Characterization of the Function of the AtST4 Subfamily Members in Cytokinin-Dependent Growth Control in Arabidopsis thaliana

Effat Khodashenas

A Thesis
in
The Department
of
Biology

Presented in Partial Fulfillment of the Requirements for the Degree of Master of Science (Biology) at Concordia University Montreal, Quebec, Canada

March 2010

© Effat Khodashenas, 2010
NOTICE:

The author has granted a non-exclusive license allowing Library and Archives Canada to reproduce, publish, archive, preserve, conserve, communicate to the public by telecommunication or on the Internet, loan, distribute and sell theses worldwide, for commercial or non-commercial purposes, in microform, paper, electronic and/or any other formats.

The author retains copyright ownership and moral rights in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

In compliance with the Canadian Privacy Act some supporting forms may have been removed from this thesis.

While these forms may be included in the document page count, their removal does not represent any loss of content from the thesis.

AVIS:

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque et Archives Canada de reproduire, publier, archiver, sauvegarder, conserver, transmettre au public par télécommunication ou par l’Internet, prêter, distribuer et vendre des thèses partout dans le monde, à des fins commerciales ou autres, sur support microforme, papier, électronique et/ou autres formats.

L’auteur conserve la propriété du droit d’auteur et des droits moraux qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

Conformément à la loi canadienne sur la protection de la vie privée, quelques formulaires secondaires ont été enlevés de cette thèse.

Bien que ces formulaires aient inclus dans la pagination, il n’y aura aucun contenu manquant.
ABSTRACT

Characterization of the Function of the AtST4 Subfamily Members in Cytokinin-Dependent Growth Control in Arabidopsis thaliana

Effat Khodashenas

The main objective of our laboratory is to characterize the function of the 18 sulfotransferase-coding genes of Arabidopsis thaliana. In this study, we describe the biochemical and biological characterization of the three members of the AtST4 subfamily (AtST4a, b and c). The analysis of published microarray data as well as transcript expression studies show that the three members of the AtST4 subfamily are expressed in roots and regulated by cytokinins. AtST4b is among the group of genes exhibiting the highest level of induction following treatment with the cytokinin trans-zeatin. In contrast, AtST4c is repressed under the same experimental conditions.

To elucidate their biological function, we isolated AtST4a, AtST4b, and AtST4c loss of function mutants. Using metabolite profiling of the knockout mutant and mass spectrometry, we demonstrate that AtST4b encodes a cadabicine (cyclic dihydroxycinnamoyl spermidine) sulfotransferase. Even though cadabicine has previously been reported to occur naturally in plants, this is the first report of its occurrence in A. thaliana and the first report of the accumulation of its sulfonated conjugate. Phenotypic analysis of the AtST4b-knockout mutant showed alternations in root, shoot and reproductive development as compared to wild type plants and a partial loss of sensitivity to cytokinins. Our results seem to indicate that AtST4b plays an important role in the cytokinin-mediated effects on growth.
Unfortunately, we could not find the endogenous substrate of AtST4a and AtST4c, and the results of our biochemical and metabolic profiling experiments suggest that they might sulfonate the same substrate or substrates with very similar properties. However, the AtST4a and AtST4c loss of function mutants show differences in their growth behavior suggesting that they have different functionalities in A. thaliana.
# TABLE OF CONTENTS

**LIST OF FIGURES** ......................................................................................... IX

**LIST OF ABBREVIATIONS** ........................................................................ XI

**CHAPTER 1- REVIEW OF LITERATURE** ...................................................... 1

1.1) Introduction ............................................................................................. 1

1.2) Sulfotransferases ................................................................................... 1
   1.2.1) Introduction .................................................................................... 1
   1.2.2) Sulfotransferase structure .............................................................. 4
   1.2.3) *Arabidopsis thaliana* sulfotransferases ......................................... 5
   1.2.4) The *AtST4* subfamily ................................................................. 6

1.3) Cytokinins .............................................................................................. 7
   1.3.1) Introduction .................................................................................... 7
   1.3.2) Cytokinin perception and signal transduction ................................ 9
   1.3.3) Two-component system ............................................................... 9
   1.3.4) Physiological responses to cytokinins ........................................... 17

1.4) Polyamines ............................................................................................. 22

**CHAPTER 2- MATERIALS AND METHODS** ............................................. 27

2.1) Materials ............................................................................................... 27

2.2) Methods ............................................................................................... 27
   2.2.1) Plant growth conditions ............................................................... 27
   2.2.2) Seed sterilization .......................................................................... 27

2.3) Regulation studies of the *AtST4* subfamily ........................................ 28
   2.3.1) Regulation study of the *AtST4* subfamily in response to cytokinins 28

2.4) Molecular characterization of T-DNA insertion mutants .................. 29
   2.4.1) Genomic DNA extraction ............................................................ 29
   2.4.2) Screening for T-DNA insertion mutants ........................................ 29
   2.4.3) RT-PCR analysis of T-DNA insertion mutants ............................ 31

2.5) Phenotype analysis of mutant plants .................................................. 31
| 2.5.1) Root growth analysis | ................................................................. 31 |
| 2.5.2) Determination of rosette diameter, seed size, seed and silique number | ................................................................. 32 |
| 2.5.3) Statistical analysis | ................................................................. 32 |

| 2.6) Enzymology | ................................................................. 33 |
| 2.6.1) Expression and purification of recombinant enzymes in *Escherichia coli* | ................................................................. 33 |
| 2.6.2) Sulfotransferase assays | ................................................................. 33 |
| 2.6.3) Preparation of *Arabidopsis thaliana* extracts for the detection of the endogenous substrate | ................................................................. 34 |
| 2.6.4) Preparation of *Arabidopsis thaliana* extracts for the detection of the endogenous enzymatic products | ................................................................. 35 |
| 2.6.5) Mass spectrometry | ................................................................. 35 |
| 2.6.6) Thin Layer Chromatography | ................................................................. 36 |

### CHAPTER 3 - RESULTS

| 3.1) Regulation study of the *AtST4* subfamily | ................................................................. 37 |
| 3.1.1) Introduction | ................................................................. 37 |
| 3.1.2) Transcript expression study of the *AtST4* subfamily in response to cytokinins | ................................................................. 38 |
| 3.1.3) Study of the presence of cytokinin-responsive elements in the upstream region of the *AtST4* genes | ................................................................. 39 |

| 3.2) Molecular characterization of T-DNA insertion mutants | ................................................................. 43 |
| 3.2.1) Introduction | ................................................................. 43 |
| 3.2.2) Isolation of T-DNA insertion homozygous mutants | ................................................................. 43 |
| 3.2.3) Transcript expression analysis of the T-DNA insertion homozygous mutants | ................................................................. 44 |

| 3.3) Biochemical characterization of *AtST4b* | ................................................................. 46 |
| 3.3.1) Introduction | ................................................................. 46 |
| 3.3.2) Expression of *AtST4b* recombinant sulfotransferase | ................................................................. 46 |
| 3.3.3) HPLC purification of *AtST4b* substrate and product | ................................................................. 47 |
| 3.3.4) Neutral loss mass spectrometry of *AtST4b* purified product | ................................................................. 49 |
| 3.3.5) LC-MS/MS analysis of *AtST4b* purified product | ................................................................. 50 |

| 3.4) Characterization of *AtST4b* biological function | ................................................................. 54 |
| 3.4.1) Introduction | ................................................................. 54 |
| 3.4.2) Primary root length analysis | ................................................................. 54 |
| 3.4.3) Number of lateral roots | ................................................................. 55 |
| 3.4.4) Rosette diameter | ................................................................. 58 |
| 3.4.5) Number of leaves | ................................................................. 59 |
| 3.4.6) Flowering time | ................................................................. 61 |
| 3.4.7) Seed production | ................................................................. 61 |
| 3.4.8) Seed size and volume | ................................................................. 63 |

| 3.5) Biochemical characterization of *AtST4a* and *AtST4c* | ................................................................. 65 |
| 3.5.1) Introduction | ................................................................. 65 |
| 3.5.2) Expression of *AtST4a* and *AtST4c* recombinant sulfotransferases | ................................................................. 66 |
| 3.5.3) HPLC purification of the substrate and product | ................................................................. 67 |
| 3.5.4) Neutral loss mass spectrometry of *AtST4a* and *AtST4c* reaction products | ................................................................. 68 |
3.5.5) Thin layer chromatography of AtST4a and AtST4c reaction products ............................................. 69

3.6) Characterization of AtST4a and AtST4c biological function .............................................................. 70
   3.6.1) Introduction ......................................................................................................................................... 70
   3.6.2) Primary root length analysis ............................................................................................................... 70
   3.6.3) Number of lateral roots ....................................................................................................................... 71
   3.6.4) Rosette diameter ................................................................................................................................. 73
   3.6.5) Number of leaves ................................................................................................................................. 74
   3.6.6) Flowering time ..................................................................................................................................... 75
   3.6.7) Seed production ................................................................................................................................. 77
   3.6.8) Seed size and volume of the AtST4c-KO plants .................................................................................. 79

CHAPTER 4- DISCUSSION AND FUTURE WORK ............................................................................................ 82

4.1) Biochemical and biological characterization of AtST4b........................................................................... 82
4.2) Biochemical and biological characterization of AtST4a and AtST4c ......................................................... 88

REFERENCES: .................................................................................................................................................. 92

ANNEX 1 ......................................................................................................................................................... 103
   Table 1. Statistical data of root length phenotype analysis ............................................................................. 103
   Table 2. Statistical data of hormone treatment analysis of root tissue ......................................................... 105
   Table 3. Statistical data of number of lateral root phenotype analysis ......................................................... 105

ANNEX 2 ......................................................................................................................................................... 107
   Table 1. Statistical data of rosette diameter analysis ..................................................................................... 107
   Table 2. Statistical data of number of leaves phenotype analysis ................................................................. 107

ANNEX 3 ......................................................................................................................................................... 109
   Table 1. Statistical data of number of seeds per silique analysis ................................................................. 109
   Table 2. Statistical data of number of siliques analysis ............................................................................... 109
   Table 3. Statistical data of seed size phenotype analysis ............................................................................. 110
LIST OF FIGURES

Figure 1. General sulfonation reaction catalyzed by SULTs .................................................. 2
Figure 2. Phylogenic tree of A. thaliana SULTs ........................................................................ 6
Figure 3. Structure of naturally occurring cytokinins ................................................................. 8
Figure 4. Schematics of two-component system in bacteria and plants ..................................... 11
Figure 5. A model for cytokinin signal transduction through his-to-asp phosphorelay ........ 14
Figure 6. General expression pattern of Arabidopsis cytokinin two-component signaling elements .......................................................................................................................... 17
Figure 7. Polyamine biosynthesis pathway in plants ................................................................. 25
Figure 8. Transcript expression profile of the members of AtST4 subfamily in response to cytokinins .................................................................................................................................. 38
Figure 9. Identification of T-DNA insertion homozygous lines of the AtST4 genes ............. 45
Figure 10. SDS-PAGE of partially purified AtST4b recombinant sulfotransferase ............... 47
Figure 11. HPLC profile of an acid hydrolyzed root extract ..................................................... 48
Figure 12. HPLC purification of the AtST4b radiolabeled enzymatic product ...................... 49
Figure 13. Mass spectrometry analysis of AtST4b-reaction product .......................................... 50
Figure 14. Identification of AtST4b-sulfated product ................................................................. 53
Figure 15. Root phenotype of the AtST4b-KO mutants ............................................................. 57
Figure 16. Shoot phenotype of the AtST4b-KO plants ............................................................... 60
Figure 17. Reproductive development is improved in the AtST4b-KO line ............................. 63
Figure 18. Seed size of the AtST4b-KO plants ........................................................................ 65
Figure 19. SDS-PAGE of AtST4a and AtST4c recombinant sulfotransferases .................... 66
Figure 20. HPLC purification of AtST4a and AtST4c endogenous substrate ......................... 67
Figure 21. HPLC purification of AtST4a and AtST4c radiolabeled product. 68
Figure 22. Thin layer chromatography of AtST4a and AtST4c sulfated products 70
Figure 23. Root development of AtST4a-KO and AtST4c-KO 73
Figure 24. Shoot phenotype of AtST4a-KO and AtST4c-KO. 77
Figure 25. Reproductive developments of AtST4a- and AtST4c-KO 79
Figure 26. Seed size of AtST4c-KO plants 81
LIST OF ABBREVIATIONS

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>AHKs</td>
<td>Arabidopsis histidine kinase receptors</td>
</tr>
<tr>
<td>AHPs</td>
<td>Arabidopsis histidine phosphotransmitter proteins</td>
</tr>
<tr>
<td>ARR</td>
<td>Arabidopsis response regulator</td>
</tr>
<tr>
<td>BAHD acyl transferase</td>
<td>First letters of the first characterized members of the family: BEAT, AHCT, HCBT and DAT</td>
</tr>
<tr>
<td>BRs</td>
<td>Brassinosteroids</td>
</tr>
<tr>
<td>CID</td>
<td>Collision induced dissociation</td>
</tr>
<tr>
<td>CKs</td>
<td>Cytokinins</td>
</tr>
<tr>
<td>CKXs</td>
<td>Cytokinin oxidases</td>
</tr>
<tr>
<td>Col-0</td>
<td>Columbia 0</td>
</tr>
<tr>
<td>C-terminus</td>
<td>Carboxyl-terminus</td>
</tr>
<tr>
<td>DAG</td>
<td>Days after germination</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethylsulfoxide</td>
</tr>
<tr>
<td>DPM</td>
<td>Disintegrations per minute</td>
</tr>
<tr>
<td>E.coli</td>
<td><em>Escherichia coli</em></td>
</tr>
<tr>
<td>eFP</td>
<td>Electronic fluorescent pictograph</td>
</tr>
<tr>
<td>FT-IMS</td>
<td>Fourier transform-ion mobility spectrometry</td>
</tr>
<tr>
<td>GARP motif</td>
<td>First letters of the members in maize GOLDEN2, Arabidopsis response regulators and <em>Chlamidomonas</em> PSR1</td>
</tr>
<tr>
<td>HCAAs</td>
<td>Hydroxycinnamic acid amides</td>
</tr>
<tr>
<td>His-tag</td>
<td>Histidine-tag</td>
</tr>
<tr>
<td>HPLC</td>
<td>Reverse phase high performance liquid chromatography</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>IPT</td>
<td>Isopentenyl dephosphate transferase</td>
</tr>
<tr>
<td>Kbp</td>
<td>Kilobase pair</td>
</tr>
<tr>
<td>Kda</td>
<td>Kilodalton</td>
</tr>
<tr>
<td>KO</td>
<td>Knockout</td>
</tr>
<tr>
<td>LB</td>
<td>T-DNA left border</td>
</tr>
<tr>
<td>LC-MS/MS</td>
<td>Reverse phase liquid chromatography-tandem mass spectrometry</td>
</tr>
<tr>
<td>LFY</td>
<td>LEAFY gene</td>
</tr>
<tr>
<td>MS media</td>
<td>Murashige and Skoog media</td>
</tr>
<tr>
<td>Ni-NTA</td>
<td>Nickel-nitrilotriacetic acid</td>
</tr>
<tr>
<td>N-terminus</td>
<td>Amino-terminus</td>
</tr>
<tr>
<td>PAPS</td>
<td>3'-phosphoadenosine 5'-phosphosulfate</td>
</tr>
<tr>
<td>PAs</td>
<td>Polyamines</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>Q-TOF mass spectrometer</td>
<td>Tandem quadrupole time-of-flight mass spectrometer</td>
</tr>
<tr>
<td>RT-PCR</td>
<td>Reverse transcription-polymerase chain reaction</td>
</tr>
<tr>
<td>SCT</td>
<td>Spermidine dicoumaroyl transferase</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate-poly acrylamide gel electrophoresis</td>
</tr>
<tr>
<td>SSAT</td>
<td>Spermidine/spermine N1-acetyltransferase</td>
</tr>
<tr>
<td>SULTs</td>
<td>Sulfotransferases</td>
</tr>
<tr>
<td>TCS</td>
<td>Two-component signaling system</td>
</tr>
<tr>
<td>T-DNA</td>
<td>Transfer DNA</td>
</tr>
</tbody>
</table>
TLC  Thin layer chromatography

\( t\)-zeatin  \( trans\)-zeatin
Chapter 1- Review of Literature

1.1) Introduction

This chapter reviews the main topics related to the characterization of the three members of the *AtST4* sulfotransferase (SULT) subfamily from *Arabidopsis thaliana*. At first, we present a brief introduction to SULTs, their structure and functions, with emphasis on plant soluble SULTs. Next, we introduce cytokinins, a class of phytohormones that regulate many physiological processes including the expression of the *AtST4* gene subfamily in *A. thaliana*.

Since the substrate of *AtST4b* is a polyamine conjugate, we conclude this chapter by reviewing the biological significance of the accumulation of polyamines in plants.

1.2) Sulfotransferases

1.2.1) Introduction

SULTs catalyze the transfer of a sulfuryl group (SO$_3$) from the ubiquitous donor 3’-phosphoadenosine 5’-phosphosulfate (PAPS), which is the active form of inorganic sulfate, to an appropriate hydroxyl group of different substrates in a process called the sulfonation reaction (Fig. 1). These enzymes have highly conserved domains and are found across all kingdoms, from bacteria to plantae and animalia (Varin *et al.* 1997). The sulfonation reaction seems to be very important for life, because a defect in sulfate metabolism or sulfate transport lead to severe skeletal disorders and/or death in human and mouse (ul Haque *et al.* 1998; Rossi and Superti-Furga 2001).
Figure 1. General sulfonation reaction catalyzed by SULTs (Chapman et al. 2004).

There are two classes of SULT proteins based on their cellular localization and function: membrane-associated SULTs and cytosolic SULTs. The membrane-associated SULTs are localized in the Golgi apparatus and catalyze the sulfonation of macromolecules such as proteins, peptides and complex carbohydrates. A large number of these enzymes catalyze the sulfonation of biological signaling molecules which are essential for life. Members of this class of enzymes have been characterized only from Mimosa pudica and Oryza sativa in plants, and are not the interest of this study. The so-called cytosolic SULTs or soluble SULTs sulfonate small organic molecules such as flavonoids, steroids, glucosinolates and hydroxyjasmonates (Honke and Taniguchi 2002; Hernández-Sebastiá et al. 2008).

The sulfonation reaction has different effects on a metabolite. The sulfated conjugate is more polar and water soluble than the original molecule, therefore the presence of a sulfate group might influence their transport and excretion. For example, mammalian
cytosolic SULTs play an important role in phase II of the biotransformation and excretion of xenobiotics in the liver (Weinshilboum and Otterness 1994; Yasuda et al. 2005). Sulfonation also modulates the biological activity of some metabolites such as steroids. For example, the sulfonation and desulfonation reactions are controlling the level of active estrogen in the blood of mammals. The sulfated estrogens are stored and eventually can be converted to the active hormone by sulfatases (Strott 1996). In contrast, the sulfonation can increase the biological activity of some metabolites. For example, the sulfonation of tyrosine residues on the chemokine receptor CCR5 (a principal HIV-1 co-receptor) is a modification that is required for its biological activity and indirectly facilitates entry of HIV-1 (Farzan et al. 1999).

In plants, SULTs are involved in the regulation of plant growth, defense and stress responses by the modulation of the biological activity of signal molecules and hormones. For example, it has been shown that the sulfonation of desulfoglucosinolates is required for the activation of their antimicrobial properties (Piotrowski et al. 2004; Klein et al. 2006). Moreover, the choline sulfate that accumulates under saline conditions is involved in salt or drought stress-tolerance in Limonium species and other plants of the Plumbaginaceae family (Klein and Papenbrock 2004). So far, the characterized plant cytosolic sulfotransferases are as follows:

The flavonol SULTs from Flaveria species (Varin et al. 1992; Varin et al. 1997), the choline SULT from Limonium sativum (Rivoal and Hanson 1994) and the brassinosteroid SULTs from Brassica napus (Rouleau et al. 1999; Marsolais et al. 2004). The cytosolic SULTs from A. thaliana are discussed in detail later in the text.
1.2.2) Sulfotransferase structure

The X-ray crystal structures of four mammalian cytosolic SULTs and one membrane-associated SULT show a globular structure composed of a single α/β domain with five-stranded parallel β sheet surrounded by α-helices (Negish et al. 2001). The amino acid sequence alignment of plant and animal cytosolic SULTs revealed four conserved regions named region I to region IV. Using various approaches, the most important structural and functional features of SULTs have been identified:

- **The PAPS-binding region:** All SULTs contain conserved domains involved in PAPS binding (region I and region IV). The consensus amino acid sequence of the PAPS-binding regions of plant SULTs are PKxGGTTLKAL for region I and FRKGxVGDW for region IV. The glycine-rich domain followed by the conserved lysine of region IV is a motif (GXXGXXK) that is found to be essential in some nucleotide binding proteins (Weinshilboum et al. 1997; Klein and Papenbrock 2004; Hernández-Sebastiá et al. 2008). The lysine residue from region I is required for the formation of the 5'-phosphosulfate loop (PSB loop). Also, the three amino acids RKG from region IV are involved in the formation of the 3'-phosphate binding loop (PB loop).

- **The catalytic binding region:** The crystal structure and site-directed mutagenesis studies have revealed the importance of conserved histidine, serine and arginine residues from region II and of a lysine residue from region I in the formation of an unstable ternary enzyme-PAPS-substrate complex and for the transfer of the sulfonate group to the substrate (Hernández-Sebastiá et al. 2008).
• **The substrate binding region:** The study of human and Flaveria flavonol SULTs have indicated two domains, in close proximity to region II, with high amino acid divergence. These domains are responsible for the substrate specificity of SULTs (Hernández-Sebastiá et al. 2008).

### 1.2.3) *Arabidopsis thaliana* Sulfotransferases

*A. thaliana* is a small flowering plant that is widely used as a model organism in plant biology. The genome of *A. thaliana* was completely sequenced in 2001 and provided useful information for biochemical and genetics studies of its over 26,500 genes. There are 18 SULT coding genes in *A. thaliana* based on sequence similarity to previously characterized SULTs. Of these, one (At3g51210) is a pseudogene that codes for a truncated protein, and seven have been fully characterized: Flavonoid SULT (At3g45070) (Gidda and Varin 2006), desulfoglucosinolate SULTs (At1g74100, At1g74090, At1g18590) (Piotrowski *et al.* 2004), hydroxyjasmonate SULTs (At5g07010) (Gidda *et al.* 2003) and brassinosteroid SULTs (At2g03760 and At2g14920) (Marsolais *et al.* 2004). The phylogenetic tree of *A. thaliana* SULTs is shown in Figure 2.
1.2.4) The AtST4 subfamily

The AtST4 subfamily is part of a family of plant SULTs designated as SULT203 (formerly SULT5), based on the recently proposed guidelines for sulfotransferase nomenclature (Blanchard et al. 2004). The AtST4 subfamily contains three members: AtST4a (At2g14920), AtST4b (At1g13420) and AtST4c (At1g13430).

Molecular studies of the AtST4 subfamily have shown that the three genes are expressed mainly in roots and are regulated by cytokinins (Marsolais et al. 2007). The AtST4a gene or SUL203A1 is located on chromosome 2 and is more related to AtST4c (80% amino acid sequence identity) than AtST4b (71% amino acid sequence identity). AtST4a has been characterized recently and was shown to code for a brassinosteroid SULT in in vitro studies (Marsolais et al. 2007). However, the presence of brassinosteroid sulfate has never been demonstrated in vivo. AtST4b and AtST4c are in tandem on chromosome 1.

Figure 2. Phylogenic tree of A. thaliana SULTs. The respective amino acids sequences of 18 SULTs were grouped using the Clustal W program (http://www.ebi.ac.uk/Tools/clustalw2). The colored box indicates the AtST4 subfamily and its members.
Despite their high amino acid sequence identity with AtST4a, neither AtST4b nor AtST4c exhibited any activity with brassinosteroids (Marsolais et al. 2007).

1.3) Cytokinins

1.3.1) Introduction

Cytokinins (CKs) are a class of phytohormones that play critical roles at different stages of growth and development. The name cytokinin refers to the ability of these molecules to promote cell division and cytokinesis. Miller et al. identified the first CK, kinetin, from the autoclaved product of herring sperm DNA as a metabolite that had the ability to promote cell division (Miller et al. 1995).

CKs are adenine derivatives and classified based on the substitution on their N\textsuperscript{6}-side-chain as aromatic or isoprenoid CKs. Both aromatic and isoprenoid CKs are naturally occurring, but the latter is by far the most abundant in plants. The common isoprenoid CKs are; trans-hydroxylated N\textsuperscript{6}-side chain or trans-zeatin (t-zeatin), N\textsuperscript{6}-(\Delta\textsuperscript{2}-isopentenyl) adenine (iP), dihydrozeatin (DZ) and cis-zeatin (c-zeatin) (Fig. 3A).

The relative abundance of different CKs depends on the plant species, tissue and developmental stages. For example t-zeatin is the most abundant CK in A. thaliana, whereas c-zeatin and its derivatives are the predominant CKs in chickpea seeds and male flower buds of Mercurialis (Mok and Mok 2001; Sakakibara 2006).

The aromatic CKs that occur naturally in plants are benzyladenine (BA), ortho-topolin and meta-topolin. They are not widely distributed and are found in a limited number of plant species (Fig. 3B). All natural CKs may also be present in the corresponding nucleoside, nucleotide and glycoside counterparts in the plant metabolome. There are
also some synthetic metabolites that possess cytokinin activity such as diphenylurea (DPU) but have not been found naturally in plants.

Figure 3. Structure of naturally occurring cytokinins. (A) Isoprenoid CKs, (B) Aromatic CKs (Sakakibara 2006).
1.3.2) Cytokinin perception and signal transduction

Since their discovery 50 years ago, scientists have tried to understand the mechanism by which CKs are sensed and perceived in the cell. Early after their discovery, scientists found that the sites of biosynthesis and action of CKs are spatially separated. Isolation of cytokinin-binding-proteins (CBPs) as CK receptors was not successful since none were shown to be a true receptor. With the advent of molecular biology, the mechanism of CK signaling became well understood. The discovery of histidine kinase receptors and downstream components indicated that plants use the two-component signaling system (TCS), initially discovered in bacteria, in the CK signal transduction pathway.

In the next section, the mode of action of TCS in prokaryotes and higher eukaryotes like plants is discussed, and then the components of this system in CK signaling in A. thaliana are explained.

1.3.3) Two-component system

Plants like bacteria utilize the two-component system (TCS) to sense and respond to their environment. Basically, TCS consists of two main components: the sensor kinase and the cognate response regulator (RR). The sensor histidine kinase is usually a membrane-bound protein that senses environmental stimuli. It consists of two domains: the “input domain” and the “transmitter domain”. The input domain dimerizes upon binding of ligands and autophosphorylates the transmitter domain at a histidine residue (H) (Fig. 4A). This phosphate group is then transferred to the RR. Thus, the sensor histidine kinase directly modulates the activity of RRs in response to stimuli by phosphorylation (Chang and Stewart 1998).
The RR, which propagates the signal through the cytoplasm, consists of two domains: the “receiver domain” that receives the phosphate group on its conserved aspartate (D) residue, and the “output” domain that is not present in all response regulators and has DNA binding activity. Phosphorylation of the receiver domain results in a conformational change in the output domain that will eventually lead to transcription of special sets of genes (Fig. 4A).

The TCS is very well studied in prokaryotes. There are more than 40 different TCSs that have been identified in Escherichia coli. Compared to bacteria, plants use a relatively complicated system. They use a hybrid sensor kinase and a multistep phosphorelay cascade in which the phosphoryl group alternates between sequential histidine and aspartate residues of different substrates (Fig. 4B). This allows for longer-lasting signaling in cells in which different pathways can be involved in response to extracellular signals. The output of such complex responses probably leads to a better adaptation to environmental changes.

The components of such a pathway in the model plant A. thaliana comprise the Histidine Kinase receptors (AHKs), the Histidine Phosphotransmitter proteins (AHPs) and the Response Regulators (ARRs).
Figure 4. Schematics of two-component system in bacteria and plants (D'Agostino and Kieber 1999). (A) Basic prokaryotic two-component system. The input domain (red) is responsible for perceiving external stimuli and modulates the histidine kinase activity of the attached transmitter domain (green). Active sensor kinases act as dimmers. For the sake of simplicity, only a monomer of the sensor kinase is depicted. The phosphate is then transferred to a conserved aspartate residue (D) in the receiver domain (blue) of a cognate response regulator, which activates or inactivates the output domain (yellow). (B) The phosphorelay system in plants. The input domain of a hybrid kinase, regulates the activity of the transmitter domain by phosphorylation of its histidine residue. The phosphate is then transferred to an aspartate residue on the fused receiver domain and then to a histidine on an AHP protein (purple) and, finally, to an aspartate residue on an ARR protein. Abbreviations: AHP, Arabidopsis histidine phosphotransfer (HPt) proteins; ARR, Arabidopsis response regulator.

Arabidopsis Histidine Kinase Receptors (AHKs)

Kakimoto identified an Arabidopsis gene, CKII that produces a cytokinin-independent phenotype in calli when overexpressed by transfer DNA (T-DNA) activation tagging. The CKII protein contains a sensor histidine kinase domain, a receiver domain, and two potential transmembrane domains in its structure. These characteristics together with the
phenotype of the mutant suggested that CKI1 is a cytokinin receptor (Kakimoto 1996). However, the binding of this protein to CKs has not been demonstrated yet.

The CRE1/ WOL1/ AHK4 gene was later identified independently in different research studies as a cytokinin receptor coding gene from mutants that exhibited reduced sensitivity to cytokinins. Using forward genetics, Inoue et al. found that mutations in the CRE1 (Cytokinin Response 1) gene produce a cytokinin-insensitive phenotype in Arabidopsis. The crel mutants were complemented by introduction of a functional CRE1 gene. Moreover, expression of CRE1 conferred a cytokinin-dependent phenotype when introduced in a yeast strain deficient in the endogenous SLN1 histidine kinase gene (Inoue et al. 2001). Molecular and genetic evidences revealed that CRE1 is allelic to WOL1 (WOODEN LEG 1) and to AHK4 genes. The WOL1 gene was initially found to be required for proper formation of root vascular tissue (Mahonen et al. 2000).

Almost at the same time, Ueguchi et al. found the homologs of CRE1 as a novel family of sensor histidine kinase genes based on amino acid sequence similarity to the transmitter domain of several sensor histidine kinases in A. thaliana. The members of the so-called AHK family (AHK2, AHK3 and AHK4) were over 60% identical in the amino acid sequence of their transmitter and receiver domains, and they all encoded plasma membrane associated proteins. Among them, AHK3 and AHK4 were shown to bind directly to cytokinins using a yeast suppression assay (Ueguchi et al. 2001; Ueguchi et al. 2001). Binding of AHK4 to different types of cytokinins was further demonstrated using heterologous yeast and E. coli phosphohorelay systems (Suzuki et al. 2001; Yamada et al. 2001). Using protoplast transient expression analysis in the presence of cytokinins, all
three putative receptors were able to induce the expression of the ARR6 promoter which is a cytokinin primary response gene (Hwang and Sheen 2001).

Furthermore, study of AHKs expression patterns and AHKs single, double and triple mutants revealed distinct, yet overlapping functions of the three receptors in regulating root and shoot growth (Higuchi et al. 2004; Nishimura et al. 2004; Riefler et al. 2006).

AHKI and AHK5 (CKI2) are also coding for transmembrane histidine kinases in Arabidopsis. However, their potential role in cytokinin signaling has not been demonstrated yet (Kakimoto 1996; Urao et al. 1999).

**Arabidopsis Histidine Phosphotransmitter Proteins (AHPs)**

Identification of sensor histidine kinases as cytokinin receptors suggested the presence of other components of the His-to-Asp phosphorelay system in *A. thaliana* including Histidine Phosphotransmitter proteins (AHPs) and Response Regulators (ARRs). There are at least five putative histidine phosphotransmitter coding genes in the *A. thaliana* genome (Miyata et al. 1998; Suzuki et al. 2000). AHPs act as cytoplasmic-nuclear shuttles and transfer the signals from AHK, which are mainly localized in the plasma membrane, to ARRs, which are mostly localized in the nucleus (Fig. 5). It was shown that some AHPs (AHP1 and AHP2) transiently translocate from cytoplasm to nucleus upon induction by cytokinins (Hwang and Sheen 2001). Additionally, the binding of AHPs to AHKs and type-B ARRs have been demonstrated using the *E. coli* phosphorelay system and the yeast two-hybrid system, respectively (Suzuki et al. 2001; Suzuki et al. 2001). Analysis of AHPs loss-of-function mutations indicates the positive and redundant function of these elements in cytokinin signaling (Hutchison et al. 2006).
Figure 5. A model for cytokinin signal transduction through His-to-Asp phosphorelay (Heyl and Schmulling 2003). Ligand binding induces receptor dimerization and autophosphorylation. Transfer of the phosphoryl group by activated receptors activates AHPs which transport the signal from the cytoplasm to type-B ARRs in the nucleus. Type-B response regulators transcribe target genes, among them type-A ARR genes. Type-A response regulators may down-regulate the primary cytokinin signal response via a negative feedback loop. Abbreviations: D, aspartate residue, H, histidine residue, P, phosphoryl group.

Arabidopsis Response Regulators (ARRs)

The Response Regulators (ARRs) are the final components of the phosphorelay circuitry in plants. There are at least 22 ARR-coding genes in the Arabidopsis genome that contain
invariant DDK residues in their receiver domain, a hallmark that is also present in
response regulators of prokaryotes and yeasts. Based on amino acid sequence similarity
and protein structure, ARRs fall into two distinct classes: type-A and type-B. Most of
type-A ARRs lack the output domain at their C-terminal ends and show rapid induction
upon treatment with cytokinins. In contrast, type-B ARRs contain an extended C-terminal
domain having the characteristics of transcriptional activators and their expression is not
influenced by cytokinins.

**Type-A ARRs: (ARR3-ARR9, ARR15-ARR17 and ARR22):** The 11 members of type-A
ARRs are mainly composed of a receiver domain with a short extension at their C-
terminus (less than 100 amino acids). Their receiver domains are more closely related to
each other (60-93% identical in amino acid sequence) than the receiver domain of type-A
and type-B ARRs (which are less than 30% identical in amino acid sequence). Type-A
ARRs are known as “primary response genes” since their transcript level is rapidly and
transiently induced by cytokinins (within 10 minutes). This induction is insensitive to
protein synthesis inhibitors (such as cycloheximide) and is specific for cytokinins
(D'Agostino et al. 2000). Type-A ARRs are located either in the cytoplasm or the nucleus
and their transcripts are present in all parts of the plant (Heyl and Schmulling 2003).
Most of the type-A ARRs are negative regulators of the cytokinin signaling pathway
(Hwang and Sheen 2001; To et al. 2004; Lee et al. 2007).

**Type-B ARRs: (ARR1, ARR2, ARR10-ARR14, and ARR18-ARR21):** Type-B ARRs
have a receiver domain and a large carboxy-terminal output domain (250-260 amino
acids). The presence of a nuclear localization signal domain (NLS), a GARP DNA-
binding domain (or B-motif) and a proline/glutamine-rich domain in the C-terminal
region of all type-B ARRs indicates that they act as transcriptional activators (D'Agostino and Kieber 1999). Recently, a core DNA sequence motif (G/A)GAT(T/C) was identified to be bound by type-B ARRs and be present in the upstream region of most of type-A ARR promoters and putative cytokinin-induced genes (Sakai et al. 2000; Rashotte et al. 2003).

Studies of several type-B ARR mutant lines suggest distinct but also redundant functions among type-B ARRs (Imamura et al. 2003; Tajima et al. 2004). The exogenous application of cytokinins does not alter the transcription of type-B ARRs (Imamura et al. 1999; Kiba et al. 1999).

The expression pattern of different components of the TCS in A. thaliana is illustrated in Fig. 6.
1.3.4) Physiological responses to cytokinins

Cytokinins are involved in many plant growth and developmental processes. In the following section, we first discuss the main downstream target genes which are regulated through the cytokinin signaling pathway. Then, we review the available microarray data showing the cytokinin regulation of the members of AtST4 subfamily in A. thaliana.

Cytokinins and cell division

The cytokinin role in cell division and proliferation was first documented in the 1950s (Miller et al. 1995). In 1999, CycD3 (a member of a D-type cyclin gene family involved
in the G1 to S transition of the cell cycle) was found to be up-regulated in Arabidopsis mutants with high levels of endogenous cytokinins and to be rapidly induced by the exogenous application of cytokinins (Riou-Khamlichi *et al.* 1999). This induction was cycloheximide independent. Moreover, transgenic cells overexpressing the CycD3 gene were cytokinin independent when grown *in vitro*. Altogether, these results suggest that cytokinins regulate cell cycle progression at the G1-S transition through the positive regulation of CycD3 expression.

**Cytokinins and shoot development**

The exogenous application of cytokinins induces shoot formation from calli grown *in vitro*. The analysis of Arabidopsis loss-of-function and gain-of-function mutants with altered cytokinin levels has confirmed the positive role of cytokinins on shoot development. Reduced cytokinin content in the cytokinin-receptor triple mutant (*ahk2,3,4*) and cytokinin-oxidase overexpressor lines (35s::CKXs) resulted in reduced size of shoot apical meristem (SAM), retarded leaf formation and reduced cell production in leaves (Werner *et al.* 2001; Werner *et al.* 2003; Higuchi *et al.* 2004; Nishimura *et al.* 2004). Furthermore, increased levels of cytokinins in transgenic plants overexpressing the bacterial cytokinin biosynthesis gene, isopentyl transferase (*IPT*), led to ectopic shoot formation, reduced apical dominance and increased mesophyll cell layers in leaves. The gene expression studies showed induction of *KNOTTED1* and *STM* homeobox genes in these plants, as well. The *KNOTTED1*-like homeobox genes are expressed exclusively in the SAM and are involved in its development and maintenance (Rupp *et al.* 1999). Moreover, the defect of a recently identified cytokinin biosynthesis enzyme in rice resulted in pre-mature termination of shoot meristem (Kurakawa *et al.* 2007).
Cytokinins and senescence

Leaf senescence is defined as an aging process that is accompanied with the degradation of chlorophyll and photosynthesis proteins. It is a complex mechanism that occurs at final stages of leaf growth and is controlled by many factors. It was shown that cytokinins delay leaf etiolation and senescence (Wingler et al. 1998). Overexpression of the cytokinin biosynthesis gene, *IPT*, under the control of a senescence-specific gene promoter delayed senescence in transgenic tobacco (Gan and Amasino 1995). How cytokinins regulate senescence at the molecular level is not known yet, but mutational studies of cytokinin signaling components indicate a primary role for *AHK3* and *ARR2* (a type-B ARR) in this process (To and Kieber 2008).

Cytokinins and root development

In contrast to their promotional role in shoot development, cytokinins reduce root meristem size and inhibit primary root elongation and lateral root formation. Mutants with reduced cytokinin sensitivity (*ahk3* and *arr1,12*) develop an enlarged root meristem and a longer primary root (Dello Ioio et al. 2007). Furthermore, cytokinin-deficient mutants that overexpress cytokinin oxidase exhibited an overall enhanced root system (Werner et al. 2001; Werner et al. 2003). However, reduced cytokinin perception in the receptor triple mutant (*ahk2,3,4*) and in the phosphotransmitter quintuple mutant (*ahp1,2,3,4,5*) resulted in reduced root development. This implies that although cytokinins have a negative regulatory effect in root development, a certain level of cytokinins is needed for proper root formation. Support for this hypothesis comes from molecular and genetic analysis of the *WOODEN LEG* (*WOL*, an allele of *AHK4/CRE1*) mutant. It was shown that the *wol* mutant develops short roots with fewer embryonic vascular tissues
that give rise only to protoxylem tissue. Interestingly, the defect in root vasculature was complemented by exogenous cytokinin application, indicating a cytokinin requirement for root vascular tissue formation and development (Mahonen et al. 2000; Mahonen et al. 2006).

**Cytokinins and the AtST4 subfamily**

As we have seen, cytokinins play an important role in many physiological and developmental processes in plants. Regulation of these processes requires changes in the expression of cytokinin responsive genes. Many genome-wide microarray analyses have been performed in order to discover the genes downstream of the cytokinin signaling pathway. A review of these studies reveals that members of AtST4 subfamily are regulated by cytokinins in A. thaliana.

Study of genome-wide gene expression of transgenic A. thaliana seedlings that carry the bacterial cytokinin-biosynthesis gene (IPT) under the control of a chemically inducible promoter showed up-regulation of AtST4b (At1g13420) transcript 6 hours and 24 hours after IPT induction (Hoth et al. 2003).

In the combinational microarray analysis performed by Kiba et al., the AtST4b transcript was increased up to 7.3-fold in 21-day old wild type Columbia (Col) seedlings treated for 3 hours with the cytokinin t-zeatin. In this study, the AtST4b transcript was also up-regulated 11.3-fold in the absence of cytokinin treatment in a transgenic line overexpressing ARR21 (a type-B ARR) as compared to wild type plants. These results are interesting because based on the kinetics and level of induction, the authors have categorized AtST4b as one of the genes that is up-regulated rapidly and specifically by
cytokinins in a manner similar to that of the type-A ARR genes, which are hallmarks of the two-component phosphorelay system (Kiba et al. 2005).

Microarray analysis performed on roots of two-week old seedlings of Columbia 0 (Col-0) and arr10/arr12 double mutant (type-B ARRs) shows that even though AtST4b is induced in both cytokinin-treated wild type and mutant line, the level of AtST4b induction is attenuated in the mutant line as compared to the wild type plants. These results suggest that ARR10 and ARR11 act as positive regulators of the downstream AtST4b gene (Yokoyama et al. 2007).

Using genome-wide expression profiling, Brenner et al. classified immediate-early (15 min) and delayed (120 min) cytokinin response genes in A. thaliana. They also analyzed gene expression in cytokinin deficient transgenic line (35S:AtCKX1). Surprisingly, AtST4b is not among the genes whose transcripts have been changed by the exogenous application of cytokinins in wild type and the 35S:AtCKX transgenic line. However, the available data indicates the repression of AtST4b transcript in cytokinin oxidase overexpression line compared to wild type plants (Brenner et al. 2005).

As mentioned earlier, type-A ARRs are generally negative regulators in the cytokinin signaling circuitry. In order to investigate the downstream components of type-A ARRs, Lee et al. produced a transgenic line overexpressing ARR7 (type-A ARR). Analysis of microarray data showed that ARR7 overexpression has a distinctively repressive impact on various groups of cytokinin-regulated genes. In particular, the expression of all type-A ARRs (except for ARR22), AHK1 and AHK4, most of the cell expansion genes and the cytokinin oxidase genes were repressed in the ARR7 overexpression line. Surprisingly, regulation of AtST4b gene was positively affected by ARR7. These results showed that
expression of *AtST4b* is not only up-regulated by cytokinins after 30 min and 2 hours treatment in wild type plants, it is also induced in *ARR7* overexpression line by cytokinin treatments. The high levels of induction in the overexpressor line indicate the great impact of *ARR7* on *AtST4b* regulation.

In the Affymetrix full genome array, the same probe set represents both the *AtST4a* and *AtST4c* genes making it impossible to differentiate these two genes. Based on the GeneChip analysis performed by Lee *et al.* (as opposed to *AtST4b*), the expression of *AtST4c* (and/or *AtST4a*) was repressed by the exogenous application of cytokinins in both wild type and *ARR7* overexpressor line (Lee *et al.* 2007).

### 1.4) Polyamines

Polyamines (PAs) are low molecular weight, nitrogen-containing cationic compounds that are found in almost all organisms. Due to their positive charge, PAs can bind directly to RNA, DNA, nucleotide triphosphates and proteins, and in some cases can modulate their function. PAs play important roles in the regulation of gene expression, modulation of certain ion channels, cell proliferation, programmed cell death or apoptosis, ribosome biogenesis, protein synthesis, membrane rigidity and embryonic development (Igarashi and Kashiwagi 2000; Thomas and Thomas 2001; Kusano *et al.* 2008). The major polyamines in plants are putrescine, spermidine and spermine. The biosynthesis of polyamines in plants is initiated by decarboxylation of the amino acid arginine to form the diamine putrescine. Putrescine is then converted to the triamine spermidine and the tetraamine spermine through the sequential addition of aminopropyl residues (Kusano *et al.* 2007) (Fig. 7). Intracellular PA content is modulated through a complex circuitry of biosynthesis, degradation, cellular uptake and efflux. These mechanisms are controlled
by different enzymes and are tightly regulated to maintain PA homeostasis. The study of loss-of-function and gain-of-function mutation of these genes has confirmed the importance of PAs in plant growth and development. Studies by Imai et al. showed that even though spermine is not essential for the survival of Arabidopsis, suppression of the genes coding for the enzymes involved in spermidine biosynthesis (SPDSI and SPDS2) either by T-DNA insertion or RNA interference is lethal in Arabidopsis and results in embryonic arrest in double mutant seeds (Imai et al. 2004; Imai et al. 2004). A double mutation of the two putrescine biosynthesis genes (ADCI and ADC2) leads to a similar developmental defect in the embryo of Arabidopsis (Urano et al. 2005). Moreover, overexpression of oat ADC in tobacco resulted in increasing endogenous putrescine levels and toxic phenotypes such as necrosis and dwarfism (Panicot et al. 2002). Altogether, these data indicate that putrescine and spermidine are essential for plant growth and embryogenesis, while spermine is not essential for normal growth in Arabidopsis. Furthermore, based on the spermidine biosynthesis pathway and requirement of putrescine for spermidine biosynthesis, it is probably the lack of spermidine that leads to embryonic defects in putrescine biosynthesis mutants.

PAs can be conjugated with a variety of compounds by formation of an amide linkage. In animals, acetylation by spermidine/spermine acyltransferase (SSAT) reduces the positive charges of PAs and their biological activity (Pegg 2008). In plants, however, the acylated conjugates of PAs play an important role in growth and development. A group of these polyamine-based alkaloids are referred as hydroxycinnamic acid amides (HCAAs) and formed by the addition of an acyl moiety (hydroxycinnamic acid in this case) to the acceptor molecules (polyamines) in presence of BAHD acyl transferase
(named after the first letters of the first characterized members of the family; BEAT, AHCT, HCBT and DAT) (D'Auria 2006). The hydroxycinnamoyl component can be replaced with coumaroyl, caffeoyl, hydroxyferuloyl or sinapoyl acyl groups. HCAAs are widely distributed in plants and implicated in several growth and developmental processes such as germination, cell division, flower formation, cell wall formation, as well as stress and defense responses (Facchini et al. 2002). For example, Luo et al. have characterized two sinapoyl spermidine derivatives that accumulate largely in seeds of Arabidopsis and seem to act as polyamine reserve during germination (Luo et al. 2009). Such a storage role for polyamine conjugates have previously been reported in rice seeds, as well (Bonneau et al. 1994). Luo et al. have also identified a gene coding for spermidine dicoumaroyl transferase (SCT) that is specifically expressed in the root tip and induced by cytokinins (Luo et al. 2009). Based on its site of expression and induction by cytokinins, the authors have assigned a role in cell division for the SCT in roots. Furthermore, spermidine conjugates in the pollen coat of Arabidopsis are responsible for its shape and autofluorescence (Grienenberger et al. 2009). Interestingly, a DNA-UV protective function have been suggested for cinnamoyl-derivatized spermidines that accumulate in the pollen (Bienz et al. 2005).

In addition to their roles in plants, HCAAs are representing an important class of antioxidant and chemotherapeutic agents that have the potential to be used in the treatment of human diseases and as insecticides (Klose et al. 2002; Park and Schoene 2006; Russo et al. 2007).
1.5) Purpose of the present study

Based on previous genetic studies and microarray analyses, three SULT-coding genes (AtST4a, AtST4b and AtST4c) were found to be expressed exclusively in roots and regulated by cytokinins in *A. thaliana*. *AtST4a* has been previously partially characterized and was shown to encode a brassinosteroid SULT in *in vitro* studies. The purpose of the present study was to characterize the biochemical and biological function of the other members of the *AtST4* subfamily. To further investigate the role of cytokinins on the
regulation of AtST4a, -4b and -4c expression, we analyzed their expression in roots of wild type Arabidopsis at different time points following treatments with t-zeatin.

The best way to study the function of a gene is to study the phenotype of plants that are lacking it. Therefore, homozygous AtST4a, AtST4b and AtST4c T-DNA insertion mutants were isolated and subjected to a number of metabolic and phenotypic analyses. To identify the biochemical function of AtST4a, -4b and -4c in vitro, their coding sequence was expressed in E. coli and their enzyme activity was tested using radioactive PAPS and purified plant metabolites. Using High Performance Liquid Chromatography (HPLC), neutral loss mass spectrometry and liquid chromatography-tandem mass spectrometry (LC-MS/MS) analyses, we were able to characterize the substrate of AtST4b. Finally, the consequence of changing the sulfated metabolome of AtST4a, AtST4b and AtST4c single mutants was analyzed by monitoring a number of growth parameters known to be under the control of cytokinins. The results of these investigations are presented in the next chapters.
Chapter 2- Materials and Methods

2.1) Materials

Seeds of wild type *A. thaliana*, ecotype Columbia 0 (Col-0) were obtained from Lehle seeds (USA). The Arabidopsis lines carrying a T-DNA insertion allele for *AtST4a* (GABI_177E08) and *AtST4b* (GABI_231G06) were obtained from Gabi-Kat (http://www.gabi-kat.de/), and *AtST4c* (FLAG_334F06) from INRA (http://www.inra.fr/). All the mutant lines were generated in a Col-0 background.

2.2) Methods

2.2.1) Plant growth conditions

The Arabidopsis plants were grown either in soil or on vertical petri dishes containing full-strength Murashige and Skoog (MS) medium (1% sucrose, 0.4% Gelrite, 0.05% MES, PH 5.7) under long day conditions (16 hours light/8 hours dark) at a light intensity of \(~130 \text{ \mu mol m}^{-2}\text{s}^{-1}\). The temperature was kept at 20 °C during night-time and gradually increased to 22 °C during day-time.

2.2.2) Seed sterilization

The Arabidopsis seeds were sterilized by a 30 seconds immersion in 70% ethanol, 5 minutes shaking in a mixture of 10% bleach and 0.02% SDS solution and several times rinsing with sterile distilled water. The seeds were then vernalized by keeping them 2-4 days at 4°C in dH2O before planting.
2.3) Regulation studies of the *AtST4* subfamily

2.3.1) Transcript expression study of the *AtST4* subfamily in response to cytokinins

For transcript expression analysis, 16-day-old plants were sprayed with 20 µM *t*-zeatin dissolved in 50% dimethylsulfoxide (DMSO) for 30 minutes, 2 hours, 4 hours and 6 hours. RNA samples were extracted from root tissue using the RNeasy Plant Mini Kit (Qiagen). For reverse transcription polymerase chain reaction (RT-PCR) experiments, 2 µg of total RNA was treated with 2 µl Expand Reverse Transcriptase Buffer 5x (Roche) and 0.2 µl DNase I 40 U/µl (Roche). The volume was adjusted to 20 µl using DEPC-treated water. The reaction was incubated 15 minutes at room temperature, followed by the addition of 2 µl of 30 mM EDTA. DNase I was heat-inactivated at 65°C for 10 minutes and tubes were put on ice for 2 minutes and centrifuged for a short time. Following DNase I treatment, 8 µl of 50 µM Oligo dT (15 mers) was added to each reaction. Then, the tubes were incubated 10 minutes at 65°C, put on ice for 2 minutes and centrifuged for a short time. A mix of 10 µl Expand RT Buffer 5x (Roche), 5 µl dTT (100 mM), 2 µl dNTP (25 mM), 2 µl Expand Reverse Transcriptase 50 U/µl (Roche) was added to each reaction, followed by 60 minutes incubation at 43°C. The synthesized cDNAs were then used in PCR reactions.

The absence of genomic DNA contamination was verified by PCR using the *Actin* primers which span introns of 8 *Actin* genes. Furthermore, the *Actin* genes were used as internal control for RNA calibration. The volumes of cDNA, used as template for PCR, were adjusted after a preliminary calibration based on the levels of *Actin* PCR products. The following thermocycling program was used: initial denaturation at 94°C for 2 minutes, followed by the specified number
of cycles at 94°C for 30 seconds, 55°C for 45 seconds (except for Actin at 60°C), and 72°C for 1 minute, and a final extension at 72°C for 5 minutes.

The following primers were used:

Actin-F (5'-GCTGATGGTGGAAGACATTCA-3')
Actin-R (5'-CATAGCAGGGGCATTGAAAG-3')
AtST4a-F (5'-CGGGATCCATGGATGAAAAGATAGACCAA-3')
AtST4a-R (5'-GGGGTACCTTAGAATTTCAA-3')
AtST4b-F (5'-TGCCATGGGTGAGAAAGATATTCCA-3')
AtST4b-R (5'-CGGGATCCCTACAATTTCAAACCAGAGCC-3')
AtST4c-F (5'-CGCTTAAGTCAATTTCAAACCACACATCA-3')
AtST4c-R (5'-AGAACAAAAACCACACATCA-3')

The primers were used at a concentration of 0.4 µM.

2.4) Molecular characterization of T-DNA insertion mutants

2.4.1) Genomic DNA extraction

The genomic DNA was extracted from leaf tissue of at least 20 individual plants according to the Extract-N-Amp plant PCR kit (Sigma-Aldrich).

2.4.2) Screening for T-DNA insertion mutants

The wild type AtST4a, AtST4b and AtST4c, and their knockout alleles were identified by means of polymerase chain reaction (PCR) using the following primers.

- The wild type AtST4a allele was identified using the following primers in PCR:
  
  AtST4a-F (5'-GGACCCCGCTTTCAAAGTACC-3')
  AtST4a-R (5'-TCAATTTTGTCTACCATTTCCAGG-3')
The AtST4a-R in combination with Gabi-Kat LB (5'-ATATTGACCATCATACTCATTGC-3') was used for detection of *AtST4a*-KO allele.

- The wild type *AtST4b* allele was identified using the following primers in PCR:
  
  AtST4b-F (5'-CTCTTATGCCACCAAAATAACAAG-3')
  
  AtST4b-R (5' - CCGGATCCCTACAATTTCAAAACCAGAGCC-3')

The AtST4b-R in combination with Gabi-Kat LB (5'-ATATTGACCATCATACTCATTGC-3') was used for detection of the *AtST4b*-KO allele.

- The wild type *AtST4c* allele was identified using the following primers in PCR:
  
  AtST4c-F (5'-CGCTTAACCTACCTTGAAG-3')
  
  AtST4c-R (5'- AGAACAAAACACCACATCA-3')

The AtST4c-F in combination with INRA LB (5'-CTACAAAATTGCCCTTTTCTTATCGAC-3') was used for detection of the *AtST4c*-KO allele.

10-15 ng of genomic DNA was used as a template in all PCR reactions. In all cases, the *actin* gene was used as a positive internal control. The thermocycling program was as follows: initial denaturation at 94°C for 2 minutes, followed by 37 cycles at 94°C for 45 seconds, 55°C for 1 minute (except for AtST4a at 58°C), and 72°C for 1 minute, and a final extension at 72°C for 7 minutes. The PCR was performed using Ex Taq DNA polymerase (Takara Biomedicals).
2.4.3) RT-PCR analysis of T-DNA insertion mutants

The two-week old wild type and mutant plants were either treated with 20 µM t-zeatin for 2 hours or mock treated. The RNA was extracted from root tissue using the RNeasy Plant Mini Kit (Qiagen) and cDNA was generated as explained in section 2.3.1. The resulting cDNAs were used as template for knockout confirmation analysis. The following gene specific primers were used in PCR:

AtST4a-F (5'-GGACCCGCTTCAAGTACC-3')
AtST4a-R (5'-TCAATTTTGCTACCATTTCAGG-3')
AtST4b-F (5'-ATGGGTGAGAAAGATATTCCA-3')
AtST4b-R (5'-CTACAATTTCAAACCAGAGCC-3')
AtST4c-F (5'-CGCTTAAACTACCCTTGAAG-3')
AtST4c-R (5'-AGAACAAAAACCACACATCA-3')
Actin-F (5'-GCTGATGGTGAAGACATTCA-3')
Actin-R (5'-CATAGCAGGGGCATTGAAAG-3')

The PCR program was as follows: initial denaturation at 94°C for 2 minutes, followed by specific number of cycles at 94°C for 45 seconds, 55°C for 1 minute (except for AtST4a at 58°C and Actin at 60°C), and 72°C for 1 minute, and a final extension at 72°C for 7 minutes.

2.5) Phenotype analysis of mutant plants

2.5.1) Root growth analysis

Seedlings of wild type and Arabidopsis mutants were grown vertically on MS media. The length of the primary root was marked on the petri dishes at around the same time every
day within the 10 days after germination (DAG) and measured by a ruler. The emerging lateral roots were counted using a Nikon dissecting microscope (Nikon SMZ1500). The plants were photographed 7 DAG with a Nikon D70 camera.

2.5.2) **Determination of rosette diameter, seed size, seed and silique number**

Rosette diameter and number of leaves of soil-grown plants were determined 14, 21 and 26 days after germination (DAG). For each plant, two measurements of rosette diameter were taken with a ruler, and the average of the two values was used for data analysis. Seed size of wild-type, *AtST4b* and *AtST4c* mutant lines were determined measuring the length and width of 40 seeds. The volume was estimated based on formula of volume = 4/3 X π X length X width X depth (Riefler *et al.* 2006). The number of seeds per silique was counted for 30 mature siliques of each genotype. The number of siliques was measured 36 and 46 DAG.

2.5.3) **Statistical analysis**

Data analysis was performed using the PASW Statistics 18 (formerly known as SPSS) software. Since the data were not generally suitable for parametric statistics, comparisons of means from multiple groups were analyzed by the Kruskal-Wallis test. If this test found a significant effect, the differences between each of independent mutant groups and wild type group were analyzed by the Mann-Whitney U. When the data were suitable for parametric analysis, we compared two groups using two-independent-sample test (t-test). A *P* value of 0.05 or less was considered significant in all cases.
2.6) Enzymology

2.6.1) Expression and purification of recombinant enzymes in *E. coli*.

The coding sequences of *AtST4a*, *AtST4b* and *AtST4c* were previously cloned in a bacterial expression system (pQE-30, Qiagen) and transformed in *E. coli* strain XL1-blue. A culture of *E. coli* harboring *AtST4a*-, 4b- and 4c- (O.D₆₀₀ = 0.7) was induced with 1mM isopropylthio-β-D-galactopyranoside (IPDG) for 10 hours at 22°C. Bacterial cells were collected by centrifugation and resuspended in lysis buffer (50mM sodium phosphate, 0.3M NaCl, 10mM imidazole and 14 mM β-mercaptoethanol, pH 8.0). The cells were lysed by sonication, and the recombinant proteins were recovered in the soluble fraction by centrifugation at 13,000 rpm for 20 minutes at 4°C. The soluble recombinant proteins were purified by affinity chromatography onto a nickel-nitrolotriacetic acid agarose matrix (Qiagen) under native condition. Protein concentration was estimated using the Bradford Reagent (Bio Rad) and bovine serum albumin as a reference protein. To verify the solubility and evaluate the level of purity of the recombinant proteins after chromatography, aliquots were subjected to 12% polyacrylamide gel electrophoresis according to the method of Laemmli. The proteins were visualized by 15 minutes staining in 0.1% Coomassie Blue. The gel was later destained with several changes of 40% methanol and 10% of acetic acid for removal of the background coloration.

2.6.2) Sulfotransferase assays

The reaction mixture (50 µl) contained 50 pmol [³⁵S] PAPS (NEN Life science products, Boston, MA), 5µl of hydrolyzed metabolite extract (dissolved in 50% methanol) and approximately 2 µg of the recombinant enzyme preparation (extracted in 50 mM Tris-Cl,
pH 7.5). The reactions were allowed to proceed for 10 minutes at 22°C and then stopped by the addition of 10 µl 2.5% acetic acid. The sulfated enzymatic product was extracted in ice-cold n-butanol and an aliquot was counted for radioactivity using a liquid scintillation counter. The remaining fraction was lyophilized for product identification by high performance liquid chromatography (HPLC), thin layer chromatography (TLC) and mass spectrometry.

2.6.3) Preparation of Arabidopsis thaliana extracts for the detection of the endogenous substrate

Roots of 16 day-old A. thaliana plants grown vertically on MS media were ground to a fine powder in liquid nitrogen, and the powder was homogenized in 50% methanol (approximately 3 ml/g of plant tissue) for 1 hour. Methanol was evaporated and the aqueous phase extracted with 1:1 volume of n-butanol. The butanol extract was lyophilized and resuspended in 50% aqueous methanol. To release the endogenous substrate, the extract was hydrolyzed in 0.1N HCl by boiling at 98°C for 10 minutes and subsequently neutralized in 0.1N NaOH. This fraction was tested directly as substrate for enzymatic assays or purified on a Novapak C18 reverse phase column equilibrated with solvent A (0.05% acetic acid in water) using a Waters 625 LC HPLC system. The column was washed for 5 minutes in 100% solvent A. Metabolites were eluted with a linear gradient of solvent A into solvent B (100% methanol, 0.05% acetic acid) in 50 minutes, followed by 10 minutes in 100% solvent B at a flow rate of 0.5 ml/min. Fractions of 0.5 ml were collected and assayed for activity with the recombinant enzymes. The fractions exhibiting the highest activity were selected and assayed to get enough purified products for HPLC, TLC and mass spectrometry.
2.6.4) Preparation of *Arabidopsis thaliana* extracts for the detection of the endogenous enzymatic products

Approximately 100,000 dpm of $^{35}$S-labeled sulfate product (dissolved in 100 µl of 50% methanol) produced according to the protocol described in section 2.6.3, was fractionated on a Novapack C$_{18}$ reverse phase column equilibrated with solvent A (0.05% acetic acid and 5 mM ammonium acetate in water) using a Waters 625 LC HPLC system. The column was washed for 5 minutes in 100% solvent A. Plant metabolites were eluted with a linear gradient of solvent A into solvent B (100% methanol, 0.05% acetic acid and 5 mM ammonium acetate) in 50 minutes, followed by 10 minutes in 100% solvent B at a flow rate of 0.5 ml/min. Fractions of 0.5 ml were collected and counted for radioactivity using a liquid scintillation counter. Using the same conditions, 1 gram of non-hydrolyzed, non-radiolabeled root extract from wild type and mutant plants were purified on the HPLC system. The fractions corresponding to the elution time of the radioactive product were lyophilized and redissolved in 100 µl 50% methanol for mass spectrometry analysis.

2.6.5) Mass spectrometry

Electrospray ionization tandem mass spectrometry (ESI-MS/MS) was used for analysis of the HPLC-purified product of wild type and T-DNA insertion mutant lines. Data acquisition and evaluation was performed using the Masslynx software. The catalyzed reaction products were analyzed using neutral loss scan in the negative and positive mode in search of a parent ion which gives a neutral loss of 80 mass unit (mass of sulfuryl group). The analyses were preformed on a Quattro triple quadrupole from Micromass using a cone voltage of 20 eV and collision induced dissociation (CID) energy of 35 eV (2.5 mTorr argon). To get structural information and accurate mass of the AtST4b-
sulfated product, the corresponding peak was further analyzed on a Q-TOF2 mass spectrometer (Micromass, UK).

2.6.6) Thin Layer Chromatography

Approximately 15,000 dpm of purified assay products were spotted on cellulose TLC plates (Analtech, 100 microns) and migrated in a mobile phase consisting of butanol, water, and acetic acid (6:2:2, V/V/V). The dried plates were later exposed to Kodak BioMax film for 7 days and then the autoradiograms were developed and scanned.
Chapter 3- Results

3.1) Regulation study of the AtST4 subfamily

3.1.1) Introduction

Genome-wide microarray analysis revealed that members of the AtST4 subfamily are regulated by cytokinins (Hoth et al. 2003; Brenner et al. 2005; Kiba et al. 2005; Lee et al. 2007; Yokoyama et al. 2007). An examination of the available microarray data showed that AtST4b is strongly induced by cytokinins in seedlings of A. thaliana (Genevestigator: ~4.92 fold induction after 1 hour and 3 hours treatment with 1µM zeatin). In contrast, AtST4a and/or AtST4c are slightly repressed by cytokinins on the same data set (Genevestigator: ~0.8 fold repression after 1 hour and 3 hours treatment with 1µM zeatin in seedlings). Furthermore, results from Gene Atlas and the electronic Fluorescent Pictograph (eFP) browser engine showed that the members of the AtST4 subfamily are expressed mainly in the elongation zone of seedling roots (Zimmermann et al. 2004; Winter et al. 2007). Based on these results, a reverse transcription polymerase chain reaction (RT-PCR) experiment was performed to evaluate the effect of cytokinins on transcript expression of the members of the AtST4 subfamily in root tissue.

In addition, since the members of the AtST4 subfamily are regulated by cytokinins, we searched for putative cytokinin response cis-elements upstream of the coding region of AtST4a, AtST4b and AtST4c. The results of these analyses are discussed in the following sections.
3.1.2) Transcript expression study of the *AtST4* subfamily in response to cytokinins

Transcript expression analysis using gene specific primers on 16-days-old Arabidopsis confirmed the regulation of the *AtST4* subfamily members by the cytokinin *trans*-zeatin (*t*-zeatin) in roots. As shown in Fig. 8, 30 minutes after treatment with *t*-zeatin, *AtST4b* transcript level was low, but gradually increased with time. In contrast to *AtST4b*, the expression of *AtST4c* was repressed by *t*-zeatin in roots. While *AtST4c* expression was high in the absence of cytokinins, it decreased gradually following *t*-zeatin treatment, until there was no detectable level of transcript 6 hours after treatment. The *AtST4a* expression level was too low to be detected under the conditions of our experiment.

![Size (bp) vs. No. Cycles for Actin, AtST4a, AtST4b, and AtST4c](image)

**Figure 8. Transcript expression profile of the members of the *AtST4* subfamily in response to cytokinins.** Total RNA was extracted from roots of 16-day-old Arabidopsis (Col-0) vertically grown on MS media treated or non-treated with 20µM *t*-zeatin for various time. The *Actin* was used to quantify the amounts of cDNAs and to test for genomic DNA contamination. Lane 1: 0 hour, lane 2: 30 minutes, lane 3: 2 hours, lane 4: 4 hours, lane 5: 6 hours and lane 6: genomic DNA.
3.1.3) **Study of the presence of cytokinin-responsive elements in the upstream region of the AtST4 genes**

As mentioned earlier in Chapter 1, type-B ARRs are classified as transcriptional factors in the multi-step phosphorelay signal transduction of the cytokinin pathway in Arabidopsis. Analysis of the structure of type-B ARRs has revealed that, in addition to the acidic domain, nuclear localization domain and glutamine-rich domain, they contain a DNA binding domain in the C-terminal extension. This domain is called the B motif or GARP motif and is composed of 60 amino acids. The GARP motif is representative of the plant Myb-related transcription factors. The plant Myb proteins are distinctly related to the well-known mammalian Myb-repeat transcription factors, which bind DNA in a sequence-specific manner (Sakai *et al.* 1998; Imamura *et al.* 1999; Riechmann *et al.* 2000). In an attempt to characterize the DNA target of the plant Myb factors, Hosoda *et al.* found that the B motif derived from ARR 10, a representative of type-B ARR binds specifically to the optimal DNA sequence 5'-AGATT-3' in *in vitro* assays (Hosoda *et al.* 2002). In a different research study, the B motif of ARRI and ARR2 was shown to have the highest binding affinity to the same core sequence (5'-AGATT-3') through gel retardation assays (Sakai *et al.* 2000). The core sequences 5'-GGATC-3' and 5'-GGATT-3' were found to be bound by the B motif of ARRI and ARR2 with less affinity. Moreover, Sakai *et al.* showed that transgenic plants expressing a reporter gene under the control of multiple copies of the sequence 5'-GGATT-3' are significantly activated by ARR1 and ARR2. Following the same procedure, Imamura *et al.* found that a truncated version of ARR11, encompassing the receiver domain and GARP motif, binds more preferably to the sequence motif 5'-GGATT-3', rather than 5'-AGATT-3' in *in vitro*
assays (Imamura et al. 2003). In addition, Rashotte et al. performed a genome-wide microarray analysis to study the genes that are regulated by cytokinins. They found that genes that are consistently induced by cytokinins (for example, type-A ARRs), contain a significantly high proportion of the core sequence motif 5'-AAGATC-3' within the 1 kbp upstream of their translation start site. Interestingly, they found that the frequency of this motif in the upstream region of type-A ARRs is positively correlated with the induction level (Rashotte et al. 2003). For example, while the sequence motif 5'-AAGATC-3' should randomly occur approximately 0.38 times per kilobase pair, cytokinin up-regulated and down-regulated genes contained on average 14 and 6 times this motif within their upstream region, respectively. Overall, the optimal binding core sequence for the type-B ARRs seems to be 5'-(A/G)GAT(T/C)-3' with a specific requirement for the central GAT. However, since this short motif occurs frequently in the Arabidopsis genome, the existence of additional factors has been proposed in order to increase the specificity of the response (Sakai et al. 2000).

Since the members of the AtST4 subfamily are regulated by cytokinins, we searched for the putative cis-acting motifs, mentioned above, in the 1,000 bp upstream of the predicted translation start site of AtST4a, AtST4b and AtST4c. The upstream coding region of ARR5 (At3g48100) and ARR15 (At1g19050) (representatives of type-A ARRs) were used as references for genes that are up-regulated by cytokinins (Genevestigator: ~5.63 and 4.05 fold induction for ARR5 and ARR15, respectively, after 1 hour and 3 hours treatment of seedlings with 1µM zeatin). The upstream coding region of a peroxidase (At5g19890) and auxin-responsive gene (At5g50760) were used as references for genes that are down-regulated by cytokinins (Rashotte et al. 2003; Lee et al. 2007) (Genevestigator: ~0.67 and
0.53 fold repression, respectively, after 1 hour and 3 hours treatment of seedlings with 1µM zeatin). The promoter regions of Actin8 (At1g49240) and a-tubulin (At1g04820) were used as references for genes that are not regulated by cytokinins (Genevestigator: ~0.92 and 1.06 fold changes for Actin8 and a-tubulin, respectively, after 1 hour and 3 hours treatment of seedlings with 1µM zeatin).

As shown in Table 1, the core sequence motif 5'-(A/G)GAT(T/C)-3' occurs more frequently in the upstream promoter region of the putative cytokinin up-regulated genes (ARR5, ARR15) than the cytokinin down-regulated and the cytokinin non-regulated reference genes. The occurrence of this motif in the upstream region of the members of the AtST4 subfamily is less than ARR5 and ARR15, but is more than that of the cytokinin non-regulated genes and close to that of the cytokinin down-regulated genes.

Altogether, the results show that the members of the AtST4 subfamily, especially AtST4b and AtST4c, contain a higher number of the potential cytokinin-response elements compared to their random occurrence in the Arabidopsis genome. Their occurrence in the AtST4 promoters is similar to what is observed for the putative cytokinin-regulated genes. Therefore, they have the potential to be regulated by cytokinin transcriptional activators such as type-B ARRs and subsequently be part of the cytokinin response.
<table>
<thead>
<tr>
<th>Sequence motif</th>
<th>Gene (AGI)</th>
<th>5'-AAGATC-3' (Rashotte et al., 2003)</th>
<th>5'-AGATT-3' (Hosoda et al., 2003; Sakai et al., 2000)</th>
<th>5'-GGATC-3' (Sakai et al., 2000)</th>
<th>5'-GGATT-3' (Imamura et al., 2003; Sakai et al., 2000)</th>
<th>5'-AGATC-3' (Rashotte et al., 2003)</th>
<th>5'- (A/G)GATC(T/C)T-3' (Sakai et al., 2000)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>AtST4a (At2g14920)</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>AtST4b (At1g13420)</td>
<td>2</td>
<td>5</td>
<td>1</td>
<td>2</td>
<td>3</td>
<td>11</td>
</tr>
<tr>
<td></td>
<td>AtST4c (At1g13430)</td>
<td>3</td>
<td>5</td>
<td>1</td>
<td>0</td>
<td>4</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>ARR15 (At1g19050)</td>
<td>6</td>
<td>7</td>
<td>4</td>
<td>2</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>ARR5 (At3g48100)</td>
<td>5</td>
<td>6</td>
<td>2</td>
<td>2</td>
<td>9</td>
<td>19</td>
</tr>
<tr>
<td></td>
<td>Peroxidase (At5g19890)</td>
<td>0</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>Auxin-responsive gene (At5g50760)</td>
<td>1</td>
<td>5</td>
<td>0</td>
<td>3</td>
<td>2</td>
<td>10</td>
</tr>
<tr>
<td></td>
<td>Act8 (At1g49240)</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
<td>3</td>
</tr>
<tr>
<td></td>
<td>α-tubulin (At1g04820)</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>2</td>
<td>1</td>
<td>4</td>
</tr>
</tbody>
</table>

Table 1. Putative cytokinin cis-acting motifs in promoters of the *AtST4* genes. The 1,000 bp upstream of the predicted translation start site of the genes were retrieved from ATTEDII (http://atted.jp/) and analyzed for the presence of cis-acting sequence motifs. The *ARR5* and *ARR15*, peroxidase and auxin-responsive gene and, *Actin8* and *α-tubulin* were used as references for the putative cytokinin up-regulated genes (pink), down-regulated genes (blue) and non-regulated genes (yellow), respectively. The numbers show the frequency of the respective motifs in forward and complementary strands.
3.2) Molecular characterization of T-DNA insertion mutants

3.2.1) Introduction

The primary aim of this study was to determine the role of the *AtST4* genes in plant growth and development. Using a reverse genetic approach, we identified loss of function mutants for *AtST4a*, *AtST4b* and *AtST4c*. Subsequently, we conducted a number of metabolic and phenotypic analyses to determine the biochemical and biological function of the *AtST4* genes in Arabidopsis.

3.2.2) Isolation of T-DNA insertion homozygous mutants

Arabidopsis lines carrying a T-DNA insertion in the *AtST4a* (GABI_177E08), *AtST4b* (GABI_231G06) and *AtST4c* (FLAG_334F06) loci were identified from the publicly available T-DNA insertion libraries. In all these Columbia 0 mutant lines, the T-DNA insertion is located within the intronless coding region. The site of T-DNA insertion was estimated by nucleotide blast using the sequencing information of the T-DNA borders retrieved from the TAIR web site (Fig. 9A). The *AtST4a* and *AtST4c* mutants carry an insertion near the end of their coding sequence (737 bp and 870 bp downstream of their translation start sites, respectively), which separates the regions encoding for the sulfotransferase catalytic domain (region I) and PAPS binding domain (region I and region IV). In the *AtST4b* mutant, the T-DNA insertion is located 74 bp downstream of the translational start site, which is adjacent to the region I involved in catalytic and PAPS binding activity. PCR analysis using gene-specific primers flanking the insertion site and the T-DNA left border (LB) primer showed that the *AtST4a*, *AtST4b* and *AtST4c* mutants are homozygous for the T-DNA insertion. As shown in Fig. 9B, while the intact
gene amplification product was not detected, the PCR-amplified T-DNA insertion product was identified in the mutant lines. In contrast, wild type Arabidopsis showed the presence of only the intact gene amplification product.

3.2.3) Transcript expression analysis of the T-DNA insertion homozygous mutants

To determine if the T-DNA insertion has affected the transcript level of the genes, RT-PCR analysis was performed on RNA samples of the homozygous *AtST4a*, *AtST4b* and *AtST4c* mutant plants. The presence of the respective gene-specific transcripts in wild type plants and their absence in the mutant plants confirmed that *AtST4a*, *AtST4b* and *AtST4c* homozygous mutant lines are null alleles (Fig. 9C). Therefore, the resultant T-DNA insertion mutants were designated as *AtST4a*-knockout (*AtST4a*-KO), *AtST4b*-knockout (*AtST4b*-KO) and *AtST4c*-knockout (*AtST4c*-KO) lines. It is important to note that in contrast with Fig. 8, we could detect amplification of the *AtST4a* transcript. This was obtained using a different primer pair generating a shorter PCR product. In the future, studies of *AtST4a* expression will have to be conducted using this primer set.
Figure 9. Identification of T-DNA insertion homozygous lines of the *AtST4* genes. (A) Schematic representation of the chromosomal region encompassing the *AtST4* genes and their T-DNA insertion site. The black boxes represent: exons, white boxes: 5' and 3'UTRs, triangles: T-DNA insertion sites, arrows:
gene-specific primers, LB: T-DNA left border. The consensus sequence from region I (PKxGTTWLKAL) and region IV (FRKGxVGDWK) were used to identify the regions involving in catalytic and PAPS binding activity. (B) PCR screening of the T-DNA insertion lines. Genomic DNA was extracted from leaves of three-week old AtST4a (GABI_177E08), AtST4b (GABI_231G06) and AtST4c (FLAG_334F06) T-DNA insertion lines and used as a template for PCR screening. The Actin gene was used as a positive control. (C) RT-PCR analysis of the AtST4 T-DNA insertion homozygous lines and wild type plants. The two-week old mutant and wild type plants were treated (+) or non-treated (-) two hours with 20µM t-zeatin. Total RNA was extracted from roots and subjected to RT-PCR analysis using gene specific primers. 4a, 4b, 4c and Act are abbreviations for AtST4a, AtST4b, AtST4c and Actin genes, respectively.

3.3) Biochemical characterization of AtST4b

3.3.1) Introduction

To identify the substrate of the AtST4b sulfotransferase, plant metabolites were extracted and purified from roots of 14-day-old Arabidopsis. Subsequently, we used radiolabeled PAPS as sulfonate donor to assay the activity of the recombinant enzyme with the plant extract. To confirm the identity of the product of the reaction and to demonstrate its accumulation in vivo, the metabolite profile of the AtST4b T-DNA insertion homozygous line was compared with that of wild type plants using neutral loss mass spectrometry and liquid chromatography-tandem mass spectrometry (LC-MS/MS) analyses.

3.3.2) Expression of AtST4b recombinant sulfotransferase

To determine the biochemical function of AtST4b, the coding sequence was cloned in an E.coli expression plasmid as a fusion protein with 6 histidine residues at the N-terminus. The His-tag was used for affinity purification of the enzyme on Ni-agarose column. The partially purified recombinant enzyme corresponded to the expected size of the protein (~37.71 KDa) on SDS-PAGE (Fig. 10).
Figure 10. SDS-PAGE of partially purified AtST4b recombinant sulfotransferase. Cultures of bacteria expressing the recombinant enzyme were grown 14 hours at room temperature, induced or non-induced by IPTG and subjected to SDS-PAGE analysis. An aliquot of the induced culture was used for purification by affinity chromatography on Ni-NTA agarose. M: protein molecular marker, US: non-induced soluble proteins, IS: induced soluble proteins and P4b: purified AtST4b.

3.3.3) HPLC purification of AtST4b substrate and product

HPLC purification of AtST4b substrate

To determine the substrate of AtST4b, enzyme assays were performed on a library of chemicals including the previously reported sulfotransferase substrates, but none of them were accepted by the recombinant enzyme.

Based on the available microarray data on Genevestigator, the enzyme is expressed almost exclusively in roots of *A. thaliana*, therefore, root extracts were used to purify the potential substrate of AtST4b (Zimmermann *et al.* 2004). In order to remove the sulfonate group of the extracted metabolites, mild acid hydrolysis was performed prior to the assay and purification. The total root extracts were later fractionated using reverse
phase High Performance Liquid Chromatography (HPLC) and the individual fractions were assayed with the AtST4b recombinant enzyme. The results show that fraction 34, 37 and 42 have the highest enzymatic activity (Fig. 11). The presence of multiple peaks can be explained in part by the nature of the substrate of the enzyme which might contain chiral carbons giving rise to stereoisomers that could be resolved on the HPLC column.

**Figure 11. HPLC profile of an acid hydrolyzed root extract.** Fraction 34, 37 and 42 showed the highest enzyme activity.

**HPLC purification of AtST4b enzymatic reaction product**

To identify the product of the reaction, enzyme assays were performed on fraction 37 that exhibited the highest enzyme activity. Following the enzyme assay, the AtST4b radiolabeled-sulfated product was purified by reverse phase HPLC. The results show that the product elutes at 27 min (Fig. 12).
3.3.4) Neutral loss mass spectrometry of AtST4b purified product

Because of their fragility, sulfonated compounds can easily be detected using neutral loss mass spectrometry. By increasing the collision energy, the sulfuryl group (SO₃⁻) is easily removed from the parent molecule. This chemical feature can be used to identify sulfonated compounds in a complex mixture by looking for parent ions losing a mass of 80 daltons (SO₃⁻) following the increase in collision energy. We used this diagnostic tool to identify the sulfonated compound(s) present in fraction 27 of a purified root extract from wild type plants. Figure 13 (upper graph) shows that several putative sulfonated metabolites are present in this fraction with a major one (more than 98% of ion count) having a mass-to-charge ratio (m/z) of 514 daltons in negative mode ([M-H]⁻). To identify the AtST4b enzymatic product in the fraction, we compared the wild type neutral loss metabolite profile with the one obtained with an AtST4b-KO root extract under the same condition. The root extract of the AtST4b T-DNA insertion homozygous line
showed almost the same profile as the wild type extract, except for the absence of the 514 [M-H] peak (Fig. 13 lower graph). These results indicate that the reaction product synthesized in vitro by the recombinant AtST4b enzyme accumulate in vivo and has a mass of m/z 514 daltons [M-H].

Figure 13. Mass spectrometry analysis of AtST4b-reaction product. Neutral loss profiles of a root extract from wild type (upper graph) and AtST4b-KO mutant (lower graph) plants in negative mode. The major peak at m/z 514 is missing in the AtST4b-KO scan.

3.3.5) LC-MS/MS analysis of AtST4b purified product

To elucidate the chemical structure of the compound having a mass of 514 [M-H] daltons, a cytokinin-treated root extract was subjected to reverse phase liquid chromatography-tandem mass spectrometry (LC-MS/MS) analysis. Initially, the fragmentation pattern was obtained using a nano liquid chromatograph coupled with a tandem quadrupole time-of-flight (Q-TOF) mass spectrometer at the Centre for
Biological Applications of Mass Spectrometry (CBAMS, Concordia University). Fig. 14A shows the fragmentation pattern of the AtST4b enzymatic product. In the positive ion electrospray mass spectrum, the AtST4b sulfonated product produced a protonated molecular ion at m/z 516 ([M+H]+). The MS/MS fragmentation of this compound gave major fragment ions at m/z 348 and 291, which correspond to the loss of spermidine or a fragment of it (Fig. 14A and B). The molecular ion at m/z 436 is due to the cleavage of the sulfonate moiety from the parent ion (m/z 516). Further fragmentation of the 436 [M+H]+ ion, produced the ions at m/z 419 (due to loss of NH3), 362, 348, 320, 291, 265, 263 and 203 which have been previously reported in the MS/MS fragmentation of cadabicine from Capparis spinosa (Khanfar et al. 2003). The structure and accurate mass of cadabicine and cadabicine-sulfate was further determined by Dr. Jurgen Schmidt from the Institute of Plant Biochemistry (Halle, Germany) using Fourier Transform Ion Mobility Spectrometry (FT-IMS) (Fig. 14B). Even though cadabicine has been reported to occur in nature in the root bark of Capparaceae (Capparis spinosa and Capparis decidua) and in seeds of Brassicaceae (Brassica napus), it is the first time that it is identified from A. thaliana (Khanfar et al. 2003; Baumert et al. 2005). Furthermore, it is the first report of the natural occurrence of its sulfonated derivative.
Figure 14. Identification of AtST4b-sulfated product. (A) LC-MS/MS spectrum of the AtST4b sulfated product in positive mode. The possible structure of some major fragments is shown. (B) Structure of cadabicine and cadabicine sulfate with exact mass determination using FT-IMS. Cadabicine contains an amide conjugate (spermidine, shown in red) and two acyl conjugates (hydroxycinnamic acid, shown in black and blue).
3.4) Characterization of *AtST4b* biological function

3.4.1) Introduction

Having identified the structure of the *AtST4b* enzymatic product, we attempted to determine its function *in planta* by comparing the growth behavior of the *AtST4b* T-DNA insertion homozygous mutant (*AtST4b*-KO) with wild type plants.

3.4.2) Primary root length analysis

Cytokinins inhibit root elongation. Furthermore, cytokinin receptor double mutants exhibit longer roots and increased number of lateral roots (Higuchi *et al.* 2004; Riefler *et al.* 2006). Moreover, as it was shown earlier, *AtST4b* expression is regulated by cytokinins and its expression is restricted to root tissue (Marsolais *et al.* 2007). Consequently, we analyzed the root growth behavior in the *AtST4b*-KO mutant in presence and absence of cytokinins.

Observation of primary root growth 10 days after germination (DAG) under *in vitro* conditions demonstrated that the root length of *AtST4b*-KO had increased 6.5 to 16.5% more than that of wild type seedlings (Fig. 15A and B). Interestingly, the *AtST4b*-KO plants had a greater growth rate in the beginning and mid (between 2 to 7 DAG with the highest increase at 4 DAG) than at the end of our analysis. To determine the role of *AtST4b* in the cytokinin response pathway, we examined root elongation in response to the exogenous application of the cytokinin i-zeatin. We compared primary root elongation of wild type and the *AtST4b*-KO line in presence of 1 and 5µM i-zeatin at different stages of development. Fig. 15C shows that 5 DAG, roots of *AtST4b* mutant plants are less sensitive to different concentrations of cytokinins. Similar differences were
also obtained 7, 12, and 14 DAG (data not shown). Such an effect has been reported for cytokinin receptor mutants and cytokinin deficient plants (Werner et al. 2001; Higuchi et al. 2004; Nishimura et al. 2004). Altogether, our results suggested that AtST4b participate in the root growth inhibition mediated by cytokinins.

3.4.3) Number of lateral roots

Formation of lateral roots is also inhibited by cytokinins (Werner et al. 2001; Werner et al. 2003). To assess the effect of AtST4b on lateral root formation, we counted the number of lateral roots in 12- and 14-day-old wild type and AtST4b-KO plants. As shown in Fig. 15D, even though the number of lateral roots had been increased 24% to 25% in the AtST4b-KO plants, there is no statistically significant difference between the wild type and AtST4b-KO lines 12 and 14 DAG with a P-value of 0.06 at 14 DAG. Even though the trend for increased lateral root is evident, the small sample size used for the analyses probably did not allow assessing the differences between the mutant line and control plants. A larger sample size will be required to evaluate the role of AtST4b in the growth of secondary roots.
Days After Germination (DAG)

Root Length (mm)

- WT
- AIST4b-KO

0.028
2.88E-4
8.69E-06
0.001
0.015
0.001
0.031
Figure 15. Root phenotype of the AtST4b-KO mutants. (A) Kinetic of root elongation in wild type (Col-0) and AtST4b-KO plants. Plants were grown vertically on MS agar plates under long day conditions and
root length was measured every day. Label for each day shows the corresponding significant P-value. Error bars represent standard deviation (SD) (27≤n≤44). (B) Root system of in vitro grown wild type (left) and AtST4b-KO (right) plants 7 DAG. (C) Relative changes of primary roots length in wild type (Col-0) and AtST4b-KO plants 5 DAG in presence of increasing concentration of t-zeatin. Root length of wild type plants in the absence of cytokinins was set at 100%. Length of roots in the absence of t-zeatin was: WT: 0.81 ± 0.11 (mean ± SD), AtST4b-KO: 0.93 ± 0.18. (n=19). (D) Number of lateral roots 12 and 14 DAG in plants grown vertically on MS agar plates. Error bars represent SD (9≤n≤12). Analysis of the significance level between wild type and the AtST4b-KO plants was performed by either Kruskal-Wallis test followed by Mann-Whitney U test or by two-independent-sample test (t-test). Only P-values with a significant level less than 0.05 are shown. See Annex 1 for details of statistical values.

3.4.4) Rosette diameter

Study of cytokinin-deficient plants and cytokinin receptor mutants indicate that shoot growth is impaired in these plants (Werner et al. 2001; Higuchi et al. 2004; Nishimura et al. 2004; Riefler et al. 2006). Analysis of the rosette structure of the AtST4b-KO showed that even though there is a statistical increase in size at 21 DAG (P-value=0.028), the significance level is relatively high and there is no general pattern of significant increase at 14 and 26 DAG (Fig. 16A). It is important to mention that this experiment was repeated earlier, in which there was a statistically significant increase in rosette size of the AtST4b-KO plants (17% to 30% increase) compare to wild type plants (P-value was equal to 2.53E-5, 4.96E-4 and 2.82E-5 at 14, 21 and 26 DAG, respectively) (see the rosette structure in Fig. 16C). The changes in the significance level could arise from a number of parameters such as sample size, variance or changes of experimental conditions such as light, soil, seeds, etc. On the other hand, based on the general trend of increase in the AtST4b-KO plants in both analyses, it can be said that AtST4b might negatively control rosette size.
3.4.5) Number of leaves

Further analysis of shoot development showed that the rate of leaf formation is affected during the vegetative growth in the *AtST4b*-KO mutants. As shown in Fig. 16B the *AtST4b*-KO plants developed approximately 6% to 9% more leaves than wild type plants of the same age (14, 21 DAG) (*P*-value is equal to 0.013 and 4.21E-4 at 14 and 21 DAG, respectively). The increase in number of leaves of *AtST4b*-KO plants was reproducible and observed in a separate experiment.
Figure 16. Shoot phenotype of the AtST4b-KO plants. (A) Rosette diameter of in-soil grown plants 14, 21 and 26 DAG. Error bars represent SD \((21 \leq n \leq 32)\). (B) Number of leaves of in-soil grown plants 14, 21 DAG. Error bars represent SD \((21 \leq n \leq 39)\). (C) Rosette of wild type (left) and AtST4b-KO (right) plants 21 DAG. Analysis of the significance level between wild type and the AtST4b-KO plants was performed by a Kruskal-Wallis test followed by a Mann-Whitney U test. Only \(P\)-values with a significant level less than 0.05 are shown. See Annex 2 for details of statistical values.
3.4.6) Flowering time

To assess the role of AtST4b during reproductive development, a number of phenotype analyses were conducted. Based on these analyses, there are no apparent changes in flowering time and flower structure between wild type Arabidopsis and the AtST4b-KO mutants.

3.4.7) Seed production

Increased grain production has been linked to the reduced level of cytokinin oxidase activity and the subsequent accumulation of cytokinins in the inflorescence of rice (Ashikari et al. 2005). In Arabidopsis, the reduced cytokinin content of cytokinin oxidase overexpression lines reduced the number of flowers and seeds in siliques, but increased the size of the seeds (Werner et al. 2003). Under long day conditions, the AtST4b-KO plants produced on average ~33% more siliques than wild type plants during reproductive development (36 and 46 DAG) (Fig. 17A and C). These results were reproducible and seen in a separate experiment (data not shown). Moreover, no significant changes were found in the number of seeds in mature siliques (P-value=0.06) (Fig. 17B).
AtST4b-KO

Days After Germination (DAG)

P-value = 1.72E-04

P-value = 0.011

No. of Siliques

Days After Germination (DAG)

No. of Seeds per Silique

Days After Germination (DAG)
Figure 17. Reproductive development is improved in the AtST4b-KO line. (A) Number of siliques per plant 36 and 46 DAG. Error bars represent SD (21<n<24). (B) Number of seeds per mature siliques 57 DAG. Error bars represent SD (n=30). (C) Photograph of wild type (left) and AtST4b-KO (right) plants 5 weeks after germination. Analysis of the significance level between wild type and the AtST4b-KO plants was performed by a Kruskal-Wallis test followed by a Mann-Whitney U test. Only P-values with a significant level less than 0.05 are shown. See Annex 3 for details of statistical values.

3.4.8) Seed size and volume

As mentioned earlier, plants with impaired cytokinin signaling (cytokinin receptor mutants) and plants with reduced cytokinin content (cytokinin oxidase overexpressor lines) produce bigger seeds (Werner et al. 2003; Riefler et al. 2006). To investigate the role of AtST4b in seed development, the size of wild type and mutant seeds were measured under a dissecting microscope. As shown in Fig. 18A and B, the average length and width of the mutant seeds are about 7.4% and 4.5% greater than that of wild type
seeds, respectively. These changes increased the seed volume up to ~17% in the AtST4b-KO plants (Fig. 18C).
Figure 18. Seed size of the AtST4b-KO plants. (A) The length and width of seeds from wild type and the AtST4b-KO plants. Error bars represent SD (n=40). (B) Calculated volume of wild type and the AtST4b-KO seeds based on the formula: volume = \( \frac{4}{3} \pi \times \text{length} \times \text{width} \times \text{depth} \) (Riefler et al. 2006). Error bars represent SD (n=40). (C) Wild type (top) compared to the AtST4b-KO seeds (bottom). Analysis of the significance level between wild type and the AtST4b-KO seeds was performed by a Kruskal-Wallis test followed by a Mann-Whitney U test. Only P-values with a significant level less than 0.05 are shown. See Annex 3 for details of statistical values.

3.5) Biochemical characterization of AtST4a and AtST4c

3.5.1) Introduction

In a previous study, Marsolais et al. showed that AtST4a has brassinosteroid sulfotransferase activity in vitro while AtST4b and AtST4c were inactive with the same substrate (Marsolais et al. 2007). This study was limited since it was conducted with a small number of pure compounds from our laboratory library collection. In order to elucidate the biochemical function of AtST4a and AtST4c in vivo, sulfated metabolite profiles of wild type and knockout plants were analyzed using a combination of HPLC
and mass spectrometry. Metabolites were also purified from *A. thaliana* and assayed *in vitro* with recombinant AtST4b and AtST4c.

### 3.5.2) Expression of AtST4a and AtST4c recombinant sulfotransferases

The coding sequences of *AtST4a* and *AtST4c* were previously cloned in the bacterial expression plasmid pQE 30 (Qiagen) containing a 6x His-tag at the N-terminus. Expression of the enzymes produced proteins with the expected size of 38.8 and 37.7 kDa for AtST4a and AtST4c, respectively (Fig. 19).

![Figure 19. SDS-PAGE of AtST4a and AtST4c recombinant sulfotransferases.](image)

**Figure 19. SDS-PAGE of AtST4a and AtST4c recombinant sulfotransferases.** IPTG-induced or control cultures of bacteria expressing the recombinant enzymes were grown 14 hours at room temperature and their protein extracts were subjected to SDS-PAGE analysis. Aliquots of the induced culture were used for purification by affinity chromatography on Ni-NTA agarose. M: protein molecular marker, Ctl: PQE30 empty vector, CE: bacterial crude extract (25 µg), US: non-induced soluble proteins (25 µg), IS: induced soluble proteins (25 µg) and P4a and P4c: purified AtST4a and AtST4c (10 µg).
3.5.3) HPLC purification of the substrate and product

HPLC purification of AtST4a and AtST4c substrate

Gene regulation studies showed that AtST4a and AtST4c are expressed only in the root system of Arabidopsis (Marsolais et al. 2007). Accordingly, root extracts were used to purify the endogenous substrate and product of AtST4a and AtST4c. Following mild acid hydrolysis (to release the potential substrate from its conjugated form), the extract was fractionated by reverse phase HPLC. Each fraction was assayed with recombinant AtST4a and AtST4c to detect the presence of the substrate. The results show that the substrate of both enzymes elute in fraction 32 (Fig. 20).

Figure 20. HPLC purification of AtST4a and AtST4c endogenous substrate. Individual fractions were assayed for activity with recombinant AtST4a and AtST4c.
HPLC purification of AtST4a and AtST4c reaction products

To identify the elution time of the reaction products of AtST4a and AtST4c, enzymatic assays were performed on the purified fraction containing the potential substrate (fraction 32) in order to produce the radiolabeled-sulfated products. Subsequently, the reaction products were purified on a reverse phase HPLC column and the individual fractions counted for radioactivity. The results show that the reaction product of both enzymes elute in fraction 24 (Fig. 21).

![Figure 21. HPLC purification of AtST4a and AtST4c radiolabeled product. The highest activity was recovered in fraction 24.](image)

3.5.4) Neutral loss mass spectrometry of AtST4a and AtST4c reaction products

Having determined the elution time of AtST4a and AtST4c enzymatic reaction products, root metabolites of wild type and AtST4a- or AtST4c-KO plants were purified under the
same conditions and their respective HPLC fractions (fraction 24) were analyzed by neutral loss mass spectrometry. We did not find any difference in the sulfated metabolite profile of the two knockout plants when compared with the wild type purified fraction 24. This result is not totally unexpected when we consider that the substrate and the product of both enzymes co-elute on HPLC suggesting a redundant function for AtST4a and AtST4c. It is also possible that the abundance of the in vivo products is too low for detection under our experimental conditions. Transgenic lines overexpressing AtST4a or AtST4c will have to be produced and analyzed to help in the detection of the endogenous sulfated product.

3.5.5) Thin layer chromatography of AtST4a and AtST4c reaction products

We further examined the purified AtST4a and AtST4c radiolabeled enzymatic products on TLC plates. As shown in Fig. 22, the reaction products are exhibiting the same chromatographic behavior, suggesting that the two enzymes sulfonate the same substrate from the purified fraction 32.
Figure 22. Thin layer chromatography of AtST4a and AtST4c sulfated products. 15,000 DPM of purified AtST4a and AtST4c products were spotted on a cellulose TLC plate.

3.6) Characterization of \textit{AtST4a} and \textit{AtST4c} biological function

3.6.1) Introduction

In order to determine their biological function, the consequence of changing the sulfated metabolome of \textit{AtST4a} and \textit{AtST4c} single mutant was analyzed by monitoring a number of growth parameters known to be under the control of cytokinins.

3.6.2) Primary root length analysis

As mentioned earlier, the three members of the \textit{AtST4} subfamily are expressed in the root system which is also the site of synthesis of cytokinins (Marsolais et al. 2007). To
determine the role of these genes in root growth and development, *AtST4a-KO* and *AtST4c-KO* plants were grown vertically on MS agar plates, and their primary root length was measured for 10 days under long day growth conditions. The results show that the roots of *AtST4c-KO* plants were statistically shorter than those of the wild type plants (9 to 37% shorter) for the first eight days after germination (Fig. 23A). The decrease in the root growth rate of *AtST4c-KO* was greater at the beginning of their lifespan (28% between 2 to 6 DAG as opposed to 9% between 6 to 8 DAG). In contrast, *AtST4a-KO* plants showed a significant decrease in root length only after 8 days of growth. Moreover, the decrease in primary root growth rate was not as great as the one observed for the *AtST4c-KO* plants (6.5 to 14% decrease between 8 DAG and 10 DAG).

3.6.3) Number of lateral roots

A loss of function mutation in *AtST4a* did not cause a strong effect on the number of lateral roots at 10 and 13 DAG. In contrast, *AtST4c-KO* plants produced 17% to 28% less lateral roots than wild type plants (*P*-value is 4.49E-4 and 0.069 at 10 DAG and 13 DAG, respectively) (Fig. 23B).
Figure 23. Root development of *AtST4a-KO* and *AtST4c-KO*. (A) Kinetic of root growth in wild type, *AtST4a-KO* and *AtST4c-KO* plants. Plants were grown vertically on MS agar plates and the root length was measured daily. Labels for each day show the corresponding *P*-value for *AtST4a* (green) and *AtST4c* (blue). Error bars represent standard deviation (SD) (17≤*n*≤44). (B) Number of lateral roots of plants grown on MS agar plates 10 and 13 DAG. Error bars represent SD (12≤*n*≤21). Statistical analysis was performed by a Kruskal-Wallis test followed by a Mann-Whitney U test. Only *P*-values less than 0.05 are shown. See Annex 1 for detailed statistical values.

3.6.4) Rosette diameter

It is well known that cytokinins promote cell division and shoot development in plants (Mok and Mok 2001; Werner *et al.* 2003). Analysis of the Genevestigator (web-browser data mining interface for Affymetrix Gene Chip data) showed that cytokinins slightly repress *AtST4a* and/or *AtST4c* expression in seedlings (~0.8 fold repression after 1h and
3h treatment with 1µM t-zeatin). Analysis of long day grown AtST4a-KO plants showed a slight increase in rosette diameter 21 DAG compared to wild type plants (P-value=0.035). However, the change is very small and the significance of the result is weak. By contrast, AtST4c-KO showed a slight decrease in rosette diameter 26 DAG compared to wild type plants (P-value=0.014) (Fig. 24A). This experiment was repeated for AtST4c-KO plants and we observed a significant decrease of 13% (P-value=0.017) to 30% (P-value =1.03E-7) in rosette size compared to wild type plants 21 and 26 DAG, respectively (data not shown). The consistent reduction in rosette size observed in both experiments suggests that AtST4c might play a positive role in the growth of the aerial tissue in Arabidopsis.

We also observed that the rosette of AtST4c-KO plants is bending towards the ground. A careful observation of the seedlings showed an increase in hypocotyl length in the AtST4c-KO plants (Fig. 24C).

3.6.5) Number of leaves

We also examined leaf formation in AtST4a-KO and AtST4c-KO plants. The number of leaves of AtST4a-KO plants was slightly increased when compared to wild type plants (~6.8% more at 21 DAG, P-value=0.010) (Fig. 24B). However, the contribution of AtST4a on growth is marginal when we consider the very small increase in leaf number and the relatively high P-value observed in this experiment. In contrast, leaf number was significantly reduced in AtST4c-KO plants with ~43% less leaves 21 DAG (P-value=9.03E-12).
3.6.6) Flowering time

Under long day conditions, \textit{AtST4c-KO} plants flowered earlier than wild type plants, while \textit{AtST4a-KO} showed no apparent changes in the time of flowering. Approximately 25 DAG, wild type plants have 12 to 14 leaves and start to flower. In contrast, \textit{AtST4c-KO} plants initiated flowering 21 DAG and have 6 to 7 leaves (Fig. 24D). The morphology of the flowers in \textit{AtST4a-KO} and \textit{AtST4c-KO} was found to be similar to wild type plants.

![Graph showing rosette diameter over time](image)
Days After Germination (DAG)

- **WT**
- **AtST4a-KO**
- **AtST4c-KO**

Number of Leaves

- **P-value = 0.010**
- **P-value = 9.03E-12**

Days After Germination (DAG)

- 14
- 21
Figure 24. Shoot phenotype of AtST4a-Ko and AtST4c-KO. (A) Rosette diameter of in-soil grown plants 14, 21 and 26 DAG. Error bars represent SD (21≤n≤27). (B) Number of leaves of in-soil grown plants 14, 21 DAG. Error bars represent SD (21≤n≤39). (C) Comparison of hypocotyl region in wild type (top) and AtST4c-KO (bottom). (D) Early flowering phenotype of AtST4c-KO (right) was compared to wild type (left) plants at 22 DAG. Analysis of significance between wild type and mutant lines was performed using a Kruskal-Wallis test followed by a Mann-Whitney U test. Only P-values with a significance level less than 0.05 are shown. See Annex 2 for detailed statistical values.

3.6.7) Seed production

The number of siliques was higher in AtST4a- and AtST4c-KO plants than wild type plants. AtST4a-KO plants produced on average 62% more siliques 36 and 46 DAG as compared to wild type plants. Silique number was even higher in the AtST4c-KO line (~360% and 76% more siliques 36 and 46 DAG, respectively) than that of AtST4a (Fig. 25A and C). Analysis of mature siliques showed that the number of seeds per silique was not changed significantly in AtST4a-KO. However, there was a ~17% reduction in seed number in the siliques of AtST4c-KO plants compared to the wild type counterparts (P-value=4.94E-7) (Fig. 25B).
Figure 25. Reproductive developments of AtST4a- and AtST4c-KO. (A) Number of siliques per plant 36 and 46 DAG. Error bars represent SD (n ≥ 23). (B) Number of seeds per mature siliques 57 DAG. Error bars represent SD (n=30). (C) Photograph of wild type (left) and AtST4c-KO (right) plants 5 weeks after germination. Analysis of significance between wild type and mutant lines was performed using a Kruskal-Wallis test followed by a Mann-Whitney U test. Only P-values with a significant level less than 0.05 are shown. See Annex 3 for detailed statistical values.

3.6.8) Seed size and volume of the AtST4c-KO plants

Analysis of seed size was only performed on the AtST4c-KO plants. The results presented in Fig. 26A and 26C show that mutant seeds are smaller than wild type seeds. AtST4c-KO seeds are approximately 6.2% (length) and 8.6% (width) smaller when compared to wild type seeds. As a consequence of the reduction in seed length and width, the AtST4c-KO seed volume is decreased by 21.5% compared to wild type seed volume (Fig. 26B).
**A**

![Graph showing seed length and width distribution](image)

- Seed Length: P-value = 0.003
- Seed Width: P-value = 8.57E-06

**B**

![Graph showing seed volume distribution](image)

- Seed Volume: P-value = 2.77E-06
Figure 26. Seed size of *AtST4c-KO* plants. (A) Width and length of seeds in wild type and *AtST4c-KO* plants. Error bars represent SD (n=40). (B) Calculated volume of wild type and *AtST4c-KO* seeds based on the formula \( \frac{4}{3} \pi \times \text{length} \times \text{width} \times \text{depth} \) (Riefler et al. 2006). Error bars represent SD (n=40). (C) Wild type seeds (right) compared to *AtST4c-KO* seeds (left). Analysis of significance between wild type and mutant lines was performed using a Kruskal-Wallis test followed by a Mann-Whitney U test. Only \( P \)-values less than 0.05 are shown. See Annex 3 for detailed statistical values.
Chapter 4- Discussion and Future work

Soluble sulfotransferases are a superfamily of enzymes that are widely distributed in plants, bacteria and mammals. Our knowledge about their function in plants is suffering from the lack of systematic investigations of plant sulfated metabolomes. Out of the 18 sulfotransferase-coding genes in A. thaliana, seven have been fully characterized. Herein, we describe the results of our studies on the biochemical and biological characterization of the three members of the AtST4 subfamily (AtST4a, AtST4b, and AtST4c) in A. thaliana.

4.1) Biochemical and biological characterization of AtST4b

AtST4b shares 72% amino acid sequence identity with AtST4a and 77% amino acid sequence identity with AtST4c (for the phylogenic tree, refer to Fig. 2). Analysis of the publicly available microarray data showed that AtST4b is specifically expressed in roots of young seedlings and up-regulated by cytokinins (Zimmermann et al. 2004). Such an up-regulation of AtST4b by cytokinins has been reported in several genome-wide microarray studies of A. thaliana (Hoth et al. 2003; Kiba et al. 2005; Lee et al. 2007; Yokoyama et al. 2007) (for a summary refer to Chapter 1). These results were confirmed by transcript expression analysis of the roots of Arabidopsis where AtST4b showed gradual up-regulation by the exogenous application of the cytokinin t-zeatin (Fig. 8). The detailed analyses of the microarray data showed that in transgenic plants that over-express ARR22 (a type-B ARR), AtST4b is up-regulated 11.3 fold compared to wild-type plants (Kiba et al. 2005). A change in expression of AtST4b is also reported in the microarray analysis of the ARR10 and ARR12 double mutant (two type-B ARRs)
(Yokoyama et al. 2007). These results are consistent with the presence of a relatively high number of cytokinin response elements recognized by type-B ARRs in the promoter region of the \textit{AtST4b} gene (Table 1). Taken together, these results support the hypothesis that the function of the \textit{AtST4b} gene might be related to the cytokinin plant response.

Metabolite analysis of \textit{AtST4b}-KO plants revealed that cadabicine (a dicoumaroyl spermidine conjugate) is the substrate of AtST4b in \textit{A. thaliana} and that the formation of cadabicine sulfate is dependent on AtST4b sulfotransferase activity (Fig. 13). Metabolite profiling of the wild type and mutant plants showed that although cadabicine sulfate is missing in the \textit{AtST4b}-KO plants, the substrate (cadabicine, m/z 436) is present in the root extract of both \textit{AtST4b}-KO and wild type plants (data not shown).

Cadabicine has only been reported to occur in members of the Capparaceae and Brassicaceae family which belong to the Brassicales order (Khanfar et al. 2003; Baumert et al. 2005). In this study, we demonstrate for the first time that cadabicine sulfate occurs naturally in roots and leaves (data not shown) of \textit{A. thaliana}. To date, there is no information available on the function of spermidine conjugates such as cadabicine on growth and development. Our results suggest that the sulfonation of cadabicine might be related to the effects mediated by cytokinins on root and shoot development.

The results of several experiments revealed the important role of the polyamine spermidine during growth and development in various organisms. For example, the study of the spermidine synthesis coding gene mutant (\textit{\text{Aspe2}}), showed that spermidine and/or spermine are absolutely required for the growth and cell division of \textit{Saccharomyces cerevisiae} cells (Chattopadhyay et al. 2002). Other studies have shown that it is the presence of spermidine and not spermine, which is absolutely necessary for the growth of
yeast cells (Hamasaki-Katagiri et al. 1997; Hamasaki-Katagiri et al. 1998). Similar results were obtained in higher plants. Imai et al. showed that even though spermine is not essential for survival, suppression of the two genes involved in spermidine biosynthesis (SPDS1 and SPDS2) is lethal in Arabidopsis, and results in embryonic arrest in the double mutant seeds (Imai et al. 2004; Imai et al. 2004).

Acylated polyamines are synthesized by members of the BAHD acyltransferase family. These enzymes are widely distributed in plants and play regulatory roles during plant growth and development (for a review refer to (Facchini et al. 2002; Bienz et al. 2005). The accumulation of hydroxycinnamic acid spermidine conjugates has only recently been reported in A. thaliana (Grienenberger et al. 2009; Luo et al. 2009). Using functional genomics and metabolic profiling, Luo et al. determined that sinapoyl and disinapoyl spermidine are the major polyamine conjugates that accumulate in Arabidopsis seeds. Furthermore, they characterized At2g25150 encoding for a spermidine dicoumaroyl transferase (SCT) that transfers two coumaric acids to its spermidine acyl acceptor molecule. Interestingly, SCT is predominantly expressed in root tips and up-regulated by \textit{t}-zeatin. Its co-regulation and co-expression with \textit{AtST4b} suggests that the N1, N8-di(coumaroyl)-spermidine produced by SCT is a precursor for the synthesis of cadabicine, the substrate of \textit{AtST4b}. In support of this hypothesis, metabolite analysis of SCT-knockout Arabidopsis plants conducted in our laboratory showed the absence of cadabicine and cadabicine sulfate in root tissue (data not shown).

It has been shown previously that the exogenous application of cytokinins increase the accumulation of free polyamines and of their conjugates in plants (Altman 1988; Sergive \textit{et al.} 1995). More recently, the up-regulation of spermidine synthesis coding genes was
reported in response to the application of cytokinins on Arabidopsis seedlings (Hanzawa et al. 2002). In addition, several recent studies show that there is an interaction between the cytokinins and polyamines signaling pathways. Different polyamines, especially spermine, prevent cytokinin-induced expression of ARR5, a type-A response regulator gene, in Arabidopsis and Amaranthus (Rakova and Romanov 2005). The antagonistic function of cytokinins and of the polyamines spermine and spermidine has been reported in other older studies as well (Naika et al. 1980; Feray et al. 1992). Considering the similarity of their physiological activity, and based on the above information, it has been proposed that cytokinins mediate their function partially through the regulation of polyamines homeostasis.

The negative role of cytokinins on root growth and development has been known for a long time (Cary et al. 1995). Analysis of the AtST4b loss-of-function mutant revealed a negative regulatory role of this gene on root growth in Arabidopsis (Fig. 15). Furthermore, AