Identification and characterization of upstream regulators of *Arabidopsis* Metacaspase 9

Maria Lundström

LiTH-IFM-A-Ex--11/2437--SE

Supervisor: Hannele Tuominen, Department of Plant Physiology, Umeå universitet

Examiner: Johan Edqvist, Linköpings universitet

Department of Physics, Chemistry and Biology

Linköpings universitet

SE-581 83 Linköping, Sweden
Programmed cell death (PCD) refers to a genetically controlled process causing the death of certain cells or tissues. In plants PCD is critical in normal development of for instance xylem vessels. A group of proteins called metacaspases are believed to play a pivotal role in PCD in plants. As Metacaspase 9 have been shown to be upregulated in Populus during xylem maturation this study attempted to identify genes affecting its expression in Arabidopsis thaliana by forward genetics using a reporter line with GFP fused to the promoter of Metacaspase 9 (AtMC9). Ethyl methanesulfonate seed mutagenesis was used to generate mutants resulting in eleven mutant lineages with a GFP expression pattern deviating from that of the reporter line. These mutants fell into two categories; low/no-signal mutants and ectopic expressors. Several of the low/no-signal mutants had longer roots at five to eight days after germination, a time point shown to be critical for metaxylem differentiation. Further studies of their roots would reveal whether the developing xylem is abnormal or not. Deep sequencing provided evidence for involvement of abscisic acid and polyamines in regulation of AtMC9 expression. Sequencing from a low/no-signal mutant suggests that AtMC9 expression might be affected also by disturbed lignin biosynthesis. Rescuing mutant lineages through transformations with fully functional forms of the candidate genes is the next step to experimentally validate that the candidate genes are involved in the observed changes in AtMC9 expression in each of the isolated mutants.
# Contents

1 Abstract ........................................................................................................................................... 1  
2 List of abbreviations .......................................................................................................................... 1  
3 Introduction ........................................................................................................................................ 1  
4 Materials and methods .......................................................................................................................... 3  
  4.1 Plant growth conditions .................................................................................................................. 3  
  4.2 Crossings ......................................................................................................................................... 3  
  4.3 Confocal microscopy ....................................................................................................................... 3  
  4.4 Genomic DNA extraction using CTAB .......................................................................................... 4  
  4.5 TA cloning, bacterial transformation and blue-white screening of recombinants ......................... 4  
  4.6 Restriction digest and sequencing ................................................................................................ 4  
  4.7 Chloral hydrate clearing ................................................................................................................... 5  
  4.8 DNA extraction for genotyping using SDS ...................................................................................... 5  
  4.9 Polymerase chain reaction settings for genotyping ....................................................................... 5  
  4.10 96-well format DNA extraction ..................................................................................................... 5  
  4.11 SSLP marker design ...................................................................................................................... 6  
  4.12 Polymerase chain reaction conditions during mapping ............................................................... 6  
  4.13 DNA extraction for deep sequencing ............................................................................................ 7  
  4.14 Data analysis .................................................................................................................................. 7  
5 Results .................................................................................................................................................. 7  
  5.1 Mutants fall into two phenotypical categories ................................................................................. 7  
    5.1.1 Screening of the EMS mutagenized reporter line ....................................................................... 7  
    5.1.2 GFP expression pattern in ectopic expressors and low/no-signal mutants .................................. 8  
  5.2 Characterization of mutants ............................................................................................................. 10  
    5.2.1 Checking GFP construct for mutations in mutants 20IV, 30I, 47X and 50IX ............................... 10  
    5.2.2 General growth phenotype of the M3 generation ..................................................................... 11  
    5.2.3 Vascular development in cotyledons appears to be unaffected by mutations ............................ 17  
    5.2.4 Ectopic expression of GFP in cotyledons is not associated with increased cell death .......... 18  
  5.3 Locating mutations through mapping and deep sequencing .......................................................... 18  
    5.3.1 Crossings and segregation of mutant traits ............................................................................... 18  
    5.3.2 Mapping of mutants 20IV and 30I ............................................................................................ 20  
    5.3.3 Deep sequencing of mutants 20IV, 30I, 47X and 50IX ................................................................. 22  
6 Discussion ........................................................................................................................................... 24  
7 Acknowledgements .............................................................................................................................. 27  
8 References ........................................................................................................................................... 27  
Appendix 1 .......................................................................................................................................... 31  
Appendix 2 .......................................................................................................................................... 36
1 Abstract

Programmed cell death (PCD) refers to a genetically controlled process causing the death of certain cells or tissues. In plants PCD is critical in normal development of for instance xylem vessels. A group of proteins called metacaspases are believed to play a pivotal role in PCD in plants. As Metacaspase 9 have been shown to be upregulated in Populus during xylem maturation this study attempted to identify genes affecting its expression in Arabidopsis thaliana by forward genetics using a reporter line with GFP fused to the promoter of Metacaspase 9 (AtMC9). Ethyl methanesulfonate seed mutagenesis was used to generate mutants resulting in eleven mutant lineages with a GFP expression pattern deviating from that of the reporter line. These mutants fell into two categories; low/no-signal mutants and ectopic expressors. Several of the low/no-signal mutants had longer roots at five to eight days after germination, a time point shown to be critical for metaxylem differentiation. Further studies of their roots would reveal whether the developing xylem is abnormal or not. Deep sequencing provided evidence for involvement of abscisic acid and polyamines in regulation of AtMC9 expression. Sequencing from a low/no-signal mutant suggests that AtMC9 expression might be affected also by disturbed lignin biosynthesis. Rescuing mutant lineages through transformations with fully functional forms of the candidate genes is the next step to experimentally validate that the candidate genes are involved in the observed changes in AtMC9 expression in each of the isolated mutants.

Keywords: Metacaspase 9, Programmed cell death, Xylem differentiation

2 List of abbreviations

ABA – Abscisic Acid
ABRE – ABA-Responsive Elements
AtMC9 – Arabidopsis thaliana Metacaspase 9
Caspase – Cys-dependent Asp-specific peptidase
Col – Columbia
EMS – Ethyl Methanesulfonate
GFP – Green Fluorescent Protein
INDEL – Short Insertions and Deletions
Ler – Landsberg
PA - Polyamine
PCD – Programmed Cell Death
PPT – Phosphinotricin
SND1 – Secondary Wall-Associated NAC Domain Protein1
SNP – Single-Nucleotide Polymorphism
SSLP – Single Sequence Length Polymorphism
TAIR – The Arabidopsis Information Resource
TE – Tracheary Element
VND6 /VND7 – Vascular-Related NAC-Domain6/7
Ws – Wassilewskija
WT – Wild Type
XCP1/XCP2 – Xylem Cysteine Protease1/2

3 Introduction

Programmed cell death (PCD) is a genetically controlled process resulting in the death of a cell or tissue and is required for normal development and growth in both animals and plants. In plants PCD is critical during, for instance, embryogenesis [Bozhkov et al., 2005],
development of leaf shape [Gunawardena et al., 2004], development and abortion of floral organs [Rogers, 2006] and the hypersensitive response induced by pathogen attack [reviewed by Hofius et al., 2007].

Differentiating xylem have long been known to undergo programmed cell death [reviewed by Fukuda 1996]. Pyo et al. (2004) found when tracing vessel differentiation in Arabidopsis by fusing GUS to the promoter of ZCP4 that differentiation of protoxylem started at the proximal region of the midvein in cotyledons and also in the distal region of the hypocotyls. The protoxylem then developed discontinuously in both cotyledon and along the hypocotyl-root axis and was found to form continuously in the area close to the root tip even after maturation in the rest of the seedling. Metaxylem differentiation in Arabidopsis have been found to start at the hypocotyl-root junction, then at the cotyledon-hypocotyl junction and develop bidirectionally [Pyo et al., 2004]. In Arabidopsis a group of NAC-domain transcription factors, the vascular-related NAC-domain genes (VND1-7), have been suggested to have a key role in controlling the differentiation of xylem vessels [Kubo et al., 2005]. When VND6 and VND7 were overexpressed under control of the cauliflower mosaic virus 3SS promoter this resulted in transdifferentiation of various cell types into xylem vessel cells in both hypocotyls and roots. Further Kubo et al. (2005) found that VND6 induced xylem vessel elements with a high resemblance to metaxylem vessels whereas VND7 induced xylem vessel elements highly similar to protoxylem vessels. Yamaguchi et al. (2008) showed that VND7 could cause transcriptional activation and found indications that it might form homodimers and heterodimers with other VNDs, and also concluded that it is pivotal in regulation of vessel differentiation in both roots and aerial parts as these areas contained defect vessels when expressing truncated VND7 proteins lacking the C-terminal transcription-activation domain.

Primary xylem contains two different cell types: xylem parenchyma and tracheary elements, TEs, that are cells specialized for transport of water and solutes. TEs develop from procambial cells into TE precursor cells, elongates, forms secondary cell walls followed by PCD through rupturing of the vacuole and digestion of the nucleus and organelles resulting in a hollow, fairly robust structure suitable for water transport [reviewed in Fukuda, 2000; Turner et al., 2007]. The collapse of the vacuole has been shown to be correlated with increased activity of proteases [Demura et al., 2002; Fukuda, 2000] and nucleases [Ito and Fukuda, 2002; Lehmann et al., 2001] in the cytoplasm, but whether this is caused by activation of cytoplasmic enzymes due to change in pH or that they are released from the vacuole is not yet known. By inducing vacuolar collapse in non-TE cells in Zinnia Obara et al. (2001) could, however, show that the resulting change in cytoplasmic pH and Ca\(^{2+}\) concentration was not enough to cause breakdown of nucleic acids in these cells.

The PCD process in animals is dependent on the caspases (Cys-dependent Asp-specific peptidases) of which there are two different kinds: the initiator caspases and the executioner caspases. The initiator caspases consist of a large and a small subunit, the p20 and p10 subunits, and also carry a N-terminal prodomain involved in interactions with other proteins. The N-terminal prodomain is absent in the executioner caspases, but they do have the p20 and the p10 subunits [reviewed by Vercammen et al., 2007]. In plants no actual caspases have been found so far, only proteins with similarities in structure or activity [reviewed by Vercammen et al., 2007; Chichkova et al., 2010]. Some of these structurally similar proteins are the metacaspases that, as the caspases, are divided into two subclasses: type I and type II metacaspases. Type I metacaspases resembles initiator caspases with p20 and p10 subunits and also a N-terminal prodomain. It is hypothesized that they may be activated in the same manner as the initiator caspases, i.e. through oligomerization and the conformational changes
that follows [reviewed by Fuentes-Prior and Salvesen, 2004], since they have been shown to
not undergo proteolytic activation. Type II metacaspases have both the p20 and the p10
subunits but, same as executioner caspases, they lack the large N-terminal prodomain and
seem to activate through autocatalytic cleavage that separates the large and the small subunits
[reviewed by Piszczek and Gutman, 2007; reviewed by Vercammen et al., 2007].

Nine metacaspase genes have been identified in *Arabidopsis thaliana* (*AtMC1-AtMC9*) of
which *AtMC1-AtMC3* are type I metacaspases and *AtMC4-AtMC9* are type II metacaspases
[Vercammen et al., 2004]. AtMC9 have, same as AtMC4, been shown to be an
arginine/lysine-specific cysteine protease [Vercammen et al., 2004]. AtMC9 has a pH
optimum at 5.0-5.5 and is deactivated at pH 7.0-7.5 [Vercammen et al., 2004] and has been
found to be upregulated in *Populus* during xylem maturation indicating potential involvement
in PCD [Courtois-Moreau et al., 2009]. There are also indications that *AtMC9* expression is
under the control of *SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN1* (*SND1*)
and *VASCULAR-RELATED NAC-DOMAIN7* (*VND7*), two master regulator genes of
secondary wall biosynthesis in fibers and vessels respectively [Zhong et al., 2010]. By
mapping and deep sequencing of *Arabidopsis* mutants based on an ethyl methanesulfonate
mutagenized reporter line carrying green fluorescent protein connected to the promoter of
Metacaspase 9 we are attempting to uncover upstream actors of *AtMC9*.

4 Materials and methods

4.1 Plant growth conditions

Seeds were sown on agar plates with Murashige and Skoog medium (Duchefa Biochemie)
and stratified in darkness at 4 °C for at least 48 h to promote uniform germination and growth
patterns. Plates were transferred to a long day growth chamber with temperatures of 23 ± 2 °C
during the day and 18 °C at night with 16 h illumination and kept in a vertical position to
enable observations of the roots as they developed. Seedlings were transferred onto soil- and
vermiculite-mixture (2:1 ratio) at an age of 10-12 days. When growing the seedlings for deep
sequencing of F2 mapping population the two mutants 47X and 50IX were plated onto
Murashige and Skoog medium containing phosphinotricin (PPT; 10 mg L⁻¹) to select for
plants being either homozygous or heterozygous for the *AtMC9::nGFP* construct carrying the Bar resistance gene.

4.2 Crossings

Inflorescences of the recipient plants were cleared of leaves, lateral shoots, mature flowers
and siliques to isolate the flower buds used for crossings and reduce the risk of pollen
contamination from nearby flowers. Crossings were performed using two pairs of fine forceps
to gently remove sepals, petals and stamens of a not yet open flower bud to prevent self-
fertilization and give access to the pistil. The pistil was dabbed with a mature anther from the
crossing partner to transfer pollen onto the stigma.

4.3 Confocal microscopy

Five-day-old M3 generation seedlings were investigated using Leica TCS SP2 AOB5
confocal microscope with argon laser. The endogenous chlorophyll signal in the cotyledons
was detected within the wavelengths 670-720 nm and GFP within 505-535 nm. Roots were
stained with propidium iodide in MS medium prior to scanning. Propidium iodide was
detected within wavelengths 600-700 nm and GFP in roots within 500-545 nm. Z-stacks were created with two reads per line and one read per plane.

4.4 Genomic DNA extraction using CTAB

Rosette leaves from M3 generation mutants were collected into individual Eppendorf tubes and flash frozen in liquid nitrogen. Leaves were ground and 2xCTAB buffer (2 % cetyltrimethylammonium bromide, 1.4 M NaCl, 100 mM Tris-HCl pH 8.0, 20 mM EDTA) added to lyse cells and assist in releasing DNA from proteins and lipids followed by incubation at 65 ºC, 10 min. Once the samples had cooled chloroform was added to assist in extraction of protein and lipids before samples were centrifuged at 16000 g (13200 rpm), 2 min, to separate the phases. The aqueous phase containing the DNA was transferred to a fresh tube containing isopropanol, mixed well and centrifuged at 13000 rpm, 5 min, to precipitate DNA. The supernatant was removed and the pellet washed using 70 % ethanol to remove residual salts and chloroform. The ethanol was removed and the pellet dried to avoid ethanol remains disturbing downstream reactions. The pellet was then redissolved in TE buffer (10 mM Tris-HCl pH 8.0, 1 mM EDTA).

4.5 TA cloning, bacterial transformation and blue-white screening of recombinants

In order to facilitate sequencing of the AtMC9::nGFP constructs PCR fragments were amplified by mixing GoTaq Green Mastermix (Promega) and ddH2O with attB1 forward (CAAGTTTGTACAAAAAAGCAG) and T-NOS reverse (ATCGCAAGACCGGCAACAGG) primers and amplified using PCR. The PCR programme started with denaturing the DNA at 95 ºC for 2 min, followed by 30 cycles of 95 ºC for 10 s, annealing at 54 ºC for 20 s followed by 1 min and 15 s elongation at 72 ºC. Final elongation lasted for 2 min at 72 ºC, then temperature was kept at 18 ºC. Samples were run on 1 % agarose gel with GelRed (BioTium) in 1xTris-acetate-EDTA buffer. Gel slices containing the fragments of the proper size were excised and DNA extracted according to E.Z.N.A gel extraction protocol (Omega Bio-Tek). Cloning was performed using pGEM-T vector system (Promega) according to the manufacturers instructions. Ligation was performed according to protocol using a 3:1 insert:vector molar ratio, but electro-competent DH5α Escherichia coli cells were used for the transformation. The ligation reactions were transferred to cuvettes together with cell suspension, placed in an electroporator and exposed to 1.5 kV given in 4 ms intervals over a time period of ~1 s. LB was added and the whole volume transferred to an Eppendorf tube. Tubes were incubated at 37 ºC with shaking, 1.5 h, thereby preventing bacteria from forming aggregates that reduce growth efficiency. The cultures were spread on LB/ampicillin/IPTG/X-Gal plates and results evaluated through blue and white selection of recombinants.

4.6 Restriction digest and sequencing

Plasmid minipreps were prepared according to Hattori and Sakaki (1986). The positive clones identified in the blue-white screening were controlled for inserts through a restriction digest. DNA was linearised using enzyme ApaI (Fermentas) in buffer blue and a double digest of DNA was performed with NcoI (4x excess) and NotI restriction enzymes (Fermentas) in buffer orange, both digestions were conducted at 37 ºC for 1 h followed by thermal inactivation at 65 ºC for 20 min. Undigested, linearized and double digested DNA was then run on a 1 % agarose gel with GelRed (BioTium). Sequencing of the DNA was performed by Eurofins MWG Operon, Germany.
4.7 Chloral hydrate clearing

Seedlings were placed in a mixture of ethanol and acetic acid in a 9:1 ratio until chlorophyll was cleared. This was followed by washing in 70% ethanol and incubation in 95% ethanol over night before clearing of tissues using chloral hydrate (2.5 g mL⁻¹) dissolved in 30% glycerol. Vascular development was assessed by studying the number of fully connected vascular loops in the cotyledons under Zeiss Axioplan microscope.

4.8 DNA extraction for genotyping using SDS

Arabidopsis rosette leaves were flash frozen in liquid nitrogen and ground to fine powder before extraction buffer (200 mM Tris-HCl pH 7.5, 250 mM NaCl, 25 mM EDTA, 0.5% SDS) was added to lyse the cells and denature proteins. The sample was allowed to thaw before it was placed in a 65 °C water-bath, 15 min, followed by centrifugation at 13000 rpm, 5 min, to pellet cell debris. The supernatant was transferred to a fresh tube with isopropanol, gently mixed and left in RT for 5 min to precipitate DNA, and centrifuged at 13000 rpm, 5 min. DNA pellet was washed in 70% ethanol followed by centrifugation at 13000 rpm, 5 min. All traces of ethanol were removed to not disturb downstream reactions before the pellet was resuspended in TE.

4.9 Polymerase chain reaction settings for genotyping

The F1 crossings were checked for contamination using two different markers: SSLP mio24 F and SSLP mio24 R (TGGTGGTGTACGATTTTACCAA and GGATCGTTTTATGCATTTCTCG) or SSLP 1830 At4g F and SSLP At4g R (AGCCATTCTAGTCAGCTT and CCTCTGTTTGAGCATTCGTG). A master mix was prepared containing GoTaq Green Mastermix (Promega), ddH2O and one of the two marker pairs. The PCR programme used started with denaturing at 95 °C for 2 min, followed by a cycle repeated 30 times of 20 s at 95 °C, annealing for 20 s at 55 °C and 30 s elongation at 72 °C. Final elongation lasted for 2 min at 72 °C and then samples were kept at 18 °C. PCR products were analysed on a 2% agarose gel with GelRed (BioTium) in 1xTris-acetate-EDTA buffer.

4.10 96-well format DNA extraction

Whole seedlings were placed individually into 8-strip cluster tubes (Corning Costar 4408) of a 96-well plate. Ball bearings and extraction buffer (0.1 M Tris pH 8, 0.05 M EDTA pH 8, 0.5 M NaCl and 1% PVP) were added to each tube to assist cell lysis and the plates placed in a bead mixer mill at 25 Hz, 6.5 min, for mechanical shearing of tissue. Samples were spun down in a microtitre plate centrifuge at 5300 g (5360 rpm) for 1 min to reduce foaming, then 10% SDS and chloroform:isoamyl alcohol (24:1) was added to each tube to extract proteins and lipids and mixed vigorously. Plates were yet again centrifuged in the microtiter plate centrifuge at 5300 g (5360 rpm) for 15 minutes to pellet cell debris. Supernatant containing the DNA was transferred to a fresh 2.2 ml storage plate (Abgene MkII, AB-0932) with 5M KOAc (KCH₃CO₂ dissolved in glacial acetic acid and ddH₂O, pH ~5.8) to precipitate protein remains, covered with sealing tape (Abgene AB-0558) and centrifuged at 5300 g (5360 rpm) for 30 min. The supernatant was transferred to another 2.2 ml storage plate with 0.7 volumes of isopropanol to precipitate DNA and mixed by vortexing, then centrifuged at 5300 g (5360 rpm) for another 30 min. Isopropanol was poured off followed by washing in 70% ethanol and yet another centrifugation at 5300 g (5360 rpm) for 10 min. All traces of ethanol were removed to not interfere with downstream reactions before the DNA was resuspended in TE.
4.11 SSLP marker design

INDEL (small insertions and deletions) polymorphisms differing between Col and Ler in the area of interest were located in Monsanto Arabidopsis Polymorphism collection release 1. Flanking sequences given by Monsanto were used to search for the full length sequence of the region of interest in The Arabidopsis Information Resource (TAIR). The genome sequence was inserted in Primer3 Input version 0.4.0 where primer size was set to minimum 18 bp, maximum 24 bp with an optimum at 20 bp and melting temperature was set at minimum 55 °C, maximum at 65 °C and optimum at 60 °C, otherwise default settings were used.

4.12 Polymerase chain reaction conditions during mapping

DNA samples were mixed with GoTaq Green Mastermix (Promega), the marker of interest (see Table 1) and ddH2O. The PCR programme started with denaturation at 95 °C for 2 min followed by a cycle repeated 30 times of 95 °C for 20 s, annealing temperature (T_a) according to Table 1 for 20 s and elongation at 72 °C according to Table 1. Final elongation lasted for 2 min at 72 °C, and then the temperature was kept at 18 °C. The samples were analysed on 3 or 4 % agarose gels with GelRed (BioTium) in 1xTris-acetate-EDTA buffer.

