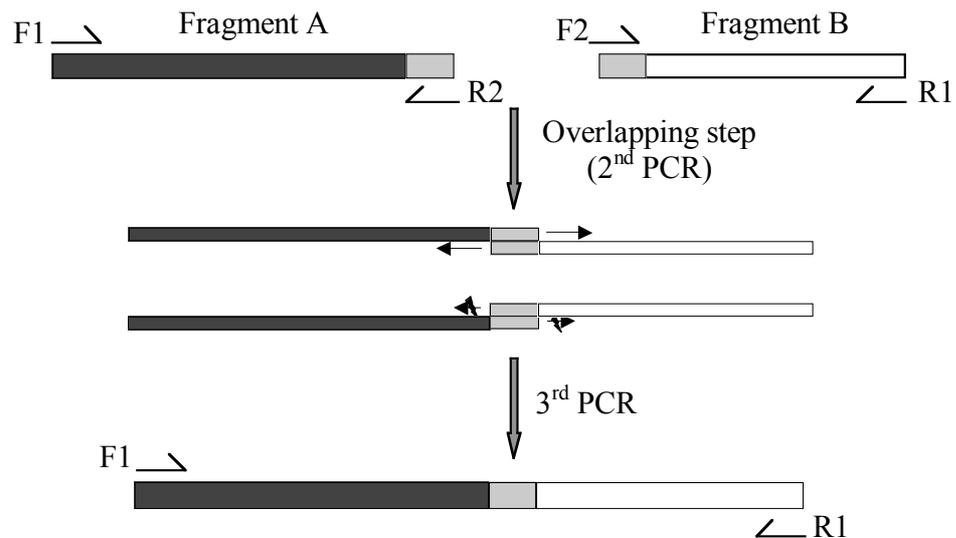
#### **2.2.3.3.1 Restriction digestion and ligation**

5' and 3' cDNA ends of *SRF2*, *SRF5* and *SRF6* were combined by this approach. In this conventional restriction enzyme cloning method, the 5' (*SRF25'\_PCRITOPPO*, *SRF55'\_PCRITOPPO*, and *SRF65'\_PCRITOPPO*) and 3' (*SRF23'\_PCRITOPPO*, *SRF53'\_PCRITOPPO*, and *SRF63'\_PCRITOPPO*) *SRF* cDNA ends to be combined were cut using restriction sites in the overlapping regions and then these fragments fused with ligation reactions. The restriction endonucleases *Clal/ApaI*, *BstBI/Spel*, and *BstXI* were used for *SRF2*, *SRF5*, and *SRF6*, respectively. Afterwards, sequencing was performed to check the ligated region.

#### **2.2.3.3.2 Overlapping PCR**

A modified version of overlapping PCR was carried out to generate full-length cDNAs of *SRF3* and *SRF8*. In this approach, two fragments to be recombined (*SRF35'\_PCRITOPPO/**SRF33'\_PCRITOPPO* for *SRF3* and *SRF85'\_PCRITOPPO/**SRF83'\_PCRITOPPO* for *SRF8*) are generated in separate PCRs by using outer flanking and internal primers (Horton *et al.*, 1989). The internal primers designed from the overlapping part of the two different constructs of the same gene are complementary to each other (Figure 2.1). In a first step, the two fragments to be fused are generated in two separate reactions. As a next step, these PCR products are mixed and used as a template in a short, five cycles PCR without primers. In this PCR, one strand from each fragment containing the overlap sequence can serve as a primer for the 3' extension of the complementary strand. Afterwards, a third PCR is performed by using only flanking primers

allowing the creation of the full-length cDNA. For all PCR reactions, proofreading *Pfu* ultra high fidelity enzyme was used (Stratagene, USA). The primers used for the overlapping PCR are listed in Table 2.2. PCR products of the third step were cloned into PCR II-TOPO vector according to the manufacturer's instructions (Invitrogen, USA).



**Figure 2.1:** Scheme of the overlapping PCR of fragment A and B. The gray box shows the overlapping region of fragments A and B. The oligonucleotide primers are denoted as black arrows. F1/R2 and F2/R1 stand for forward and reverse primers of the fragments A and B, respectively.

#### 2.2.3.3.3 Asymmetric PCR

The asymmetric PCR approach was used to generate full-length *SRF4* cDNA. Asymmetric PCR as a modification of overlapping PCR is a one-step process to combine fragments (Sandhu *et al.*, 1992). For this method four internal primers and two end primers were designed. Internal primers were designed from the overlapping part of the 3' and 5' cDNA ends of the respective *SRF4* fragments to be combined (SRF45'\_new\_PCRIITOPO/SRF45'\_PCRIITOPO/SRF43'\_PCRIITOPO). For PCR, excess amount of end primers (200  $\mu$ M), and trace amount of overlapping primers (20  $\mu$ M) were used. The primers used for asymmetric PCR are listed in Table 2.2.

#### 2.2.3.3.4 End-to-end PCR

To clone *SRF1* and *SRF7* full-length cDNAs, sequence information from 5' and 3' RACE and the EST clones were used. By using end-to-end primers full-length *SRF1* and *SRF7* cDNAs were amplified from adaptor ligated double stranded cDNA of flower tissue and subsequently cloned into PCR II-TOPO vector by using TOPO TA cloning kit (Invitrogen, USA). *Pfu* Turbo polymerase was used for these PCRs (Stratagene, USA). The sequences of *SRF1* and *SRF7* were checked by sequencing. The sequencing revealed the absence of a cytosine at position 1075 in the *SRF1B*\_PCR II TOPO (sequence according to AY518286 GenBank accession number). This nucleotide change was corrected by partially digesting the *SRF1B* construct with *Pml*I and *Spe*I double digestion and ligation with the correct fragment of the *SRF1A* digested with the same enzymes.

#### 2.2.3.3.5 Cloning of *SRF1A/B* gene from *Ler* cDNA

*SRF1A* and *SRF1B* cDNAs were reverse transcribed and amplified using Poly (A)+ RNA of the *Ler* flowers stage 1-12 (stage according to Smyth *et al.*, 1990). To achieve this, *SRF1* gene was cloned as four different fragments, fragment 1-4, into PGEM-T-easy vector (Promega, UK) by using restriction enzyme cloning method to combine all pieces (Table 2.2). To recombine fragment 1 and fragment 2, 3700 bp of *Sty*I fully digested fragment 1 including PGEM-T-easy vector and *Sty*I partially digested 1141 bp of fragment 2 were ligated. *Sac*I double digested 3867 bp of fragment 3 used as a vector and *Sac*I double digested 705 bp of the fragment 4 were recombined. Digested fragment 1 and 2 was located as 3' to 5' direction in the vector. To change the direction of the insert, the full-length insert was cut out by *Not*I restriction sites and religated to the same vector. Direction of the insert was checked with the *Nco*I restriction endonuclease. To obtain full-length *SRF1A* and *SRF1B* digested fragment 1-2 and fragment 3-4 were digested with *Sap*I restriction endonuclease and ligated. After each cloning step, sequence of new clones around ligation region was sequenced in both directions.

**Table 2.2: PCR primers used for the generation of the full-length *SRF* cDNA**

<b>Gene name</b>	<b>Primer name</b>	<b>Primer sequence</b>
<i>SRF1</i>	SRF1_RACE_F	5'-ATCGCATTACACAGTGACGCTCTCTCTC ACTCAC-3'
	SRF1_RACE_R	5'-AGAAAGAAAAAAGATTCAAATTTGAAGTA GCAGT-3'
<i>SRF1 Ler</i>	SRF1_RACE_F	5'-ATCGCATTACACAGTGACGCTCTCTCTC ACTCAC-3'
	SRS15'raceR3	5'-GGCCAAGGTCTTGAAAGACATCGGG-3'
	SRF15'F1	5'-CCATCCCAGAATCTCTTTCCC-5'
	SRF13'-R1	5'-CCACGAGTTGGACAATATTGGC-3'
	SRF13'-F1	5'-GAACAACATCCAAACCGCAC-3'
	SRF13'-R2	5'-GTCGTGAAGCTGTGGGATTGC-3'
	SRF13'-F3	5'-TATCGGGTCAACTCTTAGCGGC-3'
	SRF1_RACE_R	5'-AGAAAGAAAAAAGATTCAAATTTGAAGT AGCAGT-3'
<i>SRF3</i>	SRF3fulllengthXhoIR	5'-CCGCTCGAGTTTTACACAAATAAATATG T-3'
	SRF3fulllengthBamHI	5'-CGCGGATCCCGAAAGTTCATTGAACCTT A-3'
	SRF3overlapF1	5'-CAAAGGGAAAGAACTCTTACATACTAA-3'
<i>SRF4</i>	SRF3overlapR1	5'-TTAGTATGTGAAGAGTTCTTTCCCTTTG-3'
	SRF4fulllengthKpnIF	5'-GGGGTACCCTTCTTTGTTTCCTCTCCTTT G-3'
	SRF4fulllengthNotIR	5'-ATAAGAATGCGGCCGCTACATTTAATTTG TCGAT-3'
	SRF4overlapF1	5'-TTATAATAGCGGTGTCTTCTATAGGTG-3'
	SRF4overlapR1	5'-CTCCTAGTGTGAGAGCCTTGTT-3'
<i>SRF7</i>	SRF4overlapF2	5'-CCAAATCTGCAGCGAATCGT-3'
	SRF4overlapR2	5'-ACGATTCGCTGCAGATTTGG-3'
	SRF7fulllengthSpeIF	5'-GGACTAGTTGGCTCTTGGATTCTTAC-3'
<i>SRF8</i>	SRF7fulllengthSpeIR	5'-TGCTCTAGAGAAAAAGGGCAAATTTTT-3'
	SRF8fulllengthF1	5'-CTCTTCTTCTTCTTCTCTGTCACTGTCTT-3'
	SRF8fulllengthR1	5'-TTATTCAACAACAAGAAGAGCTCTTAA T-3'
	SRF8overlapF1	5'-GATTAAGAAGATTGATAATGCAGCACTTT C-3'
	SRF8overlapR1	5'-GAAAGTGCTGCATTATCAATCTTCTTAAT C-3'

### 2.2.3.3.6 Site-directed Mutagenesis (SDM)

The comparison of the cloned sequences with NCBI annotated gene sequences revealed several differences in nucleotide sequences. The nucleotide changes, which cause an amino acid substitution, were corrected by the site-directed mutagenesis (SDM) approach. Primers were designed by using an online program ([http:// bioinformatics.org/cgi-bin/DNA-1.cgi](http://bioinformatics.org/cgi-bin/DNA-1.cgi)). Table 2.3 shows the

primers used for SDM. *Pfu* Turbo polymerase was employed to amplify the plasmid including the mutagenesis (NEB, USA). Before the transformation of the amplified plasmid into *E. coli* DH5 $\alpha$  chemical competent cells, the methylated plasmid template was digested using *DpnI* restriction endonuclease. Single colonies were selected and both strands of the cDNA were analyzed by DNA sequencing.

**Table 2.3: Primers used to correct *SRF* cDNA by SDM**

<b>Gene name</b>	<b>Primer name</b>	<b>Primer sequence</b>
<i>SRF3</i>	SRS35'M175_F	5'-GTCGGGCTGATTAATCTTGATATATCATCT AACAAATAAGCGGC-3'
	SRS35'M175_R	5'-GCCGCTTATATTGTTAGAATATATCAAGATT AATCAGCCCGAC-3'
<i>SRF4</i>	SRS45'S283V_F	5'-CTAGGAGTTATAAATAGCGGTGTCTTCTATAG GTGGACTC-3'
	SRS45'S283V_R	5'-GAGTCCACCTATAGAAGACACCGCTATTATA ACTCCTAG-3'
<i>SRF6</i>	SRF63'F704S_F	5'-AGTTTAGACCACCAATGTCTGAAGTTGTGCA GGCT-3'
	SRF63'F704S_R	5'-AGCCTGCACAACCTCAGACATTGGTGGTCTA AACT-3'
	SRF65'ST384Q_F	5'-TTCACGAAAACAATTCCATTCAGAGTTCTT CATCAGTTGAGAC-3'
	SRF65'ST384Q_R	5'-GTCTCAACTGATGAAGAACTCTGAATGGAAT TGTTTTTCGTGAA-3'
<i>SRF7</i>	SRF7Y672C_F	5'-ATGTTATCGCCCTTTGCGTCCAGCCGGAG-3'
	SRF7Y672C_R	5'-CTCCGGCTGGACGCAAAGGGCGATAACA T-3'
<i>SRF8</i>	SRF83'G566S_F	5'-TAGGTTTCATTTGGATACAGCGCTCCAGAGTT TGCA-3'
	SRF83'G566S_R	5'-TGCAAACCTCTGGAGCGCTGTATCCAAATGA ACCTA-3'

#### **2.2.3.3.7 Alternative splicing analysis of the *SRF* genes by RT-PCR**

The same tissue types were used to create the cDNA template of the RACE-PCR. 2 $\mu$ g RNA was used to synthesize cDNA by RT-PCR. First the cDNA synthesis was checked by amplifying the glyceraldehyde-3-phosphate dehydrogenase (*GapC*) (Shih *et al.*, 1991) gene and then the putative alternative splicing region of the *SRFs* were amplified with primers shown in Table 2.4.

**Table 2.4: Primers used for alternative splicing analysis of the *SRF* genes**

<b>Gene name</b>	<b>Primer designation</b>	<b>Primer sequence</b>
<i>SRF1A</i> & <i>SRF1B</i>	SRF1_intron_F	5'-GGGAGGGCAAGGGAGGGTT-3'
	SRF1_intron_R	5'-GCTTCGAAGCACCTCCTACTCTCT C-3'
<i>SRF2</i>	SRS23'raceF3	5'-ACTGGCTATCCAACAACAGAGTCAAG TGCC-3'
	SRS25'raceR2	5'-GTTGCCAGTTGAAGCTCTGCAGCTGA G-3'
<i>SRF3</i>	SRS33'raceF2	5'-GGGAAAGTTGCTGATGGACCTTCTG-3'
	SRS35'raceR1	5'-TGACCTTCTCATCCAAAGGAGGAGG-3'
<i>SRF4</i>	SRS43'secondstrandF1	5'-GGATTTGCAGAATACCGCGTC-3'
	SRS43'secondstrandR1	5'-TTTGACACTGGAGGCCTAAGTCC-3'
<i>SRF5</i>	SRS53'raceF2	5'-CTTCGGGTTGAAAGACGGAGGAGG-3'
	SRS55'raceR2	5'-CGGTATCAGTCAAAGACATCACACGCG A-3'
<i>SRF6</i>	SRS63'RACEF2	5'-TCCAAGAAATCAGGAATCGGAGCGG G-3'
	SRS65'RACER2	5'-CCACCACTGTGGATTTCTTGACAGC A-3'
<i>SRF7</i>	SRF7_intron_F1	5'-ATCAAAGCGGTCATCATCCACG-3'
	SRF7_intron_R1	5'-CGCTCAGATGGTGGAGGACGTA-3'
<i>SRF8</i>	SRS83'RACEF2	5'-CCTCTCTGCAAGTTGCAACAAACAG C-3'
	SRS85'RACER4	5'-CGGTGCAGTATCCAGCCAAGGGAA CA-3'
<i>SUB</i>	RK5-F	5'-CCCCTGTGGCATTGGATAC-3'
	RK5-R	5'-CCACCACCATCACCACTCT-3'
<i>GAPC</i>	GapCscDNA	5'-CACTTGAAGGGTGGTGCCAAG-3'
	GapCas	5'-CCTGTTGTCGCCAACGAAGTC-3'

#### 2.2.3.3.8 Semi-quantitative RT-PCR

After the synthesis of cDNA from the respective 35S::*SRF2-5* and 35S::*SRF7* transgenic lines, standard PCR was performed by using 22, 24, 26, 28, and 30 PCR cycles. PCR products were visualized on agarose gel stained with ethidium bromide. Quantification of the bands was done by using Image J program.

#### 2.2.4 Examination of the expressed sequence tags (EST)

Available expressed sequence tags from KAZUSA (Asamizu *et al.*, 2000), RIKEN (Seki *et al.*, 2002), INRA, TAIR and Department of Life Sciences at the Aalborg University, Denmark, were examined to obtain the full-length cDNA

sequences of the *SRF* gene family members. First, bacteria cultures containing EST clones were streaked on plates containing suitable antibiotics and incubated at 37°C. After overnight incubation of the colonies in Luria-Bertani broth (Sambrook *et al.*, 1989), plasmids were isolated from the media by using the QIAprep spin miniprep kit according to the manufacturer's instruction (Qiagen, Germany). Afterwards, all plasmids were digested with suitable restriction endonucleases to release their insert. For more information please check the Table 2.5.

For each gene, the two longest ESTs were sequenced from both the 5' and 3' end using the primers M13 (-20)F (5'-CTGGCCGTCGTTTTAC-3') and M13R (5'-CAGGAAACAGCTATGAC-3'). RIKEN ESTs were sequenced using the primers T7F (5'-AATACGACTCACTATAGGG-3') and T3R (5'-TCCCTTTAGTGAGGGTTAATT-3').

## **2.2.5 Generation of expression and yeast two-hybrid constructs**

### **2.2.5.1 Overexpression constructs**

The open reading frame (ORF) of the *SRF* genes was amplified from corresponding full-length *SRF* cDNA clones by PCR. PCR amplified *SRF* cDNAs were cloned in sense orientation into a modified version of the plant transformation vector pCAMBIA2300 downstream of the CaMV 35S promoter (Dr. Ram Kishor Yadav, personal communication). The modified vector contains the three-copy myc tag, which allows the generation of *SRF* proteins that are tagged with 3xmyc at their carboxy ends. PCR fragments were cloned into the vector by using *Ascl* and *AatII* restriction endonucleases except for *SRF2* and *SRF5* for which *Ascl* and *Apal* were employed. All primers used for overexpression constructs are shown in Table 2.6. One cloned PCR product for each construct was sequenced in both directions using primers at about 500 bp intervals.

**Table 2.5: ESTs of SRF family members**

<b>Gene name</b>	<b>Accession number</b>	<b>Clone name</b>	<b>Vector name</b>	<b>Restriction site</b>	<b>Insert length (bp)</b>
<i>SRF1</i> <i>At2g20850</i> 2794 bp	AY056176	RAFL08-11-K05	λFLC-1-B	<i>SfiI</i>	<b>2500</b>
	AV530585	FB005h02	PBluescript II SK	<i>EcoRI/XhoI</i>	<b>650</b>
<i>SRF3</i> <i>At4g03390</i> 3037 bp	R30153 T43690	166I18T7 122J12T7	λ Zip-Lox λ Zip-Lox	<i>Sall/NotI</i> <i>Sall/NotI</i>	<b>450</b> <b>900</b>
	AV548626	RZL58c11F*	PBluescript II SK	<i>EcoRI/XhoI</i>	<b>1300</b>
	AV535313	FB094c11F	PBluescript II SK	<i>EcoRI/XhoI</i>	<b>450</b>
	AV529313	APZL33d07	PBluescript II SK	<i>EcoRI/XhoI</i>	900
	AV534108	FB074a11	PBluescript II SK	<i>EcoRI/XhoI</i>	550
	AV539725	RZ138f08F*	PBluescript II SK	<i>EcoRI/XhoI</i>	<b>1500</b>
<i>SRF4</i> <i>At3g13065</i> 2639 bp	AV565454 BU636210 BU635370 BU636567 AK118729	SQ224d11 005D05 021C05* 008A11* RAFL21-08-A16	PBluescript II SK PBluescript SK+ PBluescript SK+ PBluescript SK+ λFLC-1-B	<i>EcoRI/XhoI</i> <i>EcoRI/XhoI</i> <i>EcoRI/XhoI</i> <i>EcoRI/XhoI</i> <i>SfiI</i>	- 1300 <b>800</b> <b>1500</b> <b>1400</b>
	AY037178	RAFL-09-N01	λFLC-1-B	<i>SfiI</i>	<b>1700</b>
<i>SRF5</i> <i>At1g78980</i> 2492 bp	N963387 T22249	G8F12T7* 103D18T7*	PBluescript SK- pZL 1	<i>EcoRI/NotI</i> <i>Sall/NotI</i>	<b>2350</b> <b>1200</b>
<i>SRF6</i> <i>At1g53730</i> 2472 bp	AV554479 AV548397 F14013	RZ92D03 RZL53e03 ATTS4935	PBluescript II SK PBluescript II SK PHD-1	<i>EcoRI/XhoI</i> <i>EcoRI/XhoI</i> <i>NotI/NotI</i>	<b>850</b> 800 <b>1031</b>
	AV441019	APZ21g09	PBluescript II SK	<i>EcoRI/XhoI</i>	<b>1250</b>
	AV522124	APZ73b01	PBluescript II SK	<i>EcoRI/XhoI</i>	550
	AV523235	APZL21c11	PBluescript II SK	<i>EcoRI/XhoI</i>	2250
<i>SRF7</i> <i>At3g14350</i> 2446 bp	AV529118 AV535628 AV543478 T04250 AV561214 AY072139	APZL27g08 FB099d08 RZ201c07 34C8T7 SQ147f04F RAFL-09-89-K12	PBluescript II SK PBluescript II SK PBluescript II SK pZL 1 PBluescript II SK λFLC-1-B	<i>EcoRI/XhoI</i> <i>EcoRI/XhoI</i> <i>EcoRI/XhoI</i> <i>Sall/NotI</i> <i>EcoRI/XhoI</i> <i>SfiI</i>	<b>2300</b> 500 650 700 <b>550</b> <b>2500</b>
<i>SRF8</i> <i>At4g22130</i> 2567 bp	AV558045 AV561187	SQ087d02 SQ147b08F	PBluescript II SK PBluescript II SK	<i>EcoRI/XhoI</i> <i>EcoRI/XhoI</i>	<b>650</b> <b>600</b>

\*indicates a chimeric EST . ESTs which were analyzed by sequencing are written in bold.

### 2.2.5.2 Generation of a 35S::*SRF4:EGFP* construct

To generate the 35S::*SRF4:EGFP* construct the 35S::*SUB:EGFP\_pEZT-NT* clone prepared by Dr. Ram Kishor Yadav was used. The 35S::*SUB:EGFP* construct including the 3' OCS termination located between the nucleotide positions 1898 to 6942 of the pEZT-NT vector were digested with *NotI* and *SpeI*

restriction enzymes and ligated into the high copy PGREEN0029 vector using the same restriction sites. Then this construct was used to obtain the *35S::SRF4:EGFP* construct. To this end, the *35S::SRF4\_pCAMBIAxmyc* clone was digested with *Ascl* and *AatII* restriction endonuclease to release the *SRF4* ORF and ligated to *Ascl* and *AatII* digested *35S::SUB:EGFP\_PG*GREEN0029 construct lacking SUB.

**Table 2.6: Primers used for overexpression constructs of *SRF1-8***

Gene name	Primer designation	Primer sequence
<i>SRF1A-B</i>	35S_SRF1_AscIF	5'-AGGCGCGCCGGAATGAGATCGATGAGATCTGGGAGAG-3'
<i>SRF1A</i>	35S_SRF1A_AatIIR	5'-CTGACGTCAGGGCCCCCTCTCTCTTTGGTCTAGCCTCCTT-3'
<i>SRF1B</i>	35S_SRF1B_AatIIR	5'-CTGACGTCAGGGCCCCCTTTCTTCCCGTATACTGATTATC-3'
<i>SRF2</i>	35S_SRF2_AscIF	5'-AGGCGCGCCGTGATGAAAACCAAACAGCAATTGCGATTCTCGC-3'
	35S_SRF2_ApaIR	5'-CTGGGCCCCAGAGGAGAGGTAGCTGAAGGTAGGT-3'
<i>SRF3</i>	35S_SRF3_AscIF	5'-AGGCGCGCCACAATGGCTGCTAAGAGATCTAATCTAC-3'
	35S_SRF3_AatIIR	5'-AGGCGCGCCACAATGGCTGCTAAGAGATCTATCTAC-3'
<i>SRF4</i>	35S_SRF4_AscIF	5'-AGGCGCGCCATGGGACCAAATCTGCAGCGAATCGTA-3'
	35S_SRF4_AatIIR	5'-CTGACGTCAGGGCCCCCACTAGCCTCTTCAATGCCTCCACCA-3'
<i>SRF5</i>	35S_SRF5_AscIF	5'-AGGCGCGCCAAGATGACGCAGAAGCTCGTACGCCTTGTGATC-3'
	35S_SRF5_ApaIR	5'-CTGGGCCCCGTAATCATAATCGTCGTGGGCTCGA-3'
<i>SRF6</i>	35S_SRF6_AscIF	5'-AGGCGCGCCAAGATGAGGGAGAATTGGCGGTGGTGG-3'
	35S_SRF6_AatIIR	5'-CTGACGTCAGGGCCCCCATGTAATCACTGGTCGTGTCGGC-3'
<i>SRF7</i>	35S_SRF7_AscIF	5'-AGGCGCGCCAAGATGACTGAGAATCGGGTGGT-3'
	35S_SRF7_AatIIR	5'-CTGACGTCAGGGCCCCCATGTAATCATGACGCCGGAGC-3'
<i>SRF8</i>	35S_SRF8_AscIF	5'-AGGCGCGCCATGGCTATTGGAGATAGAGCTATGTT-3'
	35S_SRF8_AatIIR	5'-CTGACGTCAGGGCCCCGAATGAGATATCGACGTGCTCGTGTT-3'

### 2.2.5.3 Yeast two-hybrid constructs

For yeast two-hybrid assay, the MATCMaker two-hybrid yeast system 2 (GAL4 2H-2, Clontech) was used. This system is based on the GAL4 trans-activator properties. The intracellular part of *SRF4*, *SRF5* and *SRF6* was fused with the DNA-BD of GAL4 while the full-length cDNAs were fused to the activation domain of GAL4. The yeast two-hybrid screen was carried out in collaboration with Dr. Joachim Uhrig from University of Cologne according to Soellick and Uhrig (2001).

For construction of Y2H bait vectors, the intracellular part of *SRF4* and *SRF5* was cloned by means of the gateway cloning system (Invitrogen, USA). This system constitutes two reactions. At the first cloning step, the intracellular part of *SRF4* (corresponding amino acid residues 301-687) and *SRF5* (corresponding amino acid residues 292-699) was amplified using the SRF4-full\_PCRIITOP and SRF5-full\_PCRIITOP full-length clones as respective templates by means of suitable primers which included the attB recombination site:

SRF4\_Y2H\_gatewayF (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTGAAGAAAGAACTCTAATG-3')

and SRF4\_Y2H\_gatewayR (5'-GGGGACCACTTTGTAC AAGAAAGCTGGGTACTACTAGCCTCTTCA-3') for *SRF4* and SRS53'intraY2HF (5'-GGGGACAAGTTTGTACAAAAAAGCAGGCTCGCTCGTT TCCAAGAAAA-3') and SRS53'intraY2HR (5'-GGGGACCACTTTGTACAAGA AAGCTGGGTTCTAGTAATCATAATCGTCGTGGGCT-3') for *SRF5*.

*Pfu* polymerase was employed for all PCR reactions (Stratagene, USA). For the first reaction, BP clonase recombination reaction, PCR amplified intracellular parts of the *SRF4* and *SRF5*, respectively, including the attB1 and attB2 sites at the end of the amplified product, were purified by PEG precipitation. 100 ng purified PCR product were then incubated with 4 µl 5x BP reaction buffer, 150 ng pDONR207 entry vector, 4 µl BP clonase enzyme mixture. The reaction was adjusted to a total volume of 20µl with 1x TE buffer and incubated at 25°C for overnight. The reaction was stopped by the addition of Proteinase K and incubation for 10 minutes at 37°C, followed by the deactivation of Proteinase K by incubation for 10 minutes at 80°C. Then 3µl reaction mixture was transformed to chemical competent *E. coli* strain DH5α by heat shock. After overnight incubation of transformed competent cells at 37°C, clones were

incubated in liquid LB media and used for plasmid purification by using a alkaline lysis protocol. These new clones, also known as entry clones, were called as SRF4intra\_pDONR207 and SRF5intra\_pDONR207. After sequencing in both strands, a correct clone was used for the next cloning step, the LR clonase reaction, to transfer the DNA fragments from the entry clone to the pAS2-attr destination vector. The LR reaction consisted of 500 ng respective entry vector, *Xma*I digested 200 ng destination vector (pAS2-attr), 4 µl LR reaction buffer, 4 µl LR enzyme mixture and a suitable amount of 1x TE buffer to complete the final reaction volume (16 µl). After overnight incubation at 25°C, reaction was stopped with the addition of Proteinase K. Afterwards, 3µl reaction solutions were transformed to the chemical competent cells. To find out the correct clone, the same procedure to obtain entry clone was performed. PAS2-attr destination vector possesses the GAL4 DNA-BD and TRP1 gene allowing yeast selection on media without tryptophan amino acid. These clones were used as bait vectors in the Y2H assay.

To clone the intracellular part of the *SRF6* gene (corresponding amino acid residues 319 to 719), the SRF6-full\_PCRITOPPO full-length clone was used as a template for PCR by using the primers SRS6intraNdeIF (5'-GGGAATTCCATATGAGGTCATCACCCATGG-3') and SRS6intraBamHIR (5'-CGCGGATCCCTTGTTAGAGGCATTTAC-3') including *Nde*I and *Bam*HI restriction sites, respectively. The digested PCR fragment was then ligated into the compatible *Nde*I and *Bam*HI sites of the digested pAS2-attr vector.

## 2.2.6 PCR screening strategy for T-DNA lines

### 2.2.6.1 T-DNA insertion lines analysis

#### 2.2.6.1.1 Investigation of the *srf4* T-DNA insertion lines

Two *srf4* T-DNA insertion lines, *srf4-2* (Garlic 230E8) and *srf4-3* (Garlic 253A9), were (Eyüboğlu *et al.*, 2007). The insertions of *srf4-2* and *srf4-3* were located in the second and seventh exon, respectively. To confirm the homozygosity of *srf4-2*, PCR was carried out by using the primer for the vector pCSA110 LB3 (5'-TAGCATCTGAATTCATAACCAATCTCGATACAC-3') and for genomic DNA *srs4-2\_for* (5'-GCTGTACATTTCTCTATTCACACGTCCA-3') and *srs4-2\_rev* (5'-CACAGGTAAGTGAGCGATTCCAAGTTTC-3'); for the *srf4-3* insertion line LB3,

srs4-3\_for (5'-AATGTGTTGGTATTGCAGAGATGGCTCT-3') and srs4-3\_rev (5'-TTGACCACTTGTTTCCTCCAGTTCTAAC-3') primers were used for PCR.

Because both insertion lines are found in the *quartet* (*qrt*) mutant background (Preuss *et al.*, 1994; Rhee and Somerville, 1998), *qrt* mutant plants without T-DNA insertions were used as positive control.

#### **2.2.6.1.2 Identification of the T-DNA insertion line of *daeumling* (*At1g79870*)**

The insertion line SAIL\_1289\_A11 was ordered from Arabidopsis Biological Resource Center (ABRC). Surface sterilized seeds were placed on MS agar and treated with cold for four days. After one-week incubation in the growth chamber, all plants were transferred to the soil in greenhouse. The vector used to produce the insertion line was pDAP101. Genomic DNA was extracted from leaves using a modified CTAB method (Murray and Thompson, 1980). The T-DNA was characterized by PCR on genomic DNA by primers flanking the insertion for wild-type N848110\_F (5'-ACCCTTGTCCTGGATTCCCTAAACAT-3') and N848110\_R (5'-GGCACATCATAAGGACTCCGATTGAT-3') primers and a primer derived from the T-DNA left border of the insertion LB1 (5'-GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC-3'). We identified the position of the T-DNA using PCR primers LB1 and N848110\_F.

### **2.2.7 Nucleotide sequencing and bioinformatic analysis**

#### **2.2.7.1 Nucleotide sequencing**

The sequence reaction was performed with the ABI Prism BigDye Terminator v1.1 Cycle Sequencing Ready Reaction Kit according to the manufacturer's instruction and analyzed with an ABI 373 sequencer (PE Applied Biosystems, USA).

#### **2.2.7.2 Sequence analysis**

Sequence data were analyzed and assembled using the computer program ContigExpress (Mac OS X 10.3.9 version) and subjected to similarity search against Col genomic sequence database provided by NCBI (<http://www.ncbi.nlm.nih.gov>) using the basic local alignment search tool (BLAST) algorithm (Altschul *et al.*, 1990).

Several web sites were used to define the SRF structure. The signal peptide domains (SP), leucine-rich repeat domains (LRR), the proline-rich regions (PRR), the transmembrane domains (TM) and PEST motifs were identified using SMART (<http://smart.embl-heidelberg.de>), PROSITE (<http://www.expasy.ch/prosite>), PSORT (<http://psort.ims.u-tokyo.ac.jp/form.html>) and PESTfind (<http://embl.bcc.univie.ac.at/toolbox/pestfind/pestfind-analysis-webtool.htm>) websites, respectively. The kinase domains were determined with PlantsP database (<http://plantsp.genomics.purdue.edu>). The eleven subdomains of the kinase domain were detected according to Hanks and Quinn (1991).

Amino acid sequences of SRF1-8 were aligned with MultAlin using default parameters (Corpet, 1988).

#### **2.2.7.3 Polymorphism analysis**

To find out the polymorphisms of *SRF1* alleles between Col and *Ler* ecotypes of *Arabidopsis thaliana* plants, *Ler* genomic DNA of *SRF1* was sequenced. *Ler* genomic DNA isolated from three-week old *Ler* rosette leaves (Murray and Thompson, 1980) was used as a template to perform PCR with Taq DNA polymerase (MBI Fermentas, Germany). To get rid of the mutations resulting from the PCR reaction, four or five separate PCR amplicons were used to prepare a pool for each sequencing reaction. Sequenced *Ler* genomic sequences were compared with the Col genomic sequences obtained from NCBI database by using BLAST algorithm of NCBI.

Available Monsanto *Ler* genomic sequences and PERLEGEN substitution data from TAIR were examined and compared with Col genomic sequences from NCBI database (Jander *et al.*, 2002; Clark *et al.*, 2007). All comparative analysis was done by using BLAST tool of NCBI.

#### **2.2.7.4 Phylogenetic analysis**

Alignments for phylogenetic tree reconstruction of the *SRF* family members were created with Clustal X as implemented in DAMBE using appropriate gap penalty values (Xia and Xie, 2001). Highly divergent segments were deleted. Therefore, only the LRR, SUB and kinase domains of all SRF members were used. The resulting alignments were used to generate phylogenetic

relationships by maximum likelihood and neighbour-joining using the programs TREEFINDER ([www.treefinder.de](http://www.treefinder.de)) and MEGA 3.0 (Kumar *et al.*, 2004), respectively. The number of synonymous and nonsynonymous substitutions per site was estimated using the Nei-Gojobori method as implemented in MEGA 3.0 (Nei and Gojobori, 1986).

## 2.2.8 Phenotypic analysis

### 2.2.8.1 Phenotypic analysis of the *SRF* overexpressing transgenic lines

The sterilized seeds were displayed on MS agar plates containing suitable antibiotic. After four days of cold treatment, all seeds were grown under continuous white light at 22°C. After one to two weeks growth in the culture room, survived transgenic seedlings were transferred on soil contained in pots and kept in the greenhouse for phenotypic analysis.

Fifty T1 transgenic lines for each *SRF* gene in each ecotype were examined. To confirm the expression of the transgene in kanamycin-resistant transgenic plants RT-PCR was performed. Total RNA was extracted from confirmed transgenic plants and cDNA was produced from at least 5 transgenic plants for each line and each ecotype. The RT-PCRs were performed by using gene specific primers and myc tag specific primer. Table 2.7 presents the primers used for transgene analysis for all *SRF* genes.

**Table 2.7: RT-PCR primers of *SRF* members overexpression lines**

Gene name	Primer designation	Primer sequence
<i>SRF1A</i>	SRF13'_F5	5'-GTGCATGTCTCAGACTGCGGCTTGGCTC-3'
<i>SRF1B</i>	SRF15'_F2	5'-GGCCATTTTTTGGACCACCATC-3'
<i>SRF2</i>	SRS23'ssF1	5'-ATTTGCACTCGTCGTTTTGCC-3'
<i>SRF3</i>	SRS33'ssF1	5'-TTGCTAGCCGCTTACGGATATG-3'
<i>SRF4</i>	SRS43'ssF1	5'-GGATTTGCAGAATACCGCGTC-3'
<i>SRF5</i>	SRS53'_1F	5'-CCCTTGACTTGAACACAAGAG-3'
<i>SRF6</i>	SRS63'ssF1	5'-CAAACCGATGAAGGGTATAGCG-3'
<i>SRF7</i>	SRS7F3	5'-CAGAGCTGAATCCACACCTCTCA-3'
<i>SRF8</i>	SRS83'1F	5'-CACTAGTGAGATGGGCAACACCG-3'
<i>cmyc</i>	Cmyc_R	5'-TTATTCATTCAAGTCCTCTTCAGAAATGAGCTTTT GCTCC-3'

For phenotypic analysis, the number of rosette leaves, flowering time, stem size, stem thickness, floral organ were examined; furthermore, sepal,

petal, stamen carpel number were determined. The RT-PCR positive T2 lines were also analyzed with the same phenotypic analysis methods. Moreover, at least three RT-PCR positive T2 lines for each *SRF* gene were checked for root and hypocotyl length on MS medium. For hypocotyls length analysis, seeds of the T2 lines were sterilized and put on MS media and then treated with cold for four days. After 8 hours light treatment of the seeds in growth chamber, plates were covered with aluminium foil to provide dark conditions. At the end of 72-hour incubation, plates were scanned and hypocotyls were measured by means of ImageJ program (<http://rbs.info.nih.gov/ij/>). For root analysis, 10-day old root grown on MS media under light condition were used.

For further analysis homozygous lines of the *35S::SRF4*, 3-12 and 1-5, were used. For the homozygosity analysis, RT-PCR positive T3 lines of the *35S::SRF4* plants were checked on the MS-kan plates. 100% viable lines were used for further investigations.

#### **2.2.8.2 Phenotypic analysis of *SRF4* and *DAEUMLING (DLG)***

For all types of phenotypic analyses, the plant organ measurements were performed on Col wild-type, the homozygous T-DNA insertion line of *At1g79870 (dlg-1)*, the homozygous T3 lines of *35S::SRF4* overexpression lines, 3-12 and 1-5, homozygous *srf4-2*, and *srf4-3*, and the *quartet* mutant plants.

##### **2.2.8.2.1 Leaf size assessments**

To prevent crowding artifacts, only 5 plants per pot (7 cm x 7 cm) were grown. Plants were examined for morphological changes every two days after germination. The leaves were numbered from the first rosette leaf that appeared after the cotyledons to the last rosette leaves. Cauline leaves were numbered independently of rosette leaves.

Leaf blade measurements were made using the dissected, 16-day old cotyledons, 16- and 26-day old fifth rosette leaves, and 42-day old second cauline leaves. Petiole length of the 16-day old rosette leaves was also measured. These leaves were placed on white pages and scanned. The measurements were analyzed with ImageJ software (<http://rbs.info.nih.gov/ij/>).

#### **2.2.8.2.2 Stem size measurements**

Stem size of the 6-week old Col wild-type, *dlg-1* (*At1g79870* homozygous T-DNA insertion line), homozygous T3 lines of *35S::SRF4* overexpression lines, 3-12, homozygous *srf4-2*, *srf4-3*, and *quartet* mutant plants were assessed.

#### **2.2.8.2.3 Hypocotyl measurements**

To analyze the hypocotyl lengths of Col wild-type *35S::SRF4* (3-12), and *dlg-1*, *srf4-2*, *srf4-3*, *quartet* mutant, surface-sterilized dry seeds were plated on 0.5 x Murashige and Skoog (MS) medium, and then treated with cold for 4 days at 4°C under dark condition. Afterwards, plates were placed vertically in a growth chamber for 8 hours. All plates were wrapped with aluminium foil to provide dark conditions. After in total 72-hour incubation of plants, all plates were scanned and hypocotyl lengths were measured using the ImageJ software.

#### **2.2.8.2.4 Root measurements**

To analyze the root lengths of Col wild-type *35S::SRF4* (3-12), and *dlg-1*, *srf4-2*, *srf4-3*, *quartet* mutant, surface-sterilized dry seeds were grown on 0.5x Murashige and Skoog (MS) medium, After the cold treatment of the seeds, plates were placed vertically in a growth chamber for 10 days. Afterwards, root lengths were measured with the help of ruler.

### **2.2.9 Microscopy and art work**

Plant pictures were taken with Kodak DCS760 digital camera (Eastman Kodak). Closer pictures of siliques and flowers were taken with Color-View III CCD camera mounted on SZX-12 BINO (Olympus, Japan). Image manipulations were made by means of Adobe Photoshop 7.0 software (Adobe system Inc.).

#### **2.2.9.1 Cell size analysis**

The modified dried-gel method was used to analyze the epidermal cell sizes of the 11-day old cotyledons. For this purpose, the plants were grown on MS plates in a growth chamber (Horiguchi *et al.*, 2006). Microscope slides covered with a thin layer of 2% low-melt agarose containing 0.01% bromophenol blue

dye pre-warmed at 50°C were used. Cotyledons were arranged on agarose. After solidification of the agarose, cotyledons were carefully peeled from the slides leaving the epidermis layer on the slide and the remaining gel cast was dried 10 minutes more at room temperature. Epidermal cells in the center of leaf blade, between midvein and leaf margin, were analyzed under the microscope, BX61 (Olympus, Japan) using the differential phase contrast device. All pictures were taken by using the Color-View III CCD camera mounted on the BX61 microscope (Cell-P Olympus, Europa).

### **2.2.9.2 Analysis of the *35S::SRF4:EGFP* transgenic plants**

#### **2.2.9.2.1 Screening of EGFP transgenic lines**

To find out the EGFP transgenic lines expressing SRF4, seven-day old roots of the 100 T1 generation plants carrying the *35S::SRF4:EGFP* construct were examined under the epifluorescence microscope. Six transgenic lines displaying high, middle and weak brightness were chosen for further investigation.

#### **2.2.9.2.2 Live imaging of *35S::SRF4:EGFP* transgenic lines using confocal microscopy**

The roots of the 7-day old T1 and T2 transgenic lines carrying the *35S::SRF4:EGFP* construct were investigated by using the FV1000 IX81 FLUOVIEW confocal laser scanning microscope (Olympus, Japan). For confocal images, roots were cut into small pieces and placed on microscope slides, covered with cover slip. After excitation of the tissue at 488 nm, EGFP fluorescence was collected through a band filter (BP502-536 nm).

#### **2.2.9.2.3 Plasmolysis**

To determine the plasma membrane localization of SRF4:EGFP, the roots cells of the *35S::SRF4:EGFP* transgenic plants were treated with 1M sorbitol at room temperature for 10-15 minute. Afterwards, the retraction of the SRF4:EGFP labeled cell membrane was visualized by using confocal microscopy.

## **Chapter 3. Cloning and Structural Analysis of the *STRUBBELIG RECEPTOR FAMILY (SRF) Genes***

### **3.1 Introduction**

#### **3.1.1 Extracellular organization of LRR-RLKs**

Protein-protein interactions play an important role in many signaling pathways (Li, 2005). Studies have revealed that extracellular domains of the LRR-RLKs, such as LRRs, and PRRs are involved in the protein-protein interaction, and binding of a diverse range of biochemical substrates like steroid molecules (Shiu and Bleecker, 2001). Interestingly, the extracellular domain structure and sequence identity of RLKs varies more than the intracellular kinase domains suggesting that RLKs may function in the perception of a wide range of signals or stimuli.

##### **3.1.1.1 Leucine-rich repeats (LRRs)**

LRR-RLKs can contain between of 1-32 LRRs. Each repeat possesses approximately 24 amino acid residues in length with conserved leucines (Kajava, 1998). Ultimately, each LRR repeat forms an  $\alpha/\beta$  structure, and a number of LRRs form a horseshoe structure that is a predicted cleft for protein-protein interaction to take place. LRRs are usually rich in hydrophobic amino acids. The plant extracytoplasmic LRR motif has an LxxLxxLxxLxLxxNxLxGxIPxx consensus sequence (X represents non-conserved residues) (Torii, 2004). The most conserved region of this motif, which is underlined in the LRR consensus sequence, forms a  $\beta$  sheet with the exposed face involved in protein-protein interaction (Walker, 1994). Moreover, the differences in this region may provide the specificity of the interaction (Kobe and Deisenhofer, 1993). This is supported by the fact that the number of LRRs varies among different RLKs (Torii and Clark, 2000).

##### **3.1.1.2 Proline-rich region (PRR)**

Protein interaction domains like proline-rich regions (PRRs) are highly diverse and they play different roles in cell growth, differentiation, polarity, motility and

apoptosis (Williamson, 1994). Furthermore, these domains are highly adaptable and therefore have the essential property to contribute to the evolution of new pathways and the modulation of signaling dynamics. The PRR domains may comprise between 30-200 amino acids. These domains are widespread in prokaryotic and eukaryotic cells (Williamson, 1994). Typically, protein interactions with PRR are very weak and possess low specificity. However, weak interactions are known to play important regulatory roles in certain processes. For instance, a transient interaction between *Caenorhabditis elegans* NCK-related adaptor protein (NCK2) and PINCH1, important for muscle attachment, triggers the focal adhesion during integrin signaling via weak interactions between the third SH3 domain of NCK2 and the fourth LIM domain of PINCH1, which are cysteine-rich sequences mediating protein-protein interaction (Vaynberg *et al.*, 2005).

The importance of PRRs in protein-protein interactions is based on the unusual properties of proline (Williamson, 1994). Proline is a very unusual amino acid since its side chain is cyclized back to the backbone amid position. Among the 20 amino acids, proline is a well-known helix breaker of secondary structures like  $\alpha$ -helices or  $\beta$ -sheets, which are essential for protein folding. Having this characteristic, proline contributes to the formation of an accessible binding region for potential binding partners.

PRRs are defined based on a handful of possible conserved motifs. The PPII motif is one of the known sequence possessing a PxxP core motif (X can be any amino acid) (Li, 2005). Two or more prolines can cause a left-handed helix structure, providing a stable interaction region. This core motif is mostly found in the SH3 domain, an important family of PRRs. The SH3 domain of amphiphysin has a PxRPxR(H)R(H) consensus. Another consensus motif in the SH3 domain of Hbp (Hrs binding protein) has a Px(V/I)(D/N)RxxKP consensus motif. The xP (x mostly represents a hydrophobic amino acid) dipeptide motif provides a hydrophobic groove for interaction. Moreover, the aromatic amino acids tryptophan, tyrosine and phenylalanine are often found in PRRs as their side chains can build van der Waals bonds with the interaction proteins.

### 3.1.1.3 PEST region

Another region that could be found in the intracellular and in some cases in the extracellular domain is a PEST region. Since PEST regions are associated with proteins that have a short intracellular half-life, it is hypothesized that PEST sequences act as a signal for protein degradation. Protein regions rich in proline (P), glutamic acid (E), serine (S), and threonine (T) are called PEST region (Rogers *et al.*, 1986). These regions are often flanked by positively charged amino acids. Recent results showed that PEST sequences target proteins for degradation by the proteasome pathway (Chinchilla *et al.*, 2007). Robatzek and coworkers (2006) demonstrated that FLS2 undergoes degradation via endocytosis when it is activated with its ligand. Furthermore, a single amino acid change in the PEST motif suggests that the PEST region of FLS2 might be involved in the ubiquitin-triggered receptor endocytosis.

### 3.1.2 The intracellular kinase domain

The kinase domain of RLKs consists of 11 subdomains (I to XI) Subdomain VI is further divided into, VIa and VIb (Hanks and Quinn, 1991). Collectively, these subdomains form 2 different lobes called N- and C-terminal lobes (Huse and Kuriyan, 2002). The N lobe is composed of subdomains I to IV creating a five-stranded  $\beta$  sheet and one  $\alpha$  helix. The C lobe is larger than the N lobe and generates helical structures. The intracellular kinase domain of the plant RLK possesses serine/threonine substrate specificity with two consensus D $\Phi$ KxSN sequence in subdomain VIb and a GTxGY $\Phi$ APE sequence in subdomain VIII (x represents any amino acid residue,  $\Phi$  represents aliphatic amino acids) (Torii and Clark, 2000).

All functional kinases in plants catalyze the transfer of the gamma ( $\gamma$ ) phosphate of ATP to the hydroxyl group of serine and threonine (Huse and Kuriyan, 2002). A glycine-rich sequence motif in subdomain I (GxGxxG) located in the phosphate-binding loop allows the proper positioning of the loop to approach the phosphate of the ATP. Lysine in subdomain II interacts with the  $\alpha$  and  $\beta$  phosphates of ATP to position them for catalysis. Moreover, glutamine in subdomain III plays a role in the ATP anchoring process. Aspartate and asparagine in subdomain VIb, which is also called the catalytic loop, are highly

conserved in catalytically active kinases. Asp in this domain behaves as a base to remove a proton from the substrate hydroxyl group (Johnson *et al.*, 1996). CLV1 and BRI1, which possess a highly conserved Asp (D) and Asn (N) in the catalytic domain, exhibit active kinase domains (Clark *et al.*, 1997; Li and Chory, 1997). The activation loop located between the DFG motif in subdomain VII and the APE motif in subdomain VIII has a role in substrate recognition (Torii and Clark, 2000). The *bri1-104* and *clv1-9* mutations in the activation loop, which change the highly conserved alanine to threonine (A1031T) or valine (A839V), respectively, cause a phenotypic defect (Friedrichsen *et al.*, 2000; Clark *et al.*, 1997). These mutations affect proper substrate recognition and perturb the signal transduction.

### 3.1.3 Outlook on Chapter 3

Studies have shown, mainly by analyzing mutants exhibiting phenotypic defects, the importance of certain amino acids and motifs in kinase and other domains of RLKs with known biological functions. Therefore, comparative sequence analysis allows predicting the functionality of newly available RLKs.

In this chapter, we present the molecular cloning and structural analysis of the *Arabidopsis thaliana* STRUBBELIG RECEPTOR FAMILY members. We found that all SRF members, except SRF1B, encode leucine-rich repeat receptor-like kinases and possess six LRR repeats. Certain members of the family have a PRR either in the extracellular or intracellular domain. Analysis of the kinase domain amino acid sequences revealed that all SRF family members, except SRF2, might be atypical kinases with an inactive kinase domain. Interestingly, among the SRF members only SRF1 undergoes alternative splicing and this is predicted to encode two different isoforms. SRF1A would encode an LRR-RLK whereas SRF1B an LRR-RLP lacking most part of the intracellular domain. Moreover, we found a high number of polymorphisms for the SRF1 gene between the Col and Ler background indicating a probable evolutionary selection.

## 3.2 Results

Although the *Arabidopsis* genome has been sequenced, the functions of the high number of the genes are still unknown. For instance, the function of 2,500 genes out of 28,000 genes in *Arabidopsis* has been investigated experimentally (Hilson *et al.*, 2003). Moreover, most of the genes were inaccurate bioinformatic predictions. Approximately 30% of the genome is wrongly annotated according to a whole genome transcription analysis (Yamada *et al.*, 2003). Since most molecular experiments depend on the available correct full-length cDNA clones, the most important step is to construct reliable cDNA clones.

### 3.2.1 Generation of the full-length *SRF* members

#### 3.2.1.1 Expressed sequenced tag analysis of *SRF* genes

In order to characterize the *SRF* genes at a molecular level various ESTs were investigated. Totally, 31 ESTs for seven of the eight *SRF* genes were analyzed. There were no ESTs available for *SRF2* at that time (Table 2.5). First, the longest EST of each gene was sequenced. However, sequence analysis demonstrated that we could not obtain any full-length sequence although full-length was predicted by the sources (see Table 2.5) that provided the ESTs. We also recognized that the ESTs called RZL58c11F, 021C025, 008A11, 103D18T7 and G8F12T7 were chimeras of ESTs corresponding to different genes. Therefore, these ESTs were not used in our analysis.

#### 3.2.1.2 Cloning and annotation of the full-length *SRF* cDNA sequences

We determined the full-length sequences of the *SRF1A*, *SRF1B*, *SRF2-SRF8* cDNAs including 5'UTR, translation start site, and 3'UTR by means of 5' and 3' RACE-PCR. *SRF* family members cDNAs were amplified and cloned by various methods (see 2.2.3.3). The sequences were confirmed by comparison with the corresponding genomic sequences obtained from the NCBI *Arabidopsis* database. All *SRF* cDNAs sequence data obtained with RACE-PCR were deposited in the EMBL/GenBank. Their AGI codes, GenBank accession numbers, and map positions are given in Table 3.1.

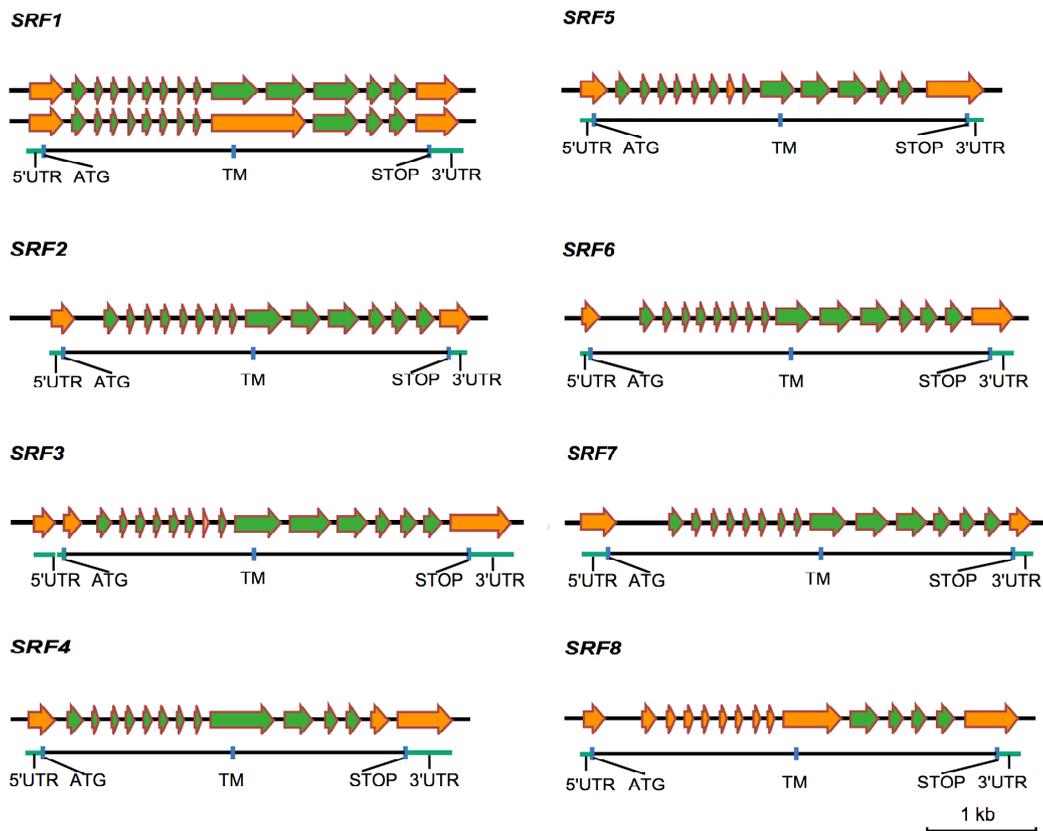
The generation of all *SRF* cDNAs showed that all *SRF* genes were wrongly annotated in the *Arabidopsis* database (Figure 3.1). The most obvious differences were found in *SRF4* and *SRF8*. Although *SRF4* was annotated

between map locations 4184783-4190867, this study identified that *SRF4* is located between map positions 4187384-4191306. The first part of the annotated gene that is located between map positions 4184590-4187092 encodes another protein. Our EST analysis also supported this result. Sequence analysis of the EST pda01557 showed that the beginning part of the annotated gene belongs to gene *At3g13062*. Interestingly, *SRF8* formerly was annotated as receptor-like cytoplasmic kinase (RLCK) lacking the extracellular domain located between map positions 11725875-11727344. Our data indicated that *SRF8* is located between map positions 11723637-11727690, possessing eight extra exons. As a consequence, *SRF8* also encodes an LRR-RLK rather than a receptor-like cytoplasmic kinase. The correct reannotated *SRF* sequence information was submitted to MIPS and TAIR *Arabidopsis* database.

**Table 3.1: Summary of information on the *SRF* family members**

<b>Gene name</b>	<b>AGI code</b>	<b>GenBank accession no.</b>	<b>Map location</b>	<b>Length of the cDNA (bp)</b>
<i>SUB</i>	<i>At1g11130</i>	AF399923	3723982-3728378 (F)	2,659
<i>SRF1A Col</i>	<i>At2g20850</i>	AY518286	8982582-8986460 (R)	2,794
<i>SRF1B Col</i>		DO914918	89824582-8986460 (R)	2,883
<i>SRF1A Ler</i>		DO914919		2,794
<i>SRF1B Ler</i>		DO914920		2,883
<i>SRF2</i>	<i>At5g06820</i>	AY518287	2112870-2116760 (F)	2,429
<i>SRF3</i>	<i>At4g03390</i>	AY518288	1490460-1494889 (R)	3,037
<i>SRF4</i>	<i>At3g13065</i>	AY518289	4187384-41911306 (F)	2,639
<i>SRF5</i>	<i>At1g78980</i>	AY518290	29712581-29716314 (R)	2,492
<i>SRF6</i>	<i>At1g53730</i>	AY518291	2065369-20069386(F)	2,472
<i>SRF7</i>	<i>At3g14350</i>	AY518292	4783067-4787243 (R)	2,446
<i>SRF8</i>	<i>At4g22130</i>	AY518292	11723637-11727690 (F)	2,567

During *SRF* cDNA sequence analysis many nucleotide alterations between annotated Col wt sequences and our results were found. The reverse transcriptase used for RT-PCR did not possess high proofreading abilities. This could have led to the formation of several mutations in the *SRF* cDNA sequences. Nucleotide alterations leading to amino acid changes were corrected with the help of the SDM method. Nucleotide alterations that did not cause amino acid changes were not corrected. All polymorphisms are shown in Table 3.2.



**Figure 3.1:** The molecular organization of *SRF1-8*. Arrows represent exons. Green arrows show correctly annotated exons whereas orange arrows indicate a corrected exon annotation based on the cDNA sequence analysis in our study. ATG and STOP represent the start and the end of the protein translation. Abbreviations: TM, transmembrane domain; UTR, untranslated region.

### 3.2.2 Alternative splicing analysis of the *SRF* genes

During the cloning of the full-length *SRF1* cDNA we found that this gene has two alternative splicing versions, called *SRF1A* and *SRF1B*. *SRF1A* encodes a receptor-like kinase (RLK) whereas the translation of *SRF1B* corresponds to a receptor-like protein (RLP) lacking the kinase domain. The *SRF1* gene contains 15 exons and 14 introns. The first 10 exons are the same for the two splice variants. Contrary to the *SRF1A* cDNA, the 10<sup>th</sup> intron of *SRF1* is not spliced out during *SRF1B* pre-mRNA formation causing a difference between *SRF1A* and *SRF1B*. Therefore, *SRF1B* is the result of the retention of intron 10 which creates a premature stop codon leading to a short protein having the SP, SUB domain, 6 LRRs, PRR, TM and a short cytoplasmic tail. *SRF1B* lacks 293 amino

acids of the C-terminal region. Hence, the two splice variants are identical until residue 482 but they are different in their carboxy terminal sequences.

**Table 3.2: Nucleotide changes in *SRF* cDNAs RACE results**

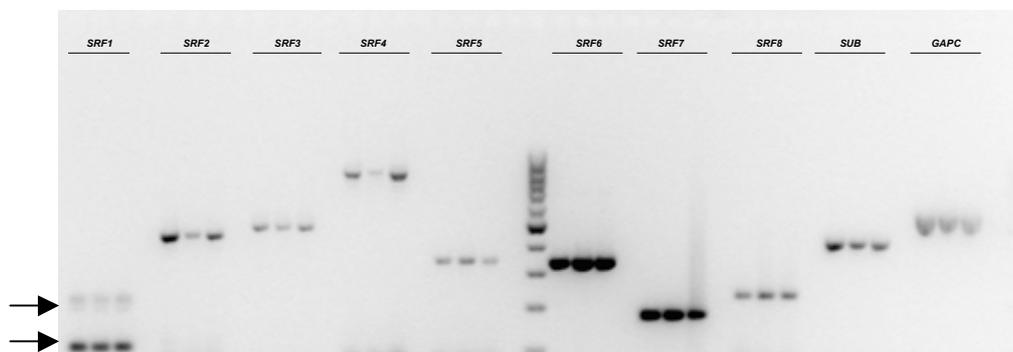
Gene name	Substitution	Nucleotide position	Amino acid change
<i>SRF1A</i>	T → C	102 (5'UTR)	
	T → C	117 (5'UTR)	
	<u>CTI</u> → <u>CTC</u>	785	Phe → Phe
	<u>CTG</u> → <u>ITG</u>	2367	Leu → Phe
	<u>TCA</u> → <u>TCG</u>	2435	Ser → Ser
<i>SRF1B</i>	T → C	102 (5'UTR)	
	T → C	117 (5'UTR)	
	<u>CTI</u> → <u>CTC</u>	785	Phe → Phe
	<u>GCT</u> → <u>CCT</u>	1164	Ala → Pro
	A → G	1786 (3'UTR)	
	C → T	2456 (3'UTR)	
	A → G	2524 (3'UTR)	
<i>SRF2</i>	<u>TTI</u> → <u>TTC</u>	608	Phe → Phe
	<u>CAA</u> → <u>CAG</u>	1907	Gln → Gln
	<u>GGA</u> → <u>GGI</u>	1949	Gly → Gly
	G → A	2423 (3'UTR)	
<i>SRF3</i>	<u>GAC</u> → <u>GAT</u>	436	Asp → Phe
	<u>ATA</u> → <u>GTA</u>	779	Ile → Met
	<u>GGT</u> → <u>GGC</u>	1333	Gly → Gly
	A deletion	2716 (3'UTR)	
	C → G	3032 (3'UTR)	
<i>SRF4</i>	<u>CGG</u> → <u>CGA</u>	160	Arg → Arg
	<u>GTG</u> → <u>GCG</u>	987	Val → Ala
	T → C	2412 (3'UTR)	
<i>SRF5</i>	<u>GCI</u> → <u>GCC</u>	893	Ala → Ala
<i>SRF6</i>	<u>CAG</u> → <u>IAG</u>	115	Gln → STOP
	<u>GGA</u> → <u>GGG</u>	1354	Gly → Gly
	<u>TCT</u> → <u>TTT</u>	2112	Ser → Phe
	<u>GAT</u> → <u>GAC</u>	2182	Asp → Asp
	C deletion	2469 (3'UTR)	
<i>SRF7</i>	C deletion	176 (5'UTR)	
	T deletion	181 (5'UTR)	
	C deletion	182 (5'UTR)	
	TGC → T_C	2253	
<i>SRF8</i>	<u>AGC</u> → <u>GGC</u>	1804	Ser → Gly

Furthermore, the alternative splicing event of *SRF1* is also detected in the *Ler* ecotype of *A. thaliana* indicating that it is not related to the *SRF1* Col/*Ler* polymorphisms (see below). PCR amplification of *Ler* cDNAs from different tissues like rosette leaves, cauline leaves, siliques, stems, roots, and seedlings showed that the two splice variants are expressed in all types of tissues with variable levels (Figure 3.2). We also checked the other *SRF* members for

alternative splicing by RT-PCR using primers flanking the 10<sup>th</sup> intron, but we did not observe the retention of exon 10 in other SRF members (Figure 3.3).



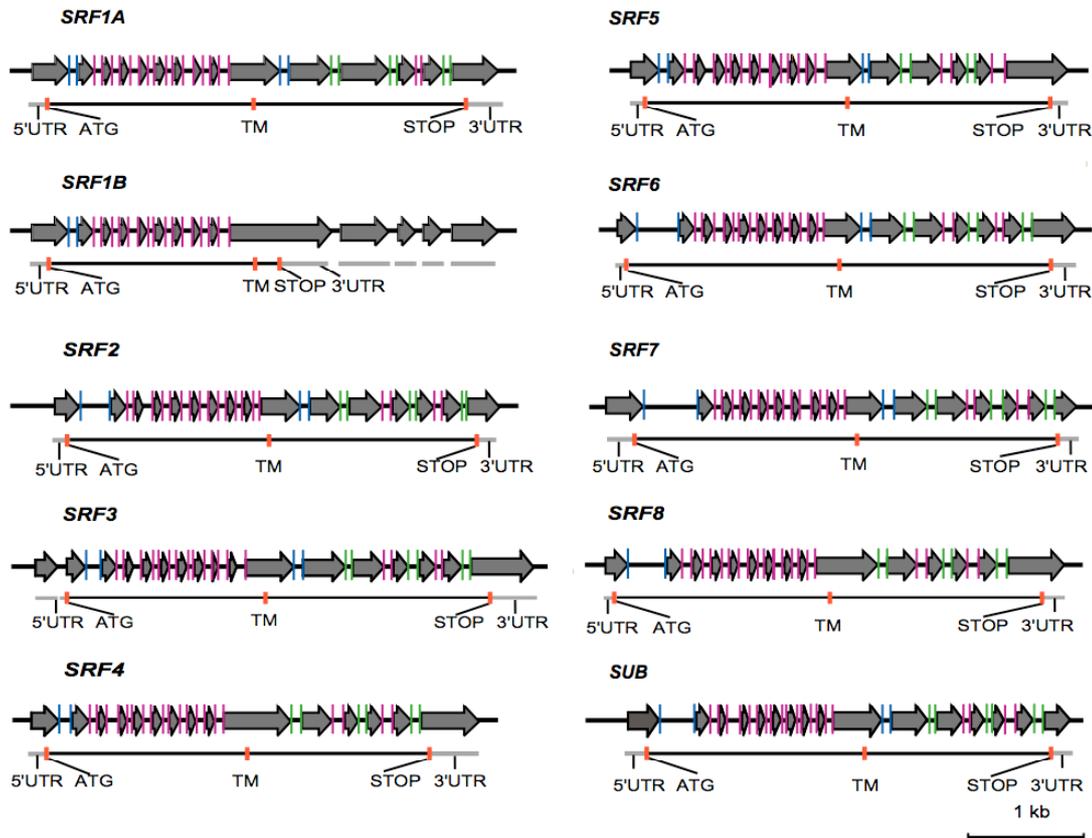
**Figure 3.2:** RT-PCR based expression profiles of SRF1A and SRF1B. GAPC was used as a control.



**Figure 3.3:** RT-PCR based alternative splicing analysis of the all SRF members. GAPC is used as a control. cDNA from Col flowers of stage 1-12 (Smyth *et al.*, 1990) (*SRF1-8*) were used.

### 3.2.3 Exon-intron organization and chromosomal distribution

Comparison of the genomic DNA and the cDNA by considering the fact that in eukaryotes the first dinucleotide of introns is generally a GT and the last dinucleotide an AG (Breathnach and Chambon, 1981) allowed the determination of the intron and exon repartition of the *SRF* genes. All exon/intron boundaries of the *SRF* genes exhibit this characteristic. After the identification of correct exon-intron positions of the *SRF* genes, we compared the exon/intron properties among *SRF* members (Figure 3.4). We could show that *SRF1A*, *SRF4*, *SRF5* and *SRF8* have 15 exons and 14 introns. In contrast, the other *SRF* sequences show variation in the exon-intron number. For example, *SRF6* has 16 exons and 15 introns while *SRF1B* has 14 exons and 13 introns, indicating gain or loss of introns. In *Arabidopsis* genomes, the average length of exons and introns



**Figure 3.4:** Exon and intron types of the *SRF* family members. Gray arrows mark exons. Type 0 introns are highlighted with a green color stripe, type 1 introns are shown with a blue color stripe, and type 2 introns are marked with a pink color stripe.

decreases with the increasing number of introns per gene (Ivashchenko and Atambaeva, 2004). It is known that the average length of all exons in a gene in the *Arabidopsis* genome containing at least 15 introns is 3195 bp. This shows that the total exon length of the *SRF* genes well matches the expected length. The average length of the exons and introns in the gene possessing 15 introns is about 152 bp and 148 bp, respectively. However, the length of exons of the *SRF* genes varies between 66 to 884 bp. Especially the length of exons encoding proline-rich region (PRR), transmembrane domain (TM), juxtamembrane domain (JM) and the beginning of the kinase domain of the respective *SRF* members ranges between 266-598 bp, a finding which is unusual for a gene bearing 15 introns.

Each *SRF* gene, excluding *SRF3*, contains 9 introns at a similar location until the region encoding the TM. Exons specifying SP, SUB and LRR are similar in size, which range from 66-133 bp. Differences start with the intron

located after the TM coding region. To understand the reason for the differences, we examined the exon-intron types of the *SRF* genes. This classification was done by using the splice frame rules (Kolkman and Stemmer, 2001). Introns can be located in the reading frame of the gene between two consecutive codons (phase 0 intron), between the first and second nucleotide of a codon (phase 1 intron), or between the second and third nucleotide (phase 2 intron) of a codon. Exons can also be classified into nine groups based on intron phases. Exons with 0-0, 1-1, and 2-2 types are called symmetric exons, and exons with 0-1, 0-2, 1-0, 1-2, 2-0 and 2-1 are named asymmetric exons. Figure 2.3 shows exon and intron types of the *SRF* genes. This examination revealed that most of the *SRF* genes have 2-1 and 1-0 exon types in the region located after the TM coding region, whereas *SRF4* and *SRF8* have 2-0 exon types indicating that there may have been an intron deletion in this part of the genes. Moreover, *SRF5* lacks the last 0-0 type of exon while at the same position the other *SRF* genes have 2-0 and 0-0 types of exons. This indicates that there can be a deletion of the intron between these two exons or they lack the last exon. In addition, this exon is very short for *SRF4* and *SUB*. *SRF4* and *SUB* also do not possess a C-terminal extension. *SRF1A* has a 1-0 and 0-0 exon type after TM. All the other *SRF* genes have 1-0, 0-2 or 2-0 exons at this position. This result indicates that there can be an intron deletion between 0-2 and 2-0 exons of the *SRF1A* gene.

Interestingly, all *SRF* members have symmetric exon types (2-2) in the extracellular domains. Kolkman and Stemmer (2001) suggested that the symmetric exons are the only ones that can be inserted into introns of the same phase, undergo tandem duplication, or be deleted without disturbing the reading frame. The highly conserved intron phasing and position of introns with respect to their nucleotide sequences also indicates the evolution of the *SRF* members from the same ancestral gene by exon shuffling.

Investigation of genomic coordinates of the *SRF* genes emphasized that *SRF1/SRF3*, *SRF4/SRF5*, and *SRF6/SRF7* gene pairs were found within segmentally duplicated blocks. We therefore propose that the sister pairs could have been originated via duplication of a DNA segment. However, *SUB*, *SRF2* and *SRF8* are not located in duplicated regions (Blanc *et al.*, 2003).

### 3.2.4 Amino acid sequence analysis

The cloning of the full-length cDNAs of all *SRF* family members allowed us to perform a detailed comparative analysis on the deduced amino acid sequences and their structural characteristics in order to infer clues about the biochemical properties and physiological functions of the SRF proteins.

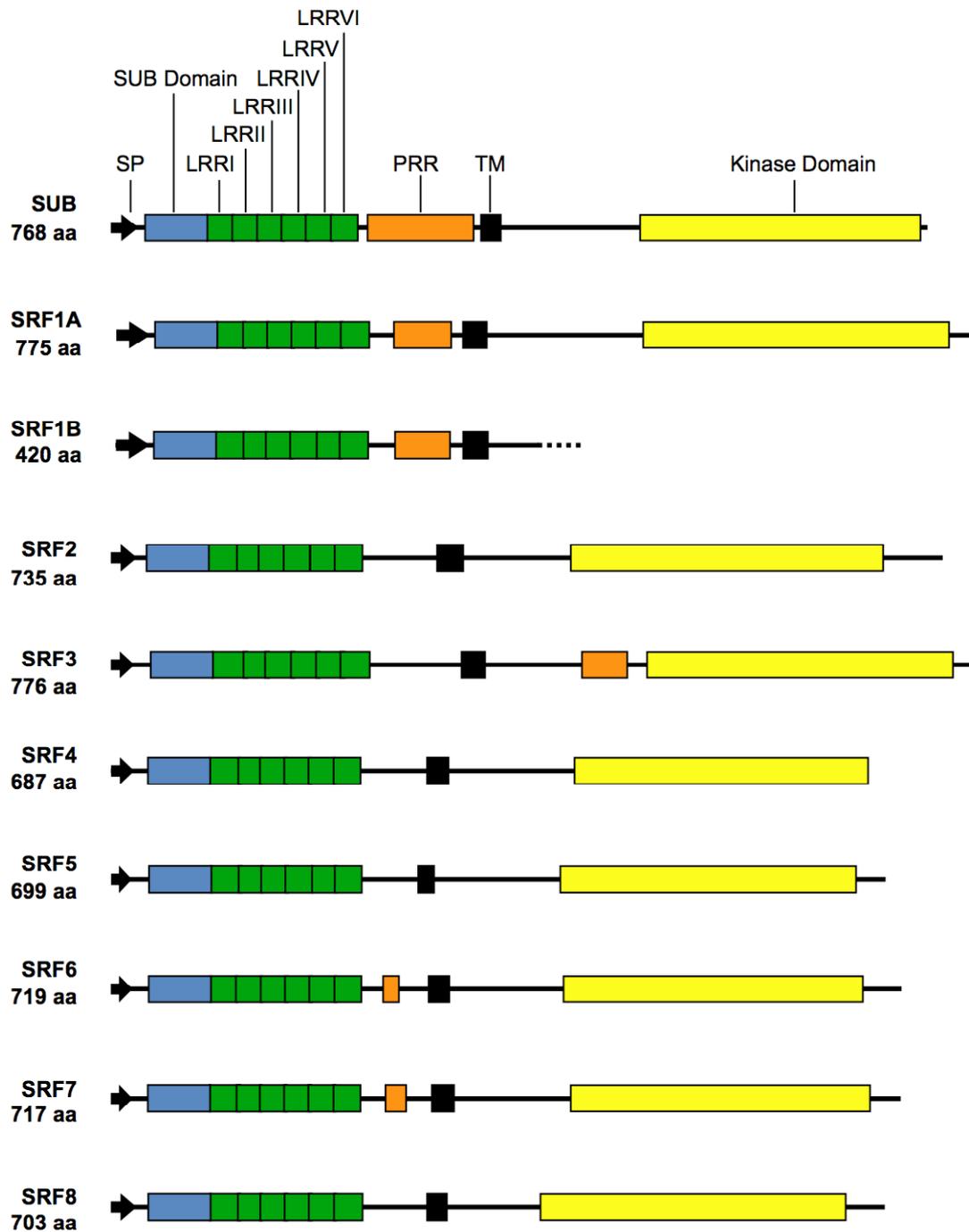
Amino acid sequence analysis revealed that full-length protein sequences range from 420 amino acids (SRF1B) to 776 (SRF3) (see Fig. 3.5), the mean length is 700 amino acids. The overall identities range from 29.6 % (SRF2/SUB) to 77.9 % (SRF6/SRF7) (Table 3.3 A). The amino acid identities of the intracellular kinase domains, however, are substantially higher (mean = 55.42 %) and increase up to 92.0 % (SRF6/SRF7) (Table 3.3 B). In contrast, there is no identity with respect to the juxtamembrane domain and the C-terminal extension.

**Table 3.3: Amino acid identities (in %) among SRF protein**

A. Overall								
	SRF2	SRF3	SRF4	SRF5	SRF6	SRF7	SRF8	SUB
SRF1	32.5	57.9	32.5	34.9	37.1	36.7	41.0	40.0
SRF2		32.8	34.4	34.9	36.0	34.8	34.8	29.6
SRF3			34.1	35.2	38.7	38.4	39.7	42.6
SRF4				55.6	43.6	44.5	40.8	32.5
SRF5					45.0	45.0	42.1	35.1
SRF6						77.9	47.5	34.2
SRF7							47.2	34.4
SRF8								34.4
B. Kinase domain								
SRF1	43.8	78.1	46.2	49.5	53.9	53.9	60.3	54.1
SRF2		43.8	43.8	43.8	45.2	45.9	48.1	41.7
SRF3			48.0	52.0	56.1	55.6	62.5	58.1
SRF4				71.0	59.3	60.0	54.0	47.1
SRF5					63.3	64.7	57.9	50.7
SRF6						92.0	66.5	51.6
SRF7							66.5	52.0
SRF8								54.3

### 3.2.5 Investigation of the extracellular domains of the SRF members

Analysis of the SRF amino acid sequences showed that all SRF members encode putative leucine-rich repeat receptor-like kinases (RLKs) except SRF1B which encodes a leucine-rich repeat receptor-like protein (RLP) lacking the kinase domain. The predicted *SRF* genes encode LRR-RLKs possessing an



**Figure 3.5:** Domain organization of SRF proteins. The different C-terminus of SRF1B is marked by a dashed line. Abbreviations: LRR, leucine-rich repeat; PRR, proline-rich region; SP, signal peptide; TM, transmembrane domain.

extracellular domain (ECD), a transmembrane domain (TM), an intracellular juxtamembrane domain (IJM), and an intracellular serine/threonine kinase

domain. In addition, some of the proteins possess an extended C-terminus (Figure 3.5).

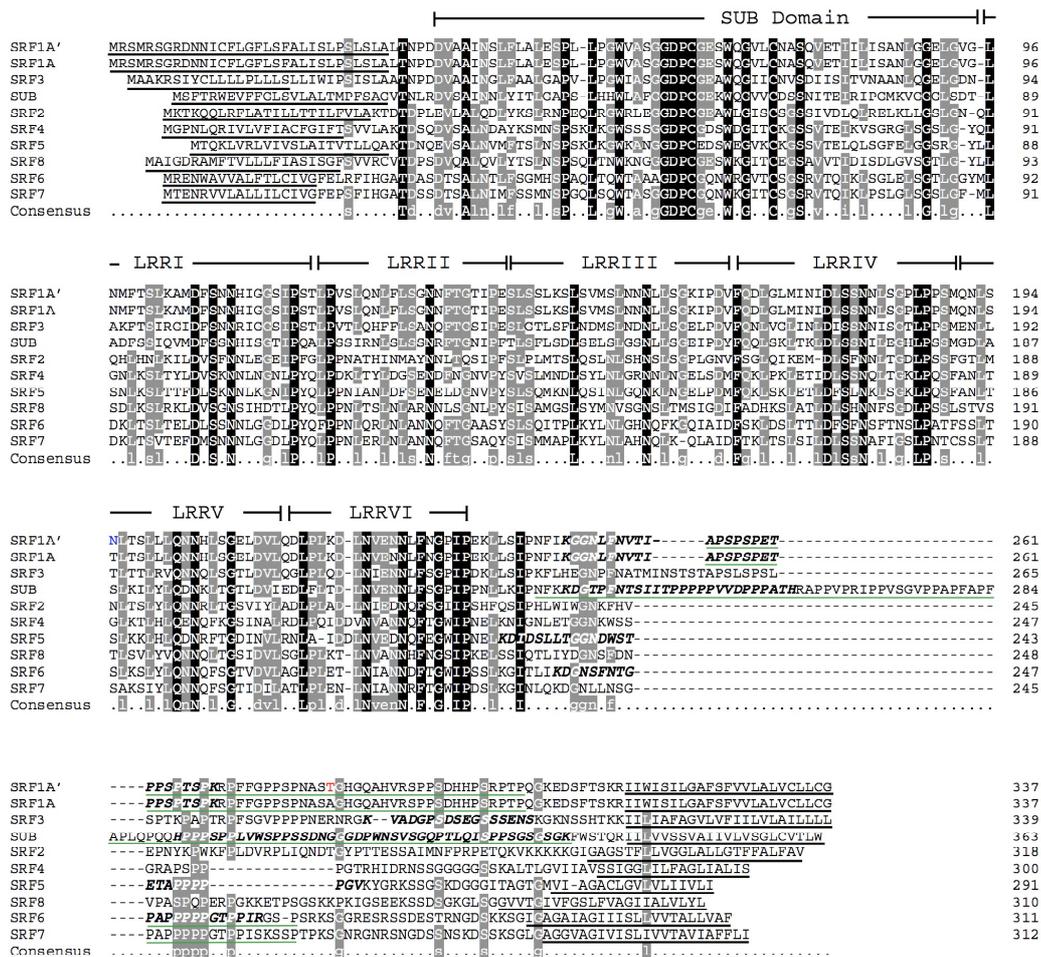
The extracellular domain of the SRF proteins comprises of a signal peptide (SP), a SUB domain, and six leucine-repeats (LRR) and certain SRFs possess a proline rich region (PRR) and PEST motif. The region encoded by the first exon of the *SRF* genes can serve as a hydrophobic signal peptide. In SRF1 and SRF7 this region corresponds to amino acid residues 1- 31 and 1-17, respectively (Figure 3.6). The residues encoding the TM domain of SRF1A and SRF7 are located between the amino acids 315-337 and 290-312, respectively. Interestingly, the hydrophobic residues of the signal peptide and the transmembrane are not conserved among family members. Between the SP and LRR domains, the SUB domain (Chevalier *et al.*, 2005), which is conserved among SRF members (amino acid identities >38%), was detected. The role of the SUB domain is unknown. However the SUB domain appears to be functionally relevant, as the *sub-3* allele that results in an amino acid substitution at a conserved position in the SUB domain cause *sub* phenotype.

### 3.2.5.1 Leucine rich repeats (LRRs)

To identify the LRRs within the SRF proteins we made comparisons to the L/ΦxxL/ΦxxL/ΦxxL/ΦD/xL/ΦS/xxNxL/ΦGxIPxx (Φ represents aliphatic amino acids) consensus sequences (Torii and Clark, 2000). According to this consensus, six LRRs (LRR I to LRR VI) are present in each SRF family member. Each LRR has a length ranging from 20 to 26 amino acids. In SRF members Asn and Gly are highly conserved in this consensus (Figure 3.6). However, SRF8 exhibits a G→D (LRR I) and G→M (LRR III) substitution. In addition, a G→Q substitution within the LRR II domain is present in SRF2 and a G→N substitution in the LRR IV domain of SRF6.

### 3.2.5.2 Putative Dimerization modules

The LRR domains of plant RLKs are often flanked by paired cysteines (Torii and Clark, 2000). However, all SRF members exhibit only one cysteine pair located at the SUB domain that is situated upstream of the LRR domains with following consensus sequence: CxxxWxGxxC.



**Figure 3.6:** Extracellular protein sequence alignment of SRF proteins. SRF1A' shows the *Ler* version of SRF1A. All other sequences correspond to the Col background. Full conservation across the alignment is marked by black columns, partial conservation by gray columns. Individual protein domains are indicated above. Red color shows a non-conservative residue exchange. The predicted signal peptide sequences and transmembrane domains are underlined with black lines. The proline-rich regions are underlined with green lines. Predicted PEST sequences are written in italic.

The length of the part located between the LRR repeats and the TM domain varies among different SRF members. In case of SRF2 this region possesses 67 residues (residues 228-294). Several SRF members display differences in the central part of this domain (residues 245-270 of SRF2). SRF1A/B, SRF3, and SUB have 10, 16, 44 amino acid insertions, respectively, in this region. However, SRF4 and SRF5 possess a 14 amino acid deletion in this part of the sequence. Certain SRF members display PRR and PEST sequences in this region.

### 3.2.5.3 PEST sequence

As in SUB, we found a putative PEST region for SRF1A, SRF1B, SRF3, SRF5, and SRF6, flanked by the LRR and TM domains. Interestingly, SRF3 and SUB bear two PEST regions. The algorithm scores to each possible PEST sequence as from -50 to +50 (Rogers *et al.*, 1986). Although a score over zero indicates a possible PEST region, a value above +5 being of particular interest. The PEST score for SRF1A/SRF1B, SRF3, SRF5, SRF6, and SUB are +6.28, +13.43/+14.11, +5.95, +6.51, +7.72/7.95, respectively. In addition, an intracellular PEST region for SRF4 having a score of +16.87 was also detected. In SRF proteins possessing proline-rich regions such as SRF1A, SRF1B, SRF6, and SUB, the PRRs are always overlapping with PEST sequences. Alignment of PEST regions of SRF members demonstrates that the PEST region does not exhibit a high conservation.

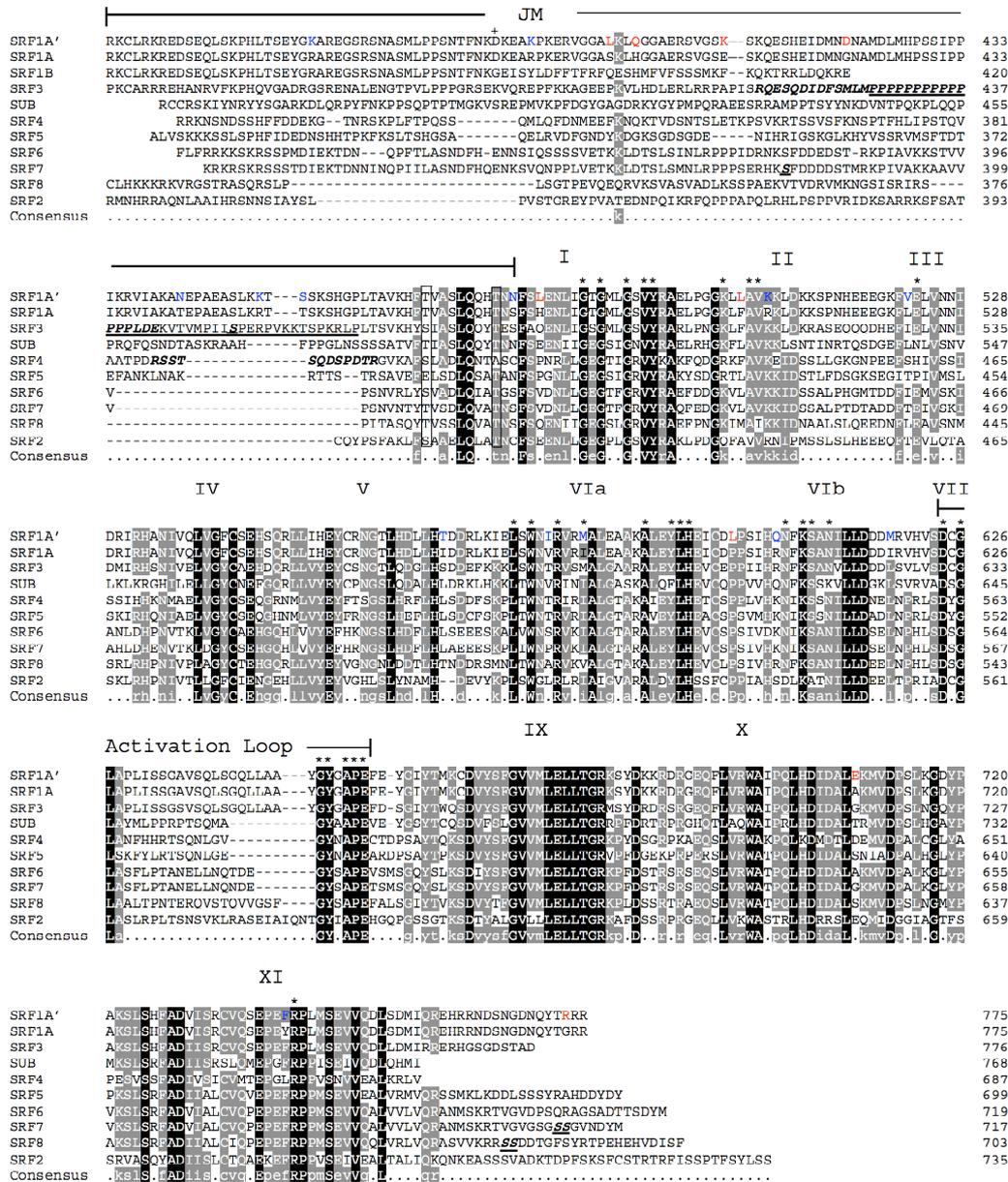
### 3.2.5.4 Proline-rich region (PRR)

The distinct region between the LRR and TM domains of certain SRF members, SUB, SRF1A, SRF1B, SRF6, and SRF7, also contains a proline-rich region (PRR). 16 of 50 amino acid residues of SRF1A PRR are proline, while SRF6 PRR possesses 9 prolines out of 16 amino acid residues. Interestingly, SRF3 presents a PRR in the intracellular part between residues 427 and 466 with 19 prolines. Moreover, SRF3 PRR contains a putative phosphorylation site of (Dr. Scott Peck, personal communication).

### 3.2.6 Examination of intracellular domains

Except for its C-terminal end, the intracellular juxtamembrane domain (JM) is highly variable among the family members (Figure 3.7). In this domain, SRF1-5, SRF8 and SUB have a highly conserved protease cleavage site with the P/GX<sub>5</sub>-<sub>7</sub>P/G consensus sequence (Yuan *et al.*, 2003). PHLTSEYG (amino acid residues 353-360) consensus of SRF1A and SRF1B, and GNDYKDG (amino acid residues 334-340) are examples for the cleavage site. Furthermore, SRF1A, SRF1B, and SUB have a YXXΦ tyrosine based endocytic sorting signature that can mediate the movement of TM proteins to lysosomes (Fritz-Laylin *et al.*, 2005). Furthermore, in the JM domain of SRF7 and SRF3, there

are putative *in vivo* phosphorylation sites, S379 and S452, respectively (Nühse *et al.*, 2004; Dr. Scott Peck, personal communication).



**Figure 3.7:** Intracellular protein sequence alignment of SRF proteins. SRF1A' shows the *Ler* version of SRF1A. All other sequences correspond to *Col* background. Full conservation across the alignment is marked by black columns, partial conservation by gray columns. Individual protein domains are indicated above. Red color shows a non-conservative residue exchange. Blue color highlights conservative or semi-conservative exchange. The proline-rich region is underlined with black lines. Predicted PEST sequences are written in italic. Asterisks highlight highly conserved kinase residues according to Hanks and Quinn (1991). The cross in the intracellular juxtamembrane region marks the point of deviation between SRF1A and SRF1B. The *in vivo* phosphorylation sites of SRF3, SRF7, and SRF8 are written in italic and are underlined.

All SRF members, except SRF1B, have an eleven-subdomained intracellular kinase domain. SRF kinase domains consist of 272 to 281 amino acid residues, corresponding to about 30 kD. The kinase domains of the SRF1A/SRF3, SRF2, SRF4/SRF5, SRF6/SRF7, and SRF8 have 279, 281, 272, 275 and 278 amino acids, respectively. The kinase domains are highly conserved, the amino acid identity of the SRF6 and SRF7 kinase domains amounts to 92.0 %.

The most significant heterogeneity is found in the C-terminal extension of the SRF members. SRF1A possesses 19 amino acid residues in C-terminal extension while SRF2 possesses 40 amino acid residues in this region. Interestingly, SRF4 and SUB proteins do not possess any C-terminus extension. Moreover, in the C-terminal extension SRF7 and SRF8 have putative *in vivo* phosphorylation sites at position <sup>710</sup>SS<sub>711</sub> and <sup>682</sup>SS<sub>683</sub>, respectively (Nühse *et al.*, 2004).

### 3.2.7 Phylogenetic analysis

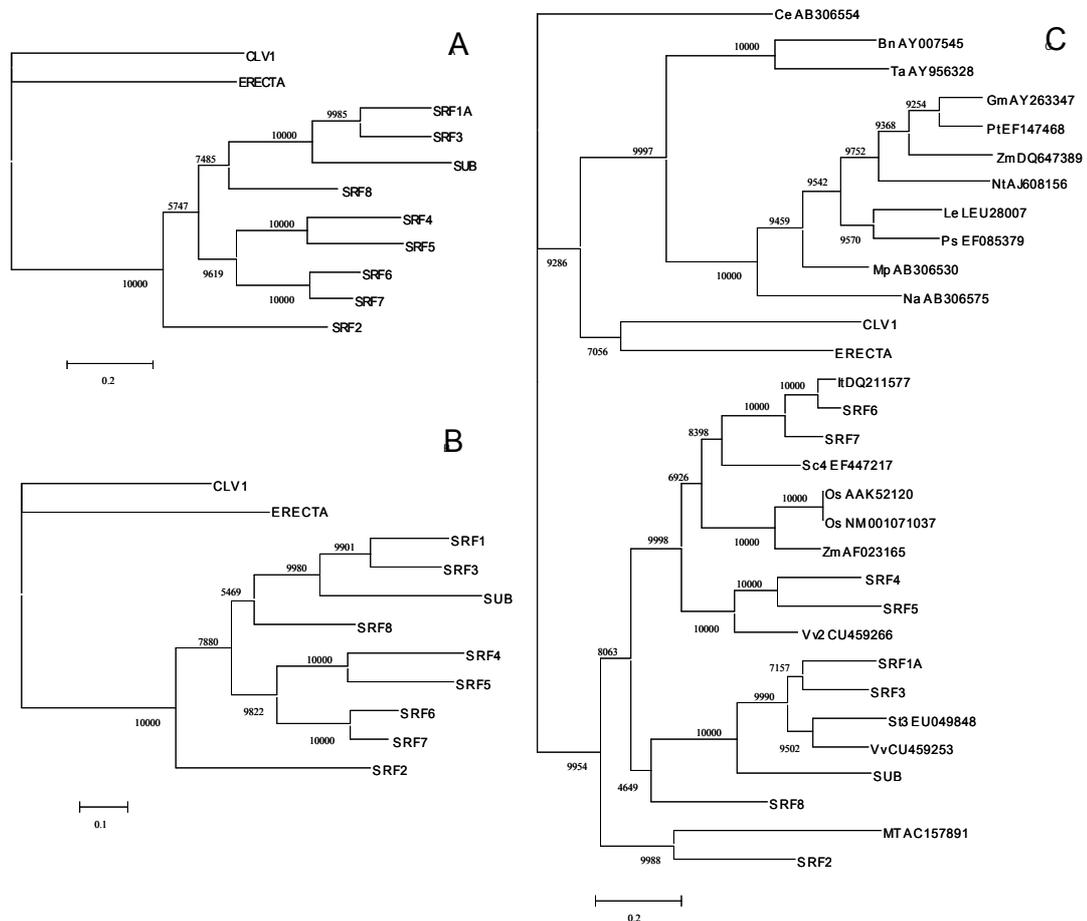
To determine the phylogenetic relationships between SRF family members, phylogenetic trees were constructed based on their cDNA sequences corresponding to the LRR, SUB and kinase domains. Figure 3.8 A shows a phylogenetic tree constructed by maximum likelihood using CLV1 and ERECTA RLK sequences as outgroups.

SRF2 is basal to the other SRFs, which means that SRF2 and the other SRFs share a common ancestor. The other SRFs form two clades, one clade consisting of SUB, SRF1 and SRF3 and the other of SRF4, SRF5, SRF6 and SRF7. SRF8 appears to be basal to the SUB/SRF1/SRF3 group, but this allocation is not well supported. SRF1/SRF3, SRF4/SRF5, and SRF6/SRF7 seem to have originated from relatively recent gene duplication events. As mentioned above, these pairs are found to be located on duplicated chromosomal blocks. The overall amino acid identity of the sister pairs amounts to 57.9 % (SRF1/3), 55.6 % (SRF4/5) and 77.9% (SRF6/7). In contrast, the genetic distance between the ERECTA and CLV1 RLKs and the SRFs is wide. The use of different substitution models, amino acid sequences, or methods (neighbor joining) delivered similar tree topologies, as well as the use of the SUB, LRR, and kinase domain sequences alone (Figure 3.8 B).

To gain a broader view at the evolutionary and taxonomical context, we chose the kinase domain and performed a BLAST search to identify further kinase sequences of other species. Besides several kinase sequences of monocot and dicot species, two sequences of the algae *Closterium ehrenbergii* and *Nitella axillaris*, one sequence of the liverwort *Marchantia polymorpha*, and one of the gymnosperm *Picea sitchensis* were found and included into the phylogenetic analysis. The resulting tree presented in Figure 3.8 C shows two main clusters, one cluster containing the CLV1 and ERECTA RLKs, and the other one the SRF RLKs. The divergence seems to have occurred very early in the evolution since the CLV1/ERECTA cluster contains also the *Nitella axillaris* kinase. The topology of the SRF cluster is essentially the same as found above. One sequence of *Isatis trinctoria* (Brassicaceae) was found to be a close relative of SRF6. The SUB/SRF1/SRF3 group now contains some other dicot sequences and the SRF4/SRF5/SRF6/SRF7 group further dicot and monocot sequences. Therefore, the divergence of the cluster containing the SRFs definitely occurred before the separation of monocots and dicots. However, since kinase sequences of *Vitis vinifera* and *Solanum chacoense*, respectively, are located basal to SRF1/SRF3, SRF4/5, and SRF6/SRF7, these SRFs may be specific for Brassicales or even for Brassicacea. Further sequences are necessary to definitely answer this question.

### 3.2.8 Assessment of substitution patterns

We estimated the rates of synonymous (Ks) and nonsynonymous (Ka) nucleotide changes in the kinase, SUB and LRR domains of the SRFs (Table 3.4). Ka indicates the number of nucleotide changes leading to changes in the amino acid sequence, while Ks reflects changes which do not affect the amino acid sequence. The ratio Ka/Ks is a measure for the selective pressure acting on a protein: a low Ka/Ks suggests the presence of purifying (negative) selection, a ratio >1 positive selection, and a ratio = 1 neutral selection which means the absence of either selective pressures.



**Figure 3.8:** Phylogenetic trees of SRFs and other RLKs. Trees were reconstructed by maximum likelihood using the HKY substitution model. Support values are indicated near the branches. Trees were reconstructed using nucleotide sequences of A: the SUB, LRR and kinase domain of SRFs, B: the kinase domain of SRFs, C: the kinase domains of SRFs and other RLKs.

Bn, *Brassica napus*; Ce, *Closterium ehrenbergii*; Gm, *Glycine max*; It, *Isatis trinctoria*; Le, *Lycopersicum esculentum*; Mp, *Marchantia polymorpha*; MT, *Medicago trunculata*; Na, *Nitella axillaris*; Os, *Oryza sativa*; Ps, *Picea sitchensis*; Pt, *Populus trichopardia*; Sc, *Solanum chacoense*; St, *Solanum toberosum*; Ta, *Triticum aestivum*; Nt, *Nicotiana tabacum*; Vv, *Vitis vinifera*; Zm, *Zea mays*.

As phylogenetic analysis demonstrated that SRF2 is basal to the other SRFs, we used the respective SRF2 cDNA sequence as an outgroup for the assessment of the substitution patterns. The results showed that the average Ks is similar for the kinase, SUB and LRR domains. But while Ks is constant or relatively constant for the kinase and LRR domains, respectively, it varies a lot with respect to the SUB domain. In contrast to Ks, the average Ka values are

different for the assessed domains. Again, the values are relatively constant with respect to the kinase domain than compared with the other two domains.

**Table 3.4: Number of substitutions in domains of SRF members**

	Kinase			SUB		
	$K_a$	$K_s$	$K_a/K_s$	$K_a$	$K_s$	$K_a/K_s$
SRF1	0.361	0.758	0.476	0.512	0.647	0.791
SRF3	0.357	0.781	0.457	0.428	0.648	0.660
SRF4	0.394	0.772	0.510	0.397	0.763	0.520
SRF5	0.378	0.780	0.485	0.379	0.919	0.412
SRF6	0.339	0.766	0.443	0.374	0.746	0.501
SRF7	0.346	0.747	0.463	0.409	0.692	0.591
SRF8	0.312	0.803	0.389	0.374	0.608	0.615
SUB	0.379	0.788	0.481	0.471	0.850	0.554
Mean	0.358	0.774	0.463	0.418	0.734	0.569
(SD)	(0.026)	(0.018)	(0.036)	(0.108)	(0.114)	(0.024)

	LRR		
	$K_a$	$K_s$	$K_a/K_s$
SRF1	0.414	0.804	0.515
SRF3	0.433	0.789	0.549
SRF4	0.400	0.675	0.593
SRF5	0.384	0.784	0.490
SRF6	0.384	0.706	0.544
SRF7	0.400	0.734	0.545
SRF8	0.355	0.850	0.418
SUB	0.417	0.847	0.492
Mean	0.398	0.774	0.515
(SD)	(0.064)	(0.018)	(0.053)

The lowest  $K_a/K_s$  ratios were obtained for the kinase domains, which indicates a purifying selective pressure acting on the maintenance of their function(s). The higher  $K_a/K_s$  ratios found for the SUB and LRR domain could reflect a lowered selective pressure acting on these domains, allowing the accumulation of an elevated number of nonsynonymous substitutions. These amino acid changing substitutions may, at least partially, have an impact on the function and therefore might contribute to the functional diversification of the SRFs. The highest  $K_a/K_s$  ratios were found for the SUB domain of SRF1 and SRF3.

All  $K_a/K_s$  ratios are markedly smaller than 1 which indicates the absence of positive selection. However, since positive selection may act on one or on a few sites only, it may not be detected by the employed method, which relies on the analysis of the (almost) complete coding sequences.

### 3.2.9 Polymorphism analysis in SRF1

The sub phenotype in above-ground organs is less prominent in the Col background (Chevalier *et al.*, 2005). The initial analysis indicated the existence of a genetic modifier of SUB linked to the *ERECTA* (*ER*) locus that is located at the middle of the chromosome 2 (Torii *et al.*, 1996). The *SRF1* homolog of *SUB* is also located close to *ER*. Therefore, it is thought that *SRF1* may be a modifier of SUB in the Col background. Hence, the entire genomic *SRF1* locus spanning nucleotides 8982429 to 8986460 (numbers as in Col) in *Ler* background including 5'UTR and 3'UTR was sequenced to identify the polymorphisms between Col and *Ler* genetic backgrounds. Moreover, the full-length sequence of the *SRF1A* and *SRF1B* cDNAs were also cloned from *Ler* and Col genetic backgrounds. Unfortunately, further experiments revealed that *SRF1* is not a modifier of SUB (Kay Schneitz and Dr. Lynette Fulton, unpublished result).

Interestingly, a very high number of polymorphisms within the 3,986 bp of *SRF1* genomic sequences was detected between the Col and *Ler* genetic background of the *SRF1* gene. These polymorphisms consist of 68 single nucleotide polymorphisms (SNPs), and 10 short insertion/deletions (indels) (Table 3.5). The 10 indels correspond to five insertions (from 2 to 6 bp) and five deletions (from 1 to 6 bp). Eight indels are located in introns and two deletions are observed in 3'UTRs. Of the single nucleotide polymorphisms, 47 are located in exons, 18 are present in introns, 2 are located in 5'UTRs, and 1 is present in a 3'UTR. Although other *SRFs* also display substitutions between Col and *Ler* ecotypes, none of them show as elevated levels of polymorphisms as *SRF1*.

**Table 3.5: Number of nucleotide polymorphisms in *SRF1A* (Col/*Ler*)**

Type of polymorphism	Exons	Introns	5'UTR	3'UTR	Total
Insertion	-	5	-	-	5
Deletion	-	3	-	2	5
SNP	47	18	2	1	68
Total	47	26	2	3	78

Ka / Ks = 1.23

The Figure 3.9A displays the frequency distribution of the *SRF1* polymorphisms at the *SRF1* locus. The analysis of *SRF1* *Ler* polymorphisms

indicates that 9 out of 47 SNPs in exons are located in the extracellular part of SRF1 whereas 38 SNPs are located in the intracellular part of SRF1 (Figure 3.9B). These results indicate that 79% of the polymorphisms affect the intracellular domain. A closer look at the amino acid sequence of SRF1 shows that 23 of the 25 non-synonymous, amino acid changing substitutions are located in the intracellular domains (Table 3.6), consisting of 10 changes in the JM domain, 12 changes in the kinase domain, and 1 in the C-terminal region. This finding suggests that the polymorphisms mostly affect SRF1A. Moreover, at least 6 of the 25 amino acid alterations in exon causes changing of the chemical and spatial properties of amino acids at the non-conserved positions among the protein sequences of the SRF family alignment (Figure 3.9C). The others are in conserved or partially conserved positions.

However, only one polymorphism located at position 600 in SRF1A causing a change from proline (Col) to leucine (Ler) is situated in a strictly conserved position among the SRF family members.

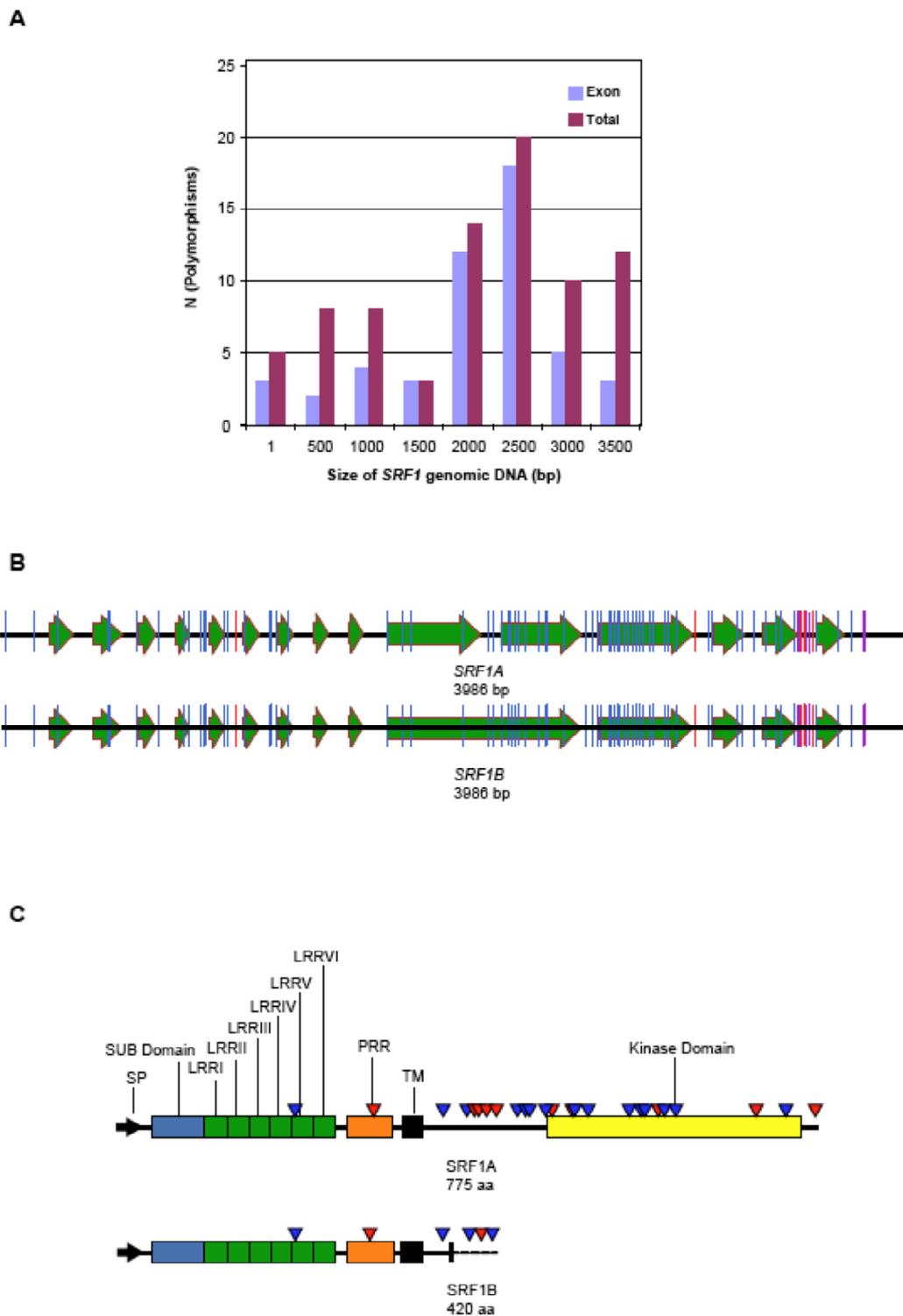
**Table 3.6: Location of single nucleotide polymorphisms in various SRF1A domains**

Type of polymorphism	SP	SUB Domain	LRR	PRR	JM	Kinase Domain	C-Terminus
Synonymous	1	2	3	1	-	13	-
Non-synonymous	-	-	1	1	11*	12	1
Total	1	2	4	2	11	25	1

\* Two of the SNPs are located in the first and third codon position of a triplet.

Ka / Ks = 0.25 (extracellular), 1.84 (intracellular)

Abbreviations: JM, juxtamembrane; LRR, leucine-rich repeat; PRR, proline-rich region; SP, signal peptide.



**Figure 3.9:** SRF1 Ler/Col polymorphisms. A) Frequency distribution of polymorphisms within the *SRF1* locus. B) Graphic representation of the polymorphisms within the *SRF1* genes. Green arrows indicate exons. Blue color shows a single nucleotide polymorphism, red color highlights an insertion and purple color denotes a deletion. C) Graphic representation of the location of the polymorphisms within the SRF1 proteins. A dashed line shows the different C-terminus of SRF1B. Polymorphisms are indicated by triangles. Red color highlights a non-conservative change, blue color a conservative or semi-conservative substitution.

### 3.2.10 Nucleotide substitutions in *SRF* gene family members

After the identification of many polymorphisms between Col and Ler ecotypes of the *SRF1* gene genomic sequences, all other *SRF* genes were also analyzed for polymorphisms. We used the Monsanto Ler sequence collection and PERLEGEN polymorphisms result (Jander *et al.*, 2002; Clark *et al.*, 2007). All *SRF* gene family substitutions are presented in Appendix A-G. However, we could not detect as many polymorphisms as in *SRF1*. Interestingly, all *SRF* genes, excluding *SRF1* and *SRF4*, possess indels in Ler background that cause the formation of early stop codon. The investigation of the *SRF2* sequences revealed that the *SRF2* gene possesses in total 25 substitutions including 20 SNPs and 5 indels of a single nucleotide. 2 indels are located in introns, and 3 indels are found in the 3'UTR of the *SRF2* gene. 10 of the SNPs are situated in introns whereas 9 SNPs are located in exons. The three of the exon SNPs cause a premature stop codon while 6 of them cause a change in the amino acid residues. According to these results, *SRF2* may not have the same function in Col and in Ler. We found only two substitutions for the *SRF2* gene in PERLEGEN substitution results and only one of them was common with Monsanto sequencing results.

Monsanto Ler ecotype partial sequencing results were obtained with the whole genome shotgun sequencing method that might create false positives, especially for insertion-deletion prediction (Venter *et al.*, 1996). Moreover, comparison of *SRF1* polymorphisms, which were checked by sequencing in our studies, with PERLEGEN substitution results does not reflect the real number of polymorphisms (see Appendix A). We detected 78 polymorphisms in *SRF1* gene whereas PERLEGEN results showed only one polymorphism. Taken together, our results indicate that none of the Ler sequencing results available in public database can yet depict the real number of polymorphism between Col and Ler background.

### 3.3 Discussion

The aim of the first part of the study was to gain full-length cDNA sequence information, including 5' and 3'UTR, of all *SRF* family members. The gained sequence information of all *SRF* members was comprehensively analyzed and

characterized at the nucleotide as well as at the amino acid level to gain knowledge on the structure and function of these genes.

### **3.3.1 Generation of full-length cDNA clones of all SRF members**

At the beginning of this study, no full-length sequence information for the *SRF* family was available. EST sequence information was screened and used to generate full-length cDNA clones via different approaches. Methodological aspects of this process are discussed in the following.

#### **3.3.1.1 EST**

Asamizu and coworkers (2000) showed that the average insert length of the EST clones was 1.28 kb. In addition, a 5' sequenced cDNA clone containing a translation initiation codon is annotated as full-length as the ESTs are synthesized from the 3' poly A tail. However, the complete 5' end of the generated ESTs is rarely analyzed by sequencing. Moreover, ESTs including a translation initiation site do not necessarily provide correct 5'UTR information that is important for the structural analysis of a gene. The results of this work confirm that it is almost impossible to obtain full-length EST clones. However, ESTs are valuable tools for cloning and characterization of genes and they provide information about the distribution of expressed (*SRF*) genes in various tissues and different developmental stages. Taken together, ESTs are valuable sources to identify genes with tissue-specific expression.

#### **3.3.1.2. Generation of *SRF* full-length cDNAs by RACE approach**

We used the RACE method to obtain *SRF* full-length sequence. By using this approach, the probability to generate the full-length 5'UTR and 3'UTR of the respective genes is increased. However, it is difficult to define the 5' end of the eukaryotic genes. To find out the exact structure and evolutionary relationship between sequences, RACE results can be used. However, in the RACE procedure, the action of T4 DNA polymerase may remove some nucleotides from the 5' end of the cDNA. Furthermore, certain secondary structures may also block the action of RT or *Taq* DNA polymerase to some extent. For this reason, to obtain the maximum possible amount of 5' end sequence of the gene, we checked the length of at least 60 5'RACE clones and sequenced the

longest one or two products. Moreover, the average size of the 5'UTR of *Arabidopsis* genes is about 131 bp. All of our 5'RACE products were longer than this value indicating that 5'UTR of the *SRF* genes could be in correct size. However, there is no method to guarantee a full-length cDNA especially with respect to the 5' end. To determine the complete 5' end of the gene the combination of RNase protection assay, primer extension assays, and cDNA or genomic sequence information is required.

Putative full-length cDNAs of the *SRF* members were successfully obtained by using four different approaches. Our approaches to create full-length *SRF* cDNA constructs included traditional restriction enzyme cloning, overlapping PCR, asymmetric PCR and end-to-end PCR. Overlapping PCR can be employed to fuse fragments more precisely and quickly than traditional restriction enzyme cloning. Therefore, overlapping PCR represents a more convenient method than traditional cloning. However, overlapping PCR is limited to the fusion of two pieces of DNA in one approach. The more powerful method called asymmetric PCR allows the fusion of more than two fragments in a one-step PCR reaction. This method was successfully applied to produce recombinant DNA by using three cDNA fragments of *SRF4*.

### **3.3.2 Sequence analysis of *SRF* genes**

Although genes of eukaryotic organisms are often interrupted by introns, they are spliced out during the transcription process. Currently, there are no appropriate algorithms available to find out the entire correct full-length sequence of eukaryotic genes (Lodge *et al*, 2007). Because genomic DNA of eukaryotic organisms possesses introns, it is difficult to identify the correct sequence of the entire gene. Even if the location of the beginning and the end of a gene are known, it is necessary to detect the exon-intron boundaries to derive the sequence of the protein-encoding region. Therefore, it is easier to identify amino acid sequences by using cDNA sequence information.

The amino acid sequences of the *SRF* members were deduced from the isolated cDNA sequences. Detailed sequence analyses of the *SRF* members in this study showed that they consist of typical elements of LRR-RLKs: an extracellular domain possessing leucine rich repeats (LRR), a transmembrane domain, and an intracellular part bearing the kinase domain. The similarities and

the differences of the single elements to other RLKs as well as the putative relationship to functional aspects are presented in the following.

### 3.3.3 Extracellular domain analysis of SRF members

The extracellular domains of RLKs, putative ligand binding domain, are highly different among members (Johnson and Ingram, 2005). Elevated levels of divergence in the extracellular domains of LRR-RLKs enable the interaction with diverse array of ligands.

#### 3.3.3.1 Leucine-rich repeats (LRRs)

The numbers of LRRs are different among different LRR-RLKs. For example, CLV1, ERECTA, BRI1 have more than 20 LRRs while SERK, and PRK1 have five LRRs (Torii and Clark, 2000). LRRV/SRF members have six LRRs. It is known that LRR domains of CLV1 RLK from *A. thaliana* are important for its biological function (Clark *et al.*, 1997). For example, G→D substitution in *clv1-4* results in a severe phenotype. It was suggested that these missense mutations effect the proper ligand receptor interactions or, alternatively, may interfere with proper receptor dimer formation. It was shown that LRR domains of Xa21 and Cf are related to ligand binding (Wang *et al.*, 1998; Thomas *et al.*, 1997). Recently, Dunning and coworkers (2007) showed that specifically LRRs 9 to 15 mediates binding of bacterial flagellin to FLAGELLING SENSING2 (FLS2) LRR-RLK. The investigation of the LRRs domain of the 14 members of LRRII RLK family revealed that certain amino acid residues differences promotes the ligand binding specificities causing functional difference of the RLKs (Zhang *et al.*, 2006). These results indicate that the different LRR domains of each RLK may provide interaction with a wide variety of ligands.

The number of the LRRs in SRF is different from the functionally characterized LRR-RLKs. Therefore, SRF members probably have different ligand specificities than other LRR-RLKs. Furthermore, as sequence conservation among SRF LRRs is low, LRRs of different SRFs may exhibit modified or different ligand specificities and different biological activities.

### 3.3.3.2 PEST motifs of SRF members

In addition to LRRs, certain SRF proteins possess other motifs in the extracellular domains such as PEST sequences. PEST sequences are found in cellular proteins like metabolic enzymes, transcription factors, protein kinases, protein phosphatases, and cyclins (Rechsteiner and Rogers, 1996). It was demonstrated that many rapidly degrading proteins have PEST regions, which contribute to their degradation (Chen and Clarke, 2002). Rechsteiner and Rogers (1996) showed that PEST domains are required for the degradation of a protein via different degradation pathways such as ubiquitine-proteasome. T-P and S-P amino acid pairs appear to represent general targets for protein kinases that initiate degradation. In addition, Yaglom and coworkers (1995) showed that phosphorylation of CLN3, the yeast G1 cyclin, precedes degradation. Moreover, Kornitzer and coworkers (1994) observed that a single amino acid mutation, T105 or P106 in the PEST region of GCN4, a yeast transcriptional activator, inhibits the degradation of the protein. We also found T-P or S-P amino acid pairs in SRF1A/SRF1B, SRF4, SRF6, and two PEST sequences of SUB. However, until so far we have not found evidence for the degradation of these SRF proteins bearing PEST sequences. It was also found that PEST motifs can be involved in protein-protein interactions, e.g. in ligand recognition. In humans, the PEST sequence of CFTR (CYSTIC FIBROSIS TRANSMEMBRANE CONDUCTANCE REGULATOR) apparently does not play a role in protein degradation (Chen and Clarke, 2002). PEST sequence might play a role in the maturation of the protein. Reverte and coworkers (2001) found that the CPEB (cytoplasmic polyadenylation element binding protein) PEST domain from *Xenopus* oocyte may have dual activity. The proteasome pathway degrades CBEP via the recognition of the PEST domain but this protein also persists for the promotion of polyadenylation. Although the exact mechanism remains unknown, this may be achieved through its PEST domain. However, until so far it is not clear whether PEST sequences of SRF proteins might be related to protein turnover. One approach to find out whether PEST sequences of SRF members have roles in protein degradation would be the examination of the SRF proteins with the help of EGFP-mediated protein visualization.

### 3.3.3.3 Proline-rich region (PRR)

It is well known that proline-rich region (PRR) have an important function in protein-protein interaction. Anggono and Robinson (2007) showed that Dynamin I which mediates vesicle fission during synaptic vesicle endocytosis, binds to the Src-homology 3 (SH3) domain of a subset of proteins via PRR. During this binding, a PxxP core motif flanked by basic amino acids on one side or the other has been shown to play an important role. PxxP motifs create a triangular structure, a left-handed polyproline type II (PPII) helix conformation that provides a proper binding site for SH3. Mutation of the two prolines in the PxxP motif reduced binding ability of Dynamin I protein.

When we close look at PRRs of the SRF members, SRF1A/SRF1B, and SUB PRR exhibit a PxxP motif. It can be postulated that this motif may provide a binding site for ligands in the frame of protein-protein interactions. To verify this hypothesis, first of all the extracellular interacting partners of the SRF members should be detected. Understanding the molecular basis of binding by PRR or other interaction domains would provide invaluable insights into the organization and regulation of the protein networks that are mediated by these domains.

### 3.3.3.4 Paired Cysteines

The LRR domains of plant RLKs are often flanked by paired cysteines spaced by six amino acids (Torii and Clark, 2000). However, all SRF family members have only one cysteine pair that is located at the strictly conserved position in the SUB domain. They lack a second cysteine pair situated downstream of LRR. In literature, there are examples of RLKs having only one pair of cysteines like SERK, which does not possess N-terminal cysteine pairs (Torii and Clark, 2000). Among SRF family members only SUB has a second cysteine pair located after the TM domain. Site directed mutagenesis of SUB paired cysteines located upstream of LRR and downstream of TM indicated that cysteine pairs are important for SUB function (Dr. Martine Batoux, personal communication). These results suggest that paired cysteines located before the LRR domain can also be important for the function of the other members of the SRF. This hypothesis can be approved with the mutation of the paired cysteines. Moreover, SRF family members have eight amino acids instead of six amino

acids between cysteine pairs with a **CxxWxGxxC** instead of a **CxWxGV(S/T)C** consensus sequence (x represents any amino acids). Other RLKs such as LePRK1 and LePRK2 have a paired cysteine spaced by 13 amino acids (Muschiatti *et al.*, 1998) as well as Xa21 second paired cysteines (Torii, 2004). This shows that paired cysteines are not always spaced with 6 amino acids. FEA2, maize ortholog of FLS2, has only a second pair of cysteines, which is sufficient for receptor dimerization. Therefore we can conclude that, although the SRF family - with the exception of SUB - exhibits only one paired cysteine, this may be sufficient to build receptor dimers.

Paired cysteines probably take part in the establishment of disulfide bonds to form a dimer between RLKs. Cysteines of CLV1 can form an intermolecular disulfide bond most probably with CLV2 (Trotochaud *et al.*, 1999; Torii, 2004). It can be suggested that SRF proteins may form homo- or heterodimers with the help of the paired cysteines. Therefore, the co-receptor of the SRF members could be an RLK like in BRI1/BAK1 example both of which are RLKs or alternatively co-receptor might be a RLP such as CLV1/CLV2 (Nam and Li, 2002; Li *et al.*, 2002; Clark *et al.*, 1997). CLV1 is a RLK and CLV2 is a RLP and they control the cell differentiation at the shoot meristem.

### 3.3.4 Transmembrane (TM) domain

In addition to cysteine pairs there is also dimerization motif in the TM domain. The GxxxG, lxxxI and GxxxT (x represents any amino acid residues) dimerization motifs are found in certain SRF TM domains. It has been shown that these motifs are involved in protein-protein interactions via intermolecular hydrogen bonds (Curran and Engelman, 2003). SRF1A, SRF1B, SRF4, SRF5, and SRF6 have an lxxxI motif, SRF5, SRF6, SRF7, SRF8 have a GxxxG motif and SRF2 and SUB has a GxxxT motif indicating that there could be protein-protein interaction via these motifs. However, SRF3 does not possess any transmembrane domain dimerization motif.

### 3.3.5 Intracellular domain analysis of SRF members

The SRF family members possess an intracellular juxtamembrane domain, a kinase domain, and a C-terminal extension, except for SRF4 and SUB (Figure 3.7).

### 3.3.5.1 Juxtamembrane domain

The juxtamembrane domain is located between the TM and kinase domain. Interestingly, there is no sequence similarity among SRF members in this domain. Although the function of this domain is unknown, recent studies have indicated that this domain could contain a putative serine/threonine phosphorylation site as well as PEST and other cleavage motifs (Nühse *et al.*, 2004). Ser452 (S452) in SRF3 and S379 in SRF7 were shown as putative serine phosphorylation site. In addition, S452 is found in the PII $\underline{S}$ PERP cleavage sequence. XA21 is a rice LRR-RLK which plays a role in pathogen resistance against *Xanthomonas oryzae* pv. *oryzae* (Xu *et al.*, 2006). Recent studies have been showed that XA21 possesses three autophosphorylation sites, located in the JM domain, positions S686, T688, and S689 of the amino acid sequence. Moreover, these putative phosphorylation sites are located in a proteolytic cleavage site (485PSRTSMKG). It is also known that many receptor tyrosine kinases have a phosphorylation site in the JM domain and it was suggested that autophosphorylation of the S and T in this cleavage site might prevent cleavage of XA21 (Hubbard and Till, 2000). As SRF3 possesses the same structure, the same scenario might be suggested for SRF3. Furthermore, SRF3 has a proline-rich region and a PEST sequence in this domain indicating that SRF3 may undergo the ubiquitination degradation process. eGFP tagged SRF3 ORF analysis *in planta* will investigate this hypothesis.

Serine and threonine (S474 residue SRF3; and T482 residue of SRF3) residues that are designated as rectangular in Figure 7 for each SRF member are suggested as probable phosphorylation targets although they do not exhibit a similarity with known phosphorylation motifs.

### 3.3.5.2 Kinase Domain

In contrast to the JM domain, all SRF members display amino acid sequence similarities in their catalytic domains and possess the eleven conserved subdomains (Hank and Quinn, 1991). However, closer investigation of the amino acid sequences of the SRF members revealed certain amino acids differences among the SRF members.

All SRF members have a highly conserved GxGxxGxVY consensus in kinase subdomain I. In this segment, which is also known as a glycine rich loop,

the function of the highly conserved valine consists of anchoring ATP that is required during phosphorylation. As all SRF members also exhibit this conserved valine, the ATP anchoring process can accurately take place. However, in subdomain II all SRF members have a highly conserved alanine while SUB exhibits a valine (V542 in the SUB amino acid sequence) substitution. This alteration is the same like in the pollen specific RLKs PRK1, LePRK1, and LePRK2. In another example, the highly conserved lysine in kinase subdomain II that is important for the transfer of phosphate is substituted with arginine in SRF2 and SRF1A. It is shown that a mutation in this residue causes loss of kinase activity (Hanks *et al.*, 1988).

Detailed comparison of kinase domains indicates that there are notable differences in a stretch of residues flanked by kinase subdomains II, III, and a region known to be variable between different protein kinases in the activation segment (Hanks and Quinn, 1991; Scheef and Bourne, 2005). The subdomains II and III are important for ATP binding and the activation segment is required for substrate binding (Johnson *et al.*, 1996). Alignment of the activation loop showed that SRF2, SRF8 and SUB have unique activation segment sequences. However, this segment is slightly more conserved within the SRF1/SRF3 and SRF6/SRF7 pairs and to some extent within the SRF4/SRF5 pair. But the activation segment sequences are still clearly distinct from each other. Comparison of the activation segment of other RLKs, BAM1/BAM2/BAM3, ERBB1/ERBB2/ERBB3/ERBB4, and ERECTA/ERL1/ERL2, showed more conserved sequences indicating the probability of redundancy between RLKs (Data not shown) (DeYoung *et al.*, 2006; Yarden and Sliwkowski, 2001; Shpak *et al.*, 2004). However, the distinct activation segments of SRF members may indicate diversity in substrate recognition and therefore diversity in function.

#### **3.3.5.2.1 Do SRF family members possess an inactive kinase domain?**

The mechanism of the plant receptor-like kinases, similar to animal receptor kinases, includes signal perception causing receptor dimerization, followed by intermolecular phosphorylation which triggers the phosphorylation of downstream signaling proteins. However, recent studies have shown that some of the RLKs have a inactive kinase domain, called atypical kinases, indicating phosphorylation-independent mechanisms. Most of the atypical kinases lack the

conserved lysine (K) in subdomain II, aspartic acid (D) in subdomain VIb, or aspartic acid in subdomain VII (Castells and Casacuberta, 2007).

*Arabidopsis* TMKL1 is an example of atypical RLK displaying a dead-kinase domain (Valon *et al.*, 1993). This protein also has a substitution in conserved amino acid residues. For instance, TMKL1 has, a Asp→Asn (D→N) substitution in subdomain VIb, and DFG→DVY substitutions in the activation loop indicating that TMKL1 can be an atypical receptor kinase.

SUB, an atypical kinase, has D625→N and N630→K substitutions in subdomain VIb. Genetic experiments indicated that catalytic activity is not necessary for SUB function (Chevalier *et al.*, 2005).

Because studies revealed that SUB is an atypical receptor-like kinase, for all SRF members, homologs of SUB, amino acid sequences were examined. There are two conserved serine/threonine indicative sequences in plant RLKs, **DΦKxSN** in kinase subdomain VIb and **GTxGYΦAPE** in subdomain VIII, respectively, which are autophosphorylated during the autoactivation of the receptor (Torii and Clark, 2000). The aspartic acid in **DΦKxSN** is very crucial because it is the acceptor of the proton from the substrate at the phosphorylation site. SUB bears two substitutions, which are decisive for the catalytic activity. The first one is the substitution of D625→N. SUB and all SRFs except SRF2 carry an asparagine instead of aspartic acid at position 625 of SUB. To the contrary, plant RLKs which exhibit kinase activity, carry an Asp at the position. However, the second serine/threonine indicative sequence in subdomain VIII (GTXGYΦAPE) is not highly conserved in SRF family. Because of these substitutions it can be hypothesized that the serine/threonine signature is destroyed. Moreover, the highly conserved DFG motif (643-645 amino acid residues in SUB) in activation loop is changed to Dc/s/yF among SRF members. All members of SRF have phenylalanine substitution in DFG motif of subdomain VIa which is involved in cation binding and orientation of the ATP gamma phosphate for phosphate transfer (Hanks and Hunter, 1995). SRF1, SRF2, and SRF3 have a F→C substitution, SUB, SRF6, and SRF7 have F→S alteration and SRF4, and SRF5 have a F→Y substitution in the DFG motif.

Detailed investigation of amino acid sequences of the SRF members suggests that SRF family members, except SRF2, might be atypical kinases. Furthermore, Castells and Casacuberta (2007) showed by means of

phylogenetic analysis of the kinase domain of the LRR-V family members that the substitution of the aspartic acid residue may have occurred after the SRF2 gene duplication giving rise to the eight genes possessing the same mutation. This interpretation is not completely correct as the phylogenetic analyses performed in this study as well as by Castells and Casacuberta (2007) clearly show that, as SRF2 is located basal to the other SRFs, the respective mutation should have occurred after the gene duplication leading to the divergence of SRF2 and the other SRFs.

If the SRF family members are atypical receptor like kinases, there might be two mechanisms for their activity. In the first case, like in ErbB3 RLK, ligand binding to most probably the heterodimer of SRFs and other RLKs causes the phosphorylation of the SRF kinase domain creating a docking site for downstream signaling proteins (Stein and Staros, 2000). Alternatively, binding of ligand to SRF members may cause a conformational change creating a kinase domain that may interact with downstream signaling protein. For the maize atypical receptor kinase MARK (maize atypical receptor kinase) RLK this scenario was demonstrated (Llompert *et al.*, 2003; Castells *et al.*, 2006). Binding of ligand to MARK mediates the conformational change which allows the interaction between MIK (Mark Interacting Kinase) and its kinase domain. In addition, interaction of MARK and MIK generates a conformational change in MIK and induces the kinase activity of MIK.

To investigate whether SRF members possess inactive kinases, *in vitro* kinase activity assay can be performed by using bacterially expressed fusion proteins. Moreover, we can use genetic approach. To this end, the crucial amino acid residues in the kinase domain of SRF4 can be changed by site-directed mutagenesis. Then, the ability of the each altered construct to rescue the mutant phenotype can be investigated. However, we have not visualized phenotype for the other *SRF* members. Therefore, at the moment second approach is not helpful for the other SRF members.

#### **3.3.5.2.1.1 Phylogenetic Analysis of SRFs**

The inclusion of kinase sequences of other monocot as well as dicot as well as of primitive plant organisms showed that the SRFs belong to a group of kinases with a monophyletic and very early origin in plant evolution. SRF2 is the most

primitive and basal of the SRF members. SUB, SRF1 and SRF3 are located in cluster of kinase sequences which diverged from a cluster containing SRF4, SRF5, SRF6, and SRF7 before the split of monocots and dicots. In contrast, while SRF8 and SUB are the results of earlier gene duplications, SRF4/SRF5, SRF6/SRF7 and especially SRF1/SRF3 appear to be the result of relatively recent gene duplications. This is also reflected by a relatively high sequence identity in this SRF pairs. Further sequences are necessary to decide the question whether these SRFs are specific for Brassicaceae. Another interesting question is whether the sequences of the other monocot and dicot species code for putative kinase-dead domains as well.

The assessment of the substitution patterns shows that evolutionary constraints are not homogenous among both different SRFs and different domains of one SRF sequence. The LRR and especially the SUB domain have an elevated rate of nonsynonymous substitutions and therefore evolve faster than the kinase domain. The homogenous substitution pattern found for the kinase domains probably reflects the constraints acting on this domain to preserve the function which may be similar or identical in all SRF members. In contrast, the elevated rate of amino acid changes in the LRR and especially the SUB domain could be driven by a lowered selective pressure or repeated positive selection events which allow the accumulation of amino acids driving the adaptation of the protein to changing environmental conditions. Generally, it can be concluded that the diversification of the SRFs is rather driven by changes in the SUB and LRR domain than by changes in the kinase domain.

### 3.3.5.3 C-terminal extension

The carboxyl terminus (C-terminus) of the SRF members shows a distinct diversity (Figure 3.7). SUB and SRF4 lack a C-terminus extension while SRF5, the closest relative of SRF4, has a 23 residues extension and SRF2 has even a 40 amino acid residues extension. Moreover, the first 12 amino acids of SRF6 and SRF7 C-terminus extensions are conserved. Interestingly, two serine residues in C-terminus of SRF7 (position 710-711) and SRF8 (position 683-684) are phosphorylated in an *Arabidopsis* suspension culture system (Nühse *et al.*, 2004). Interestingly, 75% of the plant RLK phosphorylation sites are found either in the juxtamembrane or C-terminus of the kinase domain (Nühse *et al.*, 2004).

Although the function of the SRF member's C-terminal extension is unknown, there are a couple of studies showing the probable function of the C-terminal extensions of other genes. Wang and coworkers (2005) showed that removal of the BRI1 C-terminus leads to hypersensitive receptor activity, characterized by enhanced phosphorylation of BRI1. This result indicates that phosphorylation in the C-terminus of BRI1 plays a crucial regulatory role in its activation. Vatter and coworkers (2005) demonstrated that the C-terminal extension of the mouse G-protein-coupled receptor kinase 6 (GRK6) may be required for the autoregulatory function to control the activity of this kinase towards the receptor substrate. Moreover, Akama and Takaiwa (2007) presented that the rice Glutamate decarboxylase2 (OsGAD2), which converts L-glutamate to  $\gamma$ -aminobutyric acid, possesses a C-terminal extension binding to  $\text{Ca}^{2+}$ /calmoduline to modulate enzyme activity. Taken together, C-terminal extensions can have a distinct function according to the sequence of this segment. Indeed, we can suggest that SRF7 and SRF8 bear a putative phosphorylation site in C-terminal extension, while the role of C-terminal extensions of other SRF members is still unclear.

### 3.3.6 Is SRF1 under evolutionary selection?

Interestingly, a high number of polymorphisms were detected between Col and Ler genomic sequences of the SRF1 locus. In the 3,986 bp of the *SRF1* gene sequences, 78 polymorphisms are found between Col and Ler showing about 20 polymorphisms per 1 kb, although the average number of polymorphisms between two accessions is only 4 polymorphisms in 1 kb (Nordborg *et al.*, 2005). This result indicates that *SRF1* has a very high number of polymorphisms between two accessions. However, several *Arabidopsis* genes show high number of polymorphisms. *CLAVATA2* (*CLV2*), *APETALA3* (*AP3*) and *RPS2* and the gene cluster of nucleotide-binding site leucine-rich repeat (NB-LRR) disease resistance (R) genes are examples of genes having elevated levels of polymorphisms between the accessions (Jeong *et al.*, 1999; Purugganan and Suddith, 1999; Caicedo *et al.*, 1999; Noel *et al.*, 1999).

The *RPP5* gene is a member of a clustered multigene family encoding NB-LRR proteins conferring resistance to the pathogen *Peronospora parasitica* (Noel *et al.*, 1999). Comparison of 95 kb of DNA sequence carrying Ler *RPP5*

haplotype and corresponding 90 kb Col sequences revealed that Ler and Col haplotypes exhibit an extraordinary degree of polymorphisms. Especially, LRR residues have a high number of polymorphisms since this region is very important for R gene and avirulence (Avr) protein interaction. For this reason, R genes probably evolve novel Avr recognition capacities with polymorphisms in the LRR region.

The *CLV2* locus exhibits also a high level of polymorphisms between sequences from Col, Ler and Wassilewskija (Ws) genetic backgrounds (Jeong *et al.*, 1999). In the 2163 bp of the *CLV2* gene between Col and Ler background 68 polymorphic sites were found.

Because of the high level of the polymorphisms, *SRF1* appears to be under positive selection. In addition, *SRF1* is not located in a highly polymorphic genomic region (Eyüboğlu *et al.*, 2007). Furthermore, a comparable high level of polymorphisms as in *SRF1* cannot be seen in other accessions such as Ws, although *CLV2* has elevated polymorphisms in three accessions. The ratio of non-synonymous over synonymous substitutions, the  $K_a/K_s$  ratio, may indicate the balancing selection. A ratio higher than 1 is an indicator of positive selection. This ratio for the entire *SRF1* protein is 1.25 (25 non-synonymous/20 synonymous residue substitutions). In addition, non-synonymous changes are not equally distributed in the entire *SRF1* protein. The  $K_a/K_s$  ratio is 0.29 for the extracellular domains of *SRF1* whereas this ratio is 1.85 for intracellular domains including the JM, kinase domain and C-terminal region. Taken together, *SRF1* polymorphisms mainly affect the intracellular domains of the *SRF1* protein and *SRF1* has undergone high selection.

### 3.3.7 The *SRF1* gene encodes two isoforms

Although the probability of the alternative splicing is not predicted in the database, we found out that *SRF1* mRNAs potentially code for two proteins called *SRF1A*, an RLK, and *SRF1B*, an RLP having only a short intracellular cytoplasmic tail lacking kinase domain. Recent studies have indicated that alternative splicing events in plant species are not as low as suggested in earlier studies (Wang and Brendel, 2006). In *Arabidopsis* 21.8% of the 21,641 genes (EST/cDNA) show alternative splicing events. Similarly, in rice 21.2% of investigated EST/cDNA have alternative splicing. Furthermore, 56% of the

alternative splicing events of the *Arabidopsis* genes are intron retentions (intronR) whereas for the human intronR is rare type of alternative splicing. Contrary to the mammalian exon definition mechanism in pre-mRNA splicing, in plants introns are probably recognized by intron definition mechanisms rather than exon definition mechanism. Therefore, the failure during recognition of intron might cause intron retention that is highly seen in plant (Wang and Brendel, 2006). Therefore, the generation of SRF1B by intron retention is not a rare event in plants. Similar to SRF1, MIK (Mark Interacting Kinase), a maize gene encoding for a kinase related to the GCK subgroup of MAPK4K, has undergone alternative splicing (Castells *et al.*, 2006). Three of the four splice variants were generated by intron retention.

In general one third of the intronR splicing generates premature stop codon (Wang and Brendel, 2006). The *SRF1B* splice variant creates a premature stop codon creating an RLP. In the *Arabidopsis* genome, 56 genes are predicted to encode RLPs including *TOO MANY MOUTHS (TMM)*, *CLAVATA2 (CLV2)*, and *RPP27* (Fritz-Laylin *et al.*, 2005). However, the functions of most of them are unknown. To date, the functions of only three of the *Arabidopsis* RLPs have been identified. *CLV2* is involved in meristem and organ development (Jeong *et al.*, 1999). *TMM* has a function in stomatal patterning and epidermal development (Nadeau and Sack, 2002). The recently identified *RPP27* is involved in disease resistance (Tör *et al.*, 2004). Moreover, more RLPs with known function are identified in other species like Cf-9 in tomato or Xa21D in rice, both of which are involved in disease resistance (Hammond-Kosack and Jones, 1997; Wang *et al.*, 1998).

Interestingly, RLPs such as *CLV2*, *TMM* and *Xa21D* form heterodimer with RLKs for their function. For instance, *CLV2* forms a heterodimer with LRR-RLK *CLV1* and acts in the same pathway (Kayes and Clark, 1998; Jeong *et al.*, 1999). Moreover, rice RLP *Xa21D* lacking TM and kinase domain might form a heterodimer with the one of the LRR-RLKs of the same family of *XA21* (Wang *et al.*, 1998). According to examples, we might suggest that *SRF1B* as an RLP may also form heterodimer with one of the RLK.

In contrary to *SRF1A*, *SRF1B* cytoplasmic tail possesses an Yxx $\Phi$  motif that may stimulate receptor mediated endocytosis. This motif is also found in 9 of the 56 RLPs of *Arabidopsis* (Fritz-Laylin *et al.*, 2005). It was shown by means

of site-directed mutagenesis that this motif is necessary for the function of tomato LeEIX1 (ethylene-inducing xylanase) and LeEIX2 RLPs (Ron and Avni, 2004). The authors proposed that binding of the EIX ligand to LeEIX2 protein might induce receptor-mediated endocytosis allowing the interaction with cytoplasmic proteins to generate a signal to induce plant defense. Moreover, the tomato Ve2 resistance RLP has a Yxx $\Phi$  motif in cytoplasmic tail (Kawchuk *et al.*, 2001).

To date, we have no data to hypothesize that SRF1A and SRF1B act in the same pathway or not. However, we showed that *35S::SRF1A* overexpression lines show seedling lethality while *35S::SRF1B* lines do not (see chapter 4). Taken together, it can be concluded that SRF1A and SRF1B are two splice variants, which may take part in different biological roles.

## Chapter 4. Functional analysis of the STRUBBELIG RECEPTOR FAMILY members

### 4.1 Introduction

#### 4.1.1 Anther Development

The *Arabidopsis* flowers consist of four whorls of organs, sepals, petals, stamens and carpels (Scott *et al.*, 2004). The third whorl of the flower possesses six stamens, four long and two short ones. The stamen is composed of the anther, the site of pollen development and the filament, which transmits water and nutrition to the anther (Sanders *et al.*, 1999). Anther development consists of two phases including 14 stages. The phase 1 comprises stages 1 to 8 and the phase two stages 9 to 14. Stage 1 anther is composed of three cell layers, L1, L2 and L3. The archesporial cells originating from L2 layer divide periclinally to form primary parietal and primary sporogenous cells. Further division of each layer creates secondary parietal layers. The differentiation of the secondary parietal cells leads to the formation of the endothecium, middle layer, tapetum, which surrounds the microspore mother cells, derived from sporogenous cells. Briefly, cell division takes place in the anther primordia for the formation of bilateral structures with locule, wall, connective, and vascular regions during stages 1 to 4. The L3 layer contributes to the formation of vasculature and connectives (Scott *et al.*, 2004). All anther cell types are present by stage 5. The formation of pollen starts with the meiosis event that occurs in microspore mother cells generating tetrads of haploid microspores (Sanders *et al.*, 1999). Each tetrad has a callose wall that is deposited after the second meiotic division around all tetrahedrally arranged microspores (McCormick, 2004). Then, microspores, released from the tetrad with the degeneration of the surrounding callose wall through the action of callase produced by the tapetum, are differentiated into three-celled pollen grains during stages 9 to 12. At stage 10, tapetum cell layer starts to degenerate. Finally, pollen grains are released with the degeneration of the several cell layers during dehiscence (Sanders *et al.*, 1999).

#### 4.1.1.1 Genes involved in anther development

Three classes of homeotic genes called A, B, C control the floral organ identity (Scott *et al.*, 2004). Stamen development is controlled by the B and C classes of genes, *APETALLA* (*AP*), *PISTILLATA* (*PI*), and *AGAMOUS* (*AG*), respectively. *SEPELLATA* (*SEP*) genes encode transcription factors that are also required for the full function of the homeotic genes. Mutation in one of these genes causes the conversion of the third whorl organs to another type.

Expression studies of the rice anther genes revealed that most of the genes expressed in anthers are related to protein, starch and sucrose metabolism, osmoregulation, cell wall biosynthesis and expansion, sugar transport, lipid transfer, flavonoid synthesis, and cytoskeleton structure (Scott *et al.*, 2004). The results of microarray analysis indicates that in pollen mRNA encoding signal transduction and cell wall synthesis proteins are highly present whereas transcription and translation proteins are underrepresented (McCormick, 2004).

#### 4.1.1.2 LRR-RLKs play roles in anther and pollen development

Recent studies revealed that a number of the plasma membrane localized LRR-RLKs play roles in the control of the cell fate during early anther development, e.g. *EXCESS MICROCYTOPROCYTES 1/EXTRA SPOROGENOUS CELLS* (*EXS/EMS*), *SOMATIC EMBRYOGENESIS RECEPTOR KINASE1* and *2* (*SERK1* and *SERK2*) (Canales *et al.*, 2002; Albrecht *et al.*, 2005; Colcombet *et al.*, 2005). In addition, the putative transcription factor *NZZ/SPL*, and *TAPETUM DETERMINANT1* (*TPD1*), a small secreted protein, are known proteins having a function during early anther development (Schiefthaler *et al.*, 1999; Balasubramanian and Schneitz, 2000; Yang *et al.*, 2005).

*EXS/EMS* encoding a LRR receptor-like kinase possesses an important function during anther development. *Exs/ems* mutants that generate extra meiocytes and lacks tapetal cell layer affecting anther wall organization exhibits male-sterility (Canales *et al.*, 2002; Zhao *et al.*, 2002). *SERK1* and *SERK2* that encode LRR-RLKs function redundantly in anther development and male gametophyte maturation (Albrecht *et al.*, 2005; Colcombet *et al.*, 2005). Although the single mutant of the either gene does not exhibit any phenotype, double mutant *serk1 serk2* plants are male-sterile because of the defect in

tapetum specification and formation of additional sporocytes. An increased number of the sporocytes display normal development till the tetrad stage, after which meiocytes degenerate causing the male-sterility. *TAPETUM DETERMINANT1 (TPD1)*, which encodes a predicted small-secreted protein, has a role in the specialization of the tapetal cells (Yang *et al.*, 2005) because *tpd1* mutant causes the inner secondary parietal cells to form microsporocytes instead of tapetal cells. Furthermore, overexpression of *TPD1* causes the formation of short and wide siliques, enlarging the tapetal cells via coordination with *EMS1/EXS*. As a result, tapetum cells are not separated from microspores and do not degenerate at the correct time. Due to the delay of the tapetum degeneration, plants are male-sterile.

It is suggested that *SERK1* and *SERK2* together with *EXS/EMS* may play a role in the signal perception during the development of microsporocytes to immature tapetal cells. Moreover, *TPD1*, a secreted protein, may have a role in these pathways with *EMS/EXS* and *SERK1/2*. *TPD1* may represent the ligand specification (Albrecht *et al.*, 2005).

Another example of the LRR-RLK is *BARELY ANY MERISTEM1 (BAM1)* and *BAM2* that have a redundant or overlapping function in early anther development (Hord *et al.*, 2006; De Young *et al.*, 2006). The investigation of *bam1 bam2* double mutant anthers showed that they exhibit an abnormal development at the very early stages leading to a lack of endothecium, middle layer and tapetum layers whereas single mutants do not display any phenotype, which is indicating a redundancy in function. This abnormality causes the degeneration of the pollen mother-like cells before the completion of meiosis. Closer investigation of the defect revealed that the reduced number of the L2-derived cells in double mutant anther is due to the decreased cell division. Being LRR-RLKs, *BAM1* and *BAM2* have a role in cell-cell communication mediating a developmental signal for the differentiation of the archesporial cells at stage 2 during normal anther development. Besides anther development, *BAM1* and *2* have a function in meristem size, leaf size and shape formation, and female fertility.

*POLLEN RECEPTOR-LIKE KINASE1 (PRK1)*, an LRR-RLK from *Petunia inflata*, plays a role in male-fertility (Lee *et al.*, 1996). Downregulation of *PRK1* by means of an antisense *PRK1* transgene causes a male-sterile

phenotype due to pollen abortion. It is suggested that PRK1 might play a role in the signal transduction pathway mediating the post meiotic development of microspores. Moreover, PRK1 may also be related to pollen tube growth and fertilization, because it is very highly expressed in mature pollen. Furthermore, *Lycopersicon esculantum* LePRK1 and LePRK2 are related to pollen-pistil interaction, pollen interaction and pollen tube germination or growth (Muschiatti *et al.*, 1998). These results indicate that LRR-RLKs play a crucial role in position-dependent intercellular signaling events during anther development.

#### **4.1.2 Outlook on Chapter 4**

In this chapter, functional analysis of the SRF family members was performed by using two approaches, overexpression of *SRF* genes in Col-0 and *Ler* backgrounds and yeast-two hybrid analysis of SRF4, SRF5, and SRF6 to find out their putative interacting partners. Overexpression analysis of the *SRF* members revealed that *SRF2-5* and *SRF7* might be related to anther development. In addition, SRF4 is a direct regulator of leaf size (see chapter 5). Furthermore, overexpressed-*SRF1A* and *SRF8* transgenic plants displayed seedling lethality. Yeast two-hybrid assays provided further hints about the functions of the respective SRF genes. For instance, the identification of the putative interacting candidates of SRF4 supports its role in leaf size determination.

### **4.2. Results**

In order to identify the function of the SRF family members two approaches were used, overexpression analysis and yeast-two hybrid methods.

#### **4.2.1 Overexpression analysis of the *SRF* family members**

In order to understand the functions of the SRF family members, besides T-DNA insertion line analysis (Eyueboglu *et al.*, 2007), examination of the overexpression lines of the *SRF* cDNAs under the control of Cauliflower mosaic virus (CaMV) 35S promoter was performed. Each family member was ectopically expressed in Col, *Ler* wild type, and *sub-1* backgrounds. Transgene expression was analyzed for each gene in each background. When T1 lines

were selected on MS-kanamycin plates and grown on soil, all transgenic lines except *SRF1A* and *SRF8* showed normal vegetative development. The first visible defect was detected when flowers and siliques developed.

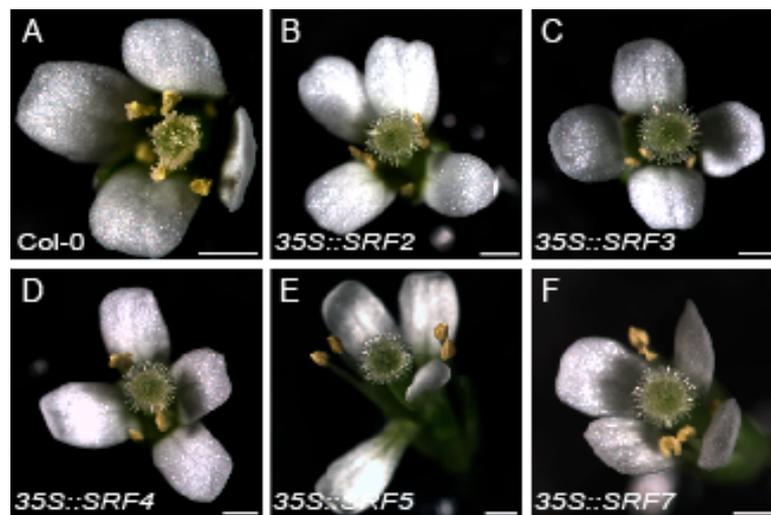
#### 4.2.1.1 Male sterility phenotype

Although, overexpressed lines developed a normal inflorescence architecture and normal floral structures, *35S::SRF2/3/4/5/7* T1 transgenic plants in Col background, interestingly, formed short siliques without seeds (Figure 4.1).

Detailed examination of flowers of these plants revealed that floral organs develop normally but stamens of the transgenic flowers lack pollen grains or sometimes possess less pollen grains than wild type flowers whereas ovules exhibit a normal appearance. In wild-type, however, the anthers at the height of the stigma dehisced and released pollen grains. These characteristics are typical for a male-sterility phenotype. Male sterility for *35S::SRF2/3/4/5/7* was visualized in 5, 10, 9, 12 and 12 transgenic plants out of 50 T1 transgenic lines in Col background, respectively (Figure 4.2).



**Figure 4.1:** Siliques of the ectopically expressed *SRF* lines. a: Wild type (Col-0), b: *35S::SRF1B*, c: *35S::SRF6*, d: *35S::SRF2*, e: *35S::SRF3*, f: *35S::SRF4*, g: *35S::SRF5*, and h: *35S::SRF7*. Scale bar: 1 mm.



**Figure 4.2:** Male sterile stamens of the ectopically expressed *SRF* lines flowers. 3.2A) A wild-type flower. 3.2B-F) Flowers of the overexpression lines. Overexpression lines flowers lack mature pollen on stamens and on stigma. Scale bars: 0.5 mm.

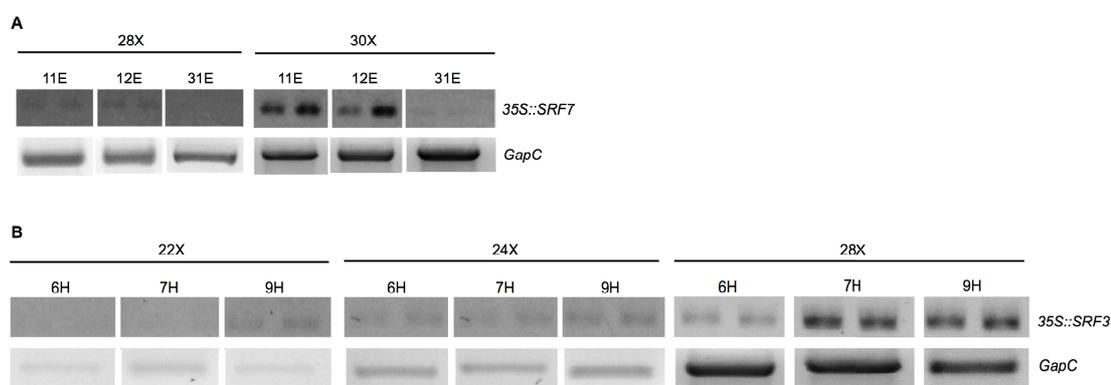
Besides different numbers in absolute male sterility, different levels of sterility can also be detected in the overexpression lines (Table 4.1). For the confirmation of the existence of a male sterility problem, sterile plants were fertilized with wild-type pollens. The sterile phenotype was rescued, allowing normal seed production. To understand whether the male sterility of the transgenic lines is related to the transgene expression level, RT-PCR analysis was performed. 11E and 12E, two RT-PCR positive lines of *35S::SRF7* transgenic plants, displayed absolute male sterility whereas 31E, also an RT-PCR positive line, did not exhibit absolute male-sterility. Semi-quantitative RT-PCR showed that 11E and 12E possess a higher transgene level than 31E. Contrary, comparison of the RT-PCR positive overexpression lines of *SRF3* plants (6H, 7H, and 9H) indicated that transgenic lines bearing absolute male-sterility (6H) possess a lower transgene expression level than the lines displaying male fertility (Figure 4.3). We found that increased transgene levels of the *35S::SRF4*, *35S::SRF5*, and *35S::SRF7* lines were correlated with male sterility whereas the *35S::SRF2* and *35S::SRF3* transgene expression level was not directly related to male sterility. Interestingly, overexpression lines of only *SRF3* and *SRF4* in *Ler* background possess male sterility, indicating the importance of genetic background in the experiments. Moreover, silique size analysis of the transgenic plants showed that in addition to *35S::SRF2/3/4/5/7*

lines, siliques of 3 transgenic *35S::SRF1A* line, 6 plants of *35S::SRF1B* line and 3 plants of *35S::SRF6* transgenic lines are shorter than Col wild-type siliques. However, absolute male sterility could not be detected for these transgenic lines.

**Table 4.1: Overexpression lines grouped by silique size.**

Silique size	14-15	11-13	8-10	5-7	3-4
Col-0	15	11	1		
<i>35S::SRF1A</i>	19	7	2	1	-
<i>35S::SRF1B</i>	24	18	2	4	-
<i>35S::SRF2</i>	7	23	8	7	5
<i>35S::SRF3</i>	6	19	9	7	10
<i>35S::SRF4</i>	10	17	8	4	9
<i>35S::SRF5</i>	3	22	5	8	12
<i>35S::SRF6</i>	11	36	2	1	-
<i>35S::SRF7</i>	1	26	8	3	12
<i>35S::SRF8</i>	5	15	-	-	-

Silique size is in mm. Plant numbers are indicated in each *35S::SRF1-8* overexpression T1 line.

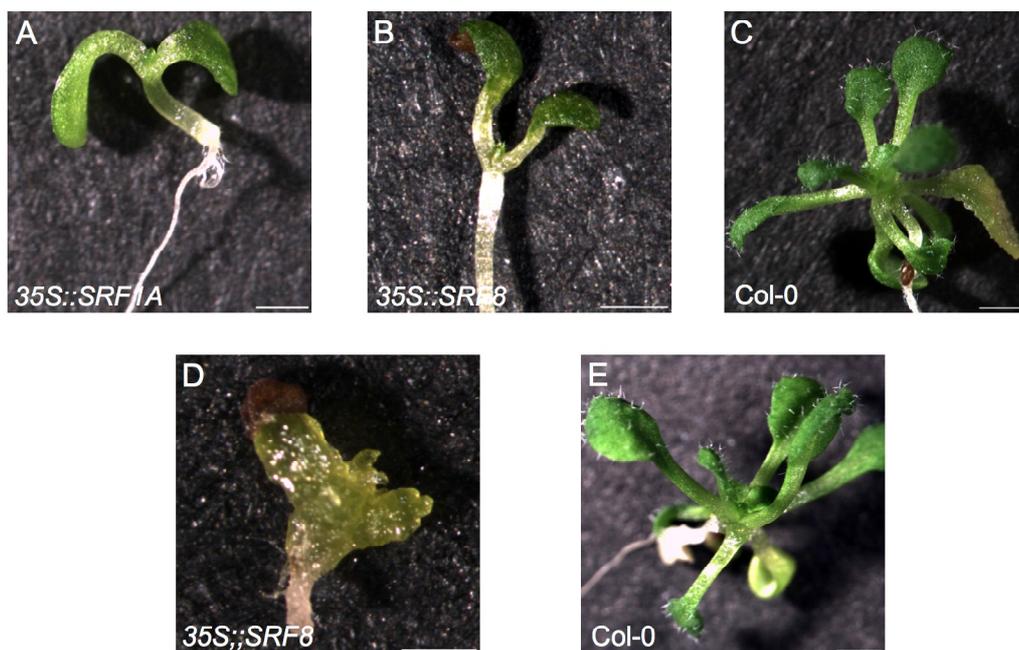


**Figure 4.3:** Semi-quantitative RT-PCR of *35S::SRF7* and *35S::SRF3* transgenic lines. A: 11E, 12E, 31E and represent RT-PCR of *35S::SRF7* transgenic lines and B: 6H, 7H, 9H of *35S::SRF3* transgenic lines. *GapC* is used as a control. 22X, 24X, 28X, 30X indicate the number of PCR cycles.

#### 4.2.1.2 Seedling lethality phenotype

In addition, T1 generation of *35S::SRF1A* and *35S::SRF8* transgenic in Col, Ler ecotypes and *sub-1* backgrounds led to plants which exhibited seedling lethality at the 2-cotyledon stage (Figure 4.4). Although *SRF1A* and *SRF8* overexpression lines exhibited seedling lethality, we could obtain 29 (Col), 38 (Ler), and 11 (*sub-1*) *SRF1A*-overexpressed transgenic plants and 20 (Col), 58

(*Ler*) and 20 (*sub-1*) *SRF8*-overexpressed transgenic plants in respective backgrounds. However, we could not detect any obvious phenotype for *35S::SRF1B* and *35S::SRF6* transgenic lines. Unfortunately, due to the normal anther development and fertility in the different *srf* T-DNA insertion lines it is presently difficult to decide whether or not the sterility and seedling lethality phenotypes relates to the wild-type function of the corresponding genes.



**Figure 4.4:** Seedling lethality of the overexpressed-*SRF1A* and *SRF8* transgenic plants in Col-0 background. 3.3A-B) T1 lines of the 17-day old transgenic plants possessing *35S::SRF1A* and *35S::SRF8* constructs, respectively. 3.3D) T2 line of the 17-day old overexpressed-*SRF8* transgenic plant. 3.3C,E) 17-day old Col-0 wild type plants. Scale bars of A-C,E are 1 mm, D is 0.5 mm.

#### 4.2.1.3 Examination of floral organs in *SRF* overexpression lines

We investigated floral organ numbers of flower stages 12 to 15 in *SRF* overexpression lines in Col and *Ler* background and wild-type Col and *Ler* plants (stages according to Smyth *et al.*, 1990). Five flowers per plant out of 50 plants were analyzed for each transgenic line. Table 4.2 and Table 4.3 display the average and standard deviation of the floral organ numbers of the respective *SRF* overexpression plants in Col background and *Ler* background, respectively. Comparison of these two parameters with the wild-type values by

means of two-tailed Student's t-test revealed that there is no statistically significant difference in floral organ numbers in the overexpression lines in Col and Ler background.

**Table 4.2: Average number of floral organ types in Col-0 wild-type and overexpression lines flowers in Col-0 background**

Genotype	Sepal Whorl	Petal Whorl	Stamens	Carpels	n	N
Col-0	4.00 ± 0.00	4.00 ± 0.00	5.93 ± 0.26	2.00 ± 0.00	20	100
35S::SRF1A	4.06 ± 0.23	4.06 ± 0.23	5.92 ± 0.31	2.00 ± 0.00	29	145
35S::SRF1B	4.00 ± 0.00	4.00 ± 0.00	5.93 ± 0.30	2.00 ± 0.00	50	250
35S::SRF2	4.00 ± 0.00	4.00 ± 0.00	5.97 ± 0.18	2.00 ± 0.00	50	250
35S::SRF3	4.00 ± 0.00	4.00 ± 0.00	5.99 ± 0.18	2.00 ± 0.00	50	250
35S::SRF4	4.00 ± 0.00	4.00 ± 0.00	5.94 ± 0.23	2.00 ± 0.00	50	250
35S::SRF5	4.00 ± 0.00	4.00 ± 0.00	5.98 ± 0.13	2.00 ± 0.00	50	250
35S::SRF6	4.00 ± 0.00	4.00 ± 0.00	5.94 ± 0.29	2.00 ± 0.00	50	250
35S::SRF7	4.00 ± 0.00	4.00 ± 0.00	5.93 ± 0.28	2.00 ± 0.00	51	255
35S::SRF8	4.00 ± 0.00	4.00 ± 0.00	5.96 ± 0.19	2.00 ± 0.00	20	100

The mean ± SD is shown. The mean values of all measurements between *SRF* overexpression lines and wild-type are not statistically significantly different ( $P > 0.1$ , Student's t-test). N depicts the number of flowers examined and n represents the number of T1 transgenic plants in Col-0 background.

**Table 4.3: Average number of floral organ types in Ler wild-type and overexpression lines flowers in Ler background**

Genotype	Sepal Whorl	Petal Whorl	Stamens	Carpels	n	N
Ler	4.01 ± 0.07	4.01 ± 0.07	5.75 ± 0.49	2.00 ± 0.00	21	105
35S::SRF1A	4.00 ± 0.00	4.00 ± 0.00	5.79 ± 0.41	2.00 ± 0.00	38	190
35S::SRF1B	4.00 ± 0.00	4.00 ± 0.00	5.72 ± 0.50	2.00 ± 0.00	50	250
35S::SRF2	4.00 ± 0.00	4.00 ± 0.00	5.75 ± 0.50	2.00 ± 0.00	49	245
35S::SRF3	4.00 ± 0.00	4.00 ± 0.00	5.77 ± 0.45	2.00 ± 0.00	48	240
35S::SRF4	4.00 ± 0.00	4.00 ± 0.00	5.79 ± 0.43	2.00 ± 0.00	50	250
35S::SRF5	4.00 ± 0.00	4.00 ± 0.00	5.79 ± 0.42	2.00 ± 0.00	50	250
35S::SRF6	4.00 ± 0.00	4.00 ± 0.00	5.79 ± 0.41	2.00 ± 0.00	50	250
35S::SRF7	4.00 ± 0.00	4.00 ± 0.00	5.77 ± 0.47	2.00 ± 0.00	50	250
35S::SRF8	4.01 ± 0.09	4.01 ± 0.09	5.72 ± 0.53	2.00 ± 0.00	50	250

The mean ± SD is shown. The mean values of all measurements between *SRF* overexpression lines and wild-type are not statistically significantly different ( $P > 0.1$ , Student's t-test). N depicts the number of flowers examined and n represents the number of T1 transgenic plants in Col background.

#### 4.2.1.4 The stem length of the respective *SRF* overexpression lines

The stem length of the 8-week old (only for *SRF3* and *SRF5* the stem length was measured in 5-week old plants) transgenic plants (T1 lines) with ectopically

expressed *SRF* members was measured from the rosette leaves until the apex of the main inflorescence. Table 4.4 presents the results. According to the Student's t-test results, we could not observe any statistically significant difference between the stem length of the overexpressed *SRF* lines in Col background and Col-0 plants.

**Table 4.4: The stem length of the 8-week old overexpressed *SRF* transgenic lines in Col-0 background and in wild-type Col-0.**

<b>Genotype</b>	<b>Stem size</b>	<b>n</b>	
Col-0	39.22 ± 4.93	21	
35S:: <i>SRF1A</i>	nd		
35S:: <i>SRF1B</i>	38.76 ± 6.62	49	
35S:: <i>SRF2</i>	36.29 ± 8.56	47	
35S:: <i>SRF3</i> <sup>a</sup>	33.63 ± 5.07	50	Values are in cm. The mean ± SD is shown. The mean values of all measurements between <i>SRFs</i> overexpression lines and Col-0 wild type are not statistically significantly different ( $P > 0.05$ , Student's t-test). N is the number of the plants. <sup>a</sup> The length of the 5-week old transgenic lines (5-week old Col-0 is 36.46 ± 4.18).
35S:: <i>SRF4</i>	40.57 ± 4.31	50	
35S:: <i>SRF5</i> <sup>a</sup>	34.74 ± 3.09	25	
35S:: <i>SRF6</i>	37.17 ± 3.85	49	
35S:: <i>SRF7</i>	38.97 ± 5.75	51	
35S:: <i>SRF8</i>	41.02 ± 7.58	20	

The same analysis was also made between the overexpressed *SRF* transgenic lines in *Ler* background and *Ler* wild-type plants (Table 4.5). Nonetheless no differences were detected among different *SRF* member's overexpression lines and *Ler* wild-type plants. Although 35S::*SRF3* transgenic lines showed a higher height than wild-type *Ler*, another analysis of the new 30 T1 lines and 3 RT-PCR positive T2 lines did not exhibit a longer phenotype.

**Table 4.5: The stem length of the 7-week old overexpressed *SRF* transgenic lines in *Ler* background and in wild-type *Ler***

<b>Genotype</b>	<b>Stem size</b>	<b>n</b>	
<i>Ler</i>	29.02 ± 7.02	22	Values are in cm. The mean ± SD is presented. The mean values of all measurements between <i>SRF</i> overexpression lines and wild-type <i>Ler</i> are not statistically significantly different ( $P > 0.05$ , Student's t-test), except 35S:: <i>SRF3</i> plants are significantly different from <i>Ler</i> plants ( $P < 0.01$ , Student's t-test). N is the number of the plants.
35S:: <i>SRF1A</i>	26.68 ± 7.23	38	
35S:: <i>SRF1B</i>	25.87 ± 4.97	50	
35S:: <i>SRF2</i>	27.90 ± 8.85	48	
35S:: <i>SRF3</i>	34.89 ± 4.84	48	
35S:: <i>SRF4</i>	30.44 ± 5.83	50	
35S:: <i>SRF5</i>	24.90 ± 5.07	50	
35S:: <i>SRF6</i>	31.08 ± 4.62	50	
35S:: <i>SRF7</i>	29.92 ± 6.33	50	
35S:: <i>SRF8</i>	31.73 ± 7.23	49	

#### 4.2.1.5 Stem size, rosette leaves number

Three RT-PCR positive T2 transgenic lines in Col and Ler background were also analyzed to check the number of the rosette leaves in 4-week old plants (Table 4.6 and Table 4.7). Furthermore, the stem length of the 6-week old overexpression T2 lines of the *SRF* members in Col-0 and Ler backgrounds and in Col-0 and Ler ecotypes was investigated (Table 4.8 and Table 4.9). However, no obvious phenotype was visualized between overexpression transgenic lines of the *SRF* family members and wild-type plants.

**Table 4.6: The rosette leaves number of 4-week old overexpressed *SRF* members T2 lines in Col-0 background**

Genotype	Rosette leaves number	n
Col-0	8.64 ± 0.91	25
35S:: <i>SRF1A</i> 10G	8.90 ± 1.12	45
35S:: <i>SRF1A</i> 12G	8.25 ± 1.12	49
35S:: <i>SRF1A</i> 17G	7.82 ± 0.38	48
35S:: <i>SRF1B</i> 37A	8.27 ± 0.87	50
35S:: <i>SRF1B</i> 38A	8.21 ± 0.63	50
35S:: <i>SRF1B</i> 40A	<u>7.63 ± 1.16</u>	48
35S:: <i>SRF2</i> 1H	8.19 ± 1.29	50
35S:: <i>SRF2</i> 4H	<u>7.55 ± 0.88</u>	49
35S:: <i>SRF2</i> 9H	8.76 ± 1.25	49
35S:: <i>SRF3</i> 13B	8.54 ± 0.87	50
35S:: <i>SRF3</i> 14B	8.60 ± 0.94	48
35S:: <i>SRF3</i> 20B	8.36 ± 1.20	49
35S:: <i>SRF4</i> 16C	9.02 ± 0.84	50
35S:: <i>SRF4</i> 17C	8.96 ± 0.93	50
35S:: <i>SRF4</i> 19C	9.22 ± 0.83	49
35S:: <i>SRF5</i> 1I	9.12 ± 1.05	50
35S:: <i>SRF5</i> 34I	<u>8.05 ± 0.76</u>	50
35S:: <i>SRF5</i> 37I	8.74 ± 1.02	49
35S:: <i>SRF6</i> 17D	8.33 ± 1.05	50
35S:: <i>SRF6</i> 19D	8.58 ± 0.90	50
35S:: <i>SRF7</i> 18E	<u>7.73 ± 1.20</u>	50
35S:: <i>SRF7</i> 20E	8.76 ± 1.28	50
35S:: <i>SRF7</i> 32E	8.53 ± 1.02	50
35S:: <i>SRF8</i> 29F	8.15 ± 1.09	48
35S:: <i>SRF8</i> 30F	<u>8.00 ± 0.92</u>	47

The mean ± SD is shown. The mean values of all measurements underlined between *SRF* overexpression lines and wild-type Col are statistically significantly different ( $P < 0.01$ , Student' t-test). N represents the number of T2 transgenic plants in Col-0.

**Table 4.7: The rosette leaves number of 4-week old overexpressed *SRF* members T2 lines in *Ler* background**

<b>Genotype</b>	<b>Stem length</b>	<b>n</b>
<i>Ler</i>	6.25 ± 1.48	40
35S:: <i>SRF1A</i> 14G	6.65 ± 1.39	39
35S:: <i>SRF1A</i> 15G	6.95 ± 0.62	44
35S:: <i>SRF1A</i> 16G	6.90 ± 0.45	46
35S:: <i>SRF1B</i> 29A	<u>7.61 ± 0.50</u>	50
35S:: <i>SRF1B</i> 31A	7.00 ± 0.71	50
35S:: <i>SRF1B</i> 33A	6.50 ± 0.79	49
35S:: <i>SRF2</i> 18H	<u>5.00 ± 1.03</u>	50
35S:: <i>SRF2</i> 10H	7.23 ± 0.67	50
35S:: <i>SRF2</i> 20H	7.33 ± 0.59	48
35S:: <i>SRF3</i> 21B	6.76 ± 0.53	50
35S:: <i>SRF3</i> 26B	7.26 ± 0.65	47
35S:: <i>SRF3</i> 29B	7.14 ± 0.53	48
35S:: <i>SRF4</i> 11C	5.99 ± 0.62	50
35S:: <i>SRF4</i> 13C	<u>5.69 ± 0.85</u>	50
35S:: <i>SRF4</i> 15C	6.84 ± 0.38	50
35S:: <i>SRF5</i> 22I	6.75 ± 0.35	50
35S:: <i>SRF5</i> 24I	7.00 ± 0.71	50
35S:: <i>SRF5</i> 25I	7.12 ± 0.50	49
35S:: <i>SRF6</i> 13D	<u>5.22 ± 0.55</u>	49
35S:: <i>SRF6</i> 14D	5.95 ± 0.85	48
35S:: <i>SRF6</i> 16D	6.81 ± 0.60	50
35S:: <i>SRF7</i> 25E	6.30 ± 0.54	50
35S:: <i>SRF7</i> 26E	6.61 ± 0.82	50
35S:: <i>SRF7</i> 31E	6.79 ± 0.46	50
35S:: <i>SRF8</i> 24F	7.22 ± 0.83	47
35S:: <i>SRF8</i> 25F	6.94 ± 0.25	46
35S:: <i>SRF8</i> 26F	7.02 ± 0.56	48

The mean ± SD is shown. The mean values of all measurements underlined between *SRF* overexpression lines and wild-type *Col* are statistically significantly different ( $P < 0.01$ , Student' t-test). N represents the number of T2 transgenic plants in *Ler* background.

#### 4.2.1.6 Examination of the hypocotyl length of the overexpression lines in *Col-0* background

Three RT-PCR positive T2 transgenic lines in *Col-0* background grown under the dark condition were investigated for hypocotyl length (Table 4.10). *SRF1B* 40A, *SRF5* 34I, and *SRF7* 32E displayed longer hypocotyl length than wild type hypocotyl while *SRF3* 13B exhibited a shorter hypocotyl length. However, other lines did not display statistically significant differences in hypocotyl length.

**Table 4.8: The stem length of 6-week old overexpressed *SRF* members T2 lines in Col-0 background**

<b>Genotype</b>	<b>Stem length</b>	<b>n</b>
Col-0	39.56 ± 5.33	25
35S:: <i>SRF1A</i> 10G	43.06 ± 4.00	47
35S:: <i>SRF1A</i> 12G	39.72 ± 5.70	45
35S:: <i>SRF1A</i> 17G <sup>a</sup>	33.82 ± 4.70	48
35S:: <i>SRF1B</i> 37A	39.02 ± 3.80	50
35S:: <i>SRF1B</i> 38A	41.15 ± 4.57	49
35S:: <i>SRF1B</i> 40A	40.63 ± 4.83	49
35S:: <i>SRF2</i> 1H	38.86 ± 3.80	50
35S:: <i>SRF2</i> 4H	42.52 ± 3.32	50
35S:: <i>SRF2</i> 9H	40.72 ± 4.04	48
35S:: <i>SRF3</i> 13B	39.42 ± 3.92	50
35S:: <i>SRF3</i> 14B	<u>32.82 ± 3.26</u>	49
35S:: <i>SRF3</i> 20B	38.92 ± 2.57	50
35S:: <i>SRF4</i> 16C	40.54 ± 4.79	49
35S:: <i>SRF4</i> 17C	36.88 ± 3.53	50
35S:: <i>SRF4</i> 19C	36.62 ± 3.69	50
35S:: <i>SRF5</i> 29I <sup>b</sup>	31.68 ± 5.02	50
35S:: <i>SRF5</i> 34I <sup>b</sup>	31.37 ± 4.52	49
35S:: <i>SRF5</i> 37I <sup>b</sup>	34.06 ± 1.91	50
35S:: <i>SRF6</i> 17D	38.42 ± 3.78	47
35S:: <i>SRF6</i> 19D	<u>33.07 ± 9.69</u>	48
35S:: <i>SRF6</i> 22D	37.46 ± 5.87	49
35S:: <i>SRF7</i> 18E	38.56 ± 3.49	50
35S:: <i>SRF7</i> 20E	38.79 ± 4.75	50
35S:: <i>SRF7</i> 32E	35.36 ± 5.61	50
35S:: <i>SRF8</i> 29F	37.89 ± 2.67	46
35S:: <i>SRF8</i> 30F	41.08 ± 3.08	47

Values are in cm. The mean ± SD is shown. The mean values of all measurements underlined between *SRF* overexpression lines and wild-type Col are statistically significantly different ( $P < 0.05$ , Student' t-test). N represents the number of T2 transgenic plants in Col-0 background.

<sup>a</sup>10 plants survived out of 40 plants. <sup>b</sup> 5-week old plants (5-week old Col-0 32.50 ± 2.64)

**Table 4.9: The stem length of 6-week old overexpressed *SRF* members T2 lines in *Ler* background**

<b>Genotype</b>	<b>Stem length</b>	<b>Stem length of <i>Ler</i></b>	<b>n</b>
35S:: <i>SRF1A</i> 14G	34.30 ± 7.36	30.72 ± 6.34	49
35S:: <i>SRF1A</i> 15G	33.97 ± 4.95	30.72 ± 6.34	45
35S:: <i>SRF1A</i> 16G	28.48 ± 5.41	30.72 ± 6.34	47
35S:: <i>SRF1B</i> 29A	33.69 ± 4.72	30.72 ± 6.34	50
35S:: <i>SRF1B</i> 31A	33.73 ± 5.76	30.72 ± 6.34	50
35S:: <i>SRF1B</i> 33A	<u>38.12 ± 8.84</u>	30.72 ± 6.34	49
35S:: <i>SRF2</i> 10H	31.00 ± 3.75	30.72 ± 6.34	50
35S:: <i>SRF2</i> 18H	<u>20.49 ± 3.46</u>	30.72 ± 6.34	47
35S:: <i>SRF2</i> 20H	30.65 ± 3.43	30.72 ± 6.34	50
35S:: <i>SRF3</i> 21B	31.73 ± 7.67	29.02 ± 7.02	46
35S:: <i>SRF3</i> 26B	28.11 ± 6.77	29.02 ± 7.02	50
35S:: <i>SRF3</i> 29B	32.34 ± 3.62	29.02 ± 7.02	49
35S:: <i>SRF4</i> 11C	28.64 ± 3.47	29.02 ± 7.02	50
35S:: <i>SRF4</i> 13C	<u>21.66 ± 3.58</u>	29.02 ± 7.02	50
35S:: <i>SRF4</i> 15C	30.69 ± 5.21	29.02 ± 7.02	48
35S:: <i>SRF5</i> 22I	30.27 ± 3.78	27.54 ± 7.61	49
35S:: <i>SRF5</i> 24I	<u>35.52 ± 2.65</u>	27.54 ± 7.61	50
35S:: <i>SRF5</i> 25I	28.38 ± 3.23	27.54 ± 7.61	48
35S:: <i>SRF6</i> 13D	25.97 ± 5.59	27.54 ± 7.61	50
35S:: <i>SRF6</i> 14D	<u>19.74 ± 2.85</u>	27.54 ± 7.61	49
35S:: <i>SRF6</i> 16D	31.06 ± 3.26	27.54 ± 7.61	48
35S:: <i>SRF7</i> 25E	31.93 ± 5.67	30.72 ± 6.34	50
35S:: <i>SRF7</i> 26E	32.74 ± 4.52	30.72 ± 6.34	50
35S:: <i>SRF7</i> 31E	33.36 ± 3.78	30.72 ± 6.34	50
35S:: <i>SRF8</i> 24F	32.65 ± 4.73	30.72 ± 6.34	47
35S:: <i>SRF8</i> 25F	34.23 ± 5.85	30.72 ± 6.34	46
35S:: <i>SRF8</i> 26F	29.77 ± 3.56	30.72 ± 6.34	49

Values are in cm. The mean ± SD is shown. The mean values of all measurements underlined between *SRF* overexpression lines and wild-type *Ler* are statistically significantly different ( $P < 0.02$ , Student' t-test). N represents the number of T2 transgenic plants in *Ler* background.

Table 4.10: The hypocotyl measurements of the 3-day old overexpression lines of the *SRF* members in Col-0 background and in wild type Col-0

<b>Genotype</b>	<b>Hypocotyl length</b>	<b>n</b>
Col-0	11.5 ± 1.8	50
35S:: <i>SRF1A</i> 9G	11.5 ± 1.6	50
35S:: <i>SRF1A</i> 10G	12.3 ± 1.5	48
35S:: <i>SRF1A</i> 12G	12.2 ± 1.4	50
35S:: <i>SRF1B</i> 37A	11.8 ± 0.2	50
35S:: <i>SRF1B</i> 38A	11.7 ± 1.7	50
35S:: <i>SRF1B</i> 40A	<u>12.9 ± 1.3</u>	50
35S:: <i>SRF2</i> 1H	11.2 ± 3.0	50
35S:: <i>SRF2</i> 4H	12.3 ± 2.3	50
35S:: <i>SRF2</i> 9H	12.6 ± 1.6	50
35S:: <i>SRF3</i> 13B	<u>7.6 ± 2.0</u>	47
35S:: <i>SRF3</i> 14B	10.8 ± 2.2	45
35S:: <i>SRF3</i> 20B	11.8 ± 2.4	48
35S:: <i>SRF4</i> 16C	12.0 ± 2.2	50
35S:: <i>SRF4</i> 17C	11.6 ± 1.4	50
35S:: <i>SRF4</i> 19C	10.7 ± 2.4	50
35S:: <i>SRF5</i> 29I	11.7 ± 1.7	50
35S:: <i>SRF5</i> 34I	<u>13.4 ± 1.7</u>	50
35S:: <i>SRF5</i> 37I	12.2 ± 1.9	50
35S:: <i>SRF6</i> 17D	12.2 ± 0.7	50
35S:: <i>SRF6</i> 19D	10.1 ± 3.1	50
35S:: <i>SRF7</i> 18E	11.9 ± 1.3	50
35S:: <i>SRF7</i> 20E	12.3 ± 2.0	50
35S:: <i>SRF7</i> 32E	<u>13.2 ± 1.5</u>	50
35S:: <i>SRF8</i> 29F	12.1 ± 1.9	50
35S:: <i>SRF8</i> 30F	12.3 ± 1.9	48

Values are in mm. The mean ± SD is shown. The mean values of all measurements between 35S::*SRF* overexpression transgenic plants and wild type Col-0 are not statistically significant ( $P > 0.05$ , Student's t-test). Statistically significant measurements between Col-0 and overexpression lines are underlined ( $P < 0.0005$ , Student's t-test). N is the number of the plants measured.

#### 4.2.1.7 Analysis of the root length, hypocotyl length and rosette leaves number of the overexpression lines

The root length, hypocotyl length, and rosette leaves number of the RT-PCR positive T2 lines of the transgenic plants grown 10 days under light condition were also analyzed. Interestingly, all 35S::*SRF4* T2 lines exhibited longer root length and more rosette leaves than wild type plants (Table 4.11). Detailed phenotypic analysis of the 35S::*SRF4* lines is given in chapter 4.

**Table 4.11: The measurements of root length, hypocotyl length, and rosette leaf number of the 10-day old *SRFs* overexpression seedlings in Col-0 background and in Col-0 plants**

Genotype	Root length	Hypocotyl length	Rosette leaf number	n
Col-0	55.9 ± 7.7	1.5 ± 0.4	4.0 ± 0.4	50
35S:: <i>SRF1A</i> 9G	53.1 ± 8.0	2.2 ± 0.4	4.0 ± 0.4	50
35S:: <i>SRF1A</i> 10G	51.8 ± 9.9	1.3 ± 0.3	<u>4.3 ± 0.5</u>	47
35S:: <i>SRF1A</i> 12G	<u>60.1 ± 6.2</u>	1.7 ± 2.1	4.1 ± 0.3	46
35S:: <i>SRF1B</i> 37A	55.1 ± 8.2	1.3 ± 0.4	3.9 ± 0.7	51
35S:: <i>SRF1B</i> 38A	57.0 ± 10.1	1.5 ± 0.5	4.0 ± 0.6	50
35S:: <i>SRF1B</i> 40A	<u>60.0 ± 3.4</u>	1.8 ± 2.8	4.0 ± 0.3	50
35S:: <i>SRF2</i> 1H	<u>46.6 ± 16.8</u>	1.6 ± 0.3	3.8 ± 0.8	50
35S:: <i>SRF2</i> 4H	58.9 ± 6.8	1.4 ± 0.2	4.1 ± 0.2	50
35S:: <i>SRF2</i> 9H	<u>61.9 ± 7.8</u>	1.8 ± 0.7	4.0 ± 0.2	49
35S:: <i>SRF3</i> 13B	<u>53.4 ± 7.2</u>	1.5 ± 0.3	3.8 ± 0.5	47
35S:: <i>SRF3</i> 14B	54.1 ± 5.9	1.5 ± 0.2	4.0 ± 0.1	50
35S:: <i>SRF4</i> 16C	<u>62.6 ± 10.0</u>	1.5 ± 0.4	<u>4.3 ± 0.4</u>	50
35S:: <i>SRF4</i> 17C	<u>66.9 ± 8.5</u>	1.6 ± 0.4	<u>4.1 ± 0.5</u>	50
35S:: <i>SRF4</i> 19C	<u>65.6 ± 9.1</u>	1.6 ± 0.4	<u>4.1 ± 0.2</u>	50
35S:: <i>SRF5</i> 29I	52.8 ± 7.9	1.4 ± 0.3	<u>3.8 ± 0.5</u>	50
35S:: <i>SRF5</i> 34I	54.9 ± 8.9	1.3 ± 0.3	4.1 ± 0.4	46
35S:: <i>SRF5</i> 37I	58.7 ± 7.5	1.6 ± 0.4	3.9 ± 0.4	50
35S:: <i>SRF6</i> 17D	55.4 ± 8.7	1.4 ± 0.4	3.9 ± 0.7	50
35S:: <i>SRF6</i> 19D	<u>43.4 ± 11.1</u>	<u>1.1 ± 1.2</u>	3.7 ± 0.8	50
35S:: <i>SRF7</i> 18E	56.4 ± 5.3	1.4 ± 0.2	4.0 ± 0.2	50
35S:: <i>SRF7</i> 20E	<u>50.9 ± 11.3</u>	1.3 ± 0.3	4.1 ± 0.6	50
35S:: <i>SRF7</i> 32E	<u>64.4 ± 5.7</u>	1.5 ± 0.2	4.1 ± 0.3	50
35S:: <i>SRF8</i> 29F	55.4 ± 8.3	1.4 ± 0.2	<u>4.3 ± 0.3</u>	48
35S:: <i>SRF8</i> 30F	<u>46.6 ± 7.8</u>	1.6 ± 0.8	<u>3.8 ± 0.4</u>	47
35S:: <i>SRF8</i> 32F	54.7 ± 6.4	1.6 ± 0.8	4.0 ± 0.4	49

Values are in mm except for the rosette leaves number. The mean ± SD is shown. The mean values of all measurements underlined between 35S::*SRF* transgenic plants and wild type are statistically significant ( $P < 0.005$ , Student's t-test). N is the number of the analyzed plants.

#### 4.2.1.8 Could *SRF* overexpression in *sub-1* background rescue the *sub-1* phenotype?

To investigate whether *SRF* genes show functional redundancy with *SUB* we generated *SRF* overexpression lines in *sub-1* background. Although we screened at least 50 transgenic lines per overexpressed *SRF* construct, no rescue of the *sub-1* phenotype was observed. Twisted stems and carpels were detected for all investigated transgenic plants. Therefore, it could be concluded that none of the *SRF* members can substitute for *SUB* function. Table 4.12 and Table 4.13 show the analysis of the floral organ numbers and stem sizes of the T1 transgenic lines, respectively.

**Table 4.12: Average number of floral organ types in *sub-1* and SRF overexpression lines flowers in *sub-1* background**

Genotype	Sepal whorl	Petal whorl	Stamens	n	N
<i>sub-1</i>	3.95 ± 0.22	3.50 ± 0.74	5.07 ± 0.69	35	155
35S:: <i>SRF1A</i>	nd	nd	nd		
35S:: <i>SRF1B</i>	3.88 ± 0.39	<u>3.31 ± 0.72</u>	<u>4.75 ± 0.79</u>	50	250
35S:: <i>SRF2</i>	3.97 ± 0.27	3.55 ± 0.56	5.15 ± 0.78	43	215
35S:: <i>SRF3</i>	3.98 ± 0.15	3.36 ± 0.66	4.93 ± 0.57	45	225
35S:: <i>SRF4</i>	3.95 ± 0.26	3.75 ± 0.72	4.75 ± 0.79	50	250
35S:: <i>SRF5</i>	3.87 ± 0.34	3.42 ± 0.75	<u>4.92 ± 0.62</u>	50	250
35S:: <i>SRF6</i>	3.91 ± 0.28	3.50 ± 0.65	5.14 ± 0.70	50	250
35S:: <i>SRF7</i>	3.90 ± 0.31	3.57 ± 0.62	5.10 ± 0.70	49	245
35S:: <i>SRF8</i>	nd	nd	nd		

The mean ± SD is shown. The mean values of measurements underlined between SRF overexpression lines and *sub-1* are statistically significantly different ( $P < 0.03$ , Student's t-test). N depicts the number of flowers examined and n represents the number of T1 transgenic plants in *sub-1* background.

**Table 4.13: Stem size measurement of 8-week old overexpressed SRF lines in *sub-1* background and *sub-1* plants**

Genotype	Stem size	n
<i>sub-1</i>	15.2 ± 1.9	33
35S:: <i>SRF1A</i>	nd	nd
35S:: <i>SRF1B</i>	16.3 ± 4.3	50
35S:: <i>SRF2</i>	15.0 ± 2.1	47
35S:: <i>SRF3</i>	14.6 ± 2.6	50
35S:: <i>SRF4</i>	16.3 ± 2.5	46
35S:: <i>SRF5</i>	<u>13.2 ± 1.9</u>	50
35S:: <i>SRF6</i>	15.2 ± 2.9	50
35S:: <i>SRF7</i>	15.4 ± 2.5	49
35S:: <i>SRF8</i>	15.9 ± 1.8	20

Values are in cm. The mean ± SD is shown. The mean values of all measurements underlined between SRF overexpression lines and *sub-1* are statistically significantly different ( $P < 0.01$ , Student' t-test). N represents the number of T1 transgenic plants in *sub-1* background.

#### 4.2.2 Identification of putative interaction partners of SRF4, SRF5 and SRF6

To find out the putative interacting partners of the SRF4, SRF5 and SRF6, we used the yeast two-hybrid system. We fused the intracellular part of the SRF4 (amino acid residues 301-687), SRF5 (amino acid residues 292-699) and SRF6 (amino acid residues 319-719) to the binding domain of the bait vector pAS2-attr. After the screening of three different cDNA libraries prepared by using three-week old green tissue, 30-day old inflorescence and root cell culture, respectively, several putative interacting partners were obtained.

59 colonies were obtained for SRF4 yeast two-hybrid screen. Since the same gene may occur in different colonies, PCR was performed to amplify the inserts of the prey vectors. Then all amplicons were digested with *TaqI* restriction endonuclease to detect identical clones. Of 59 obtained colonies, 9 different clones showed different *TaqI* digestion pattern. By sequencing and BLASTN analysis 5 different genes were found. Table 4.14 shows the putative interacting partners of SRF4.

**Table 4.14: Putative interacting partners of SRF4**

<b>AGI code</b>	<b>Gene bank accession number</b>	<b>Length of the protein<sup>a</sup></b>	<b>Description</b>	<b>Corresponding cDNA library</b>
At1g43170	NM_103469	389	Arabidopsis ribosomal protein1 (ARP1)	suspension culture
At1g54290	NM_104307	113	Eukaryotic translation initiation factor SU11	suspension culture
At1g79870	NM_106636	313	Oxidoreductase family protein (DAEUMLING)	suspension culture
At2g19640	NM_179660	398	SET-domain containing protein	suspension culture
At3g61480	NM_116013	1091	Putative protein	suspension culture
At3g12965	NC_003074	42	Putative protein	suspension culture

<sup>a</sup>The amino acid residue number of proteins.

55 positive colonies were obtained for the SRF5 yeast two-hybrid approach. After the analysis of the restriction pattern of each PCR product 15 samples were sequenced (Dr. Joachim Uhrig, personal communication). The putative interacting partners of the SRF5 are shown in Table 4.15. 4 candidates were detected after the analysis of the sequence results by means of BLASTN. Interestingly, all candidates inserted in the pGADT7 vector were obtained from the inflorescence cDNA library.

**Table 4.15: Putative interacting partners of SRF5**

<b>AGI code</b>	<b>Gene bank accession number</b>	<b>Length of the protein<sup>a</sup></b>	<b>Description</b>	<b>Corresponding cDNA library</b>
At1g05570	NM_100436	1922	Callose synthase1	Inflorescence
At1g48090	NM_179452	3427	C2-domain containing protein	Inflorescence
At1g76850	NM_106336	1090	Exocyst complex component sec5 (SEC5A)	Inflorescence
At5g61720	NM_125566	390	Expressed protein	Inflorescence

<sup>a</sup>The amino acid residue number of proteins.

Interestingly, for SRF6 only six colonies were obtained after screening of the different yeast two-hybrid libraries. The repetition of the screen did not allow to increase the number of the candidates. After obtaining prey vectors including putative interacting partners, sequence analysis was carried out. According to the results, three different candidates were observed (Table 4.16).

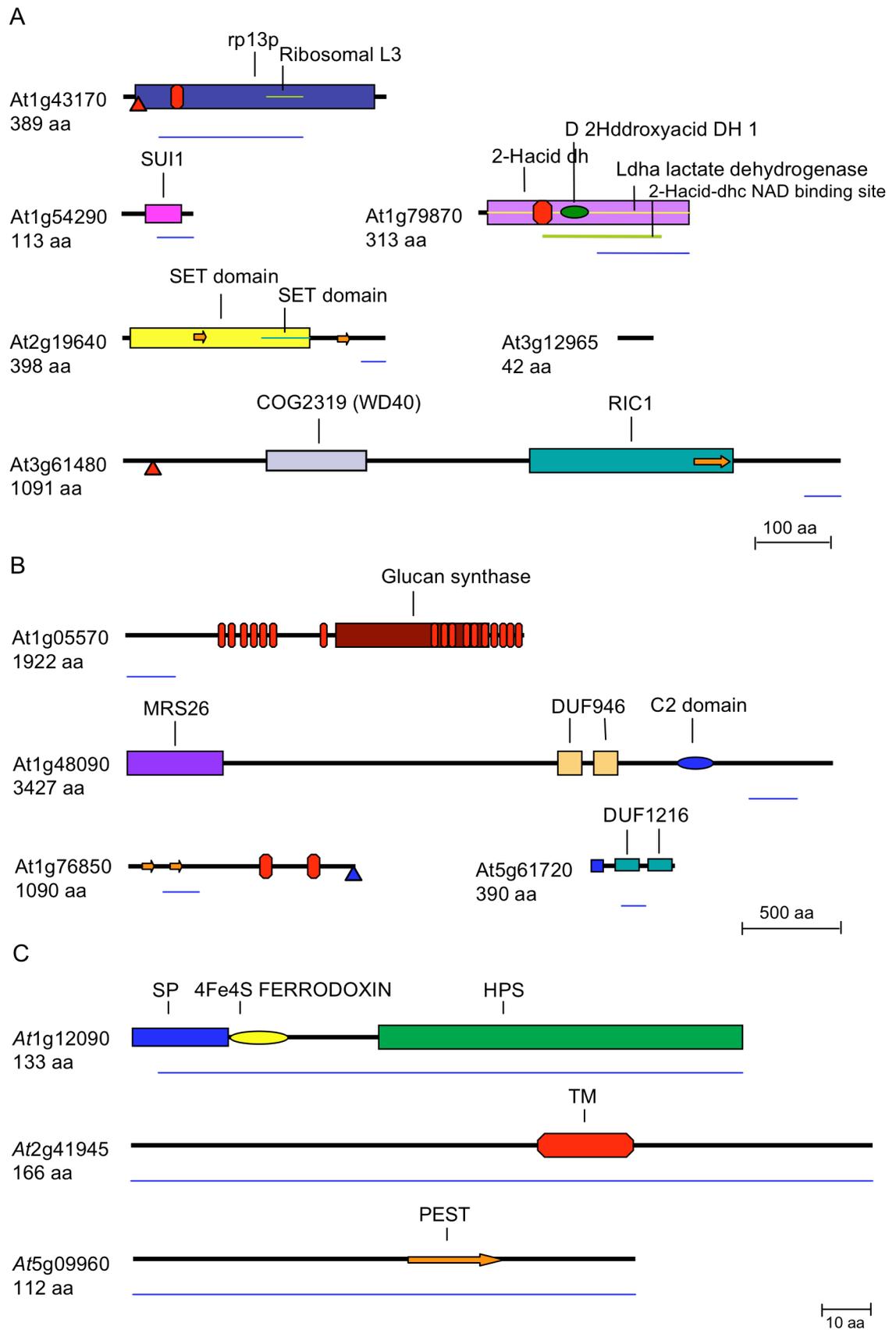
**Table 4.16: Putative interacting partners of SRF6**

<b>AGI code</b>	<b>Gene Bank Accession number</b>	<b>Length of the protein<sup>a</sup></b>	<b>Description</b>	<b>Corresponding cDNA library</b>
At1g12090	NM_101081	137	Extensin-like protein (ELP)	Green tissue
At2g41945	NM_201934	166	Expressed protein	Inflorescence
At5g09960	NM_121033	112	Expressed protein	Green tissue

<sup>a</sup>The amino acid residue number of proteins.

#### 4.2.2.1 Examination of the putative interacting partners

For the characterization of the interacting partners, the domain structure of them were analyzed by means of SMART (<http://smart.embl-heidelberg.de>), PROSITE (<http://www.expasy.ch/prosite/>), PSORT (<http://psort.ims.u-tokyo.ac.jp/form.html>), PESTfind (<https://embl.bcc.univie.ac.at/toolbox/pestfind-analysis-webtool.html>), and NCBI domain search tool web sites (<http://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi>) (Figure 4.5).



**Figure 4.5:** Domain organization of putative interacting partner candidates of

SRF. Putative interacting partners of SRF4 (A), SRF5 (B), and SRF6 (C), respectively. The part of the proteins fished with yeast two hybrid analysis is underlined with a blue stripe. A red octagon represents the transmembrane domain. An orange arrow indicates PEST sequence. Red and blue triangles exhibit mitochondrial and nuclear targeting sequences, respectively.

### 4.3 Discussion

A number of reverse genetic approaches can be used to identify the function of the genes. The most commonly used techniques are knockout or overexpression assays (Zhang, 2003). The overexpression approach can provide useful insights in the function of a gene while loss-of function approach may fail in this respect because of a redundancy of the gene of interest (Strabala *et al.*, 2006). For example, only 10% of the predicted genes of the *C. elegans* displayed phenotypes when they are knocked out (Zhang, 2003). Besides redundancy, the inability to detect small phenotypic changes can also hamper the identification of gene functions. Because of these reasons, overexpression analysis can be the alternative strategy to the knockout/knockdown approach due to the lower affection by redundancy. Moreover, even when a phenotype with the knockout strategy is obtained, overexpression analysis still can be informative due to the creation of an unexpected phenotype. Overexpression phenotypes mostly result from gain-of functions which can be classified in two types: hypermorphs, and neomorphs. Overexpressed genes displaying the same function like the endogenous gene are called hypermorphs. Due to the high amount of the introduced protein or expression in another tissue or developmental stage, overexpressed proteins can show a new phenotype. This phenomenon is called as neomorphs. Because of these reasons, it is mostly expected that overexpression phenotypes should be the opposite of the knockout phenotype. Sometimes, the overexpression strategy does not exhibit any phenotype because more than one gene have to be overexpressed to obtain a phenotype. For instance, overexpression of PI, AP3, SEP3 and AG are necessary for the transformation of the leaves into staminoid organs (Honma and Goto, 2001; Goto *et al.*, 2001; Zhang, 2003). The overexpression of a single gene is not sufficient to visualize this phenotype.

Another approach to gain hints about the function of a protein is to examine protein-protein interactions by means of yeast two-hybrid analysis. Therefore we used these two approaches to obtain some clues about the function of SRF family members.

#### **4.3.1 Ectopic expression of number of the *SRF* genes exhibit male-sterile phenotype**

The phenotypic analysis of overexpression lines of some of the *SRF* genes displayed sterile plants with the absence of the silique development after flower opening. We found out that sterility is due to a defect during the formation of the mature pollen in anther since pollination of the carpel with wild-type pollens causes the formation of normal length of siliques having seeds inside. This approach is evident for the male sterility problem generating short siliques. This analysis revealed that some of the *SRF* members could have a function during anther development. Although the exact reason for the lack of mature pollen in the anther is not known, some reasons may be suggested.

Further investigation is necessary to find out what cell types are affected with the overexpression of the respective *SRF* genes. In-situ hybridization of the *SRF* genes revealed that *SRF2*, *SRF3*, *SRF4*, *SRF5*, *SRF7* genes are expressed in tapetum (Karen Pfister, personal communication). It can be speculated that these genes affect the signaling in the tapetum causing male sterility because the tapetum provides nutrition and materials for the formation of pollen wall.

As the 35S promoter does not function in early sporogenous tissue, it can be expected that the male sterile phenotype might be due to a disturbed late gametogenesis (Jenik and Irish, 2000; Yang *et al.*, 2005). Moreover studies indicate that the 35S promoter function is not uniform in the flower but can be variable. To identify the exact reasons of the male-sterility phenotype of the overexpressed transgenic *SRF* lines, the anther of the wild type and transgenic plants should be fixed and analyzed by light microscopy or alternatively propidium iodide stained different stages of the anthers could be examined by confocal microscopy.

However, we have some hints about the probable function of *SRF5* during pollen development. We found by means of yeast-two hybrid analysis

that *SRF5* can interact with *CalS1* which synthesizes callose. Callose is also found at the wall of the microspores that is formed at the end of the meiosis. However, studies revealed that the callose wall is not required for male meiosis (Scott *et al.*, 2004).

In addition to *SERK1-2* and *EXS/EMS*, *SRF5* might be a new LRR-RLK involved in male fertility.

#### **4.3.1.1. Expression level of *SRF4*, *SRF5*, and *SRF7* affect the strength of male sterility**

RT-PCR analysis of the overexpressed-*SRF4*, *SRF5* revealed that phenotypic severity is positively correlated with the expression levels of the transgene. Possessing higher transgene expression level, severity of the male-sterility phenotype was also elevated most probably affecting pollen formation. Lower expression levels of the genes led to a lower level of pollen formation causing the formation of shorter siliques with a few seeds inside. Furthermore, overexpression lines without detectable overexpressed genes displayed no visible phenotype. The correlation between the transgene expression and the severity of the phenotype was already observed in *myb26/male sterile35* mutant plants. *MYB26/MALE STERILE35*, a putative transcription factor located in nucleus, plays a regulatory role in the determination of the endothelial cell development in the anther acting upstream of the lignin biosynthesis pathway (Yang *et al.*, 2007). No secondary thickening is visualized in *ms35* mutant whereas overexpression lines of the *MS35* have ectopic secondary thickening causing male-sterility. It is shown by RT-PCR that different levels of the expression between different lines and in different tissues affect the phenotypic severity. Lines having a very high expression of the gene are completely sterile while reduced fertility can be detected in the lines possessing reduced expression levels, suggesting that the amount of the *MYB26* protein is linked to the extend of the phenotypic changes.

#### **4.3.2 Overexpression of *SRF1A* and *SRF8* resulted in seedling lethality**

Extreme overexpression of a gene, however, sometimes can be lethal for the organism. For instance, overexpression of the *MCM1*, a cell cycle gene controlling G1/S transition, causes complete stop of cell division in the G1 phase

(Stevenson *et al.*, 2001). However, using a moderate overexpression promoter lead to the 5-10% shift in G1 phase. Although seedling lethality was observed for *SRF1A* and *SRF8* overexpression lines, we could have detected survived transgenic plants. The reason of this phenomenon might be similar to the MCM1 situation. Although we did not perform RT-PCR for the plants exhibiting seedling lethality, most probably a high expression level of the *SRF1A* and *SRF8* might cause the seedling lethality.

What could be the reason for seedling lethality in the overexpressed *SRF1A* and *SRF8* lines? It is hard to speculate about the reason of the seedling lethality phenotype of the *SRF1A*.

Regarding the probable role of *SRF8*, we have the GO description obtained with the enrichment of functional categories within the top 100 genes correlated to *SRF8* gene (Eyüboğlu *et al.*, 2007). This analysis might shed light on the probable function of *SRF8*. According to this analysis *SRF8* may act in sterol biosynthesis. It is shown that sterol biosynthesis is required for embryo and seedling development. Steroids regulate growth in animals and plants (Wang and He, 2004). *CYP51A2*, a cytochrome P450, plays a crucial role in sterol biosynthesis indirectly affecting membrane proteins (Kim *et al.*, 2005). *Cyp51a2* mutant plants displayed growth defects such as short roots, stunted hypocotyls, reduced cell elongation, and seedling lethality. Interestingly, *Arabidopsis* transgenic plants ectopically expressing *CYP51A2* had no morphological changes whereas *cyp51a2-3* mutant displayed seedling lethality.

In addition to seedling lethality phenotype, preliminary studies showed that *SRF8* might be a candidate as a disease resistance gene against *Alternaria* (Dr. Birgit Kemmerling, personal communication). All these hints suggest that *SRF8* might have a dual function: in defense response and as a putative positive regulator of PCD.

Taken together, all these results suggest that seedling lethality of the overexpressed *SRF8* might be hypermorph phenotype.

#### **4.3.3 What might be the function of *SRF4*, *SRF5* and *SRF6*?**

Protein-protein interaction is considered as crucial to understand the biological processes (Pawson and Nash, 2003). Although there are several methods to study protein interaction *in vivo* and *in vitro*, the yeast two-hybrid approach is

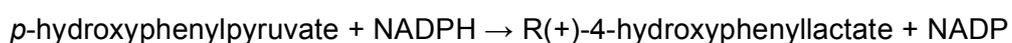
very successful in identifying putative protein interaction partners. However, the yeast-two hybrid result has to be confirmed with an *in vivo* approach. E.g. Ehler and coworkers (2006) showed that the basic leucine zipper transcription factors AtbZIP1 and AtbZIP63 closely interact in yeast whereas in the plant system this interaction appears to be not as strong as in yeast, indicating that a plant-cell specific environment affects the protein-protein interaction.

#### 4.3.3.1 Yeast two-hybrid analysis of SRF4

Six different proteins called oxidoreductase family protein, SET domain containing protein, cytoplasmic ribosomal protein and two expressed proteins with unknown functions are candidates as interacting partners of the SRF4. An understanding of their roles could help us to identify the role of SRF4.

##### 4.3.3.1.1 At1g79870, an oxidoreductase family protein

The most prominent putative interaction partner of SRF4 belongs to the oxidoreductase family protein, because the T-DNA insertion line of this interacting partner displays a similar phenotype like *srf4-2* and *srf4-3* (see chapter 5). The Oxidoreductase protein shows a high similarity to hydroxyphenylpyruvate reductase (HPPR). HPPR of *Coleus blumei* plays a role in the rosmarinic acid biosynthesis II pathway (Kim *et al.*, 2004). The HPPR catalyzes the reduction of 4-hydroxyphenylpyruvate, which originates from L-tyrosine, to 4-hydroxyphenyllactate. The following reaction shows the enzymatic reaction of HPPR.



In the next step of the biosynthesis, 4-hydroxyphenyllactate connects with coumaroyl-CoA, originating from L-phenylalaline, by a transesterification reaction leading to the formation of rosmarinic acid. Rosmarinic acid (3-(3,4-dihydroxyphenyl)-1-oxo-2E-propenyl)oxy-3,4-dihydroxy-benzenepropionic acid) has an antibacterial, antiviral, antioxidant, and antiinflammatory activities (Berger *et al.*, 2006; Petersen and Simmonds, 2003). In mammalian, it helps to prevent cell damage caused by free radicals reducing the risk for cancer.

Moreover, it acts as a defense compound. However, there is no rosmarinic acid in *A. thaliana*.

Interestingly, 4-hydroxyphenylpyruvate is also an origin for homogentisic acid which is a precursor for the formation of tocopherols and plastoquinones, both of which are important for photosynthetically active plant tissues (Kim *et al.*, 2004).

*At1g79870* has a D-isomer specific 2-hydroxyacid dehydrogenase domain. 2-hydroxyacid dehydrogenase from *Haloferax mediterranei* catalyzes the reversible NADH-dependent reduction of 2-ketocarboxylic acids into the 2-hydroxy carboxylic acid (Bonete *et al.*, 2000; Domenech and Ferrer, 2006). This reaction is very essential for the production of hydroxyacids, amino acids or alcohols.

#### **4.3.3.1.2 *At2g19640*, SET-domain containing protein**

We detected a SET-domain containing protein interacting with SRF4 in the yeast two-hybrid approach. It was shown that regulation of this protein can be mediated via protein-protein interactions affecting methylation, substrate specificity and localization of the protein (Ng *et al.*, 2007; Guitton and Berger, 2005).

Eukaryotic gene expression depends on the intrinsic regulatory metabolism and chromosomal context, such as chromatin structure (Ng *et al.*, 2007). Chemical modification of histones like acetylation, methylation, phosphorylation, ubiquitination, and SDP-ribosylation is crucial for the eukaryotic gene regulation. For instance, methylation of histone H3K9 leads to the binding of HP1, a chromodomain containing protein, causing gene repression whereas acetylation of the histone provides the binding of the bromodomain containing protein leading to open chromatin formation which triggers transcriptional activation. *At2g19640* (*ASHR2*) is similar to the ASH protein. *ASHR2* has a homolog in rice called *SDG740* (*Os08g10470*). It was shown that *ASH1* plays a role in H3K36 methylation. Moreover, *ashr2* mutant display an early flowering phenotype indicating a decrease in H3K36 methylation at the FLOWERING LOCUS C (FLC).

It is suggested that to maintain the meristem and organ identity during different stages of development, epigenetic control of gene programs is needed.

It can be hypothesized that SET-domain containing proteins can be necessary for this control (Baumbusch *et al.*, 2001).

For example, regulation of MEDEA (MEA), SWINGER (SWN), and CURLY LEAF (CLF) that are involved in development may be mediated by interacting partners (Baumbusch *et al.*, 2001). Because mutations in these proteins cause similar phenotypes, they may have a role in similar pathways. CLF, involved in the control of leaf and flower morphology and flowering time, represses the expression of *AGAMOUS* in hypocotyls, cotyledons, leaves and inflorescence stem and petals (Baumbusch *et al.*, 2001). A similar scenario may be suggested for SRF4. SRF4 interacting with SET-domain containing protein may indirectly regulate the expression of the transcription factor which is necessary during leaf development.

#### **4.3.3.1.3 At1g54290, Eukaryotic translation initiation factor SUI1**

In eukaryotic cells, protein synthesis starts with the binding of the methionyl-tRNA<sub>i</sub> (met-tRNA<sub>i</sub>) and mRNA to the 40S ribosomal subunit. This binding is followed by the junction of the 60S ribosomal subunit to create the 80S initiation complex. These reactions are mediated by about 10 proteins called initiation factors. The smallest subunit of eIF3, one of the initiation factors comprising 8 or more subunits in *Saccharomyces cerevisiae*, is called SUI1. Naranda and coworkers (1996) showed that SUI1 plays a role during the recognition of the initiation codon because a mutation in *sui1* alters the translation start site selection in *Saccharomyces cerevisiae*.

Another eukaryotic translation initiation factor is eIF4E. This protein recognizes and binds to the cap structure at the 5' end of eukaryotic mRNA (Monzingo *et al.*, 2007). Under normal conditions, mammalian eIF4E binds to suppressor called 4E-binding proteins (4E-BPs) (Raught and Gingras, 1999). Interestingly, this protein is phosphorylated by the FRAPm/TOR intracellular signaling pathway which releases eIF4E from 4E-BP and mediates the formation of a complex with eIF4G. This phosphorylation is performed by the mitogen-activated protein kinase-interacting protein kinase1 (MNK1) and MNK2 at the Ser-209 residue of eIF4E. It is shown that phosphorylation of wheat eIF4E is necessary for normal growth and development (Monzingo *et al.*, 2007) There

are two isoforms of eIF4E in *Arabidopsis* suggesting an overlapping function. Moreover, these isoforms have a role in plant virus resistance.

*Drosophila melanogaster* eIF4E phosphorylation is necessary for normal growth and development (Lachance *et al.*, 2002). Because increased phosphorylation of the protein causes elevated growth rate, the Ser-209 to Ala mutation leads to reduced growth. Examination of the compound eye of *Drosophila* revealed that reduced size is because of the reduced cell size and a minor reduction in cell number. Because mutation in the genes related to the biosynthesis of proteins and nucleic acids cause growth deficiency. Moreover, overexpression of the protein in mammalian cell cultures cause increased growth.

All these studies support that the protein translation factor SUI1 is also a good candidate as a putative interaction partner of SRF4. Because mutation in *SRF4* causes a smaller organ size and overexpression of *SRF4* leads to the formation of bigger organs, a similar effect on cell size as described for other translation initiator factors in mammalian and *Drosophila* organisms.

It could be speculated that SRF4 may control the *At1g54290* activity via its phosphorylation.

#### **4.3.3.1.4 *At1g43170*, Arabidopsis ribosomal protein1 (ARP1)**

Eukaryotic ribosomes comprise more than 70 proteins, the amount of which changes according to the environmental and developmental conditions. Cytoplasmic ribosomal proteins possess functions in both transcriptional and post-transcriptional control mechanisms (Kim *et al.*, 1990). ARP1 is a highly expressed L3 type of a ribosomal gene in leaves.

#### **4.3.3.2 Putative interacting candidates of SRF5**

Four different proteins called CALS1, C2 domain containing protein, SEC5A and expressed protein with unknown function are candidates as interacting partners of the SRF5.

##### **4.3.3.2.1 CALLOSE SYNTHASE 1**

CalS1 is a catalytic subunit of the cell plate-specific callose synthase complex (Verma and Hong, 2001). CalS1 possesses 16 transmembrane domains located

in the N-terminal region and a large central loop facing the cytoplasm and 1922 amino acids in length. Analysis of the *Arabidopsis* genome revealed that there are 12 *CalS* related genes (Hong *et al.*, 2001). Although *CalS* family members share high overall identity (92%), the N-terminal region differs markedly between *CalS* isoforms (average 32% similarity) among family members. Since we fished about 600 bp of the 5' region of the gene with the Y2H approach, this explains why we obtained only *CalS1* and not the other members of the family. Moreover, all *CalS* members have different expression profiles in plant tissues. For instance, *CalS1*, *CalS2*, *CalS3*, *CalS9*, *CalS10*, *CalS11*, and *CalS12* are highly expressed while *CalS5*, *CalS6*, *CalS7* and *CalS8* genes are expressed at a very low level or their expression is induced under certain conditions, such as pathogen infection. Therefore, these isozymes may catalyze callose synthesis in different locations and in response to different physiological and developmental signals.

Callose is located around pollen mother cells and in pollen grains. Moreover, it appears during cytokinesis (Verma and Hong, 2001). During cell plate formation, callose plays a role in filling the tubular network of the fused vesicles allowing the expansion of the cell plate in the center of the phragmoplast. Interestingly, according to the AtGenExpress data SRF5 is highly expressed in stamen and pollen. Taken these aspects and our results together, it can be suggested that SRF5, which may be located on the plasma membrane, regulates *CalS1* when the cell plate comes into contact with the plasma membrane during post-meiotic cell plate formation of pollen development.

The results of Verma (2001) suggest that kinase cascades also play a role in regulating cytokinesis (Verma, 2001). For instance, protein kinase 1 (NPK1) of *Nicotiana* affects the cell plate formation. Moreover, the *Nicotiana* MAPK (p43<sup>Ntff6</sup>) is localized on the cell plate and shows activity in dividing cells.

#### **4.3.3.2.2 C2 domain containing protein**

Another putative interaction partner of SRF5 is the C2-domain containing protein. The respective gene, *At1g48090*, consists of 10,378 bps encoding 3427 amino acids. The region between residues 3025 and 3190 of the C-terminal region was fished in the yeast two-hybrid assay. It possesses the C2 domain encoding region between amino acid residues 2658 and 2763, and the MRS6

like domain, vacuolar protein sorting associated protein. However, the function of this protein is unknown. C2 domains are conserved regions in many eukaryotic signaling proteins including protein kinases (PKCs) that are a large family of phosphatidyl dependent serine/threonine kinases playing several roles in many signaling pathways (Guerrero-Valero *et al.*, 2007). The C2 domain of the protein kinase C $\alpha$  (PKC $\alpha$ ) is characterized possessing eight antiparallel  $\beta$  strands connected by an interstrand loop to form a beta-sandwich structure (Evans *et al.*, 2006). When the intracellular Ca<sup>2+</sup> concentration is increased, C2 domain of the protein binds to two Ca<sup>2+</sup> ions in a pocket called Ca<sup>2+</sup>-binding loop. Binding of Ca<sup>2+</sup> to PKC $\alpha$  triggers the localization of the protein to the plasma membrane especially with phosphatidylinositol-4,5-bisphosphate (PIP<sub>2</sub>).

A number of genes encoding proteins with C2 domain are expressed in plants representing suitable targets for Ca<sup>2+</sup> mediated signaling pathways. Taken together, proteins containing C2 domain may have a function in protein phosphorylation, lipid modification, and membrane trafficking.

At1g48090 C2-domain containing protein also has similarity to pleckstrin homology (PH) containing protein. Pleckstrin homology domains (PH) recognize phospholipids of target membranes, providing intracellular targeting specificity of the protein (Evans *et al.*, 2006). C2 domain consolidates the translocation of protein to intracellular membranes by means of Ca<sup>2+</sup> signals.

One example of the PH domain containing protein is mice Bruton's tyrosine kinase (Btk) (Fukuda *et al.*, 1996). Cytoplasmic tyrosine kinase Btk has a role in B cell activation and development. Mutation in PH domain of the protein causes X-linked immunodeficiency (XLA). It is shown that PH domain can bind to the IP4 and IP6. Mutation in PH domain prevents the binding capacity of IP4 to protein causing a XLA disease.

At1g48090 (C2 domain containing protein) and At5g61720 (expressed protein) as a putative interacting partner of the SRF5 intracellular part possesses also domain similar to DUF. *Drosophila* DUF (*dumbfounded*) gene is required for myoblast aggregation and fusion. DUF protein has a role for the attraction occurring between two types of myoblast during the initial phase of the fusion (Cooper, 2001).

What can be the function of DUF domain containing gene during cell plate formation? May be during the cell plate formation golgi derived vesicles

come to the phragmoplast and fused with each other to form the cell plate, DUF domain containing proteins located on the golgi membrane act as attractant for vesicles to recognize the correct type of vesicles and fused to each other to form cell plate. The different domains of this protein have roles in membrane trafficking and vesicles fusion.

#### **4.3.3.2.3 SEC5A**

SEC5A is another protein that was fished with yeast two-hybrid assay. SEC5, component of the exocyst complex in *Saccharomyces cerevisiae*, is related to exocytosis that is the process to direct the secretory vesicles to the plasma membrane to secrete its ingredients such as membrane proteins, lipids or soluble proteins, hormones, antibodies (TerBush *et al.*, 1996). Some of the exocyst complex components are localized on the plasma membrane while some of them are situated on the Golgi membrane indicating that exocyst has a role to decide the site of fusion (Sommer *et al.*, 2005). Exocyst is also required for the endocytic cycling process in addition to the role of in trafficking from Golgi to plasma membrane. SEC5 of the *Drosophila melanogaster* oocytes is located on clathrin-coated pits and vesicles at the plasma membrane suggesting a role as clathrin-dependent endocytosis for the rapid recycling of vitellogenin receptor *Yolkless*, a protein of the low density lipoprotein receptor family, to take up of vitellogenin (yolk protein) into oocytes.

Although yeast and mammalian SEC5 has a role in exocytosis acting as subunit of exocyst complex, SEC5 in *Drosophila* oocyte localized in clathrin-coated vesicles may suggest an independent specialized role of the Sec5 during clathrin-mediated endocytosis during the uptake of the yolk proteins by the receptor *Yolkless* (Sommer *et al.*, 2005). In addition, Murthy and Schwartz (2004) found that SEC5 has a role for the delivery of new plasma membrane whose null mutant prevents the cell growth and division. In spite of data regarding exocyst complex function, regulation of activation process is still unknown. Taken together, these results suggest that SEC5 can have diverse function.

However, all these available data cannot explain the SRF5 and SEC5A interaction. Preliminary data of confocal microscopy investigations show that the overexpressed EGFP tagged SRF5 protein may be located on the plasma

membrane in *Arabidopsis thaliana* cells. Although *S. cerevisiae* SEC5 protein does not have any transmembrane domain (TM), *A. thaliana* SEC5A protein has two predicted TM domains. It can be located on plasma membrane or gGolgi-derived vesicles. It may be hypothesized that SRF5, located on the plasma membrane, interacts with SEC5A which may be located on vesicles like in *Drosophila* oocyte SEC5 protein during the fusion of the cell plate with the plasma membrane. Alternatively, they can interact during the golgi-derived vesicles and plasma membrane association processes such as endocytosis.

Investigation of the putative interacting partners of the SRF5 indicates that SRF5 may related to callose synthesis and associated with vesicle transport and fusion.

#### **4.3.3.3 Putative interacting partners of SRF6**

Three different proteins were fished as putative interacting partners of SRF6 with the yeast two-hybrid assay. Interestingly, all putative partners are very short peptides ranging from 112 to 166 amino acid residues.

Extensin-like protein (ELP, At1g12090) is the unique candidate for which we can find information, at least about the ortholog of the protein from tomato. *Lycopersicon esculantum* extensin-like protein (LeEXT1) has a probable role during cellular tip growth suggesting involvement in root hair expansion (Bucher *et al.*, 2002). Because *LeExt1*/GUS expression is observed in root hair-bearing cells (trichoblast), but not in root hairless cells (atrachoblast), the protein probably plays a role in root hair tip growth. Plasma membrane and cell materials are present for the synthesis of these materials during the rapid root hair growth. These materials are deposited in the expanding tip place. Moreover, a high expression of the extensin-like protein is observed in root hairs of tomato and cowpea (Bucher *et al.*, 2002). Extensin is the most abundantly found and the best-characterized cell wall protein belonging to the family of hydroxyproline (Hpy) rich glycoproteins (HRGPs) (Showalter, 1993). These proteins are rich in hydroxyproline, valine, tyrosine, lysine, and histidine. Ser-Hyp<sub>4</sub> is the repeating motif of extensin protein. Probably to stabilize the protein, carbohydrates are attached to the Hyp residues. The regulation of the extensin gene is modulated by wounding, fungal and viral infection, fungal and endogenous elicitors, ethylene, red light, heat shock, gravity and development.

In an in-situ hybridization assay with the *LeEXT1* mRNA, transcripts were observed in the differentiation zone of rhizodermal cells whereas no signals were detected in root cap, meristematic and elongating cells at the root tip.

Another extensin domain including protein is the proline-rich extensin-like receptor kinase1 (PERK1) in *Brassica napus* (Silva and Goring, 2002). This RLK has an unusual proline-rich domain exhibiting sequence similarity with the extensin cell wall protein. The suggested function of the RLK is related to the general perception and response to a wound and/or pathogen stimulus probably perceiving the changes via the extensin-like domain and starting a response signaling cascade.

Interestingly, it is shown that SRF6 might have a function with RPP27 in defense response to downy mildew pathogen (Personal communication, Dr. Mahmut Tör; Tör *et al.*, 2004). It is known that *LeEXT1* expression can be regulated by fungal elicitors. Furthermore, global expression profile and analysis of enrichment of GO term also support this suggestion.

ELP, a putative interaction partner of SRF6, possesses two domains called FERRODOXIN iron sulfur binding (4Fe-4S) domain and AAI domain. 4Fe-4S ferredoxin domain subdomain of ferredoxin mediates electron transfer in different metabolic reactions (PROSITE). Another domain called AAI domain is a plant lipid transfer protein/seed storage protein/trypsin-alpha amylase inhibitor domain. This domain creates a four-helical bundle structure stabilized by disulphide bonds.

#### **4.3.4 Functional redundancy of the *SRF* genes**

What is the function of the SRF genes and is there redundancy between the individual SRF genes in this gene family? At present there is no definitive answer to these questions, except for *SRF4* that will be discussed broadly in chapter 5. SUB/SCM affects the orientation of the cell division plane and cell number in various tissues, and root hair patterning (Chevalier *et al.*, 2005; Kwak *et al.*, 2005; Kwak and Schiefelbein, 2007). Ectopic expression of certain *SRF* genes causes male sterility (*SRF2-5* and *SRF7*) or seedling lethality (*SRF1A* and *SRF8*). However, loss-of function mutants of *SRF* genes had no morphological changes. The lack of appropriate environmental conditions might be the reason that no phenotype was observed. In the case of *SRF6*, *SRF7*,

and *SRF8*, expression of these *SRF* genes was still detectable in the respective T-DNA lines by RT-PCR suggesting that sufficient *SRF* activity could still be present in the analyzed T-DNA insertion lines.

Many representatives of the RLK families such as BAM1/BAM2, SERK1/SERK2, and ER/ERL1/ERL2 show redundancy among family members (De Young *et al.*, 2006; Albrecht *et al.*, 2005; Shpak *et al.*, 2004; Shpak *et al.*, 2005). Although redundancy among the family members of the RLKs families is not rare, our investigations suggest that the redundancy among the *SRF* members might not be the case. Our results show that none of the *35S::SRF1-8* constructs could rescue the *sub-1* phenotype indicating that *SRF* members can not functionally replace the *SUB*. Moreover, sequence differences between *SRF* members such as variable proline-rich regions, PEST sequences, differences in the region between the sixth LRR and the TM domain, juxtamembrane domain, alterations in the activation segment and C-terminal extension suggest that many of the *SRF* proteins perform separate functions (see Chapter 3). In addition to these observations, results from global pair-wise *SRF* coexpression analysis do not support the idea of redundancy among the *SRF* members. One could argue about conserved sequence and expression profiles of the *SRF1/SRF3*, *SRF4/SRF5*, and *SRF6/SRF7* gene pairs. However, these gene pairs differ in the GO term enrichments in the groups of coexpressed genes suggesting no redundancy among the *SRF* gene pairs (Eyüboğlu *et al.*, 2007). Although *SRF4* and *SRF5* appear as a sister pair in the phylogenetic tree, there is no overlap in the putative interaction partners of these two proteins as found with the yeast two-hybrid approach. In addition, the *SRF4* phenotype also supports this suggestion. The *srf4* mutants displayed smaller leaves while *srf5* single and double mutants did not show any obvious phenotype. Therefore, these investigations indicate that *SRF* genes exhibit diversity at the functional level.

The last question is what the hypothetical roles of the *SRF* genes are. We might answer this question by the help of our results, global expression profiling and the analysis of the enrichment of GO terms among genes coexpressed with *SRF* genes. According to the expression profiling of *SRF1* and *SRF3* it can be said that these genes could have a role in lignification and pectin biosynthesis.

Enrichment of GO term analysis revealed that SRF7 could have a role in primary cell biosynthesis and in the process of cellulose biosynthesis.

Yeast two-hybrid analysis showed that SRF5 could interact with Callose Synthase 1. Therefore, we can suggest that SRF5 might have a role in cell plate formation. Moreover, all genetic results suggest that SRF4 is a positive regulator of leaf size (see chapter 5).

## Chapter 5. SRF4 is a positive regulator of leaf development

### 5.1 introduction

Leaf as a fundamental unit of the shoot system, is the key organ for the understanding of plant morphogenesis. The initial form of the leaf is a cylindrical primordium originated on the flank of the shoot apical meristem (Donnelly *et al.*, 1999). The development of leaf primordium depends on the foundation of several polarities (Tsukaya, 1998). Final size and shape of a leaf is created by polarity dependent growth. However, the control of leaf shape and size especially at the cellular level is still unclear (Tsukaya, 2003). Leaf development is a very complex process because cell division and cell expansion occur simultaneously in the same region of the leaf during development. In addition to these factors, other factors at the cell-cell level and whole-plant level contribute to the determination of the leaf size.

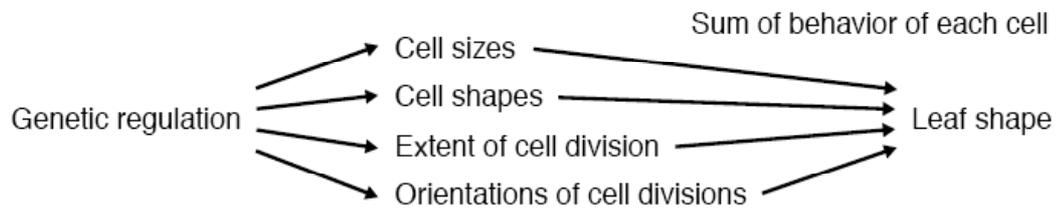
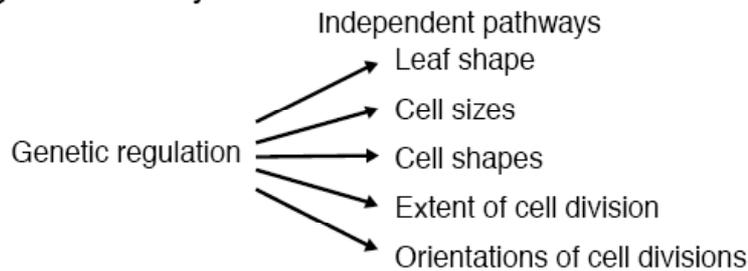
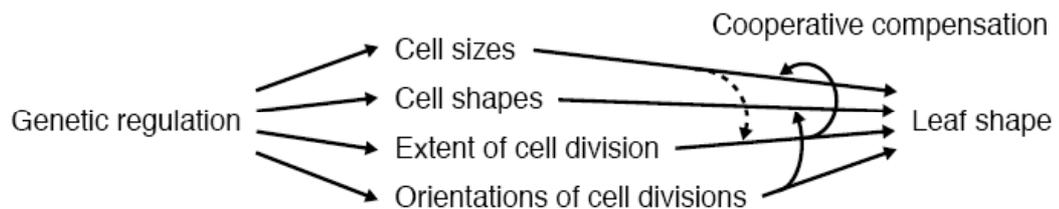
#### 5.1.1 Leaf organogenesis

There are three theories to explain leaf organogenesis, the organismal theory, the cell theory, and the neo-cell theory (Tsukaya, 2003). Figure 5.1 shows the three theories.

In the organismal theory, leaf form is created according to the genetic information. Genetic influences on shape and size of the cell do not affect the leaf size and shape. However, studies revealed that this theory does not reflect the reality. For example, a mutation affecting the arrangement of microtubules leads to the alteration in leaf shape and a mutation in the cell wall loosening protein, expansin, causes a change in cell size (Kim *et al.*, 2002; Cho and Cosgrove, 2000). Thus, these observations, which indicate that the size and shape of the cell affects the size and shape of the leaves, support the cell theory.

### 5.1.1.1 Genes involved in cell expansion

According to the cell theory, the cell, the basic subunit of multicellular organisms, is the unit of organogenesis and morphogenesis (Tsukaya, 2003). A modification of the cell could have direct consequences on leaf size shape. For instance, the *angustifolia* (*an*) mutant plants display same length but narrower and thicker cotyledons and rosette leaves than wild-type and decreased branching of trichomes (Tsuge *et al.*, 1996). Cellular analysis indicated that palisade cells of *an* mutant leaves are narrower and thicker than wild-type. ANGUSTIFOLIA (*AN*) is a new member of the carboxy-terminal binding protein/brefeldin A ribosylated substrate (CtBP/BARS) family, which regulates the polarity-dependent cell elongation along with the width of the leaves of *Arabidopsis thaliana* (Folkers *et al.*, 2002). In animals CtBP plays a role as a co-repressor of transcription regulation while BARS maintains the structure of the Golgi apparatus. Therefore, AN has a function not only as a co-repressor in the nucleus but also interacts with the kinesin like protein in cytosol (Kim *et al.*, 2002; Folkers *et al.*, 2002). AN interacts with MER15, which is involved in cell wall formation. The alteration of the cell wall by the activity of the MER15 results on the modification the orientation of microtubules that might be the reason of the leaf phenotype. Another gene, *ROTUNDIFOLIA3* (*ROT3*), controls the leaf length by regulating the cell expansion of the leaf cells in the leaf length direction (Tsuge *et al.*, 1996; Kim *et al.*, 1998).

**Cell theory (classical)****Organismal theory****Neo cell theory**

**Figure 5.1:** A comparison of the three theories about leaf development (Tsukaya, 2003)

### 5.1.1.2 Genes involved in cell proliferation

Besides *AN* and *ROT3* which affect cell expansion, studies revealed the existence of genes affecting cell proliferation rather than cell expansion during development. *ANGUSTIFOLIA3* (*AN3*), which encodes a homolog of the human transcription coactivator synovial sarcoma translocation protein (SYT), and *GROWTH REGULATING FACTOR5* (*AtGRF5*), which encodes a putative transcription factor, regulate cell proliferation (Horiguchi *et al.*, 2005). The *an3* and *atgrf5* mutant plants display a narrow leaf phenotype because of a decreased cell number while overexpressed *AN3* and *GRF5* plants exhibit larger leaves due to the increased cell proliferation in leaf primordium. Furthermore, yeast two-hybrid analyses showed that these two proteins interact with each other suggesting that *AN3* and *AtGRF5* act together regulating cell proliferation in leaf primordia. These results reveal that shape and size of each

leaf cell affects the shape and size of the whole leaf and thus support the cell theory.

Other members of the *Arabidopsis thaliana* *GROWTH REGULATING FACTOR* (*GRF*) are also known to have a role in leaf size (Kim *et al.*, 2003). Overexpression of *AtGRF1* and *AtGRF2* causes larger leaves and cotyledons and a delayed bolting of inflorescence stem whereas *atgrf1 atgrf2 atgrf3* triple mutant plants display smaller rosette leaves and cotyledons affecting cell size rather than cell number. Single mutants do not show any phenotype while double mutant plants display minor phenotype. These results reveal a redundancy of *AtGRF* proteins. Moreover, overlapping expression patterns of *AtGRF* genes support this suggestion. In addition to the *AtGRF* family, encoding putative transcription factor, extra-small sisters (*xs*) mutants cause reduced leaf size with normal leaf shape (Fujikura *et al.*, 2007).

### 5.1.1.3 Compensatory phenomena

Although studies support the cell theory, new findings reveal that leaf organogenesis could not be explained only with the behavior of each leaf cell in the organ. For instance, a loss-of-function mutation in the *aintegumenta* (*ant*) gene causes the formation of small leaves possessing fewer cells with increased cell volume compared to wild-type whereas overexpression of *ANT* leads to an elevated cell number without changing the cell volume (Mizukami and Fischer, 2000). Therefore, *ANT*, a AP-2 like transcription factor, plays a role in the coordination of cell proliferation and lateral organ development. These results suggest the existence of an compensatory system explaining that a decrease in cell proliferation results in an increase in the cell volume of the individual cell.

Moreover, overexpression of the *AUXIN-BINDING PROTEIN1* (*ABP1*) showed that an increase in cell volume leads to decreased cell number (Jones *et al.*, 1998). However, until now there are no reports showing that a decrease in cell volume causes an increased cell number or that an increase in cell number leads to decrease in cell volume. To explain the leaf development in the light of new findings, the neo-cell theory was created suggesting that cells are the unit of the leaf morphogenesis and each cell, however, is controlled by the factors

governing the organ morphogenesis. This theory supports the importance of the cell-cell communication during organ morphogenesis.

Studies have shown that there are also genes regulating both cell division and cell elongation during leaf development. The *curly leaf* mutant in *Arabidopsis* exhibits narrow and curved leaves with reduced length (Kim *et al.*, 1998). *CLF* is a member of polycomb gene family. Investigations showed that it affects both cell number and size of cells in leaf indicating a role in both cell division and cell expansion processes during leaf development.

### **5.1.2 Leaf morphogenesis not only depend on the cell morphology**

In addition to the individual leaf cell level, leaf organogenesis is also regulated at the whole-plant level, e.g. depending on the availability of nutrients or influenced by environmental factors like quantity and quality of the light or water amount. For instance, Tsukaya (2003) showed that the detachment of apical meristem from young seedlings leads to the formation of larger leaves which supports the idea of the regulation of leaf development at the whole-plant level.

#### **5.1.2.1 Control of leaf development via plant hormones**

In plants, growth and development of organ primordia are mediated by various developmental signals, such as the plant hormones auxin, brassinosteroid (BR), gibberellic acid (GA) and ethylene (Horiguchi *et al.*, 2005). Studies showed that auxin plays a role during shoot and lateral root formation, apical dominance and senescence (Davies, 1995). Auxin acts as a signal not only for cell division, expansion and differentiation but also for the determination of organ size at the cellular level (Hu *et al.*, 2003). AUXIN RESISTANT1 (AXR1) may be involved in auxin dependent cell proliferation during development since the *axr1* mutant displays smaller leaves, inflorescence stem and floral organs affecting cell proliferation (Tsukaya, 2002). ARGOS, AUXIN-REGULATED GENE INVOLVED IN ORGAN SIZE, is also involved in organ size control changing the period of cell proliferation thereby altering the cell number and the duration of organ growth (Hu *et al.*, 2003). The transgenic plants possessing overexpressed and reduced expressed ARGOS show enlarged and reduced aerial organs, respectively. Investigation of the fifth rosette leaves revealed that ARGOS affects the cell number both in leaf length and leaf width direction. In addition

ARGOS also affects floral organs, stem size, hypocotyls length and silique size of the plant. Cotyledons, however, are not affected. Studies revealed that ARGOS might be regulated by auxin via AXR1 thereby ARGOS can affect the duration of cell proliferation by regulating ANT that acts downstream of ARGOS (Hu *et al.*, 2003).

Interestingly the *ARGOS LIKE (ARL)* gene that has a sequence homology with *ARGOS*, plays a crucial role during organ growth controlling cell expansion (Hu *et al.*, 2006). Reduced expression of *ARL* causes smaller cotyledons, leaves and other lateral organs and shorter root and hypocotyl whereas plants overexpressing *ARL* exhibit the larger lateral organs and longer root and hypocotyl. Studies suggest that *ARL* acting downstream of *BRI1* receptor-like kinase mediates a BR-related cell expansion signal during organ growth (Hu *et al.*, 2006). BR is a crucial plant hormone required for developmental processes related to cell expansion and elongation. Therefore, BR might be responsible for hypocotyl elongation and other organ growth. Moreover, BR upregulates the genes affecting the cell expansion, such as *ROT3*, *TCH4* (Tanaka *et al.*, 2005).

### 5.1.2.2 Endoreduplication

Besides cell division and cell expansion, another important factor affecting final size of the cells is endoreduplication. It is the amplification of nuclear DNA without cell division. (Cookson *et al.*, 2006). Endoreduplication can play a role during cell expansion, cell differentiation, and metabolic activity. For instance, during hypocotyl elongation and trichome development endoreduplication plays a crucial role (Traas *et al.*, 1998). At the beginning of the trichome development three endocycles take place forming two branches of a trichome. In this developmental stage, *GLABRA1 (GL1)* and *TRYPTYCHON (TRY)* function as a positive and negative regulators, respectively (Larkin *et al.*, 1994; . The third branch of the trichome is formed in the next step with the fourth endocycle. *GLABRA3 (GL3)* is a positive regulator of this step. *TRY* and *KAKTUS (KAK)* play a negative regulatory role to prevent additional endoreduplication causing extra branch formation in trichomes (Folkers *et al.*, 1997). *gl3* and *try* mutants plants displayed a reduction and increase in the final cell size, respectively. Thus, an increase in the DNA content of the cell also affects the size of the cell

most probably by increasing transcriptional activity, protein synthesis and metabolism (Larkins *et al.*, 2001).

### 5.1.3 Outlook on Chapter 5

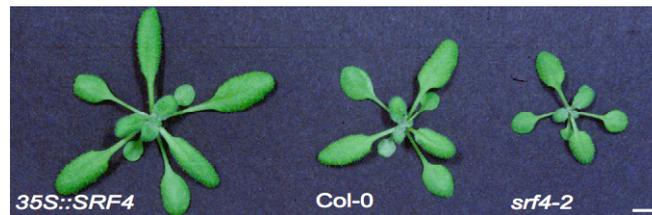
In this part of the study, to identify the function of the gene, we analyzed the overexpression and insertion lines of the *SRF4* gene displaying bigger and smaller leaves, respectively. Analysis of the mutant plants revealed that *SRF4* is a positive regulator of leaf size determination. Moreover, T-DNA insertion line of the *At1g79870*, a putative interacting partner of *SRF4*, showed similar phenotype like *srf4* mutant plants. Therefore, this insertion line was also used for phenotypic analysis. Interestingly, preliminary results of the cell size analysis of the *SRF4* mutant plants suggested that *SRF4* might affect the cell expansion rather than cell proliferation. Finally, analysis of 35S::*SRF4*:EGFP plants showed that *SRF4* could be located on the plasma membrane.

## 5.2 Results

### 5.2.1 Functional analysis of the *SRF4* gene

#### 5.2.1.1 *SRF4* affects the rosette leaves size

The T2 generation of overexpression and the insertion lines of *SRF4* were analyzed morphologically in greenhouse conditions. During this analysis, we detected leaf size differences for *SRF4* constructs. Two insertion lines, *srf4-2* and *srf4-3*, exhibited smaller leaves than the Col wild-type whereas the T2 generation of *SRF4* overexpression lines displayed bigger leaves than wild-type plants. However, not all of the T2 generation of the *SRF4* overexpression transgenic plants showed this phenotype. Therefore, the homozygous T3 generation of transgenic lines 3-12 and 1-5, which display obvious leaves phenotype, was obtained for further investigation. Leaf blade length, width, perimeter, and area measurements were done using the fifth rosette leaves of the 16-day old, respective *SRF4* transgenic plants and the Col wild-type (Tsuge *et al.*, 1996). Figure 5.2 displays the leaf of the plants with altered *SRF4* activity.



**Figure 5.2:** Leaves of the plants with the altered SRF4 activity. The *35S::SRF4* (line 1-5), Col and *srf4-2* plants are indicated. Leaves are enlarged and reduced, respectively. Scale bar is 0.5 mm.

Moreover, during phenotypic analysis of the rosette leaf size we recognized that approximately 80% of the two independent homozygous insertion lines of the *srf4* plants exhibited a smaller leaf phenotype. Both T-DNA insertions are located in exons encoding parts of the extracellular domain of SRF4. However, the reason of the reduced penetrance of the *srf4* mutants is unclear. Especially, the *srf4-2* mutant is supposed to encode only a very short part of the extracellular domain (41 amino acid residues). For this reason, it is unlikely that the only 41 amino acid part of the SRF4 might cause reduction in penetrance. Furthermore, concerning SRF5, a close relative of SRF4, the insertion lines *srf5-1* and *srf5-2*, and *35S::SRF5* transgenic plants were analyzed, and they did not display any leaf size differences. In addition, *srf4-2 srf5-1* double mutants looked like *srf4-2* single mutant. Taken together, these findings indicate that the reduced penetrance of the *srf4* phenotype is not due to the redundancy of SRF5. The reduced penetrance of the *srf4* phenotype, therefore, should be related to unknown factors.

For further phenotypic analysis of the rosette leaves, the fifth rosette leaves exhibiting the most reproducible distinguishing features were chosen (Tsuge *et al.*, 1996). Investigation of the fifth leaf blades of the 16-day old plants revealed that the length and the width of the leaf blade are decreased in *srf4-2* and *srf4-3* mutant leaves whereas overexpressed *SRF4* plants exhibit longer leaf blade length and width. *srf4-2* and *srf4-3* insertion lines displayed an about 20% reduction in length and width of the leaf blade (Table 5.1). This corresponds to an approximately 40% decrease in the surface area of the leaf blade. *SRF4* overexpressed transgenic plants in Col background exhibited a 25-30% increase in length and width of the leaf blade corresponding to a 40-50% increase in the surface area of the leaf blade. Thus, these results show that the

leaf index (leaf length/width ratio) remained constant for all types of transgenic lines under investigation. Although alteration in leaf blade length and width were detected, we could not observe an alteration in leaf shape.

**Table 5.1: Blade size of fifth rosette leaves in 16-day old plants**

Genotype	Length	Width	Leaf Index <sup>a</sup>	Perimeter	Area	n
Col-0	11.5 ± 1.8	6.8 ± 1.0	1.691	32.5 ± 5.5	63.5 ± 19.2	27
<i>35S::SRF4</i>	13.9 ± 1.9	8.2 ± 1.0	1.695	38.0 ± 4.8	87.4 ± 23.1	25
<i>myc.A</i>						
<i>35S::SRF4:myc.A</i>	14.5 ± 1.4	8.8 ± 0.8	1.648	39.8 ± 3.8	95.3 ± 15.8	24
<i>myc.B</i>						
<i>srf4-2</i>	9.2 ± 1.4	5.5 ± 0.8	1.673	25.0 ± 4.0	39.1 ± 11.1	13
<i>srf4-3</i>	8.6 ± 1.7	5.4 ± 1.0	1.593	24.8 ± 4.5	38.9 ± 12.8	22

Values are in mm except for the area column (mm<sup>2</sup>). The mean ± SD is shown. The mean values of all measurements between *srf4* mutants, *35S::SRF4* transgenic plants and wild type are statistically significant (P < 0.001, Student's t-test).

*35S::SRF4:myc.A* and *35S::SRF4:myc.B* correspond to the transgenic lines 3-12 and 1-5, respectively.

<sup>a</sup>Length/width ratio of leaf blade.

In addition to the 16-day old rosette leaf measurements, 26-day old rosette leaves were also assessed (Table 5.2). Similar to 16-day old transgenic plants, 26-day old overexpressed *SRF4* plants showed larger leaves while *srf4-2* and *srf4-3* mutants displayed smaller leaves. Although the leaf length of the overexpression lines increased approximately 18%, the width of the leaf elevated only 10% corresponding to an about 37% increase in leaf area. Thus, the leaf index value changed from 1.885 to 2.000. *srf4-2* and *srf4-3* insertion lines leaf length decreased 16% and 28%, respectively, whereas the width of the leaf decreased 15% and 20%, respectively. These decreases led to a 27% and 42% decrease in leaf area of the *srf4-2* and *srf4-3* insertion lines, respectively.

**Table 5.2: Blade size of fifth rosette leaves in 26-day old plants**

Genotype	Length	Width	Leaf Index <sup>a</sup>	Perimeter	Area	n
Col-0	18.1 ± 2.7	9.6 ± 1.2	1.885	48.6 ± 5.0	13.8 ± 2.6	27
35S::SRF4: <i>myc.A</i>	21.2 ± 1.9	10.6 ± 1.1	2.000	56.2 ± 8.4	18.1 ± 3.2	24
35S::SRF4: <i>myc.B</i>	21.6 ± 3.2	10.4 ± 1.0	2.077	55.4 ± 6.0	18.9 ± 7.6	20
<i>srf4-2</i>	15.2 ± 2.0	8.2 ± 0.9	1.854	40.3 ± 5.5	10.0 ± 2.4	14
<i>srf4-3</i>	12.9 ± 3.2	7.7 ± 1.6	1.675	32.3 ± 9.1	8.0 ± 3.8	11

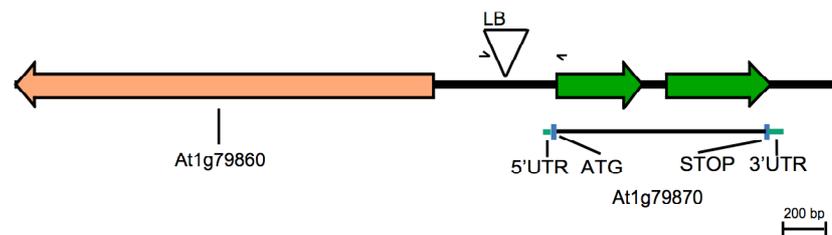
Values are in mm except for the area column (mm<sup>2</sup>). The mean ± SD is shown. The mean values of all measurements between *srf4* mutants, 35S::SRF4 transgenic plants and wild type are statistically significant ( $P < 0.001$ , Student's t-test).

35S::SRF4:*myc.A* and 35S::SRF4:*myc.B* correspond to the transgenic lines 3-12 and 1-5, respectively.

<sup>a</sup>Length/width ratio of leaf blade.

### 5.2.1.2 Phenotypic characterization of *dlg-1*

DAEUMLING (DLG) is one of the putative interacting partner of SRF4 that was showed with yeast two-hybrid approach (see chapter 4). Interestingly, we detected that *dlg-1* carrying an homozygous T-DNA insertion in the promoter region of the *DLG* gene exhibited reduced leaf blade size (Table 5.3). The T-DNA insertion is located at the 251 bp upstream of the start codon of the *DLG* gene (Figure 5.3). 54 bp insertion was detected between left border of the insertion and promoter part of the *DLG* gene.



**Figure 5.3:** Localization and orientation of the T-DNA insertion line of *At1g79870*. Green arrows indicate the exons of *At1g79870*. Orange arrow indicates *At1g79860*. Black arrows display the primers used for T-DNA insertion line analysis. Open triangle highlights the insertion position. LB: left border.

In addition, because *srf4-2* and *srf4-3* T-DNA insertion lines are in *quartet* (*qrt*) mutant background, *qrt* mutant plants were also included into the phenotypic analysis (Table 5.3). *Quartet* mutant is a sporophytic recessive mutant (Preuss *et al.*, 1994; Rhee and Somerville, 1998). In the *quartet1* mutant

all products of meiosis are held together in a tetrad during pollen development. In *quartet* mutants, because of the disruption of the callose deposition around microspores, the exine layer of the pollen tetrads are combined causing the formation of the tetrad clusters.

**Table 5.3: Blade size of fifth rosette leaves in 16-day old plants**

Genotype	Length	Width	Leaf Index <sup>a</sup>	Perimeter	Area	n
Col-0	8.1 ± 1.4	4.9 ± 0.8	1.653	21.7 ± 3.7	30.8 ± 11.2	23
35S:: <i>SRF4:myc.A</i>	10.3 ± 1.1	6.3 ± 0.7	1.634	28.2 ± 3.1	49.6 ± 9.9	26
<i>dlg-1</i> ( <i>At1g79870</i> )	5.8 ± 1.2	3.4 ± 0.6	1.706	15.6 ± 2.9	15.9 ± 6.0	24
<i>qrt</i>	9.3 ± 1.1	5.5 ± 0.6	1.691	25.1 ± 2.7	39.7 ± 8.7	25
<i>srf4-2</i>	5.1 ± 1.6	3.4 ± 0.8	1.500	14.7 ± 4.1	14.5 ± 7.8	17
<i>srf4-3</i>	7.3 ± 1.2	4.6 ± 0.6	1.587	19.8 ± 3.0	25.5 ± 7.3	25

Values are mm except for the area column (mm<sup>2</sup>). The mean ± SD is shown. The mean values of all measurements between *srf4* mutants, *qrt* mutants, 35S::*SRF4* transgenic plants, *dlg-1* mutant, and Col-0 are statistically significant ( $P < 0.005$ , Student's t-test). 35S::*SRF4:myc.A* corresponds to the transgenic line 3-12.

<sup>a</sup>Length/width ratio of leaf blade.

In addition to the blade size of the 16-day old rosette leaves, we measured their petiole length. The results showed that *SRF4* and *At1g79870* affect the petiole length (Table 5.4).

**Table 5.4: Petiole length of fifth rosette leaves in 16-day old plants**

Genotype	Length	n
Col-0	4.3 ± 1.2	24
35S:: <i>SRF4:myc.A</i>	6.2 ± 1.1	25
<i>dlg-1 At1g79870</i>	2.1 ± 0.6	25
<i>qrt</i>	5.0 ± 1.2	26
<i>srf4-2</i>	1.8 ± 1.0	15
<i>srf4-3</i>	3.5 ± 0.9	25

Values are mm. The mean ± SD is shown. The mean values of all measurements between *srf4* mutants, *qrt* mutants, 35S::*SRF4* transgenic plants, *dlg-1* mutant, and Col-0 are statistically significant ( $P < 0.00001$ , Student's t-test; *quartet*  $P < 0.05$ ). 35S::*SRF4:myc.A* correspond to transgenic line 3-12.

### 5.2.1.3. *SRF4* and *DLG* affects different type of leaves

To find out whether the leaf size difference is affected in all leaf types, we also analyzed cotyledons and cauline leaves.

### 5.2.1.3.1 Analysis of cotyledons

The blade size of the 16-day old cotyledons was also measured (Table 5.5) Results revealed that *SRF4* affects blade size of the cotyledons in a similar way to the fifth rosette leaves. The leaf index of cotyledons was also constant like in the rosette leaves indicating that both blade length and width were affected.

**Table 5.5: The blade size of the 16-day old cotyledons**

Genotype	Length	Width	Leaf index <sup>a</sup>	Perimeter	Area	n
Col-0	3.2 ± 0.3	3.1 ± 0.3	1.032	10.6 ± 0.8	7.8 ± 1.1	23
<i>35S::SRF4:myc.A</i>	3.8 ± 0.3	3.4 ± 0.3	1.118	12.4 ± 1.5	10.4 ± 1.7	26
<i>dlg-1</i> (At1g79870)	2.5 ± 0.3	2.1 ± 0.3	1.190	7.7 ± 0.9	4.3 ± 0.9	24
<i>qrt</i>	3.3 ± 0.3	2.9 ± 0.2	1.138	10.5 ± 0.8	7.5 ± 1.1	25
<i>srf4-2</i>	2.7 ± 0.4	2.4 ± 0.3	1.125	8.3 ± 1.0	4.8 ± 1.1	17
<i>srf4-3</i>	3.1 ± 0.3	2.6 ± 0.3	1.192	9.5 ± 0.8	6.2 ± 1.0	25

Values are mm except for the area column (mm<sup>2</sup>). The mean ± SD is shown. The mean values of all measurements between *srf4* mutants, *qrt* mutants, *35S::SRF4* transgenic plants, *dlg-1* mutant, and Col-0 are statistically significant ( $P < 0.001$ , Student's t-test). *35S::SRF4myc.A* corresponds to the transgenic line 3-12.

<sup>a</sup>Length/width ratio of leaf blade.

### 5.2.1.3.2 Cauline leaves measurements of transgenic plants

The second internode cauline leaves of the respective mutants were measured. The analysis revealed that there is also an alteration in cauline leaf blade size (Table 5.6). The overexpression line displayed a 16% increase in blade length while *srf4-2*, *srf4-3*, and *dlg-1* mutant lines exhibited 33.4%, 11.1% and 22.7% reduction in leaf blade length, respectively. Although the leaf length of the all analyzed plants exhibited a size change, the width of the overexpressed line and *srf4-3* cauline leaves did not display statistically significant difference. Interestingly, cauline leaf size was more affected in *dlg-1* plants than in *srf4* mutants.

**Table 5.6: Blade size of the second cauline leaves in 42-day old plants**

Genotype	Length	Width	Leaf Index <sup>a</sup>	Perimeter	Area	n
Col wt	29.29 ± 4.06	8.5 ± 2.2	3.44	70.7 ± 10.1	185.1 ± 48.5	19
<i>35S::SRF4</i>	32.78 ± 3.22	<b>9.4 ± 2.2</b>	3.48	<b>75.2 ± 6.3</b>	<b>210.3 ± 47.2</b>	20
<i>dlg-1</i>	17.10 ± 3.11	6.5 ± 1.8	2.63	41.0 ± 8.1	79.4 ± 34.3	22
<i>At1g79870</i>						
<i>qrt</i>	31.55 ± 3.35	<b>9.4 ± 1.8</b>	3.35	<b>74.5 ± 7.4</b>	<b>192.8 ± 32.7</b>	20
<i>srf4-2</i>	28.59 ± 3.31	7.4 ± 1.2	3.86	69.1 ± 7.7	159.9 ± 21.8	21
<i>srf4-3</i>	27.90 ± 2.52	<b>8.9 ± 2.2</b>	3.13	67.0 ± 7.1	<b>176.8 ± 46.5</b>	20

Values are in mm except for the area column (mm<sup>2</sup>). The mean ± SD is shown.

(P < 0.03, Student's t-test). Statistically not different values are written in bold.

*35S::SRF4:myc.A* correspond to transgenic lines 3-12. <sup>a</sup>Length/width ratio of leaf blade.

In addition to leaves size analysis, sepal and petal length, a modified leaves, were also investigated. Analysis revealed that *srf4* and *dlg-1* insertional lines displayed shorter sepal and petal whereas overexpressed *SRF4* lines did not exhibit any difference in sepal and petal length (Christine Skornia, personal communication).

#### 5.2.1.4 Investigation of stem length

Stem length of the 6-week old transgenic plants was measured a (Table 5.7). The *dlg-1* transgenic line shows almost a 50% reduced stem size than compared to Col-0. Although the stem size of the *srf4* mutants and Col-0 wild type did not showed statistically significant difference, statistical analysis of the *srf4* mutants with *qrt* mutant stem size displayed significant difference. These results indicate that *QRT* gene may affect stem size. Most probably because of this effect *srf4* mutant plant did not displayed statistical significance when compared with the stem size of the Col-0.

**Table 5.7: Stem size of the 6-week old plants**

Genotype	Stem size	n
Col-0	39.22 ± 4.93	20
<i>35S::SRF4myc.A</i>	41.85 ± 2.29	20
<i>dlg-1 (At1g79870)</i>	22.90 ± 6.20	20
<i>qrt</i>	42.01 ± 4.02	20
<i>srf4-2</i>	39.44 ± 1.96	16
<i>srf4-3</i>	39.38 ± 1.59	16

Values are in cm. The mean ± SD is shown. The mean values of all measurements between *srf4* mutants, *qrt* mutant, and between *35S::SRF4* transgenic plants, *dlg-1* mutant, and Col-0 are statistically significant (P < 0.05, Student's t-test). *35S::SRF4myc.A* corresponds to transgenic line 3-12.

### 5.2.1.5 Hypocotyl measurements of *srf4* and *dlg-1* mutants and wild-type

We also measured the hypocotyl lengths of the 72-hour old Col wild-type, *35S::SRF4*, *dlg-1* (*At1g79870* insertion line), *srf4-2*, *srf4-3*, and *quartet* mutant, which were grown under dark condition. We also observed an alteration in length of the hypocotyls. *srf4-2*, *srf4-3*, and *dlg-1* exhibited a reduced hypocotyl length when compared with their respective wild-type plants while the *35S::SRF4* transgenic line exhibited an increased hypocotyl length (Table 5.8).

**Table 5.8: Hypocotyl length of the 3-day old seedlings**

<b>Genotype</b>	<b>Length</b>	<b>n</b>
Col-0	13.74 ± 1.92	35
<i>35S::SRF4:myc.A</i>	15.99 ± 2.42	44
<i>dlg-1</i> ( <i>At1g79870</i> )	10.62 ± 1.42	39
<i>qrt</i>	14.56 ± 1.54	42
<i>srf4-2</i>	9.69 ± 3.37	25
<i>srf4-3</i>	12.95 ± 2.69	41

Values are in mm. The mean ± SD is shown. The mean values of all measurements between *srf4* mutants, *qrt* mutants, and between *35S::SRF4* transgenic plants, *dlg-1* mutant, and Col-0 are statistically significant ( $P < 0.001$ , Student's t-test). *35S::SRF4:myc.A* corresponds to transgenic lines 3-12.

### 5.2.1.6 Root size investigation

10-day old roots of the Col wild-type, *35S::SRF4*, *dlg-1* (*At1g79870* insertion line), *srf4-2*, *srf4-3*, and *quartet* mutant were measured (Table 5.9). Overexpressed and *dlg-1* mutant plants displayed significantly different root lengths compared to wild type. Although the root lengths of the *srf4-2* and *srf4-3* mutant plants were not significantly different than *qrt* mutant roots, *srf4* mutant's roots were shorter than Col-0 roots.

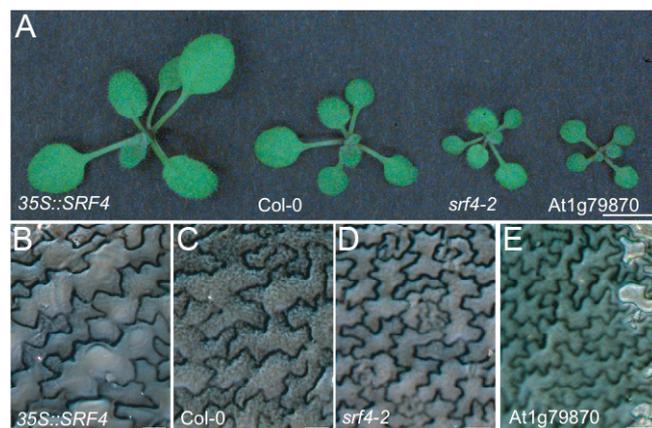
**Table 5.9: Root length of the 10-day old seedlings**

<b>Genotype</b>	<b>Length</b>	<b>n</b>
Col-0	48.4 ± 6.0	35
<i>35S::SRF4:myc.A</i>	70.1 ± 6.2	44
<i>dlg-1</i> ( <i>At1g79870</i> )	36.8 ± 7.8	46
<i>qrt</i>	40.4 ± 4.7	47
<i>srf4-2</i>	<b>39.8 ± 6.9</b>	29
<i>srf4-3</i>	<b>39.3 ± 10.7</b>	40

Values are in mm. The mean ± SD is shown. ( $P < 0.001$ , Student's t-test). *35S::SRF4:myc.A* corresponds to the transgenic lines 3-12.

### 5.2.2 Cell size measurement

The difference of the organ size could be due to the difference in cell division, difference in cell expansion or both at the same time. To understand why *srf4* and *dlg* mutant plants displayed smaller leaves than wild type leaves, we analyzed the cells of the leaf epidermis via dried-gel method (Horiguchi *et al.*, 2006). The comparison of the epidermal cells among the *35S::SRF4*, *srf4-2*, and *dlg-1* transgenic plants indicated that epidermal cells appeared smaller in *srf4-2* plants but larger in overexpression lines than compared to wild type epidermal cells. In addition, *dlg-1* mutant also displayed a smaller epidermal cell size. Thus, preliminary analysis revealed that cell size alteration might be the reason of the organ size phenotype. In Figure 5.4, the cell size differences of the transgenic and wild type lines are presented.

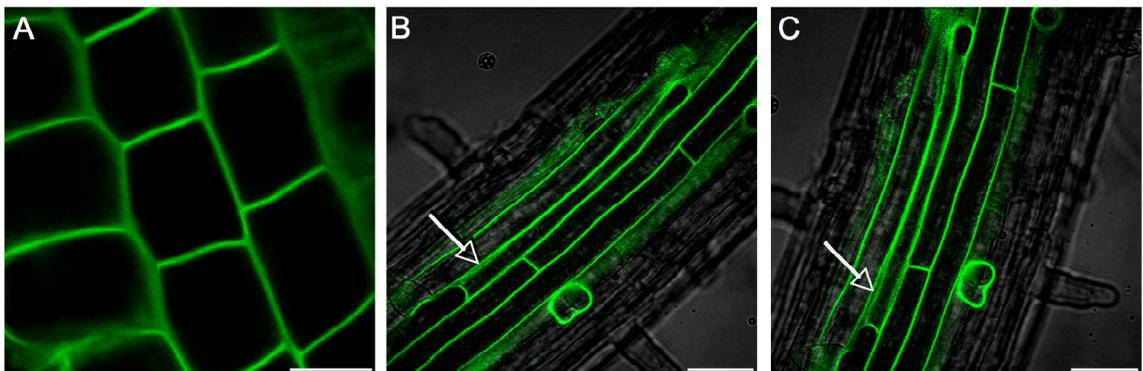


**Figure 5.4:** Phenotypic effects altering SRF4 and DLG activity. A) Leaves of plants with altered SRF4 and DLG activity. B-E) Epidermal cells of 11-day old cotyledons. Scale bar for A: 1 cm, and scale bar for B-E: 100  $\mu$ m

### 5.2.3 The cellular localization of the SRF4 protein

To detect the localization of SRF4 within the cell, we generated a transgenic lines expressing SRF4:EGFP under the 35S promoter. More than 100 independent T1 transgenic lines were checked for EGFP expression by using fluorescence microscopy. 6 of the lines showing very strong, middle, and very weak expression of the EGFP were chosen for the further analysis. The SRF4:EGFP fusion protein was detected at the cell periphery, suggesting that SRF4:EGFP might be located at the plasma membrane (Figure 5.5A). To support this observation, we performed plasmolysis experiment. Plasmolysis

exerts a negative osmotic pressure in a cell and evacuates the fluid out of a cell. This phenomena cause an invagination of the plasma membrane leaving the more rigid cell wall unaffected. After 15 minutes 1M sorbitol treatment of the root cells expressing SRF4:EGFP, shrinkage of the plasma membrane was visualized supporting the idea that SRF4:EGFP is localized on the plasma membrane.



**Figure 5.5:** SRF4:EGFP is localized in the plasma membrane. Confocal laser scanning microscopy images of Col-0 root cells expressing *35S::SRF4:EGFP*. A) Image of the SRF4:EGFP in cortex of the transgenic root . B-C) Image of the SRF4:EGFP in the stele treated with 1M sorbitol for 10 minutes and 15 minutes, respectively. Arrow indicated gaps between neighboring cell's plasma membrane. Scale bar for A-B: 0.5  $\mu$ m.

### 5.3 Discussion

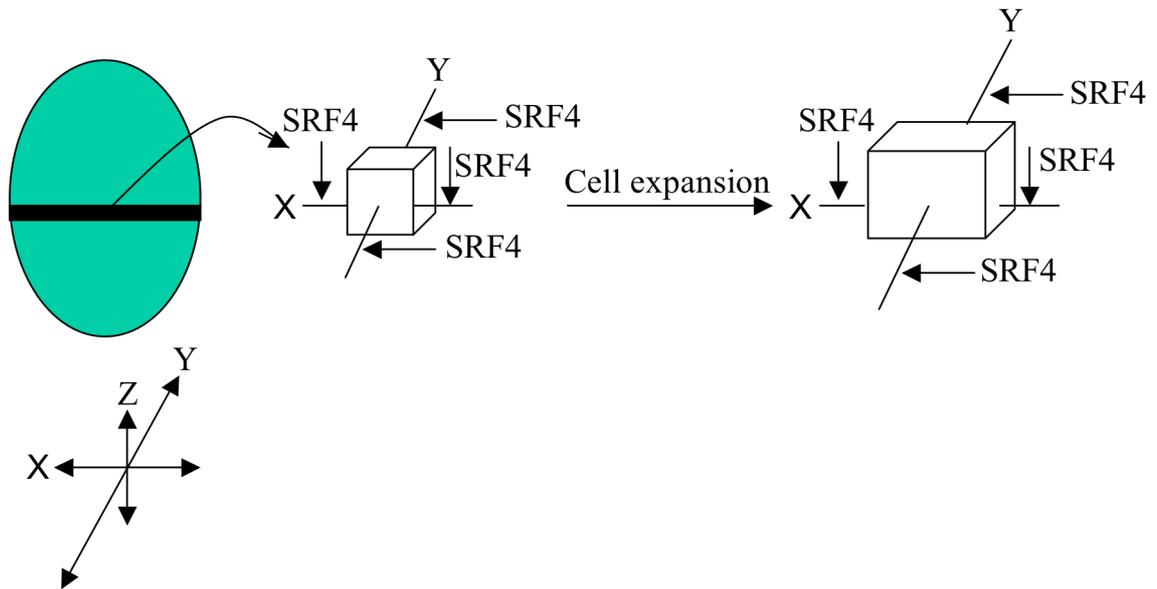
In this study we identified *SRF4* gene as a regulatory component controlling leaf width and length. Moreover, one of the putative interacting proteins of SRF4, DAEUMLING (DLG, *At1g79870*), may play a role in this regulatory process during leaf development because the phenotypes caused by loss of function of these two genes are similar to each other. In addition to the rosette leaves size, the results indicate that SRF4 and DLG affect hypocotyls length, stem length, length of the cauline leaves, and sepal and petal, which are modified leaves.

#### 5.3.1 SRF4 is a direct positive regulator of leaf size

The results of the genetic and morphological analysis indicate that SRF4 is a direct positive regulator of leaf size. SRF4 affects the leaf size by changing both the longitudinal and lateral size of the leaf.

Interestingly, the reduced expression and the overexpression of *ARGOS LIKE (ARL)*, has a sequence homology with *ARGOS*, causes smaller and larger cotyledons, leaves and other lateral organs and shorter and longer root and hypocotyls (Hu *et al.*, 2006). Studies revealed that alteration in size of the leaves and cotyledons are because of the cell size alteration rather than the cell number changes, indicating that *ARL* possesses a function in cell expansion. In contrary, *ARGOS* regulates the cell proliferation. Therefore, in spite of the sequence similarity between *ARGOS* and *ARL*, they affect different cellular processes during organ development. *ARL* is acting downstream of *BRI1* and mediates BR-related cell expansion signals during organ growth. BR is a crucial plant hormone required for developmental processes related to cell expansion and elongation. Therefore, BR might be responsible for hypocotyls elongation and other organ growth by affecting cell expansion process.

In this study we showed that *SRF4* is a positive regulator of the leaf size most probably affecting the cell expansion. Our results indicate that *SRF4* affects the size of the cotyledons, rosette leaves, cauline leaves, stem height, hypocotyls length. Overexpression and T-DNA insertion lines of the *SRF4* alter the above mentioned organs size. Our analysis revealed that the *SRF4* gene regulates the leaf morphology by affecting leaf growth directly both in longitudinal and transverse direction. Figure 5.6 summarizes the function of *SRF4*. Thus, it can be suggested that *SRF4* is a direct regulator of leaf size. Furthermore, *SRF4* also affects the length of the stem and hypocotyls.



**Figure 5.6:** Schematic representation of the SRF4 function during leaf development. X, Y, and Z represent transverse, longitudinal, and dorsiventral axes of a leaf, respectively. Cubes represent individual cells. Green sphere displays a rosette leaf.

### 5.3.2 Does SRF4 affect cell number or cell size of leaf cells?

Cell division and cell expansion are the key processes during plant growth. To find out whether SRF4 affects the cell proliferation, resulting in a changed cell number, or cell elongation altering the cell size of the leaves we performed an anatomical analysis of cotyledon epidermis and hypocotyls cells. Cotyledons are leaf-like structures developing during embryogenesis (Tsukaya, 2002). Recent studies revealed that growth of the cotyledons and hypocotyls after germination depends on the cell expansion rather than the cell proliferation. We observed shorter and longer hypocotyl and smaller and bigger cotyledons in *srf4* insertion lines and *35S::SRF4* overexpression transgenic lines compared with wild type plants. Therefore, it could be hypothesized that SRF4 affect the cell expansion rather than cell division. Moreover, our cytological analysis revealed that the epidermis cells of the 11-day old cotyledons are smaller in *srf4-2* and *dlg-1* mutants whereas they are bigger in overexpressed *SRF4* transgenic line.

In literature, certain genes affecting the cell expansion have been identified. *Angustifolia* mutant affecting cells expansion in cell-width direction bear also narrower cotyledons (Tsukaya *et al.*, 1994; Kim *et al.*, 2002). Furthermore, *ARL* overexpression and reduced expression lines affecting cell expansion also displayed increased and reduced cotyledon size, respectively

(Hu *et al.*, 2006). However, another gene, *ARGOS*, affecting cell division did not display any difference in cotyledon size. These results also support our findings. Furthermore, studies revealed that hypocotyl elongation during postembryonic growth mostly depends on the cell expansion rather than cell division (Gendreau *et al.*, 1997; Sabio *et al.*, 2003). Obtaining elongated or reduced length of the hypocotyl in *SRF4* overexpressed and insertion lines mutant plants, respectively, is also an evidence to prove that *SRF4* might affect the cell elongation process rather than the cell division process. Moreover, our cytological analysis indicated that the epidermis cells of the 11-day old cotyledons display smaller cells in *srf4-2* and *dlg-1* mutant whereas overexpressed *SRF4* transgenic lines exhibit bigger cell size.

Interestingly *SRF4* phenotypes are very similar to the *ARL* phenotype. Both of the genes affect cotyledons, rosette leaves, and hypocotyls size. *ARL* is regulated by *BR*. Thus, further investigations will reveal whether or not *BR* also play a role for *SRF4* function.

The theory suggesting that *SRF4* might affect cell expansion process is also supported by the GO term enrichment analysis showing the genes that are coexpressed with *SRF4*. This analysis allow to determine the genes that are coexpressed with *SRF4* According to this analysis, *SRF4* might be the mechanism involving pectinesterase enzyme (Eyüboğlu *et al.*, 2007). Pectinesterase is a cell wall associated enzyme facilitating plant cell wall modification (Richard *et al.*, 1994). Plant cell wall, composed of cellulose, hemicellulose, pectins, proteins and enzymes, has a role in cell size determination, growth and development, cell-cell communication, and interaction with environment (Micheli, 2001). Pectinesterase catalyses the de-esterification of methyl-esterified D-galactosiduronic acid units in pectic compounds located in the cell wall and has a function in the alteration of the cell wall mechanical stability during cell extension, stem elongation, leaves, root and hypocotyl development, pollen tube growth, pollen germination, and defense reactions against pathogens (Richard *et al.*, 1994; Pereira *et al.*, 2006). Pectinesterase activity causes the formation of stiffening of cell wall or cell wall loosening. Investigations revealed that in the hypocotyl elongation region of mung bean (*Vigna radiata*) has highly elevated levels of methyl esterified pectins (Derbyshire *et al.*, 2007). In addition, *Nicotiana tabaccum* cell

suspension cells displayed an elevated level of methyl esterification mediating cell elongation. During *Arabidopsis* hypocotyl elongation increased pectin methyl esterification is observed, indicating that the degree of pectin esterification may affect the cell elongation. Therefore, it would be interesting to analyze the level of the methyl esterified pectins in *srf4* insertional lines and in *35S::SRF4* transgenic lines to see whether SRF4 pathway is overlapped with the pectinesterase enzyme.

Since an alteration especially in size of cotyledons and hypocotyls length is not observed in cell proliferation mutants, such as *argos*, *an3* and *atgrf5*, it is unlikely that SRF4 affect the cell proliferation.

In our studies, DLG, a putative interacting partner candidates of SRF4, insertional line displayed similar leaves and hypocotyl, and stem size phenotype of SRF4. However, *dlg-1* mutant plants exhibited more severe phenotype than *SRF4* mutants have. We may assume that DLG might be involved in other pathways in addition to the SRF4 pathway.

Furthermore, putative interacting partner candidates of SRF4, such as the protein similar to translation factor SUI and cytoplasmic ribosomal proteins are also good candidates to support the theory suggesting that SRF4 is a direct regulator of leaf size. Since mutants of similar proteins revealed organ size defects indicating that they might interact with SRF4 during the regulation of different organ size determination processes.

Identification of the ligand and downstream targets of the SRF4 may reveal the signaling mechanism regulating cell expansion during organ development.

Our research allows us to discover new genes involved in leaf development and a deeper understanding of the regulation of leaf development. Moreover, increased organ mass in *35S::SRF4* plants could be applied biotechnologically to improve the yield of the agronomic crops.

## Chapter 6. Conclusion and future studies

Although the *Arabidopsis* genome has been sequenced, the functions of a high number of genes are still unknown. For instance, the function of 2,500 genes out of 28,000 genes in *Arabidopsis* has been investigated experimentally (Hilson *et al.*, 2003). In this study, we investigated the function of the eight genes belonging to the *SRF* family. Therefore, this study, provides new insights to the function of the RLKs with unknown functions.

Approximately 30% of the genome is wrongly annotated according to the whole genome transcription analysis (Yamada *et al.*, 2003). Comparison of our sequence results of the *SRF* genes and publicly available sequence data revealed that all of the *SRF* genes was wrongly annotated in database. Moreover, obtaining two isoforms for *SRF1* by means of alternative splicing also shows that not only prediction of exon-intron boundaries of genes but also the detection of the alternative splicing phenomena could not be achieved by using bioinformatic techniques.

Analysis of the protein sequences of the *SRF* members revealed that each *SRF* members might exhibit different ligand specificities since sequence conservation among the *SRF* LRRs is very low. Moreover, only certain *SRF* members have PRR and PEST sequence.

Before this study, only the role of *SUB/SCM* was known. It affects the orientation of the cell division plane and cell number in various tissues, and root hair patterning. Our investigations allowed us to obtain some valuable hints about the function of certain *SRF* members. Overexpression analyses in addition to T-DNA analysis of the family members provided first insights in the function of the *SRF* members. The most important result of this study is the identification of the function of *SRF4*. Our results revealed that *SRF4* is a positive regulator of leaf development. Cell size analysis of the T-DNA mutants and overexpressed transgenic plants of the *SRF4* indicates that *SRF4* most probably affects the cell expansion during the development of the leaf. Most of the *SRF4* putative interacting candidates obtained with yeast two-hybrid approach support the role of the *SRF4* in organ development. Especially, a T-DNA line of the *DLG*, *At1g79870*, a putative interacting partner of *SRF4* in the

yeast two-hybrid system, displayed a similar phenotype as *srf4* and strengthens our results related to *SRF4*.

Ectopic expression analysis of the *SRF* members also provides hints about the probable function of the certain *SRF* members. Because overexpression lines of *SRF2-5* and *SRF7* transgenic plants displays male-sterility phenotype, these genes might play roles in anther or pollen development. Studying anther and pollen development in the model plant *A. thaliana* can provide new knowledge about male fertility and results can be applied to the other agronomically important plants.

More hints about the function of *SRF4*, *SRF5*, and *SRF6* were obtained with the yeast-two hybrid analysis. In addition, investigation of the putative interacting partners of the *SRF5* indicates that *SRF5* might play roles in cell plate formation. Especially, the putative interacting candidates *CalS1*, which has a direct role in cell plate formation, and *SEC5A* that play a role during vesicle sorting could support the suggested role of *SRF5*. Although the yeast two-hybrid system is very successful in identifying protein interacting partners, the data obtained from yeast and plant might differ. Therefore, in planta confirmation of the interaction is necessary. For this reason, co-immunoprecipitation, fluorescence resonance energy transfer (FRET), bimolecular fluorescence complementation (BiFC) or protoplast two-hybrid approaches can be used for future experiments to confirm and extend the yeast-two hybrid results of this study. In addition to these analyses, EGFP tagged *SRF* analyses could be carried out to find out the subcellular localization of the *SRF* members.

The results obtained in our study provide a solid basis for future analysis of the *SRF* functions.

## Chapter 7. References

- Akama, K., and Takaiwa, F.** (2007). C-terminal extension of rice glutamate decarboxylase (OsAD2) functions as an autoinhibitory domain and overexpression of a truncated mutant results in the accumulation of extremely high levels of GABA in plant cells. *J Exp Bot.* **58**, 2699-2707.
- Albrecht, C., Russinova, E., Hecht, V., Baaijens, E., and Vries, S.D.** (2005). The *Arabidopsis thaliana* SOMATIC EMBRYOGENESIS RECEPTOR-LIKE KINASES1 and 2 control male sporogenesis. *The Plant Cell* **17**, 3337-3349.
- Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J.** (1990). Basic local alignment search tool. *J Mol Biol* **215**, 403-410.
- Amano, Y., Tsubouchi, H., Shinohara, H., Ogawa, M., and Matsubayashi, Y.** (2007). Tyrosine-sulfated glycopeptide involved in cellular proliferation and expansion in *Arabidopsis*. *PNAS* **104**, 18333-18335.
- Anggono, V., and Robinson, P.J.** (2007). Syndapin I and endophilin I bind overlapping proline-rich regions of dynamin I: role in synaptic vesicle endocytosis. *J Neurochem.* **102**, 931-943.
- Asamizu, E., Nakamura, Y., Sato, S., and Tabata, S.** (2000). A large scale analysis of cDNA in *Arabidopsis thaliana*: Generation of 12,028 non-redundant expressed sequence tags from normalized and size selected cDNA libraries. *DNA Research* **7**, 175-180.
- Balasubramanian, S., and Schneitz, K.** (2000). NOZZLE regulates proximal-distal pattern formation, cell proliferation and early sporogenesis during ovule development in *Arabidopsis thaliana*. *Development* **127**, 4227-4238.
- Baumbush, L.O., Thorstensen, T., Krauss, V., Fischer, A., Naumann, K., Assalkhou, R., Schulz, I., Reuter, G., and Aalen, R.B.** (2001). The *Arabidopsis thaliana* genome contains at least 29 active genes encoding SET domain proteins that can be assigned to four evolutionary conserved classes. *Nucleic Acids Res.* **29**, 4319-4333.
- Becker, D., Boavida, L.C., Carneiro, J., Haury, M., and Feijo, J.A.** (2003). Transcriptional profiling of *Arabidopsis* tissues reveals the unique characteristics of the pollen transcriptome. *FEBS Lett.* **554**, 119-126.
- Berger, A., Meinhard, J., and Petersen, M.** (2006). Rosmarinic acid synthase is a new member of the superfamily of BAHD acyltransferases. *Planta* **224**, 1503-1510.
- Blanc, G., Hokamp, K., and Wolfe, K.H.** (2003). A recent polyploidy superimposed on older large-scale duplications in the *Arabidopsis thaliana*. *Genome Res.* **13**, 137-144.
- Bonete, M.J., Ferrer, J., Pire, C., Penades, M., and Ruiz, J.L.** (2000). 2-Hydroxyacid dehydrogenase from *Haloflex mediterranei*, a D-isomer-specific member of the 2-hydroxyacid dehydrogenase family. *Biochimie* **82**, 1143-1150.
- Breathnach, R., and Chambon, P.** (1981). Organization and expression of eukaryotic split genes coding for proteins. *Annu Rev Biochem.* **260**, 5055-5060.
- Brown, M.A., Sims III, R.J., Gottlieb, P.D., and Tucker, P.W.** (2006). Identification and characterization of Smyd2: a split SET/MYND domain-containing histone H3 lysine 36-specific methyltransferase that interacts with the Sin3 histone complex. *Molecular Cancer* **5**, 26.
- Bucher, M., Brunner, S., Zimmermann, P., Zardi, G.I., Amrhein, N., Willmitzer, L., and Reismeier, J.W.** (2002). The expression of an extensin-like protein correlates with cellular tip growth in tomato. *Plant Physiol.* **128**, 911-923.
- Caicedo, A.L., Schaal, B.A., and Kunkel, B.N.** (1999). Diversity and molecular evolution of the *RPS2* resistance gene in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* **96**, 302-306.

- Canales, C., Bhat, A.M., Scott, R., and Dickinson, H.** (2002). EXS, a putative receptor kinase, regulates male germline cell number and tapetal identity and promotes seed development in *Arabidopsis*. *Current Biology* **12**, 1718-1727.
- Castells, E., and Casacuberta, J.M.** (2007). Signaling through kinase-defective domains: the prevalence of atypical receptor-like kinases in plants. *J Exp Bot.* **58**, 3503-3511.
- Castells, E., Puigdomenech, P., and Casacuberta, J.M.** (2006). Regulation of the kinase activity of the MIK GCK-like MAPK4K by alternative splicing. *Plant Mol Biol.* **61**, 747-756.
- Chen E.Y., and Clarke, D.M.** (2002). The PEST sequence does not contribute to the stability of the cystic fibrosis transmembrane conductance regulator. *BMC Biochem.* **3**:29.
- Chevalier, D., Batoux, M., Fultin, L., Pfister, K., Yadav, R.K., Schellenberg, M., and Schneitz, K.** (2005). STRUBBELIG defines a receptor-kinase mediated signaling pathway regulating organ development in *Arabidopsis*. *Proc Natl Acad Sci USA.* **102**, 9074-9079.
- Chevalier, D., and Walker, J.C.** (2005). Functional genomics of protein kinases in plants. *Briefings in Functional Genomics and Proteomics* **3**, 362, 371.
- Chinchilla, D., Zipfel, C., Robatzek, S., Kemmerling, B., Nuernberger, T., Jones, J.D.G., Felix, G., and Boller, T.** (2007). A flagellin-induced complex of the receptor FLS2 and BAK1 initiates plant defense. *Nature* **448**, 497-501.
- Cho, H.T., and Cosgrove, D.J.** (2000). Altered expression of expansin modulates leaf growth and pedicel abscission in *Arabidopsis thaliana*. *Proc Natl Acad Sci U S A* **97**, 8783-9788.
- Clark, R.M., Schweikert, G., Toomajian, C., Ossowski, S., Zeller, G., Shinn, P., Warthman, N., Hu, T.T., Fu, G., Hinds, d.A., Chen, H., Frazer, K.A., Huson, D.H., Schoelkopf, B., Nordborg, M., Raetsch, G., Ecker, J.R., and Weigel, D.** (2007). Common sequence polymorphisms shaping genetic diversity in *Arabidopsis thaliana*. *Science* **317**, 338-342.
- Clark, S.E., Running, M.P., and Meyerowitz, E.M.** (1995). CLAVATA3 is a specific regulator of shoot and flower meristem development affecting the same processes as CLAVATA1. *Development* **121**, 2057-2067.
- Clark, S.E., Williams, R.W., and Meyerowitz, E.M.** (1997). The *CLAVATA1* genes encodes a putative receptor kinase that controls shoot and floral meristem size in *Arabidopsis*. *Cell* **89**, 575-585.
- Clough, S.J., and Bent, A.F.** (1998). Floral dip: a simplified method for *Agrobacterium* mediated transformation of *Arabidopsis thaliana*. *Plant J* **16**, 335-743.
- Cock, J.M., Vanoosthuysse, V., and Gaude, T.** (2002). Receptor kinase signaling in plants and animals: distinct molecular systems with mechanistic similarities. *Current Opinion in Cell Biology* **14**, 230-236.
- Colcombet, J., Boisson-Dernier, A., Ros-Palau, R., Vera, C.E., and Schroeder, J.I.** (2005). *Arabidopsis* SOMATIC EMBRYOGENESIS RECEPTOR KINASES1 and 2 are essential for tapetum development and microspore maturation. *The Plant Cell* **17**, 3350-3361.
- Cookson, S.J., Radziejowski, A., and Granier, C.** (2006). Cell and leaf size plasticity in *Arabidopsis*: what is the role of endoreduplication. *Plant Cell Environ.* **29**, 1273-1283.
- Cooper, E.** (2001). A new role for ion channels in myoblast fusion. *The Journal of Cell Biology* **153**, F9-F11.
- Corpet, D.E.** (1988). Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.* **16**, 10881-10890.
- Curran, A.R., and Engelman, D.M.** (2003). Sequence motifs, polar interactions and conformational changes in helical membrane proteins. *Curr Opin Struc Biol.* **13**, 412-417.

- Derbyshire, P., McCann, M. C., and Roberts, K.** (2007). Restricted cell elongation in *Arabidopsis* hypocotyls is associated with a reduced average pectin esterification level. *BMC Plant Biology* **7**, 1-12.
- DeYoung, B.J., Bickle, K.L., Schrage, K.J., Patel, K., and Clark, S.E.** (2006). The CLAVATA1-related BAM1, BAM2 and BAM3 receptor kinase-like proteins are required for meristem function in *Arabidopsis*. *The Plant Journal* **45**, 1-16.
- Domenech, J., and Ferrer, J.** (2006). A new D-2-hydroxyacid dehydrogenase with dual coenzyme-specificity from *Haloflex mediterranei*, sequence analysis and heterologous overexpression. *Biochimica et Biophysica Acta* **1760**, 1667-1674.
- Donnelly, P.M., Bonetta, D., Tsukaya, H., Dengler, R.E., and Dengler, N.G.** (1999). Cell cycling and cell enlargement in developing leaves of *Arabidopsis*. *Developmental Biology* **215**, 407-419.
- Dunning, F.M., Sun, W., Jansen, K.L., Helft, L., and Bent, A.F.** (2007). Identification and mutational analysis of *Arabidopsis* FLS2 leucine-rich repeat domain residues that contributes to flagellin perception. *The Plant Cell*
- Ehlert, A., Weltmeier, F., Wang, X., Mayer, C.S., Smeekens, S., Vicente-Carbajosa, J., and Droege-Laser, W.** (2006). Two-hybrid protein-protein interaction analysis in *Arabidopsis* protoplasts: establishment of a heterodimerization map of group C and group S bZIP transcription factors. *Plant J.* **46**, 890-900.
- Evans, J.H., Murray, D., Leslie, C.C., and Falke, J.J.** (2006). Specific translocation of protein kinase Calpha to the plasma membrane requires both Ca<sup>2+</sup> and PIP2 recognition by its C2 domain. *Mol Biol Cell* **17**, 56-66.
- Eyüboğlu, B., Pfister, K., Haberer, G., Chevalier, D., Fuchs, A., Mayyer, K.F., and Schneitz, K.** (2007). Molecular characterization of the STRUBBELIG-RECEPTOR FAMILY of genes encoding putative leucine-rich repeats receptor-like kinases in *Arabidopsis thaliana*. *BMC Plant Biol.* **7**, 16.
- Fletcher, J.C., Brand, U., Running, M.P., Somin, R., and Meyerowitz, E.M.** (1999). Signaling of cell fate decision by *CLAVATA3* in *Arabidopsis* shoot meristems. *Science* **283**, 1911-1914.
- Folkers, U., Berger, J., and Huelskamp, M.** (1997). Cell morphogenesis of trichomes in *Arabidopsis*: differential regulation of primary and secondary branching by branch initiation regulators and cell size. *Development* **124**, 3779-3786.
- Folkers, U., Kirik, V., Schoebinger, U., Falk, S., Krishnakumar, S., Pollock, M.A., Oppenheimer, D.G., Day, I., Reddy, A.S.N., Juergens, G., and Huelskamp M.** (2002). The cell morphogenesis gene *ANGUSTIFOLIA* encodes a CtBP/BARS-like protein and is involved in the control of the microtubule cytoskeleton. *EMBO J.* **21**, 1280-1288.
- Friedrichsen, D.M., Joazeiro, C.A., Li, J., Hunter, T., and Chory, J.** (2000). Brassinosteroid-insensitive-1 is a ubiquitously expressed leucine-rich repeat receptor serine/threonine kinase. *Plant Physiol.* **123**, 1247-1256.
- Fritz-Laylin, L.K., Krishnamurty, N., Toer, M., Sjoelander, K.V., and Jones, J.D.** (2005). Phylogenomic analysis of the receptor-like proteins of rice and *Arabidopsis*. *Plant Physiol.* **138**, 611-623.
- Fujikura, U., Horiguchi, G., and Tsukaya, H.** (2007). Dissection of enhanced cell expansion processes in leaves triggered by a defect in cell proliferation, with reference to roles of endoreduplication. *Plant Cell Physiol.* **48**, 278-286.
- Fukuda, M., Kojima, T., Kabayama, H. and Mikoshiba, K.** (1996). Mutation of the pleckstrin homology domain of Brutons tyrosine kinase in immunodeficiency impaired inositol 1.3.4.5-tetrakisphosphate binding capacity. *J. Biol. Chem.* **271**, 30303-30306.
- Gendreau, E., Traas, J., Desnos, T., Grandjean, O., Caboche, M., and Hoeffte, H.** (1997). Cellular basis of hypocotyl growth in *Arabidopsis thaliana*. *Plant Physiol.* **114**, 2950-2955.

- Goff, S.A., Ricke, D., Lan, T.H., Presting, G., Wang, R., Dunn, M., Glazebrook, J., Sessions, A., Oeller, P., Varma, H. et al.** (2002). A draft sequence of the rice genome (*Oryza sativa* L. spp japonica). *Science* **296**, 92-100.
- Gomez-Gomez, L., Bauer, Z., and Boller, T.** (2001). Both the extracellular leucine-rich repeat domain and the kinase domain of FLS2 are required for flagellin binding and signaling in *Arabidopsis*. *Plant Cell* **13**, 1155-1163.
- Gomez-Gomez, L., and Boller, T.** (2000). FLS2: an LRR receptor-like kinase involved in the perception of the bacterial elicitor flagellin in *Arabidopsis*. *Mol Cell* **5**, 1003-1011.
- Goto, K., Kyojuka, J., and Bowman J.L.** (2001). Turning floral organs into leaves, leaves into floral organs. *Curr. Opin. Genet. Dev* **11**, 449-456.
- Gubler, U., and Hofmann, B.J.** (1983). A simple and very efficient method for generating cDNA libraries. *Gene* **25**, 263-269.
- Guerrero-Valero, M., Marin, V.C., Gomez-Fernandez, J.C., and Corbalan-Garcia, S.** (2007). The C2 domain of classical PKCs are specific PtdIns(4,5)P<sub>2</sub>-sensing domains with different domain. *J. Mol. Biol.* **371**, 608-621.
- Guittou, A.E., and Berger, F.** (2005). Control of reproduction by polycomb group complexes in animals and plants. *Int. J. Dev. Biol.* **49**, 202-216.
- Hammond-Kosack, K.E., and Jones, J.D.** (1997). Plant disease resistance genes. *Annu Rev Plant Physiol Plant Mol Biol.* **48**, 575-607.
- Hanks, S.K., and Hunter, T.** (1995). Protein kinases 6. The eukaryotic protein kinase superfamily: kinase (catalytic) domain structure and classification. *FASEB J.* **9**, 576-596.
- Hanks, S.K., and Quinn, A.M.** (1991). Protein kinase catalytic domain sequence database: identification of conserved features of primary structure and classification of family members. *Methods Enzymol.* **88**, 315-319.
- Hanks, S.K., Quinn, A.M., and Hunter, T.** (1988). The protein kinase family: conserved features and deduced phylogeny of the catalytic domains. *Science* **241**, 42-52.
- Hanson, J., Johannesson, H., Engstrom, P.** (2001). Sugar-dependent alterations in cotyledon and leaf development in transgenic plants expressing HDZip gene *ATHB13*. *Plant. Mol. Biol.* **45**, 247-262.
- He, K., Gou, X., Yuan, T., Lin, H., Asami, T., Yoshida, S., Russell, S.D., and Li, J.** (2007). BAK1 and BKK1 regulate brassinosteroid-dependent growth and brassinosteroid-independent cell-death pathways. *Current Biology* **17**, 1109-1115.
- Heese, A., Hann, D.R., Gimenez-Ibanez, S., Jones, A.M.E., He, K., Li, J., Schroeder, J.I., Peck, S.C., and Rathjen, J.P.** (2007). The receptor-like kinase SERK3/BAK1 is a central regulator of innate immunity in plants. *PNAS* **104**, 12217-12222.
- Hilson, P., Small, I., and Kuiper, M.T.** (2003). European consortia building integrated resources for *Arabidopsis* functional genomics. *Curr Opin Plant Biol* **6**, 426-429.
- Hong, Z., Delauney, A.J., and Verma, D.P.** (2001). A cell plate-specific callose synthase and its interaction with phragmoplastin. *Plant Cell* **13**, 755-768.
- Honma, T., and Goto, K.** (2001). Complexes of MADS-box proteins are sufficient to convert leaves into floral organs. *Nature* **409**, 525-529.
- Hord, C.L.H., Chen, C., DeYoung, B.J., Clark, S.E., and Ma, H.** (2006). The BAM1/BAM2 receptor-like kinases are important regulators of *Arabidopsis* early anther development. *The Plant Cell* **18**, 1667-1680.
- Horiguchi, G., Fujikura, U., Ferjani, A., Ishikawa, N., and Tsukaya H.** (2006). Large-scale histological analysis of leaf mutants using two simple leaf observation methods: identification of novel genetic pathways governing the size and shape of leaves. *Plant J* **48**, 638-644.
- Horiguchi, G., Kim, G.T., and Tsukaya, H.** (2005). The transcription factor AtGRF5 and the transcription coactivator AN3 regulate cell proliferation in leaf primordia of *Arabidopsis thaliana*. *The Plant Journal* **43**, 68-78.

- Horton, R.M., Hunt, H.D., Ho, S.N., Pullen, J.K., and Pease, L.R. (1989). Engineering hybrid genes without the use of restriction enzymes: gene splicing by overlap extension. *Gene* **77**, 61-68.
- Hu, Y., Poh, H.M., and Chua, N.H. (2006). The *Arabidopsis* ARGOS-LIKE gene regulates cell expansion during organ growth. *Plant J* **47**, 1-9.
- Hu, Y., Xie, Q., Chua, N.H. (2003). The *Arabidopsis* auxin-inducible gene ARGOS controls lateral organ size, *Plant Cell* **15**, 1951-1961.
- Hubbard, S.R., and Till, J.H. (2000). Protein tyrosine kinase structure and function. *Annu Rev Biochem* **69**, 373-398.
- Huse, M., and Kuriyan, J. (2002). The conformational plasticity of protein kinases. *Cell* **109**, 275-282.
- Ingram, G.C. (2007). Cell signaling: The merry lives of BAK1. *Current Biology* **17**, No 15.
- Ivashchenko, A.T., and Atambaeva, Sh.A. (2004). Variation in lengths of introns and exons in genes of the *Arabidopsis thaliana* nuclear genome. *Genetika* **40**, 1429-1431.
- James, P., Halladay, J., and Craig, E.A. (1996). Genomic libraries and a host strain designed for highly efficient two-hybrid selection in yeast. *Genetics* **144**, 1425-1436.
- Jander, G., Norris, S.R., Rounsley, S.D., Bush, D.F., Levin, I.M., and Last, R.L. (2002). *Arabidopsis* map-based cloning in the post-genome era. *Plant Physiol.* **129**, 440-450.
- Jenik, P.D., and Irish, V.F. (2000). Regulation of cell proliferation patterns by homeotic genes during *Arabidopsis* floral development. *Development* **127**, 1267-1276.
- Jeong, S., Trotochaud, A.E., and Clark, S.E. (1999). The *Arabidopsis* CLAVATA2 gene encodes a receptor-like protein required for the stability of the CLAVATA1 receptor-like kinase. *Plant Cell* **11**, 1925-1934.
- Johnson, K.L., and Ingram, G.C. (2005). Sending the right signals: regulating receptor kinase activity. *Current Opinion in Plant Biology* **8**, 648-656.
- Johnson, L.N., Noble, M.E.M., and Owen, D.J. (1996). Active and inactive protein kinases: structural basis for regulation. *Cell* **85**, 149-159.
- Jones, A.M., Im, K.H., Savka, M.A., Wu, M.J., DeWitt, N.G., Shillito, R., and Binns, A. (1998). Auxin-dependent cell expansion mediated by overexpressed Auxin-Binding Protein1. *Science* **282**, 1114-1117.
- Jones, D.A., and Jones, J.D.G. (1997). The role of leucine-rich repeats in plant defences. *Adv. Bot. Res.* **24**, 90-167.
- Kajava, A.V. (1998). Structural diversity of leucine-rich repeat proteins. *J. Mol. Biol.* **277**: 519-527.
- Kawchuk, L.M., Hachey, J., Lynch, D.R., Kulcsar, F., van Rooijen, g., Waterer, D.R., Robertson, A., Kokko, E., Byers, R., Howard, R.J., Fischer, R., and Pruffer, D. (2001). Tomato Ve disease resistance gene encode cell-surface-like receptors. *Proc Natl Acad Sci U S A* **98**, 6511-6515.
- Kayes, J.M., and Clark, S.E. (1998). CLAVATA2, a regulator of meristem and organ development in *Arabidopsis*. *Development* **125**, 3843-3851.
- Kemmerling, B., Schwedt, A., Rodrigues, P., Mazzotta, S., Frank, M., Qamar, S.A., Mengiste, T., Betsuyaku, S., Parker, J.E., Muessig, C., Thomma, B.P.H.J., Albrecht, C., de Vries, S.C., Hirt, H., and Nuernberger, T. (2007). The BRI1-associated kinase 1, BAK1, has a brassinolide-independent role in plant cell-death control. *Current Biology* **12**, 1116-1122.
- Kim, G.T., Fujioka, S., Kozuka, T., Tax, F.E., Takatsuto, S., Yoshida, S., and Tsukaya, H. (2005). CYP90C1 and CYP90D1 are involved in different steps in brassinosteroids biosynthesis pathway in *Arabidopsis thaliana*. *Plant J.* **41**, 710-721.
- Kim, G.T., Shoda, K., Tsuge, T., Cho, K.G., Uchimiya, H., Yokoyama, R., Nishitani, K., and Tsukaya, H. (2002). The *ANGUSTIFOLIA* gene of *Arabidopsis*, a plant

- CtBP gene, regulates leaf cell expansion, the arrangement of cortical microtubules in leaf cells and expression of a gene involved in cell-wall formation. *EMBO J.* **21**, 1267-1279.
- Kim, G.T., Tsukaya, H., and Uchimiya H.** (1998). The *ROTUNDIFOLIA3* gene of *Arabidopsis thaliana* encodes a new member of the cytochrome P-450 family that is required for the regulated polar elongation of leaf cells. *Genes Dev.* **12**, 2181-2191.
- Kim, H.B., Schaller, H., Goh, C-H., Kwon, M., Choe, S., An, C.S., Durst, F., Feldmann, K.A., and Feyereisen, R.** (2005). *Arabidopsis cyp51* mutant shows postembryonic seedling lethality associated with lack of membrane integrity. *Plant Physiology* **138**, 2033-2047.
- Kim, J.H., Choi, D., and Kende, H.** (2003). The AtGRF family of putative transcription factors is involved in leaf and cotyledon growth in *Arabidopsis*. *Plant J* **36**, 94-104.
- Kim, K.H., Janiak, V., and Petersen, M.** (2004). Purification, cloning and functional expression of hydroxyphenylpyruvate reductase involved in rosmarinic acid biosynthesis in cell cultures of *Coleus blumei*. *Plant Mol Biol.* **52**, 311-323.
- Kim, Y., Tsukaya, H., and Uchimiya, H.** (1998). The *CURLY LEAF* gene controls both division and elongation of cells during expansion of the leaf blade in *Arabidopsis thaliana*. *Planta* **206**, 175-183.
- Kim, Y., Zhang, H., and Scholl, R.L.** (1990). Two evolutionary divergent genes encode a cytoplasmic ribosomal protein of *Arabidopsis thaliana*. *Gene* **93**, 177-82.
- Kinoshita, T., Cano-Delgado, S.H., Hiranuma, S., Fijioka, S., Yoshida, S., and Chory J.** (2005). Binding of brassinosteroids to the extracellular domain of plant receptor kinase BRI1. *Nature* **433**, 167-171.
- Kobe, B., and Deisenhofer, J.** (1993). Crystal structure of porcine ribonuclease inhibitor, a protein with leucine-rich repeats. *Nature* **266**, 751-756.
- Kolkman, J.A., and Stemmer, W.P.** (2001). Directed evolution of proteins by exon shuffling. *Nat Biotechnol.* **19**, 423-428.
- Konez, C., and Schell, J.** (1986). The promoter TI-DNA gene 5 controls the tissue specific expression of chimeric genes carried by a novel type of *Agrobacterium* binary vector. *Mol Gen Genet.* **204**, 383-396.
- Kornitzer, D., Raboy, B., Kulka, R.G., and Fink, G.R.** (1994). Regulated degradation of the transcription factor Gcn4. *EMBO J.* **13**, 6021-6030.
- Kumar, S., Tamura, K., and Nei, M.** (2004). MEGA3: integrated software for molecular evolutionary genetics analysis and sequence alignment. *Brief Bioinform.* **5**, 150-163.
- Kwak, S.H., and Schiefelbein, J.** (2007). The role of SCRAMBLED receptor-like kinase in patterning the *Arabidopsis* root epidermis. *Dev. Biol.* **302**, 118-131.
- Kwak, S.H., Shen, R., and Schiefelbein, J.** (2005). Positional signaling mediated by a receptor-like kinase in *Arabidopsis*. *Science* **307**, 1111-1113.
- Lachance, P.E., Miron, M., Raught, B., Sonenberg, N., and Lasko, P.** (2002). Phosphorylation of eukaryotic translational factor 4E is critical for growth. *Mol Cell Biol* **22**, 1656-1663.
- Larkin, J.C., Oppenheimer, D.G., Lloyd, A.M., Pappozzi, E.T., and Marks, M.D.** (1994). Roles of the GLABROUS1 and TRANSPARENT TESTA GLABRA genes in *Arabidopsis* trichome development. *Plant Cell* **6**, 1065-1076.
- Larkins, B.A., Dilkes, B.P., Dante, R.A., Coelho, C.M., Woo, Y., and Liu, Y.** (2001). Investigating the hows and whys of the DNA endoreduplication. *J Exper Botany.* **52**, 183-192.
- Lee H.S., Karunanandaa, B., McCubbin, A., Gilroy, S., and Kao, T.** (1996). PRK1, a receptor-like kinase of *Petunia inflata*, is essential for postmeiotic development of pollen. *The Plant Journal* **9**, 613-624.
- Li, J., and Chory, J.** (1997). A putative leucine-rich repeat receptor kinase involved in brassinosteroid signal transduction. *Cell* **90**, 929-938.

- Li, J., Wen, J., Lease, K.A., Doke, J.T., Tax, F.E., Walker, J.C.** (2002). BAK1, an *Arabidopsis* LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling. *Cell* **110**, 213-222.
- Li, S.S.-C.** (2005). Specificity and versatility of SH3 and other proline-recognition domains: structural basis and implications for cellular signal transduction. *Biochem J.* **390**, 641-653.
- Li, Z., and Wurtzel, E.** (1998). The *ltk* gene family encodes novel receptor-like kinases with temporal expression in developing maize endosperm. *Plant Mol Biol* **5**, 749-761.
- Llompart, B., Castells, E., Rio, A., Roca, R., Ferrando, A., Stiefel, V., Puigdomenech, P., and Casacuberta J.M.** (2003). The direct activation of MIK, a germinal center kinase (GCK)-like kinase, by MARK, a maize atypical receptor kinase, suggest a mechanism for signaling through kinase-dead receptors. *J Mol Chem.* **278**, 48105-48111.
- Llorente, F., Alonso-Blanco, C., Sanchez-Rodriguez, C., Jorda, L., and Molina, A.** (2005). ERECTA receptor-like kinase and heterotrimeric G protein from *Arabidopsis* are required for resistance to the necrotrophic fungus *Plectosphaerella cucumerina*. *The Plant Journal* **43**, 165-180.
- Lodge, J., Lund, p., and Minchin, S.** (2007). *Gene cloning: Principles and applications.* MPG BOOKS limited, Bodmin, Cornwall, UK.
- Matsubayashi, Y., Ogawa, M., Morita, A., and Sakagami, Y.** (2002). An LRR receptor kinase involved in perception of peptide plant hormone, phytosulfokine. *Science* **296**, 1470-1472.
- Matsumoto, T., Wu, J., Baba, T., Katayose, Y., Yamamoto, K., Skata, K., Yano, M., and Sasaki, T.** (2001). Rice genomics: current status of genome sequencing. *Novartis Found Symp* **236**, 28-38.
- McCarty, D.R., and Chory, J.** (2000). Conservation and innovation in plant signaling pathways. *Cell* **103**, 201-209.
- McCormick, S.** (2004). Control of male gametophyte development. *The Plant Cell* **16**, 142-153.
- Micheli, F.** (2001). Pectin methylesterases: cell wall enzymes with important roles in plant physiology. *TRENDS in Plant Science* **6**, 414-419.
- Mizukami, Y., and Fischer, R.L.** (2000). Plant organ size control: AINTEGUMENTA regulates growth and cell numbers during organogenesis. *Proc Natl Acad Sci U S A* **97**, 942-947.
- Monzingo, A.F., Dhaliwal, S., Dutt, C.A., Lyon, A., Sadow, J.H., Hoffman, D.W., Robertus, J.D., and Brownings, K.S.** (2007). The structure of eukaryotic translation initiation factor-4E from wheat reveals a novel disulfide bond. *Plant Physiol.* **143**, 1504-1518.
- Murashige, T., and Skoog, F.** (1962). A revised medium for rapid growth and bioassays with tobacco tissue cultures. *Physiol Plant* **15**, 473-497.
- Murray, M.G., and Thompson, W.F.** (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **2**, 469-470.
- Murthy, M., and Schwartz, T.L.** (2004). The exocyst component Sec% is required for membrane trafficking and polarity in the *Drosophila* ovary. *Development* **131**, 377-388.
- Muschietti, J., Dircks, L., Vancanneyt, G., and McCormick, S.** (1994). LAT52 protein is essential for tomato pollen development: pollen expressing antisense *LAT52* RNA hydrates and germinates abnormally and cannot achieve fertilization. *Plant J* **6**, 321-338.
- Muschietti, J., Eyal, Y., and McCormick, S.** (1998). Pollen tube localization implies a role in pollen-pistil interactions for the tomato receptor-like protein kinases LePRK1 and LePRK2. *The Plant Cell* **10**, 319-330.
- Murray, M.G., and Thompson, W.F.** (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* **2**, 469-470.

- Nadeu, J.A., and Sack, F.D.** (2002). Control of stomatal distribution on the Arabidopsis leaf surface. *Science* **296**, 1697-1700.
- Nam, K.H. and Li, J.** (2002). BRI1/BAK1, a receptor kinase pair mediating brassinosteroid signaling. *Cell* **110**, 203-212.
- Naranda, T., Macmillan, S.E., Donahue, T.F., Hershey, J.W.B.** (1996). SUI1/p16 I required for the activity of eukaryotic translation initiation factor 3 in *Saccharomyces cerevisiae*. *Molecular and Cellular Biology* **16**, 2307-2313.
- Narita, N.N., Moore, S., Horiguchi, G., Kubo, M., Demura, T., Fukuda H., Goodrich, J., and Tsukaya, H.** (2004). Overexpression of a novel small peptide ROTUNDIFOLIA4 decreases cell proliferation and alters leaf shape in *Arabidopsis thaliana*. *Plant J* **38**, 699-713.
- Nei, M., and Gojobori, T.** (1986). Simple methods for estimating the numbers of synonymous and nonsynonymous nucleotide substitutions. *Molecular Biology and Evolution* **3**, 418-426.
- Ng, D.W-K., Wang, T., Chandrasekharan, M.B., Aramayo, R., Kertbundit, S., and Hall, T.C.** (2007). Plant SET domain containing proteins: Structure, function and regulation. *Biochim. Et Biophys. Acta* **1769**, 316-329.
- Noel, L., Moores, T.L., Van Der Biezen, E.A., Parniske, M., Daniels, M.J., Parker, J.E., and Jones, J.D.** (1999). Pronounced intraspecific haplotype divergence at the *RPP5* complex disease resistance locus of *Arabidopsis*. *Plant Cell* **11**, 2099-2112.
- Nordborg, M., Hu, T.T., Ishino, Y., Jhaveri, J., Toomajian, C., Zheng, H., Bakker, E., Clabrese, P., Gladstone, J., Goyal, R., Jakobsson, M., Kim, S., Morozov, Y., Padhukasahasram, B., Plagnol, V., Rosenberg, N.A., Shah, C., Wall, J.D., Wang, J., Zhao, K., Kalbfleish, T., Schulz, V., Kreitman, M., and Bergelson, J.** (2005). The pattern of polymorphism in *Arabidopsis thaliana*. *Plos Biol.* **3**, e196.
- Nühse, T.S., Stensballe, A., Jensen, O.N., and Peck, S.C.** (2004). Phosphoproteomics of the Arabidopsis plasma membrane and a new phosphorylation site database. *Plant Cell* **16**, 2394-205.
- Page, R.D.** (1996). TreeView: an application to display phylogenetic trees on personal computers. *Comput Appl Biosci.* **12**, 357-358.
- Park, A.R., Cho, S.K., Yun, U.J., Jin, M.Y., Lee, S.H., Sachetto-Martins, G., and Park, O.G.** (2001). Interaction of the Arabidopsis receptor protein kinase Wak1 with a glycine-rich protein, AtGRP-3. *J Biol Chem* **276**, 26688-26693.
- Pawson, T., and Nash, P.** (2003). Assembly of cell regulatory systems through protein interaction domains. *Science* **300**, 445-452.
- Pereira, L. A. R., Schoor, S., Goubet, F., Dupree, P., and Moffatt, B. A.** (2006). Deficiency of adenosine kinase activity affects the degree of pectin methylesterification in cell walls of *Arabidopsis thaliana*. *Planta* **224**, 1401-1414.
- Petersen, M., and Simmonds, M.S.** (2003). Rosmarinic acid. *Phytochemistry* **62**, 121-125.
- Preuss, D., Rhee, S.Y., and Davis, R.W.** (1994). Tetrad analysis possible in *Arabidopsis* with mutation of the QUARTET (QRT) genes. *Science* **264**, 1458-1460.
- Prugganan, M.D., and Suddith, J.I.** (1999). Molecular population genetics of floral homeotic loci. Departures from the equilibrium-neutral model at the *APETALA3* and *PISTILLATA* genes of *Arabidopsis thaliana*. *genetics* **151**, 839-848.
- Raught, B., and Gingras, A.C.** (1999). eIF4E activity is regulated at multiple levels. *Int J Biochem Cell Biol* **31**, 270-279.
- Rechsteiner, M., and Rogers, S.W.** (1996). PEST sequences and regulation by proteolysis. *Trends Biochem Sci.* **21**, 267-271.
- Reverte, C.G., Ahearn, M.D., and Hake, L.E.** (2001). CPEB degradation during *Xenopus* oocytes maturation requires a PEST domain and the 26S proteasome. *Dev Biol.* **231**, 447-457.

- Rhee, S.Y., and Somerville, C.R. (1998). Tetrad pollen formation in quartet mutants of *Arabidopsis thaliana* is associated with persistence of peptic polysaccharides of the pollen mother cell wall. *Plant J* **15**, 79-88.
- Richard, L., Qin, L.X., Gadal, P., and Goldberg, R. (1994). Molecular cloning and characterization of putative pectin methylesterase cDNA in *Arabidopsis thaliana* (L.). *FEBS Letters* **355**, 135-139.
- Robatzek, S., Chinchilla, D., and Boller, T. (2006). Ligand-induced endocytosis of the pattern recognition receptor FLS2 in *Arabidopsis*. *Genes Dev* **20**, 537-542.
- Rogers, S., Wells, R., and Rechsteiner, M. (1986). Amino acid sequences common to rapidly degraded proteins: the PEST hypothesis. *Science* **234**, 364-368.
- Ron, M., and Avni, A. (2004). The receptor for the fungal elicitor ethylene-inducing xylanase is a member of a resistance-like gene family in tomato. *Plant Cell* **16**, 1604-1615.
- Ronquist, F., and Huelsenbeck, J.P. (2003). MrBayes 3: Bayesian phylogenetic inference under mixed models. *Bioinformatics* **19**, 1572-1574.
- Ross, K.J., Franz, P., Armstrong, S.J., Vizir, I., Mulligan, B., Franklin, F.C., and Jones, G.H. (1997). Cytological characterization of four meiotic mutants of *Arabidopsis* isolated from T-DNA transformed lines. *Chromosome Res.* **5**, 551-559.
- Saibo, N.J.M., Vriezen, W.H., Beemster, G.T.S., and Van der Straeten D. Growth and stomata development of *Arabidopsis* hypocotyls are controlled by gibberellins and modulated by ethylene and auxins. *Plant J* **33**, 989-1000.
- Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989). *Molecular cloning, a laboratory manual*. Cold Spring Harbor Laboratory Press.
- Sanders, P.M., Bui, A.Q., Weterings, K., MacIntire K.N., Hsu, Y-C., Lee, P.Y., Truong, M.T., Beals, T.P., and Goldberg, R.B. (1999). Anther developmental defects in *Arabidopsis thaliana* male-sterile mutants. *Sex Plant Repro* **11**, 297-322.
- Sanders, P.M., Lee, P.Y., Biesgen, C., Boone, J.D., Beals, R.P., Weiler, E.W., and Goldberg, R.B. (2000). The *Arabidopsis* *DELAYED DEHISCENCE1* gene encodes an enzyme in the jasmonic acid synthesis pathway. *The Plant Cell* **12**, 1041-1061.
- Sandhu, G.S., Aleff, R.A., and Kline, B.C. (1992). Dual asymmetric RCR: one-step construction of synthetic genes. *Biotechniques* **12**, 14-16.
- Scheeff, E.D., and Bourne, P.E. (2005). Structural evolution of the kinase-like superfamily. *PLoS Comput Biol.* **1**, e49.
- Schieffthaler, U., Balasubramanian, S., Sieber, P., Chevalier, D., Wisman, E., and Schneitz, K. (1999). Molecular analysis of *NOZZLE*, a gene involved in pattern formation and early sporogenesis during sex organ development in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. USA* **96**, 11664-11669.
- Schmidt, E.D., Guzzo, F., Toonen, M.A., and de Vries, S.C. (1997). A leucine-rich repeat containing receptor-like kinase marks somatic plant cells competent to form embryos. *Development* **214**, 2049-2062.
- Schmidt, H.A., Strimmer, K., Vingron, M., and Von Haeseler, A. (2002). TREE-PUZZLE: maximum likelihood phylogenetic analysis using quartets and parallel computing. *Bioinformatics* **18**, 502-504.
- Schneitz, K., Huelskamp, M., Kopczak, S.D., and Pruitt, R.E. (1997). Dissection of sexual organ ontogenesis: a genetic analysis of ovule development in *Arabidopsis thaliana*. *Development* **124**, 1367-1376.
- Scott, R.J., Spielman, M., and Dickinson, H.G. (2004). Stamen structure and function. *Plant Cell (Suppl)* **16**, S46-S60.
- Schopfer, C.R., Nasrallah, M.E., and Nasrallah, J.B. (1999). The male determinant of self-incompatibility in *Brassica*. *Science* **286**, 1697-1700.
- Seki, M., Narusaka, M., Kamiya, A., Ishida, J., Satou, M., Sakurai, T., Nakajima, M., Enju, A., Muramatsu, M., Hayashizaki, Y., Kawai, J., Carninci, P., Itoh, M.,

- Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., and Shinozaki, K. (2002). Functional annotation of a full-length Arabidopsis cDNA collection. *Science* **296**, 141-145.
- Shih, M.C., Heinrich, P., and Goodmann, H.M. (1991). Cloning and chromosomal mapping of nuclear genes encoding chloroplast and cytosolic glyceraldehyde-3-phosphate dehydrogenase from *Arabidopsis thaliana*. *Gene* **104**, 133-138.
- Showalter, A.M. (1993). Structure and function of plant cell wall proteins. *Plant Cell* **5**, 9-23.
- Shpak, E.D., Berthiaume, C.T., Hill, E.J., and Torii, K.U. (2004). Synergistic interaction of three ERECTA-family receptor-like kinases controls *Arabidopsis* organ growth and flower development by promoting cell proliferation. *Development* **131**, 1491-1501.
- Shpak, E.D., McAbee, J.M., Pillitteri, L.J., and Torii, K.U. (2005). Stomatal patterning and differentiation by synergistic interactions of receptor kinases. *Science*, **309**, 290-293.
- Shiu, S.H., and Bleecker, A.B. (2001a). Plant receptor-like kinase family: diversity, function, and signaling. *Sci STKE* **113**, RE22.
- Shiu, S.H., and Bleecker, A.B. (2001b). Receptor-like kinases from *Arabidopsis* form a monophyletic gene family related to animal receptor kinases. *PNAS* **98**, 10763-10768.
- Shiu, S.H., and Bleecker, A.B. (2003). Expansion of the receptor-like kinase/Pelle gene family and receptor-like proteins in Arabidopsis. *Plant Physiology* **132**, 530-543.
- Shiu, S.H., Karlowski, W.M., Pan, R., Tzeng, Y.H., Mayer, K.F., and Li, W.H. (2004). Comparative analysis of the receptor-like kinase family in *Arabidopsis* and rice. *Plant Cell* **16**, 1220-1234.
- Silva, N.F., and Goring, D.R. (2002). The proline-rich, extensin-like receptor-like kinase-1 (PERK1) gene is rapidly induced by wounding. *Plant Mol. Biol.* **50**, 667-685.
- Shpak, E.D., Berthiaume, C.T., Hill, E.J., and Torii, K.U. (2004). Synergistic interaction of the three ERECTA-family receptor-like kinases controls *Arabidopsis* organ growth and flower development by promoting cell proliferation. *Development* **131**, 1491-1501.
- Shpak, E.D., McAbee, J.M., Pillitteri, L.J., and Torii, K.U. (2005). Stomatal patterning and differentiation by synergistic interaction of receptor kinases. *Science* **309**, 290-293.
- Singh, R., and Green, M.R. (1993). Sequence-specific binding of transfer RNA by glyceraldehyde-3-phosphate dehydrogenase. *Science* **259**, 365-368.
- Sinha, N. (1999). Leaf development in angiosperms. *Annu. Rev. Plant Physiol.* **50**, 419-446.
- Smyth, D.V., Bowman, J.L., and Meyerowitz, E.M. (1990). Early flower development in *Arabidopsis*. *The Plant Cell* **2**, 755-767.
- Snyders, S., and Kohorn, B.D. (1999). TAKs, thylakoid membrane protein kinases associated with energy transduction. *J. Biol. Chem.* **274**, 9137-9140.
- Soellick, T.R., and Uhrig, J.F. (2001). Development of an optimized interaction-mating protocol for large-scale yeast two-hybrid analyses. *Genome Biol.* **2**, 0052.
- Sommer, B., Oprins, A., Rabouille, C., and Munro, S. (2005). The exocyst component Sec5 is present on endocytic vesicle in the oocyte of *Drosophila melanogaster*. *J. Cell. Biol.* **169**, 953-963.
- Song, W.Y., Wang, G.L., Chen, L.L., Kim, H.S., Pi, L.Y., Holsten, T., Gardner, J., Wang, B., Zhai, W.X., Zhu, L.H., Fauquet, C., and Ronald, P. (1995). A receptor kinase-like protein encoded by the rice disease resistance gene, Xa21. *Science* **270**, 1804-1806.

- Stein, R.A., and Staros, J.V.** (2000). Evolutionary analysis of the ErbB receptor and ligand families. *J Mol Evol.* **50**, 397-412.
- Stevenson, L.F., Kennedy, B.K., and Harlow, E.** (2001). A large-scale overexpression screen in *Saccharomyces cerevisiae* identifies previously uncharacterized cell cycle genes. *PNAS* **98**, 3946-3951.
- Strabala., T.J., O'donell, P.J., Smit, A.M., Ampomah-Dwamena, C., Maartin, E.J., Netzler, N., Nieuwenhuizen, N.J., Quinn, B.D., Foote, H.C., and Hudson, K.R.** (2006). Gain-of-function of many CLAVATA3/ESR genes, including four new family members, correlate with the tandem variations in the conserved CLAVATA3/ESR domain. *Plant Physiol.* **140**, 1331-1344.
- Strahl, B.D., Grans, P.A., Briggs, S.D., Sun, Z-W., Bone, J.R., Caldwell, J.A., Sahana, M., Cook, R.G., Shabanowitz, J., Hunt, D.F., and Allis, C.D.** (2002). Set2 is a nucleosomal histone H3-selective methyltransferase that mediates transcriptional repression. *Molecular and Cellular Biology* **22**, 1298-1306.
- Swiderski, M.R., and Innes, R.W.** (2001). The Arabidopsis PBS1 resistance gene encodes a member of a novel protein kinase subfamily. *Plant Journal* **26**, 101-112.
- Tanaka, K., Asami, T., Yoshida, S., Nakamura, Y., Matsuo, T., and Okamoto, S.** (2005). Brassinosteroid homeostasis in Arabidopsis is ensured by feedback expression of multiple genes involved in its metabolism. *Plant Physiol.* **138**, 1117-1125.
- TerBush, D.R., Maurice, T., Roth, D., Novick, P.** (1996). The exocyst is a multiprotein complex required for exocytosis in *Saccharomyces cerevisiae*. *EMBO J.* **15**, 6483-6494.
- Thomas, C.M., Jones, D.A., Parniske, M., Harrisosn, K., Balint-Kurti, P.J., Hatzixanthis, K., and Jones, J.D.** (1997). Characterization of the tomato Cf-4 gene for resistance to *Cladosporium fulvum* identifies sequences that determine recognitional specificity in Cf-4 and Cf-9. *Plant Cell* **9**, 2209-2224.
- Torii, K.U.** (2004). Leucine-rich repeat receptor kinases in plants: structure, function, and signal transduction pathway. *International Review of Cytology* **234**, 1-46.
- Torii, K.U., and Clark, S.E.** (2000). Receptor-like kinases in plant development. *Advances in Botanical Research Incorp. Advances in Plant Pathology* **32**, 226-268.
- Torii, K.U., Mitsukawa, N., Oosumi, T., Matsuura, Y., Yokoyama, R., Whittier, R.F., and Komeda, Y.** (1996). The *Arabidopsis ERECTA* gene encodes a putative receptor protein kinase with extracellular leucine-rich repeats. *The Plant Cell* **8**, 735-746.
- Tör, M., Brown, D., Cooper, A., Woods- Tör, A., Sjölander, K., Jones, J.D.G., and Holub, E.B.** (2004). Arabidopsis downy mildew resistance gene *RPP27* encodes a receptor-like protein similar to *CLAVATA2* and tomato *Cf-9*. *Plant Physiol.* **135**, 1100-1112.
- Traas, J., Huelskamp, M., Gendreau, E., and Hoefte, H.** (1998). Endoreduplication and development: rule without dividing? *Current Opinion in Plant Biology* **1**, 498-503.
- Trotochaud, A.E., Hao, T., Wu, G., Yang, Z., and Clark, S.E.** (1999). The CLAVATA1 receptor-like kinase requires CLAVATA3 for its assembly into a signaling complex that includes KAPP and Rho-related protein. *Plant Cell* **11**, 393-406.
- Trotochaud, A.E., Jeong, S., and Clark, S.E.** (2000). CLAVATA3, a multimeric ligand for the CLAVATA1 receptor-kinase. *Science* **289**, 613-617.
- Tsukaya, H.** (1998). Genetic evidence for polarities that regulate leaf morphogenesis. *J. Plant Res.* **111**, 113-119.
- Tsukaya, H.** (2002). Leaf development. *The Arabidopsis Book*, 1-23
- Tsukaya, H.** (2003). Organ shape and size: a lesson from studies of leaf morphogenesis. *Current Opinion* **6**, 57-62.

- Tsukaya, H., Tsuge, T., and Uchimiya, H.** (1994). The cotyledon: a superior system for studies of leaf development. *Planta* **195**, 309-312.
- Tsuge, T., Tsukaya, H., and Uchimiya, H.** (1996). Two independent and polarized processes of cell elongation regulate leaf blade expansion in *Arabidopsis thaliana* (L.) Heynh. *Development* **122**, 1589-1600.
- Valon, C., Smalle, J., Goodman, H.M., and Giraudat, J.** (1993). Characterization of an *Arabidopsis thaliana* gene (*TMKL1*) encoding a putative transmembrane protein with an unusual kinase-like domain. *Plant Mol Biol.* **23**, 415-421.
- Vatter, P., Stoesser, C., Samel, I., Gierschik, P., and Moepps, B.** (2005). The variable C-terminal extension of G-protein-coupled receptor kinase 6 constitutes an accessorial autoregulatory domain. *FEBS J.* **272**, 6039-6051.
- Vaynberg, J., Fukuda, T., Chen, K., Vinogradova, O., Velyvis, A., Tu, Y., Ng, L., Wu, C., and Qin, J.** (2005). Structure of an ultraweak protein-protein complex and its crucial role in regulation of cell morphology and motility. *Mol Cell* **17**, 513-523.
- Venter, J.C., Smith, H.O., and Hood, L.** (1996). A new strategy for genome sequencing. *Nature* **381**, 364-366.
- Verma, D.P.** (2001). Cytokinesis and building of the cell plate in plants. *Annu Rev Plant Physiol Plant Mol Biol* **52**, 751-784.
- Verma, D.P., and Hong, Z.** (2001). Plant callose synthase complex. *Plant Mol. Biol.* **47**, 693-701.
- Verwoerd, T.C., Dekker, B.M., and Hoekema, A.** (1989). A small-scale procedure for the rapid isolation of plants RNAs. *Nucleic Acids Res.* **17**, 2362.
- Walker, J.C.** (1994). Structure and function of the receptor-like protein kinases of higher plants. *Plant Mol Biol.* **26**, 1599-1609.
- Wang, B.B., and Brendel, V.** (2006). Genomewide comparative analysis of alternative splicing in plants. *Proc Natl Acad Sci U S A* **103**, 7175-7180.
- Wang, G.L., Ruan, R.L., Song, W.Y., Sideris, S., Chen, L., Pi, L.Y., Zhang, S., Zhang, S., Fauquet, C., Gaut, B.S., Whalen, M.C., and Ronald, P.C.** (1998). Xa21 encodes a receptor-like molecule with a leucine-rich repeat domain that determines race-specific recognition and is subject to adaptive evolution. *Plant Cell* **10**, 765-779.
- Wang, X., Li, X., Meisenhelder, J., Hunter, T., Yoshida, S., Asami, T., and Chory, J.** (2005). Autoregulation and homodimerization are involved in the activation of the plant steroid receptor BRI1. *Developmental Cell* **8**, 855-865.
- Wang, Z.-Y., and He, J.-X.** (2004). Brassinosteroid signal transduction chooses of signals and receptors. *TRENDS in Plant Science* **9**, 91-96.
- Wang, Z.Y., Seto, H., Fujioka, S., Yoshida, S., and Chory, J.** (2001). BRI1 is a critical component of a plasma-membrane receptor for plant steroids. *Nature* **410**, 380-383.
- Williamson, M.P.** (1994). The structure and function of proline-rich regions in proteins. *Biochem J.* **297**, 249-260.
- Xia, X., and Xie, Z.** (2001). DAMBE: software package for data analysis in molecular biology and evolution. *J Hered.* **92**, 371-373.
- Xu, W.H., Wang, Y.S., Liu, G.Z., Tinjuangjun, P., Pi, L.Y., and Song, W.Y.** (2006). The autophosphorylated Ser686, Thr688, and Ser689 residues in the intracellular juxtamembrane domain of XA21 are implicated in stability control of rice receptor-like kinase. *Plant J.* **45**, 740-751.
- Yaglom, J., Linskens, M.H., Sadis S., Rubin, D.M., Futcher, B., and Finley, D.** (1995). P34Cdc28-mediated control of Cln3 cyclin degradation. *Mol Cell Biol.* **15**, 731-741.
- Yamada, K., Lim, J., Dale, J.M., Chen, H., Shinn, P., Palm, C.J., Southwick, A.M., Wu, H.C., Kim, C., Nguyen, M.** (2003). Empirical analysis of transcriptional activity in the *Arabidopsis* genome. *Science* **302**, 842-846.
- Yang, C., Xu, Z., Song, J., Conner, K., Barrena, G.V., and Wilson, Z.A.** (2007). *Arabidopsis* MYB26/MALE STERILE35 regulates secondary thickening in the

- endothecium and is essential for anther dehiscence. *The Plant Cell* **19**, 534-548.
- Yang, H., Matsubayashi, Y., Nakamura, K., and Sakagami, Y.** (1999). *Oryza sativa* PSK gene encodes a precursor of phytosulfokine- $\alpha$ , a sulfated peptide growth factor found in plants. *Proc. Natl. Acad. Sci. USA* **96**, 13560-13565.
- Yang, S.L., Jiang, L., Puah, C.S., Xie, L.F., Zhang, X.Q., Chen, L.Q., Yang, W.C., and Ye, D.** (2005). Overexpression of TAPEDUM DETERMINANT1 Alters the cell fates in the Arabidopsis carpel and tapetum via genetic interaction with excess microsporocytes1/extra sporogenous cells. *Plant Physiol.* **139**, 186-191.
- Yokoyama, R., Takahashi, T., Kato, A., Torii, K.U., and Komeda, Y.** (1998). The Arabidopsis ERECTA gene is expressed in the shoot apical meristem and organ primordial. *Plant j* **15**, 301-310.
- Yarden, Y., and Sliwkowski, M.X.** (2001). Untangling the ErbB signaling network. *Nat Rev Mol Cell Biol* **2**, 127-137.
- Yuan, C.X., Lasut, A.L., Wynn, R., Neff, N.T., Hollis, G.F., Ramaker, M.L., Rupa, M.J., Liu, P., and Meade, R.** (2003). Purification of Her-2 extracellular domain and identification of its cleavage site. *Protein Expr Purif.* **29**, 217-222.
- Zhang, J.Z.** (2003). Overexpression analysis of plant transcription factor. *Current Opinion in Plant Biology* **6**, 430-440.
- Zhao, D.Z., Wang, G.F., Speal, B., and Ma, H.** (2002). The *EXCESS MICROSPOROCYTES1* gene encodes a putative leucine-rich repeat receptor protein kinase that controls somatic and reproductive cell fates in the *Arabidopsis* anther. *Genes and Development* **16**, 2021-2031.
- Zhang, X.S., Choi, J.H., Josephine, H., and Chetty, C.S.** (2006). Domain-specific positive selection contributes to the evolution of Arabidopsis leucine-rich repeat receptor-like kinase (LRR-RLK) genes. *J Mol Evol* **63**, 612-621.
- Zipfel, C., Kunze, G., Chinchilla, D., Caniard, A., Jones, J.D., Boller, T., and Felix, G.** (2006). Perception of the bacterial PAMP EF-Tu by the receptor EFR restricts Agrobacterium-mediated transformation. *Cell* **125**, 749-760.

## APPENDIX A

**Table 1: Polymorphisms in *SRF1* (Col/Ler)**

<b>Nucleotide number</b>		<b>Polymorphisms Col→Ler</b>	<b>Amino acid change</b>	<b>Substituted nucleotide position in triplet</b>
20	(5'UTR)	<u>G→A</u>		
148	(5'UTR)	<u>C→A</u>		
156-157	(5'UTR)	<u>CT</u> Deletion <sup>a</sup>		
251	(exon)	<u>C→G</u>	L→L	3 <sup>rd</sup> nt
476	(exon)	T→G	G→G	3 <sup>rd</sup> nt
482	(exon)	<u>C→T</u>	S→S	3 <sup>rd</sup> nt
598	(intron)	<u>T→A</u>		
696	(intron)	T→G		
804	(exon)	G→A	L→L	3 <sup>rd</sup> nt
837	(intron)	T→C		
883	(intron)	A→G <sub>-</sub> *		
895	(intron)	A→G		
897	(intron)	A→C		
992	(intron)	C→G		
1008	(intron)	T→C		
1040	(intron)	TCT→ GACTTC Insertion <sup>a</sup>		
1076	(exon)	T→A	L→L	3 <sup>rd</sup> nt
1187	(intron)	C→T		
1196	(intron)	G→A		
1218	(exon)	T→A	S→S	3 <sup>rd</sup> nt
1268	(exon)	C→A	T→N	2 <sup>nd</sup> nt
1709	(exon)	T→C	N→N	3 <sup>rd</sup> nt
1781	(exon)	A→G	P→P	3 <sup>rd</sup> nt
1812	(exon)	G→A	A→T	1 <sup>st</sup> nt
2050	(exon)	G→A	R→K	2 <sup>nd</sup> nt
2158	(intron)	G→A		
2178	(intron)	T→C		
2214	(exon)	G→A	R→K	2 <sup>nd</sup> nt
2241	(exon)	C→T	S→L	2 <sup>nd</sup> nt
2251	(exon)	T→G	H→Q	3 <sup>rd</sup> nt
2279	(exon)	G→A	E→K <sup>b</sup>	1 <sup>st</sup> nt
2281	(exon)	A→G	E→K <sup>b</sup>	3 <sup>rd</sup> nt
2316	(exon)	G→A	G→D	2 <sup>nd</sup> nt
2382	(exon)	C→A	T→N	2 <sup>nd</sup> nt
2409	(exon)	G→A	R→K	2 <sup>nd</sup> nt
2414	(exon)	A→T	T→S	1 <sup>st</sup> nt
2487	(exon)	G→A	S→N	2 <sup>nd</sup> nt
2496	(exon)	A→T	H→L	2 <sup>nd</sup> nt
2593	(intron)	G→T		
2623	(intron)	G→T		
2646	(exon)	C→G	F→L	3 <sup>rd</sup> nt
2654	(exon)	G→A	R→K	2 <sup>nd</sup> nt
2694	(exon)	C→G	G→G	3 <sup>rd</sup> nt
2701	(exon)	T→G	L→V	1 <sup>st</sup> nt

2718	(exon)	T→C	N→N	3 <sup>rd</sup> nt
2730	(exon)	A→C	I→I	3 <sup>rd</sup> nt
2742	(exon)	T→C	N→N	3 <sup>rd</sup> nt
2757	(exon)	G→T	V→V	3 <sup>rd</sup> nt
2763	(exon)	C→T	F→F	3 <sup>rd</sup> nt
2802	(exon)	T→C	T→T	3 <sup>rd</sup> nt
2805	(exon)	T→C	C→C	3 <sup>rd</sup> nt
2837	(exon)	T→C	I→T	2 <sup>nd</sup> nt
2847	(exon)	G→A	R→R	3 <sup>rd</sup> nt
2872	(exon)	G→A	V→I	1 <sup>st</sup> nt
2886	(exon)	T→G	I→M	3 <sup>rd</sup> nt
2936	(exon)	C→T	P→L	2 <sup>nd</sup> nt
2951	(exon)	G→A	R→Q	2 <sup>nd</sup> nt
2991	(exon)	A→G	I→M	3 <sup>rd</sup> nt
3076-3080	(intron)	AATCT Insertion <sup>a</sup>		
3134	(intron)	T→G		
3148	(exon)	G→C	S→S	3 <sup>rd</sup> nt
3262	(exon)	T→C	G→G	3 <sup>rd</sup> nt
3288	(intron)	A→T		
3338	(intron)	A→T		
3381	(exon)	A→G	Q→Q	3 <sup>rd</sup> nt
3431	(exon)	C→A	A→E	2 <sup>nd</sup> nt
3450	(exon)	C→T	S→S	3 <sup>rd</sup> nt
3530	(intron)	G→T		
3532	(intron)	AAAGAA Deletion <sup>a</sup>		
3550	(intron)	TG Insertion		
3556-3561	(intron)	TATTAT insertion <sup>a</sup>		
3565	(intron)	A deletion <sup>a</sup>		
3573-3574	(intron)	GT deletion <sup>a</sup>		
3587-3588	(intron)	AT insertion <sup>a</sup>		
3613	(exon)	A→T	Y→F	2 <sup>nd</sup> nt
3708	(exon)	G→A	G→R	1 <sup>st</sup> nt
3769	(3'UTR)	T→A		
3818	(3'UTR)	T deletion <sup>a</sup>		
3819	(3'UTR)	T deletion <sup>a</sup>		

Star indicates the substitution data obtained from PERL. Underlined and italic letters represent data from Monsanto. Otherwise all data are from our sequence analysis.

<sup>a</sup> represents INDELS in *Ler* background.

<sup>b</sup> indicates that two polymorphisms affect the same triplet.

## APPENDIX B

**Table 2: Polymorphisms in *SRF2* (Col/Ler)**

Nucleotide number	Polymorphisms Col→Ler	Amino acid change	Substituted nucleotide position in triplet
53 (5' UTR)	<u>T→C</u> *		
246 (intron)	T→C		
301 (intron)	T→G		
307 (intron)	T→A		
324 (intron)	A insertion <sup>a</sup>		
326 (intron)	A deletion <sup>a</sup>		
391 (intron)	G→A*		
896 (exon)	G→T	E→STOP	1 <sup>st</sup> nt
2059 (exon)	C→A	H→N	1 <sup>st</sup> nt
2801 (exon)	G→T	G→C	1 <sup>st</sup> nt
3316 (intron)	G→T		
3405 (exon)	G→T	G→STOP	1 <sup>st</sup> nt
3462 (exon)	G→T	E→STOP	1 <sup>st</sup> nt
3522 (exon)	G→T	A→S	1 <sup>st</sup> nt
3555 (intron)	T→A		
3570 (intron)	C→G		
3576 (intron)	G→A		
3588 (intron)	T→A		
3612 (intron)	G→A		
3616 (exon)	G→C	E→Q	1 <sup>st</sup> nt
3629 (exon)	G→A	R→K	2 <sup>nd</sup> nt
3780 (exon)	C→G	S→R	3 <sup>rd</sup> nt
3840 (3' UTR)	G insertion <sup>a</sup>		
3868 (3' UTR)	A insertion <sup>a</sup>		
3885 (3' UTR)	G deletion <sup>a</sup>		

Star indicates the substitution data obtained from PERLEGEN. Star and underlined letters represent data from both Monsanto and PERLEGEN. Otherwise all data are from Monsanto Database.

<sup>a</sup> represents INDELS in *Ler* background.

## APPENDIX C

**Table 3: Polymorphisms in *SRF3* (Col/Ler)**

Nucleotide number	Polymorphisms Col→Ler	Amino acid change	Substituted nucleotide position in triplet
355	(exon) C→A	I→I	3 <sup>rd</sup> nt
576	(intron) A→C		
1146	(exon) C→G	P→R	2 <sup>nd</sup> nt
1148	(exon) G→A	E→K	1 <sup>st</sup> nt
1231	(intron) G→A		
1241	(intron) G→A		
1247	(intron) <u>T→C</u> *		
1291	(exon) G→A	G→E	2 <sup>nd</sup> nt
1496	(intron) C→T		
1650	(intron) C→A		
1900	(exon) <u>T→C</u> *	M→T	2 <sup>nd</sup> nt
1908-1910	(exon) TCC→AAA	S→K	triplet
1913	(exon) A insertion <sup>a</sup>	T→T <sup>a</sup>	3 <sup>rd</sup>
2382	(exon) T→A	V→D	2 <sup>nd</sup> nt
2402	(exon) AA→TT	K→L	1 <sup>st</sup> /2 <sup>nd</sup> nt
2408	(exon) GG→TT	G→L	1 <sup>st</sup> /2 <sup>nd</sup> nt
2424	(exon) T→C	V→A	2 <sup>nd</sup> nt
2435-2437	(exon) CTT→GCG	L→A	triplet
2468	(exon) A→C	R→R	1 <sup>st</sup> nt
2483	(exon) G→C	D→H	1 <sup>st</sup> nt
2486	(exon) A→C	I→L	1 <sup>st</sup> nt
2793	(intron) C→T		
3345	(intron) G→C		
3585	(intron) T→C		
3595	(intron) G→C		
3841	(intron) T→G		
3868-3869	(intron) TT→GG		
3906	(exon) A→G	E→G	1 <sup>st</sup> nt
3952	(exon) T→C	G→G	3 <sup>rd</sup> nt
3956	(exon) G→A	G→S	1 <sup>st</sup> nt
3966	(exon) C→G	N→S	2 <sup>nd</sup> nt
3985-3986	(3'UTR) AA→TC		
4149	(3'UTR) T→G		
4166	(3'UTR) T insertion <sup>a</sup>		
4185	(3'UTR) T deletion <sup>a</sup>		
4190	(3'UTR) T deletion <sup>a</sup>		
4216	(3'UTR) T→G		
4219-4220	(3'UTR) TG→GT		
4274	(3'UTR) T deletion <sup>a</sup>		
4425	(3'UTR) A→T		

Star indicates the substitution data obtained from PERL. Star and underlined letters represent data from both Monsanto and PERL. Otherwise all data are from Monsanto Database.

<sup>a</sup> represents INDELS in *Ler* background.

## APPENDIX D

**Table 4: Polymorphisms in *SRF4* (Col/Ler)**

Nucleotide number	Polymorphisms Col→Ler	Amino acid change	Substituted nucleotide position in triplet
100 (5'UTR)	T deletion <sup>a</sup>		
118 (5'UTR)	<u>T</u> →C*		
144 (exon)	<u>G</u> →C*	G→A	2 <sup>nd</sup> nt
168 (exon)	T→A	L→H	2 <sup>nd</sup> nt
199 (exon)	C→G	T→T	3 <sup>rd</sup> nt
670 (intron)	A→T		
672 (intron)	C→T		
675 (intron)	T insertion <sup>a</sup>		
947 (exon)	G→A	V→V	3 <sup>rd</sup> nt
1975 (exon)	G→A	V→N	1 <sup>st</sup> nt
1996 (exon)	A→C	D→N	1 <sup>st</sup> nt
2010 (exon)	T→G	V→V	3 <sup>rd</sup> nt
2028 (exon)	T→G	L→L	3 <sup>rd</sup> nt
2069 (exon)	T→A	F→Y	2 <sup>nd</sup> nt
3112 (intron)	C→T		
3584 (3'UTR)	C→A		
3617 (3'UTR)	T→A		
3730 (3'UTR)	A→G		

Star indicates the substitution data obtained from PERL. Star and underlined letters represent data from both Monsanto and PERL. Otherwise all data are from Monsanto Database.

<sup>a</sup> represents INDELS in *Ler* background.

## APPENDIX E

**Table 5: Polymorphisms in *SRF5* (Col/Ler)**

Nucleotide number		Polymorphisms Col→Ler	Amino acid change	Substituted nucleotide position in triplet
9	(5'UTR)	T insertion <sup>a</sup>		
57	(5'UTR)	G→A		
64	(5'UTR)	A→C		
75	(5'UTR)	T deletion <sup>a</sup>		
113	(5'UTR)	A→C		
307	(intron)	A→G		
406-408	(exon)	<u>GGT</u> → <u>TTC</u> *	G→F	triplet
419	(exon)	A→T	E→V	2 <sup>nd</sup> nt
427	(exon)	T→A	W→R	1 <sup>st</sup> nt
439	(exon)	A→T	K→STOP	1 <sup>st</sup> nt
447	(exon)	A→T	K→N	3 <sup>rd</sup> nt
481	(intron)	A→C		
587	(exon)	<u>G</u> → <u>A</u> *	G→S	1 <sup>st</sup> nt
597	(exon)	<u>G</u> → <u>T</u> *	R→L	2 <sup>nd</sup> nt
1196	(intron)	T→C		
1780	(exon)	A→G /A→G*	T→A / T→P*	1 <sup>st</sup> nt
2070	(exon)	G→A	G→D	2 <sup>nd</sup> nt
2321	(intron)	T→C		
2325	(intron)	TA→CC		
2341	(intron)	T→A		
2352	(intron)	T→A		
2751	(exon)	T→C	Y→H	1 <sup>st</sup> nt
2754	(exon)	C→T	L→L	1 <sup>st</sup> nt
2774	(exon)	T→G	S→A	1 <sup>st</sup> nt
2778	(exon)	G→A	V→M	1 <sup>st</sup> nt
2781	(exon)	GACAATG insertion <sup>a</sup>	D, N insertion <sup>a</sup>	
2792	(exon)	T→C	N→N	3 <sup>rd</sup> nt
2811	(exon)	T→G	L→V	1 <sup>st</sup> nt
2813	(exon)	G deletion <sup>a</sup>	L→V <sup>a</sup>	3 <sup>rd</sup> nt
2921	(intron)	G→C		
2967	(exon)	G→C	G→A	2 <sup>nd</sup> nt
3012	(exon)	C→G	T→R	2 <sup>nd</sup> nt
3092	(intron)	C→T		
3066	(intron)	C→G		
3612	(3'UTR)	T→G		
3615	(3'UTR)	T→G		

Star indicates the substitution data obtained from PERL. Star and underlined letters represent data from both Monsanto and PERL. Otherwise all data are from Monsanto Database.

<sup>a</sup> represents INDELS in *Ler* background.

## APPENDIX F

**Table 6: Polymorphisms in *SRF6* (Col/Ler)**

Nucleotide number	Polymorphisms Col→Ler	Amino acid change	Substituted nucleotide position in triplet
39	(5'UTR) C→T		
151	(exon) T→A	T→T	3 <sup>rd</sup> nt
201	(intron) T insertion <sup>a</sup>		
202	(intron) A→T		
249	(intron) T→A		
713	(intron) C→T		
1420-1422	(exon) GGC→AAA	P→P/A→K	3 <sup>rd</sup> /1 <sup>st</sup> , 2 <sup>nd</sup> nt
1471-1472	(intron) CC→GG		
1547	(exon) A→C	Q→P	2 <sup>nd</sup> nt
1589	(exon) A→T	E→V	2 <sup>nd</sup> nt
1605	(intron) T→C		
2047	(exon) G→A	R→K	2 <sup>nd</sup> nt
2099	(exon) C→T	N→N	3 <sup>rd</sup> nt
2472	(exon) T→A	T→T	3 <sup>rd</sup> nt
3205	(exon) G→C	E→Q	1 <sup>st</sup> nt
3226	(exon) T→G	Y→D	1 <sup>st</sup> nt
3231-3232	(exon) T→A	S→S, F→T	3 <sup>rd</sup> /1 <sup>st</sup> nt
3233	(exon) T→C	F→S	2 <sup>nd</sup> nt
3234	(exon) G→T	F→T	3 <sup>rd</sup> nt
3246	(exon) T→G	I→M	3 <sup>rd</sup> nt
3247	(exon) G insertion <sup>a</sup>	Y→V <sup>a</sup>	1 <sup>st</sup> nt
3272	(exon) A→T	E→V	2 <sup>nd</sup> nt
3484	(exon) C→A	A→D	2 <sup>nd</sup> nt
3556	(intron) A deletion <sup>a</sup>		

Star indicates the substitution data obtained from PERL. Star and underlined letters represent data from both Monsanto and PERL. Otherwise all data are from Monsanto Database.

<sup>a</sup> represents INDELS in *Ler* background.

## APPENDIX G

**Table 7: Polymorphisms in *SRF7* (Col/Ler)**

<b>Nucleotide number</b>		<b>Polymorphisms Col→Ler</b>	<b>Amino acid change</b>	<b>Substituted nucleotide position in triplet</b>
341	(intron)	C→T		
708	(intron)	A→G		
763	(intron)	C→A		
768	(intron)	T→A		
873	(exon)	C→A*	L→L	3 <sup>rd</sup> nt
1010	(intron)	A→T		
1042-1043	(exon)	AA→GG	K→G	1 <sup>st</sup> /2 <sup>nd</sup> nt
1049-1050	(exon)	CA→GG	P→R	2 <sup>nd</sup> /3 <sup>rd</sup> nt
1423	(exon)	A→T	L→M	2 <sup>nd</sup> nt
1442	(intron)	G→A		
1450	(intron)	T→A		
1464	(intron)	C insertion <sup>a</sup>		
1872	(exon)	A→T	I→F	1 <sup>st</sup> nt
1893	(exon)	G deletion <sup>a</sup>	E→K	1 <sup>st</sup> nt
2179	(exon)	C→T	P→S	1 <sup>st</sup> nt
2308	(exon)	G→A	G→R	1 <sup>st</sup> nt
2338	(exon)	T deletion <sup>a</sup>	V→E	2 <sup>nd</sup> nt
2342	(exon)	T→G	V→G	2 <sup>nd</sup> nt
3662	(exon)	C→T	D→D	3 <sup>rd</sup> nt
2940	(exon)	T→G	V→G	2 <sup>nd</sup> nt
2974	(exon)	T→A	D→E	3 <sup>rd</sup> nt
3004	(exon)	G→C	S→S	3 <sup>rd</sup> nt
3149	(exon)	T→G	C→V	1 <sup>st</sup> nt
3321	(exon)	C→T	V→V	3 <sup>rd</sup> nt
3406	(exon)	C→T	P→S	1 <sup>st</sup> nt
3415	(exon)	A→T	N→Y	1 <sup>st</sup> nt
3505	(intron)	T→G		
3516	(intron)	T→C		
3574	(exon)	C→T	G→G	3 <sup>rd</sup> nt
3586	(exon)	G→A	L→L	3 <sup>rd</sup> nt
3604-3606	(exon)	TTT→GGG	S→R/F→G	3 <sup>rd</sup> , 1 <sup>st</sup> /2 <sup>nd</sup> nt
3986	(intron)	T→G		
3994	(exon)	G→C	P→P	3 <sup>rd</sup> nt
4043-4044	(exon)	GT→TG	V→STOP	1 <sup>st</sup> 2 <sup>nd</sup> nt
4050	(exon)	T→G	V→G	2 <sup>nd</sup> nt

Star indicates the substitution data obtained from PERL. Star and underlined letters represent data from both Monsanto and PERL. Otherwise all data are from Monsanto Database

<sup>a</sup> represents INDELS in *Ler* background.

## APPENDIX H

**Table 8: Polymorphisms in *SRF8* (Col/Ler)**

<b>Nucleotide number</b>	<b>Polymorphisms Col→Ler</b>	<b>Amino acid change</b>	<b>Substituted nucleotide position in triplet</b>
689-691 (intron)	GAT→AGA		
1507 (intron)	G→A		
1593 (exon)	AA→GG	Q→R	2 <sup>nd</sup> /3 <sup>rd</sup> nt
1600 (exon)	A→G	G→S	1 <sup>st</sup> nt
1964 (exon)	T insertion <sup>a</sup>	S→F	2 <sup>nd</sup> nt
1997 (exon)	T→G	L→W	2 <sup>nd</sup> nt
2012 (exon)	T→A	V→D	2 <sup>nd</sup> nt
3655 (exon)	C deletion <sup>a</sup>	T→M	2 <sup>nd</sup> nt
3757 (3'UTR)	C→T		

Star indicates the substitution data obtained from PERL. Star and underlined letters represent data from both Monsanto and PERL. Otherwise all data are from Monsanto Database.

<sup>a</sup> represents INDELS in *Ler* background.

## Lebenslauf

### Persönliche Daten

Banu Eyüboğlu

Geboren am 10.03.1975 in Ankara

Staatsangehörigkeit: Türkisch

Familienstand: ledig

### Schule und Studium

1989 – 1992	High School, Ankara
1993 – 1997	Bachelor of Science in Biology, Hacettepe University, Ankara
1997 – 1998	English Preparatory School, Middle East Technical University, Ankara
1998 – 2001	Master of Science in Biotechnology, Middle East Technical University, Ankara

### Berufliche Tätigkeit

2001-2003	Forschungsstipendium, Young Scientist Training Group, TUBITAK, Biotechnology, Middle East Technical University, Ankara
2003-2008	Wissenschaftliche Angestellte, Fachgebiet Entwicklungsbiologie der Pflanzen, TU München/Freising-Weihenstephan