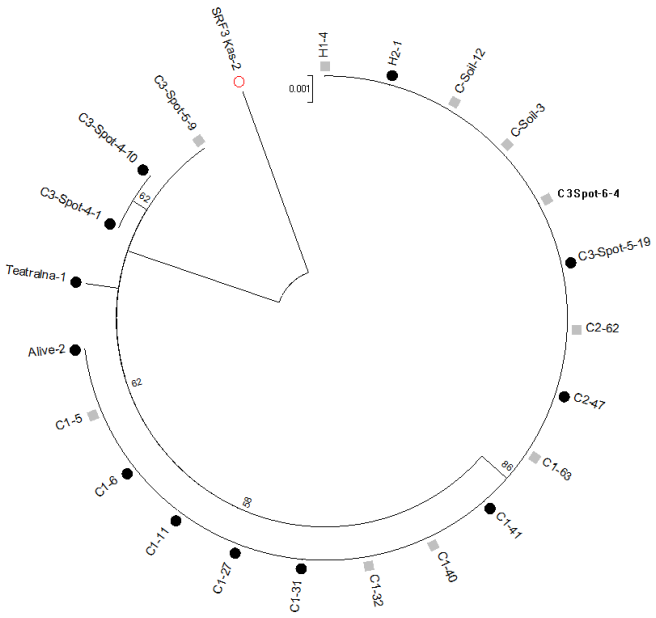
Sveriges lantbruksuniversitet
Swedish University of Agricultural Sciences

Department of Plant Biology

Characterisation of the epistatic hybrid incompatibility locus *STRUBBELIG RECEPTOR FAMILY 3* in *Arabidopsis*

Fredrik Dölfors



Department of Plant Biology, SLU
Independent project in Biology (EX0895) • 30 hec • Second cycle, A2E
Master's programme in Plant Biology
No. 169
Uppsala, 2018

Characterisation of the epistatic hybrid incompatibility locus *STRUBBELIG RECEPTOR FAMILY 3* in *Arabidopsis*

Fredrik Dölfors

Main Supervisor: German Martinez Arias, Swedish University of Agricultural Sciences, Department of Plant Biology.

Collaborator: Rubén Alcázar, Department of Biology, Healthcare and Environment. University of Barcelona. Faculty of Pharmacy. Avda Joan XXIII 27-31, 08028 Barcelona, Spain

Examiner: Christina Dixelius, Swedish University of Agricultural Sciences, Department of Plant Biology.

Credits: 30 hec
Level: Second cycle, A2E
Course title: Independent project in Biology – Master's thesis
Course code: EX0895
Programme/education: Master's programme in Plant Biology

Place of publication: Uppsala
Year of publication: 2018
Cover picture: Phylogenetic *SRF3* gene tree, image by author
Title of series: Examensarbete / Institutionen för växtbiologi, SLU
Course coordinating department: Department of Plant Biology, SLU
No: 169

Online publication: http://stud.epsilon.slu.se

Keywords: Abiotic stress, Arabidopsis, co-evolution, hybrid incompatibility, receptor-like kinase

Swedish University of Agricultural Sciences
Faculty of Natural Resources and Agricultural Sciences
Department of Plant Biology

Popular science summary

Plants have a complex innate immune system that detect and protect against pathogens and environmental sensors that, for example, help detect changes in temperature, water availability and circadian rhythm. It may happen that plants of the same population get temporarily split into two subpopulations by a physical reproductive barrier. Many generations later when the two closely related subpopulations eventually cross again, a few key changes in their DNA may have created a different problem. Two or more proteins that interact in the immune system may have changed between the subpopulations. The configuration of these proteins may no longer work together properly when combined in a hybrid offspring. This could cause autoimmune conflictions where the immune system is constantly activated, even though no disease-causing organisms are present. An active immune system costs resources and creates a competition of energy between growth and defence, leaving the hybrid plant stunted. The hybrids often fail to propagate and die prematurely. This creates a purifying selection pressure against incompatible gene variants between the loci.

Researchers have found several cases of hybrid autoimmunity occurring in wild populations of *Arabidopsis thaliana*. Rubén Alcázar and his team of researchers have discovered a few genes involved in such a case in a population in Poland. Part of my project was to study signs of negative selection against incompatible alleles between two loci. One that encodes for a cluster of immune receptors, collectively called RECOGNITION OF PERONOSPORA PARASITICA1-like (RPP1-like) and a second that produce a receptor-like protein called STRUBBELIG RECEPTOR FAMILY 3 (SRF3). At least one RPP1-like protein functions to guard SRF3 from pathogen disturbances and SRF3 acts as a receptor for the innate immune system in Central Asian ecotypes. SRF3 function in European versions of *A. thaliana* plants is unknown. I found statistical signs of negative selection in the population, but larger experiments with more individuals are needed to rule out genetic drift or neutral evolution.

While working with my master thesis I investigated the unknown function of European SRF3. My data showed European *A. thaliana* plants (Landsberg *erecta*) to have higher levels of expression of *SRF3* when the plants detected and signalled drought stress. During the early stages of drought and salt stress the cellular osmotic potential changes. It is very important that the plants detect shortages in soil water content and react to store moisture. To further understand the role of SRF3 I subjected *A. thaliana* plants that had a defective *SRF3* gene and plants that overproduces the protein to an osmotic stress. By decreasing the amount of biologically available water by dissolving an excess of sugar in the growth medium, the cells in the plant need to start producing higher amounts of solutes in order to keep osmosis stable and not lose water.

I wanted to see whether plants that overproduce SRF3 proteins had any growth advantage (tolerating the stress) and gain insights into its molecular functions in the plant but the results were inconclusive in a first trial.

Researchers can also get new ideas on how to test molecular functions of a protein by studying which other genes co-regulate together with the gene of interest. Large amounts of publicly available gene expression data can be found online and analysed by anyone. I have performed these analyses for my gene of interest (*SRF3*). I found that several genes involved in abscisic acid related signalling (a major stress induced signalling hormone) co-regulate together with *SRF3* during numerous types of stress factors. A few correlated genes functions during drought stress and production of protective sugars. These findings highlight a possible multi-functional role of a newly discovered receptor-like protein in *A. thaliana*.

Abstract

Populations of *Arabidopsis thaliana* have spread throughout temperate zones of the world and adapted to prevailing biotic and abiotic stress factors. Diverging populations of *Arabidopsis* accessions can evolve postzygotic hybrid incompatibility (HI). In some cases, HI is caused by incompatible epistatic interactions between genes that have functionally diverged between- and within species. Crosses between the European Landsberg *erecta* (L*er*) and many central Asian accessions, like Kashmir-2 (Kas-2) and Kondara (Kond) suffers from it. Genetic analysis of this incompatibility has revealed its basis in a polymorphic cluster of Toll/interleukin-1 receptor-nucleotide binding-leucine rich repeat (TNL) *RPP1* (*Recognition of Peronospora parasitica 1*)-like genes in L*er* and alleles of the receptor-like kinase *Strubbelig Receptor Family 3* (*SRF3*) in ecotypes Kas-2 and Kond, causing autoactivation of the plant innate immunity response. In this project I have analysed: I) the genetic variability of *SRF3* in a segregating population of L*er* relatives originally from Gorsów (Poland) and II) the molecular function of *SRF3*. My analysis showed that *SRF3* might be more conserved in the Gorsow population compared to the Kas-2 and Kond accessions although it does not provide insight in the evolution of incompatibility. Analysis of the role of *SRF3* using overexpressing and mutant lines and *in silico* prediction indicates that it might be involved in both drought and salt stress. In summary, this project sheds light into the role of *SRF3* as a stress-related gene and its implication in hybrid incompatibility.

Keywords: Abiotic stress, Co-evolution, Hybrid incompatibility, Receptor-like kinase

Table of contents

Table of contents **4**

1 Introduction **6**

2 Aims of this project **9**

3 Results **10**

3.1 The Gw population show signs of negative selection at *SRF3* but alleles do not co-segregate with the *RPP1*^{Ler}-like cluster 10

3.2 Prediction reveal possible *SRF3* transcriptional upregulation 24hrs post drought and salt stress 13

3.3 SRF3 does not repress the plant osmotic stress phenotype 14

3.4 Genes related to abiotic stress signalling pathways co-express with *SRF3* 19

4 Discussion **22**

5 Materials & Methods **26**

5.1 Plant Materials & Growth Conditions 26

5.1.1 Plant materials 26

5.1.2 Plant growth conditions (sequencing and abiotic stress experiment) 27

5.1.3 Plant growth conditions for the *SRF3* mutant homozygosity test 28

5.2 DNA Extraction Methods & Sample Preparations 29

5.2.1 Evolutionary analysis of *A. thaliana* Gw accessions 29

5.2.2 T-DNA mutant two-step genotyping 30

5.3 Leaf Area & Plant Weight Measurements 30

5.4 *In Silico* analysis 31

5.5 Statistical Analyses & Data Processing 31

5.5.1 Genetic variability analysis of the Gw *SRF3* 31

5.5.2 Abiotic stress experiment 32

6	**References**	**33**
7	**Acknowledgements**	**35**
8	**Supplementary information**	**36**
9	**Opposition Summary**	**43**

1 Introduction

As sessile organisms, plants need to sense and react to daily and seasonal environmental cues for healthy development, growth and fitness. Plants have evolved complex stress-response networks to tune development to the dominant stress, leading to emerging phenotypic plasticity through genotype by environment interactions (Nicotra *et al.*, 2010; Alcázar & Parker, 2011). Populations of *Arabidopsis thaliana* have spread throughout temperate zones of the world and have colonized a wide range of regions where they survive and adapt to biotic and abiotic stress factors.

As a consequence of intra-specific genetic diversity, genetic drift and natural selection, spatially separated and segregating populations may no longer be compatible. This type of incompatibility has been hypothesized to occur due to epistatic interactions between at least two segregating loci that have fixated for different alleles among subpopulations. These alleles function well together in the parental lineages but can lead to detrimental development effects in the F1 and F2 generation following hybridisation (*Figure* 1). Such post-zygotic reproductive barriers are termed Bateson-Dobzhansky-Muller (BDM) incompatibilities (Orr, 1996; Weinreich *et al.*, 2005) and have been identified in crosses between several accessions of *A. thaliana*. Some detrimental effects of hybrid incompatibility (HI) include embryo lethality (Col-0/Cvi-0 reported by Bikard *et al.*, 2009) or growth defects and chlorosis (Col-0/Bur-0 reported by Vlad *et al.*, 2010). The genetic basis for some of these incompatible accessions has been determined and involves auto-activation of immunity. Incompatible accessions include the cross between ecotype accessions UK1 and UK3 (Bomblies *et al.*, 2007; Alonso-Blanco *et al.*, 2009) and crosses between the European ecotype Landsberg *erecta* (*Ler*) and Kashmir-2 (Kas-2) and Kondara (Kond) in central Asia (Alcázar *et al.*, 2009, 2010).

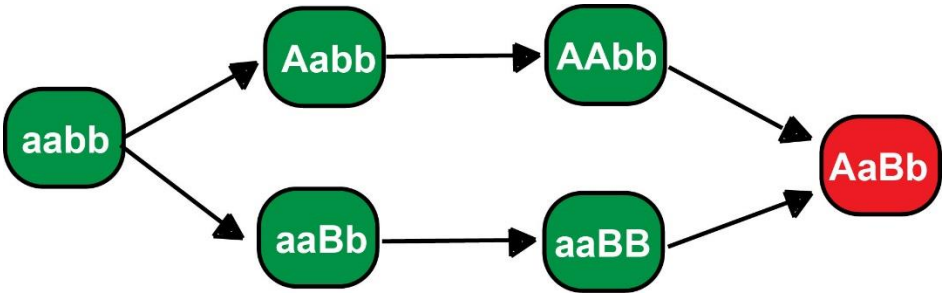


Figure 1. Illustration of the Bateson-Dobzhansky-Muller post-zygotic reproductive isolation theory. Different alleles at two or more distinct loci (a and b) fixate as a respons to local adaptation or genetic drift as a population is split into two subpopulations (green elipses). Hybridisation between the subpopulations yield epistatic interactions that may be detrimental (red elips). aa and bb = ancestral diploid genotype. A and B = new mutation, AA and BB = fixation of mutation. Adapted from Wu & Ting *et al.* (2004).

The HI reported by Alcázar (2009, 2010) is conditional and temperature dependent within the normal range experienced by *Arabidopsis* populations in the wild (14 ºC to 22 ºC). Homeostatic control is lost in incompatible hybrid plants through allelic mismatches involving immune receptors. At moderate temperature (20 ºC) HI becomes suppressed and plants develop and reproduce normally. At moderately low temperature (14ºC) hybrid seedlings exhibit transcriptionally activated salicylic acid (SA) pathway defences leading to dwarfism, necrosis and reduced reproductive fitness (Alcázar *et al*., 2009; Alcázar and Parker, 2011). Alcázar *et al*. (2009) previously fine mapped one of the interacting loci to a region in L*er* that contain a cluster of *RECOGNITION OF PERONOSPORA PARASITICA1*-like (*RPP1*-like) Toll-Interleukin1 Receptor-nucleotide binding-Leucine-rich repeat (TNL) immune receptor genes. The incompatible *RPP1*- like haplotype was found to be frequent in a population of L*er* relatives in Górzow Wielkopolski (Gw), Poland. 30% of the Gw individuals carry alleles that are incompatible to Kas-2 and Kond allelic forms of the second HI locus (Alcázar *et al.*, 2014). The second less known HI locus mapped to a membrane spanning Leucine Rich Repeat-Receptor Like Kinase (LRR-RLK) termed *STRUBBELIG RECEPTOR FAMILY3* (At4g03390) in Kas-2 and Kond (Alcázar *et al.*, 2010) (see a schematic representation in *Figure 2*).

.

Figure 2. A schematic representation of the domain structure of the hybrid incompatibility locus protein SRF3. X = unknown domain, LRR = leucine rich repeat domain, PRR = proline rich region, TM = transmembrane domain, Kinase = kinase domain. Adapted from Eyüboglu *et al.* (2007).

It has been hypothesized that trans-membrane localized RLKs can be involved in sensing the environment, including perturbations of cellular homeostasis. Several cases have been investigated previously regarding abiotic stress and abscisic acid (ABA) signalling during drought, salt, oxidative stress, cold stress and osmotic stress (see review by Osakabe *et al.*, 2013). Two examples of LRR-RLKs in *A. thaliana* that have been implicated as osmotic stress sensors include the positive regulators of tolerance attributes ATHK1 and AHK1 (Tran *et al.*, 2007; Wohlbach *et al.*, 2008). SRF3 belong to a family of 8 RLK members of which the precise molecular function is unknown (Eyüboglu *et al.*, 2007), though it has been hypothesized that the SRF is multi-functional. *SRF4* is known to be a regulator of leaf size (Eyüboglu *et al.*, 2007) and Central Asian allelic versions of *SRF3* confer disease resistance against *Hyaloperonospora arabidopsidis* (Alcázar *et al.*, 2010). Previous unpublished results by the Alcázar group have indicated that *SRF3* may also function during osmotic stress.

A basic characterisation of the RPP1^{Ler}-like incompatible *SRF3* locus is needed since very little information exist. This project aims to shed light into the natural variation and role of *SRF3*. A known SNP dense part of the *SRF3* was sequenced in the Gw population to identify natural genetic variation that may underly the *RPP1*^{Ler}-like R-gene cluster incompatibility at low temperatures. A comparison between *SRF3* alleles from accessions that contain (Gw^{+}) and others that lack the incompatible *RPP1*^{Ler}-like alleles (Gw^{-}) may show signs of co-evolution. The allelic forms of *SRF3* alleles may segregate in an incompatible *RPP1*^{Ler}-like presence/absence manner. Lower genetic variability was found than expected which could be a sign of negative selection against the presence of incompatible alleles.

Together with this, to further elucidate *SRF3* function I have investigated the role of *SRF3* as a potential stress-mediating gene. I have characterized the osmotic resistance of three Col-0 homologous knockdown mutant lines, a Col-0 homologous overexpressing line and a Col-0 line with a Gw version of *SRF3*. Within treatment comparisons between the lines provided no significant differences among the genetic lines. An *in silico* meta-data analysis revealed possible *SRF3* transcriptional upregulation after 24hrs of drought and salt stress. These results show that further studies are needed for abiotic stress involvement of *SRF3*. Publicly available transcriptome data was analysed in an additional *in silico* investigation to identify co-expressed genes during abiotic stress. This analysis identified several genes involved in ABA signalling to mutually co-express with *SRF3*, three directly involved in drought and salt stress and one annotated to osmotic stress tolerance.

2 Aims of this project

The goal of this project was to identify genetic variability at the *SRF3* locus in a European population of Landsberg *erecta* relatives and study a potential link to the evolution of an epistatic hybrid incompatibility with RPP1$^{L\textit{er}}$-like alleles. Since the basic function of the SRF3 proteins is unknown this project also aimed to identify co-expressed genes that may be linked to its function and elucidate a possible involvement in abiotic stress.

Main goals of the project

I. Identify genetic variability occurring at *SRF3* and study potential co-evolution with *RPP1*$^{L\textit{er}}$-like genes.

II. Elucidate possible SRF3 involvement in abiotic stress.

III. Identify SRF3 co-expressed genes.

3 Results

3.1 The Gw population show signs of negative selection at *SRF3* but alleles do not co-segregate with the *RPP1*^{Ler}-like cluster

To investigate the natural variation at the *SRF3* locus and possible co-evolution with the *RPP1* ^{Ler} -like cluster in the Gw population, a previously known SNP dense region between exons 8-11 in incompatible Kas-2 and Kond was chosen as region of interest. Sequencing this region of the *SRF3* gene among 22 accessions (10 Gw^{+} and 12 Gw^{-}) revealed less SNPs than previously reported inside exon 11 (Alcázar *et al.*, 2010). 1582 bases were sequenced using Sanger sequencing in two consecutive rounds using overlapping primers. Bases from 842 bp to 2423 bp downstream from the start codon of the gene were analysed. All accessions share 99% identity to the consensus sequence on TAIR except for Col-0 with 100% identity. Out of eight SNPs, five were located inside an exon and could be biologically relevant (*Figure 3*).

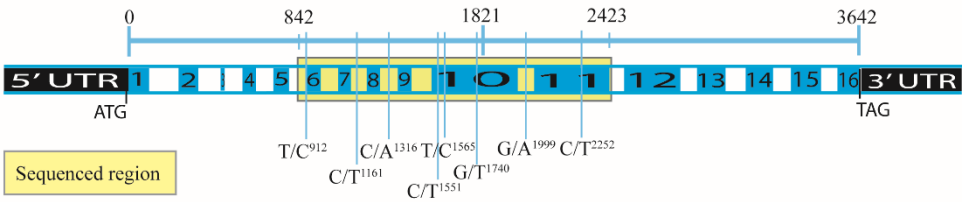


Figure 3. Schematic drawing of the *SRF3* gene. Numbers in blue boxes indicate exons, vertical blue lines indicate SNP locations and the yellow box indicate the sequenced region (842bp to 2423 bp downstream ATG. Intron/exon junctions were predicted bioinformatically. Created from data provided by Alcázar *et al.* (2010).

Most SNPs inside exons are considered transitions compared to the consensus reference sequence and there are two SNPs that deserve attention. The T/C^{912} transition and G/T^{1740} transversion lead to predicted non-synonymous amino acid changes in the protein. (*Table 1*). The T/C^{912} SNP, predicted as a tyrosine to histidine non-synonymous substitution, is however shared by all analysed accessions. The G/T^{1740} phenylalanine to leucine substitution was present in one accession that does not contain the *RPP1*^{Ler}-like cluster. SNPs located outside exons can be viewed in *Supplementary Table 1*.

Table 1. Single nucleotide polymorphisms among 22 Gw accessions inside exons.

SNP/location[a]	Exon	Substitution	Shared by
T/C 912	6	Tyr to His	All
C/T 1551	10	-	C3-Spot4-1/ C3-Spot4-10
T/C 1565	10	-	All
G/T 1740	10	Phe to Leu	Teatralna-1
C/T 2252	11	-	Teatralna-1

[a]Polymorphism location number of bases downstream from ATG

A total of four unique alleles were found among the sequenced accessions compared to the reference genome. Three of which had SNPs inside a predicted exon. No signatures of co-evolution between the *RPP1*^{Ler}-like gene cluster and the *SRF3*^{Gw} could be implied using a phylogenetic analysis by maximum composite likelihood method of judging evolutionary distances (Saitou & Nei, 1987). The *SRF3* alleles did not separate based on the presence or absence of the *RPP1*^{Ler} -like cluster (*Figure 4*). Nor were any sequences from *RPP1*^{Ler}-like incompatible *SRF3* alleles from Kas-2 or Kond compared to the *SRF3*^{Gw} alleles. Teatralna (Gw^{-}) was separated from the rest of the accessions due to the presence of 2 unique SNPs and C3-Spot-5-9 (Gw^{+}) due to having unique SNPs inside intronic regions.

The low genetic diversity (π = 0.0008) and a negative value of Tajima's D (-1.28) can be interpreted to show an excess of high or low frequency alleles in the sampled part of the population. This could indicate a purifying selection at this locus unless the population recently has undergone a genetic bottleneck event (*Table 2*). However, a Z-test of neutral evolution that compared the relative abundance of synonymous and non-synonymous substitutions show that neutral evolution can't be rejected (p-value = 0.824). No evolutionary basis could be confirmed on the genetic variation but a possible explanation for finding lower genetic variability than reported in the region of interest in the Gw population compared to Kas-2 and Kond

(Alcázar *et al.*, 2014) could be speculated upon. The presence of incompatible *RPP1* alleles might limit genetic diversity at *SRF3* by negative selection in Gw^+ individuals and limit the occurrence of autoactivation of immunity. This does however not explain why the allelic diversity is equally low in Gw^- individuals.

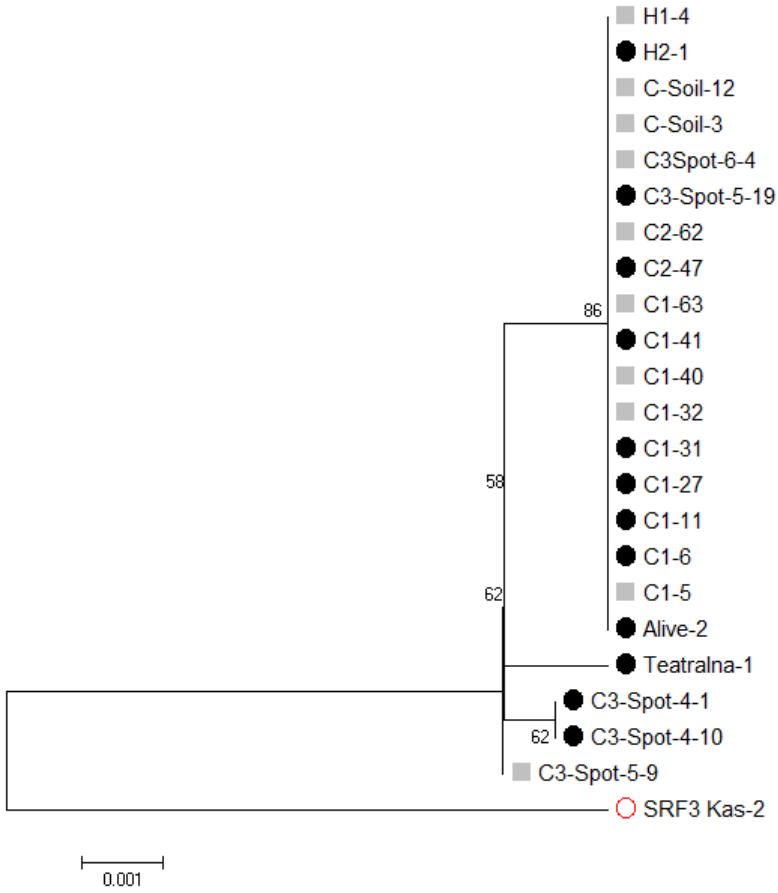


Figure 4. The evolutionary history of the sequenced *SRF3* locus within 22 accessions of the Gorzow population regarding presence or absence of the hybrid incompatibility interacting locus of *RPP1*^{Ler}-like cluster R1 to R8. Grey filled squares = *RPP1*^{Ler}-like cluster present, black filled circles = *RPP1*^{Ler}-like cluster absent, Red transparent circle = Kas-2 outgroup, bar = evolutionary distance based on single nucleotide polymorfisms. The evolutionary distances were computed using the Maximum Composite Likelihood method (Saitou & Nei, 1987).

Table 2. *Results from Tajima's Neutrality Test*[a].

m	S	Θ	π	D
22	8	0.001373	0.000831	-1.280118

m	S	Θ	π	D

[a] m = number of sequences, S = number of segregating sites, Θ = number of pairwise differences, π = nucleotide diversity, D = Tajima's test statistic.

3.2 Prediction reveal possible *SRF3* transcriptional upregulation 24hrs post drought and salt stress

Published meta-data was mined from the *Arabidopsis* eFP browser to identify possible stress factors that trigger *SRF3* expression. This analysis was based on two replicates where the data for average relative expression was published for public use (Kilian *et al.,* 2007). No standard deviation or significance could therefore be calculated. However, this data allows for an overview of how *SRF3* behaves at the transcription level for several abiotic stress situations (*Figure 5*). The fold change in expression of *SRF3* point towards that if *SRF3* has a role in sensing osmotic stress, it's function may stem from being downregulated to 60% of its expression during standard normal conditions at early stages of stress (within 6 hours) and then slowly re-establish its normal expression profile after 24hrs. This type of expression could indicate that *SRF3* is a negative regulator of osmotic stress signalling that senses early onset of osmotic stress. It seems however more likely that *SRF3* would function during drought or salt stress since the expression profile upregulates ~40% after 24hrs for both types of stress and does not show any trend of re-establishing base levels at this time point.

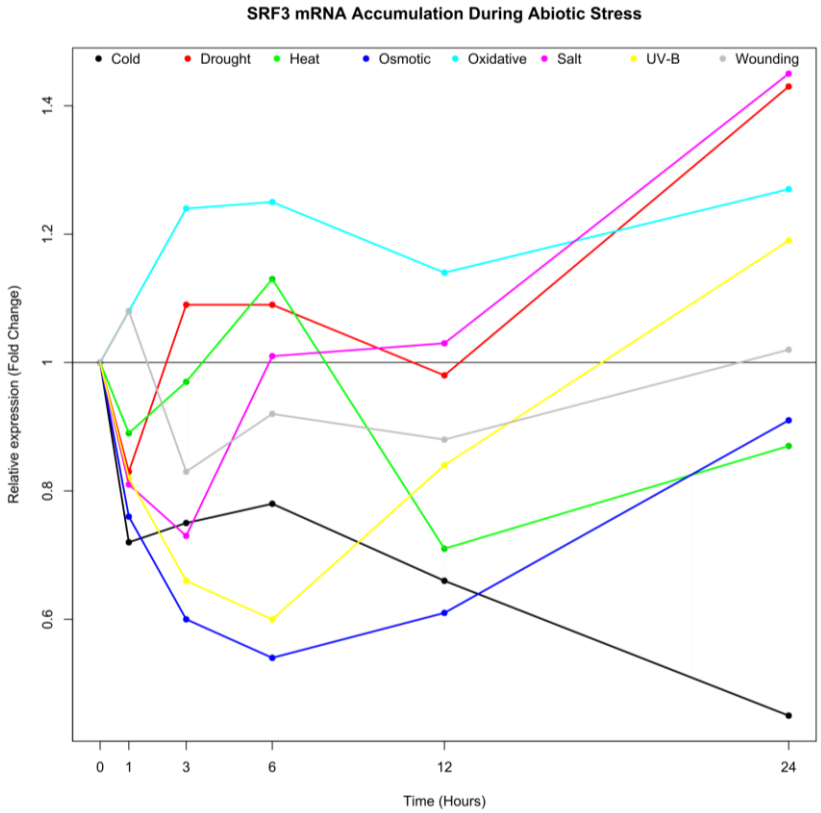


Figure 5. In silico meta-data analysis of the *SRF3* expression profile with different types of stress at different timepoints (0 – 24hrs). Black = cold, red = drought, green = heat, blue = osmotic, turquoise = oxidative, magenta = salt, yellow = UV-B and grey = wounding.

3.3 SRF3 does not repress the plant osmotic stress phenotype

SRF3 may have multiple functional roles including salicylic acid dependent defence signalling against pathogens as reported by Alcázar *et al.* (2010) and ABA-related environmental stress signalling. The preliminary *in silico* results seemed to involve SRF3 with resistance to drought and salt stress and unpublished data (from Rubén Alcázar) indicates SRF3 involvement in osmotic stress specifically. Since osmotic stress is a part of both drought and salt stress, I introduced a series of different osmotic stress levels to homozygous *SRF3* T2 mutants to elucidate a SRF3 function during abiotic stress. *SRF3*-OX (overexpressing *SRF3* in Col-0), *srf3-2*, *srf3-3 srf3-4* (individual Col-0 homolog knockdown lines which have T-DNA insertions at different locations of *SRF3*) and *SRF3*^{Ler}*:GFP* (*SRF3* from L*er* tagged with GFP introduced to Col-0) were grown *in vitro* for 10 days. They were then transferred to agar plates containing 50, 100 and 150 mM mannitol. Plants were grown on these

conditions for 14 days before determining leaf area and plant biomass accumulation as a measure of stress tolerance. Even though some growth problems were evident in the control treatment, a discernible phenotype could be visually determined at 50 mM mannitol. Stunted growth and chlorosis is visible across all genetic lines. At 100 mM – 150 mM mannitol the effects are severe, indicating that toxic conditions were reached. At 200 mM mannitol the plants died 14 days post treatment (*Figure 6*).

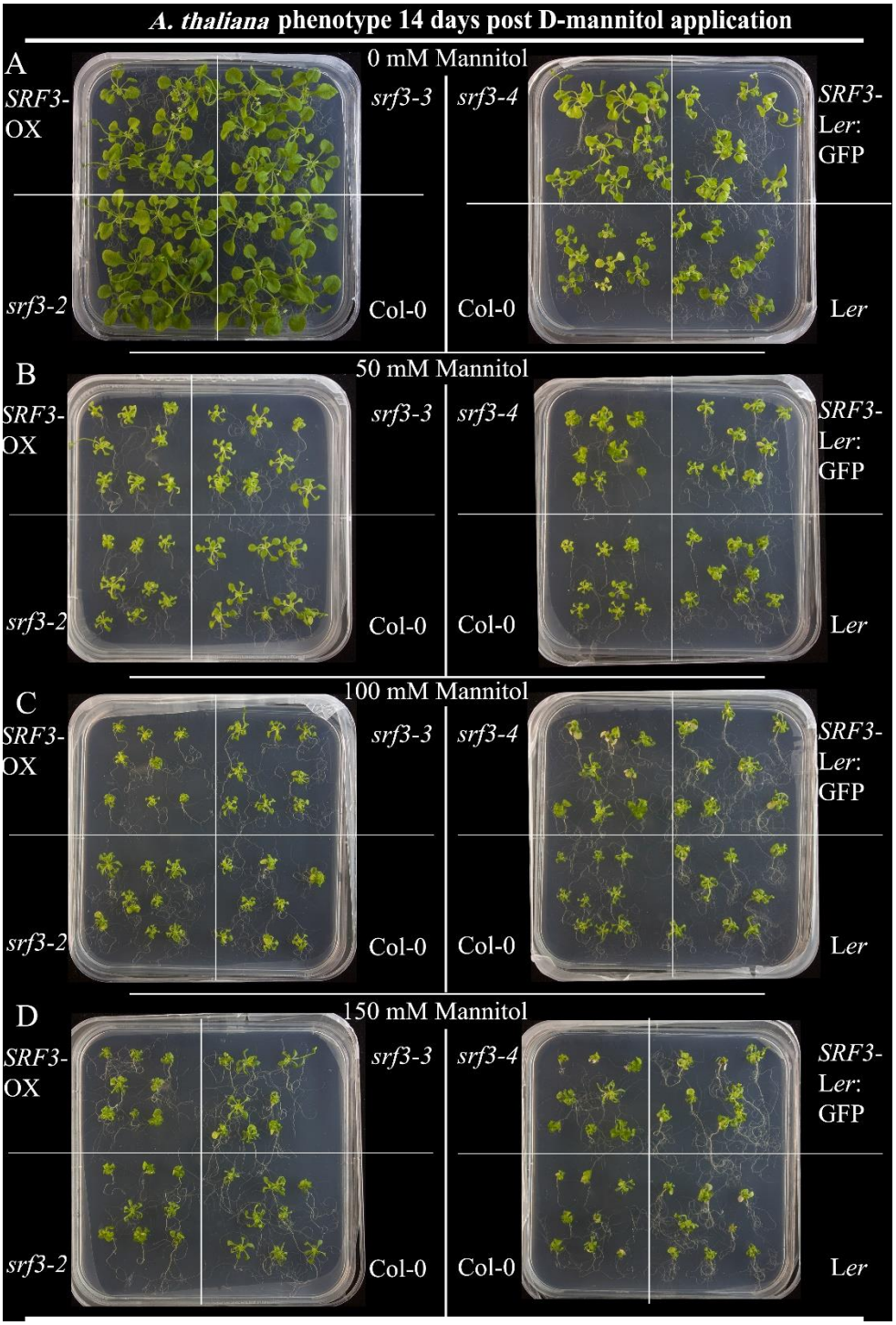


Figure 6. Phenotype of *SRF3* 14 days post mannitol stress 24 days post germination. Abbreviations: *SRF3*-OX = constitutive expression line, *srf3-2*, *srf3-3*, *srf3-4* = knockdown mutant lines, $SRF3^{Ler}$*:GFP* = L*er* version of *SRF3* fused to GFP in a Col-0 background.

A general linear model was implemented to predict which model best described the phenotypic data since it had an approximate normal distribution (QQ-plot not shown). No significances of biological meaning could be inferred on the intra-group treatment level analysing biomass accumulation and leaf area separately in a reduced model (leaf area or biomass explained by genetic line or treatment). Looking at biomass accumulation alone there was a slight increase in mass for the *srf3-4* knockdown line mock treatment compared to the other genetic lines and *srf3-3* (knockdown) had small but significant mass gain at 150 mM mannitol compared to the other lines (see discussion). *Srf3-3* has its second quantile (median) in the same range as the other genetic lines and the differences of the group means range within 0,047g ± 0,076 g (standard deviation) to Col-0 (p = 0,05) with similar results compared to *SRF3*-OX *srf3-2*, *srf3-4* and SRF3^{Ler}:GFP. This shows that a few plants grew atypically large for *srf3-3* at 150 mM. Interestingly, the overexpressing line (*SRF3*-OX) had a small but non-significant trend in biomass increase compared to the other lines at 150 mM. The *SRF3*-OX data contain several outliers that may have affected the trend (*Figure 7*).

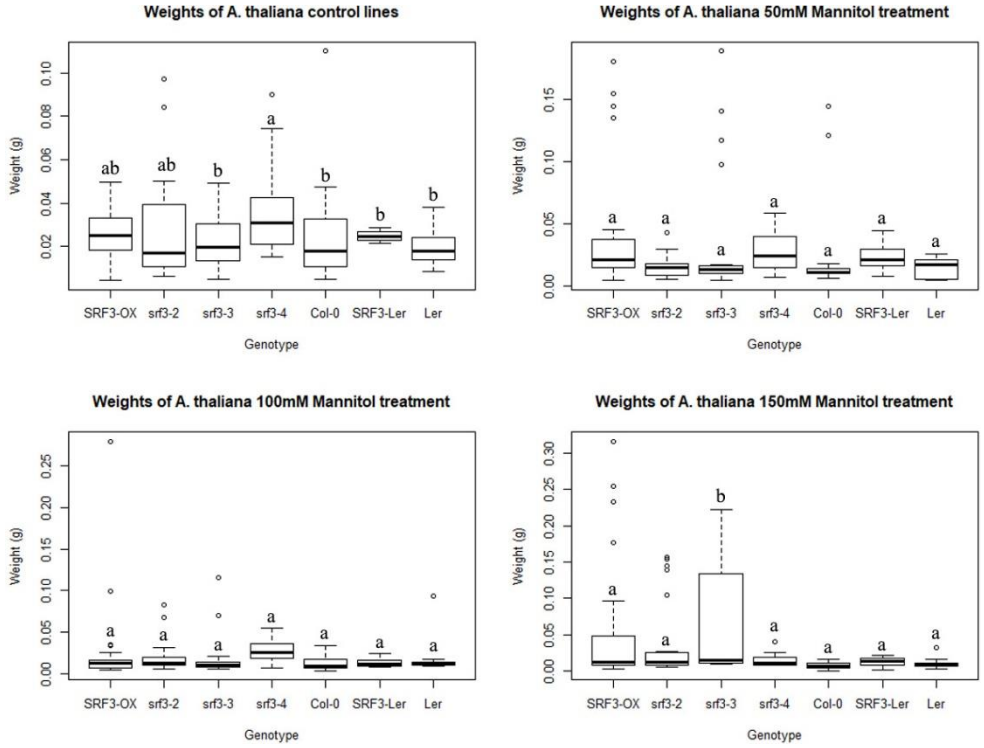


Figure 7. Boxplot of biomass accumulation of *SRF3* lines after 14 days of stress at 50 mM, 100 mM, and 150 mM mannitol treatment. Abbreviations: *SRF3*-OX = constitutive expression, *srf3-2*, *srf3-3*, *srf3-4* = knockdown lines, *SRF3*^{Ler} = L*er* version of *SRF3* fused to GFP. The data consists of three biological replicates n = 32.

The data for leaf area looks comparable to that of plant weight. Again, the *srf3-4* knockdown line seem to have a small growth advantage compared to the rest of the genetic lines under control conditions, but no line grew better than any other in intra-treatment comparisons (*Figure 8*).

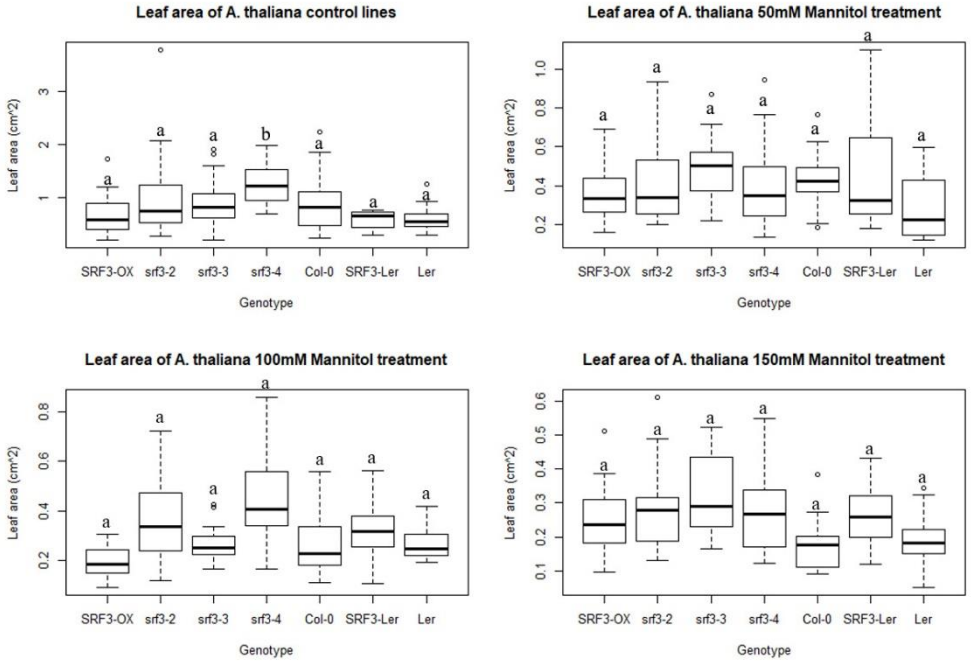


Figure 8. Boxplot of biomass accumulation of *SRF3* lines after 14 days of stress at 50 mM, 100 mM, and 150 mM mannitol treatment. Abbreviations: *SRF3*-OX = constitutive expression, *srf3-2*, *srf3-3*, *srf3-4* = knockdown lines, $SRF3^{Ler}$ = L*er* version of *SRF3* fused to GFP. The data consists of three biological replicates n = 32.

The complete picture emerges when analysing the complete data set at the same time in a full model (plant weight and leaf area explained by genetic line and treatment) (P << 0.001. R^2 = 0.54). See *Supplementary Table 2* for the complete overview. Inter-treatment comparisons are significantly different in biomass accumulation and leaf area but not in intra-treatment comparisons. These results taken together suggests that different levels of mannitol stress have a negative effect on plant growth. Overexpression of the SRF3 protein, seemingly, do not contribute to the overall health or growth capabilities of the plant 14 days post mannitol treatment.

3.4 Genes related to abiotic stress signalling pathways co-express with *SRF3*

An *in silico* meta-analysis of published Affymetrix ATH1 microarray data gathered from Genevestigator was performed to further the investigation of stress sensing by SRF3. The 25 most positively correlated genes with *SRF3* expression when sampling data from abiotic stress and their gene ontology was compiled in *Table 3*. Correlations are determined by Pearson's correlation coefficient (r) and genes biologically relevant to stress induction are highlighted with bold type. A circular correlation plot with *SRF3* in the centre show that three clusters of mutually correlated genes were found (an algorithm helps determine that correlations are not independent) (*Figure 9*).

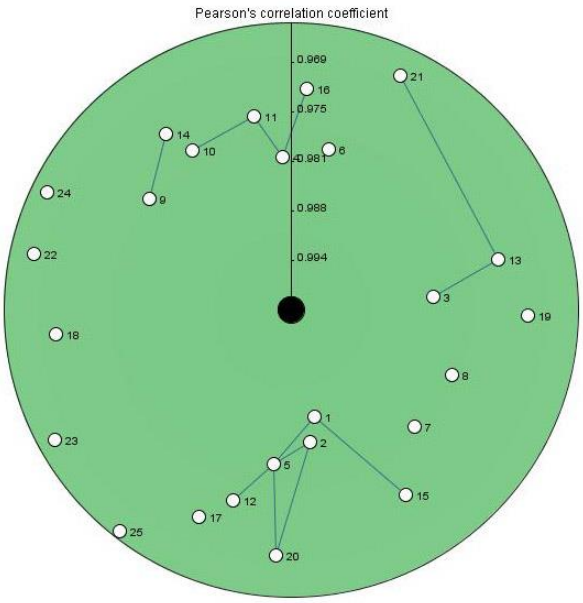


Figure 9. A circular correlation plot showing genes closely correlated to *SRF3* during abiotic stress ($r \geq 0.96$). Genes that share mutual correlation of at least $r \geq 0.985$ are connected a blue line.

It is evident that *SRF3* may be involved in a transcriptional reprogramming at the onset of stress because the list of most correlated genes changes completely when looking at genes in a standard normal condition (data not shown). Interestingly several of the most correlated genes are connected to some form of stress response pathway including the gene *TETRATRICOPEPTIDE-REPEAT THIOREDOXIN*-like *4* (*TTL4*), which is required for osmotic stress tolerance (r = 0.98). *TREHALOSE PHOSPHATASE/SYNTHASE 11* (*ATTPS11*) is presumed to be involved in trehalose biosynthesis which accumulates under heat and chilling stress in *A. thaliana* and act as a membrane stabilizer (Fernandez *et al.*, 2010). Several genes in the ABA signalling pathway co-express closely with *SRF3* including a putative *PROTEIN PHOSPHATASE 2C* (*PP2C15*, r = 0.97), a possible regulator of ABA signalling during drought stress. Two genes involved in reactive oxygen species regulation, *G-BOX BINDING FACTOR 1* (*GBF1*, r = 0.98) and *MULTIPROTEIN BRIDGING FACTOR 1C* (*MBF1C*, r = 0.97) and two genes involved in ubiquitination (*UBIQUITIN-SPECIFIC PROTEASE 25* (*UBP25*, r = 0.97) and *PHLOEM PROTEIN 2-*

A13 (*PP2*-A13, r = 0.97)). PP2-A13 is also included in the response to wounding. The MBF1C also annotates to be a transcriptional co-activator that upregulates during drought and salinity stress. Two membrane bound genes that are involved in disease resistance also coregulates closely to SRF3. A TIR-class disease resistance protein (At1G61100, r = 0.98) and another putative receptor bound kinase (AT1G51890, r = 0.96) even though only samples from abiotic stress were analysed.

Table 3. SRF3 co-expressed genes during abiotic stress in *A. thaliana*[a].

Nr[b]	Gene	Locus	R[c]	Description[d]
1	*MED19A*	AT5G12230	0.99	Upregulation of transcription from RNA polymerase II promoter
2	-	AT5G47430	0.98	Zinc ion binding, nucleic acid binding
3	-	AT2G25800	0.98	Hypothetical protein
4	*ATMPK17*	AT2G01450	0.98	MAP kinase, protein autophosphorylation
5	*GK-2*	AT3G57550	0.98	Guanylate kinase activity
6	*GBF1*	AT4G36730	0.98	**Negative regulation of gene expression, regulation of cell aging and hydrogen peroxide metabolic process**
7	-	AT1G61100	0.98	**Disease resistance protein (TIR class)**
8	*TTL4*	AT3G58620	0.98	**Required for tolerance response to osmotic stress**
9	-	AT1G61370	0.98	S-locus lectin protein kinase family protein
10	*ATTPS11*	AT2G18700	0.98	**Enzyme putatively involved in trehalose biosynthesis**
11	*ARP3*	AT1G13180	0.98	Actin filament organization, cell/ trichome morphogenesis and growth
12	*UBP25*	AT3G14400	0.97	**Ubiquitin-specific protease**
13	*RRP6L 2*	AT5G35910	0.97	rRNA processing
14	*MBF1C*	AT3G24500	0.97	**Transcriptional coactivator. Elevated expression in response to pathogen infection, salinity, drought, heat, hydrogen peroxide, and application of abscisic acid or salicylic acid**
15	*GA2*OX*6*	AT1G02400	0.97	*GIBBERELLIN 2*-OX*IDASE 6*
16	*KEA6*	AT5G11800	0.97	*K^+ EFFLUX ANTIPORTER 6*
17	*SWP*	AT3G04740	0.97	Mediator of RNA polymerase II transcription subunit 14
18	-	AT4G32160	0.97	Phox (PX) domain-containing protein, phosphatidylinositol binding
19	*PP2-A13*	AT3G61060	0.97	**Carbohydrate binding, protein binding and ubiquitination, response to wounding**
20	-	AT1G68410	0.97	**Putative protein phosphatase 2C 15**

Table 3 cont. SRF3 co-expressed genes during abiotic stress in *A. thaliana*[a].

21	-	AT4G30910	0.97	Amino-/metallopeptidase activity, manganese ion binding
22	*ELF7*	AT1G79730	0.97	*Early flowering 7*. Negative regulation of flower devel-opement
23	*SEC8*	AT3G10380	0.97	Subunit of exocyst complex 8. Pollen germination, pollen tube growth
24	*SSP4*	AT5G46410	0.97	*SCP1-like small phosphatase 4*
25	-	AT1G51890	0.96	**Putative LRR receptor-like protein kinase**

[a] Genes with a positive co-expression correlation with *SRF3* above R ≥ 0,95. Genes with unknown function are excluded
[b] Number corresponding to gene network in *Figure.9*
[c] Pearson's correlation coefficient score
[d] Description mined through NCBI Genbank and TAIR. Bold type = genes involved in abiotic stress signalling.

4 Discussion

One of the aims of this project was to identify genetic variability at the L*er*/Kas-2 and *Ler*/Kond hybrid incompatibility locus, *SRF3*, and investigate a possible co-evolutionary scenario with *RPP1*^{Ler}-like genes. The Gw population analysed in this article segregates for incompatible *RPP1*^{Ler}-like alleles that possibly shows signs of co-evolution with allelic forms of *SRF3*. The 4 *SRF3* alleles found among the 6 SNPs in the subset of the population analysed did not separate according to the basis of presence/absence of the incompatible *RPP1*^{Ler}-like alleles in the gene tree. Based on this result alone one could not conclude that the cause of the dynamics at this locus is due to co-evolution by negative frequency dependent selection. The *SRF3* locus seemed to be conserved with low levels of natural variability, as shown by the low nucleotide diversity and few segregating sites. Less diversity was found than previously reported for Kas-2 and Kond (Alcázar *et al.*, 2010) Negative selection (and therefore co-evolution) could act on the population and by that mechanism limit genetic diversity and incompatibilities. Negative selection was also indicated by the negative value of Tajima's D even though neutral evolution or drift could not be excluded by the Z-statistics. This result could be due to the low sample size of 22 individuals in the analysis. Only part of the *SRF3* gene was sequenced as well. Low frequency alleles and SNPs outside the sequenced region were likely missed. It is probable that the whole population contains more genetic variation at this locus and more allelic variants than the ones found in this small-scale experiment.

It is also possible that even though both compatible and incompatible *RPP1*^{Ler}-like alleles are present in the population, only compatible L*er* forms of *SRF3* could be present. This experiment cannot rule that hypothesis out since no sequences of alleles known to be incompatible from Kas-2 or Kond were included, only a general reference sequence from Kas-2 which placed as an outgroup in the genetic tree.

During the sequence analysis, two amino acid substitutions were found that could potentially have had an impact on the folding of the SRF3 protein. It is possible that

the hits for non-synonymous substitutions are false positives because the translation from the annotated reference coding sequence on TAIR had slightly different intron/exon boundaries compared to the bioinformatically predicted boundaries that was used. No substitutions and therefore no conformational changes were found while comparing the *SRF3* sequences to the amino acid sequence on TAIR. To confirm these results, sequencing and aligning mRNA instead of translating genomic DNA into protein sequences or perform a de novo peptide sequencing by mass spectronomy on purified proteins of interest would be more informative.

In the future, one could contrast the full gene sequence at both loci and determine if certain alleles of *SRF3* are at higher frequency in the Gw^- subpopulation but excluded from the Gw^+ subpopulation. It is also worth comparing all the $Gw^{+/-}$ *SRF3* alleles to Kas-2 and Kond versions to see whether the incompatible alleles that originated from relatives in central Asia have any similar alleles in Gw. From that basis one could speculate on co-evolution of allelic variants at these two loci. If one includes similar data for other closely related populations and accessions in Eurasia one could begin to trace the evolutionary history with a greater perspective.

The second main hypothesis was whether *SRF3* could be involved in osmotic stress tolerance as indicated by the Alcázar group (unpublished data) and the *in silico* investigation into *SRF3* expression during abiotic stress. Hints that *SRF3* could be upregulated during drought and salt stress were provided using online data from the eFP browser to survey *SRF3* expression during several types of stress. Analysing published data on co-regulated genes through the Genevestigator database placed several ABA-dependent genes in close correlation to *SRF3* expression as well. Among a gene that encode a protein required for tolerance response to osmotic stress (TTL4), a putative PP2C protein which is an important regulator of ABA signalling during drought stress (Santiago *et al.*, 2009) and an ABA dependent transcriptional co-activator during saline, drought and heat stress (MBF1C) (Zandalinas *et al.*, 2016). A third protein that could have implications in drought stress since it annotates to being involved in trehalose biosynthesis (ATTPS11) which has been reported to be a protective sugar (Fernandez *et al.*, 2010) correlated very closely as well.

All these hints lead me to investigate SRF3 function during stress induced by mannitol. This is because at early onset of drought and salt stress the osmotic pressure potential is decreased and the concentration of osmolytes increases as water availability drops. To study the potential *SRF3* involvement a series of mannitol treatments were used to simulate an osmotic stress challenge. An overexpressing line, three knockdown lines and one genetic line containing the Gw native L*er* version of

SRF3. Results from the main experiment show that inter-treatment comparisons (between lines experiencing different concentrations of mannitol) had a real and significant effect while including both biomass accumulation and leaf area in the analysis. This validates that the plants were stressed in a mannitol dose dependent manner, but it also demonstrates that overexpressing *SRF3* does not give any clear advantage (health or growth) during osmotic stress. The *SRF3*-OX (overexpressing line) did accumulate slightly more biomass than most other lines, including the Col-0 control line at 150 mM mannitol 14 days post treatment but the effect was not significant. It would be interesting to redo the experiment using completely independent observations. Up to 32 plants shared growth space in the petri dishes in this experiment, this led to that a third measurement, root length, could not be analysed since the roots grew together. This entanglement could have affected the results of the experiment since micro-competition for growth space was also a factor. If the test was to be repeated it would be interesting to include double and triple mutants of *SRF* since any effect of individual *SRF* genes could be small and masked by 7 other possibly redundant genes. It would also be good to measure the mRNA-accumulation of the knockdown and overexpressing lines. It is possible that the lines were not strongly down or up-regulated because no analysis was made on protein or RNA abundance due to time constraints. This could explain why *srf3-3* and *srf3-4* grew better in 150 mM mannitol and mock treatment respectively. They may in fact be overexpressing lines due to the positioning of the T-DNA or have been given a burst of activity since expression is stochastic, especially when the DNA sequence has been manipulated. These types of side effects could have implications in abiotic stress or development (Raj & van Oudenaarden, 2008). This reasoning goes for *SRF3*-OX as well. Since gene activity was not tested in advance, this line may not be overexpressing *SRF3*. The host lab used the line as an overexpressor because of T-DNA insertion in the 5' UTR. However, only 3% of all T-DNA insertions in the 5' UTR enhances transcription activity, as explained by Wang (2008).

For future work it would be intriguing to repeat this osmotic stress experiment due to the *a priori* reasoning but include double and triple mutants and a confirmed *SRF3* overexpressed line. One could include several more phenotypic characters for osmotic stress like root length and cellular osmopotential to gain a better understanding of the responses between the genetic lines. A complementary experiment where European wild type *Arabidopsis* plants are subjugated to stress could also be used to experimentally measure *SRF3* mRNA accumulation (Real Time qRT-PCR) to gain an insight into the transcription activities of this protein during stress,

All in all, these results show that much more work is needed to understand the intricate nature of *SRF3* signalling and molecular function. Previous results show that

it is involved in HI between ecotypes of *A. thaliana* and disease resistance in Kas-2 and Kond (Alcázar *et al.*, 2010) and the results in this study indicate that *SRF3* could be a conserved RLK in the Gw population of L*er* relatives and have multifunctional roles in drought and salt stress tolerance. Further work is needed to confirm these preliminary results.

5 Materials & Methods

5.1 Plant Materials & Growth Conditions

5.1.1 Plant materials

Different accessions and mutants of *A. thaliana* from various sources was used to produce results. For the analysis of natural variation of the *SRF3* locus a few of the recently discovered accessions of *A. thaliana* named the Gorzow accessions was primarily used. These accessions, located at two different sites in Gorzów Wielkopolski, Poland, are closely related to the Northern European L*er* ecotype and was found and donated by Alcázar *et al.* (2014). Within this population ten of the accessions that contain the *RPP1*^{Ler}-like cluster *R1* to *R8* from L*er* was chosen (Gw^{+}) and contrasted to twelve of those that do not (Gw^{-}) (*Table 4*) and the more distantly related ecotype Kas-2.

Table 4. Gorzów accessions used to elucidate natural variation at the *SRF3* locus in *A. thaliana*[a].

Gw^{+}[b]	Gw^{-}[b]
H1-4	H2-1
C-Soil-12	C-3-Spot-5-19
C-Soil-3	C2-47
C2-62	C1-41
C1-63	C1-31
C1-40	C1-27
C1-32	C1-11
C1-5	C1-6
C3-Spot-5-9	Alive-2
C3-Spot-6-4	Teatralna-1
-	C3-Spot-4-1
-	C3-Spot-4-10

[a] Accessions do not yet have seed stock ID.
[b] $Gw^{+/-}$ indicate whether the RPP1^{Ler} -like cluster is present (+) in the accession or absent (-).

T-DNA insertion mutants used to investigate *SRF3* involvement in response to drought stress were obtained from the *Arabidopsis* Biological Resource Center (ABRC) and/ or GABI-DUPLO (*Table 5*). The translational reporter mutant *SRF3*^{Ler}:GFP (native *SRF3* promoter and gene

from L*er* with a GFP tag) was previously generated and donated by Alcázar *et al*. (unpublished). Wild type seeds from Col-0 and L*er* were previously generated or obtained from the Alcázar lab stock. The double mutant (*srf1-srf3*) and the parental lines (P1 and P2) were later discarded from the experiment due to a lack of time and primers.

Table 5. Mutant knockdown, overexpressing and transgenic *strubbelig-receptor family* lines used in drought stress and transcriptome experiments.

Genotype	Effect on gene expression	NASC/Seed stock ID
SRF3-OX	Upregulated	N529908/SALK_029908
srf3-2	Knockdown[b]	N656527/SALK_057621
srf 3-3	Knockdown[b]	N501389/SALK_001389
srf 3-4	Knockdown[b]	N589934/SALK_089934
SRF3^{Ler}:GFP	Unchanged	-
srf3-srf1[a]	srf double knockdown mutant[c]	N2103287/ CS2103287
P1[a]	*srf1* double mutant parent[c]	N662999/SALK_075679
P2[a]	*SRF3* double mutant parent[c]	N438269/-[d]

[a] Only used in the transcriptome experiment
[b] Seeds obtained from the *Arabidopsis* Biological Resource Center
[c] Seeds obtained from GABI-DUPLO
[d] Only available from NASC

5.1.2 Plant growth conditions (sequencing and abiotic stress experiment)

Plants used for sequencing the *SRF3* locus in the 22 $Gw^{+/-}$ accessions and plants used in abiotic stress experiments were brought up the same way and in the same conditions. Seeds were surface sterilized using a liquid sterilisation protocol (*Supplementary Table 3*) and plated on Petri dishes with Murashige-Skoog medium containing 0,8% plant agar for sequencing and 1% plant agar for growth prior to stress treatment (*Supplementary Table 4*). The sterile seeds were stratified in a dark cold room at 8˚C for seven days prior to being put in light. The seeds were let to grow *in vitro* for 10 days post germination (DPG) in a horizontal position (in 20 ˚C, 16 h/8 h bright/dark with a light intensity of 80 ± 10 µmol × m^{-2} × s^{-1}) prior to being sampled for DNA extraction for sequencing or transferred to growth plates containing D-mannitol to simulate drought stress.

A small scale pre-experimental abiotic stress test using D-sorbitol was performed using only a few , *srf3-2*, *srf3-3*, *srf3-4 srf3-srf1*, *Col-0* SALK line control), *SRF3^{Ler}:GFP* and *L*er (control for *SRF3^{Ler}:GFP*) plants in order to evaluate at which

sugar concentration the stress became toxic to help select concentrations for the simulated drought experiment (data not shown). Seedlings for the small-scale test were surface sterilized (*Supplementary Table 3*) and plated on square growth plates with a nylon mesh atop of a 1% agarose gel that contained nutrients described in *Supplementary Table 4*, stratified for 7 days and let grown for ten days prior to transplantation. Seedlings in the same growth stage were transferred to a mock treatment and D-sorbitol treatments at 50 mM increments up to 200 mM and were stressed for seven days when toxicity became apparent. The *srf3-srf1* mutant was discarded prior to the main experiment due to poor growth during the allotted time. 24-32 plants of each remaining wild type and mutant plant lines at similar size (two to three rosette leaves with a length greater than 1 mm) were transferred to 120 mm × 17 mm square growth plates in sterile conditions. The growth plates contained 50 ml growth medium with 1% plant agar (*Supplementary Table 4*) in addition to, 0 mM, 50 mM, 100 mM, 150 mM or 200 mM D-mannitol. Three biological replicates were used for statistical reasons (yielding 32 observations per line and treatment) and the growth plate positions were switched in a randomized order every other day to remove sources of error in the ensuing statistical analysis due to undetectable differences in growth conditions. Symptoms were surveyed at 3, 5, 9 and 14 days of stress. Analysis of the stress phenotype was performed 24 days post germination (14 days post stress treatment).

5.1.3 Plant growth conditions for the *SRF3* mutant homozygosity test

Non-sterilized seeds were stratified for 7 days in small Petri dishes with a wet cotton filter inside. These were subsequently transferred directly to commercial potting soil containing 50% perlite in pots (6 cm × 6 cm (width/depth) and put in a phytotron at constant conditions. Two seeds of each line were grown together along with two technical replicates and were let to self-fertilize and complete their life cycle to amplify seeds for each line. The phytotron held a temperature of 22 ˚C with a humidity of 65 ± 5% and the light duration was set to a long day regime (16 hours bright, 8 hours dark) with a light intensity of 100 ± 15 µmol × m^{-2} × s^{-1}. This process was repeated due to an unexpected technical error in the phytotron. 28 plants per genetic line of the second generation (F_2) were propagated using the same strategy as for the first. F_2 plant growth and sampling was repeated due to a suspected technical error at a late stage. This time, six individuals of each subline was pooled per genetic line during extraction and subsequent PCR reactions with two technical replicates. The smaller sample number and pooling strategy was used due to the extra time constraint. It is possible to pool several individuals for each genetic line, because if the PCRs show homozygosity, the chance is high that all those 6 sublines in the pool was homozygous). These seeds were surface sterilized (*Supplementary Table 3*) and

plated on Petri dishes with 25 ml GM with agar concentration of 0.8%, stratified for five days and grown for 10 days. Whole plants were harvested.

5.2 DNA Extraction Methods & Sample Preparations

5.2.1 Evolutionary analysis of *A. thaliana* Gw accessions

About 20 seedlings of each accession were put in safe lock Eppendorf tubes containing 1,7 mm – 2,1 mm silica glass beads, snap freezed in liquid nitrogen and homogenized in a Silamat (Ivoclar vivadent) for 8 s. Snap freezing and tissue homogenization was repeated once. Total genomic DNA was subsequently isolated using the DNeasy Plant Mini Kit (Quiagen) as described in the kit manual except that the gDNA was eluted 2 × 50 µl in the final step. An agarose gel electrophoresis of isolated gDNA was performed to check the yield/quality and quantified using Nanodrop 2000 (Thermo Scientific). The agarose gel electrophoresis was run at 100 V for 35 min in 1 × TAE and the 100 ml gel contained 1% agarose and 5 µl ethidium bromide and 2 µl gDNA per sample was loaded with 0,5 µl loading dye. Another 20 µl ethidium bromide was added to the trays anode to enhance the signal of any DNA band. A 1 kb ladder from Invitrogen was used to estimate the size of the well bands. An optimized PCR protocol was used with a MJ Research Peltier Thermal Cycler to amplify the region of interest in the *SRF3* gene (see PCR parameters and primers in *Supplementary Figure 1* and *Supplementary Table 5* and *Supplementary Table 6*). Another agarose gel electrophoresis was performed using all of the 20 µl PCR amplified volume to purify the *SRF3* DNA from any unspecific nucleic acids and unused primers. This time without a ladder and the gel was run for 60 min at 60 V. the resulting bands were excised from the gel (~300 mg) using an E-Gel Safe Imager from Invitrogen and a scalpel. The DNA in the gel was extracted with a Nucleospin Gel & PCR Clean-Up kit from Macherey-Nagel to remove any remaining salts or impurities. No subcloning was performed prior to sequencing the DNA. Instead, 5 µl of each sample and 1 µl of a primer (diluted 1:20 from stock to 5 µM) was added directly to a PCR strip and dried in a thermal cycler at 80 ˚C for 15 min and sent to Unitat de Genomica, Serveis Cientificotecnics, Barcelona for sequencing using an automated Sanger technique. Since the region of interest in the *SRF3* gene is 2,6 kb long and Sanger sequencing gives ~1 kb quality reads another 5 µl of each sample was prepared and dried together with a different primer to cover a longer region of the gene of interest.

5.2.2 T-DNA mutant two-step genotyping

Two small, fresh leaves were cut from young F_2 plants with a scalpel, quickly dried with a paper and loaded into 2 ml deep DNAse and RNAse free 96-well polypropylene extraction plates containing two medium sized metal beads. The racks were weighed but never flash frozen in liquid nitrogen prior to tissue homogenisation in a VWR Star-beater (2 × 1 min at a frequency of 30 Hz). The ensuing DNA extraction method was based on a modified version of a protocol from Dellaporta *et al*. (1983). 300 µl extraction buffer (*Supplementary Table 7*) was added to each sample while not working on ice to keep SDS from precipitating. Samples were centrifuged at 4000 rpm for 7 min in an Orto Alresa Digicen 21 (using rotor RT 150) before transferring 200 µl supernatant to 96-well S-blocks carefully to not disturb the pellet. 20 µl of 3 M sodium acetate (NaOAc) at pH 4,8 and 400 µl ethanol 96% (EtOH) was added and mixed gently by pipetting. Thereafter the samples were stored overnight at -20 ˚C to allow precipitation of DNA and centrifuged at 4000 rpm for 7 minutes the next morning. The supernatant was discarded, 400 µl EtOH 70% added and gently mixed by pipetting before repeating a 7 min centrifugation at 4000 rpm. The supernatant was discarded by decantation and the tubes let to dry for 10 min in a fume hood to ensure that no EtOH was left in the tubes. To elute the DNA 150 µl ddH_2O was added and the pellet re-suspended by slowly pipetting up and down and left in 4 ˚C overnight. As a final step the samples were centrifuged as before the morning after and 100 µl of the supernatant transferred to new tubes to get clear samples. The recovered DNA was analysed in a Nanodrop 2000. The ensuing PCR and gel analysis was performed for all lines in order to verify that all mutant lines contained the T-DNA insert and that the lines were homozygous for the insert and not segregating. This two-step PCR genotyping verification was done twice for all lines, once using genomic primers and once using T-DNA border primers, as described by O'Malley *et al*. (2015). See *Supplementary Table 8, Supplementary Table 9* and *Supplementary Figure 2* for PCR reagents, primers and conditions respectively. The following gel electrophoresis for all amplicons was run for 90 min at 100 V with a 2% agarose gel with 25 µl EtBr × $(0.1\ dm^3)^{-1}$ gel, a 1kb ladder from Invitrogen and 0,5 µl loading dye for each well.

5.3 Leaf Area & Plant Weight Measurements

Leaf area was measured using imageJ2 (Schindelin *et al.*, 2015) after photographing the growth plates in a black box. A global distance scale was set up and leaf area detected by setting threshold colour values to precisely cover each plant.

Fresh plant weight was noted for each plant by quickly removing individual plants and measuring them on a scale (Sartorius ENTRIS124-1S lab balance). Plants were

exposed to minimal amount of time outside the plate to reduce mass loss due to water evaporation.

5.4 *In Silico* analysis

SRF3 Co-regulated (expressed) genes were mined through Genevestigator version 3 (Hruz *et al.*, 2008) using the dataset from the Affymetrix *Arabidopsis* ATH1 Genome Array. The gathered data came from abiotic stress experiments concerning anoxia, cold, drought, heat, hypoxia, oxidative, osmotic, light, salt, submergence stress and wounding. Gene ontology was mined through The *Arabidopsis* Resource (TAIR). *SRF3* expression data in different abiotic stress environments were mined through the eFP browser (Kilian *et al.*, 2007). All expression data on the various stress conditions was compared and normalized to expression profiles at the Myb domain protein (At3g27340) locus for *SRF3* expression levels as suggested by eFP browser.

5.5 Statistical Analyses & Data Processing

5.5.1 Genetic variability analysis of the Gw *SRF3*

Sequences retrieved from sequencing two stretches of the Gw accession *SRF3* locus were processed using Chromas version 2.4.4 and Bioedit version 7.2.5. Assembled reads were end trimmed and quality checked. Erroneous reads were resolved, and ghost reads deleted manually by comparing base calls to the chromatogram. The sequences were then concatenated followed by alignment to wild type Col-0 ($Gw^{+/-}$ genetic background) by the ClustalW multiple alignment algorithm and bootstrapped 1000 times. The evolutionary history was inferred using the Neighbour-Joining method, Tajima's D and a Z-test for neutral evolution based on the different SNPs using MEGA version 7.0.14. The tree is drawn to scale, with branch lengths in the same units as those of the evolutionary distances used to infer the phylogenetic tree. The evolutionary distances were computed using the Maximum Composite Likelihood method which was bootstrapped 10000 times (Saitou & Nei, 1987).
SNPs were found by pairwise comparison of *SRF3* alleles between Gw accessions and Col-0. The SNPs were then manually checked for synonymous/nonsynonymous substitutions using predicted coding regions and verified using DNAsp5. The relative SNP locations were annotated to the full length *SRF3* gene by aligning the sequenced FASTA-files with the *SRF3* consensus sequence from TAIR using NCBI blast.

5.5.2 Abiotic stress experiment

All statistical data analyses and plots were done using R Command version 3.3.1. A probability density (Q-Q) plot and a Shapiro-Wilk normality test was performed to visually and numerically check if the residuals followed an approximate normal distribution and whether parametric tests could be used. A general linear model was used to define how well the data fit my model and to define which model was best to use in further analysis. A one-way ANOVA was performed using genetic line as predictor and fresh weight and leaf area as response variables. A post-hoc test (Tukeys multiple comparisons of means) was employed using the R package Multcomp (Torsten Hothorn *et al.*, 2008) and the degree of differences of means checked by a standard Students t-test.

6 References

Alcázar, R., García, A. V., Kronholm, I., de Meaux, J., Koornneef, M., Parker, J. E. & Reymond, M. (2010). Natural variation at Strubbelig Receptor Kinase 3 drives immune-triggered incompatibilities between *Arabidopsis thaliana* accessions. *Nature Genetics*, 42(12), pp 1135–1139.

Alcázar, R., García, A. V., Parker, J. E. & Reymond, M. (2009). Incremental steps toward incompatibility revealed by *Arabidopsis* epistatic interactions modulating salicylic acid pathway activation. *Proceedings of the National Academy of Sciences*, 106(1), pp 334–339.

Alcázar, R. & Parker, J. E. (2011). The impact of temperature on balancing immune responsiveness and growth in *Arabidopsis*. *Trends in Plant Science*, 16(12), pp 666–675.

Alcázar, R., von Reth, M., Bautor, J., Chae, E., Weigel, D., Koornneef, M. & Parker, J. E. (2014). Analysis of a plant complex resistance gene locus underlying immune-related hybrid incompatibility and its occurrence in nature. *PLoS Genetics*, 10(12), e1004848.

Alonso-Blanco, C., Aarts, M. G. M., Bentsink, L., Keurentjes, J. J. B., Reymond, M., Vreugdenhil, D. & Koornneef, M. (2009). What has natural variation taught us about plant development, physiology, and adaptation? *The Plant Cell*, 21(7), pp 1877–1896.

Bikard, D., Patel, D., Metté, C. L., Giorgi, V., Camilleri, C., Bennett, M. J. & Loudet, O. (2009). Divergent evolution of duplicate genes leads to genetic incompatibilities within *A. thaliana*. *Science*, 323(5914), pp 623–626.

Bomblies, K., Lempe, J., Epple, P., Warthmann, N., Lanz, C., Dangl, J. L. & Weigel, D. (2007). Autoimmune response as a mechanism for a Dobzhansky-Müller-type incompatibility syndrome in Plants. *PLoS Biology*, 5(9), p e236.

Dellaporta, S. L., Wood, J. & Hicks, J. B. (1983). A plant DNA minipreparation: Version II. *Plant Molecular Biology Reporter*, 1(4), pp 19–21.

Eyüboglu, B., Pfister, K., Haberer, G., Chevalier, D., Fuchs, A., Mayer, K. F. & Schneitz, K. (2007). Molecular characterisation of the STRUBBELIG-RECEPTOR FAMILY of genes encoding putative leucine-rich repeat receptor-like kinases in *Arabidopsis thaliana*. *BMC Plant Biology*, 7, p 16.

Fernandez, O., Béthencourt, L., Quero, A., Sangwan, R. S. & Clément, C. (2010). Trehalose and plant stress responses: friend or foe? *Trends in Plant Science*, 15(7), pp 409–417.

Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W., Zimmermann, P., Hruz, T., Laule, O., Szabo, G., Wessendorp, F., Bleuler, S., Oertle, L., Widmayer, P., Gruissem, W. & Zimmermann, P. (2008). Genevestigator V3: A reference expression database for the meta-analysis of transcriptomes. *Advances in Bioinformatics*, 2008, e420747.

Kilian, J., Whitehead, D., Horak, J., Wanke, D., Weinl, S., Batistic, O., D'Angelo, C., Bornberg-Bauer, E., Kudla, J. & Harter, K. (2007). The AtGenExpress global stress expression data set: protocols, evaluation and model data analysis of UV-B light, drought and cold stress responses. *The Plant Journal*, 50(2), pp 347–363.

Nicotra, A. B., Atkin, O. K., Bonser, S. P., Davidson, A. M., Finnegan, E. J., Mathesius, U., Poot, P., Purugganan, M. D., Richards, C. L., Valladares, F. & van Kleunen, M. (2010). Plant phenotypic plasticity in a changing climate. *Trends in Plant Science*, 15(12), pp 684–692.

O'Malley, R. C., Barragan, C. C. & Ecker, J. R. (2015). A User's Guide to the *Arabidopsis* T-DNA Insertion Mutant Collections. In: Alonso, M. J. & Stepanova, N. A. (Eds) *Plant Functional Genomics: Methods and Protocols*. pp 323–342. New York, NY: Springer New York. ISBN 978-1-4939-2444-8.

Orr, H. A. (1996). Dobzhansky, Bateson, and the Genetics of Speciation. *Genetics*, 144(4), pp 1331–1335.

Osakabe, Y., Yamaguchi-Shinozaki, K., Shinozaki, K. & Tran, L.-S. P. (2013). Sensing the environment: key roles of membrane-localized kinases in plant perception and response to abiotic stress. *Journal of Experimental Botany*, 64(2), pp 445–458.

Raj, A. & van Oudenaarden, A. (2008). Stochastic gene expression and its consequences. *Cell*, 135(2), pp 216–226.

Saitou, N. & Nei, M. (1987). The neighbor-joining method: a new method for reconstructing phylogenetic trees. *Molecular Biology and Evolution*, 4(4), pp 406–425.

Santiago, J., Rodrigues, A., Saez, A., Rubio, S., Antoni, R., Dupeux, F., Park, S.-Y., Márquez, J. A., Cutler, S. R. & Rodriguez, P. L. (2009). Modulation of drought resistance by the abscisic acid receptor PYL5 through inhibition of clade A PP2Cs. *The Plant Journal*, 60(4), pp 575–588.

Schindelin, J., Rueden, C. T., Hiner, M. C. & Eliceiri, K. W. (2015). The ImageJ ecosystem: An open platform for biomedical image analysis. *Molecular Reproduction and Development*, 82(7–8), pp 518–529.

Torsten Hothorn, Bretz, F. & Westfall, P. (2008). Simultaneous inference in general parametric models. *Biometrical Journal*, 50(3), pp 346–363.

Tran, L.-S. P., Urao, T., Qin, F., Maruyama, K., Kakimoto, T., Shinozaki, K. & Yamaguchi-Shinozaki, K. (2007). Functional analysis of AHK1/ATHK1 and cytokinin receptor histidine kinases in response to abscisic acid, drought, and salt stress in *Arabidopsis*. *Proceedings of the National Academy of Sciences*, 104(51), pp 20623–20628.

Vlad, D., Rappaport, F., Simon, M. & Loudet, O. (2010). Gene transposition causing natural variation for growth in *Arabidopsis thaliana*. *PLoS Genetics* [online], 6(5). Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2869320/. [Accessed 2018-10-29].

Wang, Y. H. (2008). How effective is T-DNA insertional mutagenesis in *Arabidopsis*?. *Journal of Biochemical Technology*, 1(1), pp 11-20.

Weinreich, D. M., Watson, R. A., Chao, L. & Harrison, R. (2005). Perspective:sign epistasis and genetic constraint on evolutionary trajectories. *Evolution*, 59(6), pp 1165–1174.

Wohlbach, D. J., Quirino, B. F. & Sussman, M. R. (2008). Analysis of the *Arabidopsis* histidine kinase ATHK1 reveals a connection between vegetative osmotic stress sensing and seed maturation. *The Plant Cell*, 20(4), pp 1101–1117.

Wu, C.-I. & Ting, C.-T. (2004). Genes and speciation. *Nature Reviews Genetics*, 5(2), pp 114–122.

Zandalinas, S. I., Balfagón, D., Arbona, V., Gómez-Cadenas, A., Inupakutika, M. A. & Mittler, R. (2016). ABA is required for the accumulation of APX1 and MBF1c during a combination of water deficit and heat stress. *Journal of Experimental Botany*, 67(18), pp 5381–5390.

7 Acknowledgements

I would like to thank PI Rubén Alcázar for hosting me in his lab, giving excellent guidance and the opportunity to explore this fascinating topic of plant biology. I also want to show appreciation for all the practical advice and ideas I got from Kostadin Atanasov and the respective members of the Alcázar lab group and everyone at the Department of Biology, Healthcare and Environment. University of Barcelona; Faculty of Pharmacy. I also extend great gratitude to Christina Dixelius for all the motivation and help needed to make this thesis a reality, German Martinez for making time to give much appreciated and valuable feedback in a hectic time and Aybars Koc for opposing my thesis defence.
Finally, a special thanks to all my family, friends and especially Anna-Sofie for unconditional support and presence through the process of completing the thesis.

8 Supplementary information

Supplementary Table 1. Single nucleotide polymorphisms among 22 Gw accessions outisde exons.

SNP/location[a]	Shared by
C/T 1161	All except C3-Spot-4-1, C3-Spot-4-10, C3-Spot-5-9 and Teatralna1
C/A1316	All
G/A1999	All except C3-Spot-4-1, C3-Spot-4-10, C3-Spot-5-9 and Teatralna1

[a]Polymorphism location number of bases downstream from ATG

Supplementary Table 2. Significance of Tukey comparisons of plant weight and leaf area means explained by genotype/treatment (confidence level = 0.95)[a].

Genotype/ Treatment	*SRF3*-OX (0)	*SRF3*-OX (50)	*SRF3*-OX (100)	*SRF3*-OX (150)	*srf3-2* (0)	*srf3-2* (50)	*srf3-2* (100)	*srf3-2* (150)	*srf3-3* (0)	*srf3-3* (50)	*srf3-3* (100)	*srf3-3* (150)	*srf3-4* (0)
SRF3-OX (0)	N/A [c]												
SRF3-OX (50)	-	N/A											
SRF3-OX (100)	***		NA										
SRF3-OX (150)	***			NA									
srf3-2 (0)		***	***	***	NA								
srf3-2 (50)					***	NA							
srf3-2 (100)	*				***		NA						
srf3-2 (150)	***				***			N/A					
srf3-3 (0)		***	***	***	***	***	***	***	N/A				
srf3-3 (50)					***				***	NA			
srf3-3 (100)	***				***				***		N/A		
srf3-3 (150)	**				***				***			N/A	
srf3-4 (0)	***	***	***	*		***	***	***	***	***	***	***	NA
srf3-4 (50)					***				***				***
srf3-4 (100)					***				***		***		***
srf3-4 (150)	*				***				***				***
Col-0 (0)		***	***	***		***	***	***	***	***	***	***	***
Col-0 (50)					***				***				***
Col-0 (100)	***				***				***				***
Col-0 (150)	***				***				***				***
SRF3^{Ler}:GFP (0)		*	***	***	*	**	***	***			***	**	***
SRF3^{Ler}:GFP (50)					***	**			***				***
SRF3^{Ler}:GFP (100)	**				***				***				***
SRF3^{Ler}:GFP (150)	***				***				***				***
L*er* (0)			***	*	***		**	*	*		**		***
L*er* (50)	***				***				***				***
L*er* (100)	***				***				***				***
L*er* (150)	***				***				***	**			***

[a] Asteriscs indicate significance.

[b] Numbers within each parenthesis indicate concentration of Mannitol in the growth medium (mM).

[c] Significance codes: < 0.001 = ***, < 0.01 = **, < 0.05 = *, not significant = -, NA = Not applicable

Supplementary Table 2 cont. Significance of Tukey comparisons of plant weight and leaf area means explained by genotype/treatment (confidence level = 0.95)[a].

Genotype/ Treatment	*srf3-4* (50)	*srf3-4* (100)	*srf3-4* (150)	Col-0 (0)	Col-0 (50)	Col-0 (100)	Col-0 (150)	*SRF3*^{Ler}: GFP (0)	*SRF3*^{Ler}: GFP (50)	*SRF3*^{Ler}: GFP (100)	*SRF3*^{Ler}: GFP (150)	L*er* (0)	L*er* (50)	L*er* (100)	L*er* (150)
SRF3-OX (0)															
SRF3-OX (50)															
SRF3-OX (100)															
SRF3-OX (150)															
srf3-2 (0)															
srf3-2 (50)															
srf3-2 (100)															
srf3-2 (150)															
srf3-3 (0)															
srf3-3 (50)															
srf3-3 (100)															
srf3-3 (150)															
srf3-4 (0)															
srf3-4 (50)	NA														
srf3-4 (100)		NA													
*srf3-4 (*150)			NA												
Col-0 (0)	***	***	***	NA											
Col-0 (50)				***	N/A										
Col-0 (100)				***		N/A									
Col-0 (150)				***			NA								
SRF3^{Ler}:GFP (0)	**		**	***	***	***	***	NA							
SRF3^{Ler}:GFP (50)				***					NA						
SRF3^{Ler}:GFP (100)								***		N/A					
SRF3^{Ler}:GFP (150)								***			N/A				
L*er* (0)						*	***				***	NA			
L*er* (50)				***				***				*	NA		
L*er* (100)								***				*		NA	
L*er* (150)				***				***				***			NA

[a] Asteriscs indicate significance.
[b] Numbers within each parenthesis indicate concentration of Mannitol in the growth medium (mM).
[c] Significance codes: < 0.001 = ***, < 0.01 = **, < 0.05 = *, not significant = -, NA = Not applicable

Supplementary Table 3. *Aarabidopsis thaliana* Seed Sterilization Protocol.

1. Put a small amount of seeds into 2ml Eppendorf tubes, wash seeds shortly with distilled water.
2. Fill the tube with commercial bleach (30% sodium hypochlorite).
3. Add 0,5% (v/v) Triton X-100.
4. Incubate in an orbital or rotary shaker for 10 min at 160 rpm.
5. Wash with ddH2O and decant the fluid. Repeat 4 times.
6. Plate the seeds in sterile growth plates with nutrient medium or suspend in 0,1% agarose. Stratify for up to 7 days.

Supplementary Table 4. Plant growth medium (1 dm^3).

Ingredient	Ammount
Murashige-Skoog medium (MS)	4,8 g
2-(N-morpholino)ethanesulfonic acid (MES)	0,5 g
Saccharose	3,0 g
Distilled water (dH_2O)	1,0 dm^3
Plant agar	Variable
Mix MS, MES and saccharose in 90% final volume dH_2O on a magnetic stirrer.	
Adjust pH to 5,6 - 7,8 with potassium hydroxide (KOH)	
Rinse down potential edge bound nutrients with the rest of the ddH_2O and aliquot to 0,5 l flasks.	
Add plant agar. 8 g × l^{-1} for a concentration of 0,8%.	
Autoclave 30 min at 120 ˚C.	

Supplementary Table 5. PCR primers used to amplify the Gw accession *SRF3* locus and sequence the region of interest.

	Sequence (5'->3')	Length (bp)	Tm (°C)	GC (%)	Complementarity	
					Self	3'
Fwd[b]	CAGATTCGTGGACATCGTT	19	55,35	47,37	3	2
Rev[b]	ATATCCGTAAGCGGCTAGC	19	56,38	53,63	6	6
Seq1[c]	CCCATACCGGACAAATTAC	19	53,12	47,37	4	0
Seq2[c]	AACCAGAATTATGAAGAC-GAGG	22	55,90	40,19	6	0

[a] All primers were blasted with NCBI primer blast
[b] PCR primers used to amplify *SRF3*
[c] Seq1and Seq2 represents two primers used for sequencing different areas of the gene of interest

Supplementary Table 6. Optimized PCR reagent contents used to amplify SRF3.

Reagent	Ammount (µl)
ddH_2O	29,75
PCR buffer (10×)	5
$MgCl_2$ (1,5 mM)	5
dNTP (10 mM)	5
Forward primer (100 µM)	0,5
Reverse complementary primer (100 µM)	0,5
LA Taq	0,25
Template DNA	4
Total volume	50

Supplementary Table 7. DNA extraction buffer used in the transcriptional target experiment (for two-step PCR genotyping).

Ingredient	Working concentration
Tris HCl pH 8,0	200 mM
NaCl	250 mM
EDTA	25 mM
SDS[a]	0,5% (v/v)

[a] Add SDS last in order to avoid foaming

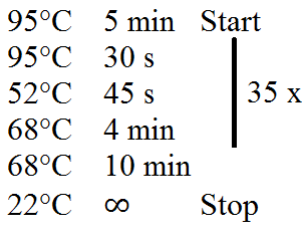
Supplementary Table 8. Optimized PCR reagent contents used for SRF3 T-DNA insertion mutant two-step PCR genotyping.

Reagent	Ammount (µl)
ddH_2O	15,6
PCR buffer (10×), incl. $MgCl_2$ (1,5 mM)	2
dNTP (1,25 mM)	2
Forward primer (100 µM)	0,1
Reverse complementary primer (100 µM)	0,1
LA Taq	0,2
Template DNA	2
Total volume	20

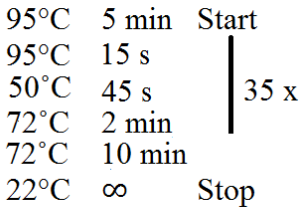
Supplementary Table 9. PCR primers used for SRF3 T-DNA insertion mutant analysis (two-step PCR genotyping).

					Complementarity	
Primer	Sequence (5'->3')	Length (bp)	Tm (°C)	GC (%)	Self	3'
$LP_{SRF3\text{-}OX}$[b]	TCCACCGAAATTTCAAGTCTG	21	56,46	42,86	8	2
$RP_{SRF3\text{-}OX}$[b]	ACTAATCCGACTCGAGACCG	20	58,70	55,00	6	6
$LP_{srf3\text{-}2}$[b]	AGCCTCTCAAGGTCATGTAAG	21	57,10	47,62	4	2
$RP_{srf3\text{-}2}$[b]	CCCATACCGGACAAATTAC	19	53,12	47,37	4	0
$LP_{srf3\text{-}3}$[b]	AGCCTCTCAAGGTCATGTAAG	21	57,10	47,62	4	2
$RP_{srf3\text{-}3}$[b]	AAGGATCGAGCTCTGAGAA	19	55,12	47,37	6	3
$LP_{srf3\text{-}4}$[b]	AAAAATTCGGCTGGAATTGTC	21	60.31	38.10	4	0
$RP_{srf3\text{-}4}$[b]	CAGAAGAGAGCGTCATGGTTC	21	60.00	52.38	4	2
LB1,3[c]	ATTTTGCCGATTTCGGAAC	19	54,69	42,11	6	4

[a] All primers were blasted with NCBI primer blast
[b] PCR primers used to amplify *SRF3*
[c] LB1,3 amplifies the left T-DNA border in conjunction with an *SRF3* LP primer



Supplementary Figure 1. PCR programme used to amplify the Gw accession *SRF3* locus from gDNA. 95 ˚C = dsDNA denaturation, 52 ˚C = primer annealing, 68 ˚C = ssDNA strand elongation, 22 ˚C = standard normal environment.



Supplementary Figure 2. PCR programme used to analyse *SRF3* T-DNA insertion mutants. 95 ˚C = dsDNA denaturation, 50 ˚C = primer annealing, 72 ˚C = ssDNA strand elongation, 22 ˚C = standard normal environment.

9 Opposition Summary

A few points were discussed after my presentation regarding the hypothesis of co-evolution between two epistatic loci and regarding my work on the functions of the SRF3 protein. For the question of co-evolution between *RPP1*^{Ler}-like alleles and *SRF3* alleles in the Gorzów population, the opposition suggested to include more gene variants from other *A. thaliana* populations and ecotypes. One could search online databases to find more alleles that would give a broader perspective of the evolution of the hybrid incompatibility. One could also sequence the whole *SRF3* locus instead of only the region that is the most SNP dense in other accessions. Also, if I would have sequenced the *RPP1*-like locus in the same individuals I could get a deeper understanding of how the alleles segregate. A fourth point that would help investigate a possible co-evolution would be to sequence more individuals of the population. In my project I only sequenced 22 individuals where approximately half had the *RPP1*^{Ler}-like alleles present. It could be that by limiting the sample number, I missed several low frequency alleles that segregate in the population, which could be a sign of negative frequency dependent selection. During the sequence analysis I also looked for amino acid substitutions. Since I only had DNA to work with and predicted intron/exon barriers, the analysis could be improved by sequencing mRNA or exome sequencing instead to get rid of most of the introns. This would improve the phylogenetic gene tree (*Figure 4*) as well by eliminating possibly unimportant intronic SNPs.

For the second hypothesis where I investigated a possible role of SRF3 in osmotic stress tolerance signalling we discussed ways to improve the experiment. By adding more phenotypic traits typical for water stress, like including measurements of root length, for example one would gain information from the growth of the whole plant, not just biomass accumulation and leaf area. For this, one would need to limit the number of plants in each *in vitro* growth dish to avoid entanglement of roots and competition for space. One could also measure the gene activity in stressed wild type plants at different time points to confirm that *SRF3* indeed upregulates during

osmotic stress. A third improvement would be to confirm the status of the knockdown and overexpressing lines. In order to know if the data represented complete knockdown lines or lines that for example still had 70% of its activity, I would have needed to extract RNA from all individual plants. Transcriptional activity was never controlled for during my stay in Alcázars lab due to time and cost constraints. A fourth idea was put forth that since *SRF3* has 7 closely related genes in the family, one could expect some functional redundancy. A double or triple mutant could reveal a phenotype that would not visible if the stress tolerance is governed by multiple genes with small additive effects.

No 169
Uppsala 2018

Department of Plant Biology
SLU
Box 7080
75007 Uppsala, Sweden