Studying the extracellular loop domain of DEK1 in evolutionary context

Utilizing a dominant negative effect and cross-species domain swap chimeras in Arabidopsis thaliana

Kostadinka Mikkelsen

Master thesis in Applied Biotechnology

Department for Education and Natural Science

HØGSKOLEN I HEDMARK

2011

Acknowledgements

This Master Thesis study was performed at the Department of Education and Natural Sciences at Hedmark University College, located in Hamar, Norway, in the period between January 2010 and August 2011, under the supervision of the Associated Professors Wenche Johansen and Robert C. Wilson, and Professor Odd-Arne Olsen.

I am very grateful to Hedmark University College for providing me with the opportunity of taking this education at excellent facilities and working in a pleasant environment with highly educated and helpful employees.

First I would like to thank Wenche Johansen for helping me to write this thesis and for teaching me precision in the laboratory; thank you Wenche, you are my rock star with pipette. I am very thankful for all the help I got from Robert C. Wilson and Professor Odd-Arne Olsen, for the great ideas and enthusiastic support. Thanks to the whole DEK1 scientific team, for the regular inspiration in our DEK1 meetings and DEK1 journal club.

I want to express my gratitude to all employees at Biohuset Hamar for all the help I have received and for the good and friendly environment they provide for the students at Hedmark University College. Thanks to my international classmates, for making this study time a culturally rich and fun experience.

And last but not least I want to thank my family for being supportive and understanding while I was working with this project.

Kostadinka Mikkelsen

Hamar, September 1, 2011

Table of Contents

Abbreviations

At	*Arabidopsis thaliana*
BLAST	Basic Local Alignment Search Tool
bp	base pair
CALP	calpain
GFP	Green Fluorescent Protein
DEK1	Defective Kernel 1
cDNA	complementary DNA
5’-RACE	Five prime Rapid Amplification of cDNA Ends
gDNA	genomic DNA
MEM	Membrane Domain of Dek1
Mv	*Mesostigma viride*
nptII	neomycin phosphotransferase II
DEL or MEM-DEL	Membrane Domain with Deleted Loop
PCR	Polymerase Chain Reaction
Pp	*Physcomitrella patens*
qRT-PCR	quantitative real-time PCR
RNAi	Ribonucleic Acid interference
Zm	*Zea mays*

Abstract

DEFECTIVE KERNEL 1 (DEK1) is a highly conserved phytocalpain, embedded in the membrane of all plant cells examined. Null alleles of *dek1* are lethal and dek1 mutants have shown that functional DEK1 protein is essential for proper development of both epidermal cell layer in plant embryos, and for positional dependant aleurone cell formation during seed development in maize, *Arabidopsis* and in tobacco plant. Transgenic *Arabidopsis* lines overexpressing the membrane-anchored domain of DEK1 have shown a dominant negative phenotype, suggesting that the AtDEK1-MEM plays role in the transmission of proposed positional cues. In this research we used the dominant negative effect and cross-species domain swap analysis in order to study conserved features of the extracellular DEK1-LOOP domain in *Arabidopsis thaliana* in an evolutionary context. Two constructs, designed to overexpress the At*MEM* part of *DEK1*, where the *LOOP* domain was replaced by the *DEK1-LOOP* sequence from other organisms, were successfully cloned and transformed into wild type *Arabidopsis thaliana* background. Transgenic plants overexpressing the *AtDEK1-MEM* construct, where the *Arabidopsis* loop sequence was replaced with loop from the unicellular alga *Mesostigma viride,* showed phenotype frequencies significantly similar to those observed in plants overexpressing the membrane-anchored domain of DEK1 that results in dominant negative phenotype. These results suggest that the DEK1-LOOP of *Mesostigma viride* most probably interacts with the same positional cues or the same family of proteins as the DEK1-LOOP in *Arabidopsis thaliana*. Transgenic lines overexpressing the *AtDEK1-MEM* construct, where the *AtDEK1-LOOP* sequence was replaced with *DEK1-LOOP* from maize (*Zea mays*), only alleviated the dominant negative effect by exerting significantly lower frequency of the respective phenotype, also compatible to phenotypic frequencies observed in plants overexpressing the AtDEK1-MEM-DEL construct, where the loop domain is deleted. These results suggest that the DEK1-LOOP domain could have functionally diverged in species-specific manner between *Arabidopsis* and maize. However, all these speculations are true only if the observed phenotypes are result of dominant negative effect caused by the DEK1-LOOP domain. Q-PCR analyses of the transgenic transcripts in two and five weeks old plants could not be correlated to the observed phenotypes.

1. Introduction

1.1 Defective Kernel 1 (Dek1)

The DEFECTIVE KERNEL 1 is a highly conserved phytocalpain, which has been shown to be essential for proper development of both the epidermal cell layer in plant embryos and for positional-dependent aleurone cell formation during seed maturation in *Zea mays*, *Arabidopsis thaliana* and *Nicotiana tabacum* (Lid et al., 2002 & 2005; Ahn et al., 2004; Johnson et al., 2005). The interest in defective kernels in maize led to the discovery of Dek1. A large class of maize mutations effecting both endosperm and embryo development of kernels was characterized already in 1920 by Jones and Mangelsdorf (Jones 1920; Mangelsdorf, 1923 & 1926) and was termed *defective kernel (dek)* mutations. The maize *DEFECTIVE KERNEL 1* (*DEK1*) gene was the first to be phenotypically characterized and cloned in studies of aleurone cell fate specification in maize endosperm development (Becraft & Asuncion-Crabb, 2000; Becraft et al., 2002; Lid et al., 2002). The aleurone cells compose the outer epidermal layer/s of the endosperm. In most cereals, including typical maize (*Zea mays*) lines, the aleurone is a single celled layer, in barley (*Hordeum vulgare*) the aleurone comprises about three cell layers and in rice (*Oryza sativa*), the number of layers varies. The most well known function of the aleurone is as a digestive tissue along with protection from and communication with the external environment (reviewed by Becraft & Yi, 2010). In *Arabidopsis thaliana* the aleurone also controls seed dormancy and germination (Bethke et al., 2007).

1.1.1 DEK1 is Membrane Anchored Member of the Calpain Gene Superfamily

The *DEK1* gene in maize encodes a 240 kDa protein predicted to be anchored in the plasma membrane of all plant cells. To identify the *DEK1* gene, Lid et al (2002) screened a population of maize lines containing a high-copy number of Mu elements, then identified and cloned a Mu element cosegregating with the dekl mutant phenotype, and used the sequences flanking the Mu insertion to clone maize *DEK1*. The predicted sequence identified *DEK1* as a member of the calpain gene superfamily by the presence of a conserved cysteine proteinase domain (II) shared by all calpains, and by domain III, found in most

calpains (Sorimachi & Suzuki, 2001). Submission of the DEK1 amino acid sequence in the TMHMM 2.0 program (Krogh et al., 2001) predicted 21 transmembrane domains (Lid et al., 2002).

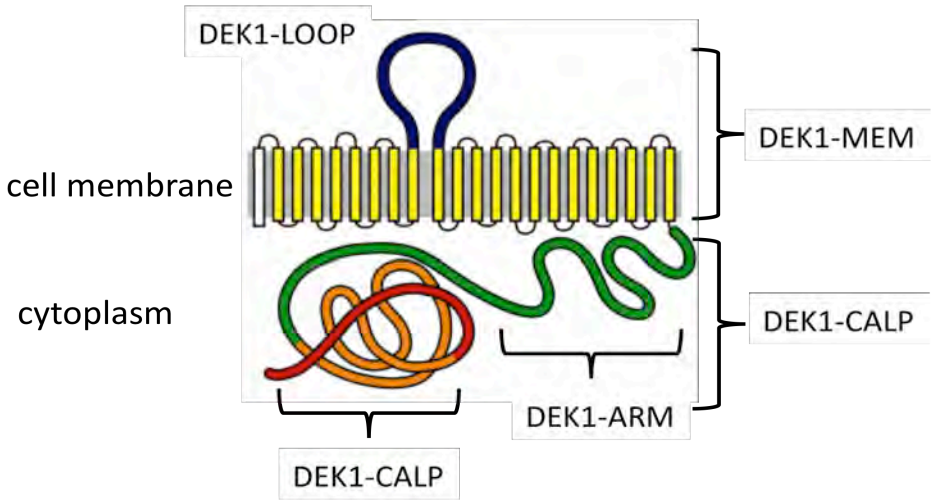


Figure 1 Predicted structure of maize DEK1, model based on the TMHMM 2.0 program (Krogh et al., 2001). DEK1-MEM consists of the membrane anchored domains (one gray and the rest yellow) and the extracellular loop domain (in blue). DEK1-CALP corresponds to the cytoplasmic part of the protein, divided into DEK-ARM and the true DEK1-CALP, the calpain (modified figure from Lid et al, 2002).

The maize (*Zea mays*) DEK1 protein consists of 2159 amino acid residues and contains about seven distinguishable domains (Figure 1). Domain A (gray) is located at the N-terminus containing a predicted endoplasmic reticulum and membrane targeting signals. Following is domain B1 (yellow) containing eight transmembrane stretches disrupted by ~ 300 amino acids long extracellular loop, the C domain (blue). Domain B2 (yellow) consist of 13 membrane-anchored stretches, followed by the DEK1-ARM, domain D (green) a hydrophilic and charged region, which connect the membrane part (DEK1-MEM) to the calpain domain (DEK1-CALP), at the C-terminal and the very end of the protein. Domain II (orange) is the highly conserved, catalytic, cystein proteinase domain and domain III (red) is found in most calpains. The presence of those two domains (II and III) was the reason why DEK1 was identified as a member of the Calpain Gene Superfamily (Lid et al., 2002; Wang et al., 2003).

1.1.2 The Calpain Domain

Calpains are a large class of intracellular, calcium dependent, cysteine proteinases that play essential roles in multiple developmental pathways in animals; like cell differentiation, proliferation and apoptosis (Perrin & Huttenlocher, 2002; Goll et al., 2003). Active calpain enzymes in mammals and invertebrates associate both with the plasma membrane and with internal membranes, like the endosomal reticulum (ER). In addition calpains associate with proteins or phospholipids in the plasma membrane and undergo autolysis (Sato & Kawashima, 2001; Shao et al., 2006). The calpain domain is the most conserved part of DEK1 in plants and the modeled maize calpain shows high structural similarity in addition to proteinase activity similar to animal milli (m)-calpains (Wang et al., 2003). There are currently 14 known human calpain isoform genes, which are defined by the presence of a protease domain that is similar to that found in micro (μ)-calpain and m-calpain which are the two most extensively studied and ubiquitously expressed isoforms (Goll et al., 2003). In contrast to m-calpain, the DEK1 calpain is active in the absence of calcium, although its activity is stimulated by the addition of calcium (Wang et al., 2003). In animals, calpains are activated in a multistep route involving translocation to the plasma membrane, activation by Ca^{2+}, and catalytic cleavage by intramolecular processes (Zalewska et al., 2004). In maize, DEK1-CALP failed to complement the loss-of-aleurone cell phenotype in homozygous maize dek1 mutant endosperm under the control of the 27-kD g-Zein promoter. The expression of AtDEK1-CALP under the control of the AtDEK1 promoter also failed to fully complement the *Arabidopsis* dek1-1 mutant phenotype. These results led to the conclusion that DEK1-MEM is necessary for the DEK1-CALP activity (Tian et al., 2007). In contrast, Johnson et al. (2008) were able to fully complement the embryo lethal mutant *dek1-3* phenotype with a DEK1-CALP construct under the control of ribosomal protein 5A (RPS5A) promoter, showing that the membrane part of the protein is not required for its function, but probably plays a regulatory role (Johnson et al., 2008). Localization of AtDEK1 to the plasma membrane, endomembrane system and the cytosol suggested that DEK1 might undergo a regulatory cleavage event, like those described in animal calpains. Further investigation implied that DEK1 undergoes autolytic cleavage within the juxtamembrane domain and at the start of the calpain domain, releasing it from the membrane into the cytoplasm and that the cleaved calpain domain is the biologically active form of the DEK1 protein (Johnson et al., 2008).

1.1.3 Localization of DEK1

The maize DEK1 protein was predicted to have 21 membrane-spanning domains, interrupted by an extracytosolic side and ending of cytosolic part by the TMHMM 2.0 program (Lid et al., 2002). The subcellular localization of DEK1 was determined by polyclonal antibodies generated in rabbits against five different regions of DEK1. The results showed that DEK1 localizes to the plasma membrane and endosomal compartments (Tian et al., 2007). To test whether posttranscriptional regulation and subcellular localization of the DEK1 protein are likely to be important for its function (as it is in animals), Johnson et al. (2008) designed a construct expressing DEK1 fused with the GFP protein and used confocal microscopy to observe the subcellular localization of DEK1-GFP protein in fully complemented *Arabidopsis* dek1 mutant plants. The GFP-fluorescence observed in all studied tissues was associated with the plasma membrane and the endoplasmic reticulum (ER). In addition the presence of a GFP-pool was detected in the cytoplasm (Johnson et al., 2008).

1.1.4 DEK1s role in aleurone fate specification and maintenance

Loss-of-function mutants in maize dek1 lack aleurone layer, indicating that functional DEK1 is required for the signaling or perception of signals that specify the development of the outer cell coat (Becraft et al., 2002; Lid et al., 2002). Similar observations have been reported in *Arabidopsis thaliana* and in rice (Lid et al., 2005; Hibara et al., 2009).

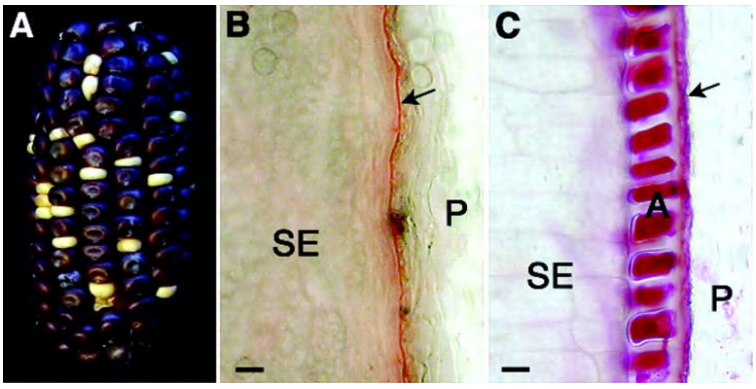


Figure 2 Phenotypes of dek1-mum1 homozygous maize kernels. A: An ear that segregates 3:1 for wild-type (dark) and dek1 (white) grains. B: Section of a dek1 kernel: starchy endosperm (SE) is located in the periphery of the endosperm close to the remnants of the nucellus (arrow); P represents the maternal pericarp. C: Hand section of a wild type kernel with aleurone cells (A,) starchy endosperm (SE) and pericarp (P). (Modified figure from Lid et al. 2002).

The aleurone layer of specific maize line that contains anthocyanin pigments, which give the kernels a dark, bluish-black color was used to visualize the defect in the aleurone cell layer. Homozygous dek1 maize kernels are white because they lack the aleurone layer and consequently the anthocyanin pigments (Figure 2 A). Homozygous *dekl-mum1* mutant endosperms initiate aleurone cell fate specification at a high frequency, indicated by the presence of peripheral cells containing darkly stained cytoplasm, and in mature endosperm, aleurone cells are almost completely lacking (Figure 2 C). These results suggested that the initiation of aleurone cell fate does not depend on DEK1, and that its function plays a role in maintaining the aleurone cell fate specified at early developmental stage in the endosperm (Lid et al., 2002).

Homozygous *Arabidopsis* dek1 mutant embryos lack proper control of mitotic divisions and the endosperm forms only a partial aleurone-like cell layer (Lid et al., 2005). Loss of AtDEK1 activity in *Arabidopsis* leads to early embryonic arrest phenotype (Lid et al., 2005; Johnson et al., 2005), which makes studying the role of DEK1 in the maintenance of epidermal identity after germination and during embryogenesis impossible. This problem was circumvented by the production of transgenic *Arabidopsis thaliana* plants expressing an AtDek1-RNAi construct under the control of the 35S promoter, which is not active during early embryo development. This approach allowed the embryo to partially complete its development before the AtDEK1 transcript have being silenced by the RNAi construct (Johnson et al, 2005).

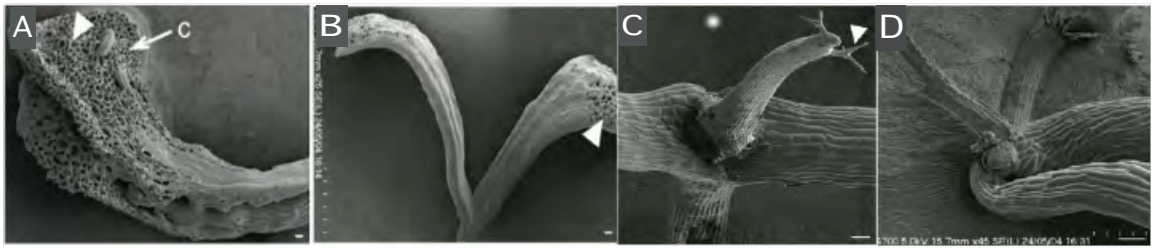


Figure 3 Phenotypic analysis of AtDek1-RNAi knockdown seedlings. A: AtDEK1RNAi seedlings showing severe developmental defects including fusion of the cotyledons (c) and mesophyll-like cells replacing the epidermis (arrowhead). B: AtDEK1-RNAi seedling with no fusion of cotyledons and zones of 'lacy' epidermis on the adaxial surface (arrowhead). C and D: Less severe AtDEK1-RNAi knockdowns produce central organs such as radialized spikes, sometimes bearing trichomes (arrowhead in C). (Modified figure form Johnson et al., 2005, scale bar 100 μm).

The transgenic RNAi *Arabidopsis* seedlings formed cotyledons with mesophyll-like cells at the lamina surface, and in the most severe lines the cotyledons were fused together (Figure 3). These results implied that AtDEK1, which is required for the initial differentiation of protodermal cell fate are also required to maintain the observation of positional signaling, continuously stimulating epidermal identity during late embryogenesis (Johnson et al., 2005).

1.1.5 DEK1 is growth regulator and plays global role in plant development

The observations of the AtRNAi phenotypes where mesophyll-like tissue occurred on the organs and stems of plants showing low levels of AtDEK1 activity suggested that when levels of AtDEK1 activity are limited, epidermal growth can no longer keep up with that of underlying tissues, leading to a loss of epidermal continuity, and thus the ability to maintain epidermal identity. The role of AtDEK1 in maintaining epidermal cell fate could be an indirect consequence of a fundamental role in coordinating growth (Johnson et al., 2008). Virus-induced silencing of the tobacco Nb*DEK1* gene in *Nicotiana benthamiana* causes arrested organ development and hyperplasia in all major organs examined. These defects include hyper proliferation of leaf and stem epidermis and reduced stomata and trichome development, inferring that DEK1 plays a global role in plant development (Ahn et al., 2004).

1.1.6 DEK1 responses to positional cues

An internalized daughter cell resulting from a periclinally divided aleurone cell loses its identity and differentiates to a starchy endosperm cell in maize and starchy endosperm cells may convert to aleurone cells if they become positioned on the surface showing that the fate of endosperm cells maintain flexible until late in their development. Dek1 mutant cells are unable to perceive or respond to the positional cues that normally specify aleurone identity, showing that the function of DEK1 seems to be required for this surface-dependent response throughout endosperm development (Becraft & Asuncion-Crabb, 2000; Becraft et al., 2002; Lid et al., 2002).

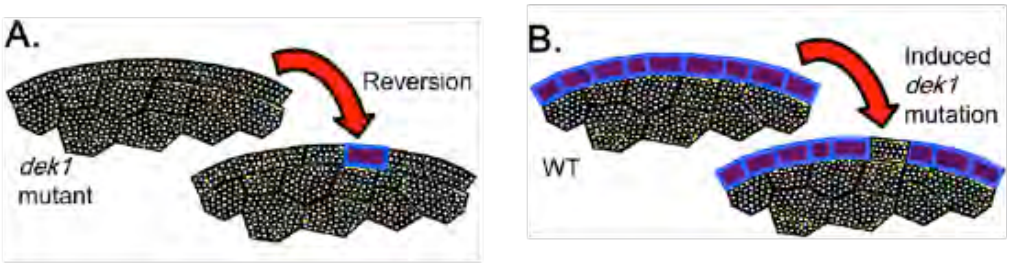


Figure 4 Developmental plasticity of endosperm cells. (A) Reversion of a dek1 mutant cell to wild type results in the transdifferentiation of the peripheral cell from starchy endosperm to aleurone. (B) Induction of a dek1 mutant cell in a wild-type background results in the transdifferentiation of aleurone to starchy endosperm. These transdifferentiation events can occur even late in development, illustrating that the cues that specify aleurone cell fate are present throughout development and required to maintain aleurone identity (modified figure from Becraft & Yi, 2010).

Transdifferentiation was similarly observed in cultured endosperms when fissures developed and starchy endosperm cells newly exposed to a surface became aleurone (Gruis et al., 2006). A maize mutant produces an endosperm composed of multiple spheroid masses of cells, each with a layer of aleurone cells, since some of these aleurone cells were not in contact with maternal tissues, proposing that surface position induces aleurone identity (Olsen, 2004). This hypothesis is supported by recent reports of aleurone differentiation on isolated endosperms grown in vitro in the absence of maternal tissues (Gruis et al., 2006; Reyes et al., 2010). The function of DEK1 has also been implicated in epidermal cell fate specification in *Arabidopsis* embryos, leaves, and meristems (Johnson et al., 2005; Lid et al., 2005), suggesting that the mechanisms involved in surface position recognition described by Gruis et al. (2006) for endosperms may be universal to epidermal cell formation in all plants.

In situ hybridization experiments show that *DEK1* is ubiquitously expressed, and could have the flexibility to respond to surface position upon cell repositioning (Lid et al., 2005; Gruis et al., 2006). Analysis of *DEK1* expression by massively parallel signature sequencing on microbead arrays from a total of 37 maize and 11 *Arabidopsis* tissues, showed that the *DEK1* gene is expressed in most maize tissues at a level ranging from 30 to 55 ppm (parts per million). In *Arabidopsis thaliana*, a low level of the At*DEK1* transcript occurred in young and germinating seeds (10 ppm), but the transcript was undetected in older seeds and seedlings. The highest level of the At*DEK1* transcript was observed in seedling roots (67 ppm). Examination of *Arabidopsis* embryo and endosperm showed that the *DEK1* gene is

expressed evenly throughout early developmental stages, declining as the seeds mature, and is present in all vegetative plant organs examined. Together, these results suggest that DEK1 functions in diverse developing tissues (Lid et al., 2002; Lid et al., 2005). In addition these results demonstrate that the expression of *DEK1* is not restricted to the outermost cell layer (L1), even though revertant sector analysis revealed that DEK1 functions cell-autonomously because wild-type cells cannot rescue the phenotype of adjacent dek1 mutant cells, and dek1 mutant cells cannot impose their phenotype onto adjacent wild-type cells (Becraft et al., 2002).

1.2 AtDEK1-MEM and the Dominant Negative Phenotype

Dominant mutations have had important consequences in evolution and are especially attractive to geneticists, as they are recognized in the T1 generation, thus simplifying phenotypic characterization. Traditional approaches for characterizing protein function have significant limitations especially in identifying the roles specific proteins play *in vivo*. An alternative approach is to engineer mutations in the protein of interest that abolish its function and that also inhibit the function of simultaneously expressed wild-type protein. Those mutations, whose gene products act antagonistically to the normal, wild-type gene products within the same cell are called dominant negative mutations or antimorphic mutations (Herskowitz, 1987). The dominant negative phenotype usually occurs if the mutant product can still interact with the same elements as the wild-type product, but block some aspect of its function. An example can be a mutation in a transcription factor that removes the activation domain, but still contains DNA binding domain, this product can block the wild-type transcriptional factor from binding the DNA site leading to reduced levels of gene activation. Dominant negative mutants have already provided insights into the molecular mechanisms of a number of protein families, including hormone receptors, oncogenes, and growth factor receptors, and have been identified as the cause of at least a few autosomal dominant diseases. Expression of dominant negative mutants under the control of highly active promoters holds great promise for the study of the roles specific proteins and protein families play in development, health, and disease (Sheppard, 1994).

Over-expression of AtDEK1-MEM under the control of the strong Cauliflower Mosaic Virus (CaMV35S) promoter gives a spectrum of dominant negative phenotypes, ranging from

plants with defective shoot apical meristems unable to produce adult leaves, to cotyledons with disorganized or no epidermal cells, to plants producing radialized rosette leaves in *Arabidopsis thaliana.* In order to provide a baseline for AtDEK1-MEM overexpression phenotypes, Tian et al (2007) generated Dek1 knockdown phenotypes, by expressing an AtDEK1-RNA-interference (RNAi) construct under the control of the CaMV35S promoter. The observed phenotypes closely resembled the AtDEK1-MEM lines, and were also compatible with the phenotypes observed previously in a similar experiment by Johnson et al. (2005).

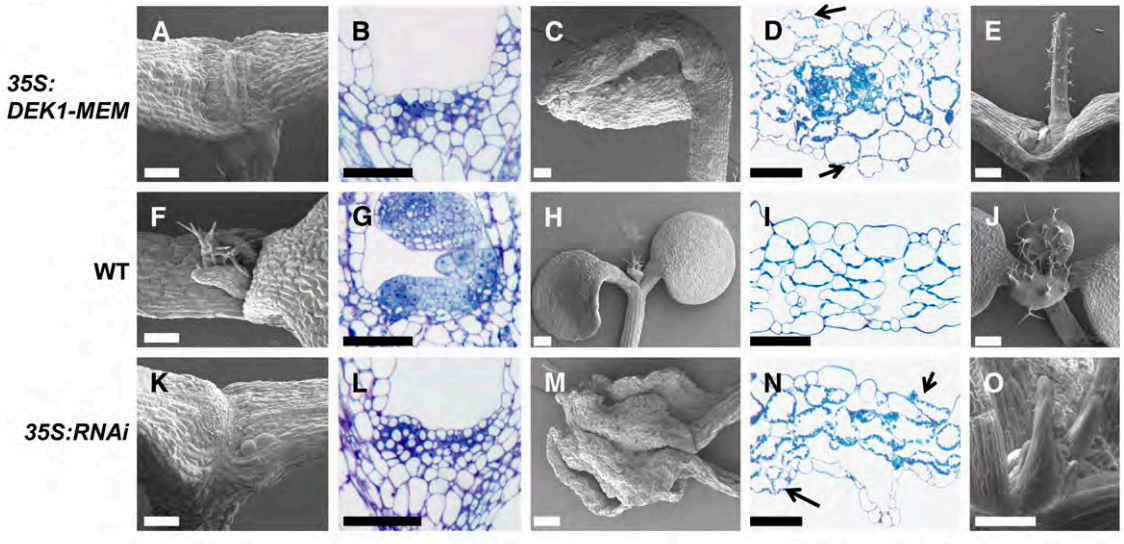


Figure 5 Phenotypic analysis and comparison of Arabidopsis lines overexpressing AtDEK1-MEM and AtDEK1-RNAi. From A to E Phenotypes of At DEK1-MEM seedlings, from F to J Comparable developmental stages and organs in wild-type seedlings, and from K to O Comparable developmental stages and organs in At DEK1-RNAi seedlings. The shoot apex is severely affected in At DEK1-MEM plants (A and B) and severely affected in RNAi lines (K and L). Cotyledons from At DEK1- MEM (C) and At DEK1-RNAi (M) seedlings are also distorted. Note in the cross sections of cotyledons from At DEK1-MEM (D) and At DEK1-RNAi (N) plants that the cells exposed to both adaxial and abaxial surfaces contain chloroplasts, suggesting that they have not completely differentiated into epidermal cells. Shoot apices from less severely affected seedlings in both At DEK1-MEM and At DEK1-RNAi lines can display one terminal radialized rosette leaf (E) or several radial rosette leaves (O). Bars 100 μm (Modified figure from Tian et al., 2007).

An explanation for the dominant negative phenotype caused by AtDEK1-MEM overexpression could be that AtDEK1-MEM interacts with either a ligand or interacting membrane proteins that normally mediate the activation of the resident AtDEK1. The overexpressed AtDEK1-MEM truncated protein may compete effectively with native Dek1 for the activators and/or interactors, and since it lacks the cytoplasmic calpain domain, it cannot trigger the downstream signaling pathways, leading to a dominant negative

phenotype. Plants expressing AtDEK1-MEM under the weaker AtDEK1 promoter produced a phenotype indistinguishable form wild type suggesting that there is a threshold level of truncated protein that has to be surpassed to be able to interfere with endogenous DEK1 function. Wild type phenotype was also observed in plants overexpressing AtDEK1-MEM-DEL, which lacks the extracellular loop region, supporting the speculation that the loop region is essential for either the perception and/or the transmission of positional signals. The *AtDEK1-MEM-DEL* construct was designed in order to rule out the possibility that the AtDEK1-MEM phenotype represents a pleiotropic effect caused by ectopic expression of a large membrane protein and also led to the conclusion that the observed dominant negative phenotype is specific to AtDEK1- MEM (Tian et al., 2007).

1.3 Proposed models for DEK1 function

Investigation of cell fate specification and development in maize, cereal grains, *Arabidopsis* and rice have identified three genes involved in cell-fate specification of aleurone cells and/or epidermal cells; Crinkly4 (Cr4), encoding a receptor-like protein kinase that shares some sequence similarity with the tumor necrosis factor receptor (TNFR), Supernumerary aleurone layers 1 (Sal1) encoding a maize homolog of the human Chmp1 gene implicated in endosome trafficking, and Defective kernel 1 (Dek1) (Becraft et al., 1996 & 2002; Sheridan & Neuffer, 1980; Howard et al., 2001; Shen et al., 2003). Defects in DEK1 or in or in CRINKLY4 (CR4), prevent the differentiation of an aleurone layer and lead to the presence of starchy endosperm cells in peripheral positions (Becraft et al., 1996; Becraft & Asuncion-Crabb, 2000). Phenotypic analysis of the dek1／cr4 double mutant suggests that the two gene products function in partially overlapping pathways, as strong *dek1* alleles were epistatic to cr4 (Becraft et al., 2002). These findings led to the initial model proposing a role for DEK1 in the release of signals that are perceived by receptor-like kinase CR4, because maize cr4 mutants share some of the dek1 phenotypes, and cr4/dek1 double mutants show dek1 phenotypes (Becraft et al., 2000). The cr4 homolog in *Arabidopsis*, acr4, and the acr4/dek1 double mutants show additive effects, which suggests that dek1 and acr4 act in different pathways of epidermis specification (Johnson et al 2005).

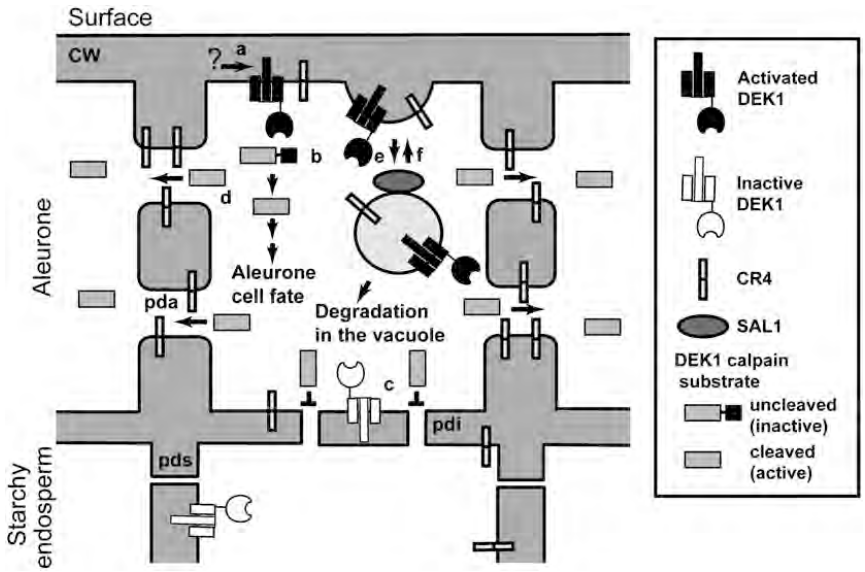


Figure 6 A Model for the Role of DEK1, CR4, and SAL1 in Aleurone Cell Specification. DEK1 at the surface of the endosperm is activated by an unknown mechanism (a), its calpain domain in the cytosol cleaving a postulated substrate (b) that leads to the specification of aleurone cell fate. DEK1 in all other positions is inactive (c). In cells with active DEK1 signaling, CR4 concentrates on plasmodesmata between aleurone cells (pda) and increases the plasmodesma exclusion limit, allowing the activated DEK1 substrate to move laterally between aleurone cells, thereby reinforcing the signal for aleurone cell fate specification (d). Plasmodesmata in cell walls between starchy endosperm cells are narrow (pds), whereas plasmodesmata in cell walls between aleurone cells and starchy endosperm cells are intermediary in width (pdi). DEK1 and CR4 are internalized by endocytosis (e) and traffic through endosomes. Whereas some DEK1 and CR4 molecules may be recycled back to the plasma membrane (f), others are sorted for degradation in the vacuole in a process that requires SAL1. Some endosomes are recycled back to the plasma membrane (f). Figure and text are cited from Tian et al. (2007).

In another model, DEK1 is proposed to cleave homeodomain–leucine zipper IV (HDZipIV) transcription factors, which regulate epidermal cell fate (Johnson et al 2005). This model is consistent with the cell-autonomous function of DEK1 and the fact that DEK1 carries nuclear targeting signals (Ahn et al 2004). At the cellular level, cell divisions are stimulated in already differentiated tissues, interfering with normal tissue differentiation. Based on the observation that protein levels of D-type cyclins, as well as transcript levels of retinoblastoma, transcription factor E2F, S-phase genes and KNOTTED1-type homeobox genes were elevated in DEK1 silenced tissue, the authors proposed that DEK1 regulates cell proliferation and differentiation during plant organogenesis, and that it acts partly by controlling the CycD/Rb pathway (Ahn et al., 2004).

Johnson et al (2008) analyzed gene expression during early leaf development of proteins known to be involved in growth regulatory pathways. The results suggested that DEK1 plays

a key role in plant growth regulation by acting downstream for the CYP78a7 gene, and for GA2Ox8 gene. The diversity of genes with changed expression due to overexpression or suppression of the CALPAIN implies that DEK1 is likely to act through a number of developmental pathways to regulate growth (Johnson et al., 2008, and references therein).

Due to the severity of mutant phenotypes, the role of DEK1 in plant development remains unclear. Null alleles of dek1 cause early embryo lethality, making this phenotype useless in making conclusions about the gene function in plant development. The RNA interference approach solved this problem, and the results showed that DEK1 is required for the maintenance of epidermal cell fate during late embryogenesis (Johnson et al., 2005).

A number of maize *dek1* alleles with specific phenotypes have been described in detail (Becraft et al., 2002). Embryos homozygous for some of these alleles develop into viable, but abnormal plants with crinkled leaves, shortened internodes and nodes that bend alternately back and forth, mutant dek1 kernels often do not germinate due to the arrest of embryo development at an early stage (Becraft et al., 2002). *Arabidopsis* dek1 mutant phenotypes are similar to maize; including lack of the aleurone layer and arrest of embryo development. Studies of dek1 mutant and transgenic phenotypes in maize, *Arabidopsis thaliana* and in tobacco have demonstrated the importance of both DEK1-MEM and DEK1-CALP in determining epidermal cell fate and are summarized in the table below (Ahl et al., 2004; Johnson et al., 2005 & 2008; Lid et al., 2002 & 2005; Tian et al., 2007).

Table 1 Summary of known and published facts about DEK1 (modified table from van der Hoorn, 2008)

Gene name	***DEK1* (maize)**	***AtDEK1* (Arabidopsis)**	***NbDEK1* (tobacco)**
Described alleles	*dek1-1…12* (Becraft et al., 2002)	*dek1- 1…4* (Johnson et al., 2005; Lid et al., 2005)	–
Knockout phenotypes	Embryo lethal, aleurone deficient (Becraft & Asuncion-Crabb, 2000; Becraft et al., 2002; Lid et al., 2002)	Embryo lethal, aleurone deficient (Johnson et al., 2005; Lid et al., 2005)	–
Knockdown phenotypes	–	Deformed plants, lack of epidermis (Johnson et al., 2005)	Callus formation on all surfaces (Ahn et al., 2004)
Overexpression phenotype	–	Loss of trichomes (only in specific ecotype), different epidermal cell shape and organization (Lid et al., 2005)	–
Endogenous expression	Ubiquitous at low levels (Lid et al., 2002; Wang et al., 2003)	Ubiquitous at low levels (Lid et al., 2005)	Ubiquitous at low levels (Ahn et al., 2004)
Localization	Plasma membrane, endogenous system (Tian et al., 2007)	Plasma membrane, endoplasmic reticulum and cytoplasm (Johnson et al., 2008)	Nuclear membrane (Ahn et al., 2004)
Putative genetic interactions	Function of receptor-like kinase CR4 depends on Dek1 (Becraft et al., 2002)	Receptor-like kinase ACR4 acts independent of Dek1 (Lid et al., 2005)	–
Proteolytic activity	Domains II and III cleave casein in vitro, stimulated by Ca^{2+} (Wang et al., 2003)	–	–
Proposed mechanisms	May cleave transcription factors in response to signals from the extracellular environment (Johnson et al., 2005 & 2008; Lid et al., 2002 & 2005; Tian et al., 2007)		

The table above summarizes described alleles, knockdown, knockout and overexpression phenotypes, expression levels, proteolytic activity, putative genetic interactions and proposed mechanisms of DEK1 studied and published so far.

1.4 Homology and Evolution of DEK1

DEK1 is highly conserved in higher plants. Sequencing data from both angiosperms and gymnosperms demonstrate a significant degree of sequence conservation. Comparison between the maize and *Arabidopsis* orthologs reveal 70 % overall identity, with highest score of 86% in the calpain domain, about 70 % in the membrane anchored domains, and lowest conservation (57 %) in the extracellular loop domain (Lid et al., 2002).

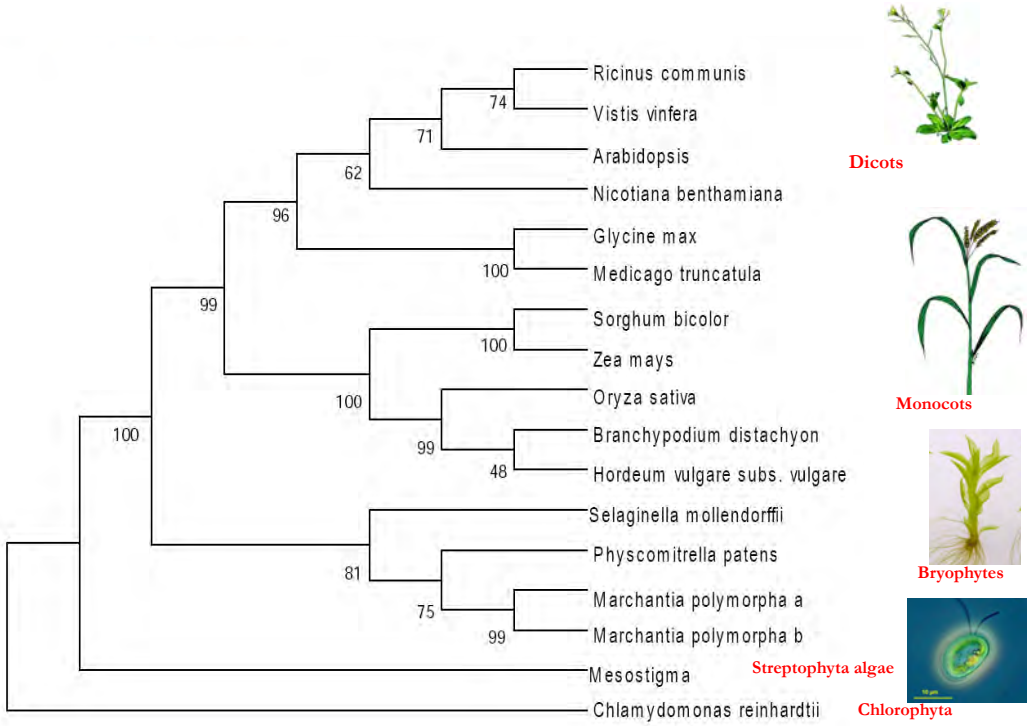


Figure 7 Phylogenetic tree based on DEK1 sequence (Modified figure from Liang, 2010, unpublished)

Johnson et al., (2008) defined DEK1 as a plant-specific phytocalpain, however *DEK1* was recently identified in the single celled green algae, *Mesostigma viride*, probably containing two or more *DEK1* genes (Johansen and Wilson 2010, unpublished results), and in the flagellated protozoan *Thecamonas trahens*, indicating that *DEK1* is not plant-specific and originated long before multicellular organisms. Newly, DEK1 was also identified in the moss *Physcomitrella patens*, separated from angiosperms by ~ 500 million years of evolution (Quatrano et al., 2007) and in the liverwort *Marchantia polymorpha* (the Kohchi lab at Kyoto University, unpublished results). Amino acid sequence comparison of DEK1 from moss and maize revealed 59 % identity and even higher identity (80 %) in the calpain domain (Tian et at. 2007).

1.5 *Arabidopsis thaliana* as model organism

Arabidopsis thaliana was discovered in the Harz Mountains, in the sixteenth century by Johannes Thal, at this time he called the plant *Pilosella siliquosa*, and the name has gone

through a number of changes since. The earliest appearance of mutant *Arabidopsis* in the scientific literature appears to be in 1873 by Alexander Broun. In 1943 F. Laibach was the first to summarize the potential of *Arabidopsis thaliana* as a model organism for genetics (Meyerowitz, 2001).

Figure 8 *Arabidopsis thaliana*. Picture to the right shoes the vegetative stage, before flowering and growth of the floral stalk (bottom left). On the centre an adult plant at full flowering/seed set. On the right, flower, floral stem and seeds. White bars represent 1 cm, except for flower and seeds: 1 mm. (Picture borrowed from http://www-ijpb.versailles.inra.fr/en/arabido/arabido.htm)

Arabidopsis thaliana, commonly known as thale-cress or mouse-ear cress, is a small dicot, angiosperm (a flowering plant) member of the mustard family (*Brassicaceae*). *Arabidopsis* may not have been of agronomic importance like its relatives cabbage and radish, for instance, but it has been widely used as model organism because of the many advantages this small plant offers to basic research in genetics and molecular biology.

Some of the significant advantages *Arabidopsis* offers as model organism is the small genome (one of the smallest in the plant kingdom) containing 115 409 949 base pairs of DNA distributed in five chromosomes. Very little of its genome is “junk DNA” which means most of the DNA collectively encodes an estimate of 25 498 genes. The short life cycle of six weeks and its prolific seed production of about 10 000 per plant make genetic studies easier. The easy cultivation and its small size of about 25 cm in high and about 5 cm in diameter make it a perfect laboratory tool and mutants are easily achieved by efficient transformation with *Agrobacterium tumefaciens* (reviewed by Meinke et al., 1998; TAIR, Sept. 2010).

The aims of this study

The aim of this study was to further investigate and understand the function of the essential plant protein Defective Kernel 1 with a focus on the extracellular loop domain of the protein. Based on the dominant negative effect caused by the loop domain in transgenic AtDEK1-MEM *Arabidopsis thaliana* plants (Tian et al., 2007) and by the creation of chimeric genes with swapped loop domain from different species, we tried to assess the functional significance of conserved residues in the AtDEK1-LOOP domain. The research question was whether the DEK1-LOOP sequences from *Zea mays*, *Physcomitrella patens* and *Mesostigma viride* could functionally replace the *Arabidopsis thaliana* DEK1-LOOP. The answer to that question should give an idea of how well this domain is conserved and whether it participates in the same signaling processes in these species separated by evolution by millions of years. A sub-aim of the study was to further investigate the pleiotropic nature of the dominant negative phenotype.

2. Materials and Methods

All laboratory procedures were performed in the laboratory facilities in Biohus, Hedmark University College, Hamar in strict accordance with the Norwegian Gene Technology Act (LOV 1993-04-02 nr 38) for production, use and application of genetically modified organisms.

2.1 Bioinformatics

2.1.1 BLAST search and LOOP identification

Sequence similarity searches were performed on the National Center for Biotechnology Information (NCBI) databases using the Basic Local Alignment Search Tool (BLAST). Based on annotations of the *Zea mays* (Zm) DEK1, the homologs LOOP regions of DEK1 in *Arabidopsis thaliana* (At) and *Physcomitrella patens* (Pp) were identified. The putative *Mesostigma viride* (Mv) DEK1-LOOP sequence was identified using a 5'-RACE cDNA cloned in our laboratory as query. The Mv sequence was submitted to the TMHMM Server v.2.0 (Prediction of transmembrane helices in proteins, at www.cbs.dtu.dk/services/TMHMM/) to identify the Mv-LOOP region. The data collected from BLAST was also used to create multiple sequence alignment (MSA) and to compare the identity percentage between the identified putative loop sequences.

2.1.2 *In silico* PCR, restriction digestion, cloning and sequence analysis

PCR primers and products were respectively designed and simulated using CLC Genomics Workbench v4.5.1. The chimeric constructs were submitted to FGENESH, HMM-based gene structure prediction server at SoftBerry.com and the predicted gene further analyzed in TMHMM Server v.2.0. to show the transmembrane forecast. The designed chimeric PCR products were cloned *in silico* in the pCR-Zero-Blunt vector (Invitrogen), enzyme restriction was simulated and further the products were cloned *in silico* in the binary vector pSEL1:35S-

AMG (Tian et al., 2007). The sequence outputs generated from sequencing reactions were analyzed by assembling the examined sequences to reference DNA in CLC Genomic Bench v4.5.1.

2.1.3 Protein structure prediction and alignments

The amino acid sequences of the Zm-, Pp- and Mv-DEK1-LOOPs were submitted to different structure, function and similarity prediction servers like PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred/), Predict Protein (http://www.predictprotein.org/), Prosite (http://prosite.expasy.org/) and others. The Phyre (Protein Homology/analogY Recognition Engine) v 0.2 engine was used to search for similarities in three dimensional structures with known function that may indicate similar functions, and further submitted to the 3DLigandSite (predicting ligand-binding sites using similar structures) as a follow up analysis of the 3D structure alignments. Search for possible substrates was performed at Calpain for Modulatory Proteolysis site (CaMPDB, http://calpain.org/).

2.2 Creation of the chimeric *AtDEK1-MEM-LOOP*-replaced construct

The lack of compatible restriction sites on the ends of the loop coding regions of the *DEK1* orthologs in the species used in this study led to the use of Overlapping Expression PCR approach (Wurch et al., 1998) for the creation of the interspecies domain swapped chimera genes. In these chimeras the sequence of At*DEK1*-extracellular-loop coding region in the AtDEK1-MEM construct was replaced with *DEK1*-loop coding regions from *Zea mays* (Zm), *Physcomitrella patens* (Pp) and *Mesostigma viride* (Mv). The strategy is represented schematically in Figure 9.

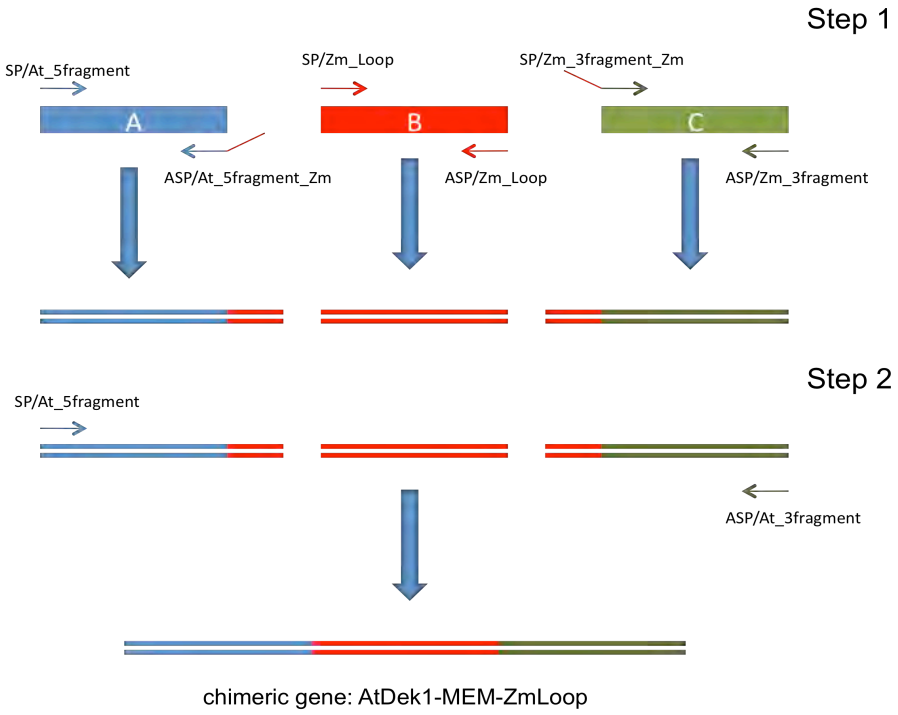


Figure 9 Strategy for creating chimera genes by replacing *AtDek1-MEM-Loop* with Zm-, Pp-, and Mv-Loops. In the first step, fragments A (5'-end of AtDek1-MEM), B (Loop part of Zm) and C (3'-end of AtDek1-MEM) are PCR-amplified separately with the indicated primers. Primers SP/At_5'-fragment, ASP/At_3'-fragmet, SP/Zm_Loop and ASP/Zm_Loop are fully identical to their respective template. Primers ASP/At_5'-fragment_Zm and SP/At_3'-fragment_Zm possess 5'-extentions (indicated in red), complementary to the 5'-end and the 3'-end, respectively, of fragment B. In the second step all three fragments are fused together in Overlapping Expression PCR reaction with the indicated primers. The product is a chimeric gene where the Loop part of AtDek1-MEM is replaced or swapped with the loop from Zm (*Zea mays*), Pp (*Physcomitrella patens*) and Mv (*Mesostigma viride*).

2.2.1 Template DNA isolation

Templates for PCR amplification of the *DEK1-LOOP* sequences (fragment B, Fig. 9) were obtained as follows: Total genomic DNA from *Physcomitrella patens* was isolated and purified with DNeasy 96 Blood & Tissue Kit (QIAGEN) following the "Purification of Total DNA from Animal Tissues, Spin-Column Protocol". A 5'-cDNA fragment containing the entire putative Loop cDNA from *Mesostigma viride* was previously cloned in our laboratory (W. Johansen, unpublished results) and used for the amplification of the *MvDEK1-LOOP*. As template for *ZmDEK1-LOOP* amplification was used a plasmid clone containing part of the genomic *Zm-DEK1* gene, harboring the *ZmDEK1-LOOP* sequence (kindly provided by UMB, Ås). The template DNA used for amplification of fragments A and C (Figure 9) was the purified plasmid pSEL1:35S-AMG (Tian et al., 2007). Primers were designed in Vector

NTI Advance™ Software (Invitrogen), indicated in the Appendix 1 (Table 1 A1) together with the PCR conditions used.

2.2.2 PCR amplification of the construct fragments

The construct fragments (Figure 9) were amplified using a High-Fidelity PCR Kit (FINNZYMES). The fragments were amplified separately in a reaction volume of 20 µl containing the following components: 1x Phusion HF Buffer (Finnzymes), 200 µM of each dNTPs, 0.5 µM Sense and Antisense primers (Appendix 1, Table 1 A1) and 0.4 U/µl Phusion Polymerase (Finnzymes). 5 ng plasmid DNA or 100 ng genomic DNA was used as templates. The cycling conditions used in the different PCR reactions are included in Appendix 1 Table 1A, along with the primer sequences. PCR products were purified using the QIAquick PCR-Purification Kit (QIAGEN).

2.2.3 Overlapping Expression PCR

The PCR amplification of the chimeric sequences was performed as described in 2.2.2. The primers used for amplification were: SP/At-5'-fragment (5'-ATACTG**AGCGCT**GAAATGTTCTCATTC-3') and ASP/At-3'-fragment (5'-GTGGGCAACTGATCATCTCTAGATTTTA-3') introducing a unique 5' *Eco*47III (*Afe*I) restriction site. The amplified PCR products contain an internal 3' *Xba*I restriction site. As template for the reactions each of the PCR-amplified fragments; A, B and C (Figure 9) were used. The cycling conditions for the PCR reactions consisted of initial denaturation at 98 °C for 30 sec., 30 repetitive cycles of 98 °C for 10 sec., 58 °C for 15 sec., and 72 °C for 60 sec., and a final elongation step at 72 °C for 10 min. PCR products were purified on 0.7 % low-melting agarose gel electrophoresis (80 V/h).

2.2.4 Cloning and verification of the chimeric genes

The purified over-lapping PCR products were cloned into vector pCR-Blunt using the PCR-Zero Blunt PCR Cloning kit (Invitrogen) according to the manufacturers instructions. One

Shot TOP10 Chemically Competent *E.coli* cells (Invitrogen) were transformed by heat shock according to manufacturers protocol and selected on Luria Bertani (LB) plates (1 % tryptone, 0.5 % yeast extract, 0.17 M NaCl and 1.5 % agar) supplemented with 50 µg/ml kanamycin. Plates were inverted and incubated at 37 °C over night.

Positive clones were selected by colony PCR using AmpliTaq Gold Polymerase (Applied Biosystems) and the appropriate primers (Appendix 1, Table 1 A1). Plasmid DNA from the selected colonies was isolated from overnight cultures using Pure Yield™ Plasmid Miniprep System Kit (Promega). The purified plasmid was verified by restriction digestion with *Eco*RI restriction enzyme (BioLabs, New England), which has restriction sites on the pCR-Blunt vector just upstream and downstream of the insert. Confirmed plasmid DNA was sequenced using the BigDye® Terminator v3.1 Cycle Sequencing Kit (Applied Biosystems) with overlapping sequencing primers (Appendix 1, Table 2 A1) and sequenced in sequencing machine Applied Biosystems 3130xl Genetic Analyzer (Applied Biosystems) to verify sequence composition.

2.2.5 Mutation of the *Xba*I restriction site on the Zm-loop

The *ZmDEK1-LOOP* sequence contains an internal *Xba*I restriction site that was removed prior to ligation into the final binary vector using the Gene Tailor™ Site-Directed Mutagenesis System Kit (Ivitrogen) with primers ZmLoopMutF and ZmLoopMutR (Appendix 1, Table 3 A1). The methylation reaction contained 1.6 µl 10x SAM, 1.6 µl Methylation Buffer, 100 ng pCR-AtMEM-ZmLoop plasmid DNA, and 1 µl 4 U/µl DNA Methylase in a final volume of 16 µl and was incubated at 37 °C for 1 hour. The mutagenesis reaction using Phusion Hot Start Polymerase contained 1x Phusion HF Buffer, 200 µM dNTPs, 0.5 µM SP/Zm-loop-MutaF, 0.5 µM ASP/Zm-loop-MutaR, 2 µl Metilation reaction (~ 12.5 ng methylated plasmid DNA), and 0.02 U/µl Phusion Hot Start II Polymerase in final reaction volume of 50 µl. PCR conditions: 98 °C for 30 sec., 25 cycles of 98 °C for 10 sec., 65 °C for 10 sec., 72 °C for 2 min., and final elongation step at 72 °C for 10 min. The site-mutated pCR-AtMEM-ZmLoop plasmid was checked on 1 % agarose gel electrophoresis and transformed into DHα-T1^{R} *E.coli* cells (Invitrogen) following manufacturer protocol. Selection, plasmid DNA isolation and sequencing was performed as in 2.2.4.

2.3 Cloning of the chimeric gene construct into pSEL1:35S-AMG vector

2.3.1 Restriction digestion

The pCR-plasmids containing the chimeric constructs and the pSEL1:35S-AMG plasmid were digested in separate reactions with *Xba*I and *Pst*I restriction enzymes (NewEngland, BioLabs). The restriction reactions contained 1 µl NEB 3 Buffer, 1 µl 10x BSA, 2 µl plasmid (~ 300 ng/ml), 0.5 µl *Xba*I and 0.5 µl *Pst*I in final volume of 10 µl. The reaction was incubated at 37 °C for 2 h. The desired fragments were separated on 0.7% low-melting agarose gel electrophoresis.

2.3.2 Ligation reaction

The ligation reaction contained 1µl Buffer for T4 DNA Ligase (NewEngland, BioLabs), 1 µl agarose bit containing *Xba*I and *Pst*I digested pSEL1:35S-AMG, 1 µl agarose bit containing digested chimeric construct (both incubated at 60 °C for 10 min.) and 1 µl of T4 DNA Ligase (NewEngland, BioLabs) in final volume of 10 µl. The reaction was incubated at room temperature for 20 min and heat inactivated at 65 °C for 10 min.

Chemically competent Top10 *E.coli* cells (Invitrogen) were transformed with 2 µl Ligation reaction and putative transformants were selected on LB-medium agar plates supplemented with 50 µl/ml kanamycin. Positive colonies were confirmed by Colony PCR, the pSEL1:35S-chimeric-gene plasmid was isolated and analyzed by restriction reaction on gel electrophoresis, and then sequenced as described earlier (2.2.4).

2.4 *Agrobacterium* mediated *Arabidopsis* transformation

2.4.1 Transformation of *Agrobacterium* by electroporation

Preparing Electrocompetent *Agrobacterium* cells: 1 ml of an over night culture of *Agrobacterium tumefaciens* C58 pGV2330 (Deblaere et al, 1985) was used to inoculate 100 ml of YEB-medium (5 % Meat Extract, 1 % Soytone Peptone, 5 % Yeast extract, 1 % Sucrose and 2 mM $MgCl_2$, pH 7.2) supplemented with 100 µg/ml Carbenicillin. The cells were incubated at 28 °C shaking 200 rpm to a density (OD_{600}) of ~ 0.4. Then the cells were harvested in sterile centrifuge tube and pelleted at 5000x g at 4 °C for 15 minutes. The pellet was washed three times with 500, 250 and 100 ml sterile ice-cold water and then two times with 50 ml sterile 10 % ice-cold glycerol. The pellet was finally resuspended in 3 ml 10 % ice-cold glycerol. Electrocompetent cell suspension was aliquoted and stored at - 85 °C.

Electroporation: 50 µl of electrocompetent *Agrobacterium* cells were gently mixed with 2 µl (~ 50 ng/µl) of the binary plasmid pSEL1:35S-AMG-NOS carrying the construct of interest in 1 mm gap electroporation cuvette (BTX, Harvard Apparatus) on ice. The cells were electroporated (Eppendorf, Electroporator 2510) at pulse 1250 V (1.25 kV/mm) and then 1 ml of SOC-medium (2 % bacto tryptone, 0.5 % bacto yeast extract, 100 mM NaCl, 2.5 M KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 % glucose) was immediately added, the cells transferred to sterile Falcon tubes and incubated at 28 °C, 200 rpm for 1 hour. Putative transformants were selected on YEB agar plates supplemented with 50 µg/ml rifampicillin, 100 µg/ml carbenicillin and 50 µg/ml kanamycin, incubated at 28 °C for 2-3 days. Colonies were analyzed using colony PCR with primers specific to the cloned gene (Appendix 1, Table 1 A1) to verify the presence of desired the constructs. Positive colonies were inoculated into 50 ml LB-medium supplemented with 100 µg/ml carbenicillin and 50 µg/ml kanamycin and incubated at 28 °C for two days (to a stationary phase, OD_{550} ~ 1.3) on rotary shaker at 200 rpm.

2.4.2 Transformation of *Arabidopsis thaliana*

Wild type *Arabidopsis thaliana* ecotype Col-0 plants were transformed by direct drop-by-drop inoculation to every flower by using a micropipette (Martínez-Trujillo et al., 2004) a

modified version of the floral-dip method (Clough & Bent 1998). Bolts of flowering healthy wt *Arabidopsis* plants were clipped once to encourage proliferation of many secondary bolts before transformation. 40 ml transformed *Agrobacterium* cells (from overnight culture) were centrifuged at 3000x g for 15 min. at room temperature. The cell pellet was resuspended in 45 ml Infiltration medium (5 % sucrose, 0.5 x M&S-medium, pH 5.8) to an OD_{600} ~ 0.8. Prior to transformation 0.05 % Silwert L-77 was added to cell suspention.

Co-transformation was performed by the same method; two *Agrobacterum* cultures at equal concentrations (OD_{600} ~ 0.8), each carrying different plasmid; pSEL1:35S-AMG (Tian et al., 2007) and pRPS5A:CALPAIN-GFP (Johnson et al., 2008) were mixed just before transformation (Buck et al., 2009).

2.4.3 Plant material, growth conditions and selection of transformed plants

Seeds from transformed and wild type *Arabidopsis thaliana* (provided by UMB, Ås, Norway) were surface sterilized as follows: about 1 mg (~ 50 seed) volume of seeds were first treated with 1 ml 75 % ethanol for 5 minutes (on inverter), and then with 1 ml 10 % chloride solution for 20 minutes, followed by five rinses with sterile water.

Surface-sterilized seeds were resuspended in approximately 5 ml 0.75 % Phyto Agar (Duchefa Biochemie) solution before sowing on 1x M&S (Murashige & Skoog, 1962) plates containing M&S Medium (Duchefa Biochemie), 1 % sucrose and 0.05 % MES, pH adjusted to 5.7, and 0.8 % Phyto Agar. Selective plates were supplemented with 50 μg/ml kanamycin (and 20 μg/ml hygromycin for selection of co-transformants). Plates were placed in dark and cold (4 °C) for 2 days (stratification period). Germination and cultivation occurred in a growth room at 22 °C with light intensity (85 μmol m^{-2} s^{-1}), 60-70 % humidity and 18 h light and 6 h dark period. With each selection of putative transgenic plants, positive and negative control were also sown: wild type on non-selective medium and wild type on selective medium respectively. Seedlings were phonotypically characterized and transferred from selective medium to sterilized soil by the age of two weeks.

2.5 Analyzing transgenic lines

2.5.1 Identification of transgenic lines

Transgene insertion in kanamycin resistant plants was verified by PCR using the Extract-N-Amp^{TM} Plant Kit (SIGMA) with primers SP/pSEL-AtMEM-GT and ASP/pSEL-AtMEM-GT (oligo sequences are included in Appendix 1, Table 4 A1) according to the manufacturer's instructions. DNA extracted from wild type plants was used as negative control. The PCR reaction contained 10 µl Extract-N-Amp Ready Mix (SIGMA), 0.4 µM form each forward and reverse primers, and 4 µl Plant extract (DNA template) in a total volume of 20 µl. PCR conditions were as follow: denaturation step at 94 °C for 3 min., 40 cycles of 94 °C for 30 sec., 58 °C for 30 sec., and 72 °C for 3 ½ min., and final elongation step at 72 °C for 10 min. PCR products were verified on 0.8 % agarose gel electrophoresis (90 V/h).

2.5.2 Phenotypical characterization

Transgenic plans were characterized phenotypically using a dissecting microscope (Nikon SMZ 1500) and pictures were taken with Camera (Nikon, Digital Sight, DS-Fi1) using the NIS-Elements F3.0 software. Statistical analysis of phenotypic frequencies was performed in R statistical analysis program.

2.5.3 Total RNA isolation from *Arabidopsis* plants

Total RNA from wild type and transgenic *Arabidopsis* lines was extracted using RNeasy Mini Kit (QIAGEN), following the Plant cell and tissue protocol. Plant tissue material (~100 mg) was gathered from two and five weeks old seedlings. The tissue was taken with sharp tweezers (sterilized between each sample) and collected into Eppendorf tubes containing 3 steel beads, pretreated with 0.1 % DEPC and autoclaved (Sambrook et al., 1989). Then 450 µl buffer RLT (QIAGENE) supplemented with β-mercaptoethanol (10 µl/ml) was added to the tubes and the tissue was disrupted for 45 seconds in the mini-bead beater machine. The Qiagene protocol was then followed from step 4 (page 54, RNeasy® Mini Handbook,

Version September 2010). RNA concentrations and purity were determined spectrophotometrically (NanoDrop 1000 v3.7.1, Thermo Scientific).

2.5.4 First-Strand cDNA Synthesis

Total RNA was first diluted to ~ 50 ng/µl. To 8 µl diluted total RNA was added 1 µl 10x DNase I Buffer (Invitrogen) and 1 µl DNase I (Invitrogen). The reaction was incubated at room temperature for 15 min. To each reaction was added 1 µl of 25 mM EDTA (Stop solution) and incubated at 65 °C for 10 min. Then 1 µl of 250 ng/µl random primer mixture $pd(N)_6$ (Amersham Pharmacia Biotech Inc) and 1 µl of 10 mM dNTPs was added to each reaction, and incubated at 65 °C for 5 min. A 7 µl of reverse transcription mixture containing 4 µl 5x First-Strand buffer, 1 µl 0.1 M DTT, 1 µl 40 U/µl RiboLock RNase inhibitor (Fermentas) and 1 µl 200 U/µl SuperScript III Reverse Transcriptase (Invitrogen) was added to each well. The reverse transcription reaction (cDNA synthesis) was carried out according to the manufacturers conditions: 25°C for 5 min, 50 °C for 30 min and 70 °C for 15 min. The control reactions were performed without Reverse Transcriptase.

2.5.5 Real-time quantitative PCR (qPCR)

Real-time qPCR was performed in the 7500 Real-time PCR System (Applied Biosystems) using Eva Green Fire Pol qPCR Mix. The qPCR reaction contained 5x Eva Green Fire Pol qPCR Mix (Solis Bio Dyne), 0.1 µM of each sense and antisense primer and 5 µl of 10-fold diluted cDNA in a total reaction volume of 20 µl. The reactions were performed in optical 96-well plates. Cycling conditions applied was at 95 °C for 15 min, and then 40 cycles with 95 °C for 15 sec, 60 °C for 15 sec and 72 °C for 20 sec. Specific primers were designed to amplify cDNA representing each inserted transgene (the different loop regions): AtDEK1-LOOP, AtDEK1-LOOP-DEL, MvDEK1-LOOP and ZmDEK1-LOOP, the DEK1-Calpain and protein phosphatase 2A (PP2A) used as reference gene. Primer sequences and respective amplicon size are represented in Appendix 1, Table 5 A1. Negative control omitting cDNA was performed for each run.

The data resulting from monitoring the qPCR reactions was analyzed and determined in the LinRegPCR program. The program uses non-baseline corrected data, performs a baseline correction on each sample separately, determines a window-of-linearity and then uses linear regression analysis to fit a straight line through the PCR data set. From the slope of this line the PCR efficiency of each individual sample is calculated. The mean PCR efficiency per amplicon and the Ct value per sample are used to calculate a starting concentration per sample (N0), expressed in arbitrary fluorescence units (Ramakers et al., 2003; Ruijter et al., 2009).

2.5.6 Verification of qPCR products

The PCR products were verified on the 1 % agarose gel electrophoresis and by sequencing using BigDye Terminator v3.1 Sequencing Kit (Applied Biosystems). Prior to the sequencing reaction, PCR products were treated with Exonuclease I (New England Biolabs) to degrade ssDNA (excess PCR primers) from the samples. The reaction contained 2 µl of PCR product, 1x Big Dye Sequencing Buffer and 4 U ExoI in a total reaction volume of 10 µl. The samples were incubated at 37 °C for 60 min and 85 °C for 15 min. The resulting PCR products were sequenced with overlapping specific primers (Appendix 1, Table 5 A1). To 5 µl of ExoI-treated PCR product 1x BigDye Sequencing Buffer, 0.32 µM gene specific forward (or reverse) primer and 0.5 µl of BigDye Terminator mix v3.1 were added in a 10 µl reaction volume. The cycling conditions: 96 °C for 1 min, 15 cycles: 96 °C for 10 sec, 50 °C for 5 sec, 60 °C for 1 min 15 sec, 5 cycles: 96 °C for 10 sec, 50 °C for 5 sec, 60 °C for 90 sec. and 5 cycles: 96 °C for 10 sec, 50 °C for 5 sec, 60 °C for 2 min. To clean up the sequencing reactions BigDye Xterminator Purification Kit was used; 10 µl of sequencing reaction was mixed with 45 µl SAM solution and 10 µl Xterminator solution. The plate containing reaction mixtures was placed on shaker for 1 hour at RT and then centrifuged prior to sequencing.

3. Results

3.1 Bioinformatics analyses

3.1.1 BLAST search showes that the DEK1 sequence is highly conserved in Streptohyta

BLASTp search in the NCBI databases using the *Arabidopsis thaliana* DEK1 sequence (Lid et al., 2002) as query identified several DEK1 homologs in other plant species including castorbean, wine grape, tobacco plant, rice, maize, barley and black cottonwood. The DEK1 sequences in these organisms is highly conserved sharing from 63 to 98 % overall amino acid sequence identity. The identities in the DEK1-LOOP domain from these organisms are between 39 - 96% (Table 2). A pairwise alignment of *Arabidopsis* DEK1-LOOP and the putative *Mesostigma viride* DEK1-LOOP sequence (identified from a cDNA cloned in our laboratory by Johansen, 2010) revealed 16 % identity.

Table 2 Dek1 homologues identified by BLASTp and their identity shared with AtDek1 and AtDek1-Loop

Organism	Accession No.	Overall Dek1 Identity	Loop Domain Identity	Calpain Domain Identity
Arabidopsis lyrata	gb/AAR72488.1/	98 %	96 %	99 %
Ricinus communis	gb/EEF38998.1/	83 %	76 %	86 %
Vitis vinifera	ref/XP002285732.1/	81 %	75%	85 %
Nicotina benthamiana	gb/AAQ55288.2/	79 %	70 %	84 %
Oryza sativa	gb/AAL38190.1/	72 %	60 %	78 %
Sorhum bicolor	gb/EER95003.1/	72 %	59%	77 %
Zea mays	gb/AAL38189.1/	71 %	60 %	77 %
Hordeum vulgare	gb/ABW81402.1/	71 %	58 %	78 %
Brachypodium distachyon	gb/ACF22702.1/	65 %	53 %	-
Selaginella moellendorffii	gb/EFJ08518.1/	64 %	46 %	72 %
Physcomitrella patens	gb/EDQ60982.1/	63 %	39 %	69 %
Mesostigma viride	-	-	16 %	-

The overall DEK1 identity percentage reported by BLASTp show that the DEK1 sequence is highly conserved in Streptohyta organisms. *A. thaliana* diverged from its close relative *A. lyrata* about 10 million years ago (Koch et al., 2000; Wright et al., 2002; Ossowski et al., 2010), yet DEK1 is highly conserved, sharing 98 % overall sequence identity and 96 % in the loop domain. The overall AtDEK1 shared percentage identity with maize and *Physcomitrella* is 71 % and 63 % respectively, and 60 and 39 % for the loop domain. The whole MvDEK1 sequence has not been annotated yet, nevertheless the MvDEK1-LOOP sequence share quite low percentage identity with *A.thaliana* loop. Compared to the percentage identity of the Calpain domain, known to be the most conserved and biologically active domain in DEK1, the LOOP domain is less conserved and does not have known function (Table 2).

3.1.2 Multiple sequence alignment of the DEK1-LOOP sequence reveals conserved amino acid residues

In order to reveal conserved residues and putative common features in the DEK1-LOOP protein-coding region, multiple sequence alignment (MSA) was performed in CLC Genomic Workbench (v 4.5.1) using the amino acid sequences recognized as homologs of AtDEK1-LOOP by the BLASTp search and the amino acid sequences of MvDEK1-LOOP generated from *Mesostigma viride* cDNA in our laboratory. The MSA was performed using the ClustalW method (Edgar and Batzoglou, 2006; Thompson et al., 1994).

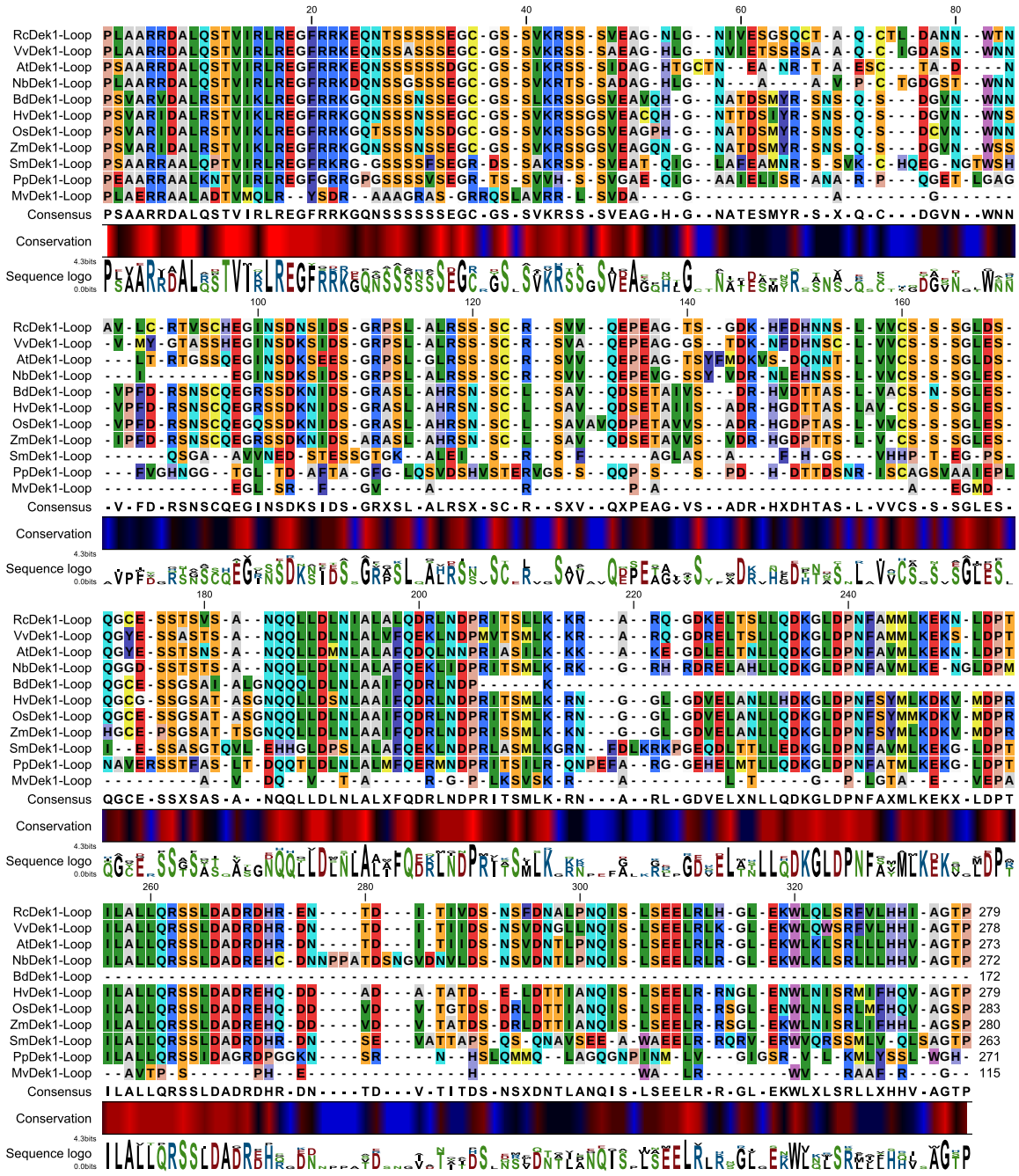


Figure 10 Multiple sequence alignment of the Dek1-Loop amino acid sequences, performed in CLC Work Bench (using ClustalW) . Abbreviations for the organisms used in the alignment: Rc (*Ricinus communis*), Vv (*Vitis vinifera*), At (*Arabidopsis thaliana*), Nb (*Nicotina benthamiana*), Bd (*Brachypodium distachyon*), Hv (*Hordeum vulgare*), Os (*Oryza sativa*), Zm (*Zea mays*), Sm (*Selaginella moellendorffii*), Pp (*Phiscomitrella patens*), and Mv (*Mesostigma viride*). The most conserved predicted residues are indicated in red and least conserved areas in blue.

The predicted multiple sequence alignment of the amino acid sequences of DEK1-LOOP domain reveals conserved clusters of amino acids shared by Streptohyta organisms. The *Mesostigma viride* (Mv) DEK1-LOOP sequence is less then half the size of DEK1-LOOP sequence in *Arabidopsis* (115 aa and 273 respectively). However, it is notable that the

MvDEK1-LOOP sequence show preserved amino acid residues in areas where the sequences of more derived plants are well conserved. For instance the Proline, Arginine, Leucine, Valine, Serine and Glysin residues in the first conserved cluster at the beginning of the alignment (Figure 10). In addition, the MvLOOP sequence was predicted to contain a few phosphorilation sites and one myristoylation site common also for the *Arabidopsis* and maize sequences predicted by Prosite (Appendix 2, Figure 2 A2). For instance at residue 40 KRSS on the At and Zm sequences and RRLS on the Mv- sequence were predicted as cAMP- and cGMP-dependent protein kinase phosphorylation sites, and on the conserved cluster at around residues 260 in the alignment, SSLD (on AtLOOP and ZmLOOP) and SPHE (on MvLOOP) were predicted as Casein kinase II phosphorylation sites (Figure 10).

3.1.3 Similarity tree

To reveal presumed evolutionary relationship between the homolog DEK1-LOOP sequences, a phylogenetic tree was constructed based on the multiple sequence alignment produced by CLC Work Bench, using the clustering algorithm Unweighted Pair Group Method using arithmetic averages (UPGMA) for distance data (Michener & Sokal, 1957; Sneath & Sokal, 1973).

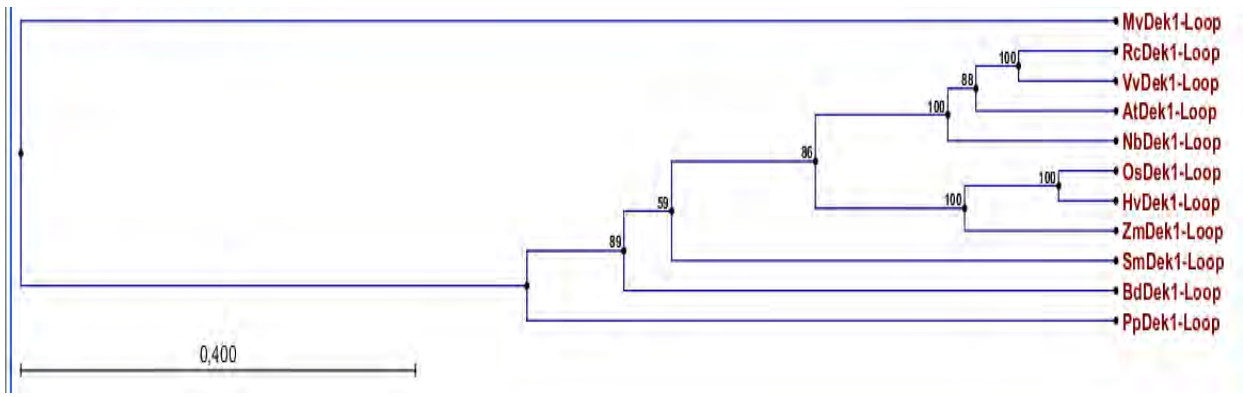


Figure 11 Phylogenetic relationship between DEK1-Loop homologs. The tree was constructed in CLC based on the MSA produced by CLC Work Bench (by the UPGMA algorithm) from amino acid sequences identified by BLASTp as AtDEK1-LOOP homologues. Abbreviations for the organisms used: Mv (*Mesostigma viride*), Rc (*Ricinus communis*), Vv (*Vitis vinifera*), At (*Arabidopsis thaliana*), Nb (*Nicotina benthamiana*), Os (*Oryza sativa*), Hv (*Hordeum vulgare*), Zm (*Zea mays*), Sm (*Selaginella moellendorffii*), Bd (*Brachypodium distachyon*), and Pp (*Phiscomitrella patens*).

Several different programs and MSAs were used to generate similarity threes and all predicting similar branching, of these the one with highest bootstrap scores is shown here (Figure 11). The predicted tree is also comparable to the Lineage Report **(**see Appendix 2, Figure 1 A2) provided by the BLAST search, which show how closely the organisms in the BLAST hitlist are related to the query sequence (AtDEK1-LOOP), according to their classification in the taxonomy database. The predicted phylogenic tree based on DEK1-LOOP sequences is also very similar to the phylogenetic tree of the whole DEK1 protein (Figure 3), suggesting that divergence has accrued between monocots and dicots from a common ancestor, and the Loop domain sequence is clearly affected by this divergence.

3.1.4 Bioinformatics anlyses did not reveal clues about DEK1-LOOP potential function

In search for homolog domains to the DEK1-LOOP, the amino acid sequences of At- Zm-, Pp- and Mv-Dek1-Loop were submitted to different structure, function and similarity prediction servers. Search for homolog sequences in the NCBI non-redundant database using the Zm-, Pp- and Mv-DEK1-LOOP as queries, did not result in significant matches, except one that was recognized as putative conserved domain argS, arginyl-tRNA synthetase (E-value = 4.79e-03) on the AtDEK1-LOOP sequence. Closer examination of the alignment of this hit suggested that this is a false positive result. The Pfam server found no hits for neither of the sequences, and the Phyre engine recognized EF-hand motif in the AtDEK1-LOOP, which could have been quite interesting, if biologically relevant. However, the sequence alignment was very weak, and the estimated precision and the percentage of identical residues in the sequence alignment were too low. This EF-hand motif was not detected in the other sequences. The At-, Zm-, Pp- and Mv-Dek1-Loop amino acid sequences were also submitted in the 3DLigandSite; predicting ligand-binding sites using similar structures (Wass et al., 2010) where sufficient homologous structures with ligands bound could not be identified.

Interesting results were generated from the Predict Protein server (http://www.predictprotein.org/). Prosite predicted a few putative N-glycosylation sites on the *Arabidopsis* and *Zea mays* DEK1-LOOP amino acid sequences. One cAMP- and cGMP-dependent protein kinase phosphorylation site was predicted common for all three

sequences. Also a few Protein kinase C phosphorylation sites, a number of Casein kinase II phosphorylation sites and some N-myristoylation sites were predicted common on all At-, Zm-, and Mv-Dek1-Loops. An extra Amidation site was predicted only on the *Mesostigma viride* DEK1-LOOP sequence (Appendix 2, Figure 2 A2).

In summary, the performed bioinformatics analyses on the DEK1-LOOP revealed that the sequence is not as conserved as the biologically active part of the DEK1 the Calpain domain. No significant results that could indicate anything about the DEK1-LOOP function were identified by searching in bioinformatics databases.

3.2 Phenotypic screening and characterization of transformants

Wild type plants transformed with the four different constructs; AtDEK1-MEM, AtDEK1-MEM-DEL (with deleted loop domain), AtDEK1-MEM-MvLoop (with replaced loop domain from *Mesostigma viride*), and AtDEK1-MEM-ZmLoop (with replaced loop domain from *Zea mays*) were selected on M&S medium supplemented with kanamycin. The putative transformants were studied under dissecting microscope at the age of two weeks, and then verified by PCR with the appropriate primers (data not shown).

3.2.1 Phenotypes of the AtDek1-MEM-over-expressing plants

The observed phenotypes of AtDek1-MEM overexpressing transformants obtained in this study closely resembled the phenotypes described by Tian et al. (2007) in a identical research. In addition these phenotypes resembled very closely the range of phenotypes described in plants overexpressing the At DEK1-RNAi construct by Johnson et al. (2005) and Tian et al. (2007). Combined these studies suggest that overexpression of the DEK1 membrane anchored domain induces a dominant negative effect comparable to the loss or down-regulation of DEK1 function. In the present experiment a total of 80 independent T1 lines were obtained and studied. Phenotypic alterations were observed in 62 % of the obtained independent T1 lines, which is consistent with the percentage (~ 50 %) observed by

Tian et al (2007) in the initial experiment. About 38 % of the transgenic plants showed phenotype nearly undistinguishable form wild type plants (wt).

The most common observed phenotype was represented by 19 plants (24 % of transgenic lines), showing partially or completely damaged epidermal cell layer on the cotyledons (DEC, Figure 12 D and F). Of these plants 12 grew up normally with wild type appearance after being transferred to soil, 1 of them was sterile and 6 died.

Another common and more severe phenotype represented by 14 plants (18 % of transgenic plants) was observed with wrinkled, damaged cotyledons, which were partially or completely fused on the adaxial faces (FC, Figure 12 J, K and L). The same phenotype has been described by Johnson et al (2005) in 25% of the RNAi knockdown mutants, identified as class 1, and by Tian et al (2007) in AtDek1-MEM overexpressing plants and in AtDek1-RNAi knockdown plants. Most of the seedlings classified as FC (fused cotyledons) died soon after being transferred to soil, however a few FC plants managed to produce several strange structured leaves from the apical meristem (Figure 12 L) before they died and one of them grew up normally and showed no further developmental defects.

A third common phenotype was represented by 9 plants (11 % of transgenic plants) with fully missing or partly defective shoot apical meristem (DSAM, Figure 12 G, H and I), which were unable to produce any other leaves beyond the cotyledons. This phenotype was most frequently combined with the DEC phenotype (8 of the 9 seedlings). All off these seedlings died shortly after being transferred to soil (by the age of around two and a half weeks).

A small number of transgenic seedlings (n = 4) could be distinguished by having only single first true leaf (SFTL, Figure 12 E), three of those also showed the DEC phenotype. Two of the SFTL classified seedlings grew up normally and the other two showed phenotypes on soil; they did not produce further organs beyond the rosette leaves, which were also abnormal (Appendix 2, Figure 4 A2, 19MEM2 and 20MEM1).

Another small phenotype group (n = 4) produced only one, single cotyledon (SC, Figure 12 M, N and O) and meristem-structure-like formation on the stalk (N), which in two of the four seedlings developed into abnormal true leaves (O). One of these seedlings managed to grow up normally with wild type appearance, but it was sterile. This phenotype was not described previously in the initial experiment performed by Tian et al. (2007).

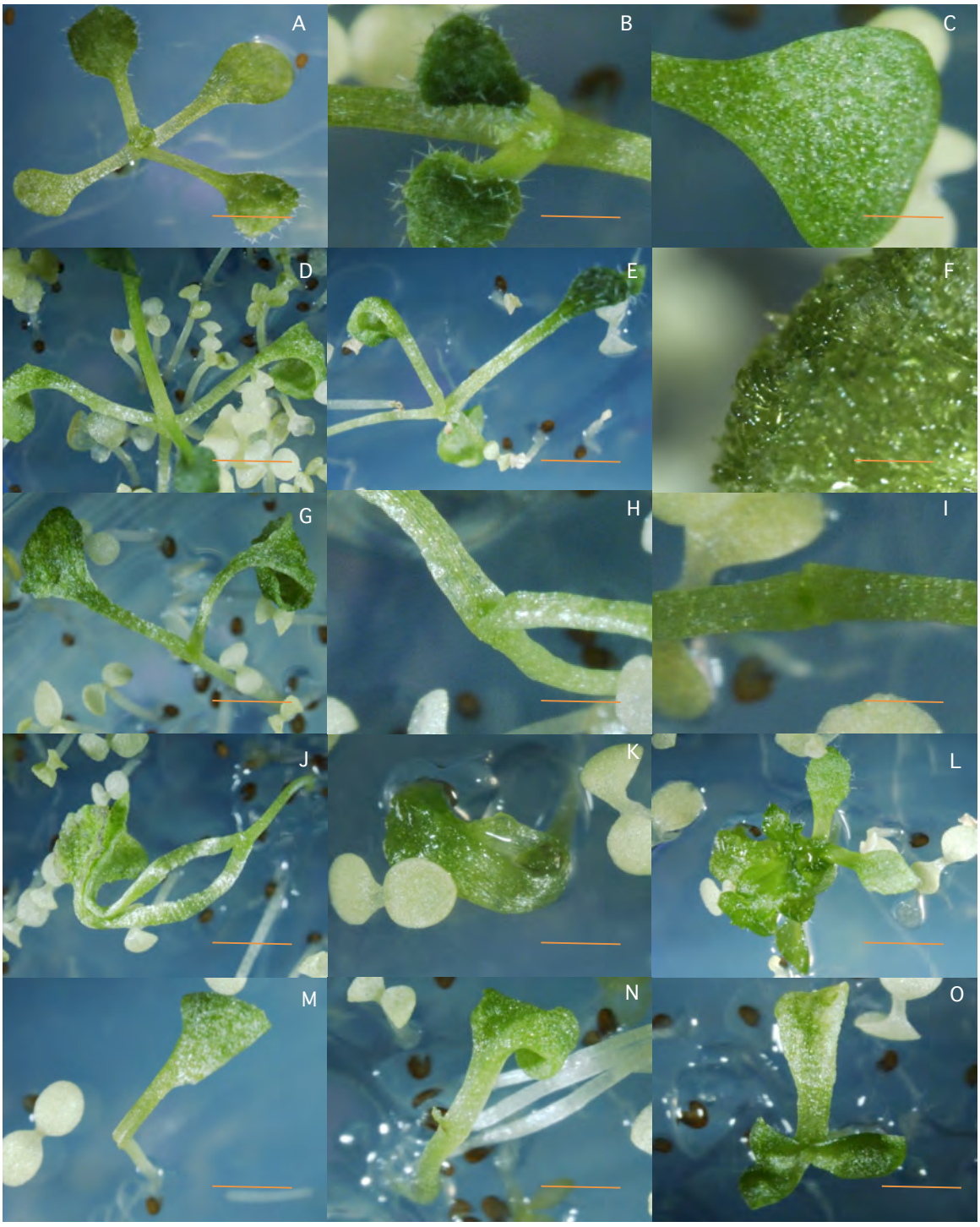


Figure 12 Examples of the most common phenotypes observed in 2 weeks old AtDEK1-MEM overexpressing seedlings. A: wild type *Arabidopsis* seedling, B: normal shoot apical meristem, C: normal cotyledon, D: defective epidermal cell layer on the cotyledons (DEC) phenotype, E: single first true leave (SFTL) phenotype combined with DEC, F: close up of defective cotyledon, G: defective shoot apical meristem (DSAM) phenotype combined with DEC, H & I: close up of the defective shoot apical maristem (DSAM), J, K and L: fused cotyledons (FC) phenotype, M, N and O: single cotyledon (SC) phenotypes. Pictures were taken under dissecting microscope Nikon SMZ 1500. Bar size is 1mm.

One transformation attempt was enough to generate a sufficient number of independent T1 lines. After being transferred to soil 73 % of the seedlings showing wild type phenotype grow up normally and formed healthy siliques with seeds, and 27 % of them were sterile. About 28 % of seedling showing abnormal phenotypes grew up normally without further developmental defects observed, 8 % were sterile and 64 % died shortly after being transferred to soil. Very few plants (5 %) showed a phenotype after being transferred to soil (see Appendix 2, figure 4 A2) by the age of two weeks.

3.2.2 Phenotypes of the At-MEM-DEK1-DEL-over-expressing plants

Three transformation attempts with the At-MEM-DEK1-DEL construct were performed with very low transformation efficiency accomplished. A total of 33 independent T1 lines were obtained and studied. About 73 % (24 seedlings) of the kanamycin resistant seedlings showed phenotype very closely mimicking or undistinguishable from wild type phenotype (Figure 13). Structural alterations with a degree of variation were observed in 27 % of the obtained independent T1 lines, which is contradictory with data published by Tian et al. (2007), where wild type phenotype was observed in all kanamycin resistant seedlings.

The phenotypes obtained in transgenic plants expressing the AtMEM-DEK1-DEL construct could be compared with the phenotypes observed in AtDEK1-MEM overexpressing seedlings. The most common phenotype shown in 6 seedlings (18 % of transgenic plants), was the one that could be described by the presence of only one or single first true leaf (SFTL, Figure 13 D and E), five of them were also combined with the DEC phenotype (Figure 13 F and I), where the epidermal layer on the cotyledons was damaged.

Three seedlings (9 % of transformants) showed the DSAM phenotype (Figure 13 G and H) with fully missing or partly defective shoot apical meristem, two of them were combined with the DEC phenotype (defects on the epidermal layer of the cotyledons).

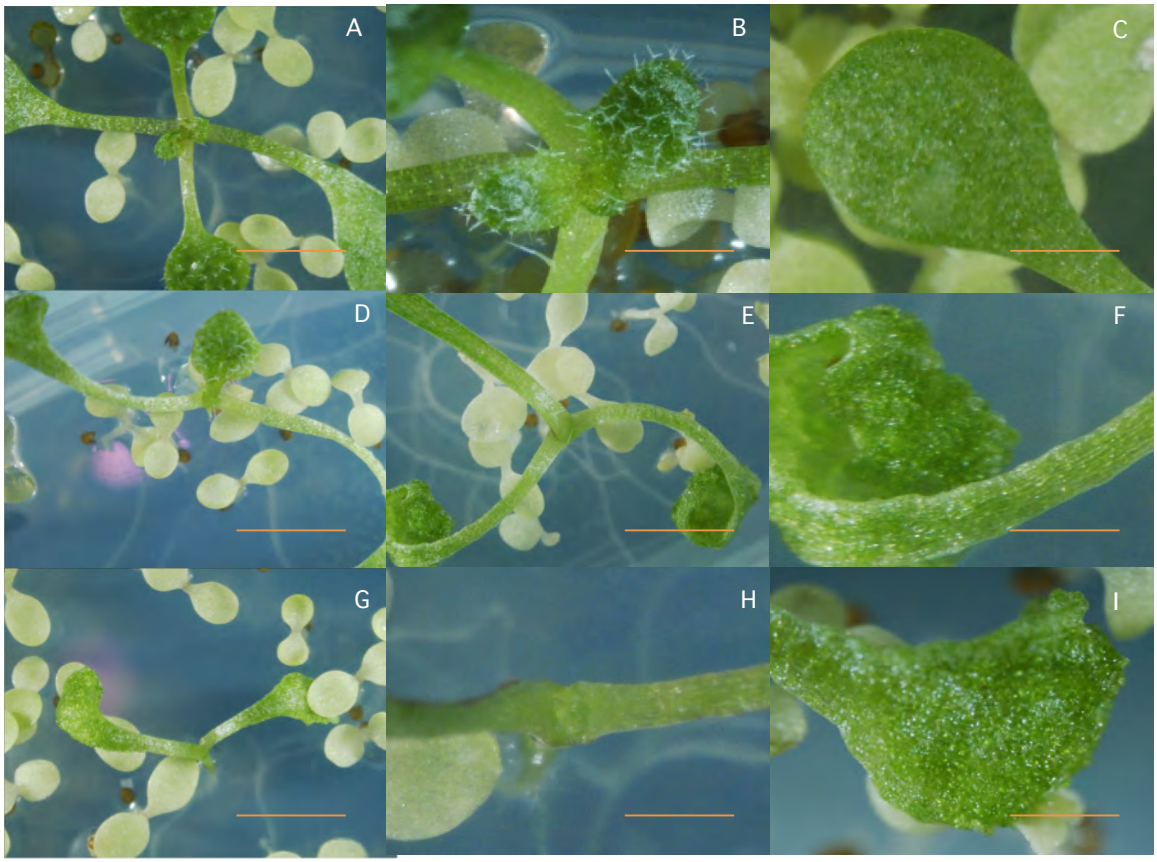


Figure 13 Examples of the most common phenotype observed in 2 weeks old AtDEK1-MEM-DEL overexpressing seedlings. A: wild type phenotype, B: normal shoot apical meristem (SAM), C: normal cotyledon, D and E: single first true leave (SFTL), F and I: defective ephidermal layer on the cotyledons (DEC), G: defective shoot apical meristem (DSAM), H: close up of defective shoot apical meristem. Pictures were taken under dissecting microscope Nikon SMZ 1500. Bar 1mm.

After being transferred to soil 62% of the seedlings showing wild type phenotype grew up normally and formed healthy siliques with seeds, and 38% were sterile. About 22 % of seedling showing abnormal phenotypes grew up normally without any further developmental abnormalities, while 22% of those were sterile and 56 % died shortly after being transferred to soil.

3.2.3 Phenotypes of AtDEK1-MEM-MvLOOP-over-expressing plants

A total of 46 independent transgenic lines were obtained and studied from one transformation attempt with the AtDek1-MEM-MvLoop construct. 17 (37 %) of the transformants produced a phenotype undistinguishable from wild type and 63 % showed a variety of developmental defects similar to those observed in seedlings overexpressing the

AtDEK1-MEM construct. Of these 69 % (20 transgenic plants) showed the most common phenotype with defective epidermal layer on the cotyledons (DEC, Figure 14 C), all of those developed normal rosette leaves, 11 grew up normally with wild type appearance, 1 was sterile and 8 died. 5 seedlings (17 %) showed the single first true leave (SFTL, Figure 14 D) phenotype combined with DEC phenotype. Two seedlings expressed the fused cotyledon (FC, Figure 14 E) phenotype. Only one seedling lacked shoot apical meristem (DSAM) and only one showed the single cotyledon (SC, Figure 14 F) phenotype.

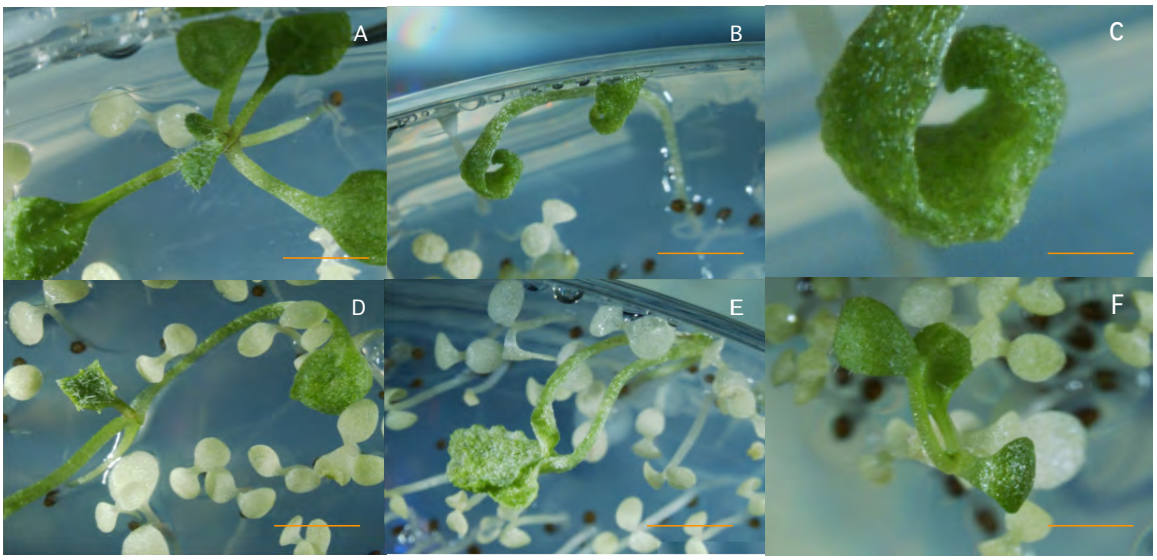


Figure 14 Examples of the most common phenotypes observed in 2 weeks old AtDEK1-MEM-MvLoop overexpressing seedlings. A: wild typet phenotype, B and C: defective shoot apical meristem (DSAM) phenotype (the actual DSAM is not shown), C: defective ephidermal layer on the cotyledon (DEC) close up, D: single first true leave (SFTL) phenotype, E: fused cotyledons (FC) phenotype, and F: singel cotyledon (SC) phenotype. Pictures were taken under dissecting microscope Nikon SMZ 1500. Bar 1mm.

After being transferred to soil 65 % of the wild type phenotype plants grew up normally and developed seeds, 29 % of them were sterile and one plant died. About 59 % of the seedlings expressing abnormal phenotypes grow up to be normal, one of them was sterile and the rest died soon after being transferred to soil.

One distinction between the AtMEM-DEK1-MV-LOOP expressing plants and those expressing the AtDEK1-MEM construct with or without the At LOOP was that the MvLOOP expressing plants grew faster and taller then the other transgenic plants. At the same age the MvLOOP overexpressing plants were growing approximately 5 cm taller then

wild type control plants and AtDEK1-MEM transgenic plants (Appendix 2, figure 5 A2). This particular characteristic was not observed in any of the other transgenic plants obtained in this study, however it was not further investigated.

3.2.4 Phenotypes of AtDEK1-MEM-ZmLOOP-over-expressing plants

Seedling overexpressing the AtDEK1-MEM-ZmLOOP where the loop of Arabidopsis was replaced with the loop from maize showed 38 % phenotypic abnormalities, which were comparable with the range of phenotypes observed in plants overexpressing the other constructs (AtDEK1-MEM, AtDEK1-MEM-DEL, and AtDEK1-MEM-MvLoop). Two transformation attempts generated 37 T1 independent transformants, of these 23 (62 %) showed wild type phenotype.

Only two seedlings showed the DEC (defective epidermal layer on the cotyledons) phenotype, which was the most common phenotype observed in AtDEK1-MEM overexpressing seedlings. Of the transgenic lines 6 (16 %) had a defective shoot apical meristem (DSAM, Figure 15 B and C). 3 seedlings produced only one first true leaf or spike-like structure classified as SFTL (Figure 15 D and E) phenotype and 3 seedlings produced the phenotype with fused cotyledons on the abaxial sides (FC, Figure 15 G and H).

One seedling showed a phenotype not observed previously, the seedling had three normally looking cotyledons and three normal first true leaves (Figure 15 I). This phenotype has been observed previously in wild type populations (Wenche Johansen, personal communication) and was therefore classified as wild type phenotype in this experiment.

Figure 15 Examples of the most common phenotypes observed in 2 weeks old AtDEK1-MEM-ZmLOOP overexpressing seedlings. A wild type phenotype, B: the defective epidermal layer on the cotyledons with defective shoot apical meristem (DEC+DSAM), C: single first true leave (SFTL), and D: fused cotyledons (FC). Pictures were taken under dissecting microscope Nikon SMZ 1500. Bar 1mm

Of the wild type expressing seedlings 13 (57 %) grew up with normal physiology, 4 (17 %) of them where sterile and 5 died, 11 (79 %) of the seedlings showing abnormal phenotypes died shortly after being transferred to soil. Two of the DEC phenotype seedlings showed no further developmental defects after being transferred to soil, two of them grew up normal and one was sterile.

3.3 Frequencies of the observed phenotypes

The observed variations in phenotypes of all the independent transformation events could be grouped into six different classes or categories; 1) wild type phenotype (wt), 2) defective epidermal cells on the cotyledon (DEC) characterized by wrinkled cotyledons with defects in the epidermal cell layer most frequently on the adaxial side, 3) defective shoot apical

meristem (DSAM) phenotype where the apical meristem was missing or in some way was defective, very often observed in combination with the DEC phenotype, 4) single first true leaf (SFTL), where the leaf could vary from normal to variation of appearance, to spike-like form. This phenotype was also very often combined with the DEC phenotype, 5) partially or completely fused cotyledons (FC) that also had damages on the epidermal layer of the cotyledons but were also partially or completely fused together on the adaxial side of the leafs, and finally 6) single cotyledon only (SC).

The frequencies of the six different classes of phenotypes were calculated to investigate if they appear in different ratios in the different lines of transformants (Table 3, Figure 16).

Table 3 Calculated frequencies of the distinctive phenotype classes observed in the different transformations

	wt	**DEC**	**DSAM**	**SFTL**	**FC**	**SC**	**Total**
AtMEM	30 (38%)	19 (24%)	9 (11%)	4 (5%)	14 (18%)	4 (5)	80
AtMEM-DEL	24 (73%)	0	3 (9%)	6 (18%)	0	0	33
MvLoop	17 (37%)	20 (43%)	1 (2%)	5 (11%)	1 (4%)	1 (2%)	46
ZmLoop	23 (62%)	2 (5%)	6 (16%)	3 (8%)	3 (8%)	0	37

Table 3 above shows the phenotypic frequencies appearing in the different transformant populations, both the number of counted representatives and the corresponding percentage calculated are presented. The total number of obtained independent lines from each transformation with the different constructs is shown in the last column (Table 3).

In order to provide better overview of table 3, the data of the calculated frequencies is presented graphically in Figure 16.

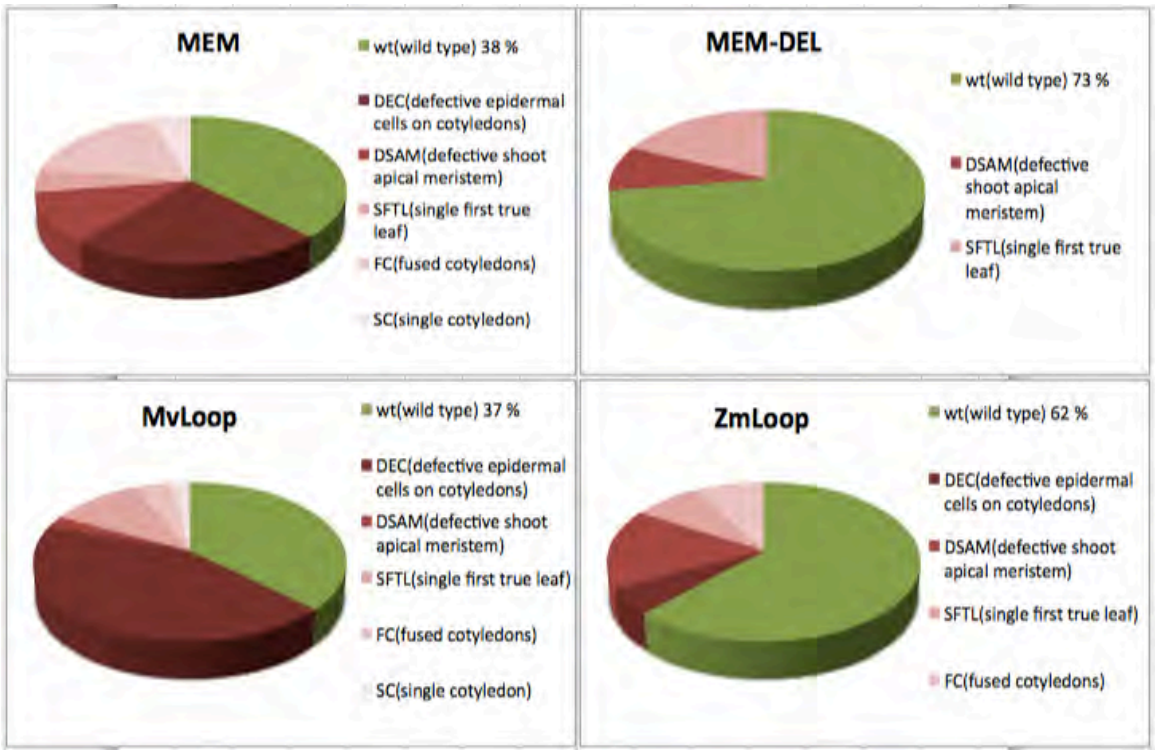


Figure 16 The frequencies of observed phenotypes in seedlings overexpressing the four different constructs. AtDEK1-MEM, AtDEK1-MEM-DEL, AtDEK1-MEM-MvLoop, and AtDEK1-MEM-ZmLoop. Wild type phenotype are represented in green, DEC in dark red, DSAM in red, SFTL in pink, FC light pink and SC in pale pink.

The figures above represents an overview of the different phenotype frequencies observed in the transgenic populations obtained from transformation with the four different constructs; AtDEK1-MEM, AtDEK1-MEM-DEL, AtDEK1-MEM-MvLoop, and AtDEK1-MEM-ZmLoop. Seedlings overexpressing the AtDEK1-MEM show 38 % wild type phenotype, 24 % DEC, 11 % DSAM, 5 % SFTL, 18 % FC and 5 % SC phenotypes. The AtDEK1-MEM-DEL overexpressing seedlings showed 73 % wt phenotype, 9 % DSAM, and 18 % SFTL phenotypes. Differences in phenotypes between the AtDEK1-MEM and AtDEK1-MEM-DEL overexpressing seedlings that may be worth noticing in addition to the main differences in wild type phenotype frequencies: The most common phenotype observed in AtDEK1-MEM seedling was defects on the epidermis of the cotyledons (DEC) with no further meristem or leaf abnormalities. This phenotype was not observed in AtDEK1-MEM-DEL seedlings. The fused cotyledons (FC) and single cotyledon (SC) phenotypes were also absent in the AtDEK1-MEM-DEL overexpressing population (figure 16). However this observed deviation could be an artifact from the much lower number of independent lines obtained.

Seedlings overexpressing AtDEK1-MEM-MvLoop closely resemble the frequencies of phenotypes observed in AtDEK1-MEM overexpressing seedlings: 37% showed wt, 43 % DEC, 2 % DSAM, 11 % SFTL; 4 % FC and 1 plant showed the SC phenotype. The AtDEK1-MEM-ZmLoop overexpressing seedling resembled the phenotype frequencies shown by the AtDEK1-MEM-DEL: 62 % showed the wt phenotype, 5 % DEC, 16 % DSAM, 8 % SFTL and 8 % FC.

Even though there is clear variation in phenotypes exerted in the different transformants, they most frequently share one common feature; the damaged epidermal layer on the cotyledons.

From Figure 16 it becomes clear that there is a difference in phenotypic frequencies between the AtDEK1-MEM lines compared to AtDEK1-MEM-DEL and AtDEK1-MEM-ZmLoop lines. However compared to the AtDEK1-MEM-MvLoop lines the phenotype frequencies are quite similar. In order to investigate if this observed differences and similarities are significant, we used χ^2 statistical test of independence, which tests the hypothesis that proportions (frequencies in our case) are the same in different groups or transformed lines. The chi-squared test was performed based on the numbers represented in table 3. The categorical variables used in the test were; seedlings showing phenotype and seedlings not showing phenotype (undistinguishable from wild type).

Table 4 Chi-squared test based on phenotype vs. no phenotype appearance in the different transformants

	MEM/DEL	**MEM/Mv**	**MEM/Zm**	**DEL/Mv**	**DEL/Zm**	**Mv/Zm**
χ^2	15.3	0.98	9.14	7.65	0.88	3.57
d.f.	1	1	1	1	1	1
p-value	**0.00009**	**0.321**	**0.00249**	**0.00568**	**0.348**	**0.059**
Min. expected value	14.6	16.4	15.5	15.0	10.8	18.3

The chi-squared test and the estimated p-values in table 4 confirm that there are significant differences in phenotypic expression between the AtDEK1-MEM and AtDEK1-MEM-DEL overexpressing seedlings, between the AtDEK1-MEM and AtDEK1-MEM-ZmLOOP seedlings and the AtDEK1-MEM-DEL and AtDEK1-MEM-MvLOOP overexpressing seedlings. The differences in observed phenotypes between AtDEK1-MEM and AtDEK1-MEM-MvLOOP, AtDEK1-MEM-DEL and AtDEK1-MEM-Zm were found to be

statistically insignificant. The minimum expected value show how well the test fits to the data, if this number is lower then 5; the test is unsuitable (Table 4).

3.4 Quantification of the transgene transcripts by qRT-PCR

Quantitative reverse-transcriptase PCR (qRT-PCR) was performed to investigate if the transgene expression levels correlated with the variation in phenotypes observed, or if the transgene is silenced in mature transgenic plants showing no abnormalities. Three representatives of each phenotype class of the AtDek1-MEM, AtDek1-MEM-DEL, AtDek1-MEM-MvLoop and AtDek1-MEM-ZmLoop lines were selected for qPCR analysis. Samples were collected from cotyledons two weeks after germination and from rosette leaves of five-week-old plants. Corresponding tissues from wild type plants grown side-by-side with the transgenic lines were also analyzed in order to compare the relative expression ratio with that of the transgenic lines.

Total RNA was isolated from sample tissue, reverse-transcribed into cDNA in two technical replicates, and then quantified by qPCR with the appropriate primers (see Appendix 1, Table 5 A1) specific towards the Loop region (the gene of interest), the reference gene used in the assay (protein phosphatase 2A gene), which is constitutively expressed (Czechowski et al., 2005), and the DEK1 calpain. Amplicon size and composition was verified by agarose gel electrophoresis and by sequencing (data not shown).

Real-time-PCR data was analyzed with the LinRegPCR program, where the PCR efficiency is calculated for each individual sample. The mean PCR efficiency per amplicon and the Ct value per sample are used to calculate a starting concentration per sample (N0), expressed in arbitrary fluorescence units. The results from the qRT-PCR analysis are presented by three different calculations; 1) as relative expression ratio of N0 between target gene and the calpain gene (Loop/Calp)), a measure of transgene expression relative to native DEK1 expression, 2) relative expression ratio of target gene and reference gene (Loop/PP2A), also a measure of transgene expression level but “normalized” to the reference gene (PP2A), and 3) relative expression ratio of native DEK1 “normalized” to the reference gene (Calp/PP2A), a measure included to investigate if the native *DEK1* expression was affected by the expression of the transgene.

The primers used to detect transgenic transcript in the AtDEK1-MEM lines will also detect endogenous DEK1 transcripts. Therefore, in wild type plants, one expects a Loop/Calp ratio of 1, which, as shown in Figure 17 (first two samples), was also experimentally confirmed.

The relative expression levels of the endogenous AtMEM-Loop (Loop/Calp and Loop/PP2A) and the Caplain (Calp/PP2A) in two- and five-week-old wild type plants were comparable (the first two samples in Figure 17). The Loop/Calp ratio in all analyzed transgenic lines are higher then 1 (the wild type ratio), confirming that the LOOP transcripts are overexpressed in these lines, and that non of the transgenic lines are subjected to transgene silencing (Figure 17). Thorough examination of the Loop/Calp expression ratios in the transgenic lines suggests that there are no correlation between the level of transgene expression and the severity of the phenotype observed. These results are confirmed by the Loop/PP2A ratios, using the expression level of constitutively expressed PP2A as reference transcript. In addition, it seems that there are no consistency in transgene expression level between individuals within a specific phenotypic class. For example, in transgenic lines classified as wild-type phenotype (lines 1, 10 and 33), the transgene expression levels vary strongly, between 3 (in line 33) to 42 (in line 10). In transgenic lines expressing the most common phenotype, the DEC phenotype, (line 14, 75 and 77) the relative expression level of the transgene also varies from approximately 2 (in lines 14 and 75) to 54 (in line 77). Relative expression ratios of the transgene (Loop/Calp and Loop/PP2A) in plants with wild type phenotype do not appear lower or different then the ratios observed in seedling showing severe abnormalities.

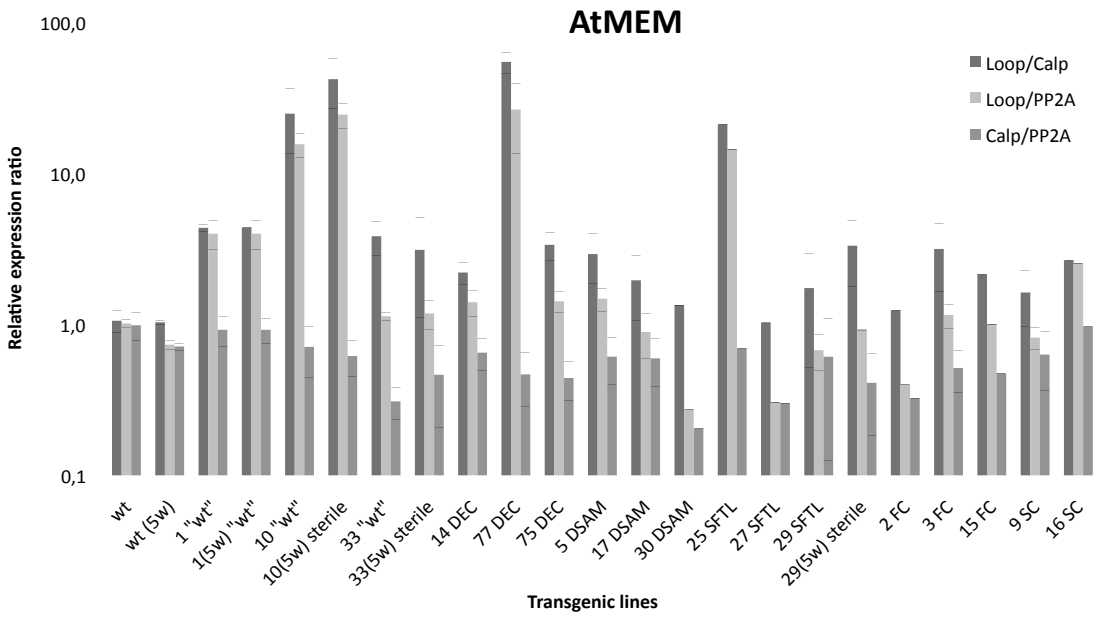


Figure 17 Quantitative analysis of the AtDek1-MEM construct in transgenic *Arabidopsis* lines. The relative expression ratio (based on starting concentration in the samples) of the Loop/Calpain, Loop/PP2A and Calp/PP2A are represented for each transgenic line. Each number represents independent transgenic line, the (5w) stands for sample taken from five weeks old plants from the same line, and the abbreviations are for the phenotypes: wt (wild type) phenotype, DEC (defective epidermal cell layer), DSAM (defective shoot apical meristem), SFTL (single first true leaf), FC (fused cotyledons) and SC (single cotyledon). The standard deviation for three technical replicates is shown.

In order to investigate if overexpression of the AtDEK1-MEM-LOOP sequence influences the expression of endogenous DEK1, the ratio of Calp/PP2A was analyzed and calculated (third column for each sample in Figure 17). In most of the transgenic lines analyzed, the result suggest that the endogenous expression of *DEK1* is not affected by the transgene insert, except for a few transgenic lines where the relative expression ratio of the calpain is lower then observed in wild type (the first two samples in Figure 17) in lines 33 (“wt”), 30 (DSAM), 27 (SFTL) and 2 (FC). Based on these inconsistent results it is not possible to conclude whether the endogenous DEK1 transcript is influenced by the overexpression of AtDEK1-MEM.

The q-PCR analysis also shows that the relative expression levels of the analyzed transcripts (all three calculations) in two-week-old seedlings (tissue from cotyledon) remain relatively unchanged or comparable to expression levels in mature plants (five-week-old, tissue from rosette leaves). Low relative expression ratio in two-week-old seedling observed in line 1 seems to be unchanged in the mature plant, and the high expression ratio in line 10 (two-weeks) is also high in the five-weeks-old plant.

Taken together, these results suggest that there are no correlation between the expression level of the transgene and the phenotypes observed in AtDEK1-MEM overexpressing plants, and it is uncertain if the endogenous DEK1 transcript is affected by the transformation.

The primers designed to detect transgene transcripts in the AtDEK1-MEM-DEL, AtDEK1-MEM-MvLoop, AtDEK1-MEM-ZmLoop lines do not amplify endogenous DEK1 transcripts, this was confirmed experimentally in a negative control for the primers, using cDNA generated from wild type plants. Q-PCR results obtained from transgenic plants transformed AtDEK1-MEM-DEL, AtDEK1-MEM-MvLoop, AtDEK1-MEM-ZmLoop constructs do not differ from the results obtained and described for the AtDEK1-MEM expression.

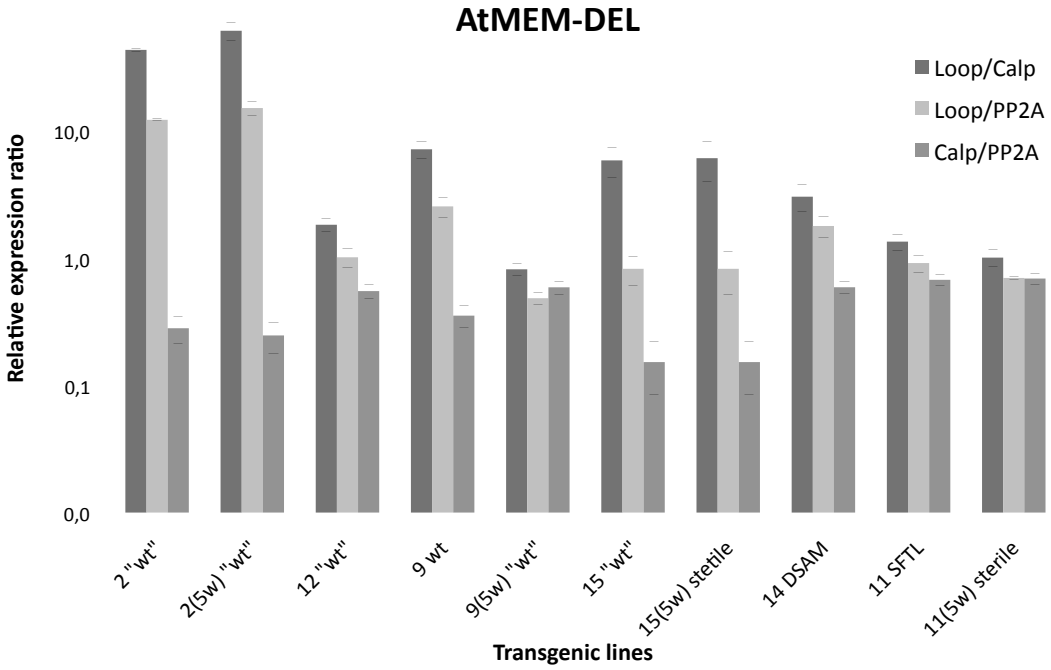


Figure 18 Quantitative analysis of the AtDek1-MEM-DEL construct in transgenic Arabidopsis lines. The relative expression ratio (based on starting concentration in the samples) of the Loop/Calpain, Loop/PP2A and Calp/PP2A are represented for each transgenic line. Each number represents independent transgenic line, the (5w) stands for sample taken from five weeks old plants from the same line, and the abbreviations are for the phenotypes: wt (wild type) phenotype, DSAM (defective shoot apical meristem), SFTL (single first true leaf), The standard deviation for three technical replicates is shown.

Q-PCR analyses of transgenic lines harboring the AtDEK1-MEM-DEL construct (represented in Figure 18) are similar to the results observed in AtDEK1-MEM transgenic

lines (Figure 17). There is no consistency in expression levels of the transgene (Loop/Calp and Loop/PP2A) in plants showing the same class of phenotype. For instance, transgenic lines 2, 12, 9 and 15 in Figure 18 represent plants with wild type appearing phenotype, where the observed relative expression ratio is quiet variable. The expression levels in 5-weeks-old lines 2 and 9, which represent fertile plants, and lines 11 and 15 represent sterile plants can not be correlated.

There is inconsistency in calpain expression (third columns for each sample, Calp/PP2A) observed here as well (as in AtDEK1-MEM lines); in lines 2 and 15 the calpain expression is much lower then expected (compared to wild type expression, Figure 17) and in the other lines no significant difference from wild type in calpain expression is observed.

Previously, in the qPCR analyses of the AtDEK1-MEM transgenic lines (Figure 17) we observed that expression in two-weeks-old seedlings was comparable to the expression levels in the mature plant (at the five weeks of age). In AtDEK1-MEM-DEL line 9 the transgene is expressed at around 7 in two-weeks-old seedling, but in the 5-weeks-old plant the transcript does not appear to be overexpressed (Figure 18), which could be a sign of transgene silencing in this mature plant. However such observation in only one sample is not convincing. Lines 14 and 11 show abnormal phenotypes DSAM and SFTL respectively, the transgenic transcript is not very highly overexpressed in these seedlings, as one would expect in a dominant negative phenotype. Line 14 was dead before the age of 5 weeks and line 11 showed sterility with no significant change in expression levels form the 2 weeks old seedling. These results suggest that there is no correlation between the expression levels of the transgene and the phenotypes observed in AtDEK1-MEM-DEL overexpressing plants.

Similar to the previous observations, the relative expression ratio of AtDEK1-MEM-MvLoop also could not be correlated to phenotypic variation in seedlings or to sterility in mature plants (Figure 19). An inconsistency in calpain expression levels are also observed here and is unclear whether the AtDEK1-MEM-MvLoop overexpression has an effect on the endogenous expression of DEK1 (Calp/PP2A columns in Figure 19).

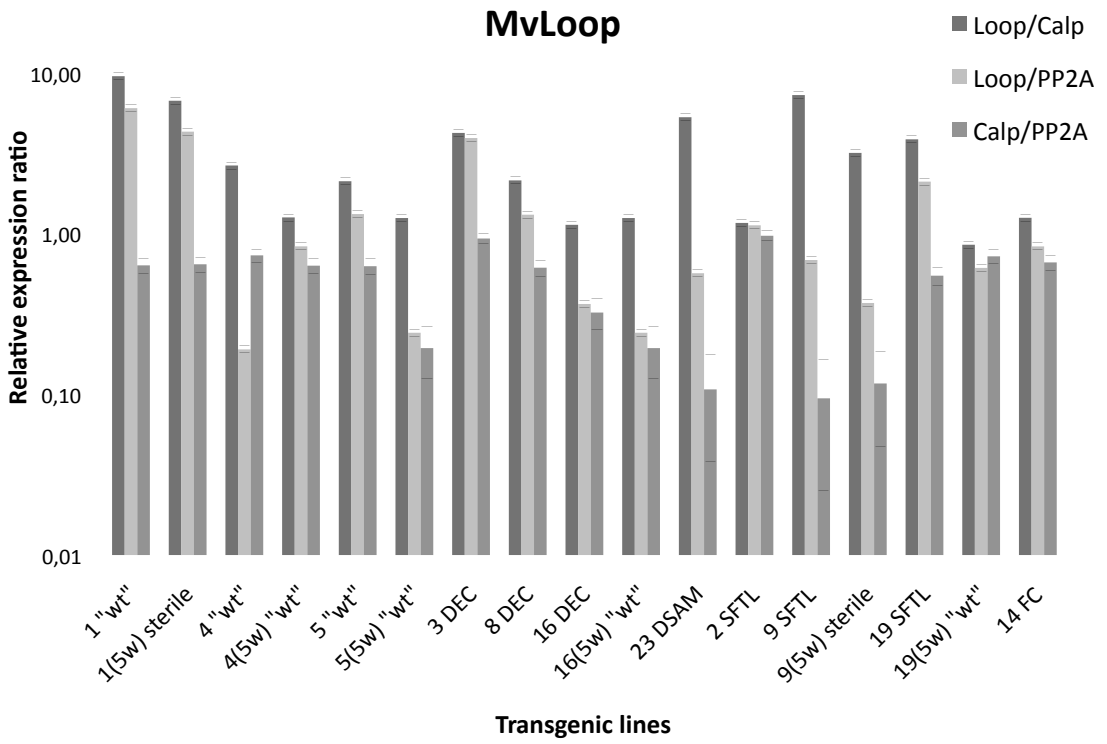


Figure 19 Quantitative analysis of the AtDek1-MEM-MvLoop construct in transgenic Arabidopsis lines. The relative expression ratio (based on starting concentration in the samples) of the Loop/Calpain, Loop/PP2A and Calp/PP2A are represented for each transgenic line. Each number represents independent transgenic line, the (5w) stands for sample taken from five weeks old plants from the same line, and the abbreviations are for the phenotypes: “wt” (wild type phenotype), DEC (defective epidermal cell layer), DSAM (defective shoot apical meristem), SFTL (single first true leaf), and FC (fused cotyledons). The standard deviation for three technical replicates is shown.

Lines 1, 4 and 5 in represent transgenic plants with wild type appearance. All three plans had wild type appearance as grown plans and line 1 was sterile. The expression ratios (Loop/Calp and Loop/PP2A) in these wild type appearing transgene plans vary strongly between these lines, lines 3, 8 and 16 represent the DEC phenotype, here is also variation in expression in-between the lines, suggesting that expression ratio can not be correlated to phenotype. In addition, the expression levels in plans showing severe phenotypes (23, 2 and 14) can not be described as different from the expression levels in wild type appearing lines (1, 4 and 5). These results suggest that there is no correlation in expression levels of the transgene and the phenotype observed in AtDEK1-MEM-MvLoop overexpressing plants.

The relative expression ratio of the AtDEK1-MEM-ZmLoop construct also could not be correlated to the phenotypes observed (Figure 20).

Figure 20 Quantitative analysis of the AtDek1-MEM-ZmLoop construct in transgenic Arabidopsis lines. The relative expression ratio (based on starting concentration in the samples) of the Loop/Calpain, Loop/PP2A and Calp/PP2A are represented for each transgenic line. Each number represents independent transgenic line, the (5w) stands for sample taken from five weeks old plants from the same line, and the abbreviations are for the phenotypes: wt (wild type) phenotype, DSAM (defective shoot apical meristem), SFTL (single first true leaf). The standard deviation for three technical replicates is shown.

The relative expression ratio of Loop/Calp (first column for each sample in Figure 20), shows that the inserted transgene is overexpressed in transgenic lines both in two-weeks-old seedlings and in five-weeks-old plants at comparable levels. These results are showing that the transgene is not subject to silencing in plants showing wild type phenotype as young seedlings or as mature plants. This is confirmed by the Loop/PP2A ratio (second column), however there is no obvious correlation in expression ratio to phenotypes observed or no similarities in expression levels in plans that show the same phenotype class. As observed in the previous qPCR analysis of the other transformants, the relative expression ratio of the endogenous DEK1 (Calp/PP2A) is also inconsistent in the different lines. Taken together, these results suggest that there is no correlation between the expression level of the transgene and the phenotypes observed in AtDEK1-MEM-ZmLoop overexpressing plants (Figure 20).

4. Discussion

Arabidopsis thaliana has been recommended as a model plant for functional studies of DEK1 due to the similarities in dek1 mutant phenotypes with maize, including lack of the endosperm aleurone layer and arrest at early embryo development. The DEK1 protein is remarkably conserved throughout the plant kingdom, and it is, therefore, tempting to speculate that it may interact with the same protein, or family of proteins in different species. It is proposed that DEK1 functions to convey positional information to individual cells by sensing their developmental context in a tissue, probably mediated by DEK1's proposed extra cellular domain (Becraft et al., 2002; Lid et al., 2002 & 2005; Johnson et al., 2005; Tian et al., 2007). This extracellular domain, also called Loop, is the least conserved part of the DEK1 protein in comparison with the membrane domains and the intracellular calpain and arm domains. It has been postulated that DEK1-LOOP function to conveying positional information from the extracellular environment to the biologically active DEK1-Calpain domain. This hypothesis is supported by the dominant negative phenotype shown by *Arabidopsis thaliana* seedlings overexpressing only the membrane part (AtDEK1-MEM) of DEK1 (Tian et al., 2007). In this study we try to investigate whether the DEK1-LOOP interacts as proposed with the same positional cues, substance/s, protein/s, or ligands in species separated by millions of years of evolution. In addition, we try to examine whether the expression levels of transgene AtDEK1-MEM transcript can be correlated to the pleiotropic nature of the dominant negative phenotype.

4.1 Relative expression levels of AtDEK1-MEM transcript could not be correlated to pleiotropic phenotypes

The dominant negative phenotype observed in AtDEK1-MEM overexpressing transformants led to the hypothesis that the Loop domain of DEK1 may interact with either a ligand or other membrane proteins, and function in the initial activation of DEK1 (Tian et al., 2007). In this model, the overexpressed truncated protein competes with the native DEK1 for DEK1 activators and/or interactors, and since it lacks the cytoplasmic calpain domain (which is the biologically active domain), it cannot trigger the downstream signaling pathways, leading to

a dominant negative phenotype. This hypothesis was supported by three lines of evidence: 1) the observation that the overexpression of AtDEK1-MEM under the weaker AtDEK1 promoter failed to give a detectable phenotype (Tian et al., 2007), 2) similar phenotypes were observed in AtDEK1-RNAi transformants where DEK1 is down-regulated (Johnson et al., 2005; Tian et al., 2007), and 3) the observations that plants overexpressing AtDEK1-MEM-DEL, which lack the loop region, show a phenotype indistinguishable from wild type plants. These evidences strengthened the speculation that the extracellular loop domain is essential for either the perception and/or the transmission of important signals and that the observed dominant negative phenotype is not a pleiotropic effect caused by ectopic expression of a large membrane protein (Tian et al., 2007). Based on these evidences and hypothesis, we decided to use the AtDEK1-MEM experimental system to investigate cross-species functionality analysis of the extracellular loop. Our hypothesis was that if the replaced loop domain from one species is able to functionally replace *Arabidopsis* DEK1-LOOP, the transformation should result in plants with the same dominant negative phenotype as observed in AtDEK1-MEM over-expressing plants.

In this study a total of 80 independent transgenic AtDEK1-MEM lines were generated and studied. Of these 30 showed phenotypes undistinguishable from wild type and the rest demonstrated certain degrees of variation in phenotypic abnormalities like defects in the shoot apical meristem, radialized rosette leaves, and most frequently defects on the epidermal cell layer of the cotyledons. Our results are consistent with results from the initial experiment performed by Tian et al (2007), but they also raise few important questions which have not been addressed earlier; 1) why as many as 38 % of the transformants expressing *AtDEK1-MEM* do not show a phenotype and are indistinguishable from wild type plants, 2) what is the reason for the phenotypic variation in abnormalities among the independent transgenic lines, and 3) why as many as 36 % of plants with abnormal phenotype managed to recover, attaining a wild-type appearance later in development, after being transferred from selective medium to soil.

To be able to give possible and reasonable answers to these questions, one should take under consideration all possible factors that may influence the expression of the transgene. For instance, the promoter and the selective marker used in the experiment. Since both the nptII-gene (the selective maker gene providing kanamycin resistance) and the *AtDEK1-MEM*-construct are each controlled by the 35S promoter one possible explanation for the variation in phenotypes could be that the 35S promoter controlling *AtDEK1-MEM* expression was

somehow silenced in transgenic lines showing the wild type phenotype and in plants recovering to a apparent normal phenotype at a later developmental stage. That transgene expression is subjected to silencing, which may also first become manifested later in plant development, is a well-known and common phenomenon, and has also been suggested to explain the apparent reversion to normal wild type phenotype in AtDEK1-RNAi lines (Johnson et al., 2005). We attempted to investigate whether the *AtDEK1-MEM* expression was silenced in two weeks old seedlings showing the wild type phenotype and in more mature plants (at the age of 5 weeks), by measuring the steady-state level of the transgene transcript by qRT-PCR analysis. We were also interested to examine if expression levels of the transgene construct could be involved in the sterility observed in about 37 % of the transgenic lines. The qPCR analyses did show that the construct was overexpressed in both two weeks old seedlings and in five weeks old plants. The construct was highly expressed in some five weeks old plants with no significant differences in expression between fertile and sterile plants. Unfortunately the expression measurements at those developmental stages did not reveal any specific pattern of expression that could be correlated with the observed variation in phenotypes. However, the qRT-PCR analyses showed that the expression of the transgene is definitely not silenced in plans that do not exert developmental abnormalities (wild type phenotype). These observations may suggest a hypothesis that developmental stage specific timing of transgene expression might be important for the dominant negative effect of the *AtDEK1-MEM*-construct. According to this model, expression of the *AtDEK1-MEM* during later developmental stages alone would not have any effects on phenotype.

Because it is postulated that the observed phenotypes are result of negative dominant effect one could expect low expression levels of the transgene in plants not showing phenotype (or wild type phenotype) and elevated expression levels in the seedlings exerting more severe phenotypes. Johnson et al (2005) performed semi-quantitative PCR of plants overexpressing the *AtDEK1-RNAi* construct and reported the *AtDEK1* expression was visibly reduced in seedlings showing the most severe phenotypes compared to wild type, while in transformants showing weaker or wild type phenotypes, no significant decrease in *AtDEK1* expression was observed (Johnson et al., 2005). Q-PCR analysis of DEK1 expression in homozygous dek1-3 lines complemented with pRPS5A:AtDEK1-GFP, which showed a range of phenotypes from similar to AtDEK1-RNAi knockdown plants to wild-type in appearance, showed that levels of DEK1-GFP transcript and fusion protein expression were variable and correlated with the phenotypes observed (Johnson et al., 2008). As already

mentioned, the qPCR analysis of the steady-state level of transgene transcripts obtained in this study did not appear to correlate to the phenotypic pleiotropy observed in AtDEK1-MEM transformants or to fertility verse sterility in grown plants.

It is important to be aware that the qPCR analyses only provide information of how well our transgene is expressed in form of mRNA and not how effectively this transcript is translated into protein or how this protein is further treated in the cells. We do not know if some kind of repair or repression system are involved and somehow regulates or demolishes the expressed truncated protein. The qPCR analyses showing that transformants with wild type appearances expressed the transgene at very high levels, suggest that AtDEK1-MEM could be regulated post-transcriptionally or that could be a subject to post-transcriptional transgene silencing (PTGS) specifically at the translation level. It is well known that transgenes expressed at high levels are subject to PTGS (De Wilde et al., 2000, and references therein). In order to further investigate this case Western blotting to detect transgene protein product must be performed.

There may be few important factors that can also explain why the qPCR analysis did not appear to correlate to the observed phenotypes as expected. In order to investigate if there exists a relationship between the level of transgene expression and phenotype observed it is important that the tissue collected from the samples are harvested from plants at similar developmental stage and that the collected tissue were comparable. This was a challenge due to the severe developmental defects in some of the analyzed plants. With the intention of allowing seedlings growth and to collect sample from identical tissue, a cotyledon was the only option at this stage. Some of the tissue sampled was not suitable for Q-PCR analysis because of poor or bad RNA isolation, most probably because of the small size/weight of the sample and unhealthy looking tissue (prerequisite for good RNA isolation results). The tissue samples gathered from plants showing the fused cotyledon phenotype most probably also contained meristem cells, making this samples inequitable to compare with samples gathered from only one cotyledon. And even though the quest was to gather samples from plants at the same developmental stage, we cannot firmly guaranty that seedlings at the same age (two weeks) were at the same developmental stage; some seedlings were developing more slowly then others in comparison to wild type control, and others had arrested development i.e. because of lack of a shoot apical meristem. The time points at which the samples were taken should also be considered, and the small window they represents. We can not assume that detection of transgene transcripts at these two time points (two and five

weeks) provides information about expression during early stages of development that may be more important for the onset of the dominant negative effect.

Another very important factor in qPCR analysis is the reference gene of choice, which is used in the calculations of the relative expression ratios. The expression of the reference gene should not be altered or effected in any way by the target gene or growth conditions used (Thellin et al., 1999; Schmittgen et al., 2000; Tong et al., 2009). In this study we used protein phospatase 2A as reference gene (Czechowski et al., 2005), which is not supposed to be affected by the overexpression of the target gene, but this is not certain. PP2A is a serine/threonine-specific Ca^{2+}-independent protein phosphatase, which is expressed ubiquitously in eukaryotic cells, and is an important component in the regulation of signal transduction, in the reversible protein phosphorilation events in plants and in the control of cell metabolism (Rundle et al., 1995; Terolet al., 2002). The PP2A gene have shown superior expression stability in addition to low absolute expression levels, which makes this gene a good candidate for quantification of transcript levels in *Arabidopsis* (Czechowski et al., 2005)), especially when analyzing low-abundant transcripts, such as DEK1. Recently, partial degradation of PP2A by m-calpain was reported in apoptotic cells (Janssens et al., 2009) In addition, it was suggested that PP2A is involved in calpain-mediated FoxO regulation and that PP2A regulatory subunits B56 alpha and gamma are *in vitro* substrates of calpain. The calpain regulates B56 alpha stability *in vivo*, suggesting a direct role of calpain in the regulation of PP2A function in animal cells (Bertoli et al., 2009). This is not necessary the situation in plant systems, but it should be considered, especially because we do not know if the endogenous DEK1 protein and therefore the calpain are affected by the incretion and expression of the transgene in analyzed plants. Probably the use of PP2A gene as reference gene to analyze the DEK1 transcript levels should be reconsidered. Even though our analyses are performed at the transcript level and calpain regulation of PP2A reported in this literature occurs at protein level.

The fact that as many as 36 % of the transgenic plants with developmental defects recovered to wild-type appearance after being transferred to soil, and that the phenotypes were observed mainly in seedlings growing on selective medium, inflate a suspicion of a probable toxic effect of kanamycin on the seedlings, or silencing of the ntpII-gene promoter in seedlings showing phenotypes or morphological abnormalities. The transformants are not under kanamycin selection before the germination stage, and even here, there is some window when it comes to the timing of ntpII expression and the ability of the seedlings to

survive on kanamycin. If this expression was silenced or not “turned on” on time (before the root reaches the selective medium) a toxic effect could result in a phenotype. In a collection of 111 transgenic *Arabidopsis thaliana* lines, silencing of the nptII gene was observed in 56 % of the lines and three distinct nptII-silencing phenotypes were identified (Meza et al., 2001). However, the phenotypes described in the nptII-silenced lines were not comparable with the phenotypes observed in our study. A parallel qPCR analysis of ntpII gene expression levels should be performed to rule out this possibility.

The *DEK1* gene has been suggested to be important for multiple developmental processes (Becraft et al., 2002; Johnson et al., 2008), and the observed phenotypic pleiotrophy may simply be a consequence of disturbance of these processes. Depending on which process was first affected by the lack of DEK1 activity, this could lead to several different malfunctions manifested by a variation in abnormalities observed. Again, this could be explained if developmental stage specific timing of transgene expression is important for the onset of the dominant negative effect shown in the AtDEK1-MEM transformants.

4.2 Expression of AtDEK1-MEM where the Loop domain is deleted only alleviates the dominant negative effect

The *AtDEK1-MEM-DEL* construct, which is identical to the *AtDEK1-MEM*-construct except that it lacks the Loop domain, was designed and transformed into *Arabidopsis* by Tian et al. (2007) to investigate the possibility that the phenotypes observed in plants overexpressing *AtDEK1-MEM* do not represent a pleiotropic effect caused by ectopic expression of a large membrane protein. All of the 51 AtDEK1-MEM-DEL T1 lines obtained by Tian et al. (2007) were indistinguishable from wild-type control plants, suggesting that the AtDEK1-Loop domain is essential for either the perception and/or the transmission of positional signals.

In our study, the insertion of the *AtDEK1-MEM-DEL*-construct in wild type background was used as a control to generate a baseline for comparing the dominant negative phenotype of plants overexpressing *AtDEK1-MEM* construct to those harboring the cross-species replaced loop construct. We obtained a total of 33 independent transgenic lines of which 27 % showed phenotypes similar to those observed in plants overexpressing AtDEK1-MEM and to those overexpressing the AtDEK1-RNAi construct (Johnson et al., 2005; Tian et al.,

2007). The variation in abnormalities was most frequently associated with damaged epidermal cells on the cotyledons combined with either missing or defective shoot apical meristem or the production of only one single first true leaf. Our results are inconsistent with the results obtained in the initial study performed by Tian et al. (2007), and depict that the deletion of the loop was not enough to eliminate the dominant negative effect, but it was enough to alleviate it, considering the lower phenotipic frequency emerged in the transformed lines.

These unexpected results bring back the question of whether the phenotypes of AtDEK1-MEM overexpressing seedlings are the result of ectopic expression of a large membrane protein. The phenotypes of AtDEK1-MEM-DEL are similar and comparable to AtDEK1-MEM overexpressing seedlings, except that they are observed in significantly lower frequency. In addition, the phenotypes observed resemble the down regulation of DEK1s function (Johnson et al., 2005) and not phenotypes of seedling overexpressing the DEK1 protein (Lid et al 2005). Considering the much lower frequency of phenotypes observed in AtDEK1-MEM-DEL transgenic lines and their nature, it is quite unlikely that they are result of pleiotropic effect of a large membrane protein. However, these conflicting results are difficult to interpret based on the present data.

One reasonable explanation for the unexpected observed phenotypes in the AtDEK1-MEM-DEL lines could be the presence of an important binding domain on the membrane part of the protein, which is only partially affected by the removal of the Loop, perhaps managing to bind reversibly or moderately to the proposed ligand, substrate, protein or positional cues. If this could explain the situation we would expect to see a higher expression levels of the transgene in seedling showing a phenotype. However the results of qPCR analysis do not support this hypothesis. As discussed earlier, we have to keep in mind the reliability of the qPCR analysis given the difficulties of obtaining tissue at the same developmental stage and the possibility that the transgene could be post-transcriptionaly regulated. One way to assess this question is to transform *Arabidopsis* with AtDEK1-MEM construct containing larger deletion region, or a series of deletions.

We did try to further investigate the dominant negative effect by co-transforming the pSEL1:35S-AMG (Tian et al., 2007) and the pRPS5A:CALPAIN-GFP (which was able to complement dek1 mutants, Johnson et al., 2008) constructs in *Arabidopsis*. If the observed phenotypes in AtDEK1-MEM transgenic plants are the result of dominant negative effect,

the insertion of an active calpain protein should in theory alleviate this effect. Seeds from co-transformed plants were germinated on M&S medium supplemented with both kanamycin and hygromycin, for selection of both constructs. However, no resistant progeny was identified. Putative transgenic seeds were then sowed on M&S medium with only kanamycin to escape probable toxic effect of hygromycin, still no resistant progeny was detected. In both experiments, some seedlings that were distinguishable from others, by their greener appearance, where detected even after three weeks of growth, suggesting antibiotics resistance, but their development was clearly arrested and they did not produce any other organs beyond the cotyledons (Figure 5 A2 in Appendix 2). Since this phenomenon was not observed in the control plates (wild type plants), there was a suspicion that expression of both constructs could possible result in a lethal phenotype. To investigate this possibility, seeds were sown on unselective medium, where all seed germinated and grow up normally showing that no lethal phenotype was detected. This result led to the conclusion that the co-transformation was most probably unsuccessful. This experiment should be carried out and completed in the further work. In fact, confirming or invalidating the dominant negative phenotype observed in AtDEK1-MEM overexpressing seedlings by any approach should be priority number one in further work.

Some small and maybe important differences in growth conditions between the initial experiment performed by Tian et al (2007) and in this study should be mentioned. It is well known that environmental conditions influence *Arabidopsis* phenotypes, and could explain the discrepancy obtained in our study and that of Tian et al. (2007). In the initial experiment, the *Arabidopsis* plants were grown in growth chambers where all conditions are constant and easily controlled. In our study the plants were grown in a plant room, where some temperature and humidity fluctuation has been observed, however this temperature fluctuation was never more then ± 2°C, which should not be of that large influence on germinating seeds or seedlings growing in closed petri-dishes. Maybe more important is the difference in light and dark periods used in the experiments; Tian et al (2007) were growing plants for 8 h of dark and 16 h of light, while in our study 18 h of light and 6 h dark period was used as growth conditions. If this factor infuenced phenotype frequencies in the AtDEK1-MEM-DEL lines, it should be considered to also influence the other transgenic lines as well.

Another important observation regarding the AtDEK1-MEM-DEL experiment is the relatively low number of transformants achieved (33 independent lines), despite the three transformation attempts made. The low number of generated transformants could simply be a result of low transformation efficiency, but could also be a consequence of a lethal phenotype. If this is the case, it is more probable that the observed phenotypes in plans overexpressing both the AtDEK1-MEM and AtDEK1-MEM-DEL constructs are a result of a pleiotropic effect of large membrane protein. Higher number of AtDEK1-MEM-DEL transformants should be generated and investigated, and the T1 generation should be thoroughly inspected for defective T1 seeds in order to assess this question.

4.3 Plants overexpressing the DEK1-LOOP domain from the single celled algae *Mesostigma veride* exert similar phenotype frequencies as AtDEK1-MEM plants

To examine species-specific interaction features of the DEK1-LOOP domain, a construct was designed where the Loop domain of AtDEK1 in the AtDEK1-MEM-construct was swapped with the Loop domain from the single-celled alga *Mesostigma viride*, and transformed into wild type *Arabidopsis* background. *Arabidopsis* and *Mesostigma* are separated by approximately 725-1200 million years of evolution according to different estimates by molecular clock methods (Hedges et al., 2004; Yoon et al., 2004; Zimmer et al., 2007). A pairwise alignment of *Arabidopsis* and *Mesostigma* DEK1-LOOP sequences revealed 16 % identity, which is relatively low. However, multiple sequence alignment of DEK1-LOOP homologs revealed that the *Mesostigma* sequence was conserved in regions where the sequences of more derived plants are well conserved (Figure 10). Further bioinformatics analyses revealed several putative conserved phosphorylation sites and one myristoylation site shared by both sequences, which could be important for DEK1-LOOP function. One of the putative phosphorylation sites was for Casein kinase II. Phosphorylation by the protein kinase CK2 promotes calpain-mediated degradation of I-kappa B-alpha in B cells (Shen et al., 2001). Myristoylation plays a vital role in membrane targeting and in signal transduction in plant, in responses to environmental stress (Podell & Gribskov, 2004). These predicted sites may be important and could be interesting for further investigation.

The cross-species domain swap experiments were build on the prospect that the phenotypes observed in AtDEK1-MEM transformants are the result of dominant negative phenotype, in witch the Loop plays an essential role (Tian et al., 2007). Our results obtained from the AtDEK1-MEM-DEL transformation (discussed earlier), suggest that the DEK1-LOOP alone may not be the only important component for the dominant negative phenotype, as previously proposed. Based on the unexpected phenotype that was observed in the control transformants with AtDEK1-MEM-DEL and the consequent complications they infer, it is difficult to make any conclusions to the results achieved from the swapped domain experiments.

A total of 46 independent transgenic lines were obtained and studied from transformation with the AtDEK1-MEM-MvLoop construct. Of these, 63 % showed variation of seedling phenotypes similar or undistinguishable from the dominant negative phenotypes observed in seedling overexpressing the *AtDEK1-MEM* construct. The phenotypic frequencies observed in seedling overexpressing the *AtDEK1-MEM* construct with *Mesostigma viride* swapped Loop domain are similar to the frequencies observed in *AtDEK1-MEM* transgenic lines and significantly different from those in *AtDEK1-MEM-DEL* overexpressing plants, where the AtDEK1-LOOP is deleted. These results may suggest that despite the low sequence identity (16 %) and the millions of years of evolution that separate these two species, the DEK1-LOOP from *Mesostigma viride* is able to functionally replace the *Arabidopsis* DEK1-LOOP sequence. This proposal further implies that there could be well-conserved functional sites or motifs in the MvDEK1-LOOP sequence. These functional sites most probably interacts with the same positional cues or the same family of proteins as in the AtDEK1-LOOP domain, and are probably essential for conveying information to the intracellular calpain domain. It is important to emphasize that the proposal that the MvLOOP can functionally replace the AtLOOP domain only holds true if the observed phenotypes are the result of dominant negative effect, caused by the Loop domain.

4.4 The replacement of *Arabidopsis*-dek1-Loop with maize-dek1-Loop also alleviates the dominant negative effect

Surprisingly, the phenotype frequencies of seedlings overexpressing the AtDEK1-MEM-ZmLoop construct, in which the AtDEK1-LOOP domain is replaced with dek1-loop domain

from maize, resembled the frequencies observed in seedlings overexpressing the AtDEK1-MEM-DEL construct, where the loop domain is deleted. In total 37 independent transgenic lines were obtained, of which 38 % showed phenotypic alternations similar to those described for the other transformants obtained in this study. The performed statistical test indicated significant differences in the observed phenotype frequencies between the AtDEK1-MEM overexpressing seedling and the AtDEK1-MEM-ZmLoop transformants, and no significant difference with the frequencies observed in AtDEK1-MEM-DEL transgenic lines. These results suggest that there may exist functional species-specific divergence between *Arabidopsis* and maize DEK1-Loops, or that the ZmDEK1-LOOP interacts with different substances, proteins or positional cues, since it could not compete effectively with the endogenous AtDEK1-LOOP. This result is in contrast to the result achieved from the MvLoop swap experiment where the LOOP sequences have actually diverged considerably. Despite the much higher sequence identity and closer evolutionary ancestor compared with *Mesostigma viride*, the DEK1-LOOP domain in maize seems to have a different function than the AtDEK1-LOOP according to the results obtained in this study.

The hypothesis that the *DEK1-LOOP* function has diverged between *Arabidopsis* and maize is also supported by the similarity tree build based on the predicted AtDEK1-LOOP homologues, suggesting that the sequence diverged between monocots and dicots from common ancestor. Another evidence supporting different functional roles for DEK1 in maize and *Arabidopsis* is the reported differences in maize and *Arabidopsis* double dek1/cr4 mutants. In contrast to CR4, which is a critically important protein for aleurone and embryo development in maize, null mutation in ACR4, one of five CR4-like RLKs (receptor-like kinases) in *Arabidopsis*, result in only a very weak null mutant phenotype characterized by disorganization of ovule integuments and sepal margins as well as weak epidermal defects. Strong *dek1* alleles are epistatic to cr4, and the phenotype in maize dek1/cr4 double mutant suggests that the two gene products function in partially overlapping pathways (Becraft et al., 2002). In contrast, the *Arabidopsis thaliana* homologues of crinkly 4 (acr4)/dek1 double mutants show additive effects, suggesting that dek1 and acr4 act in different pathways both involved in epidermis specification (Johnson et al., 2005). Complementation experiments in *Arabidopsis* dek1-1 mutant with ZmDEK1-CALP showed that although DEK1-CALP improved the growth of both endosperm and embryos of homozygous dek1-1 seeds, a full complementation of the aleurone and embryo phenotypes to wild type was not achieved

(Tian et al., 2007). DEK1 in maize could also be localized differently to DEK1 in *Arabidopsis*, since Tian et al. (2007) detected maize DEK1 largely in FM4-64-labeled endosomal compartments whereas Johnson et al. (2008) did not find this to be the case for DEK1-GFP fusion proteins in *Arabidopsis thaliana*, which where only weakly associated with BFA bodies in roots (Tian et al., 2007; Johnson et al., 2008).

Finally, it should be clear that all the discussion and speculations regarding the loop-domain-swap results are based on the assumption that the phenotypes observed in AtDEK1-MEM transformants are the result of dominant negative effect caused by the loop domain. This effect has not been validated.

4.5 Further work

4.5.1 Verification of the dominant negative phenotype

In this study we used transgenic plants transformed with the AtDek1-MEM-DEL construct as control or as a baseline. In the initial study performed by Tian et al. (2007) the AtDek1-MEM-DEL was designed and transformed into wild type background with the intention to show that the observed dominant negative phenotype observed in AtDEK1-MEM transformants was not a pleiotropic effect caused be the ectopic expression of large membrane protein. All the AtDEK1-MEM-DEL transgenic lines described by Tian et al., (2007) where undistinguishable from wild type, confirming that the loop domain is important for the dominant negative phenotype observed in AtDEK1-MEM transgenic lines (Tian et al., 2007). In the study presented in this thesis an unexpected phenotype was observed in 27 % of transformed seedlings overexpressing the AtDek1-MEM-DEL construct complicating the experimental system chosen for the cross-species domain swap analyses. There are two ways to clear this out:

1) Verification of the dominant negative effect in AtDEK1-MEM transformants. The dominant negative phenotype should be confirmed before any further work with cross-species swap-domain experiments can be done. This can be done by few approaches, for instance by co-transformation of wild type plants with the AtDek1-MEM construct and Capain construct under the promoter RPS5A, which is sufficient to complement dek1-3-

mutants (Johnson et al., 2008). This experiment was performed once in this study, but unfortunately resulted in no transformatns. Alternatively, the negative dominant effect can also be assessed by a genetic cross between a transgenic AtDek1-MEM line showing a phenotype with plant overexpressing the Calpain domain.

2) Investigation of the phenotypes observed in AtDEK1-MEM-DEL expressing plants. As discussed and proposed earlier, there could be a motif on the dek1 membrane sequence, just outside of the loop-coding region, that is involved in signal transduction. The expression of an AtDEK1-MEM construct with deletion beyond the Loop sequence or a series of deletions can help to clear this out. If this is the case, the chosen system for cross-species domain swap experiments has to be optimized according to the obtained results.

4.5.2 Validation of the qPCR analysis

The relative expression ratios of the transgene transcripts, performed by qPCR analyses in this study, could not be correlated to the phenotypic pleiotropy observed in all transformants. Earlier in the discussion we proposed that the transgene could be post-transcriptionaly regulated and that Western blotting assay can elucidate this. We also suspect that the reference gene used in the qPCR analysis can be affected by calpain expression. Since it was shown that the PP2A gene is directly involved in calpain regulation in animal cells (Xu & Deng, 2006; Jonssen et al., 2009; Bertoli et al., 2009), and as discussed earlier the PP2A gene could be inappropriate as reference gene in these experiments. Jonson et al. (2005) and Lid et al. (2005) have used the EIF2 (EUKARYOTIC TRANSLATION INITIATION FACTOR2) gene, which is considered to be constitutively expressed in *Arabidopsis* (Metz et al., 1992) in qPCR analysis of dek1 and the calpain transcript expression. Since they have achieved conclusive results, maybe this gene could be used in our experimental system as well, as long as the transgene expression is not post-transcriptionally regulated.

4.5.3 Further evolutional and cross-species domain swap analysis

As mentioned earlier it is essential that the system used for cross-species domain swap analyses is the correct one for these experiments. The results obtained in this study put this

in question, due to an unexpected phenotype shown in control AtDEK1-MEM-DEL transformants.

The results obtained from our and other experiments studies suggest that DEK1 may localize and act slightly different in *Arabidopsis* and maize, which can also lead to a suspicion that DEK1 may have slightly divergent function in monocots and dicots. In order to be more confident in such speculations, more independent transgenetic lines with both the AtDek1-MEM-ZmLoop and AtDEK1-MEM-MvLoop have to be obtained to confirm the observed statistical significant differences in the calculated frequencies. It will be quite interesting to investigate if there is functional divergence in DEK1 between monocots and dicots, which means that cross-species domain-swap experiments with different representatives of monocots and dicots have to be performed, when a proper system for these experiments is confirmed or developed.

The obtained results also bring up a curiosity about what to expect from the replacement of the AtDEK1-LOOP with the loop of the moss *Phiscomitrella patens*, which was initially one of the goals in this study. Unfortunately, due to difficulties in amplifying the Dek1-Loop from *Phiscomitrella patens*, the presence of *Pst*I restriction site on the loop sequence, the presence of unexpected second *Afe*I restriction site on the pSEL1:35S-AMG, and the time limitation of the project, *Arabidopsis* transformats with Pp-replaced-loop were not achieved. The work has come to the point where the construct can be cloned into the binary vector pSEL1:35S-AMG. Prior to this reaction the *Pst*I restriction site on the Pp-loop sequence have to be mutated as depicted in materials and methods or a new infusion cloning approach can be used. Primers for the site-directed mutagenesis and for the infusion cloning are designed and stored at Biohus, Hamar.

5. Conclusion

Utilizing the dominant negative effect in plants overexpressing the AtDEK1-MEM construct in *Arabidopsis thaliana* to study cross-species functional analysis in the DEK1-LOOP domain, is an easy and fast approach. However, in this study we met complications; unexpected phenotype emerged at low frequency in the control transformants overexpressing the AtDEK1-MEM-DEL construct, where the loop domain is deleted, inconsistent with results obtained in the initial experiment performed by Tian et al. (2007). This result compromised the experimental system chosen for the cross-species domain-swap analysis. The attempt to elucidate the dominant negative effect in a co-transformation experiment, ended with no transformats, leaving this bottleneck for the further work.

Two constructs harboring swapped cross-species DEK1-LOOP domains were designed and successfully cloned and transformed in wild type *Arabidopsis thaliana*. Phenotypic analysis of the obtained independent transgenic lines suggests that the DEK1-LOOP from *Mesostigma viride* was able to functionally replace the DEK1-LOOP in *Arabidopsis thaliana*. Surprisingly the replacement of AtDEK1-LOOP with DEK1-LOOP from maize suggests divergent function, since the observed phenotypic frequencies resemble the frequencies of transformants with deleted dek1-loop. However, before it is validated that there is a negative dominant effect exerted by the loop domain, it is difficult to conclude anything from the results obtained in the cross-species domain swap experiments.

Analysis of the expression levels of transgenic transcript by qRT-PCR confirmed that the inserted constructs were overexpressed in transgenic lines, however the expression ratios could not be correlated to phenotypic variation observed. In addition, high expression levels of the transgene in wild type appearing plants suggest that AtDEK1-MEM may be regulated post-transcriptionaly.

References

Ahn J-W, Kim M, Lim JH, Kim G-T, Pai H-S (2004) Phytocalpain controls the proliferation and differentiation fates of cells in plant organ development. Plant J 38:969–981

The Arabidopsis Genome Initiative (2000), Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*, *Nature* 408, 796-815 | doi:10.1038/35048692

Becker B & Marin B (2009) Streptophyte algae and the origin of embryophytes, Annals of Botany 103: 999–1004, doi:10.1093/aob/mcp044

Becraft PW, Asuncion-Crabb Y (2000) Positional cues specify and maintain aleurone cell fate in maize endosperm development. Development 127: 4039–4048

Becraft PW, Stinard PS, McCarthy D (1996) CRINKLY4: a TNFR-like receptor kinase involved in maize epidermal differ- entiation. Science 273:1406–1409

Becraft PW, Li K, Dey N, Asuncion-Crabb Y (2002) The maize dek1 gene functions in embryonic pattern formation and cell fate specification. Development 129: 5217–5225

Becraft B & Yi G (2010) Regulation of aleurone development in cereal grains. Journal of Experimental Botany, 1-7, doi:10.1093/jxb/erq372

Bertoli C, Copetti T, Lam EW, Demarchi F & Schneider C (2009) Calpain small-1 modulates Akt/FoxO3A signaling and apoptosis through PP2A. Oncogene 28, 721–733

Bethke PC, Libourel IGL, Aoyama N, Chung Y-Y, Still DW, Jones RL (2007). The Arabidopsis aleurone layer responds to nitric oxide, gibberellin, and abscisic acid and is sufficient and necessary for seed dormancy. Plant Physiology 143, 1173–1188

Buck De S, Jacobs A, Van Montagu M, Depicker A (1998) Agrobacterium tumefaciens Transformation and Cotransformation Frequencies of Arabidopsis thaliana Root Explants and Tobacco Protoplasts, MPMI Vol. 11, No. 6, pp. 449–457. Publication no. M-1998-0316-01R

Cheng J, Randall A, Sweredoski M, Baldi P (2005) SCRATCH: a Protein Structure and Structural Feature Prediction Server, Nucleic Acids Research, Web Server Issue vol. 33, 72-76

Clough SJ, Bent AF (1998) Floral dip: a simplified method for Agrobacterium-mediated transformation of Arabidopsis. Plant J 16: 735–743

Czechowski T, Stitt M, Altmann T, Udvardi MK, Scheible W-R (2005). Genome-Wide Identification and Testing of Superior Reference Genes for Transcript Normalization in Arabidopsis, Plant Physiology 139:5-17

Edgar RC and Batzoglou S (2006) Multiple sequence alignment. Curr. Opin. Struct. Biol., 16, 368-373

Eriani G, Dirheimer G, Gangloff J. 1989 Isolation and characterization of the gene coding for Escherichia coli arginyl-tRNA synthetase, Nucleic Acids Res.17(14):5725-36.

Deblaere R, Bytebier B, De Greve H, Deboeck F, Schell J, Van Montagu M and Leemans J (1985) Efficient octopine Ti plasmid-derived vectors for *Agrobacterium*-mediated gene transfer to plants. Nucl. Acids Res. 13 (13): 4777-4788.doi: 10.1093/nar/13.13.4777

De Wilde C, Van Houdt H, De Buck S, Angenon G, De Jaeger G and Depicker A (2000) Plants as bioreactors for protein production: avoiding the problem of transgene silencing, Plant Molecular Biology 43: 347–359

Drickamer KM & Taylor E (2006). *Introduction to Glycobiology* (2nd ed.) Oxford University Press, USA. ISBN 978-0199282784

Du Verle D, Takigawa I, Ono Y, Sorimachi H, Mamitsuka H (2009) CaMPDB: a Resource for Calpain and Modulatory Proteolysis, *Genome Informatics* 22: 202-214

Goll DE, Thompson VF, Li H, Wei W, Cong J (2003) The calpain system. Physiol. Rev. 83, 731–801

Gruis D, Guo H, Selinger D, Tian Q, Olsen O-A (2006) Surface position, not signaling from surrounding maternal tissues, specifies aleurone epidermal cell fate in maize. Plant Physiology 141, 898–909

Hedges SB, Blair JE, Venturi ML, Shoe JL (2004) A molecular timescale of eukaryote evolution and the rise of complex multicellular life, BMC Evolutionary Biology 4: 2, doi:10.1186/1471-2148-4-2

Herskowitz I (1987) Functional inactivation of genes by dominant negative mutations. Nature 329:219–222

Hibara K, Obara M, Hayashida E, Abe M, Ishimaru T, Satoh H, Itoh J, Nagato Y (2009) The ADAXIALIZED LEAF1 gene functions in leaf and embryonic pattern formation in rice. Developmental Biology 334, 345–354

Johnson KL, Degnan KA, Walker JR, Ingram GC. (2005) AtDEK1 is essential for specification of embryonic epidermal cell fate. Plant Journal 44: 114–127

Johnson KL, Faulkner C, Jeffree CE, Ingram GC. (2008) The phytocalpain defective kernel 1 is a novel Arabidopsis growth regulator whose activity is regulated by proteolytic processing. Plant Cell 20: 2619–2630

Jones DG (1920) Heritable characters of maize IV. A lethal factor-defective seeds. Heredity 11**:** 161-167

Koch MA, Haubold B, Mitchell-Olds T (2000). Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in Arabidopsis, Arabis, and related genera (Brassicaceae). Mol. Biol. Evol. 17: 1483–1498

Krogh A, Larsson B, von Heijne G, Sonnhammer EL (2001) Predicting transmembrane protein topology with a hidden Markov model: application to complete genomes. *J Mol Biol 305*:567-580

Lid SE, Al RH, Krekling T, Meeley RB, Ranch J, Opsahl-Ferstad HG, and Olsen OA (2004) The maize disorganized aleurone layer 1 and 2 (dil1, dil2) mutants lack control of the mitotic division plane in the aleurone layer of developing endosperm. Planta 218, 370–378

Lid SE, Gruis D, Jung R, Lorentzen JA, Ananiev E, Chamberlin M, Niu X, Meeley R, Nichols S, Olsen OA (2002). The defective kernel 1 (dek1) gene required for aleurone cell development in the endosperm of maize grains encodes a membrane protein of the calpain gene superfamily. Proceedings of the National Academy of Sciences, USA 99: 5460–5465

Lid SE, Olsen L, Nestestog R, Aukerman M, Brown RC, Lemmon B, Mucha M, Opsahl-Sorteberg HG, Olsen OA. (2005). Mutation in the Arabidopsis thaliana DEK1 calpain gene perturbs endosperm and embryo development while over expression affects organ globally. Planta 221: 339–351

Mangelsdorf PC (1923) The inheritance of defective seeds in maize. J. Heredity 14:119- 125

Mangelsdorf PC (1926) The genetics and morphology of some endosperm characters in maize. Conn. Agr. Exp. Sta. Bull. 279: 509-614

Margis R, Margis-Pinheiro M (2003) Phytocalpains: orthologous calcium-dependent proteinases. Trends Plant Sci 8:58–62

Martinez-Trujillo M, Limones-Briones V, Cabrera-Ponce J L and Herrera-Estrella L (2004) Improving Transformation Efficiency of *Arabidopsis thaliana* by Modifying the Floral Dip Method, Plant Molecular Biology Reporter 22: 63–70

McDonald JH (2009) Handbook of Biological Statistics, 2nd ed. Sparky House Publishing, Baltimore, Maryland, http://udel.edu/~mcdonald/statbigchart.html

Metz AM, Timmer RT and Browning KS (1992) Sequences for two cDNAs encoding Arabidopsis thaliana eukaryotic protein synthesis initiation factor 4A. Gene, 120, 313–314

Michener CD and Sokal RR (1957) A quantitative approach to a problem in classification, Evolution 11, 130-162

Meinke DW, Cherry JM, Dean C, Rounsley SD, Koornneef M (1998) *Arabidopsis thaliana:* A Model Plant for Genome Analysis, Science 282: 662-681

Meyerowitz EM (2001) Prehistory and History of *Arabidopsis* Research, *Plant Physiology*, Vol. 125, 15–19

Murashige T, Skoog T (1962) A revised medium for rapid growth and bioassays with tobacco tissue culture. Physiol Plant 15:473– 497

Neuffer MG, Johal G, Chang MT and Hake S (2009) Mutagenesis - the Key to Genetic Analysis. Maize handbook-Volume II: Genetics and Genomics. Springer Science Business Media LLC

Neuffer MG, Sheridan W, Bendbow E (1978) Maize Genet. Cooperation. NewsLett. 52, 84-88

Olsen OA (2001) Endosperm development: cellularization and cell fate specification. Annual Review of Plant Physiology and Plant Molecular Biology 52, 233–267.

Olsen O-A (2004) Dynamics of maize aleurone cell formation: the ‘surface’ rule. Maydica 49, 37-40

Ossowski S, Schneeberger K, Lucas-Lledo JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M (2010) The rate and molecular spectrum of spontaneous mutations in Arabidopsis thaliana. Science 327: 92–94

Perrin BJ and Huttenlocker A (2002) Molecules in focus: calpain. Int J Biochem Cell Biol 34: 722–725

Podell S, Gribskov M (2004) Predicting N-terminal myristoylation sites in plant proteins, BMC Genomics 17; 5(1):37, doi: 10.1186/1471-2164-5-37

Quatrano RS, McDaniel SF, Khandelwal A, Perroud PF, Cove DJ (2007) Physcomitrella patens: Mosses enter the genomic age. Curr. Opin. Plant Biol. 10: 182–189

Ramakers C, Ruijter JM, Lekanne Deprez RH, Moorman AFM. (2003) Assumption- free analysis of quantitative real-time PCR data. Neurosci Letters 339: 62-66

Reyes FC, Sun B, Guo H, Gruis D, Otegui MS. 2010. Agrobacterium tumefaciens-mediated transformation of maize endosperm as a tool to study endosperm cell biology. Plant Physiology 153, 624–631

Ruijter JM, Ramakers C, Hoogaars WM, Karlen Y, Bakker O, van den Hoff MJ, Moorman AF (2009). Amplification efficiency: linking baseline and bias in the analysis of quantitative PCR data. Nucleic Acids Research 37: 45

Sambrook J & Russell JS (1989) Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor, N.Y.

Sato K & Kawashima S (2001) Calpain function in the modulation of signal transduction molecules. Biol. Chem. 382, 743–751

Scanlon MJ, Stinard, PS, James MG, Myers AM and Robertson DS (1994) Genetic analysis of 63 mutations affecting maize kernel development isolated from *Mutator* stocks. *Genetics* 136, 281-294

Schmittgen TD & Zakrajsek BA (2000) Effect of experimental treatment on housekeeping gene expression: validation by real-time, quantitative RT-PCR, J. Biochem. Biophys. Methods 46, 69–81.

Shao H, Chou J, Baty CJ, Burke NA, Watkins SC, Stolz DB, and Wells A (2006) Spatial localization of m-calpain to the plasma membrane by phosphoinositide biphosphate binding during epidermal growth factor receptor-mediated activation. Mol. Cell. Biol. 26: 5481–5496

Shen B, Li C, Min Z, Meeley RB, Tarczynski MC, Olsen O-A (2003) Sal1 determines the number of aleurone cell layers in maize endosperm and encodes a class E vacuolar sorting protein. Proc Natl Acad Sci USA 100:6552–6557

Shen J, Channavajhala P, Seldin DC, Sonenshein GE (2001) Phosphorylation by the protein kinase CK2 promotes calpain-mediated degradation of IkappaBalpha, J Immunol. 167(9):4919-25

Sheppard D (1994), Dominant negative mutants: tools for the study of protein function in vitro and in vivo, Cell Mol. Biol., Vol 11, No. 1, 1-6

Sheridan WF, Neuffer MG (1980) Defective kernel mutants of maize. II. Morphological and embryo culture studies. Genetics 95:945–960

Sorimachi H & Suzuki K (2001) The Structure of Calpain, J.Biochem.129, 653-664

Tian Q, Olsen L, Sun B, Lid SE, Brown RC, Lemmon BE, Fosnes K, Gruis DF, Opsahl-Sorteberg HG, Otegui MS (2007) Subcellular localization and functional domain studies of DEFECTIVE KERNEL1 in maize and Arabidopsis suggest a model for aleurone cell fate specification involving CRINKLY4 and SUPERNUMERARY ALEURONE LAYER1, Plant Cell 19: 3127–3145

Thellin O, Zorzi W, Lakaye B, De Borman B, Coumans B, Hennen G, Grisar T, Igout A, Heinen E (1999) Housekeeping genes as internal standards: use and limits, J. Biotechnol. 75, 291–295

Tong Z, Gao F, Wang J, Zhou Z (2009) Selection of reliable reference genes for gene expression studies in peach using real-time PCR, BMC Mol. Biol. 10, 71

Van der Hoorn R AL (2008) Plant Proteases: From Phenotypes to Molecular Mechanisms, Annu. Rev. Plant Biol. 59:191–223

Veitia RA (2007) Exploring the Molecular Etiology of Dominant-Negative Mutations ,The Plant Cell, Vol. 19: 3843–3851

Wang C, Barry JK, Min Z, Tordsen G, Rao AG, Olsen O-A (2003) The calpain domain of the maize DEK1 protein contains the conserved catalytic triad and functions as a cysteine proteinase. J Biol Chem 278:34467–34474

Wass MN, Kelley LA, Sternberg MJ (2010) 3DLigandSite: predicting ligand-binding sites using similar structures. Nucleic Acids Res. 38: W469-73

Weigel D, Glazebrook J (2002) Arabidopsis: A Laboratory Manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp 19–39

Wurch T, Lestienne F, Pauwels PJ (1998) A modified overlap extension PCR method to create chimeric genes in the absence of restriction enzymes, Biotechnology Techniques, Vol 12, No 9, pp. 653–657

Wright S.I., Lauga B., Charlesworth D (2002) Rates and patterns of molecular evolution in inbred and outbred Arabidopsis. Mol. Biol. Evol. 19: 1407–1420

Yoon HS, Hackett JD, Ciniglia C, Pinto G, Bhattacharya D (2004) A mol- ecular timeline for the origin of photosynthetic eukaryotes. Molecular Biology and Evolution 21: 809–818

Zalewska T, Thompson VF, and Goll DE (2004) Effect of phosphatidylinositol and inside-out erythrocyte vesicles on autolysis of mu- and m-calpain from bovine skeletal muscle. Biochim. Biophys. Acta 1693: 125–133

Zimmer A, Lang D, Richardt S, Frank W, Reski R, Rensing SA (2007) Dating the early evolution of plants: detection and molecular clock analyses of orthologs. Molecular Genetics and Genomics 278: 393–402

Appendix 1

Table 1 A1 Primers and PCR-conditions used for fragment amplification prior to overlapping PCR, also used for identification of positive colonies in "colony PCR" reactions

Fragment (size)	Primers name (sequence; 5' → 3')	PCR-conditions
A: At-5´-fragment_Zm (623 bp)	SP/At_5'-fragment (ATACTGAGCGCTGAAATGTTCTCATTC)	98 ℃- 30 sec., (98 ℃- 5 sec., 58 ℃- 15 sec., 72 ℃- 30 sec.) x30, 72℃- 10 min.
	ASP/At_5'-fragment_Zm (TGGAGATCCAGCTAAATGATGGAATA)	
A: At-5´-fragment_Pp (593 bp)	SP/At_5'-fragment (ATACTGAGCGCTGAAATGTTCTCATTC)	
	ASP/At-5'-fragment_Pp (CTTTAAGGCAGCTCGACGAGCTGCTTCTGGGTCTGTTATAGATAGATGCCGAGAAAC)	
A: At-5´-fragment_Mv (621 bp)	SP/At_5'-fragment (ATACTGAGCGCTGAAATGTTCTCATTC)	
	ASP/At_5'-fragment_Mv (CGAGGGCGGCGCGGCGCTCGGCGAGAGGGTCTGTTATAGATAGATGCCGAGAAAC)	
B: Zm-Loop (840 bp)	SP/Zm_Loop (CCCTCAGTTGCAAGGATAGACGCT)	
	ASP/Zm_Loop (TGGAGATCCAGCTAAATGATGGAATA)	
B: Pp-Loop (837 bp)	SP/Pp_Loop (CCAGAAGCAGCTCGTCGAGCT)	
	ASP/Pp_Loop (TGGTGTCCCCACAAAGAACTGTAC)	
B: Mv-Loop (345 bp)	SP/Mv_Loop (CCTCTCGCCGAGCGCCGCGCC)	
	ASP/Mv_Loop (TCCACGGAAGGCAGCACGCA)	
C: At-3´-fragment_Zm (1397 bp)	SP/At_3'-fragment_Zm (CTAATATTCCATCATTTAGCTGGATCTCCAGAGAGAGCATGGGGCCTCTTTAGTC)	98 ℃- 30 sec., (98 ℃- 5 sec., 58 ℃- 15 sec., 72 ℃- 45 sec.) x30, 72 ℃- 10 min.
	ASP/At_3'-fragment (GTGGGCAACTGATCATCTCTAGATTTTA)	
C: At-3´-fragment_Pp (1367 bp)	SP/At_3fragment_Pp (GATGTTGTACAGTTCTTTGTGGGGACACCAGAGAGAGCATGGGGCCTCTTTAGTC)	
	ASP/At_3'-fragment (GTGGGCAACTGATCATCTCTAGATTTTA)	
C: At-3´-fragment_Mv (1367 bp)	SP/At_3'-fragment_Mv (TGCGCTGGGTGCGTGCTGCCTTCCGTGGAGAGAGAGCATGGGGCCTCTTTAGTC)	
	ASP/At_3'-fragment (GTGGGCAACTGATCATCTCTAGATTTTA)	

Table 2 A1 Primers used for the sequencing of the chimeric genes

Name	Oligo sequence (5'→ 3')
T7 universal forward primer	TAATACGACTCACTATAGGG
M13 universal reverse primer	CAGGAAACAGCTATGACC
SP/At_5Fragment_Seq_1	TGGCATATCCCGTCTTTTCCT
ASP/At_3Fragment_Seq_2	GGTGGAACTCACCTGTTGATGACT
SP/At_3Fragment_Seq_3	TTTTCGCACAATGGTCACAGG
ASP/At_3fragment_Seq_4	TGGGAACAGCAATAGACAGGCA

Table 3 A1 Primers used in Site-directed mutagenesis

Name	Oligo sequence (5'→ 3')
Zm_Loop_MutF	ACTAAGGAGAAGTGGTCTTGAAAAATGGTT
Zm_Loop_MutR	AGACCACTTCTCCTTAGTTCTTCTGACAGAG
Pp_Loop_MutF	GGGGAAACGAAAACTCAACCGCAGGCTTGAA
Pp_Loop_MutR	GCAGAGATCCAGGGGGAAACGAAAACTCAAC

Table 4 A1 Primers used for verification of the insert in kanamycin resistant plants

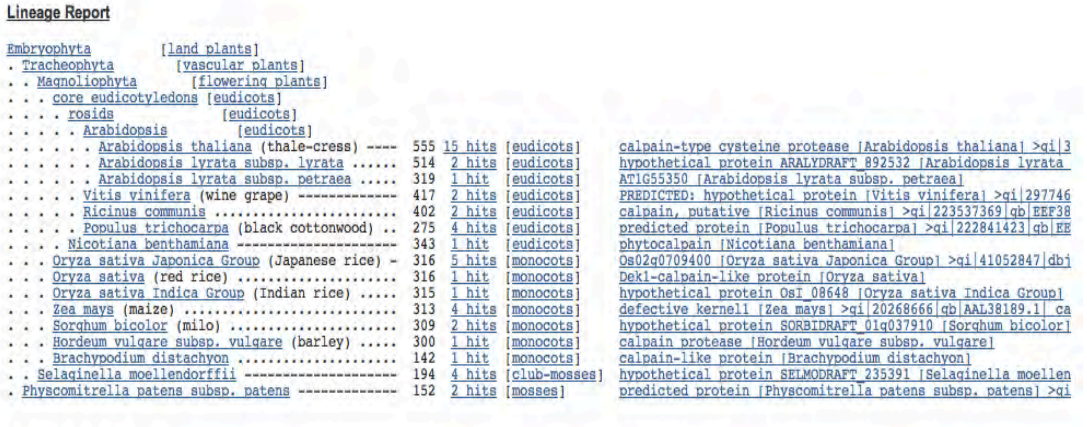
Name	Oligo sequence (5'→ 3')
SP/pSEL_AtMEM-GT	CTCTGCCGACAGTGGTCCCAAA
ASP/pSEL_AtMEM-GT	GCCATTATTATCGCCAGCCCAA

Table 5 A1 Primers used in qPCR for quantification of the transcript of the inserted gene and for its verification

Amplicon	Name	Oligo Sequence
AtDek1-MEM (size 216 bp)	AtDEK1 exon 7 qPCR SP	AAGAATTGAGACTCCGTGGACTAG
	AtDEK1 exon 8 qPCR ASP	GTGACAATAGCAGCACAGAGAAAC
AtDek1-CALP (size 167 bp)	AtDEK1 exon 27-28 qPCR SP	GCTTACTCCGTCTTACAGGTGAG
	AtDEK1 exon 27-28 qPCR ASP	CCTTCTTTTGACTGTGGAACATG
At-PP2A (size 128 bp)	AtPP2A SP	CTCCAGTCTTGGGTAAGGATG
	AtPP2A ASP	AATAACCTGGTTCACTTGGTCAAG
AtDek1-MEM-DEL (size 205 bp)	AtDEK1 exon 7a qPCR_SP	GCATCAGTTCTATCAGGTGCTG
	AtDEK1 exon8 qPCR ASP2	CAGACAACAGGTGACAATAGCAG
AtDek1-MvLoop (size 152 bp)	Mv loop-exon7b qPCR SP	TGCCTTCCGTGGAGAGAG
	AtDEK1 exon 8 qPCR ASP	GTGACAATAGCAGCACAGAGAAAC
AtDek1-ZmLoop (212 bp)	Zm loop-exon7b qPCR SP	TGGATCTCCAGAGAGAGCATG
	AtDEK1 exon 8 qPCR ASP	GTGACAATAGCAGCACAGAGAAAC

Appendix 2

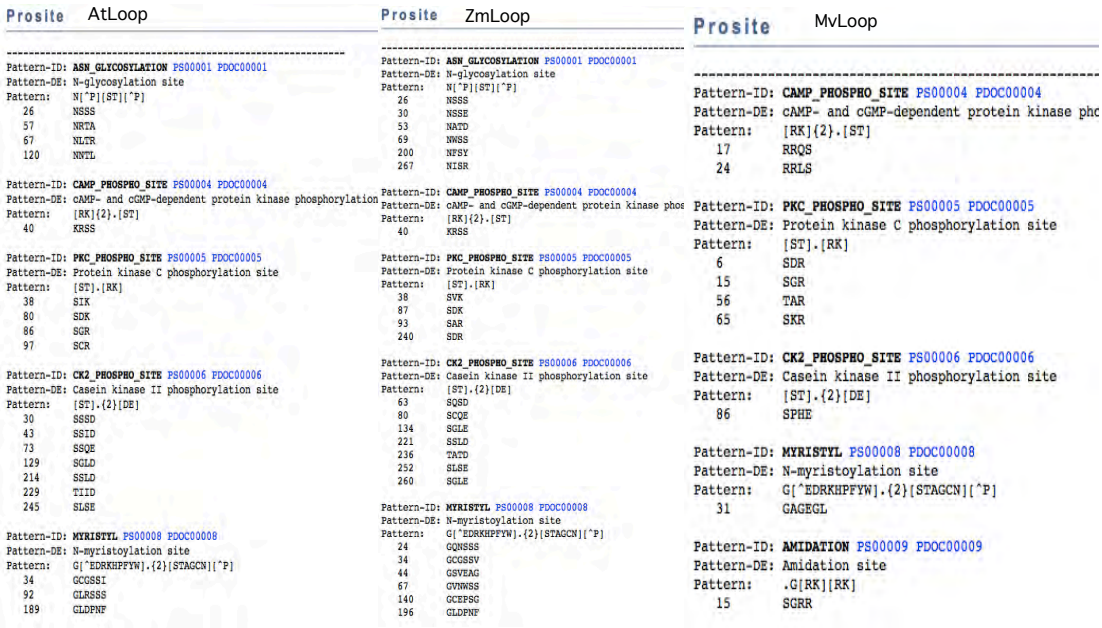
Results from bioinformatics anlyses



Lineage Report

```
Embryophyta          [land plants]
. Tracheophyta         [vascular plants]
. . Magnoliophyta        [flowering plants]
. . . core eudicotyledons [eudicots]
. . . . rosids             [eudicots]
. . . . . Arabidopsis        [eudicots]
. . . . . . Arabidopsis thaliana (thale-cress) ----  555 15 hits [eudicots]   calpain-type cysteine protease [Arabidopsis thaliana] >gi|3
. . . . . . Arabidopsis lyrata subsp. lyrata .....  514  2 hits [eudicots]   hypothetical protein ARALYDRAFT_892532 [Arabidopsis lyrata
. . . . . . Arabidopsis lyrata subsp. petraea .....  319  1 hit  [eudicots]   AT1G55350 [Arabidopsis lyrata subsp. petraea]
. . . . . Vitis vinifera (wine grape) ------------  417  2 hits [eudicots]   PREDICTED: hypothetical protein [Vitis vinifera] >gi|297746
. . . . . Ricinus communis .......................  402  2 hits [eudicots]   calpain, putative [Ricinus communis] >gi|223537369|gb|EEF38
. . . . . Populus trichocarpa (black cottonwood) ..  275  4 hits [eudicots]   predicted protein [Populus trichocarpa] >gi|222841423|gb|EE
. . . . Nicotiana benthamiana ---------------------  343  1 hit  [eudicots]   phytocalpain [Nicotiana benthamiana]
. . . Oryza sativa Japonica Group (Japanese rice) -  316  5 hits [monocots]   Os02g0709400 [Oryza sativa Japonica Group] >gi|41052847|dbj
. . . Oryza sativa (red rice) ....................  316  1 hit  [monocots]   Dek1-calpain-like protein [Oryza sativa]
. . . Oryza sativa Indica Group (Indian rice) .....  315  1 hit  [monocots]   hypothetical protein OsI_08648 [Oryza sativa Indica Group]
. . . Zea mays (maize) ...........................  313  4 hits [monocots]   defective kernel1 [Zea mays] >gi|20268666|gb|AAL38189.1| ca
. . . Sorghum bicolor (milo) .....................  309  2 hits [monocots]   hypothetical protein SORBIDRAFT_01g037910 [Sorghum bicolor]
. . . Hordeum vulgare subsp. vulgare (barley) .....  300  1 hit  [monocots]   calpain protease [Hordeum vulgare subsp. vulgare]
. . . Brachypodium distachyon ....................  142  1 hit  [monocots]   calpain-like protein [Brachypodium distachyon]
. . Selaginella moellendorffii ------------------  194  4 hits [club-mosses]  hypothetical protein SELMODRAFT_235391 [Selaginella moellen
. Physcomitrella patens subsp. patens ------------  152  2 hits [mosses]      predicted protein [Physcomitrella patens subsp. patens] >gi
```

Figure 1 A2 Lineage report provided by BLASTp search using the AtDEK1-LOOP sequence as query. Showing how closely are the organisms in the BLAST hitlist related to the query sequence according to their classification in the taxonomy database.



Prosite AtLoop

```
Pattern-ID: ASN_GLYCOSYLATION PS00001 PDOC00001
Pattern-DE: N-glycosylation site
Pattern:    N[^P][ST][^P]
   26       NSSS
   57       NRTA
   67       NLTR
   120      NNTL

Pattern-ID: CAMP_PHOSPHO_SITE PS00004 PDOC00004
Pattern-DE: cAMP- and cGMP-dependent protein kinase phosphorylation
Pattern:    [RK]{2}.[ST]
   40       KRSS

Pattern-ID: PKC_PHOSPHO_SITE PS00005 PDOC00005
Pattern-DE: Protein kinase C phosphorylation site
Pattern:    [ST].[RK]
   38       SIK
   80       SDK
   86       SGR
   97       SCR

Pattern-ID: CK2_PHOSPHO_SITE PS00006 PDOC00006
Pattern-DE: Casein kinase II phosphorylation site
Pattern:    [ST].{2}[DE]
   30       SSSD
   43       SSID
   73       SSQE
   129      SGLD
   214      SSLD
   229      TIID
   245      SLSE

Pattern-ID: MYRISTYL PS00008 PDOC00008
Pattern-DE: N-myristoylation site
Pattern:    G[^EDRKHPFYW].{2}[STAGCN][^P]
   34       GCGSSI
   92       GLRSSS
   189      GLDPNF
```

Prosite ZmLoop

```
Pattern-ID: ASN_GLYCOSYLATION PS00001 PDOC00001
Pattern-DE: N-glycosylation site
Pattern:    N[^P][ST][^P]
   26       NSSS
   30       NSSE
   53       NATD
   69       NWSS
   200      NFSY
   267      NISR

Pattern-ID: CAMP_PHOSPHO_SITE PS00004 PDOC00004
Pattern-DE: cAMP- and cGMP-dependent protein kinase phos
Pattern:    [RK]{2}.[ST]
   40       KRSS

Pattern-ID: PKC_PHOSPHO_SITE PS00005 PDOC00005
Pattern-DE: Protein kinase C phosphorylation site
Pattern:    [ST].[RK]
   38       SVK
   87       SDK
   93       SAR
   240      SDR

Pattern-ID: CK2_PHOSPHO_SITE PS00006 PDOC00006
Pattern-DE: Casein kinase II phosphorylation site
Pattern:    [ST].{2}[DE]
   63       SQSD
   80       SCQE
   134      SGLE
   221      SSLD
   236      TATD
   252      SLSE
   260      SGLE

Pattern-ID: MYRISTYL PS00008 PDOC00008
Pattern-DE: N-myristoylation site
Pattern:    G[^EDRKHPFYW].{2}[STAGCN][^P]
   24       GQNSSS
   34       GCGSSV
   44       GSVEAG
   67       GVNWSS
   140      GCEPSG
   196      GLDPNF
```

Prosite MvLoop

```
Pattern-ID: CAMP_PHOSPHO_SITE PS00004 PDOC00004
Pattern-DE: cAMP- and cGMP-dependent protein kinase phc
Pattern:    [RK]{2}.[ST]
   17       RRQS
   24       RRLS

Pattern-ID: PKC_PHOSPHO_SITE PS00005 PDOC00005
Pattern-DE: Protein kinase C phosphorylation site
Pattern:    [ST].[RK]
   6        SDR
   15       SGR
   56       TAR
   65       SKR

Pattern-ID: CK2_PHOSPHO_SITE PS00006 PDOC00006
Pattern-DE: Casein kinase II phosphorylation site
Pattern:    [ST].{2}[DE]
   86       SPHE

Pattern-ID: MYRISTYL PS00008 PDOC00008
Pattern-DE: N-myristoylation site
Pattern:    G[^EDRKHPFYW].{2}[STAGCN][^P]
   31       GAGEGL

Pattern-ID: AMIDATION PS00009 PDOC00009
Pattern-DE: Amidation site
Pattern:    .G[RK][RK]
   15       SGRR
```

Figure 2 A2 Prosite protein-protein interaction prediction on the AtDEK1-LOOP (to the left), ZmDEK1-LOOP (in the middle) and MvDEK1-LOOP (to the right) sequences

Results from phenotypic characterizations

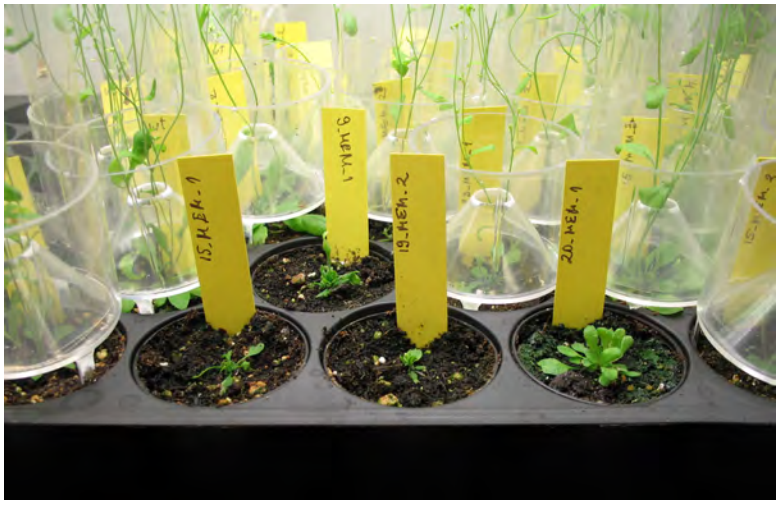


Figure 3 A2 Picture of 5 weeks old plants expressing phenotype on soil

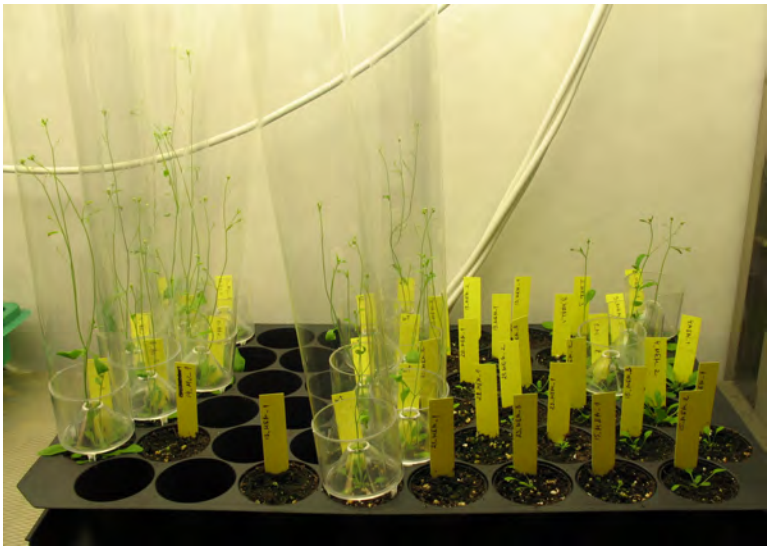


Figure 4 A2 Picture showing growth differences between AtDEK1-MEM-MvLoop transformants (to the left) and AtDEK1-MEM transformants (to the right), wild type plants are in the middle

Table 1 A2 Overview of phenotypes plants expressed after being transferred to soil

On plate	On soil	MEM	MEM-DEL	MvLoop	ZmLoop
wt	wt	22 (73 %)	15 (63 %)	11 (65 %)	13 (56 %)
	sterile	8 (27 %)	9 (37 %)	5 (29 %)	4 (17 %)
	dead	0	0	1 (6 %)	5 (21 %)
phenotype	wt	14 (28 %)	2 (22 %)	17 (58 %)	2 (14 %)
	sterile	4 (8 %)	2 (22 %)	1 (3 %)	1 (7 %)
	dead	32 (64 %)	5 (56 %)	11 (37 %)	11 (78 %)

Results from qRT-PCR analysis; data used to create histograms

Table 2 A2 Average of the relative expression ratio determined by qPCR analysis in AtDEK1-MEM transgenic lines

Name	Line/phenotype	Loop/Calp	Loop/PP2A	Calp/PP2A
Wt (2 weeks old)	-	1,1	1,0	0,98
Wt (5 weeks old)	-	1,0	0,7	0,73
1MEM1	1 "wt"	4,3	4,0	0,92
1Mem1g	1(5w) "wt"	4,3	4,0	0,92
9_MEM_1	10 "wt"	25	15	0,71
9_MEM_1g	10(5w) sterile	42	24	0,62
22_MEM2	33 "wt"	3,8	1,1	0,31
22_MEM_2	33(5wo) sterile	3,1	1,2	0,46
11_MEM_2	14 DEC	2,2	1,4	0,65
35_MEM_2	77 DEC	54	26	0,46
34_MEM_3	75 DEC	3,3	1,4	0,44
2_MEM_4	5 DSAM	2,9	1,5	0,61
14_MEM_1	17 DSAM	1,9	0,9	0,59
21_MEM_1	30 DSAM	1,3	0,3	0,20
18MEM2	25 SFTL	20,9	14,4	0,69
19MEM2	27 SFTL	1,0	0,3	0,30
20MEM1	29 SFTL	1,7	0,7	0,61
20MEM1g	29(5w) sterile	3,3	0,9	0,41
2MEM1	2 FC	1,2	0,4	0,32
2MEM2	3 FC	3,1	1,1	0,51
12MEM1	15 FC	2,1	1,0	0,47
8MEM1	9 SC	1,6	0,8	0,63
12MEM2	16 SC	2,6	2,5	0,96

Table 3 A2 Average of the relative expression ratio determined by qPCR analysis in AtDEK1-MEM-DEL transgenic lines

Name	Line/Pheno	Loop/Calp	Loop/PP2A	Calp/PP2A
1DEL1	2 "wt"	44	12	0,28
1DEL1v	2(5w) "wt"	62	15	0,25
8DEL1	12 "wt"	1,84	1,02	0,55
5DEL1	9 wt	7,21	2,56	0,35
5DEL1v	9(5w) "wt"	0,82	0,49	0,59
10DEL1	15 "wt"	5,90	0,83	0,15
10DEL1v	15(5w) stetile	6,14	0,83	0,15
9DEL1	14 DSAM	3,06	1,80	0,59
7DEL1	11 SFTL	1,35	0,92	0,68
7DEL1v	11(5w) sterile	1,01	0,70	0,69

Table 4 A2 Average of the relative expression ratio determined by qPCR analysis in AtDEK1-MEM-MvLoop transgenic lines

Name	Line/Phenotype	Loop/Calp	Loop/PP2A	Calp/PP2A
2Mv1	1 wt	9,49	6,01	0,63
2Mv1g	1(5wo) sterile	6,68	4,29	0,64
9Mv1	4 wt	2,64	0,19	0,73
9Mv1g	4(5wo) wt	1,26	0,83	0,63
11Mv1	5 wt	2,11	1,32	0,63
11Mv1g	5(5wo) wt	1,25	0,24	0,19
8Mv1	3 DEC	4,21	3,92	0,93
11Mv4	8 DEC	2,14	1,31	0,61
16Mv1	16 DEC	1,13	0,36	0,32
16Mv1g	16(5wo) wt	1,25	0,24	0,19
18Mv4	23 DSAM	5,28	0,57	0,11
6Mv1	2 SFTL	1,16	1,12	0,97
13Mv1	9 SFTL	7,24	0,68	0,09
13Mv1g	9(5wo) sterile	3,17	0,37	0,12
17Mv2	19 SFTL	3,84	2,10	0,55
17Mv2g	19(5wo) wt	0,85	0,61	0,72
14Mv5	14 FC	1,25	0,83	0,66

Table 5 A2 Average of the relative expression ratio determined by qPCR analysis in AtDEK1-MEM-ZmLoop transgenic lines

Name	Line/phenotype	Loop/Calp	Loop/PP2A	Calp/PP2A
11Zm2	24 "wt"	69,19	22,04	0,32
11Zm2v	24(5wo) "wt"	43,11	6,56	0,15
10Zm2	22 "wt"	19,57	8,74	0,45
10Zm2v	22(5wo) sterile	18,65	5,38	0,29
10Zm1	21 DSAM	1,01	2,80	2,93
18Zm1	34 DSAM	0,85	0,61	0,72
11Zm1	23 SFTL	8,10	1,88	0,23
11Zm1v	23(5wo) "wt"	5,61	1,02	0,18
20Zm1	37 SFTL	1,25	0,83	0,66

Results from co-transformation experiment

Figure 5 A2 Four weeks old seedlings on kan^{50}/ hyg^{20} selective medium from co-transformation attempt