

Lisette van Noord

Msc. Thesis Plant Sciences Plant breeding and genetic resources Wageningen university Supervisors:

dr. C.W.B. Bachem dr. J.A. Abelenda Vila

Examiners:

dr. C.W.B. Bachem dr. ir. A.B. Bonnema



Novel Insights on the effect of *StCDF1* and the *StCDF1* antisense lncRNA on drought tolerance in *Solanum tuberosum*

Analysing differences in drought stress response of both wild type and mutated StCDF1 alleles

Lisette van Noord, BSc

Reg. Nr: 911123607030

July 2017

MSc Thesis

Plant Sciences Group Laboratory of Plant Breeding Breeding for Growth and Development, Potato Group

WAGENINGEN UNIVERSITY

Supervisors:

dr. C.W.B. Bachem dr. J.A. Abelenda Vila

Examiners:

dr. C.W.B. Bachem dr. ir. A.B. Bonnema

Acknowledgements

While working on this thesis, there were a lot of people that helped me, and without them this thesis would not have been possible. Therefore, I would like to take the opportunity to thank these people. First of all, I would like to thank Christian Bachem, for the opportunity to work on this subject and his guidance during my thesis. Furthermore, I would like to thank José Abelenda for his daily supervision during this thesis, and the helpful discussions that we had. Furthermore, I would like to thank Marian Oortwijn for her help in the lab, especially in the beginning. She was always very patient, even if I did not understand something and she had to explain it multiple times. I would also like to thank Annelies Loonen for providing us with the vectors necessary for the CRISPR-Cas9 experiment. Furthermore, I would like to thank the other people in the lab that were always very helpful whenever I was looking for something or needed assistant. And finally, I would like to thank the other students. Sharing the problems that we had during our research and discussing possible solutions really helped the process of writing a thesis.

Table of Contents

Ac	knowled	lgements	
Lis	t of figu	res	3
Ab	breviati	ons	4
Ab	stract		5
1	Introd	luction	6
2	Mater	ial and methods	12
:	2.1	Available plant lines	
:	2.2	Experiment 1: non-drought comparison	13
:	2.3	Experiment 2: drought comparison	14
:	2.4	Experiment 3: inactivation of the promotor region	17
3	Result	S	20
:	3.1	Experiment 1 and 2	20
	3.1.1	Phenotypic response to drought	20
	3.1.2	Weight and number of tubers	22
	3.1.3	Water content	24
	3.1.4	Stomatal conductance	
	3.1.5	Chlorophyll	
	3.1.6	Expression profile of StCDF1 and StSP6A	
	3.1.7	Expression of SCAL1	
	3.2	Experiment 3	28
4	Discus	sion	
•	4.1	Effect of StCDF1 on drought stress tolerance	
•	4.2	StCDF1 and StSP6A gene expression profile	
•	4.3	Interaction between StCDF1 and SCAL1	
,	4.4	The effect of SCAL1 on drought tolerance	31
5	Concl	usion	33
6	Refere	ences	34
Ар	pendix /	A: protocols	38
Ар	pendix I	3: additional figures	45
Ар	pendix (C: details on the vector components	46
Αp	pendix [D: glycerol stocks	47

List of figures

Figure 1: Structure of the three different alleles of the StCDF1 gene	
Figure 2: Simplified visualization of the tuberisation pathway	
Figure 3: Phylogenetic tree of CDF genes in Arabidopsis, tomato and potato	
Figure 4: potato plants after imposing drought stress.	1
Figure 5: schematically representation of the StCDF1 gene	1
Figure 6: drought stressed plants	2
Figure 7: comparing the weight and number of tubers from drought and non-drought stressed plants	2
Figure 8: Water content of the plants and tubers after drought stress	2
Figure 9: percentage of dead and alive plants based on conductivity	2
Figure 10: percentage of dead and alive plants based on chlorophyll	2
Figure 11: relative gene expression	2 ⁻
Figure 12 relative expression of CDF1 and SCAL1	2
Figure 13: The vector created in experiment 3.	2
Figure 14: proposed model on the relationship between SCAL1. StCDF1 and drought tolerance	3:

Abbreviations

AA Amino acids
Bp Base pair

CRISPR associated protein 9

CDF Cycling Dof Factor
CDS Coding DNA Sequence

ChIP Chromatin Immuno-Precipitation

CO Constans

COLDAIR Cold Assisted Intronic Non-Coding RNA

CRISPR Clustered Regularly Interspaced Short Palindromic Repeats

DOF DNA binding with one finger

FKF1 Flavin-Binding Kelch Repeat F-Box

FLC Flowering Locus C
FT Flowering locus T

GI Gigantea

GMO Genetically Modified Organism

IQR interquartile range

LD Long days

lincRNA long intergenic noncoding RNA

IncRNA long non-coding RNA

NAT natural antisense transcripts

PHYB phytochome B

PRC2 Polycomb Repressive Complex 2
RGEN RNA guided endonucleases

RNAi RNA interference

SD Short days

SEM Standard Error of the Mean

sgRNA single guide RNA WC Water content WT Wild type

Abstract

Potato (*Solanum tuberosum*) originates from the Andes region in South America, where it will only form tubers during short-day conditions that prevail at the equator. One of the genes responsible for short-day dependency is the *StCDF1* gene. Besides the wild type (WT) allele (*StCDF1.1*) that regulates tuberisation under short day conditions, there are two mutated alleles which allow the plants to tuberize under long day conditions; *StCDF1.2* with a 7bp insertion and *StCDF1.3* with a 865bp insertion. In both cases (*StCDF1.2* and *StCDF1.3*) the mutated protein lacks the regulatory domain increasing stability throughout the day. In addition to the normal sense coding transcript of *StCDF1*, we have detected a lncRNA transcribed in a reverse antisense orientation covering the second exon of the *StCDF1* coding region. However, in *StCDF1.3*, the lncRNA is displaced by 865bp due to the large insertion and therefor unlikely to be effective.

Since *StCDF1* RNAi knockdown plants are highly tolerant to drought, we have linked the *StCDF1* locus to the drought stress response. Furthermore, due to the high levels of lncRNA expression in drought stressed *StCDF1* RNAi lines, we hypothesise that the lncRNA of *StCDF1* may be closely involved in this response. In addition to *StCDF1* RNAi knockown plants lines, various *S. tuberosum* genotypes with different *StCDF1* allele combinations were available at the start of this project. The primary aim of this research is to show the effect of drought stress on these lines to better map the effect of *StCDF1* and its lncRNA on drought tolerance. This was achieved by measuring the stomatal conductance, chlorophyll content, tuber count and weight of the tubers and plants after the imposition of drought stress. Furthermore, gene expression studies were carried out in the different genotypes to determine the levels of transcription of *StCDF1*, its lncRNA and *StSP6A*; the output of the photoperiod pathway. The results show that *StCDF1.3* is likely to be dominant over *StCDF1.1* during drought stress conditions in heterozygous plants while in the non-drought conditions. Furthermore, it appears that down regulation of the *StCDF1* up regulates its lncRNA.

Before this research, there were no lines available that contained the WT protein of StCDF1 in the absence of the IncRNA. Such a plant line can be used to analyse the effect of the WT CDF1 protein without interference of the IncRNA. Therefore, the secondary aim of this research was to engineer a *S. tuberosum* line using CRISPR-CAS9 knockout of the IncRNA. We were successful in creating an appropriate vector and transferring this vector into *Agrobacterium tumefaciens*.

Key words: Solanum tuberosum, Potato, CDF1, IncRNA, SP6A, drought tolerance, tuberisation, CRISPR-Cas9

1 Introduction

The modern crop potato, *Solanum tuberosum* L. is the most important food crop worldwide, after the cereals (FAOSTAT, 2017). Potato is grown in over 130 countries and the largest producing country is China, accounting for 22% of the world production. Potato is thought to originate from the equatorial Andes region in South America, where it will only form tubers under short-day conditions. This means that in northern latitudes, native South American genotypes such as Andigena (*Solanum tuberosum* L. subsp. *andigena* Hawkes) are not able to create tubers in spring and summer due to the local long-day conditions (Kloosterman, et al., 2013). To achieve tuberisation, these potato plants need night periods that are longer than fourteen hours. This phenomenon is also apparent in strict short-day flowering plants such as rice (Navarro et al., 2011). Tuberisation is thus likely one of the first traits that potato breeders in northern countries need to select for (Salaman, 1989).

Breeding in potato is difficult for any trait due to its tetrasomic inheritance, incompatibility barriers, and their high level of heterozygosity (Anithakumari, 2011). Furthermore, potatoes are generally vegetatively propagated and autotetraploid, limiting their genetic gain. Due to the vegetative propagation, at least 5 years are needed to collect enough seed tubers for reliable trials. These characteristics of potato also make it hard to fixate favourable alleles (Vos, 2016; uitdewilligen 2012). Therefore, the time needed to introduce a new variety of potatoes takes an average of 10-12 years after making the first crosses (Bradshaw, 2009). Due to these difficulties with potato, genetic studies are often executed on diploid genotypes and populations (Kloosterman et al., 2012). CRISPR-Cas9 technology might be very helpful for potato research and future breeding to (at least partially) overcome these problems.

CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats)-Cas (CRISPR-associated proteins) is an adaptive RNA mediated defence mechanism in many prokaryotes that offers protection to invading viruses (Wiedenheft et al, 2012; Barrangou et al, 2007). Historically, research focused mainly on the conceptual understanding of the mechanism behind CRISPR-Cas9. However, current plant researchers are beginning to use this mechanism for gene editing to improve economically important crops (Schaeffer & Nakata, 2015).

The CRISPR-Cas9 genome editing technique can provide a specific gene knockout (Wang et al., 2015). Advantages of using the CRISPR-Cas9 technique include its versatility in genome editing, the ability to cleave methylated loci and the ease of use (Belhai et al., 2013; Lozano-Juste and Cutler, 2014). Furthermore, customizing TALENs and ZFNs (other programmable nucleases) is labour intensive and time consuming, while CRISPR RNA guided endonucleases (RGENs) only need to have their RNA component replaced (Woo et al., 2015). This is because CRISPR consist of two different parts: the single guide RNA (sgRNA), and the non-specific CRISPR-associated endonuclease (Cas9) which together form a riboprotein complex. The sgRNA is responsible for the binding to the targeted DNA sequence. By changing the nucleotide sequence of the sgRNA, the genomic target changes. It is important that the sgRNA is site specific and does not occur anywhere else in the genome. The Cas9 enzyme is non-specific.

To successfully implement the CRISPR-Cas9 gene editing technology, crop specific expression constructs are needed. In potato, targeted mutagenesis initiated by the CRISPR-Cas9 technology has already been proven to be possible by Wang et al., 2015. They showed that CRISPR-Cas9 provides an efficient technique to study uncharacterized genes in potato. Due to the ease of use of CRISPR-Cas9, it creates the possibility to unravel the mechanism behind different disease and stress responses, such as drought and heat stress.

At the moment, it is essential to get a better understanding of the mechanisms potatoes employ to respond to both drought and heat stress. As from 1900 to 2100, the average temperature is estimated to rise between 1.4 and 5.8°C (Houghton et al., 2001), research into heat stress is becoming ever more important in providing a sustainable global crop yield. Furthermore,

in recent decades, the frequency and intensity of drought periods increased in Asia and Africa (Hijmans, 2003), further accentuating the importance of drought stress research in important yield crops.

Although in some countries the potato yield will increase due to the global rise in temperature, others will have a severe decrease. Overall, the global potato yield is estimated to decrease between 5 and 11% between 2010 and 2039, which worsens to an expected decrease between 9 and 18% from 2040 to 2069. From all the potato growing regions, the (sub)tropics will suffer from the largest decline in potato yield (Hijmans, 2003).

Potatoes are highly susceptible to drought stress, since potatoes have a very shallow root system (Iwama and Yamaguchi, 1996), with a system depth ranging from 0.5- 1.0m and about 85% of the roots are concentrated in the upper 0.3m layer of soil. In addition, potatoes have a very low capacity of recuperation after a period of water shortage (Monneveux et al., 2014). The severity of the effect of drought on potato production is dependent of the physiological timing, duration and severity of the stress (Jefferies, 1995). Dwayer and Boisvert (1990) concluded that drought stressed potato plants can senescence early, when compared to non-drought stressed plants. Furthermore, early season drought stress in potatoes was found to reduce the total number of tubers (Dwyer and Boisvert 1990; Haverkort et al. 1990; Mackerron and Jefferies 1988), the growth of the tubers and the final yield (Cavagnaro et al., 1971).

An important signalling pathway for the timing of tuberisation and flowering is the flowering pathway. In the model species, *Arabidopsis thaliana* FLOWERING LOCUS T (FT) plays a central role in the onset of flowering (Zeevaart, 2008). FT is activated by CONSTANS (CO), and CO is under the control of CYCLING DOF FACTORS (CDFs) and the circadian clock. Once CDF proteins are produced they can be marked for ubiquination and subsequent degradation by binding to the GIGANTEA (GI) and FLAVIN-BINDING KELCH REPEAT F-BOX (FKF1) protein complex on their third domain (Sawa et al., 2007). Under short day conditions (SD), CO is suppressed by CDFs making CO unable to induce FT expression. In contrast, during long days (LD) the stability and expression of CO is increased, promoting the expression of FT (Song et al., 2012).

The increase in daylight during LD conditions compared to SD conditions is perceived by the plant in the leaves. In the companion cells of the phloem, CO and FT are expressed. After translation the FT proteins can migrate from the leaves to the apex where it induces flowering (Corbesier et al., 2007).

The interaction between FT's and CO is highly conserved in most angiosperms (Song et al., 2015). Therefore, it is no surprise that different orthologues of FT have been discovered, including HD3A in rice, the SELF-PRUNING genes in potato and tomato, and BvFT1 in the biannual sugar beet (Abelenda et al., 2014). In potato, the *StSP6A* gene has a unique function in triggering tuberisation rather than flowering (Navarro et al., 2011).

In potato, CYCLING DOF FACTOR1 (StCDF1) binds to the CO promoter, suppressing the expression of CO. There are three different CONSTANS homologues, *StCO1*, *StCO2* and *StCO3*, likely under *StCDF1* regulation, in a tandem repeat on chromosome 2. *StCO1* and *StCO2* are important for the day length control of tuberisation, since they repress tuber formation under non-inductive LD (Navarro, Cruz-Oró, & Prat, 2015). *StCO* is known to induce *StSP5G*, and *StSP5G* in its turn represses *StSP6A* (Abelenda et al., 2016). In potato, using chromatin immuno-precipitation (ChIP), it was shown that *StCDF1* has a stronger influence on *StCO1* than on *StCO2* (Abelenda, Cruz-Oró, Franco-Zorrilla, & Prat, 2016) .

Just as with Arabidopsis, the wild type *StCDF1* allele (*StCDF1.1*) can be marked for ubiquitination by the GI-FKF1 complex, by binding to the third domain of *StCDF1*. However, this is not the case for the mutated *StCDF1* alleles; *StCDF1.2* and *StCDF1.3*. *StCDF1.2* has a 7-base pair (bp) insertion, causing a frame-shift resulting in a premature stop codon, altering the third domain. In the *StCDF1.3* allele, there is a large insertion of 865bp, causing a fusion protein with 22 additional amino acids (AAs). These

22 additional AAs replace the 52 carboxy-terminal residues, the third domain found in the WT allele. Without this third domain, ubiquitination of *StCDF1* is inhibited. Therefore, the StCDF1 protein remains more stable in the mutated alleles (Kloosterman et al., 2013; Navarro et al., 2011).

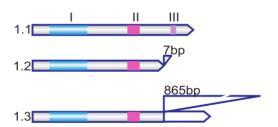


Figure 1: Structure of the three different alleles of the StCDF1 gene

(After Kloosterman et al., 2013). This cartoon shows the three different alleles of *StCDF1* with the three different domains (I, II and III). The figure shows that *StCDF1.1* has the third domain, while this is missing in *StCDF1.2* and *StCDF1.3* due to an insertion resulting in a truncated and fusion protein (represented by the non-shaded part) respectively.

Because StCDF1.2 and StCDF1.3 miss the StFKF1 binding domain, StCO is repressed during LD. Due to this repression of StCO, StSP5G no longer represses StSP6A. Therefore, the more stable StCDF1 protein encoded by the *StCDF1*.2 and *StCDF1*.3 genes, results in an early maturation and tuberisation phenotype and an independence of day length for tuberisation, enabling tuberisation under LD (Kloosterman et al., 2013; Navarro et al., 2011).

Another important factor in this pathway is phytochrome B (PHYB). It seems that PHYB can stabilise StCO, inhibiting tuber forming in LD (Navarro et al., 2015; Plantenga et al., 2016). PHYB is activated by red-light, and inactivated by far-red (Casal, 2013). An overview of this pathway under both short day and long day conditions is given in Figure 2.

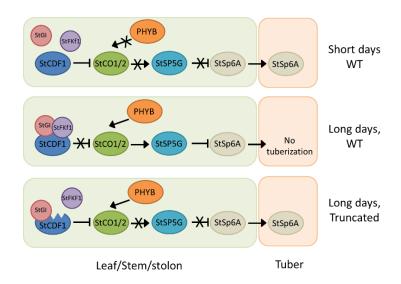


Figure 2: Simplified visualization of the tuberisation pathway

adjusted from Kloosterman et al., 2013 and Navarro et al., 2015. Under short days, StGI and StFKF1 do not bind to StCDF1, resulting in transport of StSP6A from the stolon to the tubers. During long days, StGI and StFKF1 bind to StCDF1, causing it to be ubiquinated. Therefore, there are no tubers formed. However, if the StCDF1 allele is truncated, as in the case of StCDF1.2 and StCDF1.3, StFKF1 does not bind to StCDF1, resulting in downregulation of StCO1/2, and thus tuber forming. PHYB does stabilise StCO1/2 but cannot safe the expression of StCO1/2.

StCDF1 is thus a key gene in the onset of tuberisation in potato, and mutations in this allele have a large influence on the SD dependency of the plant for tuber forming. The CDF gene family is present in multiple plant species. In Figure 3 the phylogenetic tree of different CDF genes from potato, Arabidopsis and tomato are shown. It shows that potato CDF genes StCDF1.1, StCDF2, StCDF3 and StCDF4 are orthologues to SICDF2, SICDF4, SICDF5, SICDF1 and SICDF3 of tomato respectively. In Arabidopsis, the CDF-genes are highly induced by environmental stresses like drought, extreme temperatures and abscisic acid treatments. For example, when the tomato SICDF1 or SICDF3 is overexpressed in Arabidopsis the plant will have an enhanced drought and salt tolerance. Overexpression of SICDF3 also upregulates a group of genes that encode for LEAs HSPs and DNAJ proteins. These proteins are involved in osmotic regulation, protein folding, autophagy and protection of cellular structures under abiotic stresses. Furthermore, SICDF3 is associated with protection against cellular damage caused by osmotic and low temperature stress (Corrales et al., 2017). The overexpression of SICDF3 also leads to late flowering by changing the expression of SICO. Thus, SICDF3 is likely to play a key role in abiotic stress tolerance, as well as a role in flowering time control (Corrales et al., 2017). Furthermore, Nakamichi et al., 2016 supplied evidence in Arabidopsis that up regulation of CDF genes, which result in a down regulation of CO and FT, may increase cold, drought and salt tolerance. Thus, in general it seems that CDF genes can play a role in different stress responses.

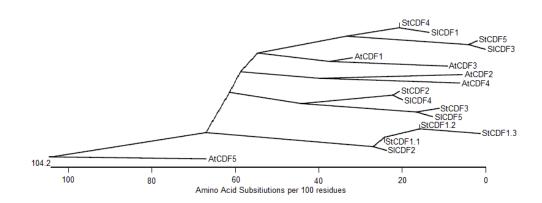


Figure 3: Phylogenetic tree of CDF genes in Arabidopsis, tomato and potato.

The figure shows that potato CDF genes CDF1.1, CDF2, CDF 3 and CDF4 are orthologs to CDF2, CDF4, CDF5, CDF1 and CDF3 of tomato respectively.

To regulate the genes in the CDF gene family, or any other gene, the plant can use a host of different regulatory mechanisms. A newly studied gene regulation tool is the long non-coding RNA (lncRNA). LncRNAs are mRNA-like transcripts larger than 200bp. LncRNAs control different biological processes by controlling expression as specific as single genes, and up to entire chromosomes (Engreitz et al., 2013; Mohammad, Mondal, & Kanduri, 2009; X. Wang et al., 2016). Previous research has shown that lncRNAs are important in the regulation of plant development and stress responses. Several lncRNAs were also found to be involved in plant reproductive developments and the response mechanism to pathogenic invasion (Zhang & Chen, 2013). However, since the field of lncRNA research is new, there still is a lot unknown about their function (Leone & Santoro, 2016).

LncRNAs are poorly conserved at the sequence level, when compared to smaller non coding RNAs, such as siRNAs, miRNAs and piRNA (Mohammad et al., 2009; K. C. Wang & Chang, 2011). In general, IncRNAs have a low expression and complex functions, making detection from computational data difficult. However, since most IncRNAs are polyadenylated, detection from transcriptome sequencing data is possible (X. Wang et al., 2016).

LncRNAs can be classified according to genomic location and context, into three different types; long intergenic noncoding RNAs (lincRNA), long intronic noncoding RNAs and natural antisense transcripts (NATs) (Kwenda et al., 2016; Zhang & Chen, 2013). The lincRNAs are located between genes and do not have any overlap with other loci (Bai, Dai, Harrison, & Chen, 2015). Sometimes, lincRNAs are located close to coding DNA sequence (CDS) of genes for proteins and are therefore occasionally referred to as adjacent-lncRNAs. They are usually associated with the CDS promoter and terminator regions. At times, the lincRNA on one strand can partially overlap with the CDS on the complementary strand. This type of lincRNA is therefore called antisense-lncRNA (Kwenda et al., 2016)

In contrast to lincRNA, long intronic noncoding RNA extends along the intronic region of a locus, and NATs overlap with exon(s) of other loci on the complement strand (Bai et al., 2015). NATs, like pseudogenes, can sometimes hybridise with overlapping genes and generate endo-siRNA. Furthermore, there are multiple examples where NATs modulate the splicing patterns of their overlapping genes (Wilusz, Sunwoo, & Spector, 2009).

It has been shown that for some specific cases, the transcription of ncRNA is sufficient to either negatively or positively influence the expression of nearby genes. In addition, the IncRNA itself can have key regulatory functions which were previously only subscribed to proteins (Wilusz et al., 2009). There are four distinct mechanisms by which an IncRNA may influence the transcriptional control of gene regulation. Firstly, the IncRNA may act as a signalling molecules for the recruitment of the chromatin modifying machinery. The use of IncRNAs as signalling molecules is beneficial to the organism because it saves time needed to translate the RNA precursors into signal proteins (Bhattacharjee et al., 2013; K. C. Wang & Chang, 2011; Wilusz et al., 2009). An example of this mechanism can be found in A. thaliana. In A. thaliana, the FLOWERING LOCUS C (FLC) gene expression is repressed by a group of IncRNAs called Cold Assisted Intronic Non-Coding RNA (COLDAIR). COLDAIR can interact with PRC2 (POLYCOMB REPRESSIVE COMPLEX 2) complexes on the FLC locus, mediating chromatin modification leading to transcription repression of FLC (Bhattacharjee et al., 2013; K. C. Wang & Chang, 2011). Secondly, IncRNAs may act as pseudogenes, interfering with the assembly of the regulatory transcript machinery by binding to protein target without exerting additional function, leading to transcriptional repression (Bhattacharjee et al., 2013; K. C. Wang & Chang, 2011; Mohammad et al., 2009; Lee, 2012). Thirdly, IncRNAs may act together with effector molecules by guiding them to their target site. Fourthly, IncRNA may act as scaffolds for the recruitment of effectors that target a specific gene (Bhattacharjee et al., 2013). These mechanisms can work in cis or trans mode. Cis-actin RNAs are limited to their transcription site and affect one or several overlapping or nearby genes (Lee, 2012). For example, cis-actin RNA may bind to sense RNA of the gene, altering the splicing patterns or generating endo-siRNAs (Wilusz et al., 2009). Trans-actin RNAs can act on genes far away from the transcription site of the IncRNA, even on other chromosomes. The distinction between cis- and trans-acting RNAs sounds distinct but can be blurred (Lee, 2012). A clear example of this is the mammalian Xist RNA. Xist RNA generally acts in cis mode, but when Xist transgenes are introduced de novo the Xist RNA can move to other regions of the genome (Guttman & Rinn, 2012; Lee, 2012). Thus, there is an abundance of non-mutually exclusive mechanisms by which IncRNA can affect gene expression (K. C. Wang & Chang, 2011).

It is known that located on the complimentary stand of *StCDF1* there is a DNA region coding for a lncRNA. Much is still unknown about this specific lncRNA. It is known, however, that the lncRNA region largely overlaps with the protein coding region of *StCDF1*. Furthermore, we know that the lncRNA of the *StCDF1* orthologue in *Arabidospsis thaliana* shows an opposite expression profile during the day, when compared with the StCDF1 protein (Dr. Rossanna Henriques, personal communication). There are different mechanisms through which this lncRNA might interfere with the *StCDF1* gene or another gene. It is likely that the lncRNA of *StCDF1* is a cis-acting RNA if and has an effect on *StCDF1*, since the protein DNA partially overlaps with the lncRNA region. Based on current knowledge, it cannot be predicted if the *StCDF1* lncRNA also functions as a trans-acting RNA.

The large insertion in *StCDF1.3* likely disconnected the promotor region from the IncRNA transcription site. Therefore, we think that the *StCDF1.3* allele no longer produces the IncRNA. *StCDF1.2* only has a small insertion, which might not interfere with the transcription of the IncRNA. However, before this research, there was no prove of this theory yet.

Little is known about the *StCDF1* gene and its role in drought tolerance. However, unpublished data from Dr. José Abelenda on the effect of *StCDF1* knockdowns in potato have shown that knocking down the *StCDF1* gene increased drought tolerance and reduces stomatal conductance. Furthermore, the leaves had fewer but larger stomatal guard cells, which were partially open. This suggests that the guard cells of *StCDF1* knockdowns were capable of gas exchange but did not enable much water evaporation (personal communication).

This research from Dr. Abelenda also showed that there is a large difference in drought stress response between heterozygote *StCDF1.1/StCDF1.2* plants and heterozygote *StCDF1.1/StCDF1.3* plants, with the former being more drought tolerant (Figure 4a and 2b). This implies that a single *StCDF1.3* allele results in an increased sensitivity to drought stress, while a single *StCDF1.2* allele does not do so (or at least to a lesser extent).





Figure 4: potato plants after imposing drought stress.

Figure a shows plants from the CE630 line. These plants are heterozygote *StCDF1.1/StCDF1.3*. In figure b, CE605 plants are shown which are heterozygote *StCDF1.1/StCDF1.2*. There is a clear difference in the severity of the drought stress response, which might be related to the difference between the *StCDF1.2* and *StCDF1.3* alleles.

The cause of this difference is currently unknown, but it might be related to the presence or the absence of an antisense IncRNA on *StCDF1*. Therefore, this research focuses on further phenotyping the *StCDF1* gene and analysing the effects of the different *StCDF1* alleles and the IncRNA, with a main focus on drought tolerance. Since only little research has been done on its effects, this IncRNA did not have a name yet. In this report we will refer to this *StCDF1* Associated IncRNA as SCAL1. The formulated research questions are;

- 1. What are the effects of the different StCDF1 alleles on drought tolerance in potato?
- What is the expression profile of StCDF1 for StCDF1.1, StCDF1.2 and StCDF1.3?
- 3. What is the relation between SCAL1 and gene expression of StCDF1?
- 4. What is the effect of SCAL1 on drought tolerance?

2 Material and methods

In this research, we performed 3 different experiments. Experiment 1 was a baseline/pilot study to determine the phenotypic differences and expression of *StCDF1* and *StSP6A* between plants that were homozygote *StCDF1.1*, heterozygote *StCDF1.1/STCDF1.3* and homozygote *StCDF1.3* without drought stress. In experiment 2 we included more plant lines and determined the response to drought stress based on phenotypic differences and *StCDF1* and *StSP6A* expression. In experiment 3 the aim was to create a transgenic plant with a wild type *StCDF1* allele but without a functional SCAL1 RNA with the use of CRISPR/cas9 technology.

2.1 Available plant lines

At the start of this research, four diploid parental lines were available; C, E, RH and SH. Line SE and C have the *StCDF1.1* and *StCDF1.2* allele while parental line E and RH have the *StCDF1.1* and *StCDF1.3* alleles (see Table 1). The differences between these alleles are described in the introduction.

Table 1: parental lines available for this research

Parental line	Allele 1	Allele 2
SE	StCDF1.1	StCDF1.2
С	StCDF1.1	StCDF1.2
E	StCDF1.1	StCDF1.3
RH	StCDF1.1	StCDF1.3

By crossing parental lines RH and E, plant line 1, 2 and 3 were created (see Table 2). These plant lines are homozygote *StCDF1.1*, heterozygote *StCDF1.3/StCDF1.1* and homozygote *StCDF1.3* respectively. Unfortunately, we will not include crosses from parental lines SE and C due to time constrains and difficulties in differentiating between the heterozygote and homozygote (*StCDF1.1* and *StCDF1.2*) plants due to the small difference of 7bp between the two alleles. Besides type 1 till 3, we included four other non-GMO plant lines; CE3027, CE3130, Andigena and Desiree. CE3027 and CE3130 are descendants of a cross between C and E and are respectively homozygote *StCDF1.1* and heterozygote *StCDF1.2/StCDF1.3*. Andigena and Desiree are commonly used tetraploid potato lines.

Table 2: non-GMO lines lines analysed for drought tolerance and tuber formation.

number	name	ploidy	CDF1 alleles	SCAL1 expected
1	RHXE_homoWT	diploid	1.1/1.1	yes
2	RHXE_hetero	diploid	1.3/1.1	Yes
3	RHXE_homoCDF1.3	diploid	1.3/1.3	No
4	CE3027	diploid	1.1/1.1	Yes
5	CE3130	diploid	1.2/1.3	Yes
6	Andigena	tetraploid	1.1/1.1/1.2/1.2	Yes
7	Desiree	tetraploid	unknown	unknown

Besides these plant lines there were 7 GMO plant lines available; plant line 8 till 14 (Table 3). Line 8, 9 and 10 are Andigena plants that have obtained a construct with the suc2 promotor and GFP, enabling us to follow the activity of *StCDF1.1* (line 10) and *StCDF1.2* (line 9). The corresponding constructs are 35s::suc2-GFP, 35s::CDF1.2::suc2-GFP and 35s::CDF1.1::suc2-GFP respectively. Type 11,12 and 13 are CE3027 plants that have taken up the same constructs as plant line 8,9 and 10.

These lines (line 8 till 13) enable us to study the differences between a stable *StCDF1* (*StCDF1*.2) and the WT (*StCDF1*.1) in both tetraploid and diploid backgrounds. Due to the GFP we can also visualise the activity of *StCDF1* and analyse the protein activity on a Western blot. In plant type 14 and 15 *StCDF1* is silenced by RNA interference (RNAi).

Table 3: GMO lines analysed for drought tolerance and tuber formation.

The "orignal genotype" is the genotype of the plant befor genetic modification. The "allele expressed" gives the allelic expression expected from the plant after genetic modification due to either the original genotype, or the plasmid insertion. All lines are expected to have a copie of SCAL1

			Original	Expected Allele
number	name	ploidy	Genotype	expressed
8	and_suc2_GFP	tetraploid	1.1/1.1/1.2/1.2	1.1
9	and_suc2_CDF1.2_GFP	tetraploid	1.1/1.1/1.2/1.2	1.2
10	and_suc2_CDF1.1_GFP	tetraploid	1.1/1.1	1.1
11	3027_suc2_GFP	diploid	1.1/1.1	1.1
12	3027_suc2_CDF1.2_GFP	diploid	1.1/1.1	1.2
13	3027_suc2_CDF1.1_GFP	diploid	1.1/1.1	1.1
14	CE3130_CDF1_RNAi	diploid	1.2/1.3	None
15	CE3027_CDF1_RNAi	diploid	1.1/1.1	None

2.2 Experiment 1: non-drought comparison

In experiment 1, plant line 1, 2 and 3 are grown. From each plant line we had five sub-lines. Sub-lines with in a plant line differ from each other because they are created during different crossing events, meaning they have different genotypes. However, they have the same alleles for *StCDF1*. From each sub-line, five plants (genetically identical) plants were grown. Also, five plants from parental line RH and E were used. This means that in total, 85 plants were used in this experiment. The plants were transferred to the greenhouse when they had around 7 leaves.

In this experiment, the number of tubers and the fresh weight of the tubers and the plant was determined. All tubers smaller than 1cm were not counted as tubers. The number of tubers and the fresh weights were measured during the harvest. The expression of *StCDF1* and *StSP6A* were determined by RNA extraction from seedlings that were grown earlier. The RNA extraction (including DNase digestion) was achieved with the RNeasy Plant Mini Kit from Qiagen, with the corresponding protocol. After that, the RNA was converted into cDNA according to the protocol 'Converting RNA into cDNA' in appendix A. The expression of *StCDF1* and *StSP6A* was determined, relative to *StNAC*, with qPCR. The primers that were used are shown in Table 4. The protocol that was used for qPCR was 'Protocol for Quantitative Real-Time PCR' and can be found on the drive on;

Table 4: Primers used for qPCR analysis of StCDF1 and StSP6A expression

target	Name of the primer	Sequence	
CDF1	qstCDF1-3UTR For qstCDF1-3UTR Rev	GCAGAAATGCAGGGTAAAGC GACACAAGAACCCGCTATGC	
SP6A	SP6A_229 for SP6A_229 rev	GACCCTAACTTGAGGGAGTA CGTCTTGATTGTCGAAATAA	
NAC	NAC_forward NAC_reverse	ATATAGAGCTGGTGATGACT TCCATGATAGCAGAGACTA	

2.3 Experiment 2: drought comparison

In experiment 2, the same lines, sub-lines and number of plants per sub-line are used as for experiment 1. Furthermore, from plant line 4 till 8 and 11, five plants were grown. From plant line 9,13 and 14 three sub-lines were used, from plan line 10 two sub-lines and from plant line 12 we used 8 sublines. From each of these sub-lines, five plants were grown. This means that at the start of the experiment we used 220 plants in total. Table 5 gives an overview of the number of sublines and number of plants used.

Table 5: number of sublines, plants per subline and total number of plants per line used in experiment 2

Plant line number	Plant line name	# sub lines	# plants per subline	Total # plants per line
1	RHXE_homoWT	5	5	25
2	RHXE_hetero	5	5	25
3	RHXE_homoCDF1.3	5	5	25
4	CE3027	1	5	5
5	CE3130	1	5	5
6	Andigena	1	5	5
7	Desiree	1	5	5
8	and_suc2_GFP	1	5	5
9	and_suc2_CDF1.2_GFP	3	5	15
10	and_suc2_CDF1.1_GFP	2	5	10
11	3027_suc2_GFP	1	5	5
12	3027_suc2_CDF1.2_GFP	8	5	40
13	3027_suc2_CDF1.1_GFP	3	5	15
14	CE3130_CDF1_RNAi	3	5	15
15	CE3027_CDF1_RNAi	2	5	10
-	RH	1	5	5
-	E	1	5	5
			total	220

Once the plants had 10 leaves, they were meant to be deprived of water till the end of the experiment. However, due to a miscommunication in the greenhouse, the plants with a GFP construct (plant line 8 till 13) did receive water. It was decided to continue measuring these plants and collect leave samples. However, we did not extract RNA from the leaf samples and the results are not used in this research.

Stomatal conductance and chlorophyll content was measured at three different moments. It was not possible to measure all plants on the same day in the morning, there for, measuring took several mornings each time. The timeline of the experiment is given in Table 6.

Table 6: timeline of experiment 2

Activity	date
Transferred plants to greenhouse	1th of December
Started drought stress	31sth of January
1st time measuring conductivity	2 nd of February - 3 th of February
1st time measuring chlorophyll	2 nd of February
Part of the plants were watered	6 th of February
2 nd time measuring conductivity	7 th of February - 8 th of February
2 nd time measuring chlorophyll	8 th of February
Leaf sampling of the drought stressed plants	10 th of February
3 th time measuring conductivity	13 th of February-15 th of February
3 th time measuring chlorophyll	13 th of February
Leaf sampling of the watered plants	16 th of February
Taking pictures of the drought stressed plants	16 th of February
Harvested drought stressed plants	17 th of February
Harvested the watered plants	21th of February

Chlorophyll was determined by taking the average chlorophyll content of the last leaf of the 5th, 6th and 7th leaflet. Conductivity was measured with by sampling the last leaf of the fifth leaflet. Measuring of the chlorophyll content and conductivity was always done in the morning, between 9:00 and 12:00.

To determine the expression of *StCDF1* and *StSP6A* the 4th and 6th leaflet were sampled of two of the five plants (plant A and B) of each subline. The sampling was performed between 9:30 and 11:40 in the morning. The leaf samples were immediately put in liquid nitrogen and stored at -80°C until further use. The grinding of the leaf samples was manually executed. The samples and the material used for grinding were kept cool during grinding with liquid nitrogen to ensure the leaf samples would not thaw. Extracting RNA and DNase digestion of the leaf samples was achieved with the RNeasy Plant Mini Kit from Qiagen, with the corresponding protocol. After RNA extraction, the RNA was transformed into cDNA. This was achieved with the protocol 'Converting RNA into cDNA' in appendix A. With the obtained cDNA the expression of *StCDF1* and SP6A was determined, relative to *StNAC*,. The primers that we used are the same as in experiment 1 and are shown in Table 4. The protocol that was used for qPCR was 'Protocol for Quantitative Real-Time PCR' and can be found in appendix A;

The expression of the SCAL1 was also determined by qPCR although the cDNA was made in a different way. The cDNA was made with the invitrogen 'SuperScript®III First-Strand Synthesis System for RT-PCR' kit, following the corresponding protocol. The only deviation from the protocol is that we halved all the advised quantities. The primers that we used for making the cDNA are shown in Table 7.

Table 7: Primers used for making cDNA for the SCAL1 analyses

target	Name of the primer	Sequence
CDF1	cDNAr1CDF	GACGTGATTGTAGTTACTAACC
CDF1	cDNAr2CDF	TGGGACGTAAAGACGTTCTC
EF3	RT_REV_349 REV	CGTTGGTGAATGCGGCAGTAGG
NAC	NAC_reverse	TCCATGATAGCAGAGACTA

We included EF3 and NAC to have the opportunity to choose between two reference genes. In the end, we only used NAC. The primers used for the qPCR are shown in Table 8.

Table 8: Primers used for the qPCR analysis of SCAL1

target	Name of the primer	Sequence
CDF1	qPCRrFCDF qPCRrRCDF	TGCAGACTCGTCGATTGAAC GAGTGCCTTTTCCTCACTCG
NAC	NAC_forward NAC_reverse	ATATAGAGCTGGTGATGACT TCCATGATAGCAGAGACTA

Since the GFP plants received water there was no use in analysing the GFP (and thus StCDF1) activity by Western Blotting for this research. However, the samples are collected and if it is needed for further research the samples can be analysed according to the protocol in Abelenda, Cruz-Oró, Franco-Zorrilla, & Prat, 2016.

Since the qPRC plates used in this analyses only have 96 wells, we used multiple plates. To correct for the difference between plates we calculated the average NAC Cq and for each sample divided this average Cq by the original Nac Cq, resulting in an correction factor. Then, each *StSP6A* and *StCDF1* Cq was multiplied by this sample specific correction factor. After the correction, the RGE expression was calculated by the following formula

$$RGE = 2^{Cq_{NAC} - Cq_{GOI}}$$

Where Cq_{NAC} is the average Cq value of NAC and Cq_{GOI} is the sample specific Cq of the gene of interest.

During harvesting of the plants, the number of tubers and the wet weight of the tubers and plant were determined. The plants and the potatoes were then stored in paper bags. The dry weight was determined by transferring the plants and tubers from the paper bags to aluminium trays and let them dry at 100°C for 44 hours. After that, the tubers and plants were weighed. The water content was determined with the following formula;

$$WC = \frac{Ww - Wd}{Ww} \times 100$$

Where WC=Water content, Ww=wet weight and Wd=dry weight.

2.4 Experiment 3: inactivation of the promotor region

The goal of experiment 3 was to create a plant with a wild type StCDF1 protein but without SCAL1. It was decided to block the expression SCAL1 in a CE3027 plant by mutating the promotor of SCAL1 with the use of CRISPR/Cas9 technology. This is done in different steps; 1, creating a vector that codes for a Cas9 protein with the right target sites, 2, transferring the vector to the plant with the use of *Agrobacterium tumefaciens* and 3, detection of CRISPR activity.

After looking at the sequence of *StCDF1* and the flanking regions three PAM sites in the promotor region of SCAL1 were found; One at around the end of the promotor region, one at 400bp and one at 1KB distance from the end of the promotor region as shown in Figure 5.

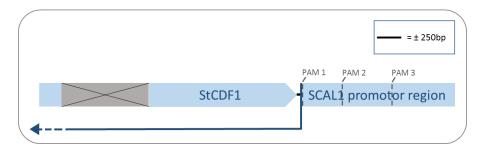


Figure 5: schematically representation of the StCDF1 gene

In this figure the protein coding region of the *StCDF1* gene and the SCAL1 promoter region are shown in light blue. The grey box with a cross represent the intron of *StCDF1*. The dark blue arrow represents the location and transcription direction of SCAL1. Since it is not known precisely how long SCAL1 is the end of the arrow is dotted. The three PAM sites, PAM1, 2 and 3, are also shown. In the right corner the approximate scale of the pictures is given.

We decided to attempt creating two different vectors that would transcribe for Cas9 proteins that each cut on two locations; at 0bp and 400bp, and at 0bp and 1kb.

For step 1, creating the right vector, we used the protocol in the syllabus of the "1st TALEN & CRIPSR Training School, Halle (saale), Germany", starting from page 12. An overview of the protocol, including deviations from the protocol will be given here. The protocol starts by amplifying the sgRNA from the selected PAM sites with a PCR reaction. SgRNA1, sgRnA2 and sgRNA3 are respectively associated with the PAM site at 0, 400 and 1000bp form the end of the promotor region. The primers we developed for the amplification of the sgRNA's are shown in Table 9. The blue part is the restriction enzyme site, capital bold shows the PAM region and the red part is the DNA binding site. The first three primers are forward primers and contain the DNA binding site. The forth primer is the reveres primer.

Table 9: The different primers used for amplifying the sgRNA needed for the construction of level 1 vector
The blue part is the restriction enzyme site, capital bold shows the PAM region and the red part is the DNA binding side.

Primer	sequence	Distance to end of the promotor region
1	5'-tgtggtctca ATT gtgaaagtaaccataatatt gttttagagctagaaatagcaag-3'	~0 base pairs
2	5'-tgtggtctca ATT aaataacataaggacatggc gttttagagctagaaatagcaag-3'	~ 400 base pairs
3	5'-tgtggtctca ATT tatgtggttcaaaaagctaa gttttagagctagaaatagcaag-3'	~1kb
Reverse	5'-tgtggtctca AGCG taatgccaactttgtac-'3	

The PCR conditions were as followed:

PCR material: PCR program: 0.5µl P966 (1 cycle) 5.0µl 5XHF buffer 98°C, 30 seconds 1μl dNTP (5mM) (35 cycles) 1µl forward primer (either 1,2 or 3) 98°C, 10 seconds 1μl reverse primer 60°C, 20 seconds 0.2μl Phusion 72°C, 20 seconds 16.3 μl Mili Q 72°C, 5 minutes

Although advised in the original protocol, we did not desalt the PCR products after the PCR reaction. After the PCR reaction the sgRNA is transferred to a level 1 vector by cut-ligation. A more detailed description of the vector components is given in Appendix C. The cut-ligation conditions were as followed:

Cut-ligation material: Assembly reaction: 2μl Buffer G (50 cycles) 1µl Bsal 37°C, 5minutes 1μl T4 DNA ligase 20°C, 5 minutes 2μl AtP (10mM) (1 cycle) $1\mu l \; Mili \; Q$ 50°C, 10 minutes 1µl P009 80°C, 10 minutes 1μl P751 (424ng/μl) or P761 (275 ng/μl) 1μl PCR product (1,2 or 3)

For PCR product 1 (sgRNA1), we used P751 and for PCR product 2 and 3 (sgRNA2 and sgRNA3) we used P761.

We then transformed the plasmids into E. coli. The steps for this transformation can be found in the protocol 'transformation into E. coli' in appendix A. To check if we indeed created the right vectors, we isolated the DNA with 'QIAprep spin MINIprep kit' from GIAgen according to the supplied protocol or by the protocol 'quick and dirty DNA isolation' in appendix A and a control digestion was performed. If the right bands appeared the DNA was prepared for sequencing. The primer that we used to multiply the DNA for sequencing was primer 40 (see Table 7). The samples were prepared for sequencing according to the protocol 'Sample preparation for GATC sequencing' in Appendix A and send to an external company.

Unfortunately, we did not succeed in cut-ligating the sgRNA2 into the vector. However, we did succeed for sgRNA1 and sgRNA3. To combine sgRNA1 and sgRNA3 we did a Golden Gate assembly, creating a level 2 vector. Details on the vector components can be found in Appendix C. The assembly reaction was as followed:

Assembly material:	Assembly reaction:
200ng 723	(50 cycles)
100ng 732	37°C, 2minutes
100ng 742	16°C, 5 minutes
100ng level1 vector sgRNA1	(1 cycle)
100ng level 1 vector sgRNA3	37°C, 5 minutes
75ng 780	80°C, 10 minutes
2μl Bpil buffer	
1μl Bpil	
2μl T4 ligase	
2μl ATP	
7.8μl Mili Q	

To check if the right vector was created DNA was isolated (in the same way as described above) and a control digestion was performed. If the control digestion showed the right size of bands on a gel the DNA was send for sequencing. The primers that we used for sequencing were; 34, 35, 36 and AL717. These primers are described in Table 7.

Table 10; Primers used for sequencing of the level 1 and level 2 vector

abbreviation	Name of the primer	Level 1 or 2 vector	Sequence
40	PDS8340	1	GAACCCTGTGGTTGGCATGCACATA
34	PDS8534	2	TTTGTGATGCTCGTCAGGGG
35	PDS8535	2	CCCGAGAATTATGCAGCATTTT
36	PDS8536	2	TCATCAGTCAATTACGGGGCT
AL717	AL717	2	GCTTGGCATCAGACAAACCGG

Primary attempts to perform this assembly reaction were unsuccessful. By comparing the sequencing results with the sequence the vector is expected to have, constructed by Marianne Oortwijn (constructed in APE, v2.0.51) we found that the promotor region in the obtained vector was missing. This appeared to be due to the 732 vector, which had a mutation inhibiting the correct assembly of the level 2 vector. Therefore, we tried to do the assembly again but replaced vector 723, 732, 742 and 780 that we used by ones that were known to be correct (provided by Annelies Loonen). This attempt was successful.

Step 2 was transforming the vector into *Agrobacterium tumefaciens*, according to the protocol "*Agrobacterium transformation of Potato*" in appendix A. After transforming the plasmid in *A. tumefaciens* the, DNA was isolated from ten colonies to check if they contained the right vector by performing a control digestion. All ten had the right bands (although very light) and four were send for sequencing. After sequencing, only one colony had the right sequence. However, this was most likely due to low DNA quality. Therefore, all four bacteria cultures were regrown in 50ml LB-medium with Kanamycin for inoculation of the potato plants. All four regrown bacteria cultures had grown well, and one was chosen for inoculation. The plant used for inoculation was CE3027, because this plant is homozygote *StCDF1.1*. From here on, the experiment was continued by Lorena Ramirez Gonzales, who will continue this research as part of her PhD research.

3 Results

3.1 Experiment 1 and 2

3.1.1 Phenotypic response to drought

In order to analyse the phenotypic differences in drought stress response the appearance of the different lines was compared and representative plants were photographed after drought stress. The pictures are shown in Figure 6. There was a clear difference in phenotype between the plants resulting from the RH x E cross (Figure 6a). The plants that were homozygous for *StCDF1.1*, seemed to cope better than the heterozygous *StCDF1.1*/*StCDF1.3* plants and the homozygote *StCDF1.3* plants. There was no distinct difference in drought stress response between heterozygote *StCDF1.1*/*StCDF1.3* plants and homozygote *StCDF1.1* plants. The parental lines, RH and E (Figure 6b) showed a similar phenotype as the homozygote *StCDF1.1* and heterozygote *StCDF1.1*/*StCDF1.3* plants.

The RNAi knockdowns of *StCDF1* in CE3130 seemed to be more tolerant compared to non-transformed CE3130 plants (Figure 6c). The same holds for CE3027 plants and their RNAi knockdowns, although the difference between the RNAi plants and CE3027 was smaller (Figure 6e). Andigena and Desiree both seemed to have a low drought tolerance based on their phenotypic response, although Andigena did seem to do a little bit better (Figure 6d).

In conclusion it seems that the *StCDF1.1/StCDF1.3* plants have a similar response to drought stress as homozygote *StCDF1.3*. Both plant types seem to cope worse with the drought stress than homozygote *StCDF1.1* plants. This indicates that the *StCDF1.3* allele is dominant over the *StCDF1.1* allele in respect to the drought stress response and furthermore, that *StCDF1.3* causes a decrease of drought tolerance. For both CE3130 and CE3027 the RNAi lines seem to cope better with the drought stress than the non-RNAi lines, meaning that both *StCDF1.1* and *StCDF1.3* have a negative effect on drought stress although the drought tolerance is lower with *StCDF1.3*.





a)





c)



Figure 6: drought stressed plants

d)

(a) shows, from left to right a homozygote *StCDF1.1*, a heterozygote *STCDF1.1/STCDF1.3*, and a homozygote *StCDF1.3* plant. (b) shows a RH plant and a E plant. (c) shows a CE3130 plant and the two right plants are RNAi_CDF1_CE3130 plants of different sublines. (d) shows a Anidgena plant (left) and a Desiree plant (right). (e) shows a CE3027 plant and the two right plants are RNAi_CDF1_CE3027 plants of different sublines.

e)

3.1.2 Weight and number of tubers

To give an indication of the effect of drought stress on the yield of the different plant lines, the weight of the tubers and the number of tuber was determined.

Figure 7 shows the total wet weight of the tubers per plant line for the non-drought stressed plants (a) and the drought stressed plants (b). This figure shows that the drought stressed plants have lower values for the tuber wet weight than the drought stressed plants. Furthermore, it shows that the homozygote *StCDF1.1/StCDF1.1*, Andigena, CE3027, RNAi CDF CE3027 and RNAi CDF CE3130 plants do not have tubers or have very small tubers.

E, RH, homozygote *StCDF1.1/StCDF1.1*, heterozygote *StCDF1.1/StCDF1.3* and homozygote *StCDF1.3/StCDF1.3* plants seem to have a similar pattern in the weight of the tubers in both the non-stressed plants as the drought stressed plants, meaning there most likely is no interaction affect between plant line and drought stress. In Figure 7c and d the number of tubers for both the non-drought stressed plants and the drought stressed plants is shown respectively. there appears to be no interaction between drought stress and plant line on the number of tubers.

In Figure 7e and f the difference in the weight of the tubers and difference in the number of tubers between drought and non-drought stressed plants is shown respectively. The difference in the number of tubers is close to 0. This means that the number of tubers is not (or only slightly) effected by drought stress. However, the weight difference between drought and non-drought plants is larger and also more variable over the different genotypes. Still, the relative decrease in tuber weight (relative to the weight of the tubers in non-drought conditions) seems to be similar, meaning that there is no three-way interaction between plant type, drought/non-drought stress and tuber weight.

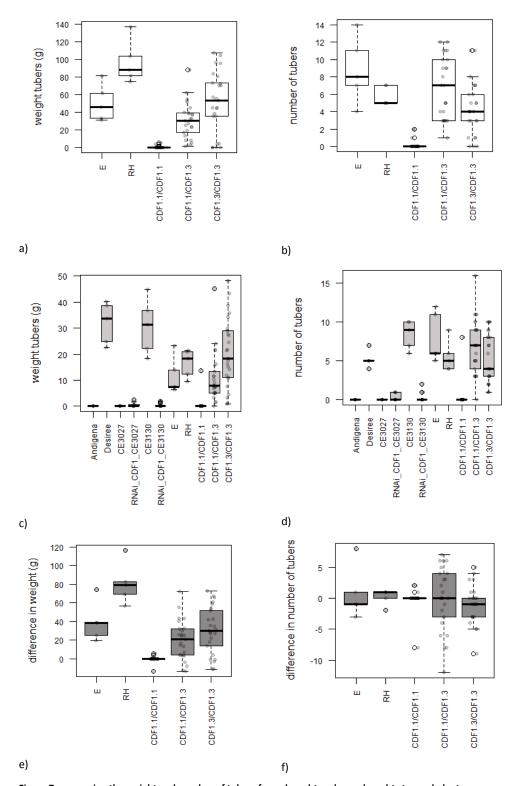


Figure 7: comparing the weight and number of tubers from drought and non-drought stressed plants

This figures shows the wet weight of the tubers and the number of tubers of non-drought stressed plants (a and b) and of the drought stressed plants (c and d), the difference in weight of the tubers (e) and difference in number of tubers (f) between non-drought and drought stressed plants of the same line. The figures are standard boxplots, with the addition that individual observations are shown. The lower whisker is the lowest value within 1.5 times the interquartile range (IQR) of the lower quartile. The upper whisker is the highest value still within 1.5 times the IQR of the upper quantile

3.1.3 Water content

Besides the weight of the plant and the weight and number of tubers, percentage of water was also calculated for the drought stressed plants to give an indication of how much water the plants lost due to drought stress. For the calculation, the following formula was used:

$$WC = \frac{Ww - Wd}{Ww} \times 100$$

Where WC=Water content, Ww=wet weight and Wd=dry weight.

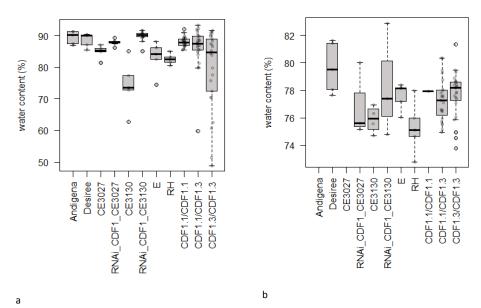


Figure 8: Water content of the plants and tubers after drought stress

Water content of (a) the plants and (b) the tubers of drought stressed plants per plant line. The two boxplots are standard boxplots, with the addition that individual observations are shown. The lower whisker is the lowest value within 1.5 times the interquartile range (IQR) of the lower quartile. The upper whisker is the highest value still within 1.5 times the IQR of the upper quantile

One value of WC tubers was below zero (RNAi_CDF1_CE3127, subline 5, plant A) and was therefore removed. Also, two values were higher than 85% and are likely to be sampling errors and removed (CE3130, subline 1, plant A and CDF1.1/CDF1.3, subline 6, plant C). Figure 8 shows the WC of the plants and the tubers. In both the plants WC and the tuber WC CE3130 has a low water content. However, the CE3130 plants with RNAi interference have a WC closer to the other lines. This means the lower WC in CE3130 is likely caused by the CDF1 gene and thus due to the combination of the *StCDF1.2* and the *StCDF1.3* allele.

3.1.4 Stomatal conductance

To give an estimate of the stomatal conductance, the conductivity of leafs was measured. Since conductivity measurements are not very precise, it was decided to define all plants with a conductivity higher than 30 as alive, while a plants with a lower conductivity were named dead. Table 11 shows the number of plants that were alive and dead at the end of the experiment according to this criterion. Since not all plant lines had the same number of plants they are difficult to compare. Therefore, the percentage of alive and dead plants per plant line are visualised in Figure 9. It shows that the heterozygote *StCDF1.1/StCDF1.3* and E plants relative perform the worst, and Desiree performs best. However, E and Desiree both only

had 5 samples. The RNAi lines seem not to cope very different with the drought stresses, based on conductivity, compared to the normal CE3027 and CE3130 plants. The homozygote *StCDF1.3/StCDF1.3* line and homozygote *StCDF1.1/StCDF1.1* plants have a similar response to drought stress, while the heterozygote *StCDF1.1/StCDF1.3* line performs less good, with 50% less alive plants.

Table 11: number of dead and alive plants based on conductivity

based on conductivity		
plant line	alive	dead
desiree	5	0
andigena	4	1
RH	4	1
E	1	4
CE3027	4	1
CE3130	4	1
RNAi CE3027	7	3
RNAi CE3130	11	4
Homo 1.3	10	15
homo 1.1	10	15
Hetro 1.1/1.3	5	20

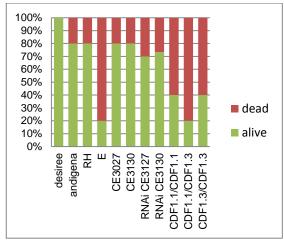


Figure 9: percentage of dead and alive plants based on conductivity

3.1.5 Chlorophyll

For chlorophyll, the dead or alive threshold was put on 43. As shown in Table 12 and Figure 10 Desiree, Andigena, E and RH are still all alive. CE3027 and CE3130 have the same number of alive and dead plants, although the difference between the RNAi lines is rather large, with the RNAi CE3130 plants generally performing better than RNAi CE3027 plants. If we compare the plants resulted from the RHxE cross it becomes clear that the hetrozygote StCDF1.1/StCDF1.3 plants perform best, followed by homozygote StCDF1.1/StCDF1.1 plants and then homozoygote StCDF1.3/StCDF1.3 plants

Table 12 number of dead and alive plants based on chlorophyll

on emorophyn		
plant line	alive	dead
desiree	5	0
andigena	5	0
E	5	0
RH	5	0
CE3027	1	4
CE3130	1	4
RNAi CE3127	0	10
RNAi CE3130	7	8
Homo 1.3	5	20
homo 1.1	16	9
Hetro 1.1/1.3	19	6

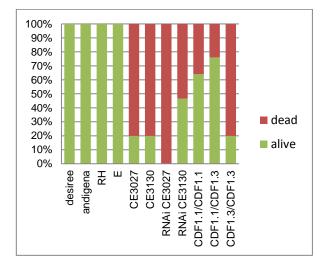


Figure 10: percentage of dead and alive plants based on chlorophyll

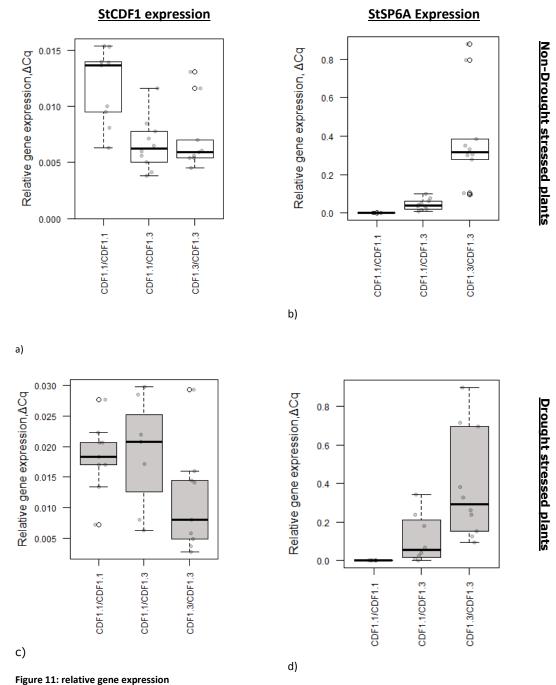
3.1.6 Expression profile of StCDF1 and StSP6A

In order to compare the activity of *StCDF1* and *StSP6A* between samples the expression of those genes was determined with qPCR. *StSP6A* was included to give an indication if the StCDF1 protein is stable.

In appendix B the expression of *StCDF1* and *StSP6A* is shown for each sample for both the non-drought stressed plants as the drought stressed plants. These figures show that there are two outliers (Hetero 2A and Trans14A) for the *StCDF1* expression of drought stressed plants and one (WT18A) for the SP6A expression of drought stressed plants. These outliers are removed and not taken into account in the further analyses, since they are most likely caused by a mistake.

Figure 11 gives the relative gene expression of *StCDF1* and *StSP6A* for both the non-drought stressed (a and b) as the drought stressed plants (c and d). Only the plants resulting from the RHxE plants are included. These figures show that the average expression of *StCDF1* in non-drought stressed plants is very similar for the heterozygote *StCDF1.1/StCDF1.3* plants and the homozygote *StCDF1.3/StCDF1.3* plants. However, in drought stressed plants this is not the case. It seems that drought stress causes a slight up regulation of *StCDF1* in homozygote *StCDF1.1/StCDF1.1* plants and a relatively high up regulation in heterozygote *StCDF1.1/StCDF1.3* plants. However, *StCDF1* in homozygote *StCDF1.3/StCDF1.3* plants does not seem to be upregulated.

In contrast to the expression of *StCDF1*, the expression of *StSP6A* is similar between drought and non-drought conditions. Furthermore, the expression of *StSP6A* is much more linked to the plant line, with the homozygote *StCDF1.3/StCDF1.3* plants having the highest overall expression, followed by the heterozygote *StCDF1.1/StCDF1.3* plants and the homozygote *StCDF1.1/StCDF1.1* plants. These figures further show that the expression of *StCDF1* is much lower than the expression of *StSP6A*.



relative gene expression of CDF1 (a) and SP6A(b) for the non-drought stressed plants, grouped per plant line, and relative gene expression of CDF1 (c) and SP6A(d) for the drought stressed plants, grouped per plant line.

3.1.7 Expression of SCAL1

We analysed the expression of SCAL1 in sample CE3130a, CE3130b, Heterozygote StCDF1.1/StCDF1.3_1b, Heterozygote StCDF1.1/StCDF1.3_4b, RNAi_CDF1_3130_1a and RNAi_CDF1_3130_2a (the letter at the end stands for the specific plant and the letter before that for the specific subline). Unfortunately, we were unsuccessful in getting a value for CE3130b, Heterozygote StCDF1.1/StCDF1.3_1b and Heterozygote StCDF1.1/StCDF1.3_4b. It might be that the expression of SCAL1 of these samples was too low to be detected.

However, for CE3130a, RNAi3130_1a and RNA3130_2a we did get vashown Figure 12. These results seem to indicate that if the expression of *StCDF1* goes up, the expression of *SCAL1* goes down. For CE3130a the relative expression of *StSP6A* was 1.322 (SEM=0.168) while for the other two the expression was 0.

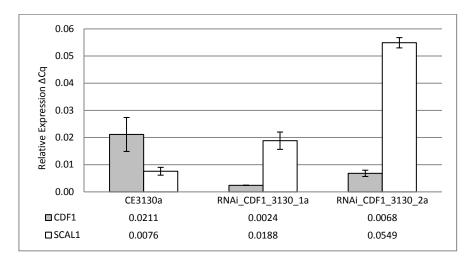


Figure 12 relative expression of CDF1 and SCAL1

This figure shows the relative expression of CDF1 and the SCAL1. Below the figure, the relative expression values are given.

3.2 Experiment 3

In experiment 3, a start was made in creating the plant material needed to determine the effect *StCDF1* has on drought tolerance, without interference of SCAL1. The created vector, that is used to infect CE3027 plants, is shown in Figure 13. If the infection is successful, and he vector works, these CE3027 plants will make a CRISPRCas9 protein that cuts out around 1kb of the promoter region of SCAL1, hopefully resulting in plants with an intact *StCDF1.1* gene but without SCAL1.

Of the bacteria cultures carrying the level1 (with either sgRNA1 or sgRNA3) vector or the level 2 vector glycerol stocks were made. Information about these glycerol stocks are in Appendix D.

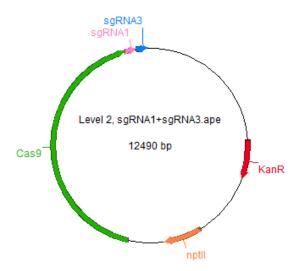


Figure 13: The vector created in experiment 3.

This vector combines sgRNA1 and sgRNA3, as well as a cas9 transcription region, nptll (plant Kanamycin resistance) and KanR (bacterial kanamycin resistance).

4 Discussion

Understanding the function of the *StCDF1* gene is of great importance as this gene is, for a large extend, responsible for the timing of tuberisation. Previous unpublished research already showed a link between this gene and drought tolerance.

However, until now, the mechanisms behind the relationship between the *StCDF1* gene and drought tolerance have not been studied.

4.1 Effect of *StCDF1* on drought stress tolerance

From the phenotypic results (figure 6) it seems that both the StCDF1.1 allele and the StCDF1.3 allele have a negative effect on drought tolerance. When looking at the pictures in figure 6, it becomes clear that knocking down StCDF1 results in an increase of drought tolerance in both CE3130 and CE3027. This is in agreement with the results of Dr. José Abelenda, who found that potatoes with a knockdown of StCDF1 cope better with drought (personal communication). Previous research has found that CDF's are involved in stress response. For example, Corrales et al., 2014 showed that overexpression of CDF's can lead to increased drought and salt tolerance in tomato and Fornara et al. 2015 found that the GI-CDF1 complex of Arabidopsis can control freezing tolerance.

Although both the StCDF1.1 and the StCDF1.3 allele have a negative effect on drought tolerance, there is a difference in severity, where the StCDF1.1 allele is performing better than the StCDF1.3 allele. Furthermore, it seems that StCDF1.3 is dominant over StCDF1.1 since both the StCDF1.3 homozygote plants and the heterozygote StCDF1.1/StCDF1.3 plants have a similar phenotypic response to drought stress.

Tuber wet weight is negatively affected by drought, which was also found by Cavagnaro et al. (1971). However, tuber number is not, which is in contrast with the results of Dwyer and Boisvert (1990), Haverkort et al. (1990) and Mackerron and Jefferies (1988). The relative reduction in tuber wet weight is equal between the plant lines, meaning that the reduction is likely not affected by the StCDF1 gene, but seems related to the weight of the tubers in non-drought conditions.

In contrast to the tuber number and weight, the water content (figure 8) of the plants after drought stress seems to be related to the StCDF1 gene, although the results are not fully conclusive. All plant lines have comparable plant water content, except for CE3130 (heterozygote StCDF1.2, StCDF1.3), which has lower plant water content. However, knocking down StCDF1 in CE3130 plants increases the plant water content a lot. This is unlike the CE3027 (which are homozygote for StCDF1.1) plants. This indicates that the StCDF1.1 allele and the mutated alleles may have a different effect on the water content of plants. If the StCDF1.3 allele is the main cause of the reduction in water content in CE3130 plants, the homozygote StCDF1.3 plants and heterozygote StCDF1.1/StCDF1.3 plants should also have a lower water content compared to the other plant lines. However, this is not the case. Therefore, we theorize that the reduced water content may be caused by the StCDF1.2 allele, rather than the StCDF1.3 allele. This is, however, in contrast to the previous research of Dr. Abelenda, mentioned in the introduction. Another theory to explain this phenomenon is an interaction effect between StCDF1.2 and StCDF1.3, implying that the low WC found in CE3130 is due to the combination of StCDF1.2 and StCDF1.3, rather than their individual effect.

However, it cannot be excluded that StCDF1.3 causes the lower water content in CE3130 plants. It might be that in homozygote StCDF1.3 plants and heterozygote StCDF1.1/StCDF1.3 plants other genes are interfering, compensating for the StCDF1.3 allele, since the genetic background of the plant (sub)lines is not equal. In addition, StCDF1 is part of the DOF family, and DOF's are often redundant in function (Cavalar et al., 2016; Morenso-Risueno et al., 2007).

The stomatal conductance and chlorophyll content of the leaves did not clearly differ between the different plant lines. During the measurements, it was noticed that the obtained values were not of high precision. It is therefore advised for further research, not to use conductivity as a measurement for stomatal conductance, but rather use thermo-imaging to give an indication of the transpiration of the plant. Another possibility would be to look at differences in the size and number of stomata, and if there is a difference in how much the stomata open. These stomata are especially interesting since unpublished data from José Abelenda (personal communication) has shown that knocking down the StCDF1 gene results in plants with fewer but larger stomatal guard cells, which were partially open (as compared to plants with an intact StCDF1).

Thus, based on the phenotypic response of the plants it is clear that StCDF1 has a negative effect on drought tolerance, although less severe for StCDF1.1 than StCDF1.3. However, quantifying the difference in drought tolerance is difficult with most of the drought stress response parameters used in this research.

4.2 *StCDF1* and *StSP6A* gene expression profile

Our results indicate that *StCDF1.3* is not able to respond to drought stress, while *StCDF1.1* can. During drought stress, homozygote *StCDF1.3/StCDF1.3* plants do not show any up regulation of *StCDF1*, as compared to normal conditions. In contrast to homozygote *StCDF1.1/StCDF1.1*, heterozygote *StCDF1.1/StCDF1.3* plants show an increase in the transcription of *StCDF1* Figure 11). We furthermore hypothesise that the up regulation of *StCDF1* seen in heterozygote *StCDF1.1/StCDF1.3* plants induced by drought stress, is caused by *StCDF1.1* rather than *StCDF1.3*. During long days, *StCDF1.1* does not result in up regulation of *StSP6A*, while *StCDF1.3* does (Figure 2), and our results show no up regulation of *StSP6A* in heterozygote *StCDF1.1/StCDF1.3* plants Figure 11). However, due to the rather large variance in *StSP6A* expression in heterozygote *StCDF1.1/StCDF1.3* plants and the small sample size it is difficult to draw definitive conclusions from these measurements.

It might be that *StCDF1.3* cannot respond to drought stress because it misses a functional copy of SCAL1. Several lines of research show that the expression lncRNA can be triggered by stress (for example; Di et al., 2014; Liu et al., 2012; Matsui et al., 2008). However, to our knowledge, no previous research shows evidence for the necessity of a lncRNA for further stress response, which could be an explanation of our results.

The results from the *StCDF1* and *StSP6A* gene expression also show that the effect of *StCDF1* on drought tolerance is most likely not regulated trough *StSP6A* and the flowering pathway, but involves another pathway that still needs to be understood. This is because an increase of *StCDF1* expression during drought stress, did not result in a change in *StSP6A* expression.

To further look into the theories proposed in this chapter, it is advised to make a primer that adheres to the 865pb insertion of *StCDF1.3* and repeat the qPCR analyses on the cDNA used in this research. With that *StCDF1.3* specific primer one can determine in homozygote plants which part of the *StCDF1* RNA is from *StCDF1.1*, and which part is from *StCDF1.3*. Another possibility is to repeat the experiment with an increased sample size and analyse the *StCDF1* and *StSP6A* expression. The expression of *StSP6A* can then be used as an indicator of the ratio of *StCDF1.1* and *StCDF1.3* expression in heterozygote plants, since StCDF1.3 proteins result in *StSP6A* expression, while StCDF1.1 proteins do not (in long day conditions). In this research, we tried to use *StSP6A* expression to explain our results of the *StCDF1* expression. However, since the sample size is too small to compensate for the variance between samples we cannot draw solid conclusions on the ratio of *StCDF1.1* and *StCDF1.3* expression in heterozygote plants.

4.3 Interaction between *StCDF1* and SCAL1

The results in Figure 12 indicate that StCDF1 can down regulate the expression of SCAL1 since if we decrease the transcription of CDF1 by RNAi interference, the expression of the SCAL1 goes up. Multiple articles explain how a lncRNA can influence the expression of a gene (for example; Wilus et al., 2009; Kwenda et al., 2016, Bhattacharjee et al., 2013; lee, 2012). In contrast, very little examples are known in which a gene influences the expression of a lncRNA. An exception to this is the research of Wibowo et la., 2016, who found evidence for a model in *Arabidopsis*, in which HS-DMR (Hyperosmotic Stress Differentially Methylated Regions), regulated by the *CNI1* (CARBON/NITROGEN INSENSITIVE 1) gene, can indirectly change the expression of *CNI1* by modulating the expression of a lncRNA. The fact that other research found that a gene can influence the expression of a lncRNA, makes our theory that StCDF1 effects SCAL1 more plausible.

In chapter 4.2 it was already mentioned that SCAL1 might be needed to increase the *StCDF1* expression in response to drought stress. The possible necessity of SCAL1 for *StCDF1* to respond to drought stress can also explain why heterozygote *StCDF1.1/StCDF1.3* plants under drought stress have an upregulated *StCDF1* expression, while homozygote *StCDF1.3* plants do not (Figure 11). Since the heterozygote *StCDF1.1/StCDF1.3* plants likely have a single copy of SCAL1 they can increase their *StCDF1* expression, in contrast to homozygote *StCDF1.3* plants which very likely have none.

If indeed SCAL1 is needed to increase the *StCDF1* expression in response to drought stress, and the *StCDF1* gene can down regulate SCAL1, there must be a negative feedback loop in which SCAL1 upregulates *StCDF1*, and *StCDF1* reduces the expression of the SCAL1. However, unpublished research of Dr. José Abelenda shows that the diurnal expression of *StCDF1* is similar to the diurnal expression of SCAL1 (personal communication). Since this result cannot be explained by a negative feedback loop between *StCDF1* and SCAL1, more research is needed to generate a better fitting model. Yet, unpublished research from Dr. Rossanna Henriques (personal communication) found that in Arabidopsis, the lncRNA of *CDF5* (for homology with *StCDF1* see Figure 3) has a contrasting diurnal expression pattern compared to the gene, as one would expect if there is a negative feedback loop. Therefore, it might be that the theory of a negative feedback loop is correct in Arabidopsis, but that in potato there are some evolutionary changes that form a different (possibly more complex) system.

The results of this report show indirectly that there likely is an interaction between SCAL1 and the DNA, RNA and/or protein of StCDF1, making SCAL1 a cis-acting lncRNA. In addition, it is possible that the SCAL1 has an effect on other genes. To look into this further, one can predict the most stable secondary structure of the SCAL1 RNA and see if it is likely to interfere with other genes. This can be done with techniques and programs mentioned by Mathews et al., (2010). It is advised to especially focus on other CDF genes since CDF's are known to often be complementary in their functions and down regulation of *StCDF1* results in up regulation of other CDF genes in potato (personal communication Dr. José Abelenda).

4.4 The effect of SCAL1 on drought tolerance

Assuming that SCAL1 has a positive effect on drought stress could explain why homozygote *StCDF1.1* plants do better under drought stress compared to homozygote *StCDF1.3* plants. This would also explain the unpublished results of Dr. José Abelenda shown in the introduction (Figure 4), which show that heterozygote *StCDF1.1/StCDF1.2* plants, which presumably have two copies of SCAL1, do better than *StCDF1.1/StCDF1.3* plants that only have one.

However, explaining why heterozygote *StCDF1.1/StCDF1.3* plants have similar drought tolerance when compared to homozygote *StCDF1.3/StCDF1.3* plants, instead of a slightly better one, is difficult. It is possible that a specific threshold of SCAL1 RNA is needed to increase drought tolerance but that this threshold cannot be reached with a single copy of SCAL1 in the genome. Should this be true, SCAL1 works in two different ways; 1, it is needed for *StCDF1* to be upregulated during

drought stress, and 2, the SCAL1 has a positive effect on drought tolerance and if more SCAL1 RNA is present or a specific threshold has been passed, the plants become more drought tolerant. This theory, combined with the theory of the negative feedback loop between SCAL1 and *StCDF1* (chapter 4.3), can be combined in a single model, explaining the relationship between SCAL1, *StCDF1* and drought tolerance. A simplified version of this model is shown in Figure 14.

However, the reason for upregulating CDF1 as a response to drought stress, which seems to have a negative effect on drought stress, cannot be explained at this point.

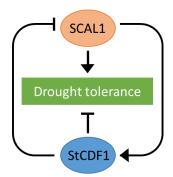


Figure 14: proposed model on the relationship between SCAL1, StCDF1 and drought tolerance This cartoon shows that SCAL1 upregulates StCDF1 and has a positive effect on drought tolerance, while StCDF1 down regulates SCAL1 and has a negative effect on drought tolerance.

To further elucidate the effect of SCAL1 on drought tolerance, and to proof or disproof the model proposed above, it is important to create a plant with a WT *StCDF1* allele but without SCAL1. In this research, we have created a vector needed for such a plant and transferred this vector to CE3027 clones but due to time constrains, we did not check if the transfer was successful. This will be done by Lorena Ramirez Gonzales, who will continue this research as part of her PhD research.

If the transformation was successful, we can see if SCAL1 is indeed necessary for the *StCDF1* gene to respond to drought stress. This can be tested by exposing both the CRISPR transformed CE3027 plants and non-transformed CE3027 plants, to both drought stress and normal conditions. If the non-transformed CE3027 plants have an increase in *StCDF1* in drought stress conditions compared to normal conditions, while the CRISPR transformed CE3027 do not, it is proven that SCAL1 is indeed needed.

Another important plant to create is the *StCDF1.2* homozygote plant from a CXSE cross. These plants most likely will have SCAL1 but with a StCDF1 protein that cannot be ubiquinated by the GI-FKF1 complex. If both these plant lines are available, it is possible to study the effect of the SCAL1 and the *StCDF1* protein separately, although the difference in genetic background might still be a problem. Therefore, it would be advised to use a large sample size for these studies.

Furthermore, if we get the *StCDF1.2* homozygote plant, we can compare the drought stress response of *StCDF1.2* homozygote and *StCDF1.3* homozygote plants. If *StCDF1.2* homozygote plants performs better after drought stress, as expected based on this research and previous unpublished research from Dr. José Abelenda (Figure 4), this can be taken into account when breeding for a long day independent tuber forming potato with higher drought tolerance.

5 Conclusion

In this research, we proposed a model for the interaction between *StCDF1* and SCAL1, the antisense lncRNA of *StCDF1*, and their effect on drought tolerance. In this model, *StCDF1* which seems to have a negative effect on drought tolerance, upregulates SCAL1 and SCAL1, which likely has a positive effect on drought tolerance, is needed for *StCDF1* to respond to drought stress. However, more research is needed to substantiate this model. For future research homozygote *StCDF1.2/StCDF1.2* plants as well as CE3027 plants without a functional SCAL1 promoter would be more than useful. Elucidating the role of SCAL1 would further deepen our still limited knowledge about lncRNA's, especially how a gene can influence the expression of a lncRNA. In addition, further research will likely clarify if *StCDF1.2* and *StCDF1.3* have a different effect on drought tolerance, this should be taken into consideration when breeding potato plants for northern latitudes.

6 References

- Abelenda, J. A., Navarro, C., & Prat, S. (2014). Flowering and tuberisation: a tale of two nightshades. Trends in plant science, 19(2), 115-122.
- Abelenda, J. A., Cruz-Oró, E., Franco-Zorrilla, J. M., & Prat, S. (2016). Potato StCONSTANS-like1 Suppresses Storage Organ Formation by Directly Activating the FT-like StSP5G Repressor. Current Biology, 26, 872–881. doi:10.1016/j.cub.2016.01.066
- Anithakumari, a M. (2011). Genetic Dissection of Drought Tolerance in Potato.
- Bai, Y., Dai, X., Harrison, A. P., & Chen, M. (2015). RNA regulatory networks in animals and plants: A long noncoding RNA perspective. Briefings in Functional Genomics, 14(2), 91–101. doi:10.1093/bfgp/elu017
- Barrangou, R., Fremaux, C., Deveau, H., Richards, M., Boyaval, P., Moineau, S., ... & Horvath, P. (2007). CRISPR provides acquired resistance against viruses in prokaryotes. Science, 315(5819), 1709-1712.
- Belhaj, K., Chaparro-Garcia, A., Kamoun, S., & Nekrasov, V. (2013). Plant genome editing made easy: targeted mutagenesis in model and crop plants using the CRISPR/Cas system. Plant methods, 9(1), 39.
- Bhattacharjee, S., Renganaath, K., Mehrotra, R., & Mehrotra, S. (2013). Combinatorial control of gene expression. BioMed Research International, 2013, 407263. doi:10.1155/2013/407263
- Bradshaw, J. E. (2009). Potato breeding at the Scottish plant breeding station and the Scottish Crop Research Institute: 1920–2008. Potato research, 52(2), 141-172.
- Casal, J. J. (2013). Photoreceptor signaling networks in plant responses to shade. Annual review of plant biology, 64, 403-427.
- Cavagnaro, J. B., De Lis, B. R., & Tizio, R. M. (1971). Drought hardening of the potato plant as an after-effect of soil drought conditions at planting. Potato Research, 14(3), 181-192.
- Cavalar, M., Phlippen, Y., Kreuzaler, F., & Peterhänsel, C. (2007). A drastic reduction in DOF1 transcript levels does not affect C 4-specific gene expression in maize. Journal of plant physiology, 164(12), 1665-1674.
- Corbesier, L., Vincent, C., Jang, S., Fornara, F., Fan, Q., Searle, I., ... & Coupland, G. (2007). FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. science, 316(5827), 1030-1033.
- Corrales, A. R., Nebauer, S. G., Carrillo, L., Fernández-Nohales, P., Marqués, J., Renau-Morata, B., ... & Medina, J. (2014). Characterization of tomato Cycling Dof Factors reveals conserved and new functions in the control of flowering time and abiotic stress responses. Journal of experimental botany, 65(4), 995-1012.
- Corrales, A.-R., Carrillo, L., Lasierra, P., Nebauer, S. G., Dominguez-Figueroa, J., Renau-Morata, B., ... Medina, J. (2017). Title: Multifaceted Role of Cycling Dof Factor 3 (CDF3) in the regulation of flowering time and abiotic stress responses in Arabidopsis. Plant, Cell & Environment. doi:10.1111/pce.12894
- Di, C., Yuan, J., Wu, Y., Li, J., Lin, H., Hu, L., ... & Lu, Z. J. (2014). Characterization of stress-responsive lncRNAs in Arabidopsis thaliana by integrating expression, epigenetic and structural features. The Plant Journal, 80(5), 848-861.
- Dwyer, L. M., & Boisvert, J. B. (1990). Response to irrigation of two potato cultivars (Solanum tuberosum L.'Kennebec'and'Superior'). Canadian Agricultural Engineering, 32(2), 197-203.
- Engreitz, J. M., Pandya-Jones, A., McDonel, P., Shishkin, A., Sirokman, K., Surka, C., ... Guttman, M. (2013). The Xist IncRNA Exploits Three-Dimensional Genome Architecture to Spread Across the X Chromosome. Science, 341(6147). Retrieved from http://science.sciencemag.org/content/341/6147/1237973.full
- FAOSTAT. 2017, June 20. Retrieved from http://www.fao.org/faostat/en/#data/QC.
- Fornara, F., Montaigu, A., Sánchez-Villarreal, A., Takahashi, Y., Ver Loren van Themaat, E., Huettel, B., ... & Coupland, G. (2015). The GI–CDF module of Arabidopsis affects freezing tolerance and growth as well as flowering. The Plant Journal,

- 81(5), 695-706.
- Guttman, M., & Rinn, J. L. (2012). Modular regulatory principles of large non-coding RNAs. Nature, 482(7385), 339–46. doi:10.1038/nature10887
- Haverkort, A. J., Van de Waart, M., & Bodlaender, K. B. (1990). The effect of early drought stress on numbers of tubers and stolons of potato in controlled and field conditions. Potato Research, 33(1), 89-96.
- Hijmans, R. J. (2003). The effect of climate change on global potato production. American Journal of Potato Research, 80(4), 271–279. doi:10.1007/BF02855363
- Houghton, J., Ding, Y., Griggs, D., Noguer, M., Linden, P. van der, Xiaosu, D., ... Johnson, C. (2001). Climate Change 2001. The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change (IPCC). Cambridge University Press. Retrieved from https://www.ipcc.ch/ipccreports/tar/wg1/pdf/WG1_TAR-FRONT.PDF
- Iwama, K., & Yamaguchi, J. (1994). Difference in root system development among crop species and genotypes. O. Ito C. Johansen JJ Adu-Gyamfi, 173.
- Jefferies, R. A. (1995). Physiology of crop response to drought. In Potato ecology and modelling of crops under conditions limiting growth (pp. 61-74). Springer Netherlands.
- Kloosterman, B., Abelenda, J. A., del Mar Carretero Gomez, M., Oortwijn, M., de Boer, J. M., & Kowitwanich, K. (2013). Naturally occurring allele diversity allows potato cultivation in northern latitudes. Nature, 495(7440), 246–250. doi:10.1017/CBO9781107415324.004
- Kloosterman, B., Anithakumari, A., Chibon, P.-Y., Oortwijn, M., van der Linden, G. C., Visser, R. G., & Bachem, C. W. (2012).

 Organ specificity and transcriptional control of metabolic routes revealed by expression QTL profiling of source--sink tissues in a segregating potato population. BMC Plant Biology, 12(1), 17. doi:10.1186/1471-2229-12-17
- Kwenda, S., Birch, P. R. J., & Moleleki, L. N. (2016). Genome-wide identification of potato long intergenic noncoding RNAs responsive to Pectobacterium carotovorum subspecies brasiliense infection. BMC Genomics, 17(1), 614. doi:10.1186/s12864-016-2967-9
- Lee, J. T. (2012). Epigenetic Regulation by Long Noncoding RNAs. Science, 338(6113). Retrieved from http://science.sciencemag.org/content/338/6113/1435.full
- Leone, S., & Santoro, R. (2016). Challenges in the analysis of long noncoding RNA functionality. FEBS Letters, 590(15), 2342–2353. doi:10.1002/1873-3468.12308
- Liu, J., Jung, C., Xu, J., Wang, H., Deng, S., Bernad, L., ... & Chua, N. H. (2012). Genome-wide analysis uncovers regulation of long intergenic noncoding RNAs in Arabidopsis. The Plant Cell, 24(11), 4333-4345.
- Lozano-Juste, J., & Cutler, S. R. (2014). Plant genome engineering in full bloom. Trends in plant science, 19(5), 284-287.
- Mackerron, D. K., & Jefferies, R. A. (1988). The distributions of tuber sizes in droughted and irrigated crops of potato. I.

 Observations on the effect of water stress on graded yields from differing cultivars. Potato Research, 31(2), 269-278.
- Mathews, D. H., Moss, W. N., & Turner, D. H. (2010). Folding and finding RNA secondary structure. Cold Spring Harbor perspectives in biology, 2(12), a003665.
- Matsui, A., Ishida, J., Morosawa, T., Mochizuki, Y., Kaminuma, E., Endo, T. A., ... & Satou, M. (2008). Arabidopsis transcriptome analysis under drought, cold, high-salinity and ABA treatment conditions using a tiling array. Plant and Cell Physiology, 49(8), 1135-1149.
- Mohammad, F., Mondal, T., & Kanduri, C. (2009). Epigenetics of imprinted long noncoding RNAs. Epigenetics: Official Journal of the DNA Methylation Society, 4(5), 277–286. doi:10.4161/epi.4.5.9242
- Monneveux, P., Ramírez, D. A., Khan, M. A., Raymundo, R. M., Loayza, H., & Quiroz, R. (2014). Drought and Heat Tolerance

- Evaluation in Potato (Solanum tuberosum L.). Potato Research, 57(3-4), 225-247. doi:10.1007/s11540-014-9263-3
- Moreno-Risueno, M. Á., Martínez, M., Vicente-Carbajosa, J., & Carbonero, P. (2007). The family of DOF transcription factors: from green unicellular algae to vascular plants. Molecular Genetics and Genomics, 277(4), 379-390.
- Nakamichi, N., Takao, S., Kudo, T., Kiba, T., Wang, Y., Kinoshita, T., & Sakakibara, H. (2016). Improvement of Arabidopsis Biomass and Cold, Drought and Salinity Stress Tolerance by Modified Circadian Clock-Associated PSEUDO-RESPONSE REGULATORs. Plant and Cell Physiology, 57(5), 1085–1097. doi:10.1093/pcp/pcw057
- Navarro, C., Abelenda, J. A., Cruz-Oró, E., Cuéllar, C. A., Tamaki, S., Silva, J., ... Prat, S. (2011). Control of flowering and storage organ formation in potato by FLOWERING LOCUS T. Nature, 478(7367), 119–122. doi:10.1038/nature10431
- Navarro, C., Cruz-Oró, E., & Prat, S. (2015). Conserved function of FLOWERING LOCUS T (FT) homologues as signals for storage organ differentiation. Current Opinion in Plant Biology, 23, 45–53. doi:10.1016/j.pbi.2014.10.008
- Plantenga, F. D. M., Siakou, M., Bergonzi, S., Heuvelink, E., Bachem, C. W. B., Visser, R. G. F., & Marcelis, L. F. M. (2016).

 Regulating flower and tuber formation in potato with light spectrum and day length. Acta Horticulturae, 1134, 267–275. doi:10.17660/ActaHortic.2016.1134.36
- Salaman, R. N., & Burton, W. G. (1985). The history and social influence of the potato. Cambridge University Press
- Sawa, M., Nusinow, D. A., Kay, S. A., & Imaizumi, T. (2007). FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. Science, 318(5848), 261-265.
- Schaeffer, S. M., & Nakata, P. A. (2015). CRISPR/Cas9-mediated genome editing and gene replacement in plants: Transitioning from lab to field. Plant Science, 240, 130–142. doi:10.1016/j.plantsci.2015.09.011
- Song, Y. H., Smith, R. W., To, B. J., Millar, A. J., & Imaizumi, T. (2012). FKF1 conveys timing information for CONSTANS stabilization in photoperiodic flowering. Science, 336(6084), 1045-1049.
- Song, X., Duan, W., Huang, Z., Liu, G., Wu, P., Liu, T., ... Hou, X. (2015). Comprehensive analysis of the flowering genes in Chinese cabbage and examination of evolutionary pattern of CO-like genes in plant kingdom. Scientific Reports, 5(March), 14631. doi:10.1038/srep14631
- Uitdewilligen, J. G. A. M. L. (2012). Discovery and genotyping of existing and induced DNA sequence variation in potato.
- Vos, P. (2016). Development and application of a 20K SNP array in potato.
- Wang, K. C., & Chang, H. Y. (2011). Molecular mechanisms of long noncoding RNAs. Molecular Cell, 43(6), 904–14. doi:10.1016/j.molcel.2011.08.018
- Wang, S., Zhang, S., Wang, W., Xiong, X., Meng, F., & Cui, X. (2015). Efficient targeted mutagenesis in potato by the CRISPR/Cas9 system. Plant Cell Reports, 34(9), 1473–1476. doi:10.1007/s00299-015-1816-7
- Wang, X., Ai, G., Zhang, C., Cui, L., Wang, J., Li, H., ... Ye, Z. (2016). Expression and diversification analysis reveals transposable elements play important roles in the origin of Lycopersicon-specific IncRNAs in tomato. New Phytologist, 209(4), 1442–1455. doi:10.1111/nph.13718
- Wibowo, A., Becker, C., Marconi, G., Durr, J., Price, J., Hagmann, J., ... & Weigel, D. (2016). Hyperosmotic stress memory in Arabidopsis is mediated by distinct epigenetically labile sites in the genome and is restricted in the male germline by DNA glycosylase activity. Elife, 5, e13546.
- Wiedenheft, B., Sternberg, S. H., & Doudna, J. A. (2012). RNA-guided genetic silencing systems in bacteria and archaea. Nature, 482(7385), 331-338.
- Wilusz, J. E., Sunwoo, H., & Spector, D. L. (2009). Long noncoding RNAs: functional surprises from the RNA world, 1494–1504. doi:10.1101/gad.1800909.complexity.
- Woo, J. W., Kim, J., Kwon, S. II, Corvalán, C., Cho, S. W., Kim, H., ... Kim, J.-S. (2015). DNA-free genome editing in plants with

preassembled CRISPR-Cas9 ribonucleoproteins. Nature Biotechnology, 33(11), 1162–1164. doi:10.1038/nbt.3389

Zeevaart, J. A. (2008). Leaf-produced floral signals. Current opinion in plant biology, 11(5), 541-547.

Zhang, Y. C., & Chen, Y. Q. (2013). Long noncoding RNAs: New regulators in plant development. Biochemical and Biophysical Research Communications, 436(2), 111–114. doi:10.1016/j.bbrc.2013.05.086

Appendix A: protocols

Converting RNA into cDNA

- Determine concentration of RNA with the nanodrop
- Calculate how much RNA solution is needed to get 500ng
- put this amount of RNA solution on a 96-wells plate and add:
 - 1μl RT
 - 4μl i-script mix
 - O Mili-Q till a total volume of 20μl
- Place plate in pcr machine and perform the following reaction program:
 - o 25°C, 5 minutes
 - o 42 °C, 30 minutes
 - o 85 °C, 5 minutes
 - o 4 °C, 5 minutes
 - o 85 °C, 5 minutes
 - o 10 °C, ∞
- Add 180 μl RNAse free water to each well

Transformation in to E. Coli

- Add 3 μl of the ligation-mix to 50 μl competent cells (from the box named 'Top 10')
- Incubat on ice for 30 minutes
- Incubate at 42°C for 30 seconds
- Incubate on ice for 2 minutes
- Add 950 μl liquid LB-medium
- Incubate at 37°C for 1 hour, at 225rpm
- \bullet Plate 100 μ l of the transformation mixture on LB-Carb with X-Gal plates
- Centrifuge the remaining part of the transformation mixture for 5 minutes at 10,000 rpm
- Through away ± 800 μl of the supernatant
- Ressuspend pellet in what is left to the supernatant ($\pm 100 \mu$ l)
- Plate everything on on LB-Carb with X-Gal plates
- Place at 37°C over night

Protocol for Quantitative Real-Time PCR

Material:

- iQ SYBR GREEN super mix 2* (BioRad 25*5 ml tubes #172-5006 CUST)
- Optical quality 96-well thin wall PCR plates (BioRad 25 plates skirted #223-9441)
- Micro seal B Optical quality sealing tape (BioRad 100 pieces # MSB-1001)
- Primers Biolegio, dissolve in MQ with DNA free pipettes and tips
 - Scale: 40 nmols
 - Method: Economy
 - No purification
 - o Quality check: Normal no extra
 - Chemistry DNA
 - o Dissolved: No
- Rainin Tips for DNA free work (Mettler Rainin)
 - 0.5-20 μl (#17005860)
 - 10-200 μl (#17005859)
 - o 100-1000 μl (#17007081)
- DNA free Rainin pipettes (Drawer nearby qPCR machine)

Method:

Prepare RT-PCR mixture. The PCR is performed in a volume of 20 μ l per well. Therefore the reaction volume is 44 μ l (duplo) or 65 μ l (triplo)! Work clean avoid cross contamination.

Duplo		Triplo
22.0 μΙ	32.50 μl	2*iQ SYBR GREEN super mix
11.0 μΙ	16.25 μl	MQ
4.4 µl	6.50 µl	Forward primer (3 µM)
4.4 µl	6.50 µl	Reverse primer (3 μM)

Add per sample:

```
2.2 μl 3.25 μl cDNA (10 ng/μl)
44.0 μl 65.00μl total volume
```

- 3. Fill the 96 wells plate according to your scheme with 20µl reaction mixture per well (duplo or triplo). For every well use a fresh tip! When dividing the mixture, cover the other wells you are not working with!
- 4. Carefully seal the RT-PCR plate with special film (Micro seal, adhesive sealer). When taking out the film, only hold its edge which is not transparent. Do not touch the transparent part of the film! After sealing the film PCR plate, use a small plastic board (can be found in the top drawer near RT-PCR machine) to make certain every well is sealed properly.
- 5. Calibrate the plates including the centrifuging trays. Centrifuge the RT-PCR plate in plate centrifuge for 2 minutes at 2500 rpm.
- 6. Run RT-PCR machine.
 - Turn on the RT-PCR machine and then the computer to establish the communication between the device and the computer.
 - The PCR machine and optical mode must be prewarmed at least 10 min in advance. Put your RT-PCR plate in the
 - When using the machine for the first time:
 - o Find the qPCR program file under C:\ Users\\Public\Public Documents\Bio-Rad\ CFX\ Plant Breeding
 - Here you find the PCR program needed. We used; CFX 2stepAMP+melt abe
 - Press 'OK' and go to Tab 'Plate'
 - o Here you select 'Create New' and fill in the 96 well scheme according your design.
 - o In 'Sample Type' Unknowns or NTC (No Template Control) can be selected.
 - o Select all sample wells and at 'Load' check the 'SYBR' box. Now you gave your samples a sample type.
 - Press 'OK' Save your File and press"Next>>'
 - Be sure that you select the proper qPCR machine.
 - o Press 'Open Lid'. Place your plate in the machine and press 'Close lid'
 - When this is done press 'Start Run'
 - O Save your data on the N drive or on a USB stick for analyzing your data later on.
 - o An overview will appear of the progress of the run.
 - o One RT-PCR run will take 100 minutes

Analysis of the data

Analysis of the Ct values and PCR products

- If the Bio-Rad iQ 5 program is not available on your computer you can install this by pressing START (Left Down your screen)→WUR→!Available Software and select CFXManager 3.1
- Activate the program.
- Open 'File' open 'Data File' and find your saved data file
- The Plots will appear in the program including a common threshold cycle (green line). The Ct's appear in the
 panel right under.
- A specific Threshold cycle is needed for your samples with one primer pair combination. Therefore select the
 wells done with one primer pair combination. You will see the Threshold cycle and Ct values will change a bit.
 This makes the values specific for the selected primer pair.
- Select the samples then copy the result table and paste this in a new worksheet called "raw data" in the Excel file you already made for the plate designed at the start of your experiment for further data analysis.
- Press Melt Curve/Peak. Two plots appear. One with the melting curve and the second is the first derivative of the
 first plot. If there is only one peak for each sample, that means there is a unique PCR product and the primers
 you designed are unique for the gene you selected.

Calculation Relative Gene Expression

- The raw data are copied in a worksheet 'Raw Data" Fill in the names in the column Identifier to name every well.
- Copy the columns "Well Identifier, Ct mean and Ct Std. Dev." from the raw data to another Excel sheet, and name it "RGE" (relative gene expression). Do not work on the raw data directly.
- The calculation starts now. It depends on the data obtained how to continue.
 - O Data without wild type material $\rightarrow \delta$ Ct calculation
 - Data with wild type material \rightarrow δδCt calculation.

δCt calculation:

This is how to calculate the Relative Gene Expression compared to a reference gene.

- The δ Ct is calculated as follows: δ Ct = Ct_(gene of interest) Ct_(reference)
- The RGE is calculated as: RGE= POWER(2;-δCt)*factor
 - o A factor (f.i.100.000) is needed to enlarge the figures needed for the Log calculation
- The Log 2 of the RGE is calculated as: LOG = LOG(RGE;2)
- Expression pattern can be depicted as follows:

$\delta\delta Ct$ calculation:

This is how to calculate the Gene Expression Transformant (Tr) compared to wild type(Wt) material.

- The $\delta Ct_{(Tr)}$ is calculated as follows: $\delta Ct = Ct_{(gene\ of\ interest,\ Tr)} Ct_{(reference,\ Tr)}$
- The δCt_(Wt) is calculated as follows: δCt = Ct_(gene of interest, Wt) Ct_(reference, Wt)
- $\delta\delta Ct = \delta CT_{(Tr)} \delta Ct_{(Wt)}$
 - o A factor is needed to enlarge the figures needed for the Log calculation
- The Log 2 of the RGE is calculated as: LOG = LOG(RGE;2)

calculate Standard deviation

Excel formula:

 $\mathsf{SEM} = \mathsf{SQRT}((\mathsf{POWER}(\mathsf{SD}_{\mathsf{ref}}/\mathsf{SQRT}(\mathsf{n}_{\mathsf{ref}});2)) + (\mathsf{POWER}(\mathsf{SD}_{\mathsf{GOI}}/\mathsf{SQRT}(\mathsf{n}_{\mathsf{GOI}};2)))$

Mathematical formula:

 $V(SD_{(ref)} / Vn_{(ref)})^2 + (SD_{(GOI)} / Vn_{(GOI)})^2$

Literature

"Kenneth J. Livak and Thomas D. Schmittgen . Analysis of Relative Gene Expression Data Using Real-Time Quantitative PCR and the $2^{-\delta\delta Ct}$ CT Method. METHODS **25**, 402–408 (2001)"

Quick and dirty DNA isolation

- 1. $\pm 1\mu$ l culture in an Eppendorf tube and centrifuge full speed for 10 minutes.
- 2. Remove the supernatant
- 3. Repeat step 1 and 2
- 4. Add 100 μl P1+RNAse. Resuspend pellet without vortexing
- 5. Add 200 µl P2. Mix by turning 4-6 times
- 6. Add 200 µl P3. Mix by turning 4-6 times
- 7. Centrifuge for 10 minutes (full speed)
- 8. Transfer supernatant to new Eppendorf tube
- 9. Add 300 μl isopropanol and mix by turning 4-6 times
- 10. Centrifuge 5 minutes at full speed
- 11. Remove supernatant
- 12. Wash with 200 μl 70 ethanol
- 13. Spin for 2 minutes at full speed
- 14. Remove as much from the ethanol as possible without touching the pellet
- 15. Let pellet dry
- 16. Resuspend pellet in 40 μl mili-Q

Sample preparation for GATC sequencing

- 1. Isolate plasmid using the miniptep kit from Qiagen (Use plasmids grown in E. Coli)
- 2. Measure concentration by nanodrop
- 3. take 2.5 μl of 80-100ng/ $\!\mu l$, depending on size of the insert
- 4. Add 2.5 μ l primer form a 10 μ M stock, add MQ to 5 μ l total volume
- 5. Place samples in the GATC sequencing service fridge

Used primers for sequencing level 1 and level 2 plasmid for CRISPR-Cas9:

Name	sequence	Forward or reverse?	Level 1 or level 2
			plasmid?
PDS 8340	GAACCCTGTGGTTGGCATGCACATAC	Forward	Level 1
PDS8534	TTTGTGATGCTCGTCAGGGG	Forward	Level 2
PDS8538	CCCGAGAATTATGCAGCATTTT	Forward	Level 2
PDS8536	TCATCAGTCAATTACGGGGCT	Forward	Level 2
AL717	GCTTGGCATCAGACAAACCGG	Reverse	Level 2

Agrobacterium transformation of Potato (S. andigena)

In vitro plant propagation using internodal explants

- 1. Stem node cuttings about 5 mm long are dissected from developed plants and placed onto CPM (clonal propagation media) for plant micropropagation.
- 2. Four week old plantlets are used as source of internodal explants for the transformation.

Transform Agrobacterium strain

- 1. Take the tubes out of -80°C (box Y9) and defrost them on ice. Each tube contains 40 μ l competent cells.
- 2. Add 110 µl cold MQ and use 50 µl per transformation.
- 3. add 1 µl plasmid DNA, mix, and put on ice for 15 minutes
- 4. transfer the mixture in a precooled cuvet.
- 5. Place it in the Gene Pulser and pulse it (1.8 kv, $25 \mu\text{F}$)
- 6. Add 1ml LB in the cuvet and transfer the contents to an eppendorf tube.
- 7. Incubate cell at 28°C for 3 hours on a shaker.
- 8. Plate cells on LB supplemented with kanomycin and rifampicin and incubate for 2 days at 28°C. select colonies and grow in 3 µl LB-Kan for 1 night at 28°C.

Preparation of Agrobacterium Culture

- 1. (DAY 1-2)To initiate the transformation take a loop full of culture from a transformed stock of *Agrobacterium* and streak it on LB solid medium containing the appropriate antibiotics
- (AGL1: Rifampicine + Gentamicine, but sure only Rif? + depending on plasmid). Incubate the plate at 28°C for two days.
- 2. (DAY 3) Inoculate a freshly grown single colony of *Agrobacterium* in 2 mL LB with specific antibiotics. Incubate the culture in a shaker (250 rpm) for one day at 28°C.
- 3. (DAY4) Add 1 mL of liquid *Agrobacterium* culture to 25 mL LB with specific antibiotics, and incubate at 28°C with shaking until OD600 0.6-0.8 (0.4 also good just later add less AIM).
- 4. Spin down the culture at 5000 rpm for 10 min and resuspend the pellet in 10 mL of LB (without antibiotics). Adjust to a final OD600 0.6.

Preparation of Explants for Agrobacterium Inoculation

- 1. (DAY 2-3) Excise internodal explants of 5 mm length from 4 week-old propagated *in vitro* plants, by removing the nodal segments with a sterile sharp scalpel blade on a sterile paper plate with humid paper towels.
- 2. Transfer the internodal explants onto Whatman filter paper placed over CIM medium in petri dishes. Add on the filter 2 ml of PACM media to keep wet.
- 3. Seal the plates using plastic wrap, and incubate them for 2 days in the dark in the Transformation fridge.

Agrobacterium Infection and Co-Cultivation

- 1. (DAY 4- afternoon- Day5/6) Transfer the internodal explants into a falcon tube with \sim 10 mL of the *Agrobacterium* suspension.
- 2. Incubate the explants with *Agrobacterium* cells for max 5 min. Shake the tubes gently during the incubation time.3. Remove the *Agrobacterium* suspension with a autoclaved drainer and transfer the explants back on the plate. Take care to remove as much Agro as possible. The same LB Agro suspension instead of being discarded can be used to transform another genotype with the same construct.
- 4. Seal the plates using plastic wrap, and incubate the explants for two days in darkness in the transformation fridge.

Selection of Transgenic callus and shoot induction

- 1. After two days of co-cultivation, transfer the explants to CIM containing 250 mg/mL cefotaxime, and 50 mg/L kanamycin (or hygromycin depending on the selection).
- 2. Seal the plates with plastic wrap and incubate them in the transformation fridge (25° C under fluorescent light at 60 μ EM/m2/s with a photoperiod of 16/8 h light/dark). Leave for 2 days.

Shoot Elongation and Root Induction

- 1. After 2 days transfer explants to shoot induction medium (SIM) plates supplemented with the same antibiotics. Change to new SIM every week until elongated shoots grow out.
- 2. After four weeks on SIM, transfer elongated shoots that are at least 2 cm to MS plates including antibiotics and claforal. This step should be repeated twice every time selecting healthy looking plants.

Culture Media

MS

Murashige and skoog media (+ vit) with 1.6% glucose

pH 5.8 (with KOH)

component	1 L	500 ml	250 ml
MS+ vit	4.4 g	2.2 g	1.1 g
saccharose	16 g	8 g	4 g
agar	8 g	4 g	2 g

PACM medium/ liter (Marian liquid media)

pH 6.5

component	1 L	500 ml	250 ml
MS + vit	4.4 g	2.2 g	1.1 g
Caseine hydrolysate	2 g	1 g	0.5 g
Saccharose	30 g	15 g	7.5 g
2.4 D (1mg/ml)	1 ml	0.5 ml	0.250 ml
Kitine (0.5 mg/ml)	0.5 ml	0.250 ml	0.125 ml

CIM (or R3B medium)

component	1 L	500 ml	250 ml
MS + vit	4.4 g	2.2 g	1.1 g
Saccharose 16%	16 g	8 g	4 g
NAA (1mg/ml)	5 ml	2.5 ml	1.25 ml
BAP (1mg/ml)	100 μΙ	50 μl	25 μΙ
plant agar	5.5 g	2.75 g	1.375 g
Claforan (200mg/ml)	1,25 ml	625 μl	312.5 μl
Kanamycin(100mg/ml)	500 μl	250 μΙ	125 μΙ

MS media +

5mg/L NAA (Naftalene Acetic Acid, Duchefa)

0.1 mg/L BAP (6-Bezylaminopurine, Duchefa)

250 mg/L claforan/ cefotaxime (only in the + antib version)

50 mg/L kanamycin or 3 mg/L hygromicin (only in the + antib version)

plant agar: Duchefa (kanamycin) or gelrite Duchefa (hygromicin)

SIM

component	1 L	500 ml	250 ml
MS + vit	4.4 g	2.2 g	1.1 g
Saccharose 16%	16 g	8 g	4 g
NAA (1mg/ml)	20 µl	10 μl	5 μΙ
GA3 (1mg/ml)	20 μΙ	10 μl	5 μΙ
plant agar	5.5 g	2.75 g	1.375 g
Zeatinriboside (1mg/ml)	2 ml	1 ml	0.5 ml
Claforan (200mg/ml)	1,25 ml	625 μl	312.5 µl
Kanamycin(100mg/ml)	500 μΙ	250 μl	125 μΙ

MS media +

0.02 mg/L NAA (1-Naphtalene Acetic Acid, Duchefa)

0.02 mg/L GA3, Duchefa or SIGMA (fresh prepared)

2 mg/L Zeatinriboside (fresh prepared)

250 mg/L claforan

50 mg/L kanamycin or 3 mg/L hygromicin

5.5 gr/L plant agar Duchefa

The hormones are added prior to autoclaving, but zeatin, and all antibiotics are added after autoclaving, when the temperature of themedium has dropped to 55°C. Sterile medium is poured into petri dishes in a laminar flow hood. All media (liquid or solid) can be stored for several weeks at 4°C, but media with the antibiotics must be fresh.

Autoclave beforehand:

at least 1 drainer/ transformation Round filters Paper trays Pieces of paper trays Forcepts

Tips

blades

Media

Glass pots for drainer

References

Martha L. Orozco-Cárdenas, Venkateswari Jagantha Chetty, and Javier Narváez-Vásquez Plant Transformation Research Center, University of California Riverside, CA 92507-0124

Appendix B: additional figures

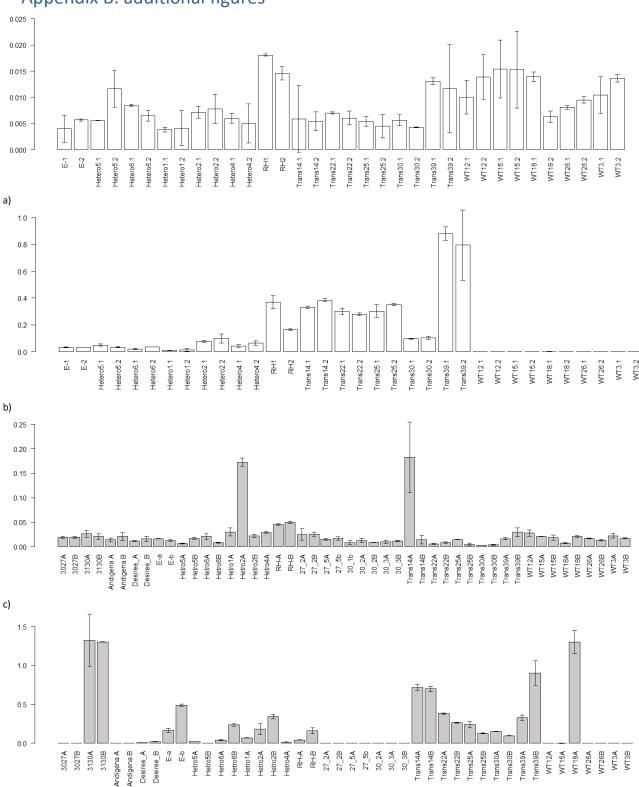


Figure 1: Relative gene expression per sample

CDF1 and SP6A expression for non-drought stressed plants (a and b respectively) and drought stressed plants (c and d respectively), for each individual plant. The error bars show the standard error of the mean (SEM).

Appendix C: details on the vector components

Below a table with the details on the vectors used in experiment 3.

abbreviation	Add gene name	Letter code	CZN	Needed for level	resistance	concentration	description	BP	Ng for 60fm
966	pICH86966	А	2774	-	Kanamycin	50	Template used for initial PCR	6552	120
009	pICSL01009	В	2772	1	Spectinomycin	83	Plasmid containing AtU6	2323	42.6
751	pICH47751	С	2769	1	Carbenicillin	424	Plasmid to clone guide, position 1	4968	92
761	pICH47761	D	2767	1	Carbenicillin	275	Plasmid to clone guide, position 2	4968	92
723	pAGM4723	L	2766	2	Kanamycin	170	Binary vector for transfomation	12919	237.2
732	pICH47732	G	2770	2	Carbencillin	111	NOS::NPT2 OCST	6234	114
742	pICH47742	Н	2771	2	Carbencillin	225	45Sp::Cas9- Nost	9623	176
780	pICH41780	J	2773	2	Spectinomycin	150	Linker for 2 guide reaction	3318	61

Appendix D: glycerol stocks

From several bacteria cultures glycerol stocks were made, which were stored in freezer G, in a box labelled: 'CRISPR CAS, glycerol stocks. Lisette, Marian Oorwijn, Abe, Christian. Feb 2017'. Below, a table is given with information on these bacteria cultures.

Name	level	date	Bacteria	Description
CRISPR_3A	1	22/3/17	E. Coli	Level 1 plasmid with sgRNA3
1.6 CRISPR	1	31/3/17	E. Coli	Level 1 plasmid with sgRNA1
L2.2	2	16/5/17	E. Coli	Level 2 plasmid with sgRNA1 and sgRNA3
L2.7	2	16/5/17	E. Coli	Level 2 plasmid with sgRNA1 and sgRNA3