Table 1. SSLP markers distinguishing between Columbia and Landsberg ecotypes used for rough and fine mapping of mutants 20IV and 30I.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Marker</th>
<th>Sequence</th>
<th>Chromosome</th>
<th>AGI position (bp)</th>
<th>PCR programme</th>
</tr>
</thead>
<tbody>
<tr>
<td>20IV</td>
<td>NGA1107 F</td>
<td>CGACGATTCACAGAATTTAGG</td>
<td>4</td>
<td>18096137 – 18096288</td>
<td>Ta: 54 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>NGA1107 R</td>
<td>GCGAAAAACAAAAAATCCA</td>
<td>4</td>
<td>16444151 – 16444264</td>
<td>Ta: 54 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>NGA1139 F</td>
<td>TTTTTCCTTGTGTGGCATCCCTG</td>
<td>4</td>
<td>16535546 – 16535805</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>NGA1139 R</td>
<td>TCGATAGGGTTGAGTTGACGGC</td>
<td>4</td>
<td>16816588 – 16816774</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>F23E12 F</td>
<td>TGCTTCTCCTGATCTTACTG</td>
<td>4</td>
<td>16853041 – 16853338</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>F23E12 R</td>
<td>GCCCTCTCTGACACCTTCTCAA</td>
<td>4</td>
<td>17110449 – 17110728</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>4_UPSC_16533 (F)</td>
<td>CGACGATTCACAGAATTTAGG</td>
<td>4</td>
<td>18570651 – 18570791</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>4_UPSC_16533 (R)</td>
<td>GCGAAAAACAAAAAATCCA</td>
<td>4</td>
<td>18570651 – 18570791</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>4_UPSC_16853 (F)</td>
<td>TCGATAGGGTTGAGTTGACGGC</td>
<td>4</td>
<td>17334688 – 17334927</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>4_UPSC_16853 (R)</td>
<td>GCCCTCTCTGACACCTTCTCAA</td>
<td>4</td>
<td>18570651 – 18570791</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>F21M12 F</td>
<td>TTTTTCCTTGTGTGGCATCCCTG</td>
<td>4</td>
<td>17334688 – 17334927</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>F21M12 R</td>
<td>TCGATAGGGTTGAGTTGACGGC</td>
<td>4</td>
<td>17334688 – 17334927</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>F3F19 F</td>
<td>GGGGATTGGGTTTGTTCG</td>
<td>4</td>
<td>3212189 – 3212189</td>
<td>Ta: 53 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>F3F19 R</td>
<td>GGGGATTGGGTTTGTTCG</td>
<td>4</td>
<td>3212189 – 3212189</td>
<td>Ta: 53 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>F9H16 F</td>
<td>AAAAACTGCGACGCGCGGAAT</td>
<td>1</td>
<td>7294861 – 7295100</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>F9H16 R</td>
<td>AAAAACTGCGACGCGCGGAAT</td>
<td>1</td>
<td>7294861 – 7295100</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>F36_1 F</td>
<td>TTGGTCCAGTTACTTTTGCTG</td>
<td>1</td>
<td>8604005 – 8604142</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>20IV</td>
<td>F36_1 R</td>
<td>TTGGTCCAGTTACTTTTGCTG</td>
<td>1</td>
<td>8604005 – 8604142</td>
<td>Ta: 55 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>30I</td>
<td>CIW12 F</td>
<td>AGGTTTTATGCTTCTCCACCA</td>
<td>1</td>
<td>9621357 – 9621484</td>
<td>Ta: 51 °C Elongation time: 15 s</td>
</tr>
<tr>
<td>30I</td>
<td>CIW12 R</td>
<td>AGGTTTTATGCTTCTCCACCA</td>
<td>1</td>
<td>9621357 – 9621484</td>
<td>Ta: 51 °C Elongation time: 15 s</td>
</tr>
<tr>
<td>30I</td>
<td>NGA280 F</td>
<td>GGGCTCCATAAAAAGTGACCC</td>
<td>1</td>
<td>20873698 – 20873802</td>
<td>Ta: 53 °C Elongation time: 20 s</td>
</tr>
<tr>
<td>30I</td>
<td>NGA280 R</td>
<td>GGGCTCCATAAAAAGTGACCC</td>
<td>1</td>
<td>20873698 – 20873802</td>
<td>Ta: 53 °C Elongation time: 20 s</td>
</tr>
</tbody>
</table>
4.13 DNA extraction for deep sequencing

Roughly 200 seedlings, with a total weight of 0.4-0.5 g, were ground in liquid nitrogen using mortar and pestle and transferred to Falcon tubes with ice cold Nuclei extraction buffer (10 mM Tris-HCl pH 9.5, 10 mM EDTA pH 8.0, 100 mM KCl, 500 mM sucrose, 4 mM spermidine, 1 mM spermine, 0.1 % 2-mercaptoethanol) and mixed thoroughly. The suspension was filtered through two layers of Miracloth into ice cold Falcon tubes and mixed with Lysis buffer (10 % Triton X-100 in Nuclei extraction buffer) for 2 min on ice. The nuclei were pelleted by centrifugation at 2000 g for 10 min. Pellet was resuspended in CTAB extraction buffer (100 mM Tris-HCl pH 7.5, 0.7 M NaCl, 10 mM EDTA, 1 % CTAB, 1 % 2-mercaptoethanol) to start denaturing proteins and release nuclear DNA before being transferred to a fresh Eppendorf tube. Each sample was incubated at 60 ºC for 30 min and allowed to cool in RT, 5 min, before chloroform:isoamylalcohol (24:1) was added and mixed by inversion for ~5 min to dissolve remaining proteins and lipids. The sample was centrifuged at 6000 rpm for 10 min and the upper phase containing nuclear DNA transferred to a fresh tube with isopropanol. DNA was pelleted by centrifugation at 13000 rpm for 3 min, followed by removal of the supernatant and then the pellet was washed in 70 % ethanol. Each sample was centrifuged at 13000 rpm for another 3 min, ethanol removed and DNA pellet allowed to dry before being resuspended in ddH2O with RNaseA (10μg mL⁻¹) and incubated at 65 ºC for 20 min to destroy DNases.

4.14 Data analysis

General growth phenotype measurements were performed in ImageJ 1.43 using pictures taken with Canon EOS 450D camera. Results were analysed using Dunnett's test in IBM SPSS Statistics 19 software as this ANOVA post hoc test allows for comparisons of several groups with one assigned control group [Zar, 2010]. The control group in this case was the AtMC9::nGFP reporter line.

5 Results

5.1 Mutants fall into two phenotypical categories

5.1.1 Screening of the EMS mutagenized reporter line

Mutants were generated through seed mutagenesis using ethyl methanesulfonate (EMS) of a Columbia (Col) reporter line with nuclear green fluorescent protein (nGFP) fused to the Metacaspase 9 promoter; AtMC9::nGFP reporter line see Figure 1. A total number of 10000 seeds were used to ensure that enough mutants with changes affecting the expression of AtMC9 were generated. In Arabidopsis the genetically functional cell number, or the number of cells in the embryonic shoot meristem that contributes to seed output, has been found to consist of two cells. This has been found to result in a chimeric M1 generation with two different sectors in which the mutations segregate differently depending on in which sector they are located. Therefore a recessive mutation in the M2 generation will segregate in a ratio of 7:1, and a dominant mutation with a ratio of 5:3, as described by Page and Grossniklaus (2002).

The M1 seeds were divided into 450 different pools containing approximately 20 plants each. From each of the 420 pools containing viable M1 plants the M2 seeds were harvested. Approximately 200 five-day-old M2 seedlings from each pool were screened searching for
seedling with changes in the GFP pattern. Based on the segregation pattern of recessive mutations and the 20 M1 plants per pool a total of 200 seeds would be needed to have any chance of finding homozygous mutants during the screening. M2 seedlings with a GFP pattern differing from that of AtMC9::nGFP were selected for further analysis. When mutants of interest reached ~30 days of age their phenotypes were documented. A mutant with a severely altered growth phenotype would be more likely to carry mutations in genes affecting plant fitness or growth in general without acting directly on expression of AtMC9, and were therefore excluded from this study. Those mutants fulfilling these requirements were allowed to self fertilize. In the end eleven mutant lineages, named 20IV, 30I, 47X, 50IX, 70I, 70II, 70III, 70IV, 80II, 81I and 81II after the pool number in which they were found, were chosen for further investigations.

Figure 1. Schematic picture of AtMC9::nGFP reporter line showing GFP signal in root cap, protoxylem and metaxylem.

5.1.2 GFP expression pattern in ectopic expressors and low/no-signal mutants

The screening of five-day-old M2 seedlings revealed that the mutant lineages seemed to fall into two different categories; the ectopic expressors with additional GFP signal in the cotyledons and the low/no-signal mutants that displayed a much weaker GFP signal in root cap, protoxylem and metaxylem as compared to that of the reporter line. Mutants 20IV, 30I, 70I, 70II and 70IV were all ectopic expressors while mutants 47X, 50IX, 70III, 80II, 81I and 81II showed only a very weak GFP signal in their roots. The low/no-signal mutants could further be divided into low-signal mutants consisting of mutants 50IX and 80II, and no-signal mutants where hardly any GFP was found in the root during the initial screening. This latter group consisted of mutants 47X, 70III, 81I and 81II.

Further examination of the GFP expression pattern in five-day-old M3 seedlings using confocal microscope could confirm that the AtMC9::nGFP reporter line lacked GFP signal in the cotyledons, see Figure 2A, but showed a clear signal in root cap, protoxylem and metaxylem, see Figure 3A. Mutant 20IV displayed GFP in their cotyledons, apparently randomly distributed on the same depth within the tissue as the endogenous chlorophyll signal, see Figure 2B.
Figure 2. Confocal pictures confirming that GFP was absent in cotyledons of AtMC9::nGFP but present in an ectopic expressor. The chlorophyll autofluorescence in five-day-old seedlings was detected within the wavelengths 670-720 nm (seen as red in the pictures), and GFP within 505-535 nm (seen as green stains in the pictures). A) AtMC9::nGFP did not show any specific GFP signal in the cotyledons while B) mutant 20IV did, in this case especially at the base and near the edges of the cotyledon. Arrows point out locations of GFP.

In mutant 47X only a weak GFP signal was found in root cap and protoxylem whereas signal was absent in metaxylem, as depicted in Figure 3B.
Figure 3. Confocal pictures showing the difference in root GFP pattern between AtMC9::nGFP and a no-signal mutant. Roots of five-day-old seedlings were stained with propidium iodide before examination. Propidium iodide was detected within wavelengths 600-700 nm (seen as red in the pictures), and GFP within 500-545 nm (seen as green stains in the pictures). A) AtMC9::nGFP reporter line with GFP signal in root cap, protoxylem and metaxylem. B) Mutant 47X, a no-signal mutant with GFP found mainly in root cap and protoxylem and lacking in metaxylem. Arrows point out locations of GFP.

5.2 Characterization of mutants

5.2.1 Checking GFP construct for mutations in mutants 20IV, 30I, 47X and 50IX

To ensure that the disrupted GFP pattern found in the mutant lineages was not caused by mutations in the GFP construct itself mutants 20IV, 30I, 47X and 50IX as well as the AtMC9::nGFP were investigated through sequencing. Constructs were amplified using PCR, followed by TA cloning to ensure sufficient DNA material and positive clones selected through blue and white screening of recombinants. To confirm that the positive white-coloured clones of the screening were carriers of the insert restriction digests were performed. The gel analysis of the restriction digestions together with one undigested sample showed all the expected fragments, where the coiled plasmid appeared as a band at 3.4 kb and the insert at 2.3 kb (see Figure 4 for details), confirming the presence of inserts from mutants 20IV, 30I, 47X and 50IX as well as the AtMC9::nGFP reporter line.

Figure 4. Restriction digests could confirm successful insertion of GFP constructs from mutants 20IV, 30I, 47X, 50IX and AtMC9::nGFP into the pGEM-T vector. A) Double-digest with NotI and NcoI showed fragments of 3.4 kb representing the plasmid, 2.3 kb being the insert and a fragment of 1.8 kb being a super coiled rest of the plasmid. B) Linearisation with ApaI showed fragments of 6 kb representing the plasmid with the insert and 1.8 kb again being the super coiled rest of the plasmid. C) The undigested sample showed one band at 10 kb and another at 7 kb both representing relaxed forms of the plasmid, and a band of 3.2 kb representing the coiled version of the plasmid and 1.8 kb the super coiled plasmid.
The sequenced fragments were fused together using the overlapping ends of each fragment to form full length sequences, as displayed in Figure 5. These sequences were compared with the full promoter sequence of At5g04200.1 in TAIR (http://www.arabidopsis.org/) fused with the sequence of the GFP insert in a MUSCLE alignment. The alignment, included in full length in Appendix 1, did contain some potential point mutations of which some would result in amino acid exchange. Mutant 20IV carried two potential point mutations, both only present in the forward sequencing reactions. One of these potential point mutations were located in the beginning of the attB1 sequence and therefore lacking a reverse sequence and in the other case the reverse sequence did not show any mutations at that position indicating that both these probably are due sequencing errors. Both mutants 30I and 47X had one potential point mutation each that would result in amino acid exchange in the GFP region of the construct. In mutant 50IX and AtMC9::nGFP reporter line two and three potential mutations respectively were found that would result in amino acid exchange. These deviations could be due to either actual mutations or sequencing errors.

Figure 5. The full length sequences of the GFP constructs were assembled using the overlapping ends of the fragments generated with attB1 (0.90 kb), attB2 (0.47 kb), eGFP (0.72 kb) and T-NOS (0.18 kb). pAtMC9 represents the promoter of the AtMC9 gene and had a size of approximately 1.4 kb. The GFP region had a size of approximately 0.9 kb. NOTE: Figure is not to scale.

5.2.2 General growth phenotype of the M3 generation

To study the potential effects of the mutations on overall growth two sets of plants were prepared with M3 generation mutants 20IV, 30I, 47X, 50IX, 70I, 70II, 70III, 70IV, 80II, 81I, 81II and AtMC9::nGFP of which one set was grown in darkness and the other in light. Root lengths were measured daily between days 4-8 after germination of the light grown plants and daily growth rate calculated as a mean over the whole measuring period. Three no-signal mutants grew significantly faster than the reporter line: 70III (t=-3.32078, p=0.012), 81I (t=-3.20307, p=0.017) and 81II (t=-4.33862, p<0.001), depicted in Figure 7. The mutants 20IV (t=3.28151, p=0.014), 30I (t=3.87041, p=0.002) and 70I (t=5.02053, p<0.001), all ectopic expressors, had significantly lower growth rates than AtMC9::nGFP over the studied time period, see Figure 6. Growth rates of mutants 47X, 50IX, 70II, 70IV and 80II did not differ significantly from that of the reporter line.
Figure 6. Roots of no-signal mutants generally have higher growth rates than AtMC9::nGFP, whereas ectopic expressors often grow more slowly. Error bars indicate standard deviation for mean growth rate during days four to eight after germination. Asterisks indicates p-values < 0.05 when comparing mutant lineages to AtMC9::nGFP. Degrees of freedom within groups were 128, and between groups 11.

To gain more specific information on which days that showed the largest variations between lineages root lengths at the different days were compared. At four days of age mutants 47X (t=-3.02640, p=0.029), 70III (t=-3.97603, p=0.001) and 81II (t=-3.62835, p=0.004) had longer roots than the reporter line. At five days after germination mutants 70I (t=2.95019, p=0.037) and 70III (t=-3.46077, p=0.008) both differed significantly from AtMC9::nGFP, where mutant 70I had a shorter root and 70III was longer. When comparing six-day-old seedlings mutants 47X (t=-3.38083, p=0.010), 70III (t=-4.22071, p<0.001), 81I (t=-3.48472, p=0.007) and 81II (t=4.16147, p=0.001) all had longer roots than the reporter line, while mutant 70I (t=3.11787, p=0.022) had significantly shorter roots. At seven days of age mutants 47X (t=-3.35819, p=0.011), 70III (t=-4.47621, p<0.001), 81I (t=-3.30068, p=0.013) and 81II (t=-4.35376, p<0.001) still were significantly longer than AtMC9::nGFP whereas mutants 30I (t=2.94700, p=0.037) and 70I (t=4.54518, p<0.001) both had shorter roots. Root lengths at eight days after germination differed significantly from the reporter line for mutants 47X (t=-3.14839, p=0.020), 70I (t=4.31892, p<0.001), 70III (t=-4.34937, p=0.001), 81I (t=-3.38016, p=0.010) and 81II (t=-4.98448, p<0.001) of which only mutant 70I had shorter roots. Altogether this means that a pattern starts to appear; the no-signal mutants 47X, 70III, 81I and 81II all show tendencies of having longer roots than AtMC9::nGFP reporter line while mutants 30I and 70I both appear to have shorter roots during the studied time period, as displayed in Figure 7.
Figure 7. No-signal mutants appear to have significantly longer roots compared to the reporter line. Total root lengths at eight days after germination with increments as indicated on the right side. Error bars indicate standard deviation. Asterisks indicate p-values < 0.05 when comparing mutant lineages to AtMC9::nGFP at eight days of age, p-values for the other days could be found in the text. Degrees of freedom within groups were 128, and between groups 11.

The heights of dark-grown hypocotyls of seven-day-old seedlings were also studied in all M3 generation mutant lineages and compared to AtMC9::nGFP. The height of mutant 80II (t=3.01011, p=0.032) was significantly smaller than that of the reporter line, see Figure 8, while the rest of the mutants did not differ significantly from AtMC9::nGFP.

Figure 8. Lengths of dark-grown hypocotyls only differs significantly in one low-signal mutant at seven days after germination. Error bars indicate standard deviation for mean height of hypocotyls and asterisk indicate p-value < 0.05 when comparing mutant lineages to AtMC9::nGFP. Degrees of freedom were 108 between groups and 11 within groups.

The leaf rosettes diameters, assessed by making three measurements across each rosette and calculating a mean for each plant, were also compared. Mutants 20IV, 47X, 50IX, 70III, 81I and 81II did not differ significantly in size while mutants 30I (t=7.396220, p<0.001), 70I (t=5.262222, p<0.001), 70II (t=4.833288, p<0.001), 70IV (t=5.330697, p<0.001) and 80II
(t=7.558367, p<0.001) all had significantly smaller rosette diameters than AtMC9::nGFP, as displayed in Figure 9.

![Figure 9](image1)

**Figure 9.** Several ectopic expressors and one low-signal mutant had significantly smaller rosette diameters at time of bolting. Error bars indicate standard deviation for mean rosette diameter. Asterisks indicates p-values < 0.05 when comparing mutant lineages to AtMC9::nGFP. The degrees of freedom within groups were 48 and that between groups were 11.

The first plant lineages started bolting at 30 days after germination, and all had bolted at 37 days of age, see Figure 10. When making comparisons between the current diameter of AtMC9::nGFP at the time points when the separate mutant lineages bolted nine mutant lineages had significantly smaller rosette diameters than the reporter line. These lineages were mutants 30I, 47X, 50IX, 70I, 70II, 70IV, 80II, 81I and 81II, see Appendix 2 for details. Mutant lineages 20IV and 70III had rosette diameters that did not significantly differ from that of AtMC9::nGFP, but they were also the first two mutants who bolted (at 32 and 30 days of age respectively) while the others bolted later when the reporter line had grown even bigger.

![Figure 10](image2)

**Figure 10.** Bolting times for mutant lineages varied between 30 to 37 days after germination. AtMC9::nGFP reporter line bolted at 30 days of age.
Even though the diameters of the rosettes varied at bolting time none of the mutant lineages showed colours that deviated distinctly from that of the reporter line, see Figure 11. Two out of five representatives of mutant 70I could possibly show a slight tendency towards yellowing at the leaf tips, but considering that these patterns are not observed in all plants they are likely not caused by the mutation.

Figure 11. Rosette diameters vary between plant lineages at time of bolting. Plants were considered to have bolted once they had a 1-3 cm high shoot. A) AtMC9::nGFP at 30 days of age. B) 20IV at 32 days of age. C) 30I at 33 days of age. D) 47X at 34 days of age. E) 50IX at 32 days of age. F) 70I at 37 days of age. G) 70II at 36 days of age. H) 70III at 30 days of age. I) 70IV at 37 days of age. J) 80II at 33 days of age. K) 81I at 34 days of age. L) 81II at 37
days of age. Scale bar indicates 2 cm.

When the plants were 40 days old one flower was taken from the primary shoot from one plant of each lineage and examined, again comparing the mutant lineages to the reporter line. No apparent developmental aberrations were found in any of the mutant lineages, see Figure 12.

Figure 12. None of the flowers from the mutant lineages showed any major phenotypical deviations from AtMC9::nGFP. A) AtMC9::nGFP. B) 20IV. C) 30I. D) 47X. E) 50IX. F) 70I. G) 70II. H) 70III. I) 70IV. J) 80II. K) 81I. L) 81II. Scale bar indicates 1 mm.
5.2.3 Vascular development in cotyledons appears to be unaffected by mutations

The vasculature of M3 generation mutants 20IV, 47X, 50IX, 70I, 70II, 70III, 70IV, 80II, 81I, 81II and homozygous mutants from F2 (Ler cross) 30I and AtMC9::nGFP were examined in six-day-old seedlings by clearing tissues with chloral hydrate. Two seedlings per plant lineage were examined comparing the mutants to AtMC9::nGFP reporter line, as shown in Figure 13. The number of fully connected vascular loops found in their cotyledons varied between two to five, but none of the mutant lineages deviated distinctly from the reporter line.

Figure 13. The number of fully connected vascular loops revealed that there was little difference between mutant lineages and AtMC9::nGFP. Six-day-old seedlings were cleared using chloral hydrate, mounted in chloral hydrate dissolved in glycerol and studied under Zeiss Axioplan microscope. A) AtMC9::nGFP with 2-4 fully connected vascular rings, or areoles which are the areas defined by the vascular loops. B) 20IV with 3-4 areoles. C) 30I

5.2.4 Ectopic expression of GFP in cotyledons is not associated with increased cell death

Cotyledons of ectopic expressors were examined searching for signs of increased levels of cell death. The M3 generation mutants 20IV, 70I, 70II, 70IV and homozygous mutants from F2 (Ler cross) of 30I and AtMC9::nGFP were stained with Trypan blue according to Weigel and Glazebrook (2002) at six days of age. The staining of the dead cells in root vasculature functioned as positive controls for the protocol, as can be seen in Figure 14. The cotyledons of the ectopic expressors did not show any signs of staining though, strongly indicating that the GFP signal in the cotyledons was not associated with increased cell death.

5.3 Locating mutations through mapping and deep sequencing

5.3.1 Crossings and segregation of mutant traits

The M2 generation mutants were allowed to self fertilize giving rise to an M3 generation. Investigations of these generations revealed that all eleven mutant lineages were carrying recessive mutations as all plants showed the mutant phenotypes. M3 generation plants were crossed into Landsberg (Ler) or Wassilewskija (Ws) ecotypes to generate mapping populations, where Ws crossings were performed as a backup in case problems with second-site modifiers would appear in the Ler crossing. Crossings were also performed with the AtMC9::nGFP reporter line for background cleaning, serving the purpose of removing mutations located in the background of the one affecting the phenotype to ensure that no other mutation contributed to the observed phenotype. As the M3 generation was found to contain only recessive mutations the F1 generations would be heterozygous for the mutant trait, not
showing the mutant phenotype at all. In case of the Ler and Ws crossings the GFP construct would also be heterozygous.

The success of F1 generation crossings were controlled by searching for contamination of other ecotypes through PCR with SSLP markers. Two different SSLP markers were used; SSLP mio24 distinguishing between Col and Ler and SSLP 1830 At4g distinguishing between Col and Ws. The AtMC9::nGFP backcrossings would be homozygous Col for both markers, the Ler crossings heterozygous for SSLP mio24 and homozygous for SSLP 1830 At4g and the Ws crossings heterozygous for both markers. Figure 15 shows the SSLP marker analysis for mutant 30I, containing the expected fragments indicating successful crossings and lacking signs of contaminations. The same experiment was performed with mutants 20IV, 47X and 50IX with similar results.

Figure 15. Genotyping of a selection of mutant 30I plants showed that the F1 crossings into AtMC9::nGFP (Col ecotype), Ler and Ws were successful and that no contaminations were present. Crossings of M3 mutant lineages into Ler and Ws were performed to generate mapping populations, where Ws would function as a backup in case that problems would arise in the Ler crossing. Backcross into AtMC9::nGFP would function as background cleaning, removing mutations other than the one of interest that could potentially affect the phenotype. Two SSLP markers were used to confirm successful crossings; SSLP mio24 separating between Col and Ler and SSLP 1830 At4g distinguishing between Col and Ws. A) SSLP mio24 gave fragments of 320 bp (upper band) in Col and 250 bp (lower band) in Ler and Ws. The two middle samples of Ws crossing did not show proper separation of the Col and Ws fragments, the upper band seems to be preferred. B) SSLP 1830 At4g resulted in fragments of 346 bp (upper band) in Col and Ler while showing bands at 279 bp (lower band) in Ws. Pure Col, Ler and Ws samples were included as controls (underlined in the pictures).

The F2 mapping population, based on the Ler crossings, was generated through self fertilization of F1 resulting in 25 % (4/16) with no GFP construct, 18.75 % (3/16) homozygous for GFP construct but heterozygous or non-carriers of the mutant trait, 18.75 % (3/16) homozygous mutants with GFP construct (either homozygous or heterozygous) and 37.5 % (6/16) being heterozygous for the GFP construct and either heterozygous or non-carriers of mutant trait, all displayed in Figure 16.
Figure 16. Inheritance pattern of GFP construct and mutant trait in F2 mapping population (Ler cross). T=with GFP construct. t=no GFP construct. M=no mutant trait. m=with mutant trait. F2 generation with white background (4/16) shows individuals lacking GFP construct entirely. Dark grey background (3/16) have both GFP construct and mutant phenotype. Light grey background (3/16) indicates individuals homozygous for GFP construct and heterozygous or non-carriers of mutation not showing mutant phenotype. Grey (6/16) boxes show individuals being heterozygous or non-carriers of mutation, so they are not showing mutant phenotype, and have a weaker GFP signal due to a heterozygous construct.

When studying F2 generation seedlings of mutant 20IV, an ectopic expressor with GFP signal appearing also in cotyledons, a total of 20.4 % (232/1137) were found to show the mutant phenotype which is slightly higher than the expected values for a recessive mutation. Corresponding number of mutants were counted also in 30I, another ectopic expressor, resulting in 19.4 % (230/1183) showing a mutant phenotype. In mutants 47X and 50IX, both with reduced GFP signal in the root compared to the reporter line, only 14.0 % (226/1620) and 14.8 % (215/1451) respectively with mutant phenotype could be found. These low numbers could probably be explained by the difficulties to accurately separate between the weak GFP signal in the mutants caused by the mutation and the weaker signals in seedlings being heterozygous for the GFP construct itself. Both mutants 47X and 50IX were grown on MS growth medium containing the herbicide phosphinotricin, PPT, to help distinguishing the seedlings completely lacking the GFP construct from the low/no-signal mutants. The GFP construct contained a Bar resistance gene making carriers of the construct resistant to PPT [Thompson et al., 1987]. In mutants 47X and 50IX a total of 28.1 % (456/1620) and 27.4 % (397/1451) respectively of the seedlings showed growth patterns that were affected by PPT. The expected number of seedlings that should not have been carriers of the GFP construct and thereby been affected by PPT was 25 %.

5.3.2 Mapping of mutants 20IV and 30I

In order to identify the the genes causing the altered GFP expression and thereby acting upstream of AtMC9 the positions of the mutations had to be located. The process of locating the mutations in mutants 20IV and 30I started with mapping. The mapping process was divided into two parts: the rough mapping to identify the approximate location of the mutation and the fine mapping narrowing down the area of interest even further. Mapping was performed using DNA material from whole ~10-day-old F2 generation (Ler cross) seedlings homozygous for the trait of interest. Homozygous mutants were identified on the stereo microscope with epifluorescence using filterset for GFP detection. Rough mapping was performed using the simple sequence length polymorphism (SSLP) markers described in
Lukowitz et al., (2000), polymorphic between the ecotypes Col and Ler, spread across the genome with 10-40 % recombinations distance. The markers were used to retrieve information about recombinations in the area of interest, where no recombinations were associated with homozygosity for Col, one recombination with a heterozygous Col/Ler genotype and two recombinations with homozygosity for Ler. DNA from ~20 homozygous mutant seedlings was used to locate the rough position of the mutation by searching for markers showing a high level of homozygosity for Col; the ecotype in which the mutation was induced. In mutant 20IV such an accumulation of Col background was found on the lower arm of chromosome 4, and for mutant 30I on upper arm of chromosome 1 strongly indicating that the mutations were also located in these areas.

The fine mapping started by finding two markers surrounding the mutation and then investigating the area between them using a third marker narrowing down the area of interest by three-point mapping. The mapping population had to be increased for the fine mapping to still find recombinations in proximity of the mutation; a total of 95 individuals were used for mutant 20IV and 56 individuals for mutant 30I. Three-point mapping was repeated with new markers until the mutation could be traced to an area of approximately 283 kb for mutant 20IV and 2800 kb for mutant 30I, placing the mutation between markers 4_UPSC_16533 and F23E12 in Figure 17 for mutant 20IV and between F3F19 and F9H16 for mutant 30I.

Figure 17. Overview of chromosomes and the SSLP markers used for mapping of mutants 20IV and 30I. First two markers flanking the mutation were located, then the area of interest could be narrowed down using three-point mapping studying a third marker located in the middle of the area of interest. A) Chromosome 4, showing the markers used for rough and fine mapping of mutant 20IV. B) Chromosome 1 with the markers used for mapping of mutant 30I. Markers labelled in green are those surrounding the mutations in the two mutants.
5.3.3 Deep sequencing of mutants 20IV, 30I, 47X and 50IX

Fine mapping could have proceeded further by expanding the mapping population and developing new SSLP markers within the regions of interest, but analysis of mutant lineages instead continued with deep sequencing as it was more cost-effective. Genomic DNA was extracted from ~12-day-old F2 generation seedlings of mutant lineages 20IV, 30I, 47X and 50IX and sent to Max Planck Institute in Tübingen, Germany, for deep sequencing. DNA from the two parental lines, that is AtMC9::nGFP and Ler that were crossed to generate the mapping populations, were also included as references to distinguish between naturally occurring single-nucleotide polymorphisms (SNPs) and those induced by EMS.

The results of the deep sequencing and SHOREmap analysis [briefly described in Schneeberger et al., 2009] showed the expected accumulation of Col ecotype alleles on the lower arm of chromosome 4 in mutant 20IV and upper arm of chromosome 1 in mutant 30I, as displayed in Figure 18. Mutations in mutants 47X and 50IX was both located on chromosome 5, on lower and upper arm respectively, see Figure 18.

![Figure 18. Deep sequencing of mutants A) 20IV and B) 30I revealed a high frequency of Col alleles on lower arm of chromosome 4 and upper arm of chromosome 1 respectively, confirming mapping results. C) Mutant 47X showed accumulation of Col alleles on lower arm of chromosome 5 whereas D) mutant 50IX appeared to have a slight accumulation of Col on lower arm of chromosome 5. The allele distribution peak of 50IX revealed that identification of homozygous mutants had not been entirely successful. Black lines indicate areas that have been removed in the diagrams of mutants 47X and 50IX.](image-url)
The mutant genomes were also investigated for presence of SNPs in proximity to the allele distribution peak where those located close to the peak were considered more likely to have caused the observed phenotype. Nonsynonymous mutations were also considered more interesting as these alter the amino acid sequence and are therefore more likely to cause changes in the phenotype. In mutant 20IV two of the nonsynonymous mutations found were located close to the allele distribution peak. The two genes, here referred to as 20GENE1 and 20GENE2, were located at 22116 and 26023 bp distance from the peak and both resulted from transition mutations replacing the original cytosine with a thymine. 20GENE1 encoded a protein involved in the biosynthesis of polyamines and 20GENE2 encoded a protein with acyltransferase activity. For mutant 30I two interesting mutations were found that resulted in amino acid exchange. The genes affected were 30GENE1 involved in the synthesis of abscisic acid (ABA) located 14842 bp from the peak and 30GENE2 encoding a transcription factor-related protein positioned 400607 bp from the peak. Both genes carried transition mutations were the original guanine had been replaced by adenine. Mutant 47X had two nonsynonymous mutations in proximity to the allele distribution peak: 47GENE1 encoding a protein involved in carbohydrate metabolism and 47GENE2 playing a role in lignin biosynthesis. In the first case a cytosine had been replaced by an adenine, and in the second gene an adenine had replaced the original cytosine. Table 2 contains some information retrieved from the deep sequencing and the amino acid exchanges resulting from the nonsynonymous mutations. Regarding mutant 50IX deep sequencing clearly revealed that there had been problems to correctly identify the homozygous mutants. Some potentially interesting candidate genes could still be found even though they could not be assigned to a certain distance from the allele distribution peak as the peak could only be detected when the resolution was rather low and no longer a safe measure to use. Genes of interest are listed in Table 2.

Table 2. Deep sequencing results showing candidate genes with nonsynonymous mutations.

<table>
<thead>
<tr>
<th>Mutant</th>
<th>Gene</th>
<th>Encoding</th>
<th>Distance from allele distribution peak (bp)</th>
<th>Number of reads with base change</th>
<th>% of reads with base change</th>
<th>Quality value (maximum 40)</th>
<th>Amino acid exchange</th>
</tr>
</thead>
<tbody>
<tr>
<td>20IV</td>
<td>20GENE1</td>
<td>Protein involved in polyamine biosynthesis</td>
<td>22116</td>
<td>1</td>
<td>100</td>
<td>9</td>
<td>Alanine → Threonine</td>
</tr>
<tr>
<td>20IV</td>
<td>20GENE2</td>
<td>Protein with acyltransferase activity</td>
<td>26023</td>
<td>1</td>
<td>100</td>
<td>3</td>
<td>Alanine → Valine</td>
</tr>
<tr>
<td>30I</td>
<td>30GENE1</td>
<td>Protein involved in ABA synthesis</td>
<td>14842</td>
<td>13</td>
<td>100</td>
<td>40</td>
<td>Glycine → Aspartic acid</td>
</tr>
<tr>
<td>30I</td>
<td>30GENE2</td>
<td>Transcription factor-related protein</td>
<td>400607</td>
<td>8</td>
<td>100</td>
<td>40</td>
<td>Alanine → Threonine</td>
</tr>
<tr>
<td>47X</td>
<td>47GENE1</td>
<td>Protein involved in carbohydrate metabolic process</td>
<td>9468</td>
<td>1</td>
<td>100</td>
<td>3</td>
<td>Aspartic acid → Tyrosine</td>
</tr>
<tr>
<td>47X</td>
<td>47GENE2</td>
<td>Protein involved in lignin biosynthesis</td>
<td>12179</td>
<td>1</td>
<td>100</td>
<td>9</td>
<td>Valine → Glutamic acid</td>
</tr>
<tr>
<td>50IX</td>
<td>50GENE1</td>
<td>Protein involved in lignin catabolic process</td>
<td>-</td>
<td>1</td>
<td>100</td>
<td>9</td>
<td>Threonine → Proline</td>
</tr>
<tr>
<td>50IX</td>
<td>50GENE2</td>
<td>RING/U-box superfamily protein</td>
<td>-</td>
<td>1</td>
<td>100</td>
<td>2</td>
<td>Isoleucine → Asparagine</td>
</tr>
</tbody>
</table>
6 Discussion

This study aimed at locating and start characterizing upstream actors of AtMC9 by forward genetics using a reporter line with GFP fused to the promoter of Metacaspase 9. Thereby the expression pattern of GFP, and indirectly also of AtMC9, could be examined in plant tissues of the mutant lineages generated through EMS seed mutagenesis of the reporter line. A total of eleven mutant lineages were examined of which five were ectopic expressors showing additional GFP signal randomly distributed in cotyledons and six were low/no-signal mutants with a much weaker GFP signal in the root compared to the original reporter line. These mutants were chosen because their GFP expression patterns were altered when comparing them to the reporter line, but at the same time had a growth phenotype that was similar to that of the AtMC9::nGFP reporter line.

The GFP signal in the AtMC9::nGFP reporter line was found in xylem tissue and root cap, both known to undergo cell death. The low/no-signal mutants of this study showed severely altered GFP expression patterns in protoxylem, metaxylem and/or root cap compared to the AtMC9::nGFP reporter line, indicating an altered developmental pattern of root xylem vessels in these mutants. When examining the root lengths of seedlings the four no-signal mutant lineages had longer roots than AtMC9::nGFP, at least between days six to eight after germination. This coincides with the time of metaxylem differentiation [Pyo et al., 2004] suggesting that especially the metaxylem should be examined closer in the no-signal mutants.

Deep sequencing of mutant 47X resulted in one candidate gene, 47GENE2, encoding a protein involved in lignin biosynthesis. Staining mutant seedlings with phloroglucinol and examining them under the microscope would reveal whether lignification in the roots is affected by a mutation or not. Lignin synthesis have been shown to be under control of SECONDARY WALL-ASSOCIATED NAC DOMAIN PROTEIN1 (SND1) and its homologs, like VASCULAR-RELATED NAC-DOMAIN6 (VND6) and VND7, known to be master regulators of secondary wall biosynthesis [McCarthey et al., 2009] and also controlling the expression of AtMC9 [Zhong et al., 2010]. Ohashi-Ito et al., (2010) found that SND1 appeared to regulate genes involved in lignin monomer synthesis, like PAL1 (phenylalanine ammonia-lyase1) and being more specific to xylem fibers whereas VND6 was controlling genes specifically expressed in xylem vessels as for instance XCP1 (xylem cysteine protease1) and XCP2, and together they regulated genes involved in the common mechanisms of secondary wall biosynthesis like the two transcription factors MYB83 and MYB46 [McCarthey et al., 2009] and peroxidases and laccases [Ohashi-Ito et al., 2010]. Further they found indications that VND6 could recognize tracheary-element-regulating cis-element (TERE) sequences and activate the promoters they were located in. This could to some extent explain how secondary
wall formation, including lignification, and PCD is coordinated in TEs [Ohashi-Ito et al., 2010; Pyo et al., 2007] even if it does not give full insight into the exact mechanism connecting lignification and expression of \textit{AtMC9}. Yamaguchi et al., (2008) found in their study that a large fraction of plants carrying a truncated form of \textit{VND7} connected to the \textit{VND7} promoter and yellow fluorescent protein had a dwarf phenotype in the aerial tissues and discontinuous protoxylem in the roots. Several plants also had a disrupted metaxylem formation in root and abnormal vessel formation in aerial tissues. Chloral hydrate clearing of the mutant lineages examined here revealed that the vasculature in cotyledons appeared to be unaffected by the mutations, but further studies of root xylem would be needed with focus on developmental pattern and lignification of the secondary cell walls. Interesting to notice is that deep sequencing of another low/no-signal mutant, 50IX, also resulted in a candidate gene involved in lignin processing; \textit{50GENE1}.

Mutations in the GFP construct could also explain the altered GFP expression in the mutant lineages. TA cloning and sequencing of mutants 20IV, 30I, 47X and 50IX as well as the \textit{AtMC9}::nGFP reporter line did indicate the presence of potential point mutations. In mutant 20IV the two point mutations were only found in forward sequencing reactions, where in one case a reverse sequence was lacking and in the other the reverse sequencing reaction did not have this potential mutation which could mean that these are due to sequencing errors. In mutants 30I, 47X and 50IX one or two point mutations were found that were present in both forward and reverse fragments. This could again be due to mistakes during the sequencing or it could indicate presence of mutations in the GFP constructs. It should however be emphasized that also the reporter line carried potential point mutations in the GFP construct. These potential mutations were not present in the mutant lineages, even though they were generated through EMS mutagenesis of the very same reporter line, which would again indicate sequencing errors.

Trypan blue staining of the ectopic expressors showed that the GFP signal in the cotyledons was not associated with cell death. Using confocal microscopy the GFP signal was located in the parenchyma, as it was found at the same depth in the tissue as chlorophyll. Further studies (pers. comm., Jakob Prestele, Umeå Plant Science Centre) revealed that the signal was only visible when looking at the adaxial side of the cotyledon and not at all from the abaxial side, placing the signal in the palisade parenchyma. Deep sequencing of the ectopic expressor mutant 20IV resulted in two potentially interesting candidate genes, both located within the region identified during the mapping process. One of them was involved in the synthesis of polyamines. In Arabidopsis which lacks the ability to produce polyamines via the ornithine decarboxylase pathway there is only one pathway for synthesis of polyamines (PAs), the arginine carboxylase (ACD) pathway [Hanfrey et al., 2001]. In the ACD pathway putrescine is formed from arginine via agmatine and N-carbamoylputrescine [Tiburcio et al., 1997]. Putrescine can then be used to synthesize spermidine and spermine. Muñiz et al., (2008) suggested a role for the polyamine thermospermine in repressing PCD in vessel elements until their differentiation is complete. \textit{ACL5}, encoding a thermospermine synthase [Kakehi et al., 2008; Knott et al., 2007], have been shown to have a major influence over xylem specification and PCD [Muñiz et al., 2008]. The \textit{acl5} mutant show expression of \textit{XCP2}, a xylem-specific papain-type cysteine protease involved in cell death [Funk et al., 2002; Zhao et al., 2000], also in immature vessel elements unlike the wild type (WT) plants [Muñiz et al., 2008]. Further, levels of polyamines appear to increase in response to \textit{Puccinia striiformis} infection in wheat, and are suggested to be involved in formation of reactive oxygen species (ROS) through the activities of diamine oxidase and polyamine oxidase. ROS are critical for polymerisation of lignin, and there are indications that lignin formation
increases in response to infection [Asthir et al., 2010; Passardi et al., 2005]. Genes involved in PA synthesis have been shown to also play a role in abiotic stress responses, for example some of these genes are upregulated in response to cold in the non-freezing range [reviewed in Alcázar et al., 2011; Usadel et al., 2008]. This allows for physiological tests of mutant 20IV by exposing plants to cold followed by examinations of expression of PA genes through qRT-PCR to see if they are upregulated, as would be the case in WT plants, or not.

Two candidate genes were found in mutant 30I. Based solely on the distance from the Col accumulation peak 30GENE1 involved in the biosynthesis of abscisic acid would be the most likely candidate. Previous studies of aba mutants have shown that these plants often have a smaller wilt phenotype, higher transpiration rate and have reduced accumulation of ABA during drought stress, often reduced seed dormancy and are more sensitive to freezing than wild type plants [Nambara et al., 1998; Gilmour and Thomashow, 1991; Léon-Kloosterziel et al., 1996; Rock, 2000]. When comparing this to what is known about general growth phenotype of mutant 30I at least one obvious coinciding trait could be found in the form of the small plant size here assessed as rosette diameter at time of bolting. A strong indication of disruptions in ABA biosynthesis is the lower levels of ABA in plant tissues during drought stress [Léon-Kloosterziel et al., 1996]. ABA is critical for stomatal closure during such stress conditions to reduce water loss [Mustilli et al., 2002] making a mutant with disrupted ABA synthesis quite sensitive to drought. Exposing mutant 30I to drought, measuring the water loss and comparing that to how much water that is lost in a WT plant could indicate whether mutations in 30GENE1 is causing the phenotype in this mutant lineage or not. Would the mutant also recover and show a WT-like phenotype when supplied with exogenous ABA it would confirm an abnormal ABA biosynthesis. ABA has also been shown to affect the polyamine metabolism during water stress [Alcázar et al., 2006a]. Genes important for polyamine synthesis were upregulated in WT Arabidopsis plants during water stress, whereas a much smaller increase was found in aba mutants. Some of the genes involved in PA biosynthesis have also been shown to contain ABA-responsive elements (ABRE) or ABRE-related motifs [reviewed in Alcázar et al., 2006b]. This connection between ABA signalling and PA synthesis could explain the similarities in GFP expression patterns between mutants 20IV and 30I.

The three genes discussed above are the most interesting of the candidate genes at this stage, especially as previous studies have revealed potential connections between ABA and polyamine synthesis, and also between polyamines and lignification through the formation of ROS and their role in lignin deposition. The involvement of these genes in AtMC9 regulation is however not confirmed. In order to find out exactly which genes that are causing the observed phenotypes in mutants 20IV, 30I, 47X and 50IX all candidate genes must be experimentally validated. Study of mutants carrying T-DNA inserts, for instance SALK mutants, disrupting the expression of the gene of interest would be an alternative. By ordering one T-DNA mutant for each gene of interest listed in Table 2 and crossing these with the AtMC9::nGFP reporter line and allowing the resulting plants to self fertilize approximately 18% of the offspring would be homozygous for the mutation and also carry at least one copy of the GFP construct, as previously described for the F2 mapping population (see Figure 16). Close examination of the GFP pattern in these homozygous mutant plants using stereo microscope with epifluorescence and filterset for GFP detection or confocal microscope would reveal which mutants that have a similar AtMC9 expression pattern as the mutant lineage in which the candidate gene was identified. Further confirmation of the candidate genes would be gained by conducting transformations to reintroduce fully functional forms of the candidate genes into mutants 20IV, 30I, 47X and 50IX to see which genes that can rescue
the phenotype, that is restore the WT phenotype, in the mutant lineage. This would for certain identify the genes controlling the \textit{AtMC9} expression. Once the identities of the involved genes have been found the next step will be to find out exactly how they interact to control the process of vessel differentiation and programmed cell death.

7 Acknowledgements

Many thanks to my supervisor Hannele Tuominen and the helpful people at Umeå Plant Science Centre providing guidance and advice whenever needed, in particular I would like to mention Jakob Prestele, Daniel Pacurar and Kjell Olofsson whose assistance and advice during the practical work were of great help. Also to my professors and teachers at Linköping University, especially Lars Westerberg for advice on the statistical methods.

8 References


Appendix 1

*MUSCLE* alignment of GFP construct in mutant lineages and AtMC9::nGFP compared to the promoter sequence of At5g04200.1 (TAIR) fused with GFP (pAt5g04200.1_eGFP in the alignment). Alignment was run in BOXSHADE 3.21 for easier visualization of similarities and differences. Shaded areas indicate similar bases at a certain location. attB2 sites are labelled in turquoise and as they overlap with the GFP promoter sequence the regions unique to GFP promoter are displayed in dark blue. eGFP forward sites are displayed with red shading, while the eGFP reverse sites are indicated in magenta. The letters above the alignment (found only in the GFP-region) indicates amino acids and possible amino acid changes. The asterisk indicates a potential frameshift in mutant 20IV, but considering that this loss of a thymine is only present in the forward transcript and not the reverse (not shown) it is most likely a sequencing error.

<table>
<thead>
<tr>
<th>Lineage</th>
<th>AtMC9nGFP</th>
<th>30I</th>
<th>20IV</th>
<th>pAt5g4200.1_eGFP</th>
<th>47X</th>
<th>50IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtMC9nGFP</td>
<td>1</td>
<td>36</td>
<td>30I</td>
<td>43</td>
<td>42</td>
<td>60</td>
</tr>
<tr>
<td>30I</td>
<td>1</td>
<td>43</td>
<td>30I</td>
<td>42</td>
<td>41</td>
<td>41</td>
</tr>
<tr>
<td>20IV</td>
<td>1</td>
<td>103</td>
<td>30I</td>
<td>102</td>
<td>101</td>
<td>101</td>
</tr>
<tr>
<td>pAt5g4200.1_eGFP</td>
<td>1</td>
<td>120</td>
<td>30I</td>
<td>120</td>
<td>101</td>
<td>101</td>
</tr>
<tr>
<td>47X</td>
<td>1</td>
<td>180</td>
<td>30I</td>
<td>180</td>
<td>101</td>
<td>101</td>
</tr>
<tr>
<td>50IX</td>
<td>1</td>
<td>156</td>
<td>30I</td>
<td>156</td>
<td>101</td>
<td>101</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lineage</th>
<th>AtMC9nGFP</th>
<th>30I</th>
<th>20IV</th>
<th>pAt5g4200.1_eGFP</th>
<th>47X</th>
<th>50IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtMC9nGFP</td>
<td>216</td>
<td>323</td>
<td>222</td>
<td>240</td>
<td>221</td>
<td>221</td>
</tr>
<tr>
<td>30I</td>
<td>216</td>
<td>323</td>
<td>222</td>
<td>240</td>
<td>221</td>
<td>221</td>
</tr>
<tr>
<td>20IV</td>
<td>216</td>
<td>323</td>
<td>222</td>
<td>240</td>
<td>221</td>
<td>221</td>
</tr>
<tr>
<td>pAt5g4200.1_eGFP</td>
<td>216</td>
<td>323</td>
<td>222</td>
<td>240</td>
<td>221</td>
<td>221</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Lineage</th>
<th>AtMC9nGFP</th>
<th>30I</th>
<th>20IV</th>
<th>pAt5g4200.1_eGFP</th>
<th>47X</th>
<th>50IX</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtMC9nGFP</td>
<td>276</td>
<td>283</td>
<td>281</td>
<td>300</td>
<td>281</td>
<td>281</td>
</tr>
<tr>
<td>30I</td>
<td>276</td>
<td>283</td>
<td>281</td>
<td>300</td>
<td>281</td>
<td>281</td>
</tr>
<tr>
<td>20IV</td>
<td>276</td>
<td>283</td>
<td>281</td>
<td>300</td>
<td>281</td>
<td>281</td>
</tr>
<tr>
<td>pAt5g4200.1_eGFP</td>
<td>276</td>
<td>283</td>
<td>281</td>
<td>300</td>
<td>281</td>
<td>281</td>
</tr>
</tbody>
</table>
AtMC9nGFP 876 TAGTTGGCAATGCAATTGACATTATTTTCAACAATGTATACGTTATATACATTCGTATATCAT
30I 883 TAGTTGGCAATGCAATTGACATTATTTTCAACAATGTATACGTTATATACATTCGTATATCAT
20IV 882 TAGTTGGCAATGCAATTGACATTATTTTCAACAATGTATACGTTATATACATTCGTATATCAT
pAt5g4200.1_eGFP 900 TAGTTGGCAATGCAATTGACATTATTTTCAACAATGTATACGTTATATACATTCGTATATCAT
47X 881 TAGTTGGCAATGCAATTGACATTATTTTCAACAATGTATACGTTATATACATTCGTATATCAT
50IX 881 TAGTTGGCAATGCAATTGACATTATTTTCAACAATGTATACGTTATATACATTCGTATATCAT
AtMC9nGFP 936 AGCATTTAAAGTTAATTAGCTCCTCTCTTCCCTTAGATAGCTAGATGGATTAGGTCTAAA
30I 943 AGCATTTAAAGTTAATTAGCTCCTCTCTTCCCTTAGATAGCTAGATGGATTAGGTCTAAA
20IV 942 AGCATTTAAAGTTAATTAGCTCCTCTCTTCCCTTAGATAGCTAGATGGATTAGGTCTAAA
pAt5g4200.1_eGFP 960 AGCATTTAAAGTTAATTAGCTCCTCTCTTCCCTTAGATAGCTAGATGGATTAGGTCTAAA
47X 941 AGCATTTAAAGTTAATTAGCTCCTCTCTTCCCTTAGATAGCTAGATGGATTAGGTCTAAA
50IX 941 AGCATTTAAAGTTAATTAGCTCCTCTCTTCCCTTAGATAGCTAGATGGATTAGGTCTAAA
AtMC9nGFP 996 TATTTGGGAAAATGACTTCCCTTTTGAAGTTACGTATATATCCAAGTTAACCAAATTAAT
30I 1003 TATTTGGGAAAATGACTTCCCTTTTGAAGTTACGTATATATCCAAGTTAACCAAATTAAT
20IV 1002 TATTTGGGAAAATGACTTCCCTTTTGAAGTTACGTATATATCCAAGTTAACCAAATTAAT
pAt5g4200.1_eGFP 1020 TATTTGGGAAAATGACTTCCCTTTTGAAGTTACGTATATATCCAAGTTAACCAAATTAAT
47X 1001 TATTTGGGAAAATGACTTCCCTTTTGAAGTTACGTATATATCCAAGTTAACCAAATTAAT
50IX 1001 TATTTGGGAAAATGACTTCCCTTTTGAAGTTACGTATATATCCAAGTTAACCAAATTAAT
AtMC9nGFP 1056 CATGAATCCATGACTTAGGATTAATTATGTCTTTCTCAAAAGTCTGATTTAAGGTTTAGT
30I 1063 CATGAATCCATGACTTAGGATTAATTATGTCTTTCTCAAAAGTCTGATTTAAGGTTTAGT
20IV 1062 CATGAATCCATGACTTAGGATTAATTATGTCTTTCTCAAAAGTCTGATTTAAGGTTTAGT
pAt5g4200.1_eGFP 1080 CATGAATCCATGACTTAGGATTAATTATGTCTTTCTCAAAAGTCTGATTTAAGGTTTAGT
47X 1061 CATGAATCCATGACTTAGGATTAATTATGTCTTTCTCAAAAGTCTGATTTAAGGTTTAGT
50IX 1061 CATGAATCCATGACTTAGGATTAATTATGTCTTTCTCAAAAGTCTGATTTAAGGTTTAGT
AtMC9nGFP 1116 TTCTTGCTTGAAAGCATTAATAAGTCAAGAGTAAGCCATTAAAAAATGTTCCATAGCTTA
30I 1123 TTCTTGCTTGAAAGCATTAATAAGTCAAGAGTAAGCCATTAAAAAATGTTCCATAGCTTA
20IV 1122 TTCTTGCTTGAAAGCATTAATAAGTCAAGAGTAAGCCATTAAAAAATGTTCCATAGCTTA
pAt5g4200.1_eGFP 1140 TTCTTGCTTGAAAGCATTAATAAGTCAAGAGTAAGCCATTAAAAAATGTTCCATAGCTTA
47X 1121 TTCTTGCTTGAAAGCATTAATAAGTCAAGAGTAAGCCATTAAAAAATGTTCCATAGCTTA
50IX 1121 TTCTTGCTTGAAAGCATTAATAAGTCAAGAGTAAGCCATTAAAAAATGTTCCATAGCTTA
AtMC9nGFP 1176 AAAGCAAACTAGACTTGAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
30I 1183 AAAGCAAACTAGACTTGAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
20IV 1182 AAAGCAAACTAGACTTGAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
pAt5g4200.1_eGFP 1200 AAAGCAAACTAGACTTGAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
47X 1182 AAAGCAAACTAGACTTGAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
50IX 1182 AAAGCAAACTAGACTTGAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
AtMC9nGFP 1236 ACGAAACCTCAATCATAAATACACATTAGCTCCAAAACCACAACACTGCACATGAAAACT
30I 1243 ACGAAACCTCAATCATAAATACACATTAGCTCCAAAACCACAACACTGCACATGAAAACT
20IV 1242 ACGAAACCTCAATCATAAATACACATTAGCTCCAAAACCACAACACTGCACATGAAAACT
pAt5g4200.1_eGFP 1260 ACGAAACCTCAATCATAAATACACATTAGCTCCAAAACCACAACACTGCACATGAAAACT
47X 1241 ACGAAACCTCAATCATAAATACACATTAGCTCCAAAACCACAACACTGCACATGAAAACT
50IX 1241 ACGAAACCTCAATCATAAATACACATTAGCTCCAAAACCACAACACTGCACATGAAAACT
AtMC9nGFP 1296 AACACACAGTAGACTGAAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
30I 1303 AACACACAGTAGACTGAAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
20IV 1302 AACACACAGTAGACTGAAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
pAt5g4200.1_eGFP 1320 AACACACAGTAGACTGAAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
47X 1301 AACACACAGTAGACTGAAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
50IX 1301 AACACACAGTAGACTGAAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
AtMC9nGFP 1356 TAAACACAGTAGACTGAAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
30I 1363 TAAACACAGTAGACTGAAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
20IV 1362 TAAACACAGTAGACTGAAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
pAt5g4200.1_eGFP 1380 TAAACACAGTAGACTGAAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
47X 1361 TAAACACAGTAGACTGAAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC
50IX 1361 TAAACACAGTAGACTGAAAAATATAAATACTAATATCGTATGTAACTTGAACTTCACTAC

33
<table>
<thead>
<tr>
<th>Region</th>
<th>Start</th>
<th>End</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>AtMC9nGFP</td>
<td>1416</td>
<td>1476</td>
<td>TATGTTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGG</td>
</tr>
<tr>
<td>AtMC9nGFP</td>
<td>1476</td>
<td>1536</td>
<td>AGCGCGACGTAAACGGCCACAAGTTCAGCGTGTCGGCGAGGGCGAGGGCGATGCCACCT</td>
</tr>
<tr>
<td>AtMC9nGFP</td>
<td>1536</td>
<td>1596</td>
<td>ACCCTCGTGACCACCTTCACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGA</td>
</tr>
<tr>
<td>AtMC9nGFP</td>
<td>1596</td>
<td>1656</td>
<td>AGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCT</td>
</tr>
<tr>
<td>AtMC9nGFP</td>
<td>1656</td>
<td>1716</td>
<td>TCTTCAAGGAACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCC</td>
</tr>
<tr>
<td>AtMC9nGFP</td>
<td>1716</td>
<td>1776</td>
<td>TGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGC</td>
</tr>
<tr>
<td>AtMC9nGFP</td>
<td>1776</td>
<td>1836</td>
<td>AGAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGA</td>
</tr>
<tr>
<td>AtMC9nGFP</td>
<td>1836</td>
<td>1916</td>
<td>ACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGA</td>
</tr>
<tr>
<td>pAt5g4200.1_eGFP</td>
<td>1440</td>
<td>1500</td>
<td>ATGGTGAGCAAGGGCGAGGAGCTGTTCACCGGGGTGGTGCCCATCCTGGTCGAGCTGG</td>
</tr>
<tr>
<td>pAt5g4200.1_eGFP</td>
<td>1500</td>
<td>1560</td>
<td>ACCCTCGTGACCACCTTCACCTACGGCGTGCAGTGCTTCAGCCGCTACCCCGACCACATGA</td>
</tr>
<tr>
<td>pAt5g4200.1_eGFP</td>
<td>1560</td>
<td>1620</td>
<td>AGCAGCACGACTTCTTCAAGTCCGCCATGCCCGAAGGCTACGTCCAGGAGCGCACCATCT</td>
</tr>
<tr>
<td>pAt5g4200.1_eGFP</td>
<td>1620</td>
<td>1680</td>
<td>TCTTCAAGGAACGACGGCAACTACAAGACCCGCGCCGAGGTGAAGTTCGAGGGCGACACCC</td>
</tr>
<tr>
<td>pAt5g4200.1_eGFP</td>
<td>1680</td>
<td>1740</td>
<td>TGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGC</td>
</tr>
<tr>
<td>pAt5g4200.1_eGFP</td>
<td>1740</td>
<td>1800</td>
<td>AGAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGA</td>
</tr>
<tr>
<td>pAt5g4200.1_eGFP</td>
<td>1800</td>
<td>1860</td>
<td>ACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCAGAAGA</td>
</tr>
</tbody>
</table>

Note: The sequence information is presented in uppercase letters, which is common in biological research to denote DNA.
**AtMC9nGFP**

```
1896  TASA  -TTSRSATTSRATATAC
1896  ACTGAGGACGGCAGCGTGCAGCTCG
1902  ACTTCAAGATCCGCCACAAC
1902  AtMC9nGFP
1902  ACGGCATCAAGGTGA
1902  30I
1903  ACTTCAAGATCCGCCACAAC
1903  AtMC9nGFP
1903  ACGGCATCAAGGTGA
1903  20IV
1902  ACTGAGGACGGCAGCGTGCAGCTCG
1902  AtMC9nGFP
1902  ACGGCATCAAGGTGA
1902  47X
1901  ACTTCAAGATCCGCCACAAC
1901  AtMC9nGFP
1901  ACGGCATCAAGGTGA
1901  50IX
1901  ACTTCAAGATCCGCCACAAC
1901  AtMC9nGFP
1901  ACGGCATCAAGGTGA
```

```
1956  TCTCCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
1962  AtMC9nGFP
1962  CCACCCAGGATCGAGCTGAAG
1962  30I
1963  TCTCCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
1963  AtMC9nGFP
1963  CCACCCAGGATCGAGCTGAAG
1963  20IV
1962  TCTCCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
1962  AtMC9nGFP
1962  CCACCCAGGATCGAGCTGAAG
1962  47X
1961  TCTCCTGAGCACCCAGGATCGAGCTGAAG
1961  AtMC9nGFP
1961  CCACCCAGGATCGAGCTGAAG
1961  50IX
1961  TCTCCTGAGCACCCAGGATCGAGCTGAAG
1961  AtMC9nGFP
1961  CCACCCAGGATCGAGCTGAAG
```

```
2016  TTTT   -APSPP   -AKTPTRSTSAITW
2023  AtMC9nGFP
2023  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2022  30I
2022  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2022  AtMC9nGFP
2022  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2022  20IV
2021  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2021  AtMC9nGFP
2021  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2021  47X
2021  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2021  AtMC9nGFP
2021  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2021  50IX
2021  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2021  AtMC9nGFP
2021  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
```

```
2196  TTTT   -APSPP   -AKTPTRSTSAITW
2203  AtMC9nGFP
2203  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2202  30I
2202  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2202  AtMC9nGFP
2202  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2202  20IV
2201  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2201  AtMC9nGFP
2201  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2201  47X
2201  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2201  AtMC9nGFP
2201  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2201  50IX
2201  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2201  AtMC9nGFP
2201  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
```

```
2256  AATTT-----------------
2263  AtMC9nGFP
2263  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2262  30I
2262  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2262  AtMC9nGFP
2262  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2262  20IV
2261  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2261  AtMC9nGFP
2261  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2261  47X
2261  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2261  AtMC9nGFP
2261  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2261  50IX
2261  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
2261  AtMC9nGFP
2261  ACTACCTGAGCACCCAGTCCGCCCTGAGCAAAGACCCCAACGAGAAGCGCGATCACATGG
```
Appendix 2

Statistical tests on the diameter of leaf rosette at the time of bolting in eleven different mutant lineages compared to AtMC9::nGFP at the corresponding age. StDev=Standard deviation. n.s.=not significant. D.f.=degrees of freedom.

<table>
<thead>
<tr>
<th>Plant lineage</th>
<th>Age at time of bolting (days)</th>
<th>Mean rosette diameter (cm) ± StDev</th>
<th>Test value</th>
<th>D.f. (within groups; between groups)</th>
</tr>
</thead>
<tbody>
<tr>
<td>20IV</td>
<td>32</td>
<td>8.43600±0.580926</td>
<td>Dunnett's test, t=1.498272, n.s.</td>
<td>12; 2</td>
</tr>
<tr>
<td>30I</td>
<td>33</td>
<td>3.63760±0.194739</td>
<td>Dunnett's test, t=45.270450, p&lt;0.001</td>
<td>12; 2</td>
</tr>
<tr>
<td>47X</td>
<td>34</td>
<td>8.84480±0.521442</td>
<td>Dunnett's test, t=5.083458, p&lt;0.001</td>
<td>16; 3</td>
</tr>
<tr>
<td>50IX</td>
<td>32</td>
<td>6.88660±0.512714</td>
<td>Dunnett's test, t=3.329123, p&lt;0.001</td>
<td>12; 2</td>
</tr>
<tr>
<td>70I</td>
<td>37</td>
<td>5.07740±2.798297</td>
<td>Dunnett's test, t=5.495572, p&lt;0.001</td>
<td>12; 2</td>
</tr>
<tr>
<td>70II</td>
<td>36</td>
<td>5.36680±1.587870</td>
<td>T-test, t=7.497, p=0.001</td>
<td>4.997</td>
</tr>
<tr>
<td>70III</td>
<td>30</td>
<td>8.30180±0.465364</td>
<td>T-test, t=1.092, n.s.</td>
<td>8</td>
</tr>
<tr>
<td>70IV</td>
<td>37</td>
<td>5.03120±0.943939</td>
<td>Dunnett's test, t=5.537709, p&lt;0.001</td>
<td>12; 2</td>
</tr>
<tr>
<td>80II</td>
<td>33</td>
<td>3.52820±0.157320</td>
<td>Dunnett's test, t=46.076640, p&lt;0.001</td>
<td>12; 2</td>
</tr>
<tr>
<td>81I</td>
<td>34</td>
<td>7.44700±0.722549</td>
<td>Dunnett's test, t=8.402308, p&lt;0.001</td>
<td>16; 3</td>
</tr>
<tr>
<td>81II</td>
<td>34</td>
<td>7.77380±0.714803</td>
<td>Dunnett's test, t=7.626374, p&lt;0.001</td>
<td>16; 3</td>
</tr>
<tr>
<td>AtMC9::nGFP</td>
<td>30</td>
<td>8.62780±0.478650</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AtMC9::nGFP</td>
<td>32</td>
<td>8.88560±0.273871</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AtMC9::nGFP</td>
<td>33</td>
<td>9.78080±0.274657</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AtMC9::nGFP</td>
<td>34</td>
<td>10.98580±0.684775</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AtMC9::nGFP</td>
<td>36</td>
<td>11.01780±0.564918</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>AtMC9::nGFP</td>
<td>37</td>
<td>11.10280±0.542567</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>