Production and characterization of transgenic Arabidopsis and soybean (*Glycine max* (L.) Merrill.) plants over-expressing oryzacystatin I (OC-I)

By

MATOME EUGENE MAKGOPA

Submitted in accordance with the requirements for the degree of
PHILOSOPHIAE DOCTOR

Forestry and Agricultural Biotechnology Institute (FABI)

Department of Plant Science

In the

Faculty of Natural and Agricultural Sciences

University of Pretoria

Supervisors:

PROF. K. KUNERT

PROF. C.H. FOYER

March, 2014

DECLARATION

I, Matome Eugene Makgopa, declare that this work is my original work and has not been submitted elsewhere. Furthermore, no plagiarism exists in the chapters and conclusions reached.

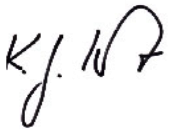
Date……

Signed…….....

MATOME EUGENE MAKGOPA

This thesis has been submitted for examination with my approval as the University supervisor.

Date……24/3/2014……



Signed…….....

PROF. KARL KUNERT

DEDICATION

I dedicate this thesis to my fiancé Marietta Muller. I could go on and on describing all that you have done for me but I am sure you know how much I appreciate having you in my life so I will be brief. You are the love of my life, my pillar, MY EVERYTHING. ICH LIEBE DICH VON GANZEM HERZEN.

ACKNOWLEDGEMENTS

First of all, I would like to thank my entire family for all their love, care and support. My mother, Pauline Makgopa, and father, Moses Makgopa, for imbibing discipline in me from an early age and teaching me how to treat others as I would like to be treated. Thank you to my brother, Martin Makgopa, and sister, Beaulah Makgopa, for all their support and fun moments.

I would like to express my deepest gratitude to my supervisor Prof. Karl Kunert for all his words of encouragement, guidance and advice. You were not only a supervisor to me but also my mentor and someone I one day hope to be half as good as professionally. I would also like to thank you for creating the opportunity for me to carry out my research in the UK. Never did I think I will ever live and study abroad and for that I am FOREVER thankful to you.

I would also like to express profound gratitude to my co-supervisor, Prof. Christine Foyer, for moulding me to be a critical thinker and for equipping me with the scientific skills I have gained. Moreover, thank you for accepting me to carry out my research at the University of Leeds (UK) and exposing me to a different environment of scientific research which has been immensely value in the completion of my PhD study.

I would like to thank my lab mates in the UK, Daniel Schnaubelt, Gloria Comadira, Marian D. Quain and Belén Márquez García, Guillaume Queval, Barbara Kapinska and Simon Driscoll, for the warm welcome, providing a friendly working environment and all the outings.

A special thank you to the winning team, Marian D. Quain and Belén Márquez García, I really appreciate all your help and support.

I am also thankful to my SA lab mates, Stefan van Wyk, Priyen Pillay, Rynhard Smit, Magdeleen du Plessis, Berhanu Fenta, Tintswalo Maluleke for always being willing to help and for all the good times.

Finally, I am forever thankful to The EU LEGIM project which funded my PhD study in the UK and the NRF for funding this PhD study in my stay in SA. The funding from both organisations is highly appreciated.

SUMMARY

In legumes, drought causes early senescence due to loss of the symbiotic relationship between the plant and the rhizobacterium. Senescence is characterized by increases in proteolytic enzymes required for protein recycling to other plant tissues. Transgenic soybean plants over-expressing a cysteine protease inhibitor (OC-I) were successfully generated and characterized. Plants of transgenic lines had differential transgene expression. Transgenic plants had lower protease activity determined by both an in-gel assay and in a fluoremetric assay using SDS-PAGE and fluorogenic protease substrate, respectively. Transgenic Arabidopsis plants over-expressing OC-I, generated by floral dip method, were more drought-tolerant compared to non-transgenic Arabidopsis plants. Seeds of different transgenic soybean lines had a lower germination rate and these transgenic lines had fewer leaves and shorter stems. Under drought stress, plants of transgenic lines performed better than wild-type non-transgenic plants with CO_2 assimilation (photosynthesis) and better instantaneous water-use efficiencies (IWUE). In particular, plants of one transgenic line (line 57) appeared to be more drought-tolerant when compared to plants of all other tested lines. These transgenic plants retained more soil water and also had fewer leaves when compared to wild-type non-transgenic plants. Results obtained in this study have provided evidence that preventing cysteine protease activity by over-expressing a protease inhibitor causes phenotypic changes of the plant demonstrating an important role of cysteine proteases in plant growth and development and plant stress. Future work will focus on identifying these OC-I sensitive proteases and investigating their individual function in plant growth and development and stress.

THESIS COMPOSITION

Chapter 1 of this thesis provides a summary of the importance of grain legumes and the role plant proteases and their inhibitors play during plant development processes and their involvement in plant stress responses is also discussed. An overview of previous and current research on the effect of drought stress on legumes and Arabidopsis is also summarized. The rationale, aims and objectives of the study are presented at the end of the chapter. **Chapter 2** reports about designing various vectors for soybean transformation. This chapter deals in particular with cloning of protease inhibitor coding sequences into plasmids of choice and verification of the cloning process by restriction enzyme analysis, PCR and sequencing. **Chapter 3** focuses on the characterization of the transgenic soybean plants for presence, gene expression and protein expression of the transgene in the soybean genome using PCR, quantitative real-time PCR and Western blotting, respectively. Protease activity profiles were also established by in-gel gelatine-containing SDS-PAGE and by protease activity measurement following hydrolysis of the specific cathepsin L-like substrate Z-Phe-Arg-MCA to liberate fluorescent 7-amino-4-methylcoumarin with a fluorometer. **Chapter 4** reports on the transformation and characterization of transgenic Arabidopsis plants under well-watered and drought conditions to find any to morphological or phenotypic changes. **Chapter 5** reports on the phenotypic characterization of transgenic soybean plants under well-watered and drought conditions. In particular, this chapter deals with the identification of easily measurable morphological traits, such as germination rate, number of leaves, biomass production, as well as physiological traits. **Chapter 6** summarizes the findings and relevant information generated from this PhD study and provides an outlook of possible future actions. This is finally followed by the **reference list** of the citations used in this thesis, **supplementary data** and **annexures**.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABBREVIATIONS

µg	Microgram
µL	Microlitre
A_{max}	Maximal photosynthesis
A_n	Net photosynthesis
CaMV	Cauliflower mosaic virus
CP	Cysteine protease
CPI	Cysteine protease inhibitor
DMSO	Dimethyl sulfoxide
E-64	Trans-epoxysuccinyl-L-leucylamido-(4-guanidino) butane
EDTA	Ethylenediaminetetraacetic acid
EtOH	Ethanol
Fv/Fm	PSII maximum quantum yield
FP	Forward primer
GSH	Glutathione (reduced)
GST	Glutathione S-transferase
GSTF	GST fusion
hrs	Hours
H_2O	Water
IPTG	Isopropylthio-β-galactoside
LB	Luria-Bertani
mM	Millimolar
Mmt	Million metric tonnes
NC	Negative control
nm	Nanometer
O.D.	Optical density
O/N	Overnight

OC-I	Oryzacystatin I
PC	Positive control
PCD	Programmed cell death
PCR	Polymerase chain reaction
qRTPCR	Quantitative real-time PCR
RP	Reverse primer
rpm	Revolutions per minute
RT	Room temperature
rubisco	Ribulose-1,5-bisphosphate carboxylase/oxygenase
s	Seconds
sd	Sterile distilled
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SP	Serine protease
SPI	Serine protease inhibitor
TAE	Tris-acetate EDTA
TLCK	N-alpha-tosyl-L-lysinyl-chloromethylketone
v/v	Volume per volume
w/v	Weight per volume
WT	Wild-type
x	Any amino acid
x*g*	Times the force of gravity
X-gal	5-bromo-4-chloro-3-indolyl-beta-D-galactopyranoside
Z-Phe-Arg-MCA	Z-Phenylalanine-Arginine-7-amido-4-methylcoumarin

Chapter 1

Introduction

1.1 The importance of soybean

Soybean [*Glycine max* (L.) Merrill.] is one of the most important leguminous seed crops as it is an economic source of both oil and protein. Soybean is widely used in human and animal food, plastics adhesive, as well as in a variety of items of processed food (Mello-Farias and Chaves, 2008). Global production of soybean in 2012 was around 268 million metric tons (mmt). The country with the greatest production was the Brazil with 83.5 mmt, followed by USA (82 mmt), Argentina (51.5 mmt) and China (12.6 mmt) (Soystats 2012). The consumption of soybean-containing products is increasing worldwide because of the reported beneficial effects including lowering of cholesterol, prevention of cancer, diabetes and obesity, and protection against bowel and kidney diseases (Friedman and Brandon, 2001). It is also used as a source of nitrogen in agriculture particularly in Africa (Keyser and Li, 1992). The ability of soybean to fix nitrogen is due to the symbiotic relationship between soybean and the soil bacteria rhizobia resulting in the formation of nodules (Puppo et al., 2005).

Stress on the plant, or the bacteria, may abolish this relationship resulting in early senescence of nodules. Early senescence is characterized by changes in proteolytic enzymes involved in various processes in the plant's responses to stress (Demirevska et al., 2008, Mosolov and Valueva, 2008). Drought is one such stressor known to induce proteolysis and constitutes about 40% of yield loss (Chaves et al., 2003, Wang et al., 2003). Understanding the mechanism of stress tolerance along with the genes involved in stress signalling network is therefore important for crop improvement.

1.2 Proteolyic enzymes

Proteolytic enzymes are hydrolytic enzymes involved in many aspects of plant growth and development. Their mode of action can be classified into two categories, limited proteolysis and unlimited proteolysis (Fan and Wu, 2005). In limited proteolysis, a protease cleaves one or a limited number of peptide bonds of a particular protein resulting in activation of the protein e.g. pro-hormones to hormones. This type of proteolysis is necessary for correct protein assembly and it controls the activity of enzymes, regulatory proteins and peptides. In unlimited proteolysis, proteins that are damaged, mis-folded and potentially harmful are degraded into their amino acid constituents. This degradation is important because it provides free amino acids required for the synthesis of new proteins. Proteases play a key role in the regulation of plant growth, development and defence as they are involved in all stages of the plant life cycle starting from remobilization of storage proteins during seed germination to the initiation of cell death and senescence (Solomon et al., 1999, Schaller, 2004, Fan and Wu, 2005).

1.2.1 Classification of proteolytic enzymes

Peptidases, peptide hydroxylase, proteases, proteinases and proteolytic enzymes are similar names given to a group of enzymes that hydrolyze peptide bonds (Rawlings and Barrett, 1999). According to the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology (NC-IUBMB, 1992), the term peptidase refers to all enzymes that hydrolyze peptide bonds. Peptidases can be grouped into two groups: exo-peptidases and endo-peptidases. Exo-peptidases are enzymes that cleave one or more amino acids from the

N- or C-terminal and those that cleave the internal peptide bonds are termed endo-peptidases (Rawlings and Barrett, 1999, Grudkowska and Zagdanska, 2004, Fan and Wu, 2005).

The term “protease”, which encompasses both exo-peptidases and endo-peptidases, will be used throughout this thesis outline. Proteases have been divided into 7 mechanistic classes by the International Union of Biochemistry. These include the cysteine, serine, aspartic, metallo, threonine, and glutamate proteases and the newly identified asparagine peptide lysases (Rawlings et al., 2011, Deu et al., 2012) (Figure 1.1A). Their mode of action involves polarization of the carbonyl group of the substrate peptide bond by stabilizing the oxygen in an oxyanion hole. This makes the carbon atom more vulnerable to attack by an activated nucleophile (Figure 1.1B). Cysteine and serine proteases along with their inhibitors are discussed in more detail.

A

Catalytic class	Nucleophile
Cysteine proteases	Cys-His
Serine proteases	Ser-His
Metalloproteases	H_2O-Me^{2+}
Aspartic proteases	H_2O-Asp
Threonine proteases	Thr
Glutamate proteases	H_2O-Glu
Asparagine proteases	Asn

B

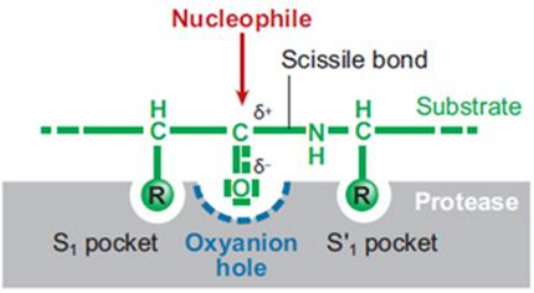


Figure 1.1 Classification and cleavage mechanism of the 7 major catalytic classes of proteases. (a) The major differences between the catalytic classes are the nature of the nucleophile. Threonine, cysteine and serine proteases use a Thr, Cys or Ser residue as nucleophile, respectively. Metallo-, aspartic and glutamate proteases use water as nucleophile, activated by electrostatic interactions with the metal ion (Me^{2+}) or aspartate (Asp), respectively. The seventh and newest protease class, the asparagine peptide lyases, uses an asparagine residue as the nucleophile. (b) The substrate protein (blue) binds via amino acid residues (R) to the substrate binding site of the protease (gray) by interacting with substrate (S) pockets of the enzyme. The scissile peptide bond (red) is adjacent to a carbonyl group. The carbonyl carbon is vulnerable to nucleophilic attack (Modified from van der Hoorn, 2008).

1.2.2 Cysteine proteases

The most studied plant proteases belong to the cysteine class. They are found in the vacuoles and are involved in the mobilization of storage proteins during seed germination (Oliveira et al., 2003). Proteases are classified based on their structural and evolutionary relationships and theses enzymes are grouped into families and clans (Rawlings and Barrett, 1999, Barrett et al., 2001). A MEROPS database (http://merops.sanger.ac.uk) of proteases has been developed which includes listings of all peptidases sequences from different families and clans. The database is continuously being updated with the addition of new members (Dubey et al., 2007).

1.2.2.1 Families of cysteine proteases

Cysteine proteases (CPs) family are designated with the prefix C and comprises of six major families: the papain family, calpains, clostripains, streptococcal cysteine proteases, viral cysteine proteases and caspases (Dubey et al., 2007). Most plant proteases belong to the papain and legumain families. Families C1, C2 and C10 can be described as papain-like, C3, C4, C5, C6, C7, C8, C9, C16 C18, and C21 are represented in viruses while C11, C15 and C20 are from bacteria (Grudkowska and Zagdanska, 2004, Dubey et al., 2007). Family C1 has been subdivided into subfamily C1A and C1B. Subfamily C1A consists of proteases that contain disulfide bridges and accumulate in vesicles, the vacuole, or the apoplast while subfamily C1B consists of proteases that lack disulfide bridges and accumulate in the cytoplasm (van der Hoorn, 2008).

1.2.2.2 Structural features of plant C1A family proteases

Plants only have C1A. Approximately 30 papain-like proteases, subdivided into 8 subfamilies, have been identified in Arabidopsis (van der Hoorn, 2008). Papain-like proteases are the most well-characterized cysteine proteases (Grudkowska and Zagdanska, 2004). Plant papain-like enzymes (C1) are synthesized as pre-propeptides of 40-50 kDa that undergo proteolytic processing of the pre- and pro-domains to yield the mature, active protease of 22-35 kDa (Beers et al., 2004, Grudkowska and Zagdanska, 2004). C1A proteases are characterized by presence of a catalytic dyad, Cys25 and His159 and an Asn175 residue required for proper orientation of the His side chain (using papain numbering, Figure 1.2A) (Beers et al., 2004). Arabidopsis C1A enzymes possess an N-terminal prodomain containing the non-contiguous Glu/Asp-Arg-Phe/Tyr/Leu-Asn-Ile/Ala/Val-Asn/Gln (ERFNIN) signature also found in human cathepsin L (Beers et al., 2004). This unique feature has been used to distinguish between cathepsin L and B as cathepsin B lacks this signature. Some proteases carry a KDEL signal at their C-terminal for retention into the endoplasmic reticulum (ER) while others carry a signal at the N terminal for vascular targeting. Granulin-like domains are also found at the C-terminal of some proteases (Grudkowska and Zagdanska, 2004).

C1A proteases have been classified based on their closest animal counterparts: cathepsin B, L, F and H (Martinez and Diaz, 2008). Using comparative genomic analysis of C1A proteases across different taxons of the plant kingdom, the authors found cat L-like proteases to have a striking pattern across taxon i.e. first appearing in algae and having increased in moss followed by a sharp increase from spike moss to angiosperms. A previous phylogenetic assembly of 138 plant C1A proteases revealed that these proteases can be classified into 8 subfamilies (C1A-1-C1A-8) based on the conserved regions within genes (Beers et al., 2004).

More recently, a phylogenetic analysis of the pro-protease domain from a broader sequence database of C1A proteases has identified 9 subfamilies of these proteases (Figure 1.3) (Richau et al., 2012). Subfamilies 1-6 cathepsin L-like that are closely related, subfamily 7 are representative of cathepsin F-like proteases, subfamily 8 the cathepsin H-like and subfamily 9 are cathepsin B-like proteases which are distantly separated from the other proteases (Figure 1.3) (Richau et al., 2012).

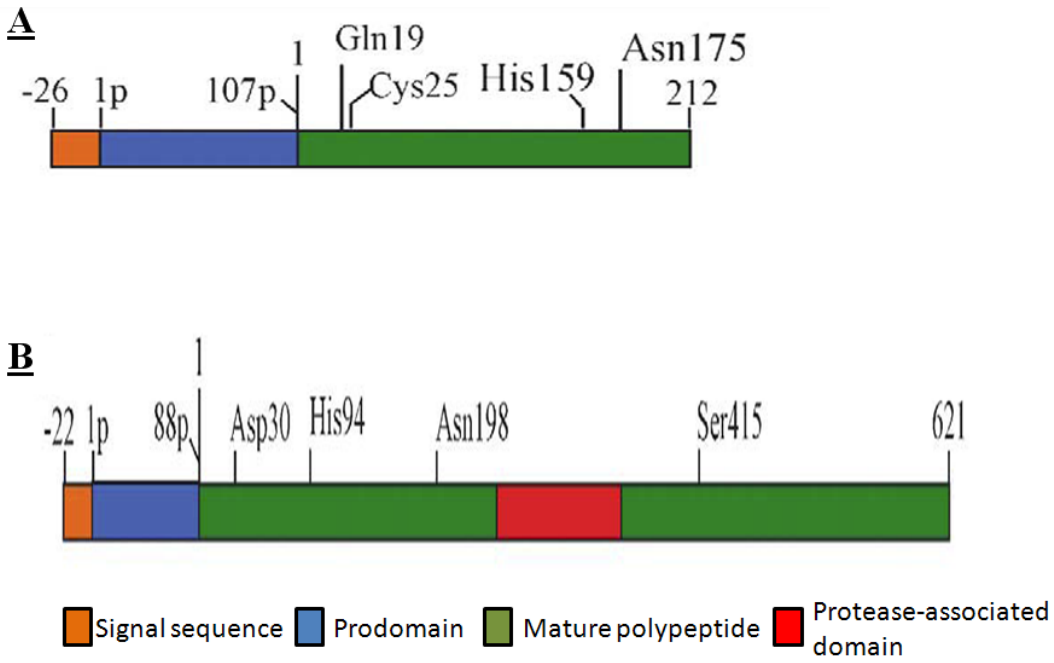


Figure 1.2 Schematic representation of the structure of pre-propapain and pre-procucumisin. (A) Signal sequences are indicated by negative numbers, and each prodomain is indicated using numbers followed by ‘‘p’’. Numbering for the mature enzyme indicates positions of the N and C termini. The active site residues Cys25, His159 and Asn175 are shown along with the conserved Gln19 residue. (B) Residues of the catalytic triad, Asp30, His94 and Ser415, are labelled, as is a conserved Asn (Asn198) involved in oxyanion hole stabilization. The core conserved region of the protease-associated (PA) domain, predicted to be involved in protease-substrate interaction (Modified from Beers et al., 2004).

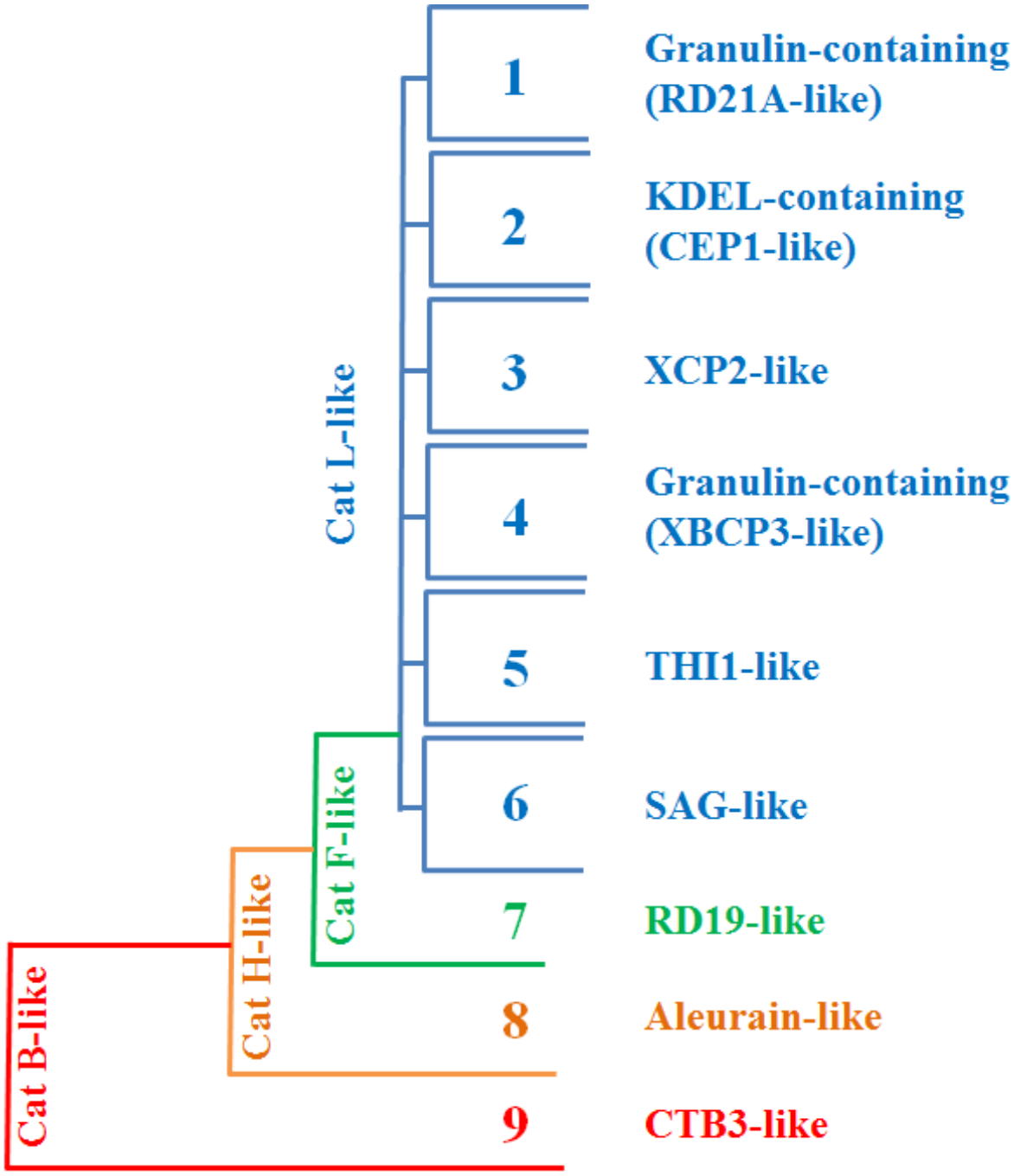


Figure 1.3 A simplified phylogenetic tree representative of the 9 subfamilies of papain-like cysteine proteases. The left text represents the cathepsin class while the right texts are examples of some of the Arabidopsis C1A proteases that have been studied (Modified from Beers et al., 2004; Martinez and Diaz, 2008; Richau et al., 2012).

1.2.2.3 Attributed functions of cysteine proteases

Cysteine proteases (CPs) have been reported to be involved in a number of functions such as leaf senescence and programmed cell death (PCD). The cysteine protease, senescence-associated gene 12 (SAG12), was markedly expressed during leaf senescence (Pontier et al., 1999, Belenghi et al., 2003, Beyene et al., 2006). Solomon et al. (1999) demonstrated the involvement of CPs in PCD through ectopic expression of a cystatin in suspension-cultured soybean Williams 82 cells that inhibited CP activity and blocked PCD triggered by either an avirulent strain of *Pseudomonas syringae* pv *glycinea* or directly by oxidative stress. Constitutive expression of a *AtCys1*, a cysteine protease inhibitor, suppresses PCD triggered by either avirulent pathogens or oxidative and nitrosative stresses in both *A. thaliana* suspension cultures and in transgenic tobacco (Belenghi et al., 2003). Another function attributed to CP is their involvement in fruit ripening (Fan et al., 2009). The authors identified and cloned a cysteine protease, designated *CsCP*, belonging to the aleurain group of the C1 family. Expression increased with postharvest peel pitting of citrus fruits. The cysteine protease papain has also been found to inhibit growth against Lepidoptera larvae. Inhibition of the activity of papain by leaf painting with E-64 abolished the toxicity thus highlighting the importance of the CP in providing protection against herbivorous insects (Konno et al., 2004). A plant disease resistance (R) protein, required for *Cladosporium* 3 (Rcr3), which encodes a cysteine protease has been shown to be important for the function of *CF-2* gene (Kruger et al., 2002). This is a *Lycopersicon pimpinellifolium* gene bred into cultivated tomato (*Lycopersicon esculentum*) that confers resistance to the pathogen *Cladosporium fulvum.* Compensation experiments to mutants of the *rcr3* results in restored *Cf-2*-dependent resistance in tomato leaves (Krüger et al., 2002).

1.2.3 Serine proteases

Another class of proteases are the serine proteases (SPs). Serine proteases are the largest class of proteases with more than 200 members divided into 14 families (van der Hoorn, 2008). These families are divided into 9 clans which are not evolutionarily related. The largest serine protease families in plants are the S8, S9, S10 and S33 each constituting about 60 members (van der Hoorn, 2008).

1.2.3.1 Structural features of plant S8 family proteases

Subtilases (family S8, clan SB) are well-known representatives from serine proteases (Beers et al., 2004). Similar to cysteine proteases, these enzymes are encoded as pre-proproteins that are proteolytically processed into mature proteins. The subtilisin-like proteases consists of a conserved catalytic triad in the order of Asp30, His94 and Ser415 (mature cucumisin numbering, Figure 1.2B) (Beers et al., 2004). An interesting feature of these enzymes is the high sequence variability within their sequence i.e. virtually one or more residues can be replaced by distinct residues with the exception of the conserved catalytic triad (Antao and Malcata, 2005).

1.2.3.2 Attributed functions of serine proteases

Serine proteases have been shown to have similar functions to CPs i.e. they are also involved in the remobilization of resources in germinating seeds (Antão and Malcata, 2005). Serine protease activity was increased during post-germinative seedling growth of the major storage organ of the white spruce (*Picea glauca*) seed, megagametophytes (radical, cotyledons and

hypocotyl) (He and Kermode, 2003). In the late post-germinative stage of the megagametophytes (> 25 mm) the activity of serine and cysteine protease activities could be discriminated with the PMSF and E-64 inhibitors using zymogram. Large quantities of mucilage are deposited into the outer wall of the seat coat during seed development. Loss-of-function mutants of the *AtSBT1.7* allele encoding one of 56 subtilisin-like serine proteases do not release mucilage resulting in altered seed development (Rautengarten et al., 2008). It has also been demonstrated that Deg1, a SP, is involved in the repair of the photosystem II from photoinhibition in Arabidopsis (Kapri-Pardes et al., 2007). The authors demonstrated that RNA interference knock-down of the Deg1 in Arabidopsis plants resulted in transgenic plants that were smaller than wild-type plants with early flowering and increased sensitivity to photo-inhibition.

1.3 Protease inhibitors

Proteases are found in almost all organisms and are encoded for by approx. 2% of all genes (Barrett et al., 2001). Proteases are important for the survival of all organisms but they are also potentially damaging thus they have to be tightly regulated by their inhibitors (Rawlings et al., 2004). Protease inhibitors (PIs) refer to naturally occurring proteins that specifically inhibit the activity of proteases. Inhibition occurs when the active-site of the substrate region of a protease binds to the corresponding substrate-like region (reactive site) on the surface of the inhibitor thus blocking the site for the actual substrate (Habib and Fazili, 2007). Inhibitors have been found for each of the classes of proteases however; only those of cysteine and serine proteases will be discussed further.

1.3.1 Cysteine protease inhibitors (cystatins)

Proteins that specifically inhibit the activity of papain and related cysteine proteases (cathepsin B, H and L, ficin, bromelain) are referred to as cystatins (Oliveira et al., 2003). These proteins are evolutionary related forming the cystatin super family. The super family comprises of three family members that are animal cystatins and one super family from plant cystatins (phytocystatins) representing almost all plant cysteine protease inhibitors (Oliveira et al., 2003, Grudkowska and Zagdanska, 2004, Habib and Fazili, 2007). The super families are classed as follows:

Family 1 comprises of stefins (stefin A and B). They are unglycosylated inhibitors of ~11 kDa, they lack a signal sequence and disulfide bonds and are expressed intra-cellular. Stefins contain the conserved pentapeptide sequence, Gln-Val-Val-Ala-Gly.

Family 2 is known as cystatin family. They have molecular masses in the range of ~13–14 kDa, contain signal sequence and two disulfide bonds at the carboxyl terminus of the molecule. All members of this family contain a conserved tripeptide with the sequence Phe-Ala-Val near the C-terminus and a conserved dipeptide, Phe-Tyr, near the N-terminus.

Family 3 is referred as kininogens. They are glycosylated inhibitors of ~88–114 kDa and have three family-2 cystatin domains. These members have the same conserved pentapeptide sequence as family 1.

Family 4 (phytocystatins). This family of proteins is characterized by features common to the family-1 and -2 cystatins. They lack disulfide bonds, have molecular mass ~11.5 kDa. It

has a highly conserved region of G58 residue, Glu-Val-Val-Ala-Gly (QVVAG) motif and a PRO-TRP (PW) motif. Another feature of the phytocystatin family is a consensus sequence ([LVI]-[AGT]-[RKE]-[FY]-[AS]-[VI]-x-[EDQV]-[HYFQ]-N) that conforms to the α-helix structure.

Phytocystatins may be grouped into different but similar evolutionary groups based on their structural and functional features. This grouping is mainly due to mutations that may have occurred overtime resulting in proteases with different and/or duplication of single domain cystatins as illustrated in Figure 1.4 (Benchabane et al., 2010). An overview of known phytocystatins, their relationship amongst themselves and other species has been obtained from sequence databases: EMBL: European Molecular Biology Laboratory, NCBI: National Centre for Biotechnology Information, PIR: Protein Information Resource, SP: SwissProt and GB: GeneBank and a phylogenetic tree generated. The tree is shown in Figure 1 (supplementary data). Five clades were distinguished with the rice OC-I grouping in clade 1 and the soybean phytocystatins in clade 2. Cystatins (~20) from soybean have been further identified using bioinformatics tools. Sequence alignment and a phylogenetic tree was generated using ClustalW with the inclusion of the rice cystatin (OC-I) to show the relationship between the soybean cystatins and the rice cystatin as shown in Figure 2 (supplementary data).

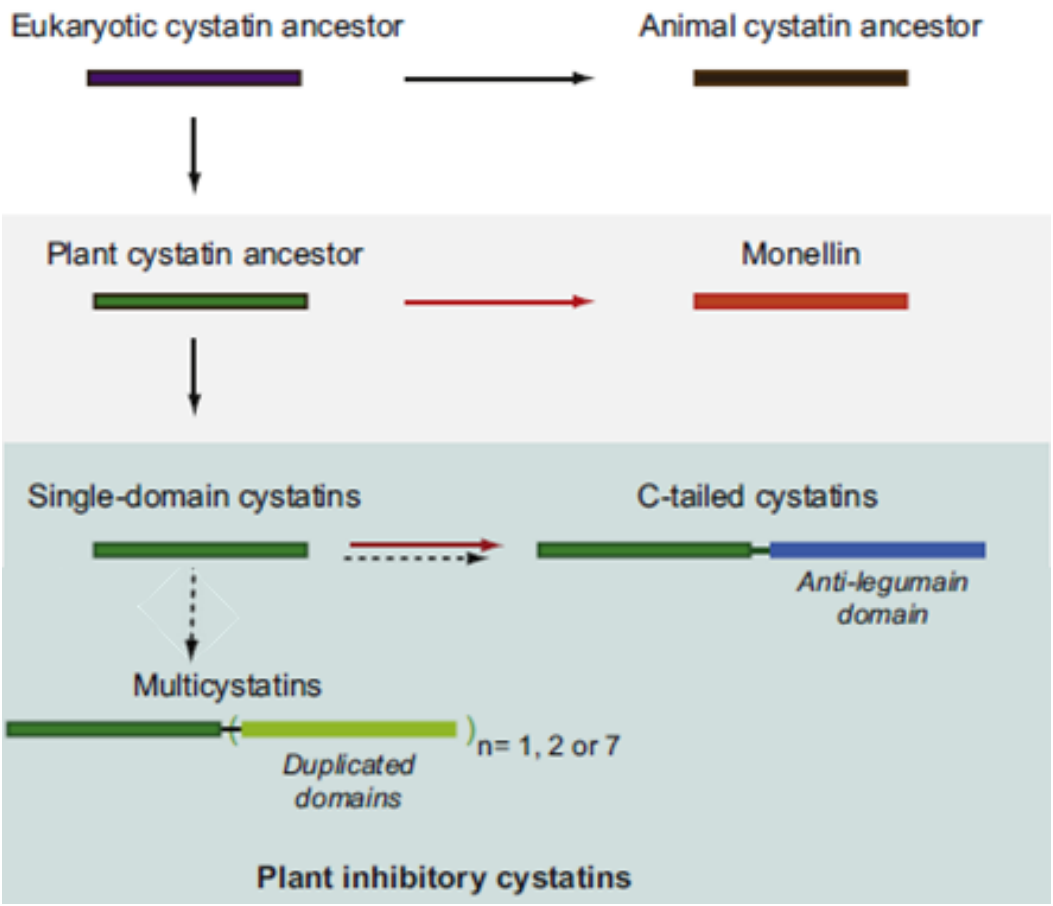


Figure 1.4 Illustration of a possible evolutionary scheme for plant cystatins. Plain arrows indicate mutations-based evolutionary processes (including adaptive evolution). Dashed arrows indicate gene duplication events leading to multi-domain cystatins. Red arrows underline functional diversification, along with the loss of inhibitory activity against papain-like Cys proteases (Modified from Bechambane et al., 2010).

Phytocystatins have been shown to be involved in the regulation of endogenous proteases activities in seeds during seed development (Oliveira et al., 2003). Hwang et al. (2009) showed that cysteine protease activity was inhibited by the cystatin AtCYS6 in transgenic Arabidopsis. Oryzacystatins, first isolated plant cysteine protease inhibitors from rice, have been shown to inhibit the activities of oryzains α, β and γ cysteine proteases that are produced during seed germination (Watanabe et al., 1991). Solomon et al. (1999) showed that

phytocystatins were required for the regulation of PCD through inhibition of the activity of CPs as outlined in section 1.2.2.3. A further study supporting the involvement of cystatins in PCD (also mentioned briefly in section 1.2.2.3) was carried out using a cystatin from *A. thaliana*, *AtCYS1*, which is constitutively expressed in roots and in developing siliques. AtCYS1 is strongly induced in leaves by wounding, challenge with an avirulent pathogen or nitric oxide (NO). The over-expression of *AtCYS1* blocks cell death activated by either avirulent pathogens or by oxidative and nitrosative stress in both *A. thaliana* suspension cultured cells and in transgenic tobacco plants (Belenghi et al., 2003).

Another function attributed to phytocystatins is that of providing protection against insects and nematodes (Urwin et al., 2001, Cowgill et al., 2002). Urwin et al. (2001) showed that transgenic potato plants were more resistant to nematodes such as the potato cyst nematode *Globodera pallida* through over-expression of a cysteine protease inhibitor. Artificial diets of the oryzacystatin-I and chicken egg white cystatin significantly reduced the survival and growth of the peach potato aphid *Myzus persicae*. However; these effects were not observed in the transgenic lines of both cystatins possibly due to the use of a 35S promoter rather than a phloem specific promoter (Cowgill et al., 2002). Members of the coleoptera order are tuber-, seed-, root- and leaf-eating insects that utilize plant storage proteins with cysteine proteases in their gut. Two plant cystatins, oryzacystatin I (OC-I, or OsCYS1) and papaya cystatin (CpCYS1), caused a reduced growth rate of the banana weevil *Cosmopolites sordidus* (coleopteran) *in vitro* using protein extract from the midgut with a fluorescent assay and a gelatine-containing SDS-PAGE (Kiggundu et al., 2010). Moreover, reduction in growth rate was also observed when the banana weevils were artificial fed cystatin-infiltrated banana stem disks. Phytocystatins further inhibit growth and have lethal effects against members of the coleopteran (*Callosobruchus chinensis*) and heminoptera (*Riptortus clavatus*) order

(Leple et al., 1995, Kuroda et al., 1996). Kuroda et al. (1996) found that by adding low concentrations of oryzacytatins I and II to the diet of different insect bean pests caused growth inhibition against insect pests from the coleoptera and hemiptera order, higher concentrations were fatal. In another study, it was shown that transgenic potato plants expressing a barley cystatin (HvCPI-1, a potent inhibitor of papain and cathepsin B) resulted in significantly reduced larval weight gains of the Colorado potato beetle (Álvarez-Alfageme et al., 2007). Further, cystatin isolated from chestnut seeds inhibited digestive cysteine proteases from two economically important agricultural pests, *Tribolium castaneum* and the mite *Dermatophagoides farinae*, *in vitro* (Pernas et al., 1998). HvCPI-6 artificial administration to the aphid species *Acyrthosiphon pisum* was toxic against the aphid (Carrillo et al., 2011). The authors also showed that over-expression of the HvPCI-6 in transgenic Arabidopsis resulted in delayed developmental time to adulthood. In addition to protection against insect pests, cystatins can be applied against viruses. Transgenic tobacco plants expressing a cysteine protease inhibitor (oryzacystatin I) had enhanced resistance against different poty-viruses namely, tobacco etch virus (TEV) and potato virus Y (PVY) and possibly viruses whose replication involves cysteine protease activity (Gutierrez-Campos et al., 1999).

Studies have also been carried out addressing the functions of cystatins under abiotic stress (Van der Vyver et al., 2003, Zhang et al., 2008, Hwang et al., 2010b). Van der Vyver et al. (2003) demonstrated that OC-I expressing tobacco plants were better protected from chilling stress. Photosynthesis was better protected against chilling-induced photoinhibition in OCI expressing transgenic tobacco plants compared to control plants (Van der Vyver et al., 2003). Over-expression of the two cystatins *AtCYSa* and *AtCYSb*, synonymous with *AtCYS3* and *AtCYS6*, respectively in transgenic yeast and Arabidopsis increased resistance to high salt,

drought, cold and oxidative stress (Zhang et al., 2008). Furthermore, the expression of two cystatins (*AtCys1* and *AtCys2*) from Arabidopsis was enhanced under high temperatures and wounding stresses (Hwang et al., 2010b).

1.3.2 Serine protease inhibitors

Serine proteinase inhibitors (SPIs) are universal throughout the plant kingdom and have been described in many plant species. There are a large number of known and partially characterized inhibitors of serine proteases (Haq et al., 2004). Of these, the best-characterized families of plant SPIs are the Kunitz-type (soybean serine trypsin inhibitor) and Bowman-Birk inhibitors (soybean serine protease inhibitor).

There are several structural features that are conserved in most Kunitz-type inhibitors. These include a molecular mass of 18–22 kDa, one or two polypeptide chains, a low cysteine content (usually with four Cys residues in two disulfide bridges), and one reactive site (Arg-Ser or Arg-Lys) involved in trypsin inhibition (Oliva and Sampaio, 2009). Most Kunitz inhibitors have four conserved Cys residues forming two disulfide bonds in a single (Figure 1.5A) or double chain polypeptide (Figure 1.5B) (Oliva et al., 2010). However; there are some Kunitz inhibitors that do not follow the Cys-Cys pattern of serine trypsin inhibitor (STI) (Figure 1.5C-F) (Oliva et al., 2010).

In contrast, Bowman–Birk type inhibitors (BBIs) have a lower molecular mass (8–10 kDa), a high cysteine content, and two independent reactive sites one at Lys 16-Ser 17 against trypsin and the other at Leu 43-Ser 44 against chymotrypsin (Figure 1.6) (Birk, 1985, Habib and Fazili, 2007). BBIs function as specific pseudo-substrates for digestive proteases resulting in

the formation of a stable complex where proteolysis occurs (Fernandez et al., 2007). They typically display a double-headed structure containing an independent protease-binding loop capable of binding and inhibiting trypsin, chymotrypsin and elastase (Fernandez et al., 2007). A third amino acid residue, in the loop is responsible for the specificity of the inhibition (Rahbé et al., 2003).

Serine proteases inhibitors are widespread and their physiological roles are similar to those of cystatins. The protease inhibitor II (PIN2) proteins (serine proteases) are important for regulating seed development (Sin et al., 2006). This was evident when expression of siRNA from the PIN2-RNAi construct resulted in transgenic *S.americanum* lines that were impaired in seed development (Sin et al., 2006). Chymotrypsin serine proteases inhibitors isoforms from pea seeds have been shown to inhibit the growth and development of the aphid *Acyrthosiphon pisum* (Rahbé et al., 2003).

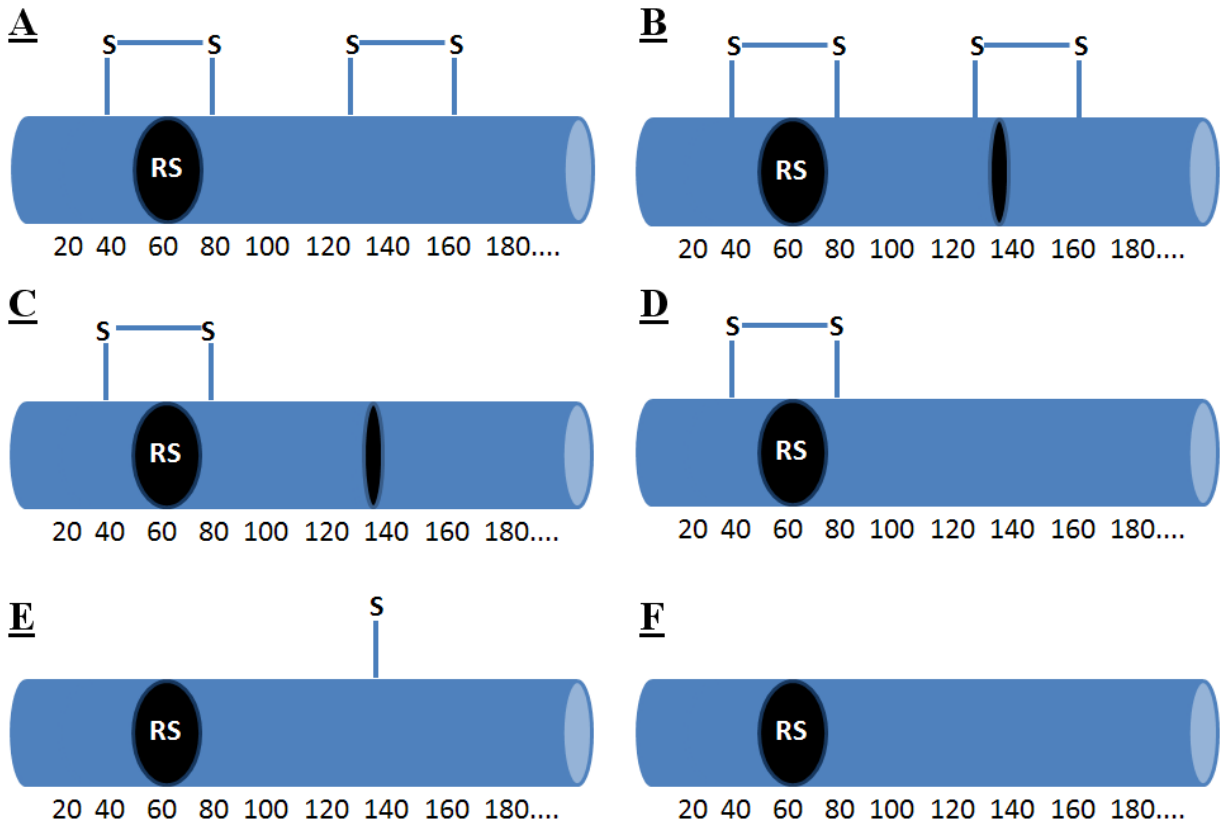


Figure 1.5 Scheme of primary structure of plant Kunitz inhibitors based on the number of disulfide bonds. (A) Inhibitors with two disulfide bonds and one polypeptide chain like the classical plant Kunitz STI; (B) Inhibitors with two disulfide bonds and two polypeptide chains like the *Mimosoideae* Kunitz inhibitors; Inhibitors with one disulfide bond and two polypeptide chains (C) or one polypeptide chain (D); (E) Inhibitors with one cysteine residue and one polypeptide chain; (F) Inhibitor devoid of cysteine residue and one polypeptide chain. RS-reactive site (Modified from Oliva et al., 2010).

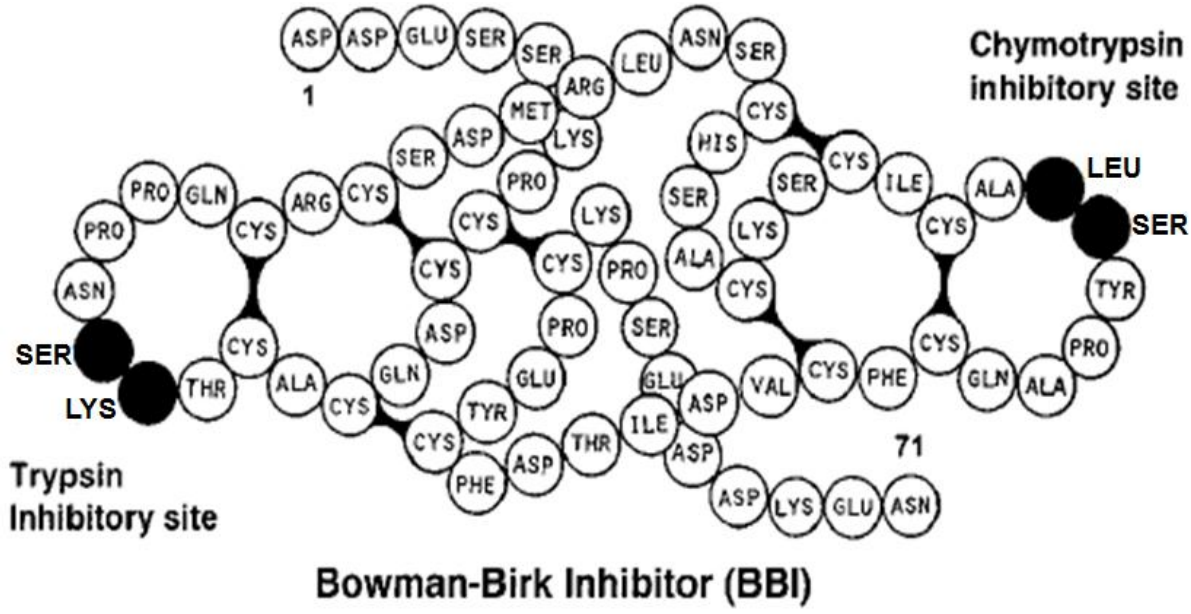


Figure 1.6 Structure of a major protease inhibitor, Bowman–Birk serine protease inhibitors (BBI), of soybeans. The positions of the reactive sites are indicated by closed circles (Friedman and Brandon, 2001).

However; the exact cause of this toxicity still needs to be investigated as no chymotrypsin activity was detected in the gut of the aphids. Dunse et al. (2010) demonstrated that co-expression of two potato type I, *Solanum tuberosum* potato type I inhibitor (StPin1A), and II, *Nicotiana alata* protease inhibitor (NaPI) provided cotton plants with protection against damage from the insect pest *Helicoverpa punctigera* (Dunse et al., 2010). The expression of a serine protease inhibitor from rice was shown to be strongly induced under drought conditions (Huang et al., 2007). The authors also showed that over-expression of the inhibitor resulted in significantly higher grain yield and seed setting. Transgenic tobacco lines over-expressing a typsin inhibitor gene from the serine PI family conferred tolerance against multiple stressors such as sodium chloride, variable pH and sorbitol (Srinivasan et al., 2009).

1.4 Plants Stress

Plants are constantly affected by abiotic stressors, such as low temperature, high salt, drought, flooding, heat, oxidative stress and heavy metal toxicity, during their development. Moreover, plants also face challenges from biotic factors including bacteria, fungi and viruses as well as from herbivores. All these stress factors prevent plants from reaching their full genetic potential and limit the crop productivity worldwide with abiotic stressors constituting more than 50% of crop failure worldwide (Mahajan and Tuteja, 2005).

1.4.1 What is stress and how is it perceived?

Stress in physical terms is defined as a response to a strain from an "object" that has undergone a mechanical force. The object would then undergo a change in response to the stress. In biological terms, it is difficult to measure the exact force exerted by stresses because plants are sessile. This is mainly because plants react differently to different environmental factors i.e. a biological condition which may be a stressor for one plants, may be favourable for another plant. In this regard, a broad definition of a biological stress is an adverse force or a condition, which inhibits the normal functioning and well-being of a biological system such as plants (Mahajan and Tuteja, 2005).

Plants undergo a variety of morphological, physiological, biochemical and molecular changes under abiotic stress. Abiotic stressors are often interconnected and may occur simultaneously. Drought is usually combined with high temperatures as well as high light (Chaves et al., 2003) and they can affect individually, or in combination, various plant processes associated with photosynthesis and elimination of reactive oxygen species (ROS) (Demirevska et al.,

2010). Various stressors often activate similar cell signal pathways and cellular pathways (Wang et al., 2003). Plant responses to such stressors are complex process that involves many genes and biochemical-molecular mechanisms as shown in Figure 1.7 (Huang et al., 2012, Wang et al., 2003). A stress signal, an osmotic or ionic effect or membrane fluidity changes, is first perceived by membrane receptors of the plant cells inducing a cascade of downstream processes. The signal generates second messengers including calcium and reactive oxygen species and inositol phosphates. The level of intracellular Ca^{2+} increases and this increase is sensed by calcium binding proteins (Ca^{2+} sensors). The sensors interact with various kinases and/or phosphatases causing gene expression. Cellular homeostasis is re-established through functional repair and protection of damaged proteins and membranes. Figure 1.7 represents a generic, simplified schematic representation of plant response to abiotic stress (Huang et al., 2012, Wang et al., 2003, Mahajan and Tuteja, 2005). Further, more details to the complexity of the plant's responses can be found in recent reviews by Yang et al. (2010) and Huang et al. (2012).

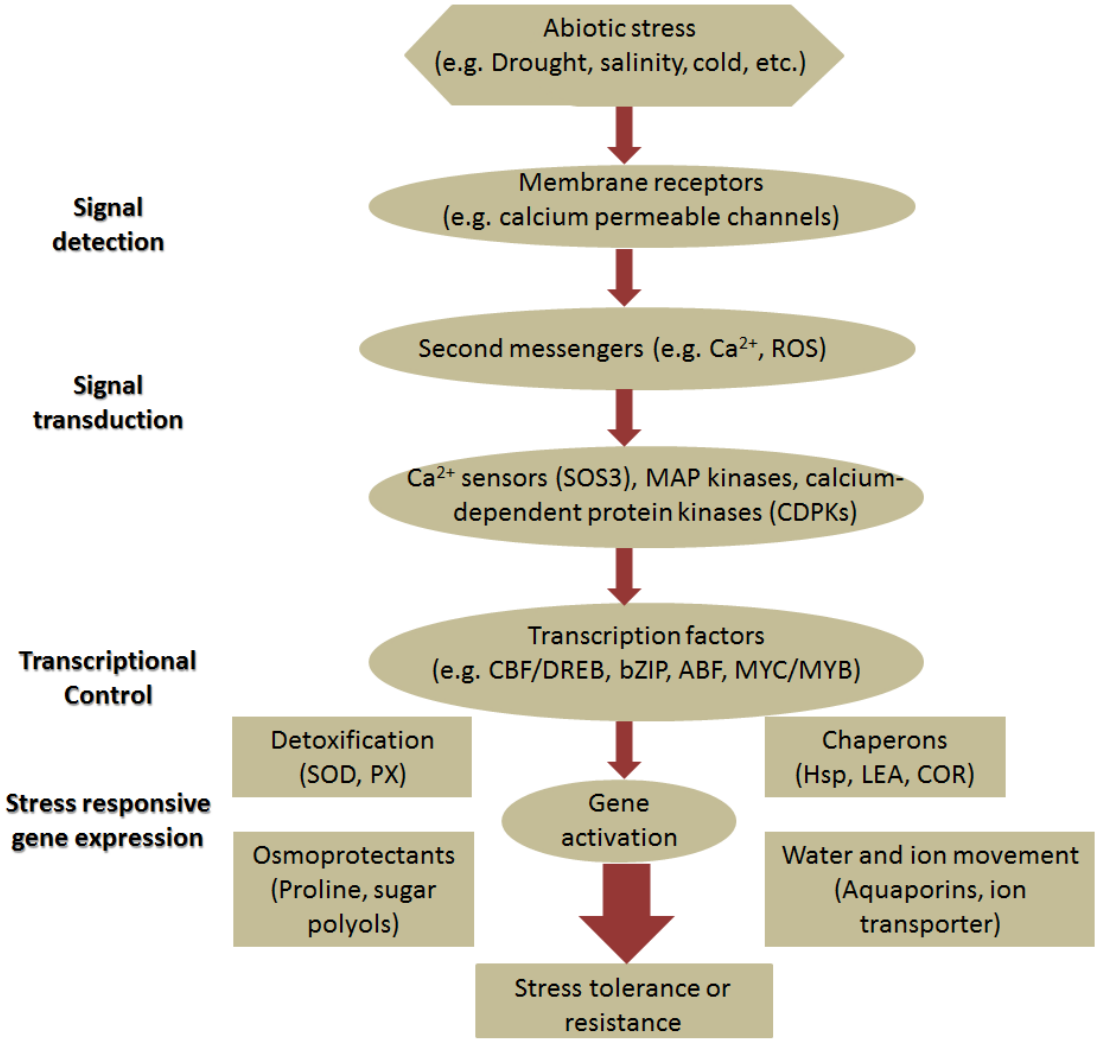


Figure 1.7 Generic pathway for the plant response to stress. The extracellular stress signal is first perceived by the membrane receptors and then activates large and complex signalling cascade intra-cellular including the generation of secondary signal molecules. The signal cascade results in the expression of multiple stress responsive genes. The products of this can provide stress tolerance directly or indirectly (Modified from Wang et al., 2003, Mahajan and Tuteja, 2005; Manavalan et al., 2009; Huang et al., 2012).

The various stress responsive genes can be broadly categorized as early and delayed induced genes (Figure 1.8) (Mahajan and Tuteja, 2005). Early genes are transiently expressed within minutes of stress perception in contrast to delayed genes which are activated hours following stress and are often sustained. These genes include the major stress responsive genes, such as RD (responsive to dehydration)/Kin (cold induced)/COR9 (cold response), modulating the production of various osmolytes, antioxidants, molecular chaperones and LEA-like proteins (late embryogenesis abundant), which function in stress tolerance.

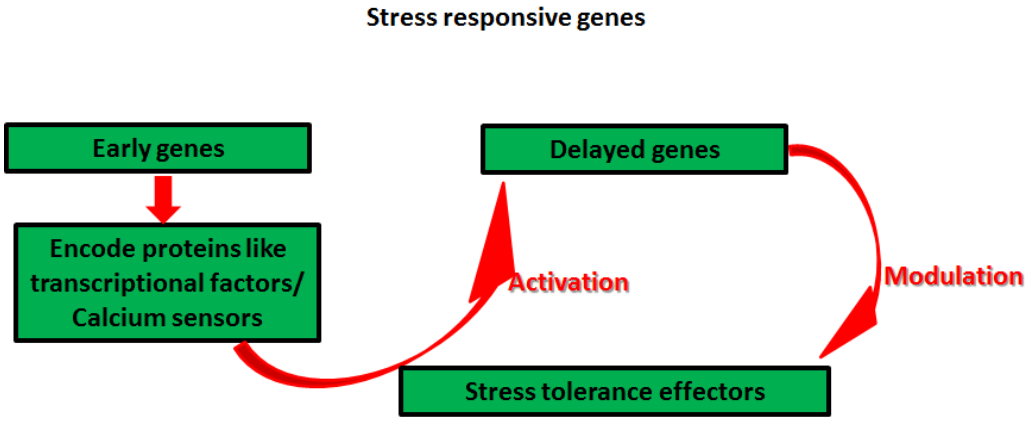


Figure 1.8 Early and delayed gene expression in response to abiotic stress signalling. Early genes encode for the transcription factors that activate the major stress responsive genes (delayed genes). The expression of major stress genes like RD/KIN/COR/RAB18/RAB29B results in the production of various osmolytes, antioxidants, molecular chaperones and LEA-like proteins, which function in stress tolerance (Modified from Mahajan and Tuteja, 2005).

1.4.2 Drought stress and plants

Drought is an important abiotic stresses inhibiting photosynthesis and decreasing plant growth and productivity. It is one of the major causes of worldwide crop losses reducing yield by more than 40% (Manavalan et al., 2009, Zlatev and Lindon, 2012). For example, polyethylene glycol (PEG) induced drought dramatically reduced (more than 50%) the yield (g/plant) of soybean plants (Hamayun et al., 2010). Plants undergo a number of molecular, biochemical and physiological processes in response to drought. These responses include stomatal closure through the production of abscisic acid (ABA), a chemical stress signalling compound, reduction in transpiration as well as decreases in the water potential of plant tissues, intracellular CO_2 and photosynthesis as well as growth inhibition (Figure 1.9) (Yordanov et al., 2003, Xu et al., 2010). Further, an indirectly proportional relationship exists between stomatal conductance and ABA production i.e. decreases in stomatal conductance coincides with increases in relative ABA. These changes in conductance and ABA production occur before any significant changes in leaf turgor (Liu et al., 2003).

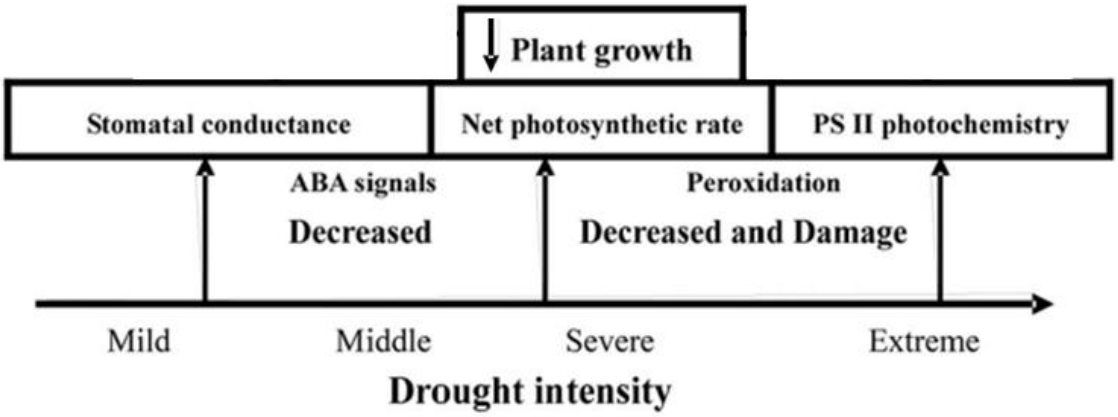


Figure 1.9 Simplified representation of some response routes in plants subjected to drought. Stomatal conductance and net photosynthetic rate are reduced as plants are subjected to different extents of drought stress. The photosynthetic apparatus is damaged under severe/extreme drought, e.g., leading to declines in PSII photochemical efficiency and enhanced peroxidation. Plant growth rate decreases gradually under water deficit (Modified from Xu et al., 2010).

1.4.3 Drought effect on plant growth

Studies have shown that drought inhibits plant growth and development (Ohashi et al., 2006, Fazeli et al., 2007, Nakayama et al., 2007, Gorai et al., 2010). Desirable traits for plants in a dry environment are therefore a high potential growth rate and efficient water-use. These traits are, however, influenced by the allocation of biomass to the different organs and the physiological and morphological properties of these organs (Zlatev and Lindon, 2012). This

allocation is dependent on a variety of environmental factors i.e. availability of water, nutrients, low light, CO_2 (Poorter et al., 2012).

Under drought conditions, a higher allocation of biomass to root/shoot ratio is usually observed suggesting an increase in the capacity for water uptake (Cregg and Zhang, 2001, Achten et al., 2010). When Scot pine (*Pinus sylvestris* L.) seedlings from Asia and Europe were compared, Asian Scot pine seedlings survived longer under drought conditions having a smaller phenotype with more biomass allocated to the roots and also having a higher water-use efficiency (WUE) (Cregg and Zhang, 2001). It has been found that even though drought dramatically decreases parameters such as seedlings height, basal diameter, leaf number, leaf area, root length, and biomass there is an increase in below-ground biomass indicating a higher root/shoot ratio (Wu et al., 2008). Allocation of biomass shifts to the shoots if the limiting factor is above-ground light or CO_2. However; biomass is allocated evenly if all above- and below-ground resources that a plant requires are limited with equal growth, a concept referred to as balanced-growth hypothesis (Poorter et al., 2012).

1.4.4 Drought effect on photosynthesis

Decrease in photosynthetic rate is attributed to stomatal and non-stomatal limitations (i.e. biochemistry) (Yordanov et al., 2003). During stomatal limitations, the stomata close rapidly avoiding water loss through transpiration. This is followed by restriction of CO_2 availability at the assimilation site in the chloroplasts. As a consequence, there is a decrease in net photosynthetic rate (A_n) that might be due to lowered C_i or inhibition of photosynthetic enzymes such as rubisco or ATP synthase (Zlatev and Lindon, 2012). Non-stomatal limitations are associated with reduced caboxylation efficiency, reduced ribulose-1,5-

bisphosphate (RuBP) regeneration, reduced amount of rubisco or inhibited functional activity of PSII (Zlatev and Lindon, 2012). Drought has been shown to have a marked reduction in both the initial slope of the A_n/C_i curve and A_{max} (Zlatev and Yordanov, 2004). Moreover, there was a reduction in maximal carboxylation efficiency and CO_2 assimilation in one of the cultivars of over five fold.

1.5 Adaptation to drought

Plants have different mechanisms for coping with drought stress (Figure 1.10) (Chaves et al., 2003). These mechanisms can be classified into 3 groups:

1.5.1 Drought escape

Drought escape is a strategy employed by the plant that involves completing its life cycle i.e. flower early during the period of sufficient water supply before the onset of drought. A practical example of drought escape is the Early Soybean Planting System employed in the southern USA. In this system, early maturing cultivars (short life cycle) are grown months before the onset of the period of possible drought thus set pods and completing the reproductive stage before the onset of drought (Manavalan et al., 2009). Recent studies have suggested that the Arabidopsis Flowering Locus T (FT) gene is a candidate for encoding "florigen", a photoperiod-regulated factor that acts to induce flowering. Thus, leaf-produced florigen has been genetically characterized in diverse crop species (Tamaki et al., 2007, Corbesier et al., 2007). Tamaki et al. (2007) showed that a protein encoded by H3da, a rice homologue of FT, moves from the leaf to the shoot apical meristem inducing flowering in rice in a similar manner as Arabidopsis *FT* gene. Over-expression of the *FT* homologous

gene from apple also accelerates flowering in Arabidopsis, poplar and apple compared to wild-type plants (Tränkner et al., 2010). Use of such candidate genes may contribute to engineering plants that escape drought through early flower.

1.5.2 Drought avoidance

Plants can also endure drought by using strategies that help them avoid tissue dehydration either by reducing water loss from aerial parts or by efficient water absorption from roots (Manavalan et al., 2009). Water loss can be minimized by closing stomata or by reducing light absorbance through rolled leaves (Chaves et al., 2003). Water uptake can be improved by increasing investment in roots by developing a longer taproot thus reaching deeper soil layers where water is available. Moreover, extensive fibrous roots can be used for foraging subsoil surface moisture and nutrients such as phosphorous (Chaves et al., 2003, Manavalan et al., 2009). Lian et al. (2004) observed that upland rice (*Oryza sativa* L. spp *indica* cv. Zhonghan 3) responded to drought stress by rolling the young leaves resulting in a decline in transpiration. There was also increased expression of the *RWC3* mRNA, a well-known aquaporin, involved in drought avoidance. Furthermore, over-expression of *RWC3* in a different rice cultivar resulted in increased leaf water potential and higher root osmotic conductivity in transgenic lines compared to the wild-type plants (Lian et al., 2004).

1.5.3 Drought tolerance

This refers to the ability of the plant to maintain turgor and continue metabolism at low leaf water status by osmotic adjustment, more rigid cells, or smaller cells, or synthesis of osmoprotectants, osmolytes or compatible solutes (Manavalan et al., 2009). Mittler et al. (2001) showed that plants of evergreen legume *Retama raetam* adopt a strategy of partial dormancy to survive drought. In these plants, a low level of a number of essential proteins, including the large and small subunits of rubisco, ascorbate peroxidase and the D1 subunit of the reaction centre of photosystem II is evident during the drought period. Upon water application, the "photosynthetically suppressed" stems recovered and accumulated essential proteins within 6-24 hrs (Mittler et al., 2001). Drought tolerant plants are able to maintain of high leaf water-use efficiencies, high Fv/Fm ratios and biomass as was the case with the soybean genotype, Prima (Fenta et al., 2012).

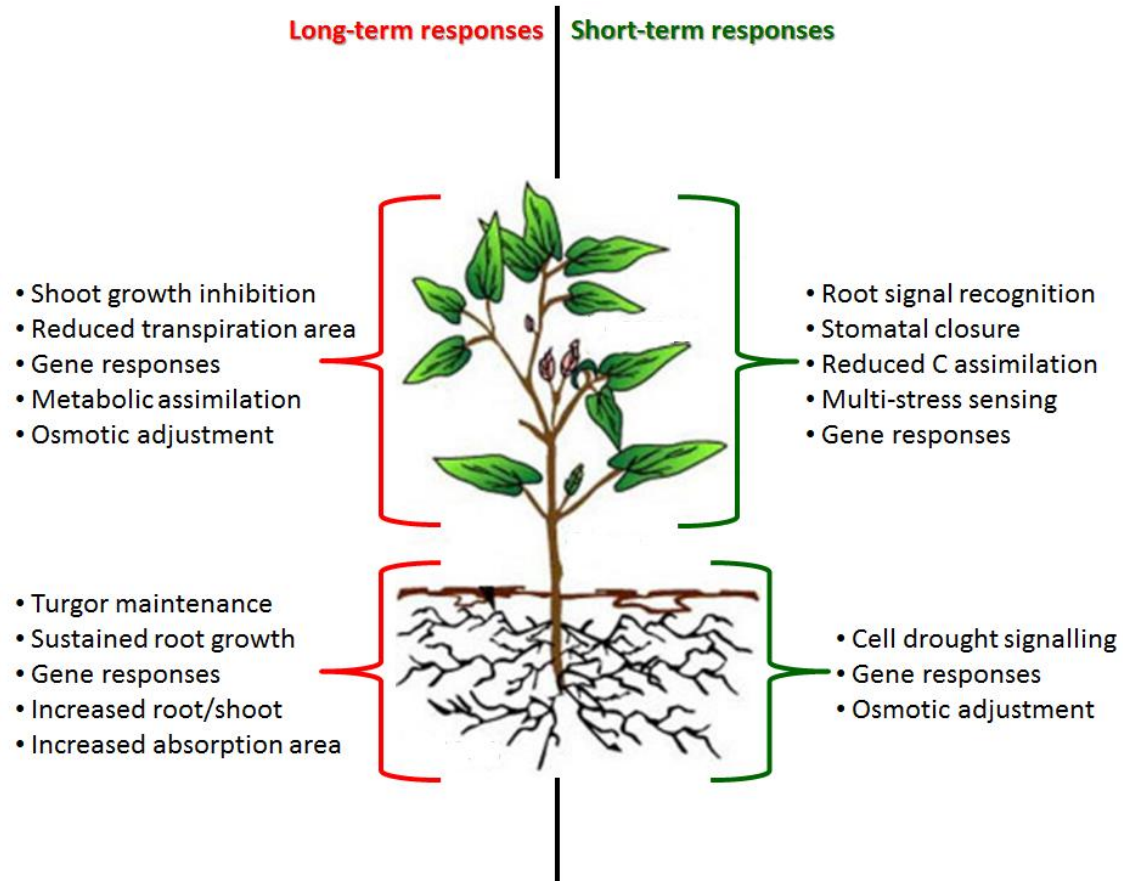


Figure 1.10 Whole plant responses to drought stress. Left: long-term or acclimation responses; right: short-term responses (Modified from Chaves et al., 2003).

1.6 Problem statement

Plants are subjected to a variety of stresses that affect plant growth and development. With climate change, in particular drought and heat/cold stress, photosynthetic CO_2 assimilation and water uptake might change affecting soybean yield and productivity (Manavalan et al., 2009). Several biotechnology strategies have been employed for the selection and characterization of more drought-tolerant plants. This includes utilizing soybean genome sequence for molecular marker development and genetic engineering of plants for the improvement of drought tolerance (Manavalan et al., 2009). However; any genetic engineering approach, as applied in this PhD study, requires the selection of an appropriate gene, cloning of the respective construct for gene transfer, efficient transformation of plants, characterization of the transformed plants for expression of the gene transferred, establishment of a population of plants stably expressing the transferred gene and finally, investigating these transformed plants for a specific modified trait.

One of the major constraints for using soybean in this approach is the low transformation efficiency with *in vitro* soybean regeneration highly genotype dependent. Soybean cultivar Williams 82, which has been used in this study, has been previously successfully transformed (Paz et al., 2006) and a soybean genomic database has been recently established for this cultivar. Further, any biotechnology project, which includes characterizing transgenic plants, requires the designing and cloning of gene constructs to be used for transformation of plants and investigating any phenotype change. Attempts to transform soybean were previously unsuccessful in our laboratory due to problems that were encountered including contamination and extremely low regeneration and would have not been possible in the time frame of a PhD project. Transformation was therefore carried out at Iowa State University

where a well-established and optimized protocol is being used allowing for sufficient regeneration. Moreover, since there was uncertainty as to whether and how many positive transgenic soybean plants would be obtained; ectopic expression of protease inhibitors was also carried out in Arabidopsis plants to monitor any phenotypic change caused by protease inhibitor expression. Transformation of Arabidopsis was carried out using the floral dip method because it is inexpensive, simple and quick unlike soybean transformation which is costly, difficult and slow.

1.7 Study aim and objectives

The overall PhD study on soybean was motivated by previous work at the University of Pretoria, South Africa and the Universities of Newcastle and Leeds, United Kingdom, on transgenic tobacco over-expressing the rice cystatin OC-I (van der Vyver et al., 2003; Prins et al., 2008). In their study, OC-I over-expressing tobacco plants had higher biomass and protein content and recovered better from chilling stress than non-transformed tobacco plants but were not more drought-tolerant. Therefore, this PhD study was aiming to replicate the phenotype of the transgenic tobacco plants with another commercially important crop like soybean and compare these results with results derived from the model plant Arabidopsis. This model plant has the advantage that molecular and gene function data are available and the ease of making crossings between different transgenic plants.

In this PhD study, transgenic soybean and Arabidopsis plants will therefore be produced over-expressing various protease inhibitors to determine if these inhibitors will cause any phenotypic changes by targeting specific classes of proteases sensitive to inhibitor action.

This might also provide information about the role of these proteases in developmental and stress reactions leading to a better understanding to limiting the effects of stress.

To achieve the aim of the study, the following objectives were set:

1. Construction of various vectors for plant transformation allowing over-expression of the rice cysteine protease inhibitor-I (OC-I) directed to the cytosol and chloroplast in soybean and Arabidopsis as well as soybean nodules and allowing over-expression of the maize serine protease inhibitors (Bowman-Birk inhibitor, BBI, and the endopeptidase type II inhibitor, TI) directed to the cytosol of soybean and Arabidopsis plants.
2. *Agrobacterium*-mediated transformation of soybean (was carried out by researchers at Iowa State University). Arabidopsis transformation using the floral dip method.
3. Molecular and biochemical characterization of the OC-I expressing transgenic soybean and Arabidopsis plants to confirm gene expressions so as to facilitate selection of transgenic plants expressing protease inhibitors in order to perform phenotypic analysis.
4. Phenotypic analysis of drought-treated transgenic soybean and Arabidopsis plants to determine a possible protective effect of OC-I expression on plant performance under drought.

Chapter 2

Construction of plasmids for transformation of soybean and *Arabidopsis thaliana*

2.1 Introduction

An important aspect of understanding the molecular mechanisms of economically important crops, such as soybean, is an efficient transformation system. There are several methods used for soybean transformation. However; *Agrobacterium*-mediated transformation is the preferred method used due to its simplicity, low cost and usually results in the insertion of a single copy of the transgene (Gelvin, 2003, Mello-Farias and Chaves, 2008). One of the key factors in soybean transformation is the construction of a suitable expression cassette for expression of the transgene (Mello-Farias and Chaves, 2008). The expression cassette usually comprises of the transgene, a selectable marker, a promoter to drive expression of the transgene and a transcription terminator.

Objective of this part of the study was to construct various expression vectors for protease inhibitor expression in different cellular compartments and in nodules. For cysteine protease inhibitor expression, three different expression cassettes were prepared. The first expression cassette comprising of cauliflower mosaic virus 35S promoter (CaMV P35S), OC-I gene and a cauliflower mosaics virus terminator was generated. CaMV P35S is well established as a promoter to express ectopic genes in plants and has been successfully used since the 1980s for a myriad of plants (Odell et al., 1988). The P35S is a strong constitutive promoter resulting in high levels of gene expression in plants such as soybean (Olhoft et al., 2003, Paz et al., 2004, Podevin and du Jardin, 2012). A second expression cassette consisting of the leghemoglobin promoter from soybean, OC-I gene (targeted to the nodules) and CaMV terminator was generated. A leghemoglobin promoter has been previously used for specific gene expression targeted to the nodules (Carvalho et al., 2003). The third expression cassette targeted expression of the OC-I to the chloroplast thus consisted of a CaMV P35S, a

phosphoribulokinase (PRK) signal sequence, the OC-I gene and a terminator. It has been shown that it is possible to target specific reporter genes to the different cellular sub-compartments through the inclusion of target specific signal peptides in the expression cassette (Mehlmer et al., 2012). In addition, two more expression cassettes were generated expressing the maize serine proteases inhibitors: Bowman-Birk inhibitor (BBI) and the type II endopeptidase inhibitor (TI) both under the control of the CaMV P35S.

The various expression cassettes were cloned into the plasmid pTF101.1. This plasmid has the phosphinothricin acetyl transferase (*bar*) gene from *Streptomyces hygroscopicus* that confers resistance to the herbicide phosphinothricin and its derivatives thus allowing for easy differentiation between transgenic and non-transgenic plants (Paz et al., 2004). The plasmid pTF101.1 has been previously successfully introduced into the *Agrobacterium tumefaciens* strain EHA101 and has been used for soybean transformation (Paz et al., 2004, Paz et al., 2006). All the expression vectors constructed were sent off to Iowa State University to be used for *Agrobacterium*-mediated transformation of soybean. Arabidopsis plants were also transformed using the floral dip method as described by Clough and Bent (1998).

2.2 Materials and Methods

2.2.1 OC-I plasmids

The plasmids pLBRCys-I and pLegCys-I were already available at the University of Pretoria. These plasmids both conferred resistance to ampicillin, pLBRCys-I consisted of a cystatin (OC-I) gene under the control of a double CaMV P35S and a CaMV terminator while

pLegCys-I was under the control of a leghemoglobin promoter and a CaMV terminator (Annex 1A and B).

The presence of the transgene (OC-I) was verified by amplification of the OC-I gene using forward primer (5'-ATG TCG AGC GAC GGA GGG CC-3') and reverse primer (5'-GAT GGG CCT TAG GCA TTT GC-3') (Annex 2A). A 25 µL PCR mixture consisted of 12.5 µL 2X BioMix Red (Bioline, UK), 0.5 µM of each primer and 30-40 ng of template plasmid. The PCR was initiated with a 5 min hot start at 94°C followed by 30 cycles each consisting of 30 s at 94°C, 30 s at 60°C and 1 min at 72°C. Final elongation occurred at 72°C for 5 min. The resultant PCR products were visualized on a 1% agarose gel. PCR was also carried out under the same conditions for pLegCys-I using the following primer sequences leghemoglobin forward primer (5'-CAC TCT TCA AGC CTT CTA TAT-3') and OC-I reverse primer (5'-GAT GGG CCT TAG GCA TTT GC-3') with annealing at 55ºC (Annex 2B). PCR products were visualized on a 1% agarose gel. The gained PCR fragments were sequenced (GATC Biotech, Germany) and verified using online tools such as NCBI blast and multalign.

2.2.2 Construction of plasmid LBRPRK

2.2.2.1 Phosphoribulokinase (PRK) amplification

The coding sequence of phosphoribulokinase (PRK) gene was amplified as a *Hind*III and *Bam*HI (underlined sequences) flanked fragment from the template plasmid pTLT-PRK-Nter-YF (kindly provided by Norbert Mehlmer, Germany) using forward primer (5'-TCCT GGA TCC ATG GCT GTC TCA ACT ATC TAC TCA AC-3') and reverse primer (5'-CTGT GGA TCC GCC AGA GTC AGC AGC TAG T-3'). A 50 µL PCR mixture which consisted

of the plasmid (50-100 ng/µL), 5 µL 10X buffer containing 1X Mg^{2+} to a final concentration of 2 mM, 0.4 µL dNTPs to a final concentration of 0.2 mM, 0.25 µL of each primer (100 µM) was added, sdH_2O was added to make up the mixture to 49 µL. Finally, 1 µL Pfu DNA polymerase (2.5 U) (Agilent technologies, UK) was added to the PCR mixture. The PCR cycles were 5 min at 94°C followed by 30 cycles each consisting of 30 s at 94°C, 30 s at 60°C and 1 min at 72°C. Final elongation occurred at 72°C for 5 min. An aliquot of the PCR product was visualized on a 1% agarose gel and an expected 200 bp band was observed.

The remainder of the PCR amplification was purified using a QIAquick PCR purification kit (Qiagen, Germany) according to the manufacturer's instructions. Briefly, five times the volume of buffer PB was added to the PCR mix. The mixture was transferred to a QIAquick column and centrifuged. The column was washed with buffer PE. Following the wash, purified PCR product was eluted with sdH_2O.

2.2.2.2 Cloning into vector pLBR

The purified PCR product and its destination vector pLBR 19 (Annex 3A) were digested with 1 U of the restriction enzymes *Hind*III HF™ and *Bam*HI HF™ (New England BioLabs, USA) for 2 hrs at 37ºC in a 20 µL reaction. The digested PCR product and linearized pLBR 19 were purified using the QIAquick PCR purification kit as mentioned above (2.2.2.1).

The linearized pLBR 19 destination vector was dephosphorylated by alkaline phosphatase treatment (Fermentas, Canada) for 30 min at 37°C. The ligation was carried out in a total volume of 10 µL containing 1 µL 10X ligase buffer (New England BioLabs, USA), 1 µL T4 ligase (1 U) (New England BioLabs, USA), 6 µL purified PCR product and 2 µL

dephosphorylated plasmid. The ligation was carried out at room temperature (RT) for 1 hr. The reaction was terminated by heating for 10 min at 65ºC.

2.2.2.3 E.coli transformation

For *E. coli* transformation, 5 µL of the ligation mix was added to 50 µL of commercially available DH5α competent cells (Promega, USA) using heat shock treatment. The mixture (ligation and competent cells) was incubated on ice for 30 min. Heat shock was performed at 42ºC for 30-45 s. Cells were immediately transferred to ice for 2 min. SOC medium (Sigma-Aldrich, USA) (250 µL) was added and the mixture was shaken at 200 rpm at 37ºC for 1 hr. The transformation mixture (100 µL) was plated onto a Luria broth (LB) plate (1.5% agar (w/v), 1% bacto-tryptone (w/v), 0.5% bacto-yeast extract (w/v), 1% sodium chloride (w/v) pH 7.5) containing 100 µg/mL ampicillin and 40 µg/mL X-Gal (5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside). The plate was incubated at 37ºC overnight (O/N). Single colonies were selected and inoculated in 5 mL LB (1% bacto-tryptone (w/v), 0.5% bacto-yeast extract (w/v), 1% sodium chloride (w/v) pH 7.5) media containing ampicillin at 37ºC with 200 rpm shaking O/N.

2.2.2.4 Plasmid extraction

Plasmid DNA was extracted using the GeneJET™ Plasmid Miniprep Kit according to the manufacturer's instructions (Fermentas, Canada). Cells were centrifuged at 13 000 rpm and the pellet was re-suspended in 250 µL re-suspension buffer. The cells were then lysed in 250 µL lysis buffer which was followed by the addition of 350 µL neutralization buffer. The

mixture was centrifuged at 13 000 rpm for 5 min. The supernatant was loaded on a GeneJET™ column, centrifuged, washed and plasmid DNA eluted in sdH_2O.

The presence of the PRK sequence was determined by PCR amplification of the 200 bp band from the clones using forward primer (5'-ATG GCT GTC TCA ACT ATC TAC TCA AC-3') and reverse primer (5'-GCC AGA GTC AGC AGC TAG T-3'). An expected 200 bp band was visualized on a 1% agarose gel and sequenced to verify identity. The positively transformed plasmid was named pLBRPRK.

2.2.3 Construction of pLBRPRKCys-I

The OC-I gene was amplified from plasmid pLBRCys-I using the following primers: forward primer (5'-TCCT GAA TTC ATG TCG AGC GAC GGA GGG CC-3') and reverse primer (5'-TCCT GAA TTC GAT GGG CCT TAG GCA TTT GC-3') containing *Eco*RI restriction sites (underlined). PCR purification was carried out identical to the procedure outlined in section 2.2.2.1. The purified PCR product together with the vector pLBRPRK were digested with 1 U each of the restriction enzyme *Eco*RI HF™ (New England BioLabs, USA) under the same conditions as outlined in section 2.2.2.2. The digested PCR product and linearized pLBRPRK were purified and dephosphorylated, respectively. Ligation and transformation was carried out according to the procedure outlined in section 2.2.2.2 and 2.2.2.3, respectively. The presence of the OC-I gene was determined by amplification of the OC-I (300 bp) using forward primer (5'-ATG TCG AGC GAC GGA GGG CC-3') and reverse primer (5'-GAT GGG CCT TAG GCA TTT GC-3'). PCR was also carried using the PRK forward primer (5'-ATG GCT GTC TCA ACT ATC TAC TCA AC-3') and OC-I reverse

primer (5'-GAT GGG CCT TAG GCA TTT GC-3') to determine the orientation of the cloned genes. The produced plasmid was designated pLBRPRKCys-I.

2.2.4 Construction of pLBR-BBI and pLBR-TI

2.2.4.1 RNA extraction and cDNA synthesis

Leaf discs were harvested from 4 weeks old maize plants, frozen in liquid nitrogen and stored at -80ºC until required. Samples were ground in liquid nitrogen using a pestle and mortar. Total RNA was isolated from 50 mg of the sample powder using the RNeasy® Mini Kit (Qiagen, Germany) following the manufacturer's instructions. The RNA quantity and purity was determined with a spectrophotometer (Nanodrop® Thermo Scientific, USA).

Extracted RNA was used to synthesize single-stranded cDNA by random hexamer priming with the QuantiTect® Reverse transcription kit (Qiagen, Germany) according to the manufacturer's instruction. In short, RNA of up to 1 µg was treated with gDNA Wipeout buffer in a final volume of 14 µL and incubated at 42°C for 2 min. The reverse-transcription mix containing 4 µL 5X Quantiscript RT buffer, 1 µL RT primer mix and 1 µL Quantiscript reverse transcriptase was added to the genomic DNA eliminated RNA to make a total volume of 20 µL. cDNA synthesis was carried out under the following conditions: 42ºC for 15 min and 95ºC for 3 min.

The cDNA quantity and quality was determined by spectrophotometer and visualized on a 2% agarose gel following PCR reaction with BBI-*Hind*III forward primer 5'-CCCG AAG CTT GCC AGG ACA GGA GAA ACA AA-3' and BBI-*Bam*HI reverse primer 5'-TCCT

GGA TCC CAT GCC GTA CGT CAG AAG AA-3' with primer annealing at 60°C. The PCR reaction was also carried out for the serine inhibitor, TI, with the following primers TI-*Hind*III forward primer 5'-GGCC AAG CTT TGC CGT AAG CAG ATC GAC TA-3' and TI-*Bam*HI reverse primer 5'-TCCT GGA TCC ATC ACG CAC ACT TTG GTT CA-3'. PCR products of 360 bp for BBI and 300 bp for TI were purified and cloned into pLBR 19 as outlined in sections 2.2.2.2 to 2.2.2.4. The success of the cloning was confirmed by PCR and sequencing. The final plasmids were designated pLBR-BBI and pLBR-TI.

2.2.5 Cloning into pTF101.1

Expression cassettes of plasmids pLBRCys-I, pLegCys-I, pLBRPRKCys-I, pLBR-BBI and pLBR-TI (Annex 3 B-F) were all cloned into the plasmid pTF101.1 (Annex 4A). This vector has a spectinomycin-resistant marker gene (aadA) for bacterial selection. The plant selectable marker gene cassette consists of a double 35S promoter (2X P35S) of CaMV, (2) a *bar* gene for resistance to the herbicide phosphinothricin and its derivatives. Plasmid pTF101.1 contains a multiple cloning site (MCS) allowing subcloning of the gene of interest between the right border region and the plant selectable marker cassette. Constructs from the above mentioned plasmids were cloned using the *Sac*I HF™ and *Xba*I restriction sites. The five plasmids, including pTF101.1, were individually digested with 1 U each of the restriction enzymes *Sac*I and *Xba*I and double digested with *Sac*I + *Xba*I (New England BioLabs, USA) for 2 hrs at 37ºC in a 20 µL reaction. Double digested pTF101.1 was dephosphorylated and the constructs (Annex 1A-E) from the plasmids were gel purified with the QIAquick® Gel Extraction Kit (Qiagen, Germany) according to the manufacturer's instructions. Ligation and transformation was carried out according to section 2.2.2.2 and 2.2.2.3, respectively. Successful transformation was confirmed by double restriction enzyme digestions of the

plasmids which were extracted using the GeneJET™ Plasmid Miniprep Kit (Fermentas, Canada).

2.3 Results

2.3.1 Verification of pLBRCys-I and pLegCys-I

A PCR was carried with plasmids LBRCys-I and LegCys-I and with OC-I specific primers (Figure 2.1A). Bands with a size of 300 bp were visualized on a 1% agarose gel confirming presence of the OC-I coding sequence in plasmids LBRCys-I and LegCys-I (Figure 2.1A). The PCR product of the OC-I was sequenced and aligned with the Genbank OC-I (Annex 5A). PCR was also carried out with plasmid LegCys-I using a forward primer which binds to part of the leghemoglobin promoter and an OC-I specific reverse primer (Figure 2.1B). Amplified products with a band size of 450 bp were visualized on a 1% agarose gel. The sequence of the leghemoglobin promoter along with the OC-I was obtained (Annex 6A).

2.3.2 Construction of the plasmid LBRPRKCys-I

The PRK gene was amplified, digested with restriction enzymes and ligated into the plasmid LBR 19 that was also digested with the same restriction enzymes (Figure 2.2A). Following ligation, the resultant recombinant DNA (Figure 2.2B) was transformed into competent *E. coli* cells (Figure 2.2C). The mixture was then plated onto an ampicillin-containing selection medium (Figure 2.2D). Transformants were cultured in liquid medium and plasmid DNA was purified (Figure 2.2E). A PCR was carried out with purified recombinant plasmid DNA and positive transformants were confirmed by sequencing (Annex 5B). For construction of

LBRPRKCys-I, the OC-I gene was amplified using primers with *Eco*RI restriction sites flanking the OC-I and cloned downstream of the PRK gene in the plasmid LBRPRK in a similar way as indicated in Figure 2.2. Positive transformants were confirmed by PCR of the PRK gene and OC-I (Figure 2.3A). Amplification was also carried out using the PRK forward primer and OC-I reverse primer and plasmids that amplified were designated LBRPRKCys-I (Figure 2.3B).

2.3.3 Construction of the plasmid LBRBBI and LBRTI

The BBI and TI serine protease inhibitor coding sequences were amplified from maize cDNA with gene primers containing restriction enzyme sites (*Bam*HI and *Hind*III) (Figure 2.4). Cloning was carried out identical to the procedure outlined in section 2.3.2 (Figure 2.2). The resultant recombinant plasmids i.e. LBR-BBI and LBR-TI were amplified and sequenced (Annex 6B and C).

2.3.4 Cloning of constructs into pTF101.1

The plasmids LBRCys-I (A); LegCys-I (B); LBRPRKCys-I (C) which express the OC-I gene in the cytosol, nodules and chloroplast, together with the plasmids LBR-BBI and LBR-TI that express the serine inhibitors in the cytosol were digested with different restriction enzymes (Figure 2.5). The constructs of interest (indicated by the red arrow in Figure 2.5) were cloned into the vector pTF101.1 to create pTF101.1-Cys-I; pTF101.1-LegCys-I; pTF101.1-PRKCys-I; pTF101.1-BBI and pTF101.1-TI (Figure 2.6). These plasmids were sent off to Iowa University (USA) for soybean transformation. Plasmids were also used for Arabidopsis transformation except for the pTF101.1-LegCys-I.

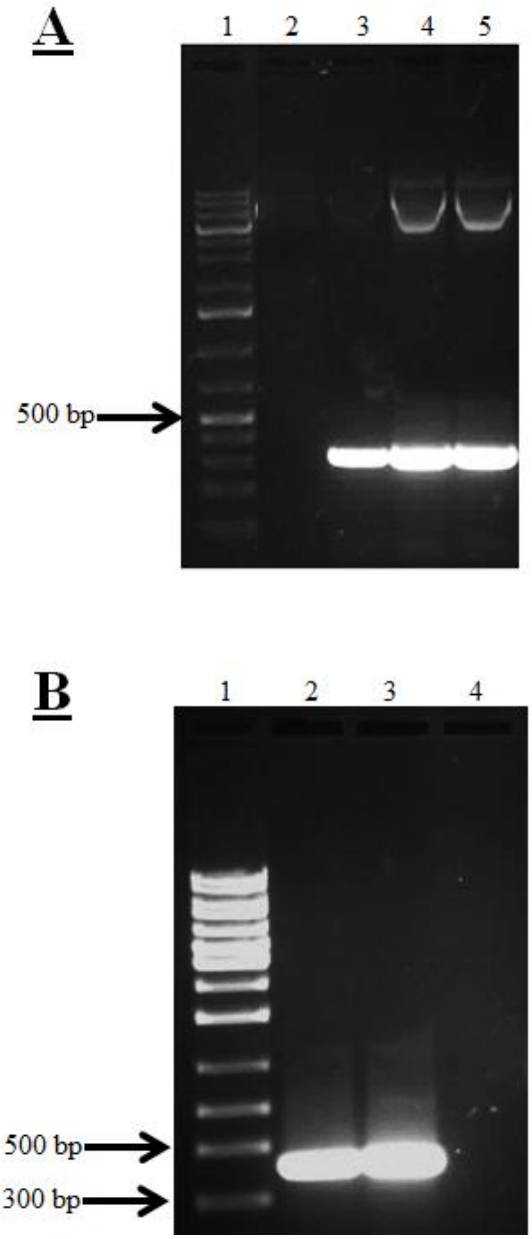


Figure 2.1 PCR amplification of the OC-I coding sequence with OC-I primers. (A) (1) 1 Kb marker; (2) Negative water control (NC); (3) Positive plasmid control (PC); (4) pLBRCys-I PCR product and (5) pLegCys-I PCR product. (B) PCR amplification of the OC-I gene using leghemoglobin promoter FP and OC-I RP. (1) 1 Kb marker; (2) pLegCys-I PCR product 1; (3) pLegCys-I PCR product 2; (4) NC (water).

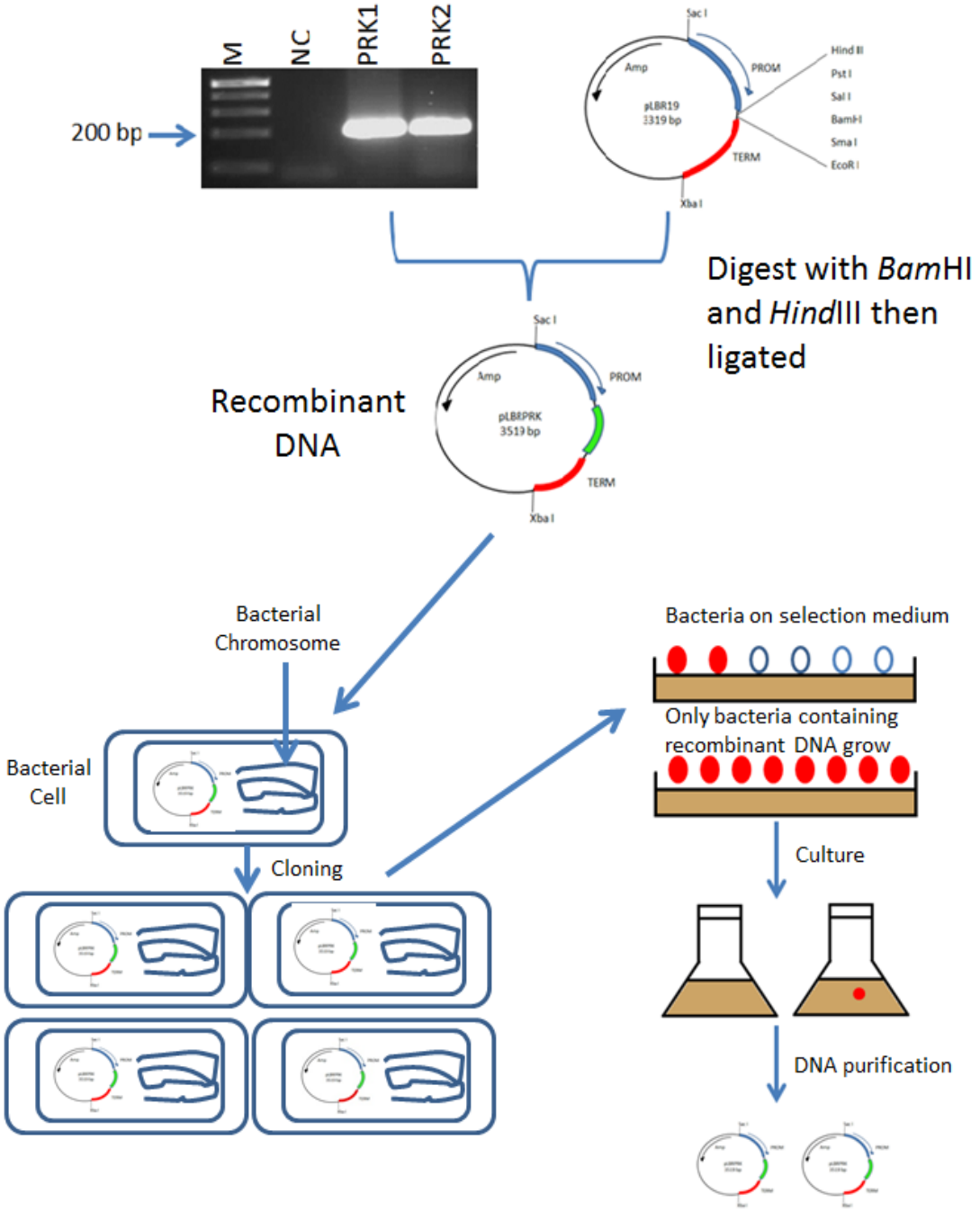


Figure 2.2 Diagrammatic illustration showing the cloning of the PRK gene into pLBR 19. (A) The PCR product and plasmid were digested with *Bam*HI *and Hind*III and ligated. (B) The resultant recombinant DNA was transformed into bacterial cells (C). (D) The bacterial mixture was plated out. (E) Positive colonies were cultured and plasmid DNA purified (Modified from http://www.accessexcellence.org/RC/VL/GG/plasmid.php).

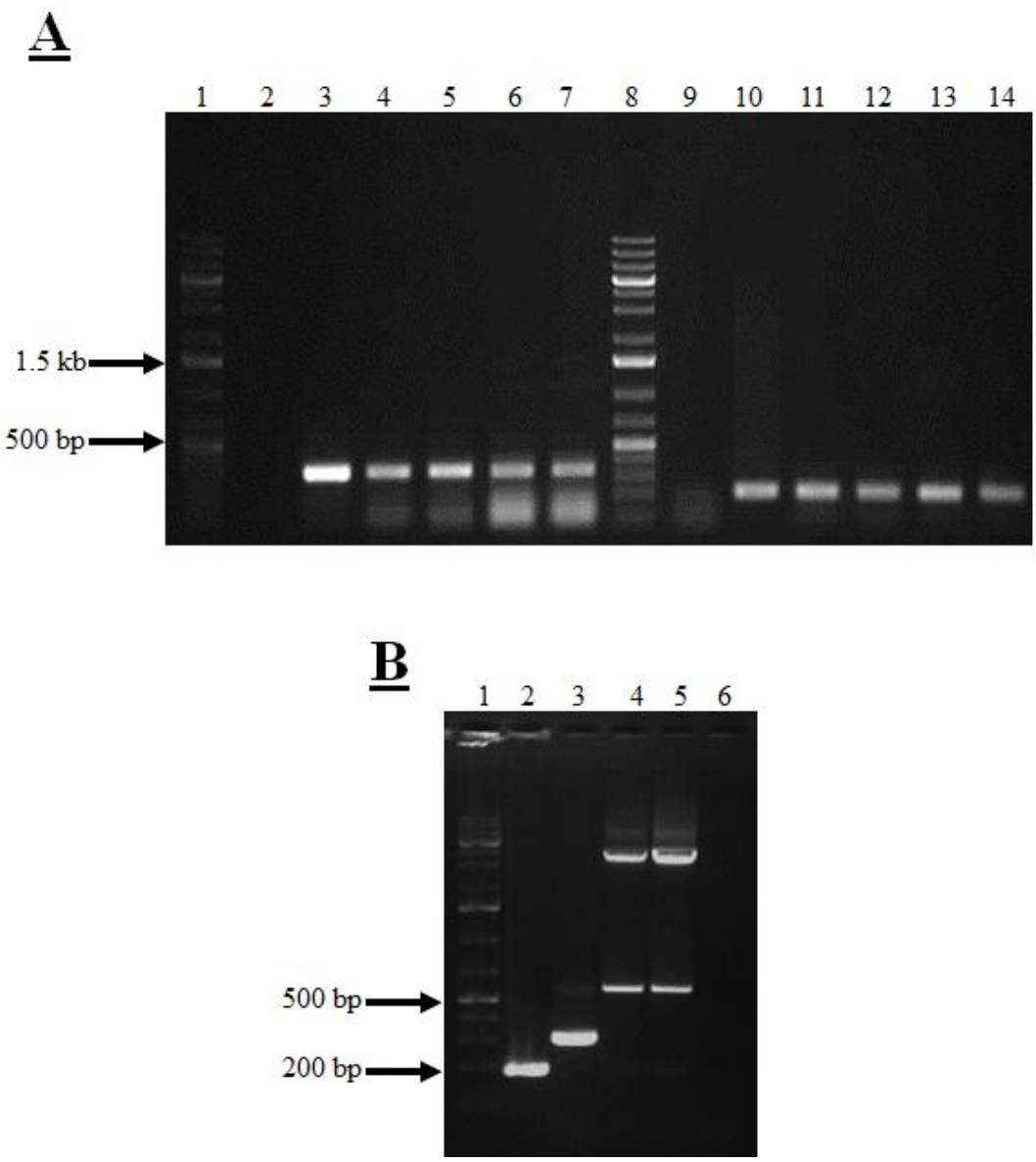


Figure 2.3 PCR of the OC-I and PRK coding sequences from pLBRPRKCys-I plasmids. (A) (1) 1 Kb marker; (2) NC (water); (3) PC (pLBRCYS-I plasmid); (4-7) pLBRPRKCys-I OC-I PCR products 1-4; L (8) 1 Kb marker; (9) NC (water); (10) PC; (11-14) pLBRPRKCys-I PRK PCR products 1 to 4. (B) (1) 1 Kb marker; (2) PRK PCR product; (3) OC-I PCR product; (4 and 5) PRKCys-I PCR products from pLBRPRKCys-I plasmids 1 and 2, respectively and (6) NC (water).

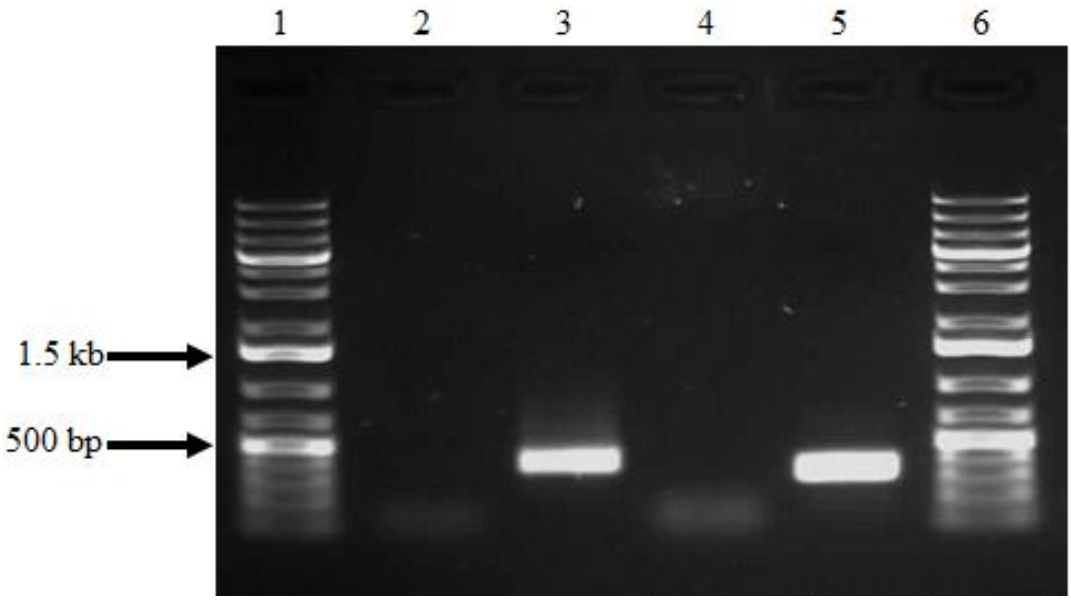


Figure 2.4 PCR of maize cDNA with BBI and TI primers. (1) 1 Kb marker; (2) NC (water); (3) BBI PCR product; (4) NC (water); (5) TI PCR product and (6) 1 Kb marker.

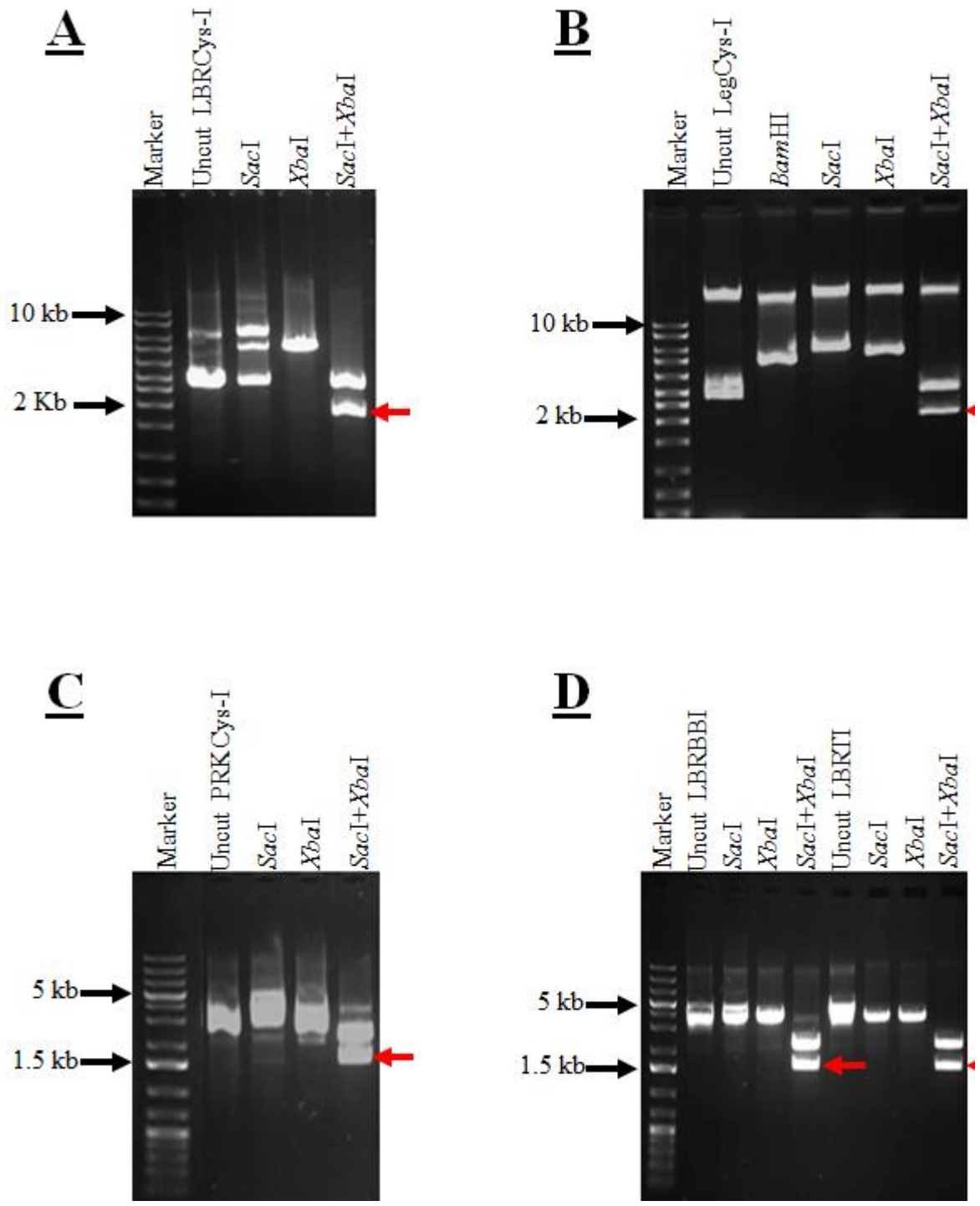


Figure 2.5 Restriction enzyme digestions of the plasmids. LBRCys-I (A); LegCys-I (B); LBRPRKCys-I (C) and the serine inhibitors BBI and TI (D). Red arrows indicate the construct of interest.

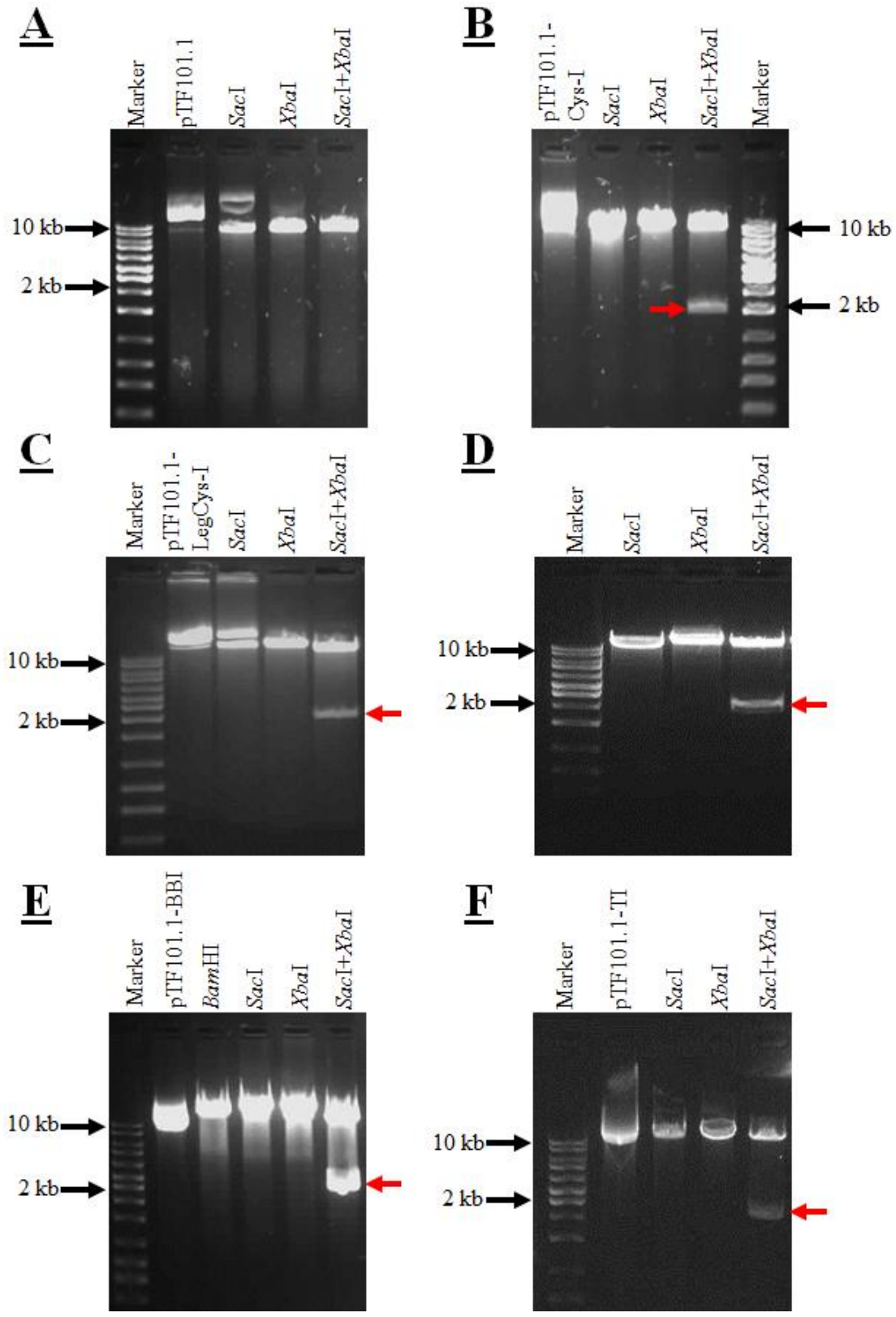


Figure 2.6 Verification of cloning by restriction enzyme digestion of the plasmids. pTF101.1 (A); pTF101.1-Cys-I (B); pTF101.1-LegCys-I (C); pTF101.1-PRKCys-I (D); pTF101.1-BBI (E) and pTF101.1-TI (F). Red arrows indicate the construct of interest.

2.4 Discussion

Transformation offers a strategy for introducing new genes or manipulating endogenous gene expression thereby allowing for the generation of changed phenotypes. This is invaluable in basic research to investigating gene function and in applied research for crop improvement (Somers et al., 2003). A total of five expression vectors were generated, three of which were designed to target over-expression of OC-I to the cytosol, nodules and chloroplast while the other two were designed to over-express the serine protease inhibitors (SPI), BBI and TI, in the cytosol. OC-I targeted to the cytosol has been previously done for tobacco (Gutiérrez-Campos et al., 2001, Van der Vyver et al., 2003). Cysteine proteases (CPs) have been shown to be involved in rubisco turnover in the leaves of tobacco plants under optimal and stress conditions (Prins et al., 2008). Successful OC-I expression vector construction to target the cystatin to the cytosol, the chloroplast and nodules and further expression of OC-I in a major legume crop species like soybean will improve our understanding of cysteine protease function particularly in the chloroplast and nodules.

Serine protease inhibitors (SPIs), preventing serine protease activity, have been previously explored for their involvement in defence against insects (Haq et al., 2004, Habib and Fazili, 2007). A recent study on *Solanum nigrum* (Hartl et al., 2010) applying RNA interference (RNAi) showed that this type of inhibitors defend plants against herbivores but seemingly did not influence plant growth and development. The isolated maize SPIs (BBI and TI) will have the potential for investigating the correctness for this finding by future investigation of phenotypic changes in soybean and Arabidopsis due to SPI over-expression. This will also allow studying in more detail the function of serine proteases targeted by their inhibitors

particularly under stress. The next chapter will report on the soybean transformation and characterization and selection of transformed soybean plants.

Chapter 3

Characterization and selection of transformed soybean plants

Part of this chapter has been published in the Plant Biotechnology journal entitled "Ectopic phytocystatin expression leads to enhanced drought stress tolerance in soybean (*Glycine max*) and Arabidopsis thaliana through effects on strigolactone pathways and leads to improved seed traits" By M. Quain; M. Makgopa; F. Magama; B. Márquez García; A. Montrose; D. Schnaubelt; N. Fernández Garcia; E. Olmos; K. Kunert; C. Foyer.

3.1 Introduction

When putative transgenic plants are obtained which show growth on a selective medium due to the expression of a selectable marker gene, it is important to characterize and select the transformed transgenic plants for stable integration of the transgene. PCR is a fast, simple and rapid first method of the detection of presence of the transgene in the genome following DNA extraction (Edwards et al., 1991, Gupta, 2008). The transcription of the gene can be detected using a more precise and sensitive method such as quantitative real-time PCR (qRTPCR) (Peirson et al., 2003). This method allows for rapid analysis of gene expression from low quantities of starting template. Analysis of data involves comparison of the transcription of the target gene relative to one or more reference genes i.e. house-keeping genes and a calibrator (Livak and Schmittgen, 2001, Schefe et al., 2006). Detection of gene expression can also be carried out at the protein level with Western blotting. This technique involves detecting the protein of interest using an antibody raised against the respective protein (Adugna and Mesfin, 2008).

The objective of this part of the PhD study was to characterize plants of different putative transgenic soybean lines generated from constructs that were provided by the University of Leeds to Iowa State University after production of T_0 generation of transformed plants selected on a BASTA-containing selection medium. Plants were selfed once to obtain T_1 generation seeds. Only plants expressing OC-I in the cytosol were selfed to obtain T_1 and T_2 generation plants. Plants derived from these seeds (T_2) were used to confirm the presence of the transgene in the plant genome using PCR. The expression level of the transgene on both the transcript and protein level was determined by qRTPCR and Western blotting with antiserum raised against OC-I, respectively. The activity of protease was monitored by SDS-

PAGE using mildly denaturing conditions and inclusion of gelatine in gels. In addition, fluorescence measurements were carried out with a fluorogenic substrate for cysteine proteases following specifically hydrolysis of the cathepsin L-like substrate Z-Phe-Arg-MCA to liberate fluorescent 7-amino-4-methylcoumarin.

3.2 Materials and Methods

3.2.1 Plant material and growth conditions

Seeds of soybean (*Glycine max* (L.) Merr.) cultivar Williams 82 together with transformed transgenic seeds of the cultivar over-expressing all the constructs (see chapter 2) were obtained from Iowa State University (USA) (Table 3.2). However; the cytosol OC-I seeds were used for all experiments because they were already on T_2 generation. The OC-I seeds (T_2) were sown in plastic pots (7.5 cm height and 8 cm diameter (top) and 6 cm diameter (bottom)) containing potting soil. Seedlings (10 days old) were transferred to larger pots (17.5 cm height and 20 cm diameter (top) and 13.1 cm (bottom). Growth was performed in controlled environment chambers with 15/9hrs light/darkness (intensity 800-1000 µmol $m^{-2}s^{-1}$) and a 26ºC/20ºC day/night temperature cycle and 60% relative humidity. Plants were grown for carrying out the experiments and they were grown to the sixth trifoliate stage before starting the experiments. Leaf material was harvested and weighed and harvested material was stored in -80ºC freezer for long-term storage.

3.2.2 DNA extraction and PCR

DNA was extracted from plant material using the method described by Edwards et al. (1991). Briefly, 30-50 mg leaf material was ground in liquid nitrogen and extraction buffer containing 200 mM Tris-HCl, pH 7.5, 250 mM NaCl, 25 mM EDTA and 0.5% SDS. The mixture was centrifuged at 13000 rpm and the supernatant was precipitated with isopropanol (Sigma-Aldrich, USA) and centrifuged again at 13000 rpm for 5 min. The DNA pellet was washed with 70% EtOH and re-suspended in 20 µL sdH_2O mixed with RNase A (25 µg/mL). PCR reaction was carried out on the extracted DNA according to section 2.2.1 using the OC-I primers FP (5'-ATG TCG AGC GAC GGA GGG-3') and RP (5'-TTG CAC TGG CTA CGA CAG GC-3') with primers annealing at 60ºC.

3.2.3 RNA extraction, cDNA synthesis and quantitative RTPCR

RNA was extracted from leaf material of soybean plants and cDNA synthesis was carried out following the method described in section 2.2.4.1. Quantitative RTPCR (qRTPCR) was performed using the QuantiFast™ SYBR® Green PCR kit (Qiagen, Germany) following the manufacturer's instructions and an ICycler® real-time PCR system (BioRad, USA). All reactions were conducted in triplicate with each PCR reaction consisting of 5 µL of diluted template cDNA (1/100), 1 µM of each primer, and 12.5 µL QuantiFast™ SYBR® Green master mix. The reaction volume was adjusted to 25 µL with nuclease-free water. Non-template control reactions were included. The excitation and emission of the SYBR® Green fluorescence was 494 nm and 521 nm, respectively. Cycling conditions consisted of a denaturation phase of 5 min at 95°C followed by 40 cycles each consisting 95°C for 10 s and annealing/elongation at 60°C for 30 s. Melting point analysis of the PCR products were

performed at 50°C to determine the specificity of the primers amplifying a single product. OC-I expression was normalized to an endogenous control (ELF1B) and primer sequences were designed with NCBI primer blast and Primer3 input (*http://www.ncbi.nlm.nih.gov/tools/primer-blast*) and (*http://frodo.wi.mit.edu*), respectively (Table 3.1).

Table 3.1 Primers used in quantitative real-time PCR (qRTPCR) of soybean plants.

Gene	Accession	Gene description	Oligonucleodes (5'-3')
Glycine max ELF1B	BM731462	Eukaryotic elongation factor 1-beta	Forward: GTT GAA AAG CCA GGG GAC A Reverse: TCT TAC CCC TTG AGC GTG G
Glycine max cystatin		Cysteine protease inhibitor	Forward: TCA CCG AGC ACA ACA AGA AG Reverse: CAT CGA CAG GCT TGA ACT CC

The output data from the qRTPCRs were processed with the CFX Manager software (Bio-Rad, UK). Quantifications were conducted using Microsoft Excel whereby the fold-change in gene expression is normalized to an endogenous reference gene and relative to the untreated control. Calculations were carried out with the $2^{\Delta\Delta Ct}$ method whereby $\Delta\Delta Ct = (Ct._{untransformed}-Ct._{reference\ gene})-(Ct._{transformed}-Ct._{reference\ gene})$ as described by Livak and Schmittgen (2001).

3.2.4 Construction of pGEX-OCI

The OC-I coding sequence was amplified by PCR (see section 2.2.1) with the following primers: FP 5'-ATAT GGA TCC ATG TCG AGC GAC GGA GGG CC-3' and RP 5'-ACTT GAA TTC TTA GAT GGG CCT TAG GCA TTT GC-3'. The underlined sequence indicates the position of the restriction sites *Bam*HI and *Eco*RI. The PCR product was cloned into the protein expression vector pGEX-4T-3 with the procedure outlined in 2.2.2 (Figure 2.2). DNA insertion was confirmed by restriction enzyme digestions (see section 2.2.2.2) and sequencing (GATC Biotech, UK) using the GST FP 5'-AAC GTA TTG AAG CTA TCC C-3' to show that the gene was in-frame with the fusion partner GST. The resultant recombinant plasmid was designated pGEX-OCI.

3.2.5 Protein production and analysis

3.2.5.1 Protein expression

Plasmid pGEX-OCI (5 µL) was used to transform (see section 2.2.2.3) 50 µL *E.coli* BL21 cells (Sigma-Aldrich, USA) for protein expression. A single colony was selected from the bacterial plate and inoculated in 5 mL LB medium supplemented with 100 µg/mL ampicillin. The culture was incubated O/N at 37ºC with 200 rpm shaking. One mL O/N culture was inoculated in 100 mL LB medium supplemented with 100 µg/mL ampicillin. Bacterial growth was allowed to take place until the OD_{600nm} of the culture was 0.6. An aliquot of this un-induced culture was harvested prior to protein expression with 0.1 M IPTG. The culture was then incubated for 5 hrs at 37ºC with 200 rpm shaking under the same conditions. The

induced culture (~80 mL) was harvested by centrifuging at 14 000 rpm at 4ºC for 10 min. The un-induced and induced cell pellets were stored at -20ºC.

3.2.5.2 Protein purification

Cell pellets from the induced cultures were freeze-thawed 3-times and re-suspended in 5 mL lysis buffer (50 mM Tris-HCl, pH 8.0, 5% sucrose (w/v), 50 mM EDTA and 0.1% Triton X-100). The lysis buffer was supplemented with freshly prepared lysozyme (1 mg/mL) and 1X protease inhibitor mix (Sigma-Aldrich, USA) before use. The suspension was incubated on ice for 5 min then centrifuged at 10 000 x *g* at 4ºC for 10 min. The supernatant was incubated with 250 µL of Glutathione Sepharose™ 4B (GE Healthcare) for 2-3 hrs at 4ºC with 35 rpm shaking. Sepharose was washed 3-times with PBS (140 mM NaCl, 2.7 mM KCl, 10 mM Na_2HPO_4, 1.8 mM KH_2PO_4, pH 7.3). An aliquot of the GST bound protein was eluted by adding 0.2 mL elution buffer (50 mM Tris-HCl pH 8.0, 10 mM reduced glutathione) and incubated at room temperature for 10 min with 35 rpm shaking. The mixture was centrifuged at 500 x *g* and the supernatant was collected and stored at -20ºC. The remainder of the GST bound protein was incubated with 250 µL PBS, pH 7.3, containing 5 µL thrombin enzyme (Sigma-Aldrich, USA) O/N with 35 rpm shaking. After incubation, the Sepharose was centrifuged at 500 x *g* for 5 min and the supernatant (purified protein) was collected and stored at -20ºC. The Sepharose was washed again with PBS, pH 7.3 and the supernatant collected separately.

3.2.5.3 Protein extraction and yield determination

The cotyledons and leaf material were ground into a powder in liquid nitrogen using a pestle and mortar. Protein was extracted by adding 50 mM Tris-HCl, pH 8, to the powder. The extract was centrifuged at 13 000 rpm for 10 min at 4ºC and the supernatant containing the protein was stored at -80ºC for further analysis. An aliquot of the protein samples was taken for calculation of the total protein yield.

Total protein content was determined with the Bradford (Bio-Rad, UK) assay reagent and BSA (Sigma-Aldrich, USA) as a standard. The OD_{595} (optical density) measurements were done with a spectrophotometer. Five concentrations of the BSA (1 mg/mL, 0.5 mg/mL, 0.25 mg/mL, 0.125 mg/mL and 0.0625 mg/mL) were prepared through a serial dilution in order to obtain a standard curve. From each standard, 5 µL was loaded in triplicate onto a clear 96 well plate and 250 µL of the Bradford reagent was added. The reaction was left at RT for 10 min and the optical density determined at 595 nm with a micro-plate reader (BMG FluoroStar Omega, Germany). The same method was applied for the protein extracts and water was used as a blank in the reaction.

3.2.5.4 Protein separation using SDS-polyacrylamide gel electrophoresis

Un-induced and induced GST-bound protein and purified protein were mixed with 4X NuPAGE® LDS Sample Buffer (Invitrogen, UK) containing 0.1% β-mercaptoethanol and boiled at 95ºC for 5 min. The protein samples were separated on a 15% polyacrylamide gel which was prepared from a 15% acrylamide/bis solution (29:1), 0.1% SDS, 0.1% ammonium persulfate (APS), 0.01% TEMED (Sigma-Aldrich, USA), 1.5 M Tris-HCl pH 8.8 and 0.5 M

Tris-HCl, pH 6.8). Electrophoresis was carried using the Bio-Rad system in SDS running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS) at 100 V for 3 hrs. After electrophoresis, the gel was stained in staining solution (0.1% Coomassie R-250, 10% acetic acid, 40% methanol) O/N. Gels were finally de-stained in de-staining solution (10% acetic acid, 40% methanol).

3.2.5.5 Western blotting for OC-I detection

Leaf protein samples (50 µg) were separated on a 15% SDS-PAGE. After electrophoresis, a wet transfer was carried out with the Hybond™-C Extra mixed ester nitrocellulose membranes (Amersham Biosciences, UK) soaked in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol). Transfer was carried out for 60 min at 110 V in a transfer chamber (Bio-Rad, UK). After transfer, the membrane was blocked in TBS-T (25 mM Tris, 150 mM NaCl, 2 mM KCl, 0.1% Tween, pH 7.4) with 5% fat free dry-milk and incubated on a horizontal shaker for 30 min. The blocking solution was replaced with the OC-I primary antibody (kindly provided by D. Michaud, Laval University, Quebec, Canada). The antibody was raised in rabbit by injecting the animal twice within 7 days with the affinity chromatography purified OC-I protein in the presence of incomplete Freund's adjuvant. Production of antibodies was monitored starting from six weeks after injection. The antibody was diluted 1:1000 in TBS-T with 5% fat free dry-milk and incubated overnight at 4°C under light shaking. The membrane was washed 3-times for 10 min with TBS-T to eliminate unspecific bound antibody. The membrane was incubated with a goat anti-rabbit secondary antibody (Sigma-Aldrich, USA) coupled to horseradish peroxidase (HRP) diluted 1:5000 in TBS-T with 5% fat free dry-milk for 60 min. The membrane was washed 3-times for 10 min with TBS-T. The presence of OC-I protein was detected using the Immun-Star™ WesternC

Chemiluminesent according to the manufacturer's instructions (Bio-Rad, UK). Briefly, the membrane was covered with equal volumes of enhanced chemiluminescence (ECL) solutions A and B for 5 min. The membrane was transferred to a cassette and a CL-XPosure (ThermoScientific, UK) placed on top in the dark. The film was developed with a Konica SRX-101A (Konica Minolta).

3.2.4.6 Protease activity

In-gel protease assay was first carried out according to a method described previously (Michaud et al., 1996). Briefly, for detection of any major protease activity in the cotyledons and leaf samples, a 15% SDS-PAGE which had been co-polymerized with 0.1% gelatine was carried out to provide a rough indication of any major difference in protease activity. Prior to electrophoresis, extracts (20 µg) were mixed with 4X NuPAGE® LDS sample buffer without β-mercaptoethanol and were not boiled before loading to prevent destruction of secondary structure proteins. After electrophoresis on ice at 100 V, the gel was incubated in 2.5% Triton X-100 for 30 min at room temperate to re-nature the proteases. The gel was rinsed 3-times and incubated in proteolytic buffer (0.1 M citrate phosphate, pH 6.0 containing 10 mM L-cysteine) at 37ºC for 3 hrs. After incubation, the gels were stained in staining solution O/N. Gels were de-stained and protease activity was visualized as clear bands on a blue background. For cysteine protease activity, protein extract from non-transformed plants was pre-incubated with 1% E-64 (trans-epoxysuccinyl-L-leucylamido-(4-guanidino) butane), a cysteine protease inhibitor, for 15 min at 37ºC before loading. The peak areas of the protease activity visualized on the gels generated were analysed using the software ImageJ (*rsbweb.nih.gov/ij*). The gel image was converted to 8-bit and the different protease bands

(clear bands) were selected. A plot was then generated from which the peak area was obtained.

Protein extracts from the cotyledons, 6 and 18 weeks old plants were also used for measuring cysteine protease activity with the synthetic substrate Z-Phe-Arg-MCA for detecting cathepsin L-like activity. This fluorometric substrate, when hydrolysed by a protease, releases bound fluorescent α-amino 4-methylcoumarin (MCA). The MCA is measured with fluorescence spectrophotometer (BMG FluoroStar Omega, Germany). Hydrolysis of the substrate by the protease was monitored with progress curves as described by Salvesen and Nagase (1989). For detection of cathepsin L-like activity, the reaction contained 10 µL (1 µg) total soluble protein extract, 8 µL from a 100 µM substrate stock dissolved in DMSO and 82 µL of 100 mM sodium phosphate buffer, pH 6.5 with 10 mM L-cysteine freshly added for cysteine-like activity. Cysteine protease activity was monitored with excitation and emission at 360 nm and 450 nm, respectively. Reaction rates represented by the slope of the curve were recorded as Fluorescence Units (FU) per unit time. All reactions were carried out in triplicate. As a control, the reaction contained the buffer and substrate but no protein extract.

3.2.6 Statistical analysis

Comparisons were made between transgenic and wild-type plants. Mean values of the gene expression and fluorometric protease activities were taken from measurements of three replicates and the standard error of the means was calculated using the Student's *t*-test at stringency level of $p < 0.05$. One-way ANOVA was used for all other statistical analysis. Tukey HSD (honestly significant difference) was used as a post-hoc test at stringency level of

$p < 0.05$. All statistical analysis was performed using SPSS v.13 for windows (Statistical Package for Social Sciences, Chicago).

3.3 Results

3.3.1 OCI fusion protein production and purification

The OC-I gene was cloned into the pGEX-4T-3 for protein expression. This plasmid has a GST-tag which allows for the formation of fusion proteins which can be used for purification. The success of the cloning process was verified by restriction enzyme digestions of the pGEX-OCI construct with restriction enzymes *Bam*HI and *Eco*RI and a combination of the two enzymes to confirm presence of OC-I (Figure 3.1A, Lane: 6). Presence of a 300 bp band after digestion confirmed OC-I cloning into pGEX-4T-3 to create pGEX-OCI. Sequencing of the plasmid verified that OC-I was in frame with GST (Figure 3.1B). Expression of the OC-I fusion in *E.coli* BL21 cells was carried out by inducing cells with IPTG after cell density had an OD_{600nm} of 0.6. Induction was carried out for 5-6 hrs. Production of an expected 37 kDa GST-fusion protein was found in induced cell after SDS-PAGE. This fusion protein was not found in un-induced cells (Figure 3.2, Lane: 3).

Purification of the GST-fusion proteins was carried by applying the freeze-thaw method after lysozyme treatment of the cells. The supernatant carrying the soluble fusion protein was mixed with GSH-Sepharose™ 4B and binding of the GSH and the GST-fusion protein was allowed to take place. Bound GSH-Sepharose-GST fusion protein was collected by centrifugation. The GST tail was released from the GST-fusion protein by enzymatic cleavage with thrombin protease and purified protein (GSTF and OC-I) were separated on a

15% SDS-PAGE. Presence of a 37 kDa (Figure 3.2, Lane: 5) and 11 kDa product (Figure 3.2, Lane: 6) for GSTF and OC-I, respectively, confirmed purification of OC-I. As a control, purified GST was included which gave a 26 kDa product (Figure 3.2, Lane: 4).

3.3.2 OC-I detection and transcription in OC-I plants

Genomic DNA extraction was carried from wild-type plants and plants of transgenic soybean lines 55, 56 and 57 for PCR amplification using OC-I primers. Expected 300 bp fragments were amplified and sequence analysis confirmed the OC-I coding sequence (Figure 3.3). Transcription of OC-I was confirmed by qRTPCR analysis and plants of the three transgenic lines had higher transcript amounts than wild-type control plants. Since OC-I primers used were possibly not highly specific for OC-I, they possibly also amplified closely related endogenous soybean cystatins and detection of transcription of these were set at 1 (Figure 3.4). However; OC-I transcription was different in plants of the same line and also different between selected lines (Figure 3.4A and B). A significantly higher ($p < 0.05$) transcription was found in plants of transgenic line 55 and plants of this line were later used for phenotypic analysis. Overall, no significant differences were found in all the lines when comparing all the data using ANOVA due to the high variation between the lines.

3.3.3 Immuno-detection of OC-I

Protein extracts from the control plants and transgenic lines 55, 56 and 57 were separated on a 15% SDS gel and OC-I expression was detected by Western blotting (immuno-blotting) with an OC-I antibody raised in rabbit. A hybridization product with the size of 11 kDa was detected in plant extracts of transgenic plants which was absent in extracts derived from wild-

type control extracts (Figure 3.5). Purified OC-I, as described in section 3.2.5.2, was used as a positive control for OC-I hybridization.

3.3.4 Proteolytic activity

3.3.4.1 Protease activity gels

Protease activity was visualized by separating protein extracts from control plants and plants of transgenic lines 55, 56 and 57 on a 15% gelatine-containing SDS gel. Gelatine is a substrate that when degraded by proteases from the protein extract results in clear bands on a blue poly-acrylamide gel. After electrophoresis, two clear bands (protease activity P1 and P2) were observed (Figure 3.6). Figure 3.6A shows protease activity of control plants and of lines 55 and 56. The intensity of band P2 was less intense in plants of transgenic line 57 than in the control indicating some OC-I activity blocking cysteine protease activity (Figure 3.6B). When E-64, an inhibitor of papain-like cysteine proteases and also trypsin, was added to control extracts, band intensity of P2 greatly decreased indicating that P2 represents cysteine protease activity.

The peak areas of protease activity (P1 and P2) were determined using ImageJ software. This software selects areas of interest and determines peak areas. Figure 3.7A shows peak areas obtained for protease activity P1. All samples (control, lines 55, 56, 57 and E-64) had similar peak areas as calculated by ImageJ indicating that P1 does not represent OC-I or E-64 sensitive cysteine proteases. Transgenic line 57 displayed a trend of lower peak area similar to that of E-64 supporting the previous findings that P2 represents OC-I and E-64 sensitive cysteine proteases.

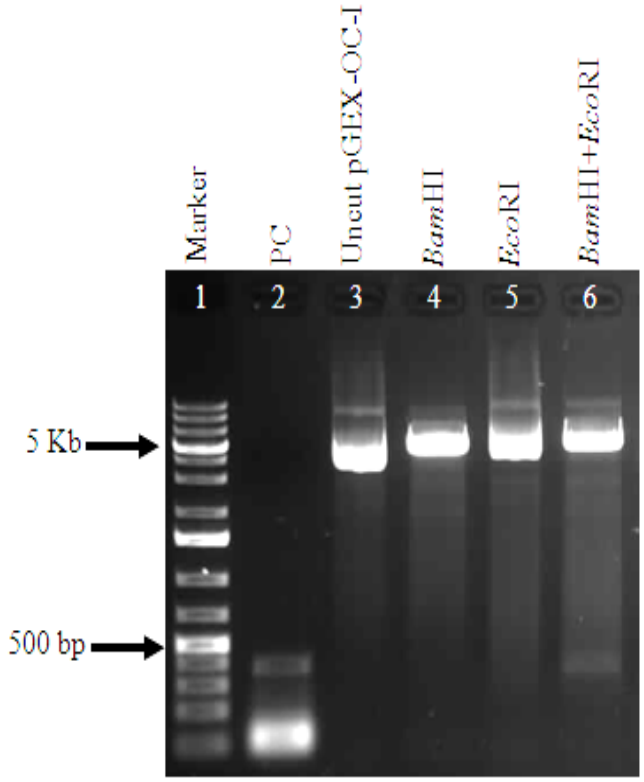
3.3.4.2 Fluorometric assay

Hydrolysis of Z-Phe-Arg-MCA was measured to detect cathepsin L-like (cysteine protease) activity in extracts of control plants and plants of transgenic lines. No significant differences were found in the different leaf tri-foliates (TF) 1-6 of 6 weeks old plants (Figure 3.8 B-E). Senescing leaves from transgenic plants, identical to cotyledons, had lower cathepsin L-like activity than the control.

Table 3.2 Various types of transgenic lines produced. Lines with OC-I expressed in the cytosol were used in this study.

T1 transgenic lines		
Soybean lines	**No. of lines**	**No. of seeds**
OC-I cytosol	8	348
OC-I chloroplast	9	1098
OC-I nodules	15	860
BBI cytosol	9	988
TI cytosol	10	732
T2 transgenic lines		
OC-I cytosol	4	562

A



B

ATCAGCAGTATATAGCATGGCCTTTGCAGGGCTGGCAGCCAC
GTTTGGTGGTGGCGACCCCTTTAAAAATCGGATCTGGTTCCG
CGTGGATCCATGTCGAGCGACGGAGGGCCGGTGCTTGGCGG
CGTCGAGCCGGTGGGGAACGAGAACGACCTCCACCTCGTCG
ACCTCGCCCGCTTCGCCGTCACCGAGCACAACAAGAAGGCC
AATTCTCTGCTGGAGTTCGAGAAGCTTGTGAGTGTGAAGCA
GCAAGTTGTCGCTGGCACTTTGTACTATTTCACAATTGAGGT
GAAGGAAGGGGATGCCAAGAAGCTCTATGAAGCTAAGGTCT
GGGAGAAACCATGGATGGACTTCAAGGAGCTCCAGGAGTTC
AAGCCTGTCGATGCCAGTGCAAATGCCTAAGGCCCATCTAAG
AATTC

Figure 3.1 Verification of the presence of OC-I in pGEX-OCI. Agarose gel showing restriction enzyme digestion of pGEX-OCI (A). Sequence of the OC-I gene in pGEX-OCI using the GST primer (B). The underlined sequence is part of the GST tag, the blue and purple is the *Bam*HI and *Eco*RI restriction enzymes, respectively. The sequence highlighted in red is the OC-I gene. The positive control (PC) is the OC-I amplicon.

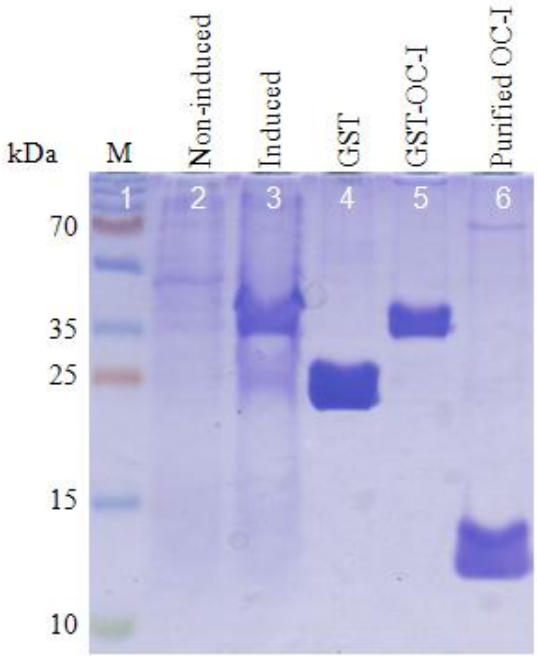


Figure 3.2 A 15% SDS-PAGE showing protein expression and purification of the OC-I.

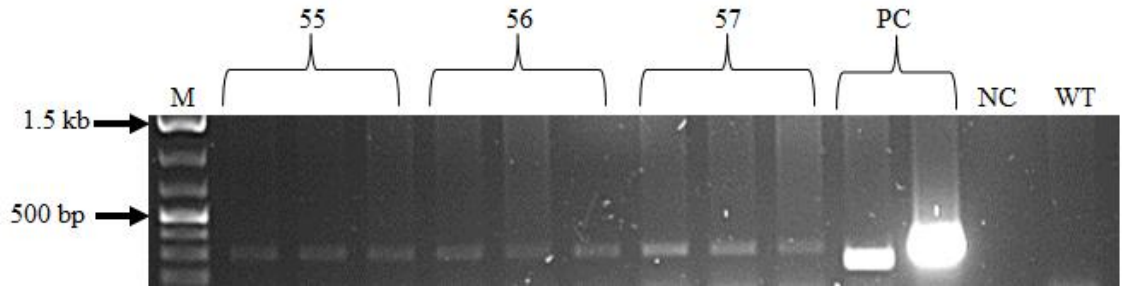


Figure 3.3 Representative agarose gel of PCR amplification of the OC-I gene. Amplification was carried out on the transgenic lines 55, 56, 57 and positive plasmid control (PC). Water negative control (NC) and wild-type (WT) soybean were also included.

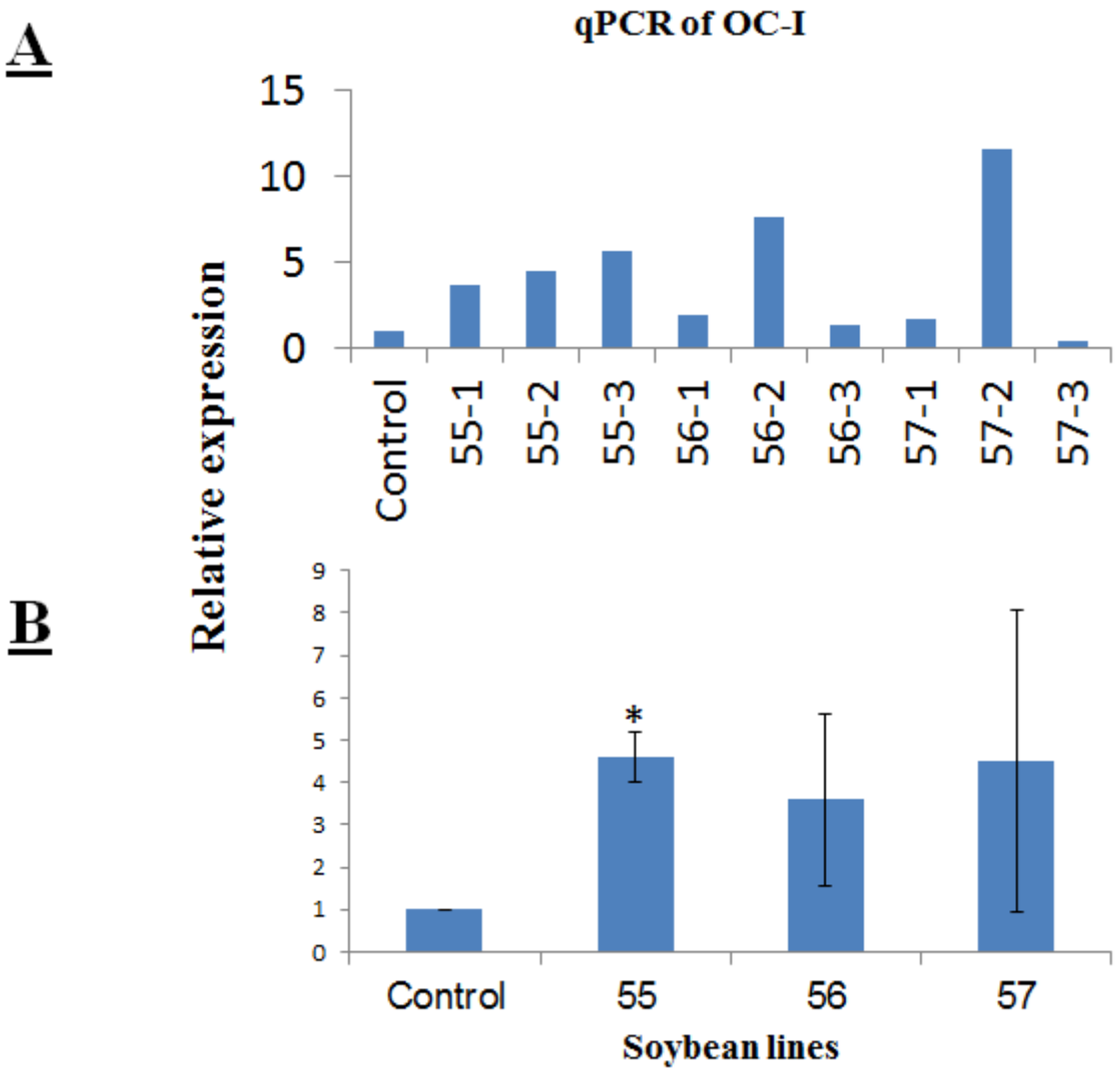


Figure 3.4 Relative expression of the OC-I transgene in leaves of soybean lines. Wild-type control plants and the three transgenic lines (55, 56, and 57) individually represented (A) and pooled (B). Bars are means ± SE of three plants (*$p < 0.05$ significant).

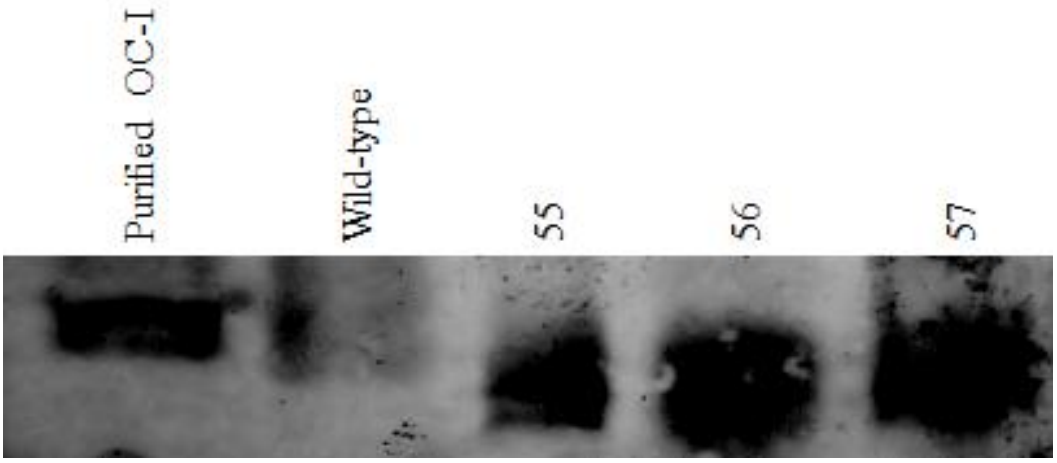


Figure 3.5 Western blot analysis for the detection of the oryzacystatin I (OC-I) in transgenic soybean leaves using the antibody raised against OC-I in rabbit.

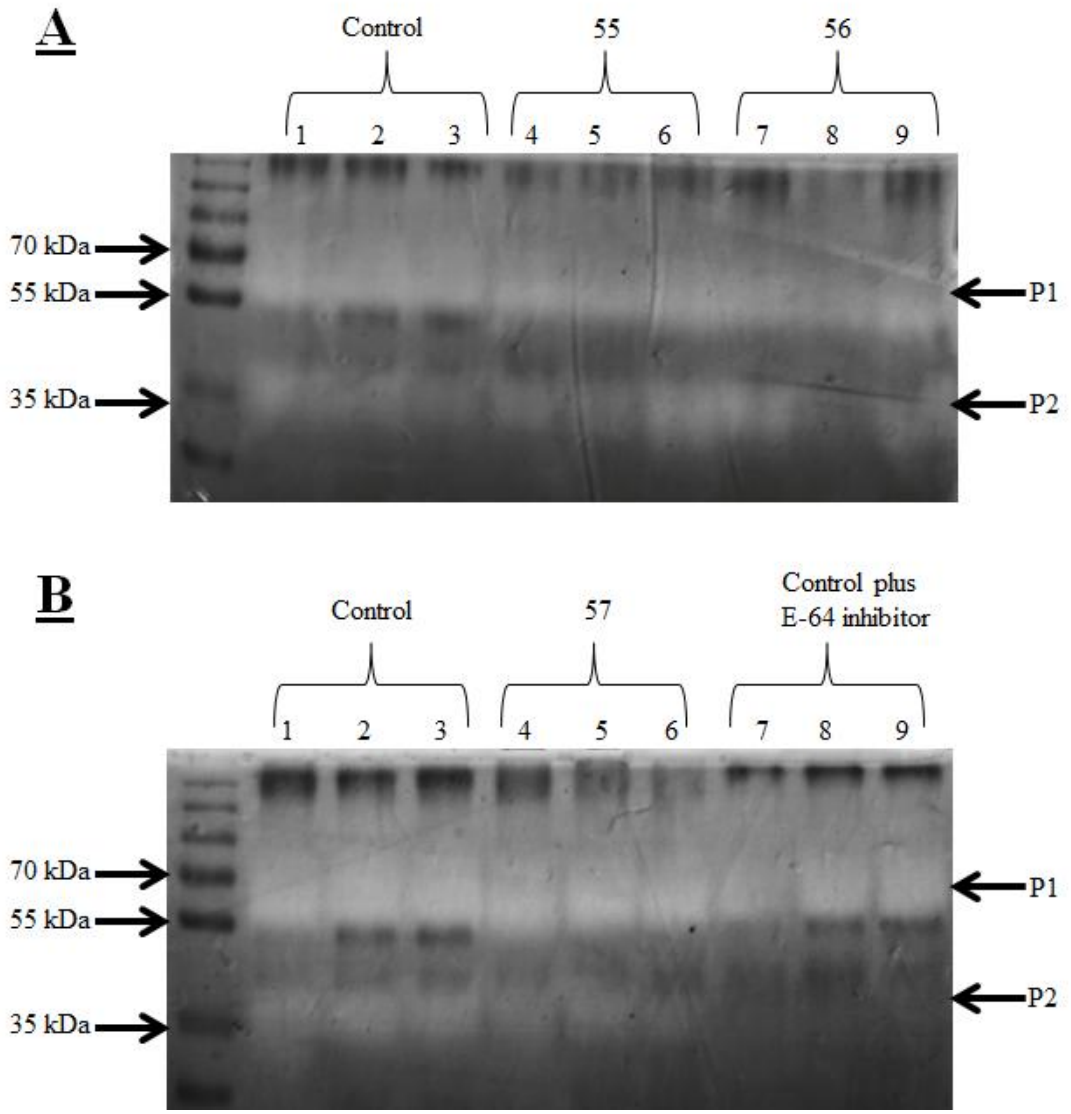


Figure 3.6 Gelatine-containing SDS-PAGE gels of control and transgenic soybean leaves from lines 55 and 56 (A) and 57 (B). Clear bands (P1 ~60 kDa and P2 ~35 kDa) on the background indicate proteolysis of gelatine by proteases in the extracts. P1 = protease activity 1, P2 = protease activity 2. For comparison control samples were also treated with the cysteine protease inhibitor E-64.

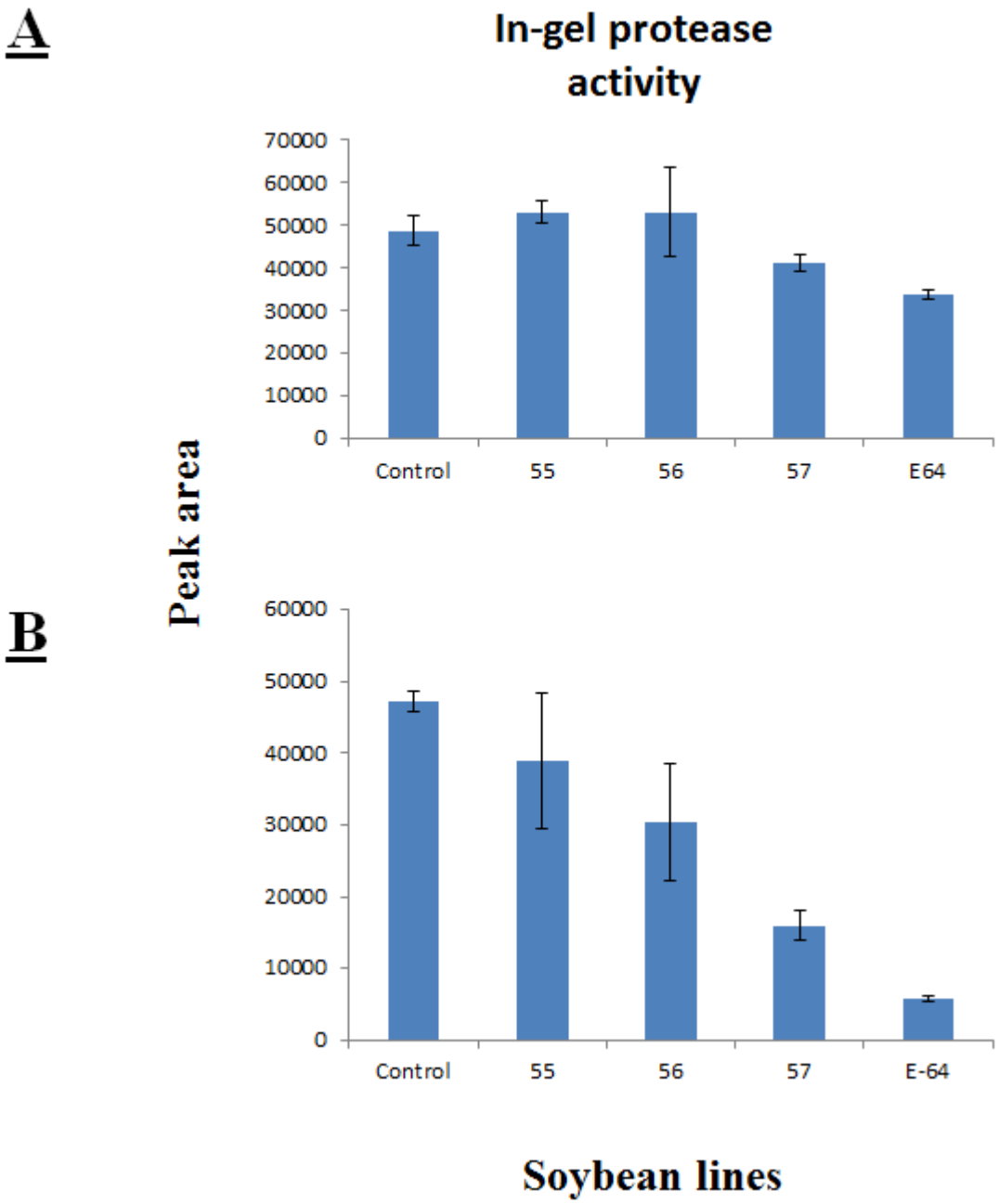


Figure 3.7 In-gel protease determination. (A) Gelatine-containing SDS-PAGE quantification of the upper protease band P1 (~60 kDa) and the lower protease band P2 ~35 kDa (B) and E-64 treated control. Error bars indicates mean ± SE of three replicates.

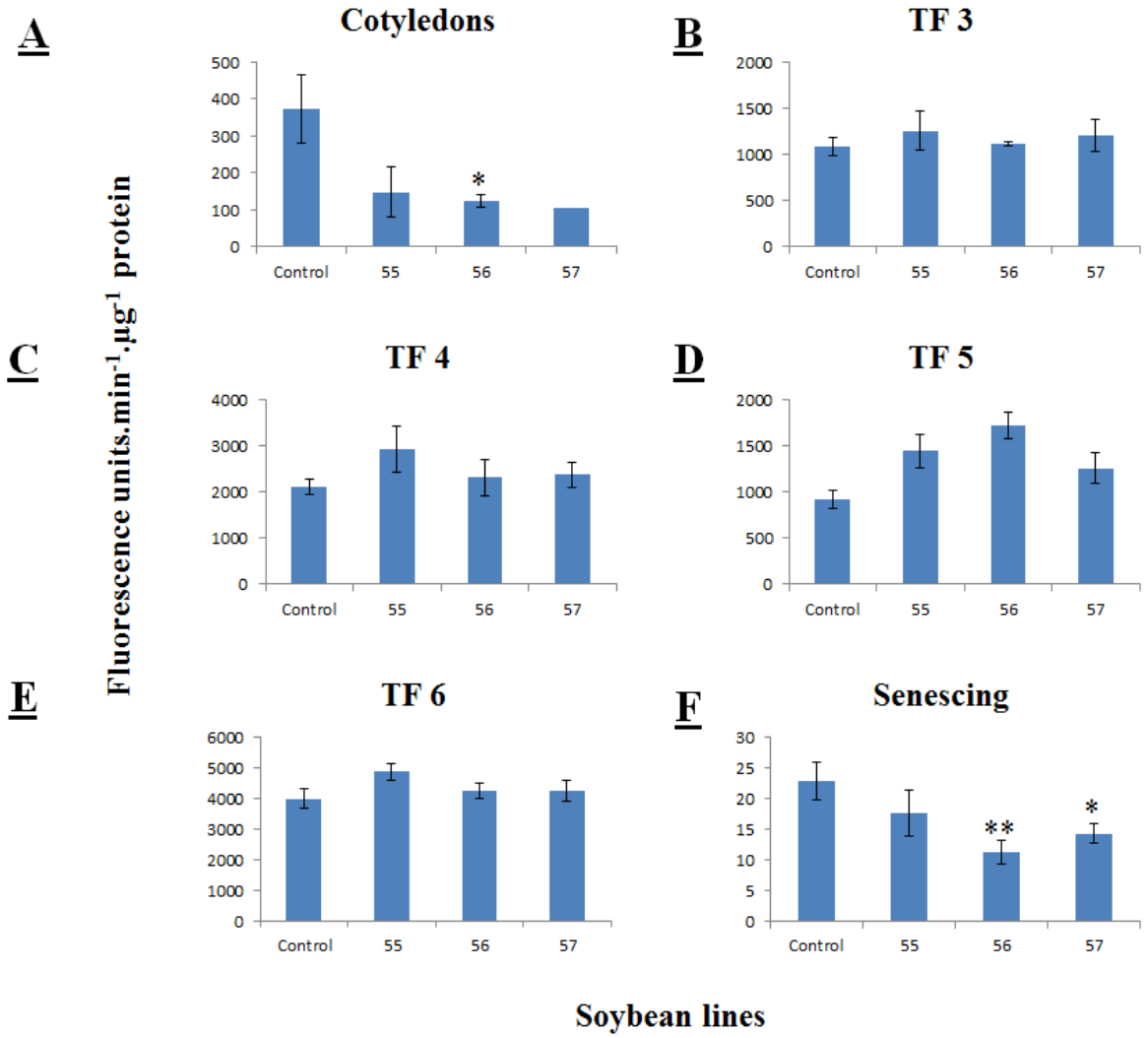


Figure 3.8 Soybean cathepsin L-like cysteine protease activity quantification of cotyledons (A), leaf tri-foliate (TF) 3, TF 4, TF 5, TF 6 and senescing leaves (B-F) (*$p < 0.05$ significant and **$p < 0.01$ highly significant). Error bars indicate mean ± SE of three replicates.

3.4 Discussion

This part of the study showed that the OC-I gene was present and stably integrated into the soybean genome. This was confirmed by PCR amplification of the OC-I coding sequence allowing selection of three different transgenic lines containing the OC-I coding sequence. Expression analysis applying qRTPCR and Western blotting further confirmed that OC-I was expressed in the plants of these three lines, which were named lines 55, 56 and 57. Expression of a transgene can be influenced by a number of factors which include copy number, location in the genome, methylation, re-arrangement of the transgene (Muskens et al., 2000, Lechtenberg et al., 2003, Tang et al., 2007). Differential expression in transgenic OC-I lines, as found in this PhD study, has been also observed in an earlier study on tobacco plants were different lines had different OC-I protein amounts after Western blotting (Van der Vyver et al., 2003).

Transgenic line 57 exhibited the strongest hybridization signal when compared to the other two lines, 55 and 56. The molecular weight of the transgenic lines was found to be smaller in size when compared to purified OC-I. This could be possibly due to partial degradation (or modification) of OC-I which is not native to soybean. However; if degradation occurs in soybean and if this affects activity has not been investigated yet. Another possibility may be due to the protein structure and composition that may affect both detergent-loading levels and polypeptide-SDS-PAGE migration rates (Guenoune et al., 2002, Rath et al., 2009). Also, disulfide bonds reduce SDS binding to globular proteins and have been linked to the fast migration of unreduced vs reduced protein possibly due to the imposition of a more compact shape by the disulfide bonds (Rath et al., 2009). Plants of line 57 also had lower cysteine protease activity detected by gelatine-containing SDS gels. In contrast, no significant

difference was observed in tri-foliates of all tested lines when protease activity was determined using a cathepsin L-like substrate. This might be attributed to the fact that the substrate used is able to react with several proteases having cathepsin L-like activity, but are insensitive to OC-I. Therefore, OC-I sensitive cysteine proteases with cathepsin L-like activity might not be the dominant cysteine proteases in soybean leaves. Interestingly, TF 3 (oldest leaf) for all lines had small amounts of cathepsin L-like protease activity while TF 6 (youngest) had almost four times more cathepsin L-like protease activity compared to that of TF 3 for all lines. This may suggest that the papain-like activity is required at different stages of leaf development similar to the result observed of papain-like proteases at different plant developmental stages (Beyene et al., 2006). In addition, their inhibitors have also been found to be differentially expressed during plant development when a GUS reporter system was used (Hwang et al., 2010). Furthermore, plants of transgenic lines also had lower cysteine protease activity in the cotyledons and senescing leaves compared to wild-type plants confirming earlier work by Ling et al. (2003) and Beyene et al. (2006), respectively. This suggests that OC-I in the transgenic lines was active and inhibited the activity of cysteine proteases in the cotyledons and senescing leaves similar to the cysteine protease inhibitor E-64.

In the next chapter, Arabidopsis plants will be characterized under well-watered and drought stress conditions are described.

Chapter 4

Characterization of transgenic Arabidopsis plants under well-watered and drought conditions

Part of this chapter has been published in the Plant Biotechnology journal entitled "Ectopic phytocystatin expression leads to enhanced drought stress tolerance in soybean (*Glycine max*) and Arabidopsis thaliana through effects on strigolactone pathways and leads to improved seed traits" By M. Quain; M. Makgopa; F. Magama; B. Márquez García; A. Montrose; D. Schnaubelt; N. Fernández Garcia; E. Olmos; K. Kunert; C. Foyer.

4.1 Introduction

Arabidopsis plants have been extensively used as a model plant for studying and understanding the molecular and biological processes in other plant species. Arabidopsis plants have many advantages such as a short life cycle, can be produced in the lab, easily transformable, availability of mutants, easy crossings, genomics and functional genomics data available and have a rather small genome (Clough and Bent, 1998, Meinke et al., 1998, Koornneef and Meinke, 2010). The genome of Arabidopsis has over 800 protease sequences (MEROPS peptidase database, http://merops.sanger.ac.uk) that are distributed over almost 60 families belonging to 30 different clans (van der Hoorn, 2008). Several senescence-associated proteases have been identified and shown to be up- or down-regulated by various stresses and natural senescence (Roberts et al., 2012).

Subsequently, protease inhibitors are also up- or down-regulated due to the imposition of stress from biotic and abiotic factors as a means of controlling protease expression. Over-expression of two papain-like cysteine protease inhibitors have been shown to improve tolerance to salt, drought and cold stress in Arabidopsis plants (Zhang et al., 2008). Furthermore, two Arabidopsis phytocystatins had enhanced expression when plants were exposed to abiotic stress during plant development (Huang et al., 2007, Hwang et al., 2010a). Conversely, over-expression of a sweet potato cysteine protease resulted in increased sensitivity to drought in Arabidopsis plants (Chen et al., 2013).

The aim of this part of the PhD study was to transform and characterize OCI-expressing Arabidopsis plants to confirm the previous soybean results with the model plant Arabidopsis. The OC-I transgene was expressed under the control of a 35S promoter and the objectives of

this study were to: (i) characterize the transgenic lines by selecting for BASTA resistance; (ii) determine the presence and expression of the transgene using PCR and qRTPCR, respectively; (iii) determine the effects of drought stress on the physiological and biochemical processes in transgenic plants.

4.2 Materials and Methods

4.2.1 Plant growth and maintenance

Arabidopsis ecotype col-0 plants were grown in a controlled environment at 20-25°C with a photoperiod of 16/8 hrs light/dark cycle. Pots of 64 cm^2 were prepared containing a mixture of wet soil and vermiculite (1:1). Seeds were sprinkled onto the pots. Primary inflorescence buds were clipped to allow formation of secondary inflorescence buds and increase the transformation process by obtaining more flower buds per plant. The plants were grown for eight weeks before being used for floral dip transformation.

4.2.2 Plasmid and *Agrobacterium* strain

The plasmid pTF101.1-Cys-I and *Agrobacterium tumefaciens* GV3101::pM90 strain was used. The plasmid conferred resistance against spectinomycin and expression of the OC-I was under the control of a 35S promoter (Annex 1A).

4.2.3 Transformation of Arabidopsis

4.2.3.1 Generation of Agrobacterium competent cells

The plasmid was transformed into the *A.tumefaciens* strain GV3101::pM90 using a modified freeze-thaw method described previously (Hofgen and Willmitzer, 1988). Briefly, the GV3101::pM90 strain was grown on an agar plate supplemented with 50 µg/mL rifampicin and 50 µg/mL gentamycin at 28ºC for two days. A single colony was picked and used to inoculate 2 mL selective LB (supplements as above) which was shaken at 200 rpm O/N at 28ºC. Half a millilitre O/N culture was used to inoculate 50 mL selective medium LB and allowed to grow at 28ºC with 200 rpm shaking until OD_{600nm} = 0.5-1.0. The culture was chilled on ice for 5 min and the cells were collected by centrifuging at 3000 rpm for 5 min at 4ºC. The cells were re-suspended in 1 mL ice-cold 20 mM $CaCl_2$ (dissolved in H_2O). Aliquots of 0.1 mL were dispensed into tubes and kept on ice for an hr.

4.2.3.2 Agrobacterium-transformation

The plasmid DNA (1 µg) was added to a tube of competent cells and mixed gently. The tubes were frozen in liquid nitrogen and thawed for 5 min at 37ºC. This was followed by the addition of 1 mL LB and incubation at 28ºC with shaking for 2 hrs. Cells were collected by spinning for 5 min at 4000 rpm and re-suspended in 0.1 mL LB. The suspension was plated out on a selection plate containing rifampicin, gentamycin and 100 µg/mL spectinomycin and incubated for 2 days until transformed colonies were visible.

4.2.3.3 Generation of transgenic plants

Agrobacterium tumefaciens strain GV3101 carrying the plasmid pTF101.1-Cys-I was used to transform Arabidopsis plants (Clough and Bent, 1998). *Agrobacterium* cells containing pTF101.1-Cys-I were grown in LB liquid medium supplemented with 50 µg/mL rifampicin and 100 µg/mL spectinomycin at 25-28ºC, 200 rpm O/N. Cells were harvested by centrifugation at 5000 rpm and re-suspended in 5% sucrose containing 0.05% silwett L-77 (Lehle seed, USA). The bacterial suspension was transferred into a beaker and the secondary inflorescences of the Arabidopsis plants (section 5.2.1) were submerged into the bacterial suspension for 5-10 s by inverting the Arabidopsis pots. Pots were placed in an upright position and covered O/N. The cover was removed the following day and Arabidopsis plants were allowed to grow in the greenhouse 8/16 hrs light/dark cycle at 20-25°C. Harvesting of the seeds was done when siliques were dry and seeds were stored at 4^{o}C.

4.2.3.4 Selection of transgenic lines

4.2.3.4.1 BASTA selection and PCR

Four weeks old plants (T_0) were selected using 200 µM glufosinate ammonium (BASTA) (Sigma-Aldrich, USA) by leaf painting. Presence of the transgene was determined by PCR (section 2.2.2.1) using the following primers FP: 5'-TCA CCG AGC ACA ACA AGA AG-3' and RP: 5'-AGC TCC TTG AAG TCC ATC CA-3'.

4.2.3.4.2 RNA extraction, cDNA synthesis and quantitative real-time PCR (qRTPCR)

RNA extraction and cDNA synthesis were carried out following the method described in section 2.2.4.1 while quantitative RTPCR was carried out as outlined in section 3.2.3. The output data from the qRTPCRs were quantified according to section 3.2.3.

The expression of the OC-I was normalized to an endogenous Arabidopsis control: putative mRNA gene (PDF2). Accession and primer sequences that were designed using NCBI primer blast and Primer3 input (*http://www.ncbi.nlm.nih.gov/tools/primer-blast*) and (*http://frodo.wi.mit.edu*), respectively are listed below: Table 4.1.

Table 4.1 Primers used in real-time quantitative real-time PCR (qRTPCR) of Arabidopsis plants.

Gene	Accession	Gene description	Oligonucleodes (5'-3')
Arabidopsis PDF2	At1g13320	Protein phosphatase 2A regulatory subunit	Forward: TAA CGT GGC CAA AAT GAT GC Reverse: GTT CTC CAC AAC CGC TTG GT
Arabidopsis cystatin		Cysteine protease inhibitor	Forward: TCA CCG AGC ACA ACA AGA AG Reverse: AGC TCC TTG AAG TCC ATC CA

4.2.4 Drought tolerance analysis of transgenic plants

The T_3 pTF101.1-Cys-I transgenic and non-transgenic Arabidopsis seedlings were transplanted into pots containing soil. The seedlings were grown in a growth chamber for 4 weeks under short day conditions (8/16 hrs, light/dark) at 22ºC and 50-60% relative humidity. Plants from three independent transgenic lines and non-transgenic plants were used to carry out the experiments. At week 4, half the plants were maintained under water-replete conditions and water was withheld from the other half for 15 days. After which leaf samples were collected for independent experiments. Overall survival percentage was calculated as (number of plants survived/number of plants tested) x 100. No statistical data are available for this assay as survival was counted as the overall number of plants.

4.2.5 Leaf number, biomass, leaf and soil water content

The number of leaves was counted individually from three representatives of each line before the leaves were weighed for biomass determination. Dry biomass was determined after exposure of plant parts (shoots) to a temperature of 80ºC for 48 hrs in a drying oven (Type U 40, Mommert, Germany). For measurement of leaf and soil water content, the fresh weight of leaf material was measured using a Model B-502-S Metter Toledo balance (Switzerland). The samples were then placed into a drying oven at a temperature of 80^{o}C for 48 hrs. The same was done for the soil samples corresponding to each plant. Leaf and soil water contents were calculated as the difference between the first and second measurements as: WC (%) = [(fresh weight - dry weight)/fresh weight)] x 100.

4.2.6 Protein extraction, yield determination and protease activity

Leaf materials (whole rosette) from well-watered and drought treated plants were ground into a powder in liquid nitrogen using a pestle and mortar. Protein extraction and yield determination were carried out according to section 3.2.5.3. In-gel protease assay and fluorometric assay were determined as described in section 3.2.4.6.

4.2.7 Statistical analysis

Statistical significance between parameters on well-watered and drought treated plants was determined using the Student's *t*-test at stringency level of $p < 0.05$ as outlined in section 3.2.6. Mean values were taken from three replicates of each line.

4.3 Results

4.3.1 Screening and selection of transformed Arabidopsis plants

Arabidopsis plants were transformed using the *Agrobacterium*-mediated method described by Clough and Bent (1998). The plants were transformed with a plasmid encoding the OC-I open-reading frame conferring resistance to the herbicide BASTA. Figure 4.1A shows a representative image of the Arabidopsis leaves after leaf painting. Some of the leaves of the Arabidopsis plants were slightly yellow with some partially herbicide-resistant green patches (Figure 4.1A, arrows 2 and 4). Those that were completely yellow (Figure 4.1A, arrow 3) were classed as susceptible to the herbicide i.e. did not carry the plasmid. Those that stayed

green (Figure 4.1A, arrows 1) following leaf painting were classed as resistant to the herbicide and most likely carry the OC-I expressing plasmid.

DNA was extracted from leaves of the plants which showed partial and complete resistance to the herbicide and PCR was carried out. For selection of transgenic lines, OC-I specific primers were designed to differentiate between plants with the transgene and those without. Amplification of the 200 bp band was indicative of the presence of the gene in the genome of the plant (Figure 4.1B). DNA from plants that did not amplify was used as control (non-transgenic) plants throughout the experiments. For determination of transcript levels of the OC-I, primers were designed that amplified OC-I possibly along with endogenous Arabidopsis cystatins to determine if transgenic plants expressed overall more cystatin transcripts. (Figure 4.1C). There was differential expression of the OC-I gene amongst the samples using non-transgenic plants as control samples. These samples were taken through to T_3 generation and were tested for drought tolerance.

4.3.2 Phenotypic analysis of drought-treated plants

Non-transgenic (control) and transgenic plants (4-wks-old) were divided into two groups, well-watered and the drought-treated. Drought treatment was carried out by withholding water for 15 days. Figure 4.2 shows a photograph that corresponds to plants at day 15 of water-deficit stress. The well-watered group of plants (left) are those that were watered throughout the experiment whilst the drought group (right) are those were water was withheld. The transgenic lines had higher survival percentage compared to the non-transgenic (control) plants with line Cys 4 having the highest survival percentage of 93.3%.

The water content of the detached leaves was determined (see 4.2.5). The leaf water content was identical in the well-watered control and transgenic plants ($p > 0.05$). The non-transgenic control plants showed a highly significant ($p < 0.01$) decrease in water content after drought (Figure 4.3A) compared to the transgenic lines subjected to the same drought conditions. The transgenic lines (Cys 1, 3 and 4) had only little decrease in water content after drought as compared to the water content under well-watered conditions. However; in Cys 3 this decrease was significantly lowered under drought compared to well-watered conditions ($p < 0.05$).

The soil water content did not differ amongst well-watered non-transgenic and transgenic plants. All plants showed a highly significant ($p < 0.01$) decrease in soil water content after drought (Figure 4.3B) confirming the imposition of drought throughout all plants. Comparing the soil water content in the non-transgenic and transgenic plants after drought revealed a significantly higher ($p < 0.05$) content for the transgenic plants.

All leaves were detached from the well-watered group of non-transgenic control plants and transgenic plants and immediately counted. Overall, the non-transgenic plants had more leaves than the transgenic lines (Figure 4.4A). Amongst the transgenic lines, Cys 1 had significantly more leaves (60) than Cys 3 and Cys 4 both having around 35 leaves.

The effects of drought on the biomass of the plants was also determined (Figure 4.4). In the well-watered plants, the fresh weight biomass of leaves of the non-transgenic plants was greater than that of transgenic line Cys 1 but not significantly so for transgenic lines, Cys 3 and Cys 4 (Figure 4.4B). Drought significantly ($p < 0.05$) decreased the fresh weight biomass of the non-transgenic and transgenic plants, but to a greater extent in the non-transgenic lines

(Figure 4.4B). The analysis of dry weight biomass showed no significant differences ($p > 0.05$) for non-transgenic and transgenic lines under well-watered and drought conditions (Figure 4.4C).

Total protein content was determined from well-watered and drought-treated plants (Figure 4.5A). Under well-watered conditions, the protein content of the transgenic lines was higher than that of the non-transgenic lines with Cys 3 being significantly higher ($p < 0.05$). However; the control plants had significantly higher protein content after drought treatment while the total protein content in the transgenic lines remained unchanged. The papain-like cysteine protease activity of the non-transgenic lines greatly decreased under drought conditions compared to the well-watered non-transgenic lines. However; the transgenic lines had unchanged cysteine protease activity after drought (Figure 4.5B) and cysteine protease activity measured might not represent all cysteine protease activity as the substrate used only detects cathepsin L-like activity.

In-gel protease activity determination was carried out using protein extract from non-transgenic plants (Figure 4.5C). To differentiate between cysteine and serine proteases, samples were incubated with specific inhibitors E-64 and TLCK, respectively. Both agents showed an inhibition of protease activity which implies presence of both cysteine and serine protease activities.

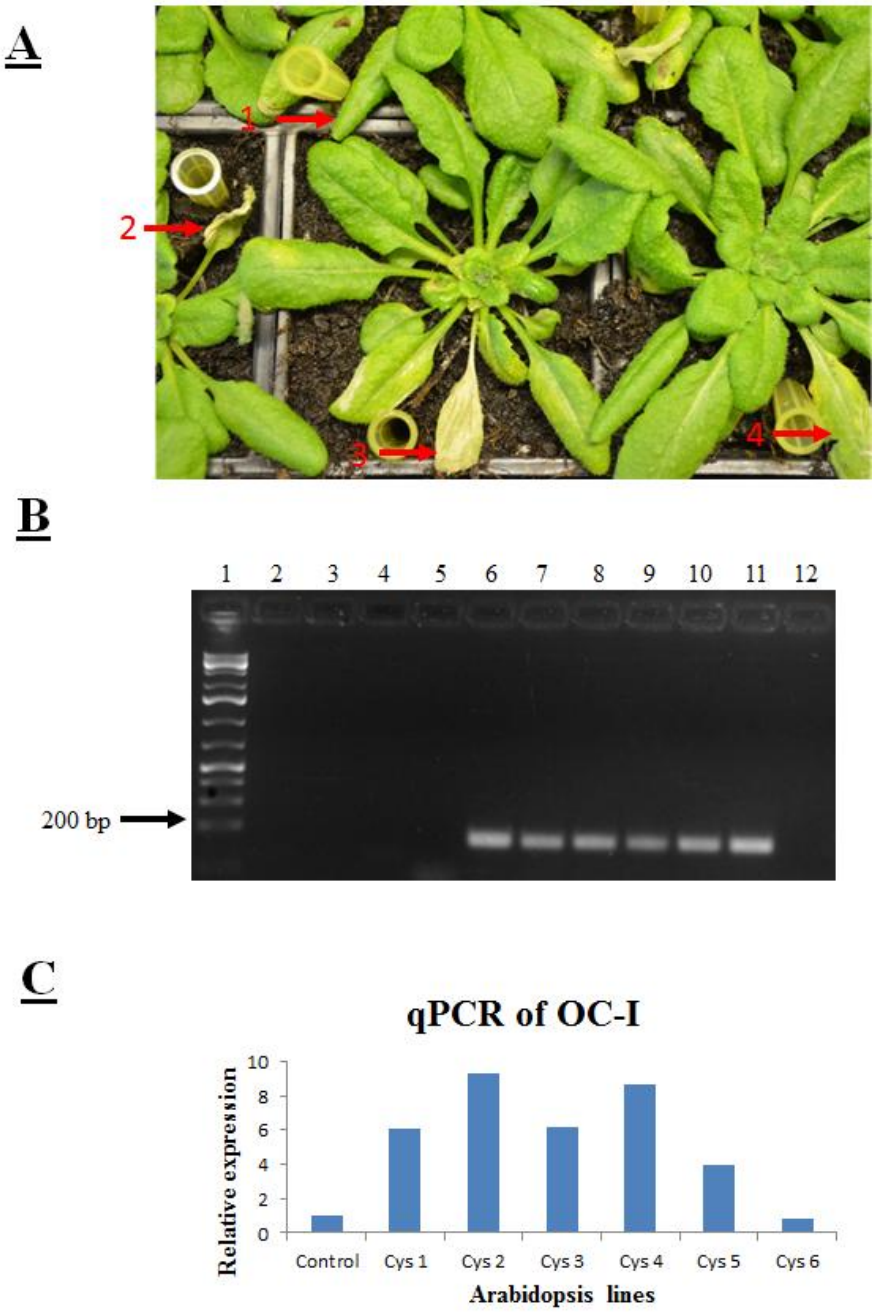


Figure 4.1 Selection of transgenic Arabidopsis plants. (A) Representative picture of BASTA selection of Arabidopsis plants. (1) Resistance; (2 and 4) partial resistance and (3) susceptible. (B) PCR amplification of the OC-I from the transformed Arabidopsis plants. Lane 1: 1 Kb marker; Lane 2: NC (negative control); Lanes 3-12 transgenic lines; (C) relative expression of the OC-I transgene in the non-transgenic (control) and transgenic lines.

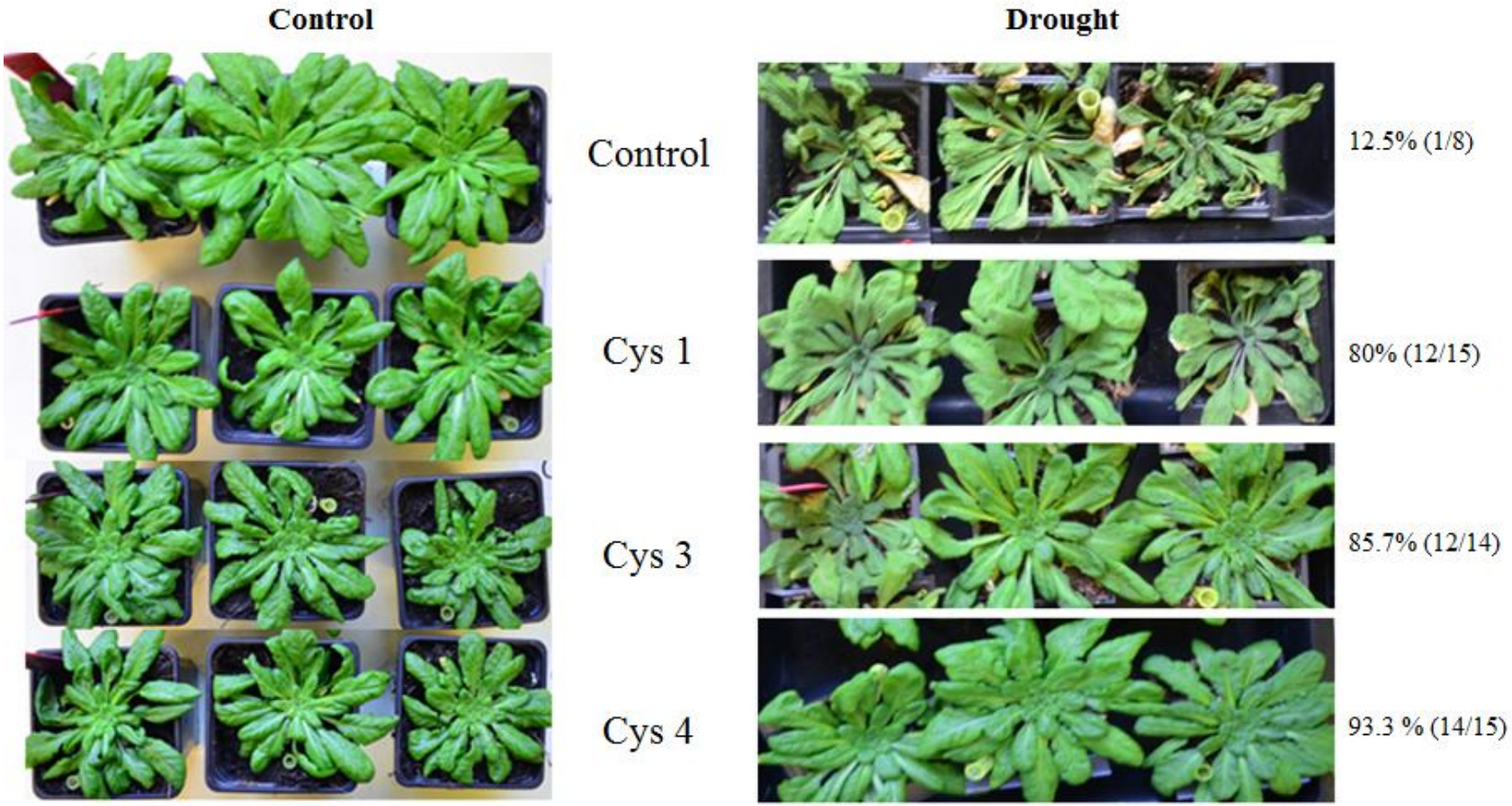


Figure 4.2 Drought tolerance of 4 weeks old Arabidopsis plants from which water was withheld for 15 days. Survival percentage for control, Cys 1, Cys 3 and Cys 4 were 12.5%, 80%, 85.7% and 93.3%, respectively.

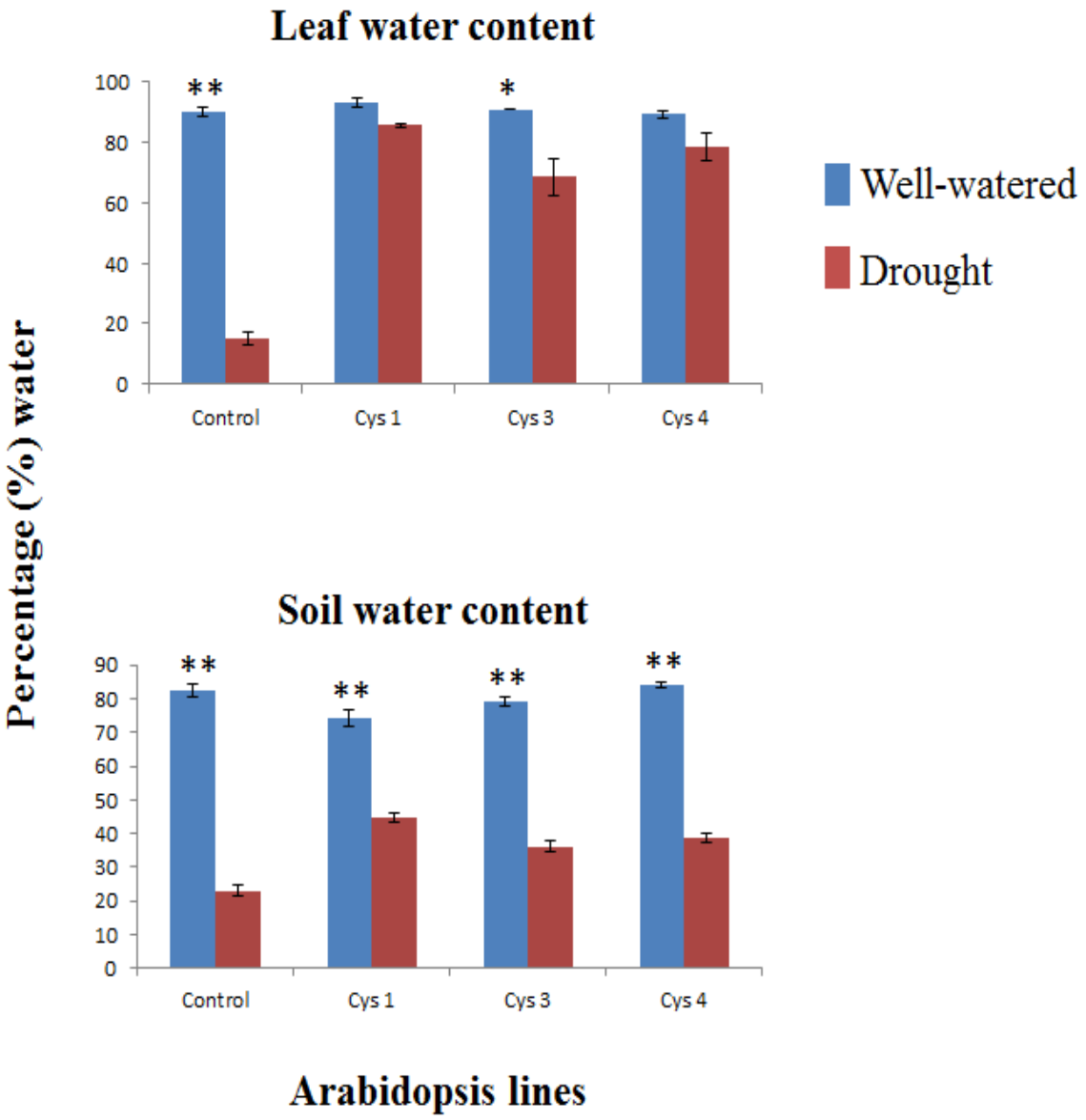


Figure 4.3 Water-loss in 4 weeks old detached leaves (A) and soil (B). Data represent the mean ± SE of 3 different plants of each line (*p < 0.05 significant and **p < 0.01 highly significant).

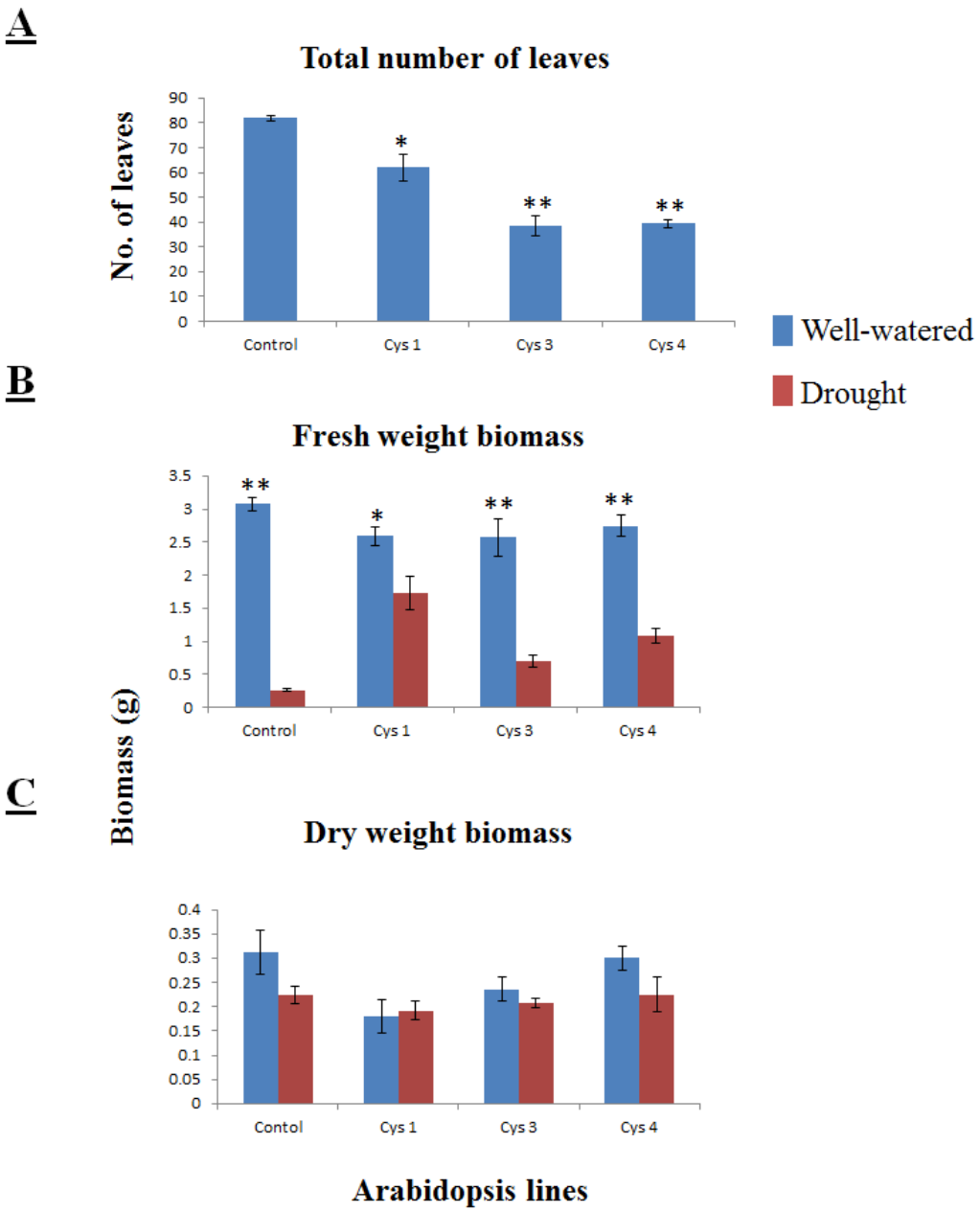


Figure 4.4 Characterization of control and transgenic lines. (A) Total number of leaves of control and transgenic lines. Drought effects on the leaf biomass of fresh (B) and dry (C) detached leaves. Data represent the mean ± SE of 3 different plants of each line (*$p < 0.05$ significant and **$p < 0.01$ highly significant).

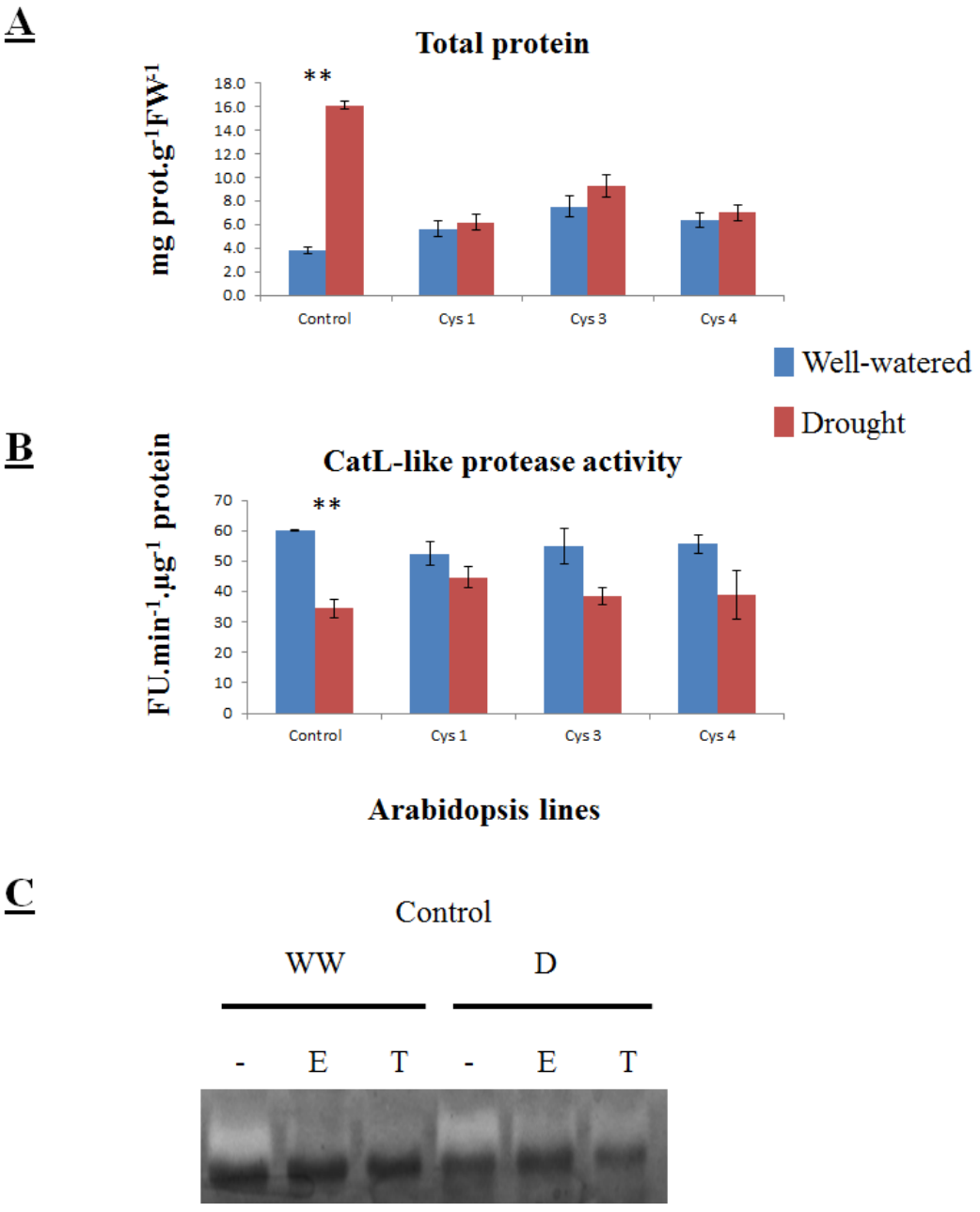


Figure 4.5 Proteolytic activity of leaves of control and transgenic Arabidopsis plants. Total protein content (A) and cysteine protease activity (B) was determined in plants grown under optimum and drought conditions. (C) In-gel protease activity on protein extracts from control samples. Data represent the mean ± SE of 3 different plants of each line (*$p < 0.05$ significant and **$p < 0.01$ highly significant). WW = well-watered, D = drought, - = protein alone, E = E-64 and T = TLCK.

4.4 Discussion

Transgenic Arabidopsis plants were selected by leaf painting with glufosinate, an ammonium salt of phosphinothricin, which is the main active ingredient of the herbicide BASTA (Paz et al., 2006). This method has been previously applied to distinguish between transgenic and non-transgenic plants with different tolerance to the herbicide (Chen et al., 2008, Wang et al., 2009). In the present study, three phenotypes i.e. susceptible, partial resistance and complete resistance were observed in relation to BASTA action. Since susceptibility to BASTA spraying does not necessarily mean non-integration of the bar gene and/or transgene into the Arabidopsis genome but might be due to low gene expression in the leaves resulting in high mortality of the progenies (Ikea et al., 2004), both PCR and qRTPCR were carried out to determine the presence and expression of the transgene in the Arabidopsis plants, respectively. Results obtained in this study clearly show differential expression of the transgene amongst plants of different tested transgenic lines.

In this Arabidopsis study, drought had a significantly dramatic effect on the phenotype of all plants with non-transgenic plants having higher biomass and more leaves more affected than transgenic plants. The leaf water content also significantly decreased under drought in non-transgenic plants when compared to transgenic plants. This confirms results obtained by other research groups where over-expression of a cystatin resulted in enhanced tolerance to abiotic stress including drought (Zhang et al., 2008). Plants with lower biomass and less leaf number, caused by blocking cysteine protease activity possibly by OC-I, are less sensitive to drought treatment very likely having a delayed onset of the drought effect (Lawlor, 2013).

OC-I expressing plants have also been reported to have higher total protein compared to control (Van der Vyver et al., 2003, Prins et al., 2008). Similar results were observed in this Arabidopsis study under well-watered conditions with transgenic Arabidopsis plants having more total protein compared to non-transgenic plants. This is possibly due to less protein degradation due to protease action. There are also reports suggesting that drought increases the total protein content through synthesis of drought proteins, such as dehydrins (Demirevska et al., 2008, Mohammadkhani and Heidari, 2008). An interesting aspect to investigate would be the effect of OC-I over-expression on these particular proteins and if limiting of cysteine protease activity results in more dehydrins.

There is considerable evidence that plant senescence, but also exposure of plants to abiotic stress including drought, induces expression of proteases such as cysteine proteases (Khanna-Chopra et al., 1999, van der Hoorn et al., 2004, Hwang et al., 2010a, Simova-Stoilova et al., 2010). In general, proteases are involved in the regulation of protein breakdown (Callis, 1995) and Chen et al. (2013) previously reported that ectopic expression of a protease from sweet potato enhances drought sensitivity of transgenic Arabidopsis lines. Protease inhibitors, such as a cystatin, which was investigated in this study, are also induced by environmental stress (Pernas et al., 2000; Hwang et al., 2010) and can prevent protease activity. Indeed, Demirevska et al. (2010) and van der Vyver et al. (2003), demonstrated that tobacco plants over-expressing the rice cystatin OC-I were better protected against several abiotic stresses. In this PhD study, OC-I over-expression in Arabidopsis plants also resulted in more drought-tolerant plants where OC-I very likely interacted with cysteine proteases causing a change in the plant phenotype which allowed better survival under drought.

The next chapter had the objective to determine if the findings with the model plant Arabidopsis exposed to well-watered and drought conditions were also observed in the economically important soybean crop using the transgenic lines already obtained.

Chapter 5

Phenotypic characterization of transgenic soybean under well-watered and drought conditions

5.1 Introduction

Plants are subjected to a variety of environmental (biotic and abiotic) stresses that trigger a wide range of plant responses. This ranges from altered gene expression and cellular mechanisms to changes in growth rates and crop yields (Reddy et al., 2004). Amongst the environmental stresses, drought is one of the major causes of crop losses constituting about 50% reduction of average yields (Wang et al., 2003). Drought stress decreased CO_2 assimilation rates due to reduced stomatal conductance (Cornic, 2000, Flexas and Medrano, 2002). The stomatal closure is subsequently followed by decreases in internal CO_2 (C_i) concentration inhibiting photosynthesis under mild to moderate drought (Reddy et al., 2004, Flexas and Medrano, 2002). Drought stress also negatively affects photosynthetic enzymes such as rubisco (Zlatev and Lindon, 2012, Flexas and Medrano, 2002).

Ideal traits for agricultural crops are both a high potential growth rate and efficient use of water. However; drought severely limits these traits (Zlatev and Lindon, 2012). Wu et al. (2008) showed that drought stress dramatically decreased seedling height, basal diameter, leaf number, leaf area, root length and biomass production in the perennial shrub, *Sophora Davidii*. Furthermore, drought decreased relative water content (RWC) and water-use efficiency (WUE) (Wu et al., 2008). Similar results were observed when two legume plants were drought treated i.e. reduced whole plant growth (biomass, leaf number, stem elongation rate and shoot growth rate) and reduced RWC (Gorai et al., 2010).

The aim of this part of the PhD study was to characterize OC-I expressing plants of three selected transgenic lines (soybean lines 55, 56, 57) under well-watered and drought conditions by (i) measuring morphological traits such overall germination percentage,

number of leaves, stem length and biomass, (ii) photosynthesis and photosynthetic parameters such as stomatal conductance, transpiration rate and C_i rates amongst others were also determined.

5.2 Materials and Methods

5.2.1 Plant growth and maintenance

Seeds of wild-type control and the T_2 seeds (selfed twice) of three independent OC-I expressing transgenic soybean lines (55, 56 and 57) were germinated on soil and the percentage germination determined. Percentage germination was determined as the number of seeds (T_2 generation) that germinated in soil divided by the number of seeds that were initially planted multiplied by 100. Seedlings were watered twice per week with an equal amount of demineralised water. Seedlings were grown in large pots (17.5 cm height and 20 cm diameter (top) and 13.1 cm diameter (bottom)) in potting soil for 6 weeks in a temperature controlled growth chamber with 15/9hrs light/dark cycle (light intensity of 800-1000 µmol $m^{-2}s^{-1}$) and a 26ºC/20ºC day/night temperature cycle and 60% relative humidity.

5.2.2 Morphological parameters

For each experiment morphological parameters, such as overall: leaf number, stem height, shoot length, fresh and dry weights of upper plant organs, were compared in 6 weeks old soybean plants from each line. Leaf number was measured by counting the number of leaves from base to tip throughout plant development till the 6^{th} week. Stem length was determined

by measuring the distance from ground surface to the apical meristem. Shoot length was determined by measuring the shoot from the base to the tip of the shoot.

5.2.3 Drought treatment

For drought exposure, 6 weeks old plants that were grown as described above (section 5.2.1) were deprived of water for 5 days until the soil water content had decreased to 40%. For measurement of leaf and soil water content, the fresh weight of leaf material was measured using a Model B-502-S Metter Toledo balance (Switzerland). The samples were then placed into a drying oven at a temperature of 80ºC for 48 hrs. The same was done for the soil samples corresponding to each plant. Leaf and soil water contents were calculated as outlined in section 4.2.5.

5.2.4 Photosynthesis, instantaneous water-use efficiency and chlorophyll a fluorescence

Photosynthetic CO_2 assimilation, stomatal conductance, intracellular CO_2 and transpiration rates were determined daily. These parameters were determined using on the central leaflet in each case of attached first trifoliate (bottom), fourth trifoliate (middle) and seventh trifoliate (top) leaves between 4 and 6 hrs after the start of the photoperiod using the LI-6400XT portable photosynthesis system (LI-COR, USA). Instantaneous water-use efficiency (IWUE) values were calculated as the ratio between CO_2 assimilation rates and stomatal conductance values as described previously (Soares-Cordeiro et al., 2009). The quantum yield of photosystem II (PSII) reaction centres, calculated as the Fv/Fm ratio, was determined using a Fluorpen fp100 (Photon Systems Instruments, Czech Republic) in light adapted leaves.

5.2.5 Statistical analysis

Statistical analysis between wild-type control and transgenic plants was carried out as outlined in section 3.2.6. For percentage germination calculations, mean values were taken from three seed batches of all the lines (wild-type, lines 55, 56 and 57).

5.3 Results

5.3.1 Morphological characterization

The shoot growth and development of the OCI-expressing transgenic plants (lines 55, 56 and 57) and wild-type plants were compared (Figure 5.1). The shoot phenotypes were compared in 6 weeks old plants with the wild-type visually taller than the plants of transgenic lines (Figure 5.1A). The stem length was also determined with all the transgenic lines significantly shorter than the wild-type ($p < 0.01$) (Figure 5.1B). The increase in the number of leaflets was also determined over time (Figure 5.1C). There was no significant difference ($p > 0.05$) in the number of leaflets from weeks 2 to 5. However; the wild-type plants had significantly ($p < 0.01$) more leaflets than the plants of transgenic lines after 6 weeks.

The effect of drought on the leaf phenotype of OCI-expressing plants (lines 55, 56 and 57) and wild-type controls was determined after water was withheld from the plants for 5 days (Figure 5.2). At 5 days drought treatment, the soil water content was 40% for wild-type plants and 48% for transgenic plants (Belén Márquez García, unpublished result). At 40% soil water content, wild-type plants started visually, in comparison to transgenic plants, to wilt (Figure 5.2A). The shoot length of well-watered and drought treated plants were also compared

between the wild-type and transgenic lines (Figure 5.2B). Under well-watered conditions, the shoots of transgenic lines (56 and 57) were smaller when compared to shoots of the wild-type and line 55. Drought significantly reduced the shoot length of wild-type plants ($p < 0.05$) and also in plants of transgenic line 55 ($p < 0.01$). However; no significant ($p > 0.05$) change were found for lines 56 and 57 and these two lines were therefore less affected by drought.

5.3.2 Physiological characterization

Photosynthesis (CO_2 assimilation) measurements were taken daily in 3 leaf ranks (top, middle and bottom) at soil water contents of 80% (well-watered) and 40% (drought conditions) (Figure 5.3). Photosynthetic CO_2 assimilation was similar in young leaves (top and middle) of all the lines under well-watered conditions but with line 57 having lower CO_2 assimilation values. However; the oldest leaves (bottom) of the transgenic lines had significantly lower ($p < 0.05$ and $p < 0.01$) rates of photosynthesis compared to wild-type plants under well-watered conditions. The transgenic lines (56 and 57) maintained significantly ($p < 0.05$) higher photosynthesis rates under drought conditions compared to wild-type plants and plants of transgenic line 55 under identical conditions.

Stomatal conductance values were also determined in the 3 leaf ranks in a similar way as 5.2.4 (Figure 5.4). Under well-watered conditions, the wild-type control together with the transgenic lines (55 and 56) had similar values in the young leaves with the exception of line 57 having significantly lower ($p < 0.01$) stomatal conductance in the young leaves (top) and the old leaves (bottom). The opposite was observed under drought conditions, there was no difference in the stomatal conductance values in young and old leaves of the wild-type and

transgenic lines 55 and 56 but with line 57 having significantly higher ($p < 0.05$) stomatal conductance values in young and old leaves.

Intracellular CO_2 (C_i) concentration was similar in all leaves (top, middle and bottom) of all the lines under well-watered conditions (Figure 5.5). There was a highly significant difference ($p < 0.001$) between the CO_2 levels of the wild-type plants and the plants of transgenic lines 56 and 57 in the young leaves under drought conditions. The C_i concentration of these lines (56 and 57) was not affected by drought as they maintained a C_i concentration similar to that of the well-watered plants (~300 µmol CO_2 mol^{-1}). In contrast, plants of the wild-type and line 55 had about 600 µmol CO_2 mol^{-1} and 450 µmol CO_2 mol^{-1}, respectively.

The transpiration rates for plants of transgenic lines 56 and 57 were lower than for the wild-type and line 55 in the young leaves (top) and old leaves (bottom) under well-watered conditions (Figure 5.6). The transgenic lines maintained higher transpiration rate in the young and old leaves under drought conditions while the wild-type control plants had fallen to very low levels under the same conditions.

Instantaneous water-use efficiency (IWUE) of leaves of wild-type and transgenic lines (55, 56 and 57) under well-watered and drought conditions were also measured (Table 5.1). IWUE values were similar in the young and old leaves of all the wild-type and transgenic lines (55, 56 and 57). However; drought dramatically decreased IWUE for wild-type plants in young and old leaves compared to transgenic lines (55, 56 and 57) with line 57 maintaining the highest IWUE in the young and old leaves.

The Fv/Fm ratios were used to determine any photo-inhibitory effects on the electron transport system (Figure 5.7). The Fv/Fm ratios from the young (top and middle) and old (bottom) leaves were compared in wild-type and transgenic lines. There was no significant ($p < 0.05$) differences between the wild-type and transgenic lines under well-watered conditions (Figure 5.7). However; the Fv/Fm ratios were significantly lower ($p < 0.05$) in the top leaves of transgenic lines 57 after the imposition of drought.

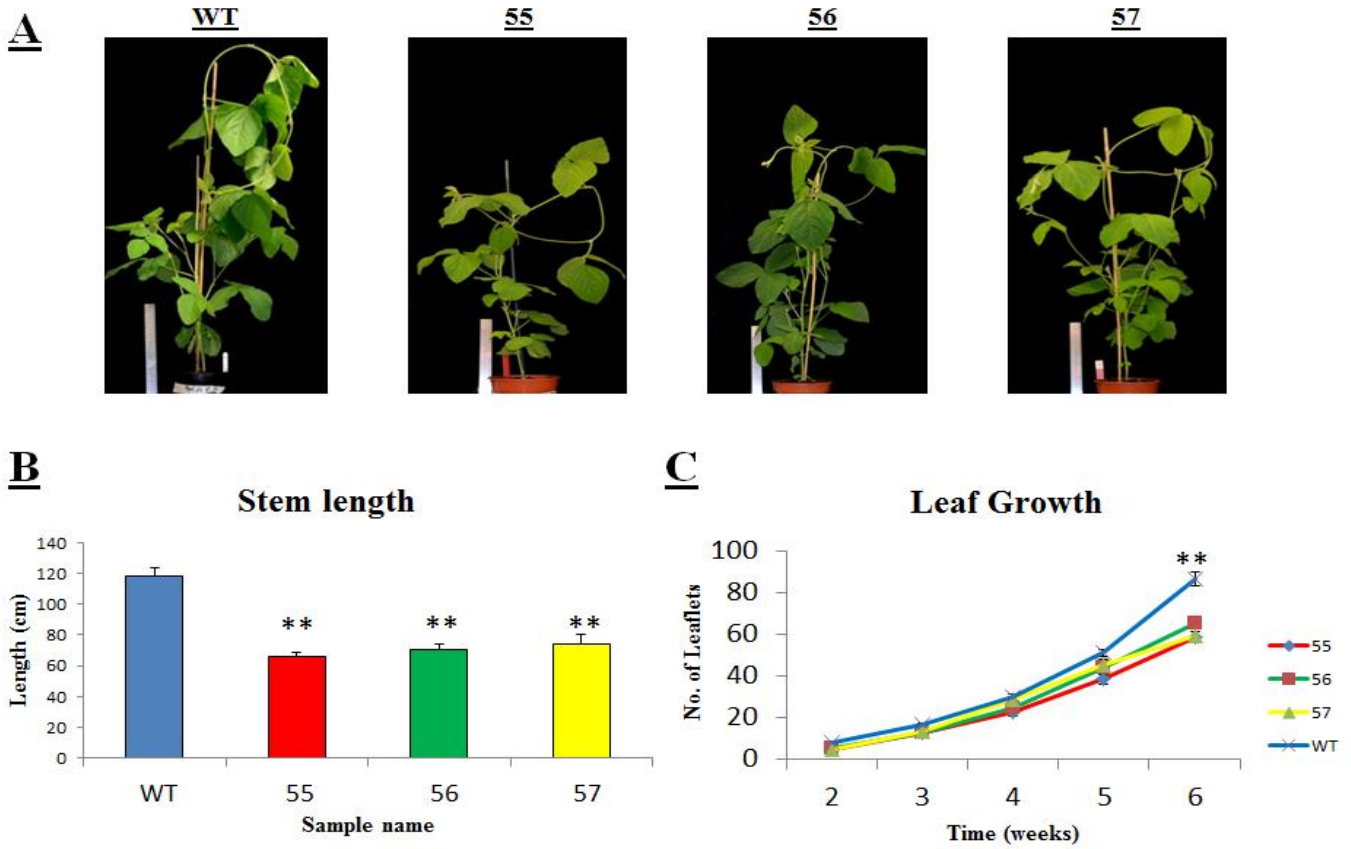


Figure 5.1 A comparison of shoot growth and development in OCI-expressing transgenic lines (55, 56 and 57) and wild-type (WT). Shoot phenotypes (A)**,** stem length (B) of WT and OC-I lines are shown at 6 weeks, together with the increase in the number of leaflets with time (C). Data represent the mean ± SE of 3 individual plants (*$p < 0.05$ and **$p < 0.01$).

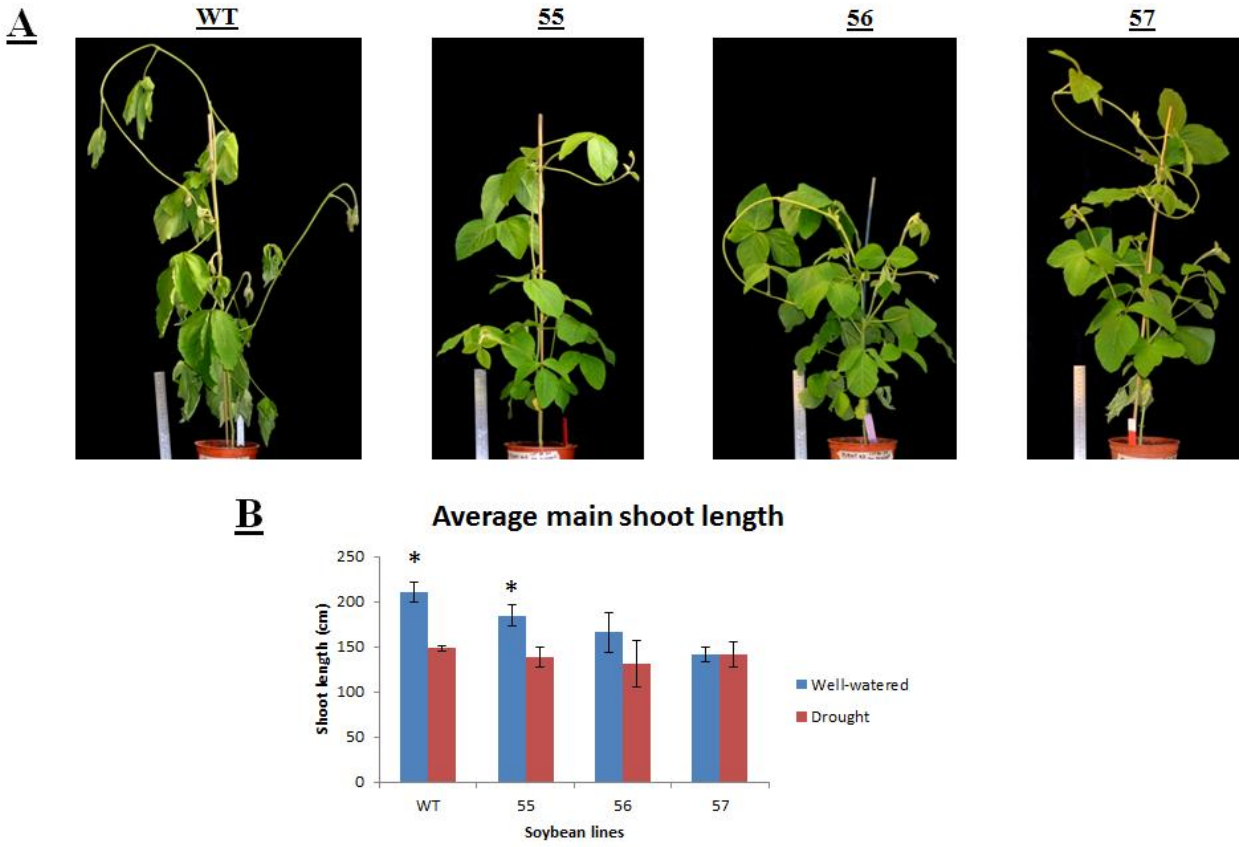


Figure 5.2 Comparison of the effects of drought on the phenotype of the leaves of OCI-expressing transgenic lines (55, 56 and 57) and wild-type (WT). Shoot phenotypes (A) and length (B) 5 days after plants were deprived of water. Data represent the mean ± SE of 3 individual leaves (*$p < 0.05$ and **$p < 0.01$).

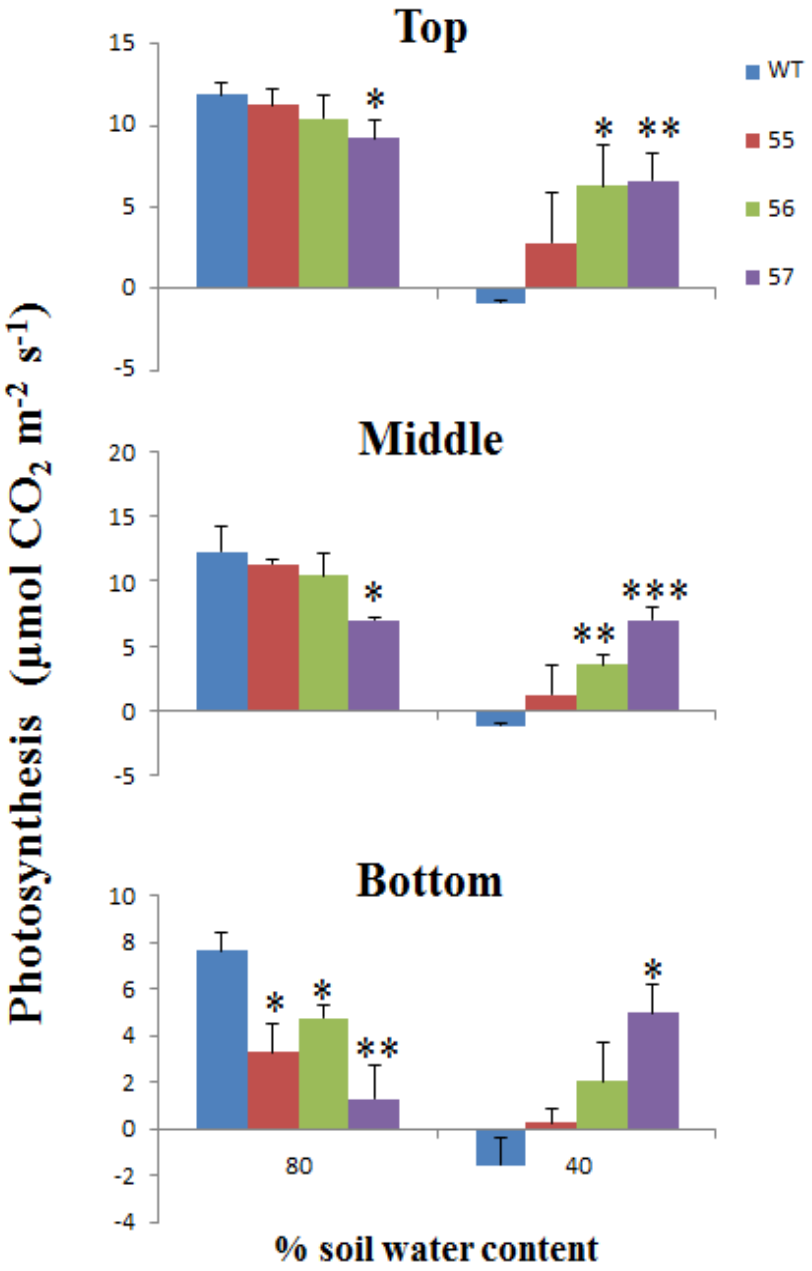


Figure 5.3 The effects of photosynthetic gas exchange on the leaves of OCI-expressing transgenic lines (55, 56 and 57) and wild-type (WT) controls after plants were deprived of water. Measurements were taken in 3 leaf ranks (top, middle and bottom) at soil water contents of 80% (well-watered) and 40% (drought conditions). Data represent the mean ± SE of 3 individual leaves (*$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$).

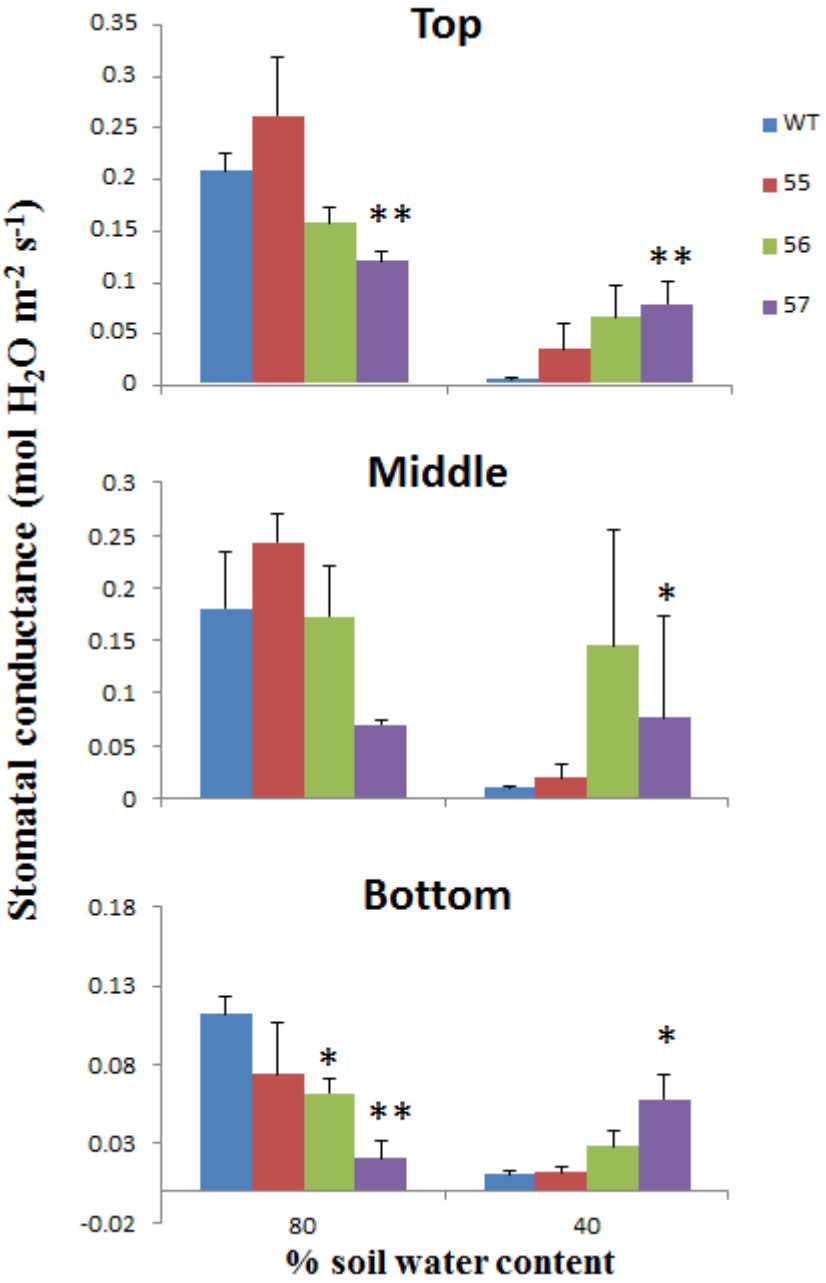


Figure 5.4 Stomatal conductance of the leaves of OCI-expressing transgenic lines (55, 56 and 57) and wild-type (WT) controls after plants were deprived of water. Measurements were taken in 3 leaf ranks (top, middle and bottom) at soil water contents of 80% (well-watered) and 40% (drought conditions). Data represent the mean ± SE of 3 individual leaves (*$p < 0.05$ and **$p < 0.01$).

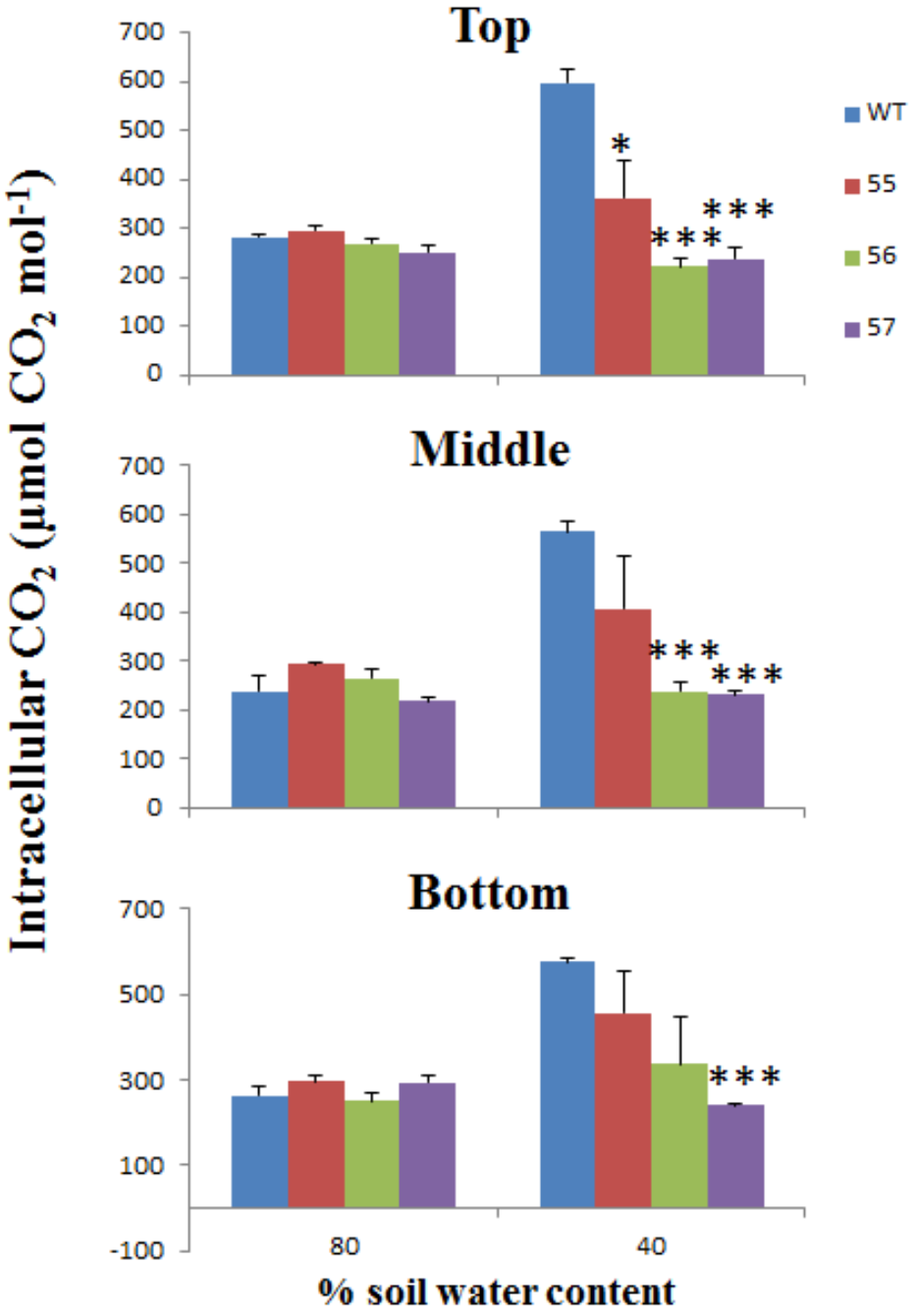


Figure 5.5 Effects of drought on intracellular CO_2 of the leaves of OCI-expressing transgenic lines (55, 56 and 57) and of wild-type (WT) plants. Measurements were taken in 3 leaf ranks (top, middle and bottom) at soil water contents of 80% (well-watered) and 40% (drought conditions). Data represent the mean ± SE of 3 individual leaves (*$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$).

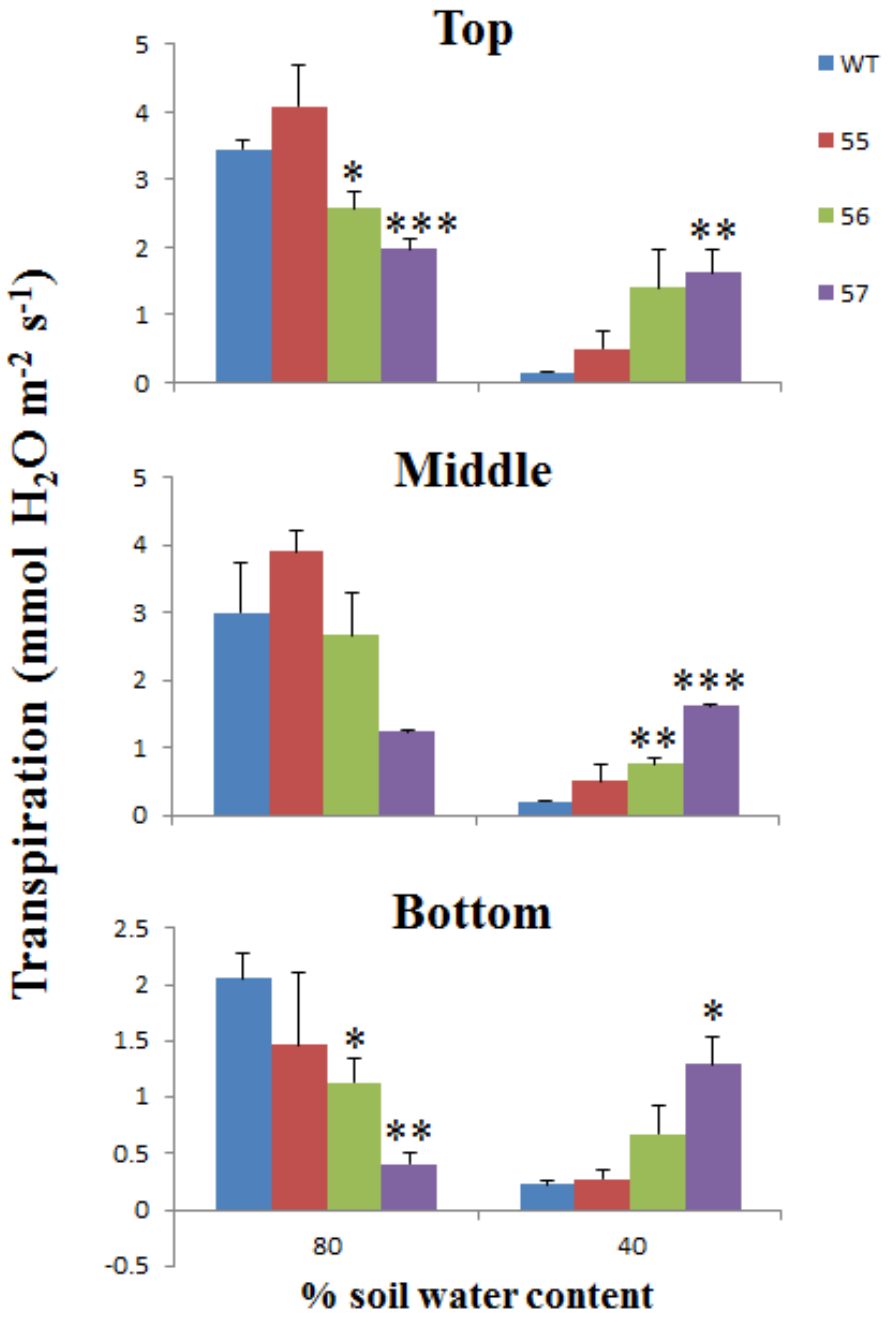


Figure 5.6 Transpiration rate of the leaves of OCI-expressing transgenic lines (55, 56 and 57) and of wild-type (WT) plants after plants were deprived of water. Measurements were taken in 3 leaf ranks (top, middle and bottom) at soil water contents of 80% (well-watered) and 40% (drought conditions) Data represent the mean ± SE of 3 individual leaves (*$p < 0.05$, **$p < 0.01$ and ***$p < 0.001$).

Table 5.1 Instantaneous water-use efficiency (IWUE) values of leaves of wild-type and transgenic lines (55, 56 and 57) under well-watered and drought conditions. Data is the mean ± SE of three independent plants (*p < 0.05, **p < 0.01 and ***p < 0.001).

IWUE (µmol CO_2 per mol H_2O)								
	Plants							
Tissue	Well-watered				Drought			
	WT	55	56	57	WT	55	56	57
TF 1	58.1 ± 16.5	61.2 ± 18.5	80.7 ± 14.1	57.1 ± 10.4	-140.5 ± 19.5	-4.33 ± 46.0	27.8 ± 73.5	85.6 ± 1.78 ***
TF 4	80.6 ± 21.1	47.3 ± 3.69	67.5 ± 12.6	101 ± 4.49	-114 ± 12.9	-175 ± 192.5	66.1 ± 30.6 **	91.2 ± 4.7 ***
TF 7	57.2 ± 2.08	45.7 ± 6.87	45.2 ± 5.18	77.7 ± 9.24	-128.5 ± 18.3	12.6 ± 46.5 *	98.7 ± 11.5 **	87.5 ± 14.8 ***

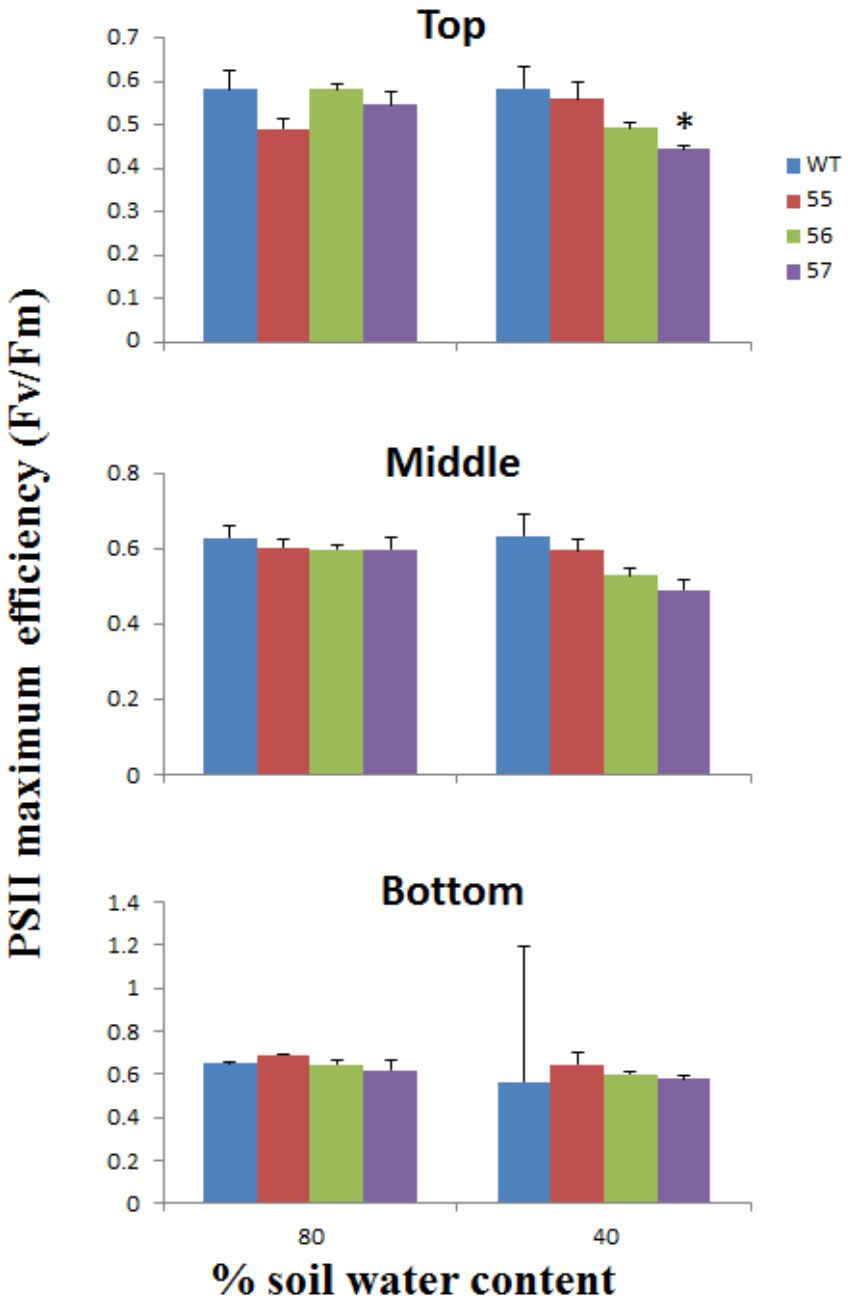


Figure 5.7 PSII maximum efficiency of light adapted leaves (Fv/Fm) of transgenic lines (55, 56 and 57) and wild-type (WT) plants. Water was withheld for 5 days. Data represent the mean ± SE of three individual leaves (*$p < 0.05$).

5.4 Discussion

Cysteine proteases play an important role during the maturation and germination of seeds. Endogenous cysteine proteases have to be regulated by phytocystatins, which are plant cysteine proteases inhibitors (Abe and Arai, 1991). Oryzacystatins, including OC-I which has been used in this study, have been considered to be inhibitors of the α, β and γ cysteine proteases that are produced during rice seed germination (Watanabe et al., 1991). In this study, there was also delayed seed germination of transgenic lines (55, 56 and 57) compared to the wild-type control seeds (data not shown). This finding confirms a previous observation made in my research group with tobacco seeds (C. Van der Vyver, personal communication). Papain-like cysteine proteases have been found in germinating seeds (Ling et al., 2003) and results of this study suggest that over-expression of the oryzacystatin I (OC-I) inhibited papain-like cysteine proteases that are required for seed germination and shoot growth. As a consequence, transgenic soybean seeds have been previously found to contain more protein and less cysteine protease activity than wild-type seeds (Quain, unpublished data). Phenotypically, plants of transgenic lines had further reduced stem length and less leaves than wild-type plants after 6 weeks of growth. A similar finding has been previously reported for OCI-expressing tobacco plants with slower plant development and growth and less stem elongation compared to non-transformed or wild-type plants after 7 weeks of growth (Van der Vyver et al., 2003).

In this study, wild-type plants further showed changes in the phenotype when compared to plants of the transgenic lines after exposure to drought (withholding water for 5 days). One of the well-known physiological effects of drought is a reduction of shoot growth (Schuppler et al., 1998). Slower cell division as well as inhibition of growth is observed under water deficit

with reduced cyclin-dependent kinase activity in wheat seedlings (Schuppler et al., 1998). In this PhD study, drought treatment was found to significantly decrease the main shoot length of both wild-type plants and in particular plants of line 55. This possibly indicates a protective effect by OC-I in plants of these lines by preventing cysteine protease activity thus resulting in smaller plants that lost less water under drought conditions.

Photosynthetic CO_2 assimilation was similar in young leaves of wild-type plants and of transgenic lines. However; under abiotic stress i.e. chilling, photosynthetic CO_2 assimilation was better protected in tobacco plants over-expressing OC-I (Van der Vyver et al., 2003) and better photosynthesis under abiotic stress conditions has also been reported (Prins et al., 2008, Demirevska et al., 2010). Findings of this soybean study indicated that the CO_2 assimilation of the control plants was severely affected by drought in young and old leaves compared to the transgenic lines (55, 56 and 57) with CO_2 assimilation higher in the transgenic lines 56 and 57. An interesting observation in this PhD study was that plants of transgenic line 57 had lower CO_2 assimilation when compared to all other lines (wild-type, line 55 and 56) under well-watered conditions. Plants of this line also had the highest hybridization signal with the OC-I antibody (Figure 3.5). This result confirms previous tobacco results where transgenic tobacco plants with highest OC-I over-expression had also lower CO_2 assimilation when compared to non-transgenic plants (Van der Vyver et al., 2003). Under well-watered conditions, stomatal conductance in young leaves was similar in leaves of the wild-type and of lines 55 and 56. Under drought conditions though, only plants of transgenic line 57 maintained significantly higher stomatal conductance in young and old leaves compared to plants from the other lines tested. This result indicates that plants with higher OC-I expression, although less performing under well-watered conditions, are possibly better protected against drought exposure. However; if a direct correlation exists between the

amounts of OC-I expressed and change of phenotype this has to be still investigated in greater detail characterizing a larger number of transgenic lines.

Intracellular CO_2 was similar in all tested lines under well-watered conditions. Under drought conditions, the C_i was significantly higher in wild-type plants than in young and old leaves of transgenic lines indicating that plants are respiring and less photosynthesizing. Drought results in stomatal closure which causes an increase in intercellular CO_2 (Brodribb, 1996). Lawlor (2002) suggested that a decrease in photosynthesis is due to limitations in RuBP regeneration, not CO_2, with C_i values up to 2500 pp. In this PhD study, wild-type plants had significantly ($p < 0.05$) lower stomatal conductance due to stomatal closure resulting in increased C_i under drought conditions. Transgenic lines, however, maintained C_i at 300 ppm because they had higher stomatal conductance and CO_2 assimilation. Reduced transpiration in plants of transgenic lines 56 and 57 under well-watered conditions in both young and old leaves was observed, but line 57 maintained significantly higher transpiration in young and old leaves under drought conditions. Moreover, since IWUE is dependent on photosynthesis and stomatal conductance, plants of transgenic line 57 had the highest water-use efficiency in the young and old leaves. It has been recently reported that a more drought-tolerant soybean cultivar, Prima 2000, also had the ability, similar to the transgenic plants tested in this study, to maintain high leaf water-use efficiencies (Fenta et al., 2012) suggesting that the OC-I transgenic lines might also be drought-tolerant. The Fv/Fm ratios of the wild-type and transgenic lines remained relatively unchanged before and after drought stress except for the top leaf of transgenic line 57. This leaf had a significantly lower Fv/Fm ratio after drought exposure possibly due to photo-inhibition. OC-I expression possibly affected in this leaf proteases involved in PSII protein turnover but further investigations are required to better understand this photo-inhibitory effect. Overall, this part of the PhD study has shown that

OC-I expression affected the phenotype of soybean particularly germination, stem length and leaf number possibly due to a limitation in cysteine protease (papain-like cysteine proteases) activity. Targeting these proteases with an inhibitor has allowed investigation into the cellular function of these proteases and their involvement in plant growth and development in greater detail. In addition, OC-I expressing transgenic soybean lines were more drought-tolerant with better photosynthesis possibly due to less protein degradation. Previous experiments with cessation of watering in transgenic plants have demonstrated that drought tolerance may be due to late stress development compared to non-transgenic plants. This is very likely caused by slower total water loss due to smaller total leaf area (LA) and/or decreased stomatal conductance (*gs*), associated with thicker laminae (denser mesophyll) and smaller cells (Lawlor, 2013).

Transgenic line 57 had the highest OC-I hybridization signal with much lower in-gel protease activity when compared to the other tested lines and lower cathepsin L-like activity was found in cotyledons and senescing leaves. This transgenic line also had better photosynthetic CO_2 assimilation under drought with better IWUE. This suggests that transgenic line 57 was more drought-tolerant despite that the Fv/Fm of the young leaves (top) of this line being significantly affected by drought treatment.

The next chapter will summarize the conclusions from the overall study and will provide recommendations for future studies.

Chapter 6

General conclusions and recommendations

In this PhD study, transgenic soybean and Arabidopsis plants have been produced over-expressing various protease inhibitors with the aim to determine if such inhibitor over-expression causes phenotypic changes in plants, in particular, to improve their drought tolerance. Since inhibitor over-expression targets a specific class of proteases lowering their activity, transgenic plants created further offer the opportunity to study the specific function of this class of proteases in developmental processes and in responses to stress.

The overall work carried out has clearly contributed with several actions to achieve this aim establishing the required tools to understand protease function in plant growth and development and stress tolerance. In general, any plant biotechnology project aimed to understand protein function using a transgenic approach firstly requires the designing and cloning of the gene constructs to be transferred, secondly the transformation of plants, and thirdly the selection and finally characterization of plants for gene expression and investigating any phenotypic change. Since none of the above mentioned requirements were established for soybean before the onset of this PhD study in our laboratory, the production of transgenic plants expressing various protease inhibitors to target specific classes of proteases was the main task of this PhD study. Moreover, there was also uncertainty as to whether and if so, how many positive transgenic soybean plants would be obtained, ectopic expression of protease inhibitors was also carried out in Arabidopsis plants to monitor any phenotypic change (s) caused by protease inhibitor expression.

A first outcome of this study was the successful construction of five plasmids namely: pTF101.1-Cys-I, pTF101.1-LegCys-I, pTF101.1-PRKCys-I, pTF101.1-BBI and pTF101.1-TI useful to target different plant tissues in Arabidopsis and soybean over-expressing a cysteine or serine protease inhibitor. This addressed the set objective of constructing various vectors

for plant transformation of protease inhibitors in various plant tissues including soybean nodules. In particular, inhibitors expressed in the chloroplast may function to protect chloroplast proteins from protease degradation resulting in better protection of photosynthesis under stress (Prins et al., 2008). However; the exact protection mechanism needs to be investigated using created transgenic plants as a potential tool for such investigations.

A second outcome of this PhD project was the establishment of various transgenic Arabidopsis and soybean plants expressing a protease inhibitor integrated into the plant genome. These transgenic plants will now build the basis for further extensive phenotypic characterization to explore the exact function of individual proteases in Arabidopsis and soybean plant development and performance under stress. Arabidopsis, as a model plant for transgene characterization, has the advantage that molecular and gene function data are available and plant crossings can be carried out. In this regard, future work is also planned to cross Arabidopsis plants to test if over-expression of inhibitor combinations will for example provide better protection against multiple environmental stresses when compared to plants over-expressing only one type of inhibitor. In contrast to Arabidopsis, studies on transformed soybean have a time constraint disadvantage which has also been experienced in this study. This is mainly due to the fact that soybean transformation is strain specific; there is low regeneration of putative transgenic shoots and low transformation efficiencies. Initial failure to set up the transformation system at the University of Pretoria severely affected rapid progress in this PhD project. Unlike soybean transformation, Arabidopsis transformation using the floral dip method is inexpensive, simple and quick. The success of soybean transformation still remains a skilled art (Paz et al., 2004, Mello-Farias and Chaves, 2008). Soybean cultivar Williams 82 has been shown to be transformable (Paz et al., 2006) and this cultivar has also been recently used for the establishment of a soybean genomic database. In

this PhD project, soybean transformation could ultimately be achieved with the help of Iowa State University for plants of the soybean cultivar Williams (Paz et al., 2006).

Transgenic soybean plants expressing various protease inhibitors created in this PhD project will also allow evaluation of any particular role of specific classes of protease inhibitors in the regulation of protease activity in soybean nodules. By silencing protease activity in nodules due to inhibitor action, the specific role of individual proteases during nodule development and under environmental stress conditions will also be a main target for detailed future investigations. In general, early senescence in soybean is characterized by protein breakdown in the nodules of which proteases play a key role leading to plant death (Solomon et al., 1999, Beyene et al., 2006). Therefore, studies on transgenic plants over-expressing protease inhibitors in the nodules will contribute to a better understanding of the mechanisms responsible for nodule senescence. In this PhD study, transgenic soybean plants have also been produced to over-express a serine protease inhibitor. Any action of this type of inhibitors on the plant phenotype, when over-expressed, has so far not been carried out. Future work on characterization of transgenic plants over-expressing serine protease inhibitors will therefore provide an insight into the specific role of serine proteases and their inhibitors in plant growth and development and under environmental stress conditions.

Several biotechnology strategies have been previously employed for the selection and characterization of plants with higher drought tolerance. This includes utilizing soybean genome sequence for molecular marker development and genetic engineering of plants for the improvement of drought tolerance (Manavalan et al., 2009). A third outcome of this PhD study was therefore successfully establishing transgenic soybean plants suitable for molecular and biochemical characterization and testing for better drought tolerance and exploring the

reason for better drought tolerance. This addressed the set objective of molecular and biochemical characterization of transgenic plants to study any phenotypical changes occurring due to protease inhibitor over-expression in non-stressed and stressed plants. Papain-like protease activities have been found in both cotyledons of growing seedlings (Ling et al., 2003) and in senescent leaves (Gepstein et al., 2003, Beyene et al., 2006). Cotyledons of transgenic sprouting plants and senescing leaves in this PhD study had lower protease activity compared to wild-type non-transgenic plants. This suggests that in both transgenic Arabidopsis and soybean plants expression of a protease inhibitor has limited and/or reduced papain-like protease activity. This result provides the opportunity to study the role of these papain-like proteases in young and senescent leaves when activity of these proteases is blocked by an inhibitor. In addition, these transgenic plants offer the opportunity to study the role of this class of proteases during senescence and if over-expression of an inhibitor delays premature senescence due to stress resulting in better recovery of transgenic plants from stress.

A fourth outcome of this PhD study was providing first evidence that transgenic soybean plants over-expressing a cysteine protease inhibitor like OC-I are more tolerant to drought when compared to non-transgenic plants. Since this PhD study was based on the hypothesis that over-expression of protease inhibitors in soybean results in delayed senescence caused by abiotic stress, support for this working hypothesis was found in this PhD study by showing that transgenic soybean plants over-expressing OC-I were better protected against drought with better water-use efficiencies and better photosynthesis when compared to wild-type non-transgenic plants. A critical issue to be further studied will be if particularly changes in growth and development (smaller biomass and lower leaf number) due to lower protease activity are the main reason for better drought tolerance delaying the onset of any severe

drought effect. With climate change (amongst other stresses), drought stress will affect crop yields and productivity (Manavalan et al., 2009). Better drought tolerance, as found in this PhD study, might therefore have an application potential for a crop like soybean. However; this has to be explored in greater detail in future studies monitoring plant performance and specifically yield productions after exposure to drought conditions. In a previous study with tobacco better drought tolerance, as found in this PhD study, was not found when OC-I was expressed (Van der Vyver et al., 2003). This contradictory result certainly requires a more detailed future analysis. Several studies have recently suggested that ectopic protein expression results in pleiotropic effects *in planta* (Gutiérrez-Campos et al., 2001, Munger et al., 2012) which might be different in tobacco and soybean and therefore partially explain the observed difference between the two types of plants expressing OC-I.

Soybean is a commercially interesting crop plant in South Africa used as a food and feed plant but with plant performance severely affected by abiotic stresses including drought. Research undertaken in this PhD study has provided a first insight into the possible potential of engineered plants expressing a protease inhibitor for improved plants performance in particular under stress. Furthermore, transgenic plants created in this PhD study will offer a potential future tool to unravel the function of specific classes of proteases in plant development and the function of protease inhibitors in stress protection.

REFERENCES

ABE, M. & ARAI, S. 1991. Some properties of a cysteine proteinase inhibitor from corn endosperm. *Agricultural and Biological Chemistry,* 55**,** 2417-2418.

ACHTEN, W. M. J., MAES, W. H., REUBENS, B., MATHIJS, E., SINGH, V. P., VERCHOT, L. V. & MUYS, B. 2010. Biomass production and allocation in *Jatropha curcas* L. seedlings under different levels of drought stress. *Biomass and Bioenergy,* 34**,** 667-676.

ADUGNA, A. & MESFIN, T. 2008. Detection and quantification of genetically engineered crops. *Journal of SAT Agricultural Research,* 6**,** 1-9.

ÁLVAREZ-ALFAGEME, F., MARTÍNEZ, M., PASCUAL-RUIZ, S., CASTAÑERA, P., DIAZ, I. & ORTEGO, F. 2007. Effects of potato plants expressing a barley cystatin on the predatory bug *Podisus maculiventris* via herbivorous prey feeding on the plant. *Transgenic Research,* 16**,** 1-13.

ANTÃO, C. M. & MALCATA, F. X. 2005. Plant serine proteases: biochemical, physiological and molecular features. *Plant Physiology and Biochemistry,* 43**,** 637-650.

BARRETT, A. J., RAWLINGS, N. D. & O'BRIEN, E. A. 2001. The *MEROPS* database as a protease information system. *Journal of Structural Biology,* 134**,** 95-102.

BEERS, E. P., JONES, A. M. & DICKERMAN, A. W. 2004. The S8 serine, C1A cysteine and A1 aspartic protease families in Arabidopsis. *Phytochemistry,* 65**,** 43-58.

BELENGHI, B., ACCONCIA, F., TROVATO, M., PERAZZOLLI, M., BOCEDI, A., POLTICELLI, F., ASCENZI, P. & DELLEDONNE, M. 2003. AtCYS1, a cystatin from *Arabidopsis thaliana,* suppresses hypersensitive cell death. *European Journal of Biochemistry,* 270**,** 2593-2604.

BENCHABANE, M., SCHLÜTER, U., VORSTER, J., GOULET, M.-C. & MICHAUD, D. 2010. Plant cystatins. *Biochimie,* 92**,** 1657-1666.

BEYENE, G., FOYER, C. H. & KUNERT, K. J. 2006. Two new cysteine proteinases with specific expression patterns in mature and senescent tobacco (*Nicotiana tabacum* L.) leaves. *Journal of Experimental Botany,* 57**,** 1431-1443.

BIRK, Y. 1985. The Bowman Birk inhibitor. Trypsin and chymotrypsin inhibitor from soybeans. *International Journal of Peptide and Protein Research,* 25**,** 113-131.

BRODRIBB, T. 1996. Dynamics of changing intercellular CO_2 concentration (C_i) during drought and determination of minimum functional Ci. *Plant Physiology,* 111**,** 179-185.

CARRILLO, L., MARTINEZ, M., ALVAREZ-ALFAGEME, F., CASTANERA, P., SMAGGHE, G., DIAZ, I. & ORTEGO, F. 2011. A barley cysteine-proteinase inhibitor reduces the performance of two aphid species in artificial diets and transgenic Arabidopsis plants. *Transgenic Research,* 20**,** 305-319.

CARVALHO, H. G., LOPES-CARDOSO, I. A., LIMA, L. M., MELO, P. M. & CULLIMORE, J. V. 2003. Nodule-specific modulation of glutamine synthetase in transgenic *Medicago truncatula* leads to inverse alterations in asparagine synthetase expression. *Plant Physiology,* 133**,** 243-252.

CHAVES, M. M., PEREIRA, J. S. & MAROCO, J. P. 2003. Understanding plant responses to drought-from genes to the whole plant. *Functional Plant Biology,* 30**,** 239-264.

CHEN, H.-J., TSAI, Y.-J., SHEN, C.-Y., TSAI, T.-N., HUANG, G.-J. & LIN, Y.-H. 2013. Ectopic expression of sweet potato cysteine protease SPCP3 alters phenotypic traits and enhances drought stress sensitivity in transgenic *Arabidopsis* plants. *Journal of Plant Growth Regulation,* 32**,** 108-121.

CHEN, R., XUE, G., CHEN, P., YAO, B., YANG, W., MA, Q., FAN, Y., ZHAO, Z., TARCZYNSKI, M. C. & SHI, J. 2008. Transgenic maize plants expressing a fungal phytase gene. *Transgenic Research,* 17**,** 633-643.

CLOUGH, S. J. & BENT, A. F. 1998. Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *The Plant Journal,* 16**,** 735-743.

CORBESIER, L., VINCENT, C., JANG, S., FORNARA, F., FAN, Q., SEARLE, I., GIAKOUNTIS, A., FARRONA, S., GISSOT, L., TURNBULL, C. & COUPLAND, G. 2007. FT protein movement contributes to long-distance signaling in floral induction of Arabidopsis. *Science,* 316**,** 1030-1033.

CORNIC, G. 2000. Drought stress inhibits photosynthesis by decreasing stomatal aperture-not by affecting ATP synthesis. *Plant Science,* 5**,** 187-188.

COWGILL, S. E., WRIGHT, C. & ATKINSON, H. J. 2002. Transgenic potatoes with enhanced levels of nematode resistance do not have altered susceptibility to non-target aphids. *Molecular Ecology,* 11**,** 821-827.

CREGG, B. & ZHANG, J. 2001. Physiology and morphology of *Pinus sylvestris* seedlings from diverse sources under cyclic drought stress. *Forest Ecology and Management,* 154**,** 131-139.

DEMIREVSKA, K., SIMOVA-STOILOVA, L., VASSILEVA, V., VASEVA, I., GRIGOROVA, B. & FELLER, U. 2008. Drought-induced leaf protein alterations in sensitive and tolerant wheat varieties. *General Applied Plant Physiolology,* 34**,** 79-102.

DEMIREVSKA, K., SIMOVA-STOILOVA, L., FEDINA, I., GEORGIEVA, K. & KUNERT, K. 2010. Response of oryzacystatin I transformed tobacco plants to drought, heat and light stress. *Journal of Agronomy and Crop Science,* 196**,** 90-99.

DEU, E., VERDOES, M. & BOGYO, M. 2012. New approaches for dissecting protease functions to improve probe development and drug discovery. *Nature Structural & Molecular Biology,* 19**,** 9-16.

DUBEY, V. K., PANDE, M., SINGH, B. K. & JAGANNADHAM, M. V. 2007. Papain-like proteases: applications of their inhibitors. *African Journal of Biotechnology,* 6**,** 1077-1086.

DUNSE, K. M., STEVENS, J. A., LAY, F. T., GASPAR, Y. M., HEATH, R. L. & ANDERSON, M. A. 2010. Co-expression of potato type I and II proteinase inhibitors gives cotton plants protection against insect damage in the field. *Proceedings of the National Academy of Sciences,* 107**,** 15011-15015.

EDWARDS, K., JOHNSTONE, C. & THOMPSON, C. 1991. A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. *Nucleic Acids Research,* 19**,** 1349.

FAN, J., YANG, Y.-W., GAO, X., DENG, W., FALARA, V., KANELLIS, A. K. & LI, Z.-G. 2009. Expression of a senescence-associated cysteine protease gene related to peel pitting of navel orange (*Citrus sinensis* L. Osbeck). *Plant cell, tissue and organ culture,* 98**,** 281-289.

FAN, S. & WU, G. 2005. Characteristics of plant proteinase inhibitors and their applications in combating phytophagous insects. *Botanical Bulletin of Academia Sinica,* 46**,** 273-292.

FAZELI, F., GHORBANLI, M. & NIKNAM, V. 2007. Effect of drought on biomass, protein content, lipid peroxidation and antioxidant enzymes in two sesame cultivars. *Biologia Plantarum,* 51**,** 98-103.

FENTA, B., DRISCOLL, S., KUNERT, K. J. & FOYER, C. H. 2012. Characterization of drought-tolerance traits in nodulated soya beans: the importance of maintaining photosynthesis and shoot biomass under drought-induced limitations on nitrogen metabolism. *Journal of Agronomy and Crop Science,* 198**,** 92-103.

FERNANDEZ, J., MELLO, M., GALGARO, L., TANAKA, A., SILVA-FILHO, M. & NESHICH, G. 2007. Proteinase inhibition using small Bowman-Birk type structures. *Genetics and Molecular Research,* 6**,** 846-858.

FLEXAS, J. & MEDRANO, H. 2002. Drought-inhibition of photosynthesis in C3 plants: stomatal and non-stomatal limitations revisited. *Annals of Botany,* 89**,** 183-189.

FRIEDMAN, M. & BRANDON, D. L. 2001. Nutritional and health benefits of soy proteins. *Journal of Agricultural and Food Chemistry,* 49**,** 1069-1086.

GELVIN, S. B. 2003. *Agrobacterium*-mediated plant transformation: the biology behind the "gene-jockeying" tool. *Microbiology and Molecular Biology Reviews,* 67**,** 16-37.

GEPSTEIN, S., SABEHI, G., CARP, M. J., HAJOUJ, T., NESHER, M. F. O., YARIV, I., DOR, C. & BASSANI, M. 2003. Large-scale identification of leaf senescence-associated genes. *The Plant Journal,* 36**,** 629-642.

GORAI, M., HACHEF, A. & NEFFATI, M. 2010. Differential responses in growth and water relationship of *Medicago sativa* (L.) cv. Gabès and *Astragalus gombiformis* (Pom.) under water-limited conditions. *Emirates Journal of Food and Agriculture,* 22**,** 1-12.

GRUDKOWSKA, M. & ZAGDANSKA, B. 2004. Multifunctional role of plant cysteine proteinases. *Acta Biochimica Polonica-English Edition***,** 609-624.

GUENOUNE, D., AMIR, R., BADANI, H., WOLF, S. & GALILI, S. 2002. Combined expression of S-VSPα in two different organelles enhances its accumulation and total lysine production in leaves of transgenic tobacco plants. *Journal of Experimental Botany,* 53**,** 1867-1870.

GUPTA, P. 2008. *Elements of biotechnology*, Rastogi Publications.

GUTIÉRREZ-CAMPOS, R., TORRES-ACOSTA, J. A., DE JESÚS PÉREZ-MARTÍNEZ, J. & GÓMEZ-LIM, M. A. 2001. Pleiotropic effects in transgenic tobacco plants expressing the oryzacystatin I gene. *HortScience,* 36**,** 118-119.

GUTIERREZ-CAMPOS, R., TORRES-ACOSTA, J. A., SAUCEDO-ARIAS, L. J. & GOMEZ-LIM, M. A. 1999. The use of cysteine proteinase inhibitors to engineer resistance against potyviruses in transgenic tobacco plants. *Nature Biotechnology,* 17**,** 1223-1226.

HABIB, H. & FAZILI, K. M. 2007. Plant protease inhibitors: a defense strategy in plants. *Biotechnology and Molecular Biology Review,* 2**,** 68-85.

HAMAYUN, M., KHAN, S. A., SHINWARI, Z. K., KHAN, A. L., AHMAD, N. & LEE, I.-J. 2010. Effect of polyethylene glycol induced drought stress on physio-hormonal attributes of soybean. *Pakistan Journal of Botany,* 42**,** 977-986.

HAQ, S. K., ATIF, S. M. & KHAN, R. H. 2004. Protein proteinase inhibitor genes in combat against insects, pests, and pathogens: natural and engineered phytoprotection. *Archives of Biochemistry and Biophysics,* 431**,** 145-159.

HARTL, M., GIRI, A. P., KAUR, H. & BALDWIN, I. T. 2010. Serine protease inhibitors specifically defend *Solanum nigrum* against generalist herbivores but do not influence plant growth and development. *The Plant Cell Online,* 22**,** 4158-4175.

HE, X. & KERMODE, A. R. 2003. Proteases associated with programmed cell death of megagametophyte cells after germination of white spruce (*Picea glauca*) seeds. *Plant Molecular Biology,* 52**,** 729-744.

HUANG, G.-T., MA, S.-L., BAI, L.-P., ZHANG, L., MA, H., JIA, P., LIU, J., ZHONG, M. & GUO, Z.-F. 2012. Signal transduction during cold, salt, and drought stresses in plants. *Molecular Biology Reports,* 39**,** 969-987.

HUANG, Y., XIAO, B. & XIONG, L. 2007. Characterization of a stress responsive proteinase inhibitor gene with positive effect in improving drought resistance in rice. *Planta,* 226**,** 73-85.

HWANG, J., HONG, J., LIM, C., CHEN, H., JE, J., YANG, K., KIM, D., CHOI, Y., LEE, S. & LIM, C. 2010a. Distinct expression patterns of two Arabidopsis phytocystatin genes, AtCYS1 and AtCYS2, during development and abiotic stresses. *Plant Cell Reports,* 29**,** 905-915.

HWANG, J. E., HONG, J. K., LIM, C. J., CHEN, H., JE, J., YANG, K. A., KIM, D. Y., CHOI, Y. J., LEE, S. Y. & LIM, C. O. 2010b. Distinct expression patterns of two Arabidopsis phytocystatin genes, AtCYS1 and AtCYS2, during development and abiotic stresses. *Plant cell reports***,** 1-11.

IKEA, J., INGELBRECHT, I., UWAIFO, A. & THOTTAPPILLY, G. 2004. Stable gene transformation in cowpea (*Vigna unguiculata* L. walp.) using particle gun method. *African Journal of Biotechnology,* 2**,** 211-218.

KAPRI-PARDES, E., NAVEH, L. & ADAM, Z. 2007. The thylakoid lumen protease Deg1 is involved in the repair of photosystem II from photoinhibition in Arabidopsis. *The Plant Cell Online,* 19**,** 1039-1047.

KEYSER, H. H. & LI, F. 1992. Potential for increasing biological nitrogen fixation in soybean. *Plant and Soil,* 141**,** 119-135.

KHANNA-CHOPRA, R., SRIVALLI, B. & AHLAWAT, Y. 1999. Drought induces many forms of cysteine proteases not observed during natural senescence. *Biochemical and Biophysical Research Communications,* 255**,** 324-327.

KIGGUNDU, A., MUCHWEZI, J., VAN DER VYVER, C., VILJOEN, A., VORSTER, J., SCHLÜTER, U., KUNERT, K. & MICHAUD, D. 2010. Deleterious effects of plant

cystatins against the banana weevil *Cosmopolites sordidus*. *Archives of Insect Biochemistry and Physiology,* 73**,** 87-105.

KONNO, K., HIRAYAMA, C., NAKAMURA, M., TATEISHI, K., TAMURA, Y., HATTORI, M. & KOHNO, K. 2004. Papain protects papaya trees from herbivorous insects: role of cysteine proteases in latex. *The Plant Journal,* 37**,** 370-378.

KOORNNEEF, M. & MEINKE, D. 2010. The development of Arabidopsis as a model plant. *The Plant Journal,* 61**,** 909-921.

KRÜGER, J., THOMAS, C. M., GOLSTEIN, C., DIXON, M. S., SMOKER, M., TANG, S., MULDER, L. & JONES, J. D. G. 2002. A tomato cysteine protease required for Cf-2-dependent disease resistance and suppression of autonecrosis. *Science,* 296**,** 744-747.

KURODA, M., ISHIMOTO, M., SUZUKI, K., KONDO, H., ABE, K., KITAMURA, K. & ARAI, S. 1996. Oryzacystatins exhibit growth-inhibitory and lethal effects on different species of bean insect pests, *Callosobruchus chinensis* (Coleoptera) and *Riptortus clavatus* (Hemiptera). *Biosciences Biotechnology and Biochemistry,* 60**,** 209-212.

LECHTENBERG, B., SCHUBERT, D., FORSBACH, A., GILS, M. & SCHMIDT, R. 2003. Neither inverted repeat T-DNA configurations nor arrangements of tandemly repeated transgenes are sufficient to trigger transgene silencing. *The Plant Journal,* 34**,** 507-517.

LEPLE, J. C., BONADE-BOTTINO, M., AUGUSTIN, S., PILATE, G., LE TAN, V. D., DELPLANQUE, A., CORNU, D. & JOUANIN, L. 1995. Toxicity to *Chrysomela tremulae* (Coleoptera: Chrysomelidae) of transgenic populars expressing a cysteine proteinase inhibitor. *Molecular Breeding: New Strategies in Plant Improvement,* 1**,** 319-328.

LIAN, H.-L., YU, X., YE, Q., DING, X.-S., KITAGAWA, Y., KWAK, S.-S., SU, W.-A. & TANG, Z.-C. 2004. The role of aquaporin RWC3 in drought avoidance in rice. *Plant and Cell Physiology,* 45**,** 481-489.

LING, J.-Q., KOJIMA, T., SHIRAIWA, M. & TAKAHARA, H. 2003. Cloning of two cysteine proteinase genes, *CysP1* and *CysP2*, from soybean cotyledons by cDNA representational difference analysis. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression,* 1627**,** 129-139.

LIU, F., JENSEN, C. R. & ANDERSEN, M. N. 2003. Hydraulic and chemical signals in the control of leaf expansion and stomatal conductance in soybean exposed to drought stress. *Functional Plant Biology,* 30**,** 65-73.

LIVAK, K. J. & SCHMITTGEN, T. D. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods,* 25**,** 402-408.

MAHAJAN, S. & TUTEJA, N. 2005. Cold, salinity and drought stresses: an overview. *Archives of Biochemistry and Biophysics,* 444**,** 139-158.

MANAVALAN, L. P., GUTTIKONDA, S. K., PHAN TRAN, L.-S. & NGUYEN, H. T. 2009. Physiological and molecular approaches to improve drought resistance in soybean. *Plant and Cell Physiology,* 50**,** 1260-1276.

MARTINEZ, M. & DIAZ, I. 2008. The origin and evolution of plant cystatins and their target cysteine proteinases indicate a complex functional relationship. *BMC evolutionary biology,* 8**,** 198.

MEHLMER, N., PARVIN, N., HURST, C. H., KNIGHT, M. R., TEIGE, M. & VOTHKNECHT, U. C. 2012. A toolset of aequorin expression vectors for in planta studies of subcellular calcium concentrations in *Arabidopsis thaliana*. *Journal of Experimental Botany,* 63**,** 1751-1761.

MEINKE, D. W., CHERRY, J. M., DEAN, C., ROUNSLEY, S. D. & KOORNNEEF, M. 1998. Arabidopsis thaliana: a model plant for genome analysis. *Science,* 282**,** 662-682.

MELLO-FARIAS, P. C. & CHAVES, A. L. S. 2008. Advances in *Agrobacterium*-mediated plant transformation with enphasys on soybean. *Scientia Agricola*, 65**,** 95-106.

MITTLER, R., MERQUIOL, E., HALLAK-HERR, E., RACHMILEVITCH, S., KAPLAN, A. & COHEN, M. 2001. Living under a 'dormant' canopy: a molecular acclimation mechanism of the desert plant *Retama raetam*. *The Plant Journal,* 25**,** 407-416.

MOHAMMADKHANI, N. & HEIDARI, R. 2008. Effects of drought stress on soluble proteins in two maize varieties. *Turkish Journal of Biology,* 32**,** 23-30.

MOSOLOV, V. & VALUEVA, T. 2008. Proteinase inhibitors in plant biotechnology: A review. *Applied Biochemistry and Microbiology,* 44**,** 233-240.

MUNGER, A., COENEN, K., CANTIN, L., GOULET, C., VAILLANCOURT, L.-P., GOULET, M.-C., TWEDDELL, R., SAINSBURY, F. & MICHAUD, D. 2012. Beneficial 'unintended effects' of a cereal cystatin in transgenic lines of potato, *Solanum tuberosum*. *BMC Plant Biology*, 12**,** 1-12.

MUSKENS, M. M., VISSERS, A. A., MOL, J. M. & KOOTER, J. 2000. Role of inverted DNA repeats in transcriptional and post-transcriptional gene silencing. *Plant Molecular Biology*, 43**,** 243-260.

NAKAYAMA, N., SANEOKA, H., MOGHAIEB, R. E. A., PREMACHANDRA, G. S. & FUJITA, K. 2007. Response of growth, photosynthetic gas exchange, translocation of 13C-labelled photosynthate and N accumulation in two soybean (*Glycine max* L. Merrill) cultivars to drought stress. *International Journal of Agriculture and Biology,* 9**,** 669-674.

OHASHI, Y., NAKAYAMA, N., SANEOKA, H. & FUJITA, K. 2006. Effects of drought stress on photosynthetic gas exchange, chlorophyll fluorescence and stem diameter of soybean plants. *Biologia Plantarum,* 50**,** 138-141.

OLHOFT, P. M., FLAGEL, L. E., DONOVAN, C. M. & SOMERS, D. A. 2003. Efficient soybean transformation using hygromycin B selection in the cotyledonary-node method. *Planta,* 216**,** 723-735.

OLIVA, M. L. V. & SAMPAIO, M. U. 2009. Action of plant proteinase inhibitors on enzymes of physiopathological importance. *Anais da Academia Brasileira de CiÃªncias,* 81**,** 615-621.

OLIVA, M. L. V., SILVA, M. C. C., SALLAI, R. C., BRITO, M. V. & SAMPAIO, M. U. 2010. A novel subclassification for Kunitz proteinase inhibitors from leguminous seeds. *Biochimie,* 92**,** 1667-1673.

OLIVEIRA, A. S., XAVIER-FILHO, J. & SALES, M. P. 2003. Cysteine proteinases and cystatins. *Brazilian Archives of Biology and Technology,* 46**,** 91-104.

PAZ, M., MARTINEZ, J., KALVIG, A., FONGER, T. & WANG, K. 2006. Improved cotyledonary node method using an alternative explant derived from mature seed for efficient *Agrobacterium*-mediated soybean transformation. *Plant Cell Reports,* 25**,** 206-213.

PAZ, M. M., SHOU, H., GUO, Z., ZHANG, Z., BANERJEE, A. K. & WANG, K. 2004. Assessment of conditions affecting *Agrobacterium*-mediated soybean transformation using the cotyledonary node explant. *Euphytica,* 136**,** 167-179.

PEIRSON, S. N., BUTLER, J. N. & FOSTER, R. G. 2003. Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. *Nucleic Acids Research,* 31**,** e73-e73.

PERNAS, M., SÁNCHEZ-MONGE, R., GÓMEZ, L. & SALCEDO, G. 1998. A chestnut seed cystatin differentially effective against cysteine proteinases from closely related pests. *Plant Molecular Biology,* 38**,** 1235-1242.

PODEVIN, N. & DU JARDIN, P. 2012. Possible consequences of the overlap between the CaMV 35S promoter regions in plant transformation vectors used and the viral gene VI in transgenic plants. *GM Crops and Food: Biotechnology in Agriculture and the Food Chain,* 3**,** 296-300.

PONTIER, D., GAN, S., AMASINO, R., ROBY, D. & LAM, E. 1999. Markers for hypersensitive response and senescence show distinct patterns of expression. *Plant Molecular Biology,* 39**,** 1243-1255.

POORTER, H., NIKLAS, K. J., REICH, P. B., OLEKSYN, J., POOT, P. & MOMMER, L. 2012. Biomass allocation to leaves, stems and roots: meta-analyses of interspecific variation and environmental control. *New Phytologist,* 193**,** 30-50.

PRINS, A., VAN HEERDEN, P. D. R., OLMOS, E., KUNERT, K. J. & FOYER, C. H. 2008. Cysteine proteinases regulate chloroplast protein content and composition in tobacco leaves: a model for dynamic interactions with ribulose-1,5-bisphosphate carboxylase/oxygenase (Rubisco) vesicular bodies. *Journal of Experimental Botany,* 59**,** 1935-1950.

PUPPO, A., GROTEN, K., BASTIAN, F., CARZANIGA, F., SOUSSI, M., M.M., L., FELIPE, M., HARRISON, J., VANACKER, H. & FOYER, C. H. 2005. Legume nodule senescence: roles for redox and hormone signalling in the orchestration of the natural aging process. *New Phytologist,* 165**,** 683-701.

RAHBÉ, Y., FERRASSON, E., RABESONA, H. & QUILLIEN, L. 2003. Toxicity to the pea aphid *Acyrthosiphon pisum* of anti-chymotrypsin isoforms and fragments of

Bowman–Birk protease inhibitors from pea seeds. *Insect Biochemistry and Molecular Biology,* 33**,** 299-306.

RATH, A., GLIBOWICKA, M., NADEAU, V. G., CHEN, G. & DEBER, C. M. 2009. Detergent binding explains anomalous SDS-PAGE migration of membrane proteins. *Proceedings of the National Academy of Sciences,* 106**,** 1760-1765.

RAUTENGARTEN, C., USADEL, B., NEUMETZLER, L., HARTMANN, J., BÜSSIS, D. & ALTMANN, T. 2008. A subtilisin-like serine protease essential for mucilage release from Arabidopsis seed coats. *The Plant Journal,* 54**,** 466-480.

RAWLINGS, N. D. & BARRETT, A. J. 1999. MEROPS: the peptidase database. *Nucleic Acids Research,* 27**,** 325-331.

RAWLINGS, N. D., BARRETT, A. J. & BATEMAN, A. 2011. Asparagine peptide lyases: a seventh catalytic type of proteolytic enzymes. *Journal of Biological Chemistry,* 286**,** 38321-38328.

RAWLINGS, N. D., TOLLE, D. P. & BARRETT, A. J. 2004. Evolutionary families of peptidase inhibitors. *Biochemical Journal,* 378**,** 705.

REDDY, A. R., CHAITANYA, K. V. & VIVEKANANDAN, M. 2004. Drought-induced responses of photosynthesis and antioxidant metabolism in higher plants. *Journal of Plant Physiology,* 161**,** 1189-1202.

RICHAU, K. H., KASCHANI, F., VERDOES, M., PANSURIYA, T. C., NIESSEN, S., STÜBER, K., COLBY, T., OVERKLEEFT, H. S., BOGYO, M. & VAN DER HOORN, R. A. L. 2012. Subclassification and biochemical analysis of plant papain-like cysteine proteases displays subfamily-specific characteristics. *Plant Physiology,* 158**,** 1583-1599.

ROBERTS, I. N., CAPUTO, C., CRIADO, M. V. & FUNK, C. 2012. Senescence-associated proteases in plants. *Physiologia Plantarum,* 145**,** 130-139.

SCHALLER, A. 2004. A cut above the rest: the regulatory function of plant proteases. *Planta,* 220**,** 183-197.

SCHEFE, J. H., LEHMANN, K. E., BUSCHMANN, I. R., UNGER, T. & FUNKE-KAISER, H. 2006. Quantitative real-time RT-PCR data analysis: current concepts and the novel "gene expression's CT difference" formula. *Journal of Molecular Medicine,* 84**,** 901-910.

SCHUPPLER, U., HE, P.-H., JOHN, P. C. L. & MUNNS, R. 1998. Effect of water stress on cell division and Cdc2-like cell cycle kinase activity in wheat leaves. *Plant Physiology*, 117**,** 667-678.

SIMOVA-STOILOVA, L., VASEVA, I., GRIGOROVA, B., DEMIREVSKA, K. & FELLER, U. 2010. Proteolytic activity and cysteine protease expression in wheat leaves under severe soil drought and recovery. *Plant Physiology and Biochemistry,* 48**,** 200-206.

SIN, S.-F., YEUNG, E. C. & CHYE, M.-L. 2006. Downregulation of *Solanum americanum* genes encoding proteinase inhibitor II causes defective seed development. *The Plant Journal,* 45**,** 58-70.

SOARES-CORDEIRO, A. S., DRISCOLL, S. P., PELLNY, T. K., OLMOS, E., ARRABAÇA, M. C. & FOYER, C. H. 2009. Variations in the dorso-ventral organization of leaf structure and Kranz anatomy coordinate the control of photosynthesis and associated signalling at the whole leaf level in monocotyledonous species. *Plant, Cell and Environment,* 32**,** 1833-1844.

SOLOMON, M., BELENGHI, B., DELLEDONNE, M., MENACHEM, E. & LEVINE, A. 1999. The involvement of cysteine proteases and protease inhibitor genes in the regulation of programmed cell death in plants. *Plant Cell*, 11**,** 431-444.

SOMERS, D. A., SAMAC, D. A. & OLHOFT, P. M. 2003. Recent advances in legume transformation. *Plant Physiology,* 131**,** 892-899.

SRINIVASAN, T., KUMAR, K. R. R. & KIRTI, P. B. 2009. Constitutive expression of a trypsin protease inhibitor confers multiple stress tolerance in transgenic tobacco. *Plant and Cell Physiology,* 50**,** 541-553.

TAMAKI, S., MATSUO, S., WONG, H. L., YOKOI, S. & SHIMAMOTO, K. 2007. Hd3a protein is a mobile flowering signal in rice. *Science,* 316**,** 1033-1036.

TANG, W., NEWTON, R. J. & WEIDNER, D. A. 2007. Genetic transformation and gene silencing mediated by multiple copies of a transgene in eastern white pine. *Journal of Experimental Botany,* 58**,** 545-554.

TRÄNKNER, C., LEHMANN, S., HOENICKA, H., HANKE, M.-V., FLADUNG, M., LENHARDT, D., DUNEMANN, F., GAU, A., SCHLANGEN, K., MALNOY, M. & FLACHOWSKY, H. 2010. Over-expression of an FT-homologous gene of apple induces early flowering in annual and perennial plants. *Planta,* 232**,** 1309-1324.

URWIN, P. E., TROTH, K. M., ZUBKO, E. I. & ATKINSON, H. J. 2001. Effective transgenic resistance to *Globodera pallida* in potato field trials. *Molecular Breeding,* 8**,** 95-101.

VAN DER HOORN, R. A. L. 2008. Plant proteases: from phenotypes to molecular mechanisms. *Annual Review of Plant Biology,* 59**,** 191-223.

VAN DER HOORN, R. A. L., LEEUWENBURGH, M. A., BOGYO, M., JOOSTEN, M. H. A. J. & PECK, S. C. 2004. Activity profiling of papain-like cysteine proteases in plants. *Plant Physiology,* 135**,** 1170-1178.

VAN DER VYVER, C., SCHNEIDEREIT, J., DRISCOLL, S., TURNER, J., KUNERT, K. & FOYER, C. H. 2003. Oryzacystatin I expression in transformed tobacco produces a

conditional growth phenotype and enhances chilling tolerance. *Plant Biotechnology Journal,* 1**,** 101-112.

WANG, W., VINOCUR, B. & ALTMAN, A. 2003. Plant responses to drought, salinity and extreme temperatures: towards genetic engineering for stress tolerance. *Planta,* 218**,** 1-14.

WANG, X., WANG, Y., TIAN, J., LIM, B. L., YAN, X. & LIAO, H. 2009. Overexpressing *AtPAP15* enhances phosphorus efficiency in soybean. *Plant Physiology,* 151**,** 233-240.

WATANABE, H., ABE, K., EMORI, Y., HOSOYAMA, H. & ARAI, S. 1991. Molecular cloning and gibberellin-induced expression of multiple cysteine proteinases of rice seeds (oryzains). *Journal of Biological Chemistry,* 266**,** 16897-16902.

WU, F., BAO, W., LI, F. & WU, N. 2008. Effects of drought stress and N supply on the growth, biomass partitioning and water-use efficiency of *Sophora davidii* seedlings. *Environmental and Experimental Botany,* 63**,** 248-255.

XU, Z., ZHOU, G. & SHIMIZU, H. 2010. Plant responses to drought and rewatering. *Plant Signaling and Behavior,* 5**,** 649-654.

YORDANOV, I., VELIKOVA, V. & TSONEV, T. 2003. Plant responses to drought and stress tolerance. *Bulgarian Journal of Plant Physiology***,** 187-206.

ZHANG, X., LIU, S. & TAKANO, T. 2008. Two cysteine proteinase inhibitors from *Arabidopsis thaliana*, *AtCYSa* and *AtCYSb*, increasing the salt, drought, oxidation and cold tolerance. *Plant Molecular Biology,* 68**,** 131-143.

ZLATEV, Z. & LINDON, F. C. 2012. An overview on drought induced changes in plant growth, water relations and photosynthesis. *Emirates Journal of Food and Agriculture,* 24**,** 57-72.

ZLATEV, Z. & YORDANOV, I. 2004. Effects of soil drought on photosynthesis and chlorophyll fluorescence in common bean plants. *Bulgarian Journal of Plant Physiology*, 30**,** 3-18.

SUPPLEMENTARY DATA

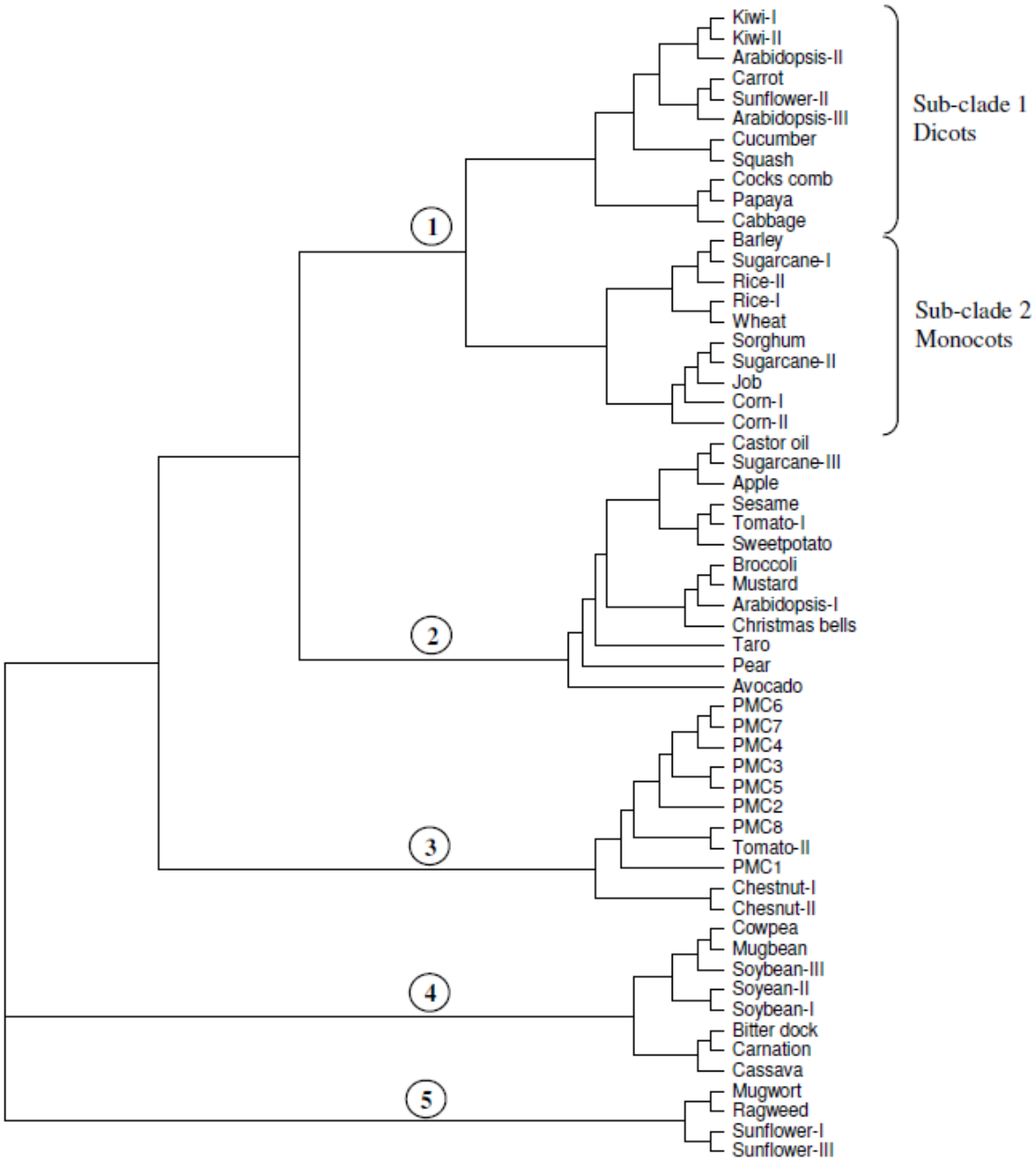
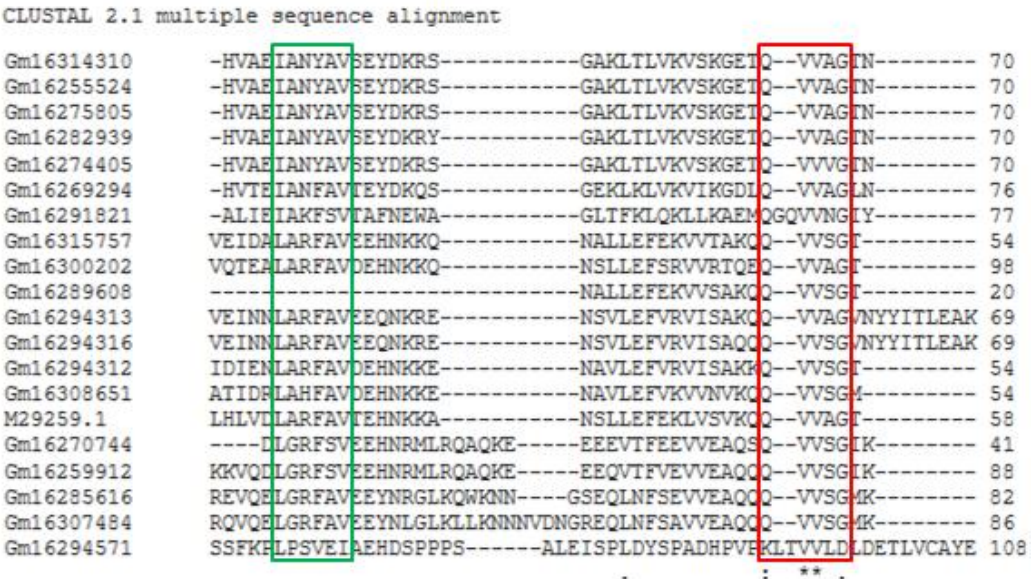


Figure 1 Phylogenetic tree of known cystatins extracted from various sequence databases

A



B

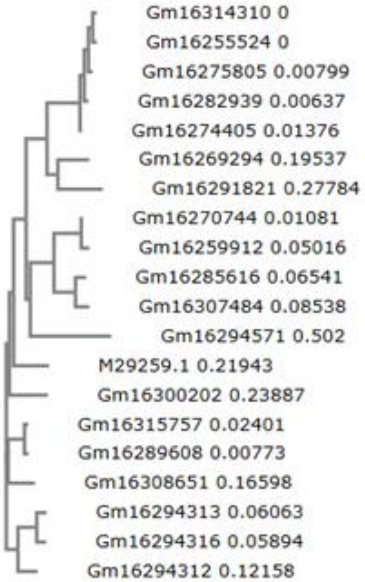
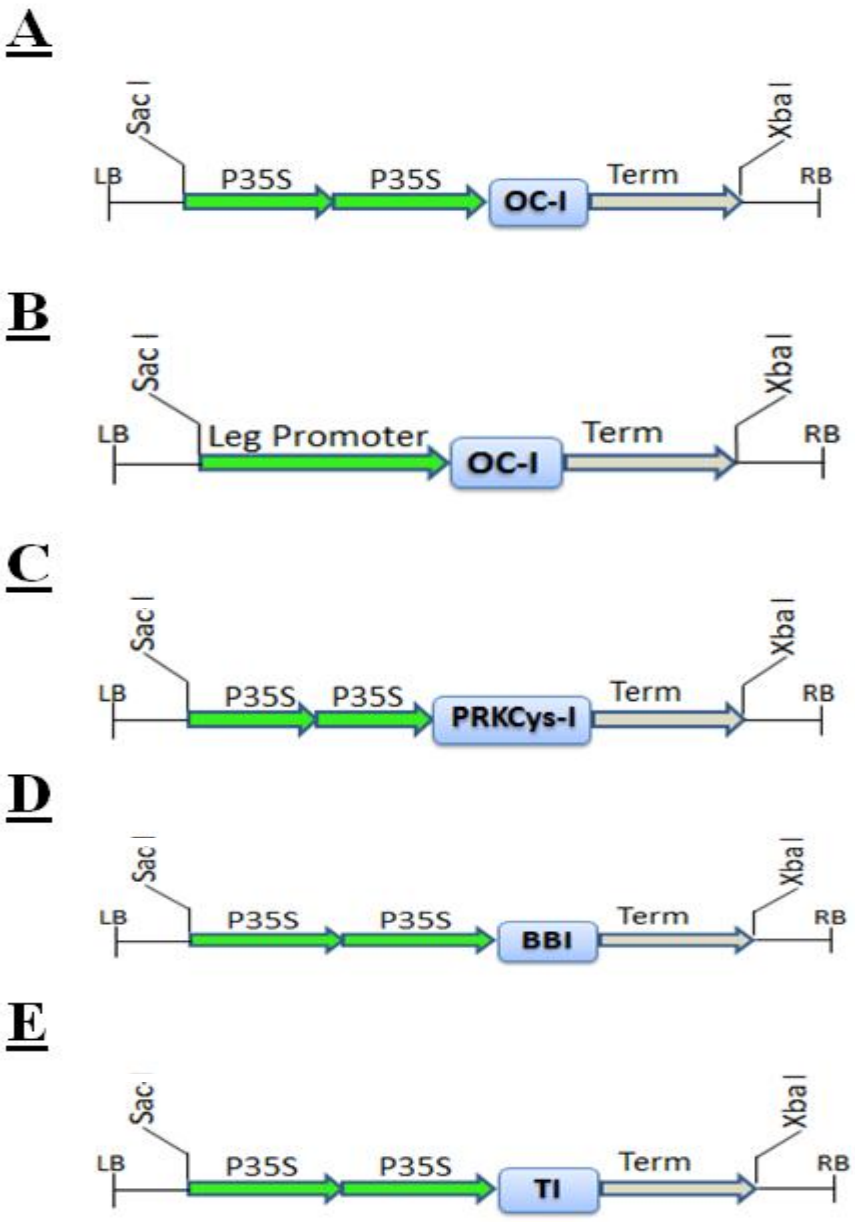


Figure 2 Part of the sequence alignment of the soybean cystatins and the rice OC-I cystatin (M2259.1) (A) and the phylogenetic tree. The green box highlights the LARFAV region and the red box highlights the QVVxG region.

ANNEXURES



Annex 1 Schematic representation of the constructs used for transformation. (A), (B) and (C) OC-I constructs directed to the cytosol, chloroplast and nodules, respectively. (D) and (E) are the serine protease inhibitors (BBI and TI) directed to the cytosol.

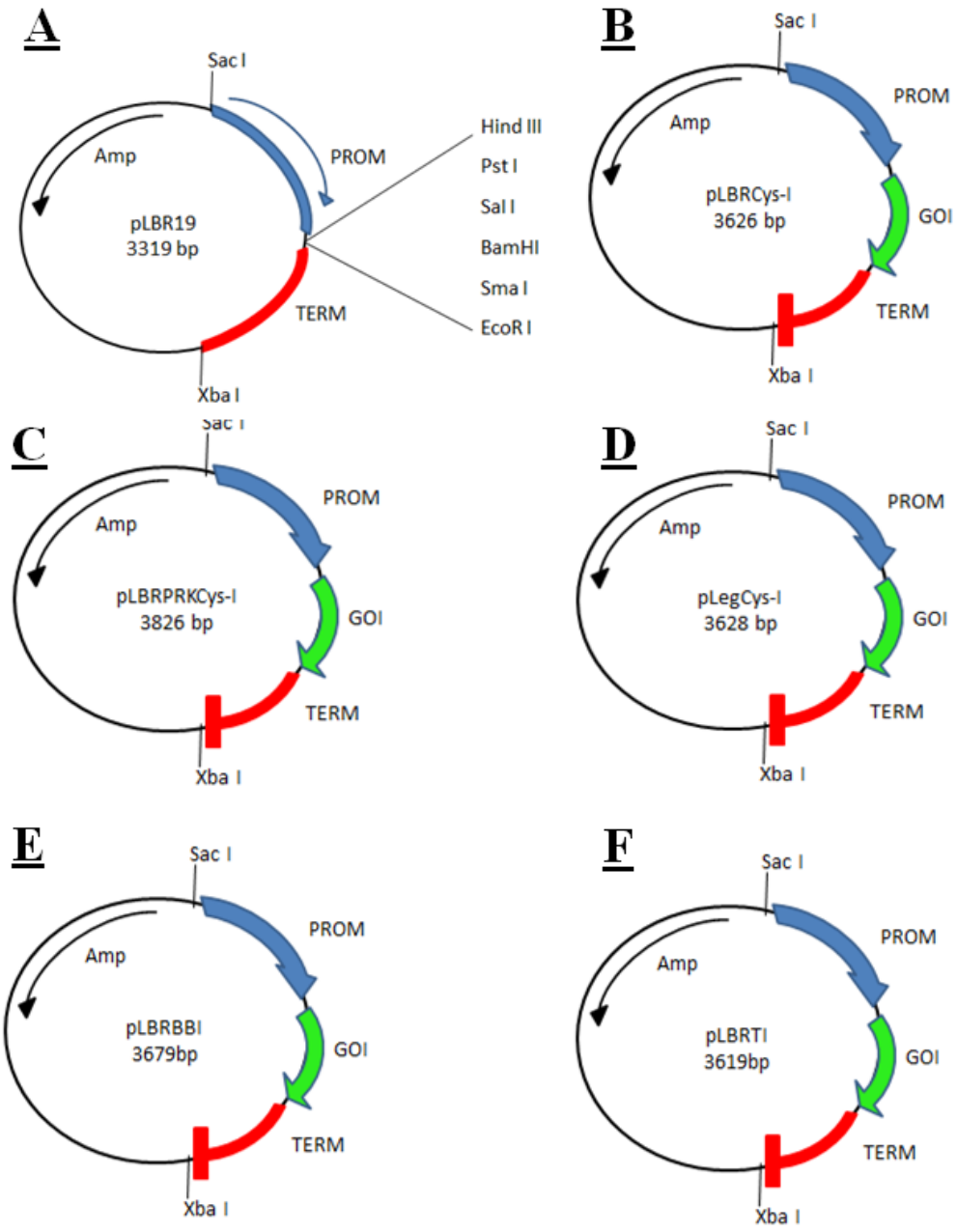
A

```
  1 ATGTCGAGCGACGGAGGGCCGGTGCTTGGCGGCGTCGAGCCGGTGGGGAACGAGAACGAC
    >>>>>>>>>>>>>>>>>>>>

 61 CTCCACCTCGTCGACCTCGCCCGCTTCGCCGTCACCGAGCACAACAAGAAGGCCAATTCT

121 CTGCTGGAGTTCGAGAAGCTTGTGAGTGTGAAGCAGCAAGTTGTCGCTGGCACTTTGTAC

181 TATTTCACAATTGAGGTGAAGGAAGGGGATGCCAAGAAGCTCTATGAAGCTAAGGTCTGG

241 GAGAAACCATGGATGGACTTCAAGGAGCTCCAGGAGTTCAAGCCTGTCGATGCCAGTGCA
                                                                <<<

301 AATGCCTAAGgCCCATc
    <<<<<<<<<<<<<<<<<
```

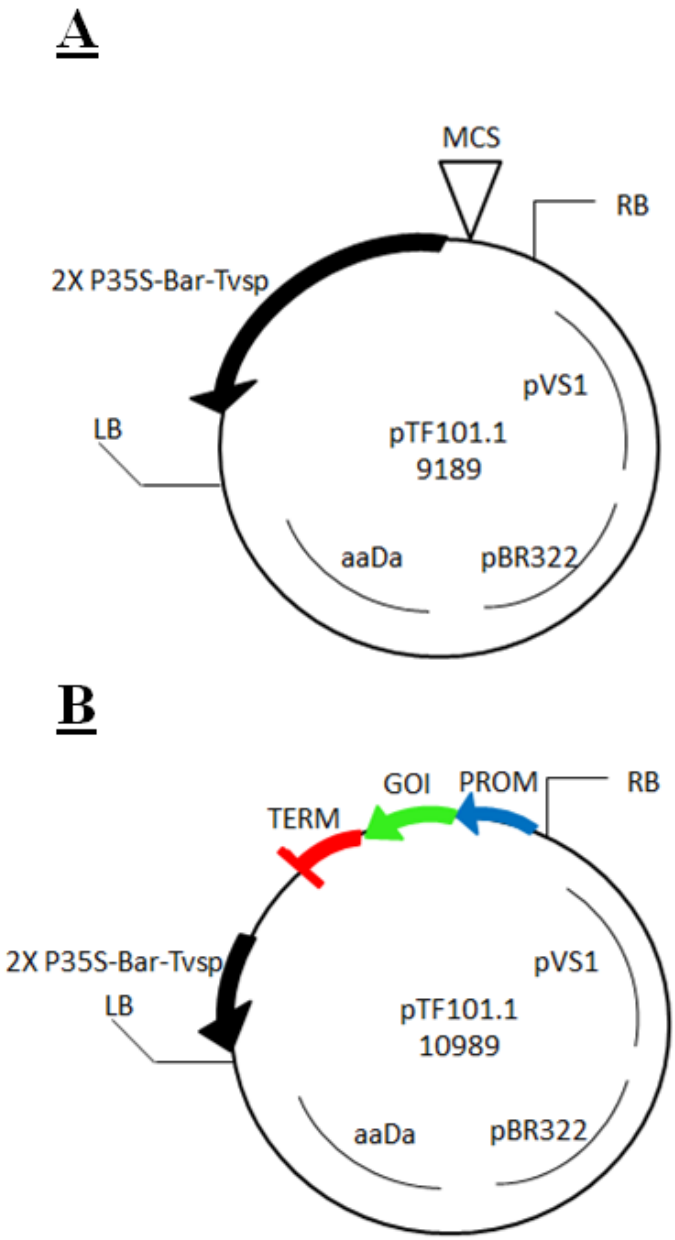
B

```
  1 TCACTCTTCAAGCCTTCTATATTAATAAGTATTGGATGTGAAGTTGTTGCATATCTTGCA
    >>>>>>>>>>>>>>>>>>

 61 GACTCTAGAGGATCCCCATCGAATTCCTGCAGTCTAGAGGCCTCCATGtCGAGCGACGGA

121 GGGCCGGTGCTTGGCGGCGTCGAGCCGGTGGGGAACGAGAACGACCTCCACCTCGTCGAC

181 CTCGCCCGCTTCGCCGTCACCGAGCACAACAAGAAGGCCAATTCTCTGCTGGAGTTCGAG

241 AAGCTTGTGAGTGTGAAGCAGCAAGTTGTCGCTGGCACTTTGTACTATTTCACAATTGAG

301 GTGAAGGAAGGGGATGCCAAGAAGCTCTATGAAGCTAAGGTCTGGGAGAAACCATGGATG

361 GACTTCAAGGAGCTCCAGGAGTTCAAGCCTGTCGATGCCAGTGCAAATGCCTAAGgCCCA
                                                    <<<<<<<<<<<<<<<<<<

421 Tc
    <<
```

Annex 2 (A) Primer alignment on the OC-I sequence and on the leghemoglobin promoter as well as the OC-I sequence (B).



Annex 3 Schematic representation of the plasmids used for subcloning the oryzacystatin I (OC-I) and serine protease inhibitors (BBI and TI) genes. (A) Parental plasmid pLBR 19, (B) and (C) OC-I constructs directed to the cytosol, chloroplast and nodules, respectively. (E) and (F) are the serine protease inhibitors (BBI and TI) directed to the cytosol.



Annex 4 Plasmid vector pTF101.1 map (A). The plasmid has a spectinomycin resistant marker gene (aadA) for bacterial selection. The plant selectable marker gene cassette consists of a double 35S promoter (2x P35S) of the cauliflower mosaic virus (CaMV). (B) The plasmid vector after addition of the construct of interest. Prom, promoter; GOI, gene of interest and term, terminator.

A

```
CLUSTAL 2.0.12 multiple sequence alignment

OCI_amplicon            ATGTCGAGCGACGGAGGGCCGGTGCTTGGCGGCGTCGAGCCGGTGGGGAACGAGAACGAC 60
_gi|169806_539-652_     ATGTCGAGCGACGGAGGGCCGGTGCTTGGCGGCGTCGAGCCGGTGGGGAACGAGAACGAC 60
                        ************************************************************

OCI_amplicon            CTCCACCTCGTCGACCTCGCCCGCTTCGCCGTCACCGAGCACAACAAGAAGGCCAATTCT 120
_gi|169806_539-652_     CTCCACCTCGTCGACCTCGCCCGCTTCGCCGTCACCGAGCACAACAAGAAGGCCAATTCT 120
                        ************************************************************

OCI_amplicon            CTGCTGGAGTTCGAGAAGCTTGTGAGTGTGAAGCAGCAAGTTGTCGCTGGCACTTTGTAC 180
_gi|169806_539-652_     CTGCTGGAGTTCGAGAAGCTTGTGAGTGTGAAGCAGCAAGTTGTCGCTGGCACTTTGTAC 180
                        ************************************************************

OCI_amplicon            TATTTCACAATTGAGGTGAAGGAAGGGGATGCCAAGAAGCTCTATGAAGCTAAGGTCTGG 240
_gi|169806_539-652_     TATTTCACAATTGAGGTGAAGGAAGGGGATGCCAAGAAGCTCTATGAAGCTAAGGTCTGG 240
                        ************************************************************

OCI_amplicon            GAGAAACCATGGATGGACTTCAAGGAGCTCCAGGAGTTCAAGCCTGTCGATGCCAGTGCA 300
_gi|169806_539-652_     GAGAAACCATGGATGGACTTCAAGGAGCTCCAGGAGTTCAAGCCTGTCGATGCCAGTGCA 300
                        ************************************************************

OCI_amplicon            AATGCCTAAGGCCCATC 317
_gi|169806_539-652_     AATGCCTAAGGCCCATC 317
                        *****************
```

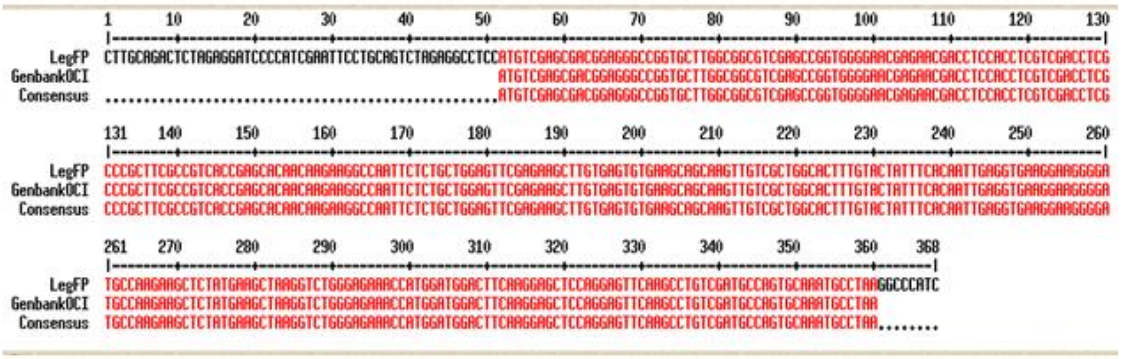
B

PRK sequence

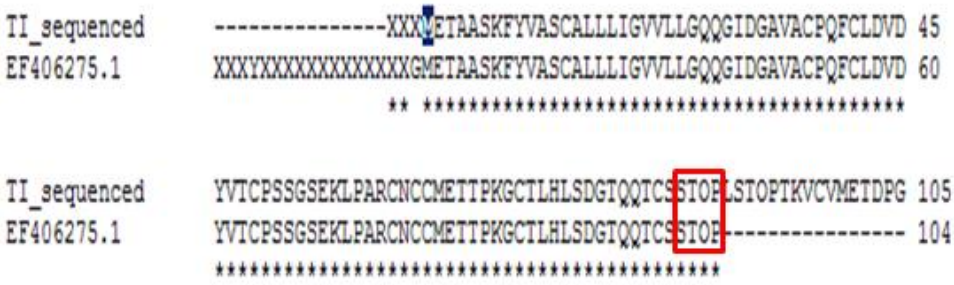
ATGGCTGTCTCAACTATCTACTCAACACAAGCTCTCAATTCAACTCATTTCTTAACC
TCTTCTTCCTCCTCCAAACAAGTCTTCCTCTACCGTCGTCAACCACAAACCAACCGT
AGATTCAACACACTCATCACTTGCGCACAAGAAACCATCGTGATCGGACTAGCTGC
TGACTCTGGC

Annex 5 (A) Sequence alignment of OC-I with the Genbank OC-I (accession no. M29259.1). (B) sequence of the PRK.

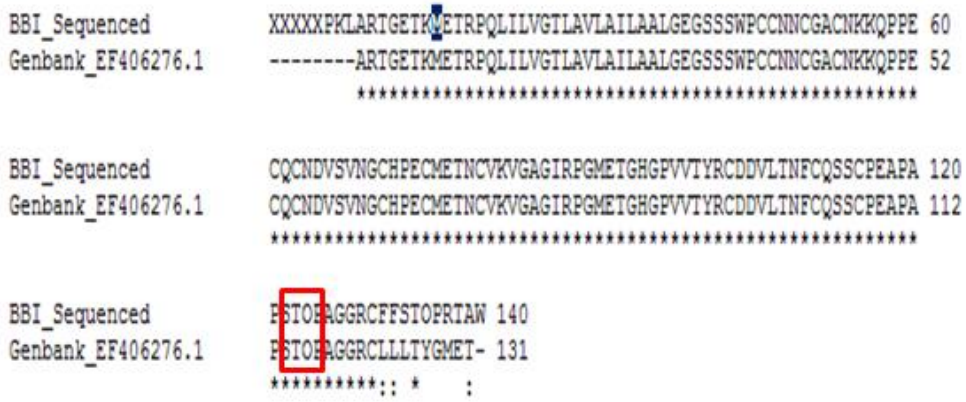
A



B



C



Annex 6 Sequence alignment of the LegCys-I (A), translated serpin (B), translated BBI (C) with Genbank OC-I, serpin and BBI, respectively. For (B) and (C), the start (methionine) and stop (TGA, TAG or TAA) codons are indicated with blue and red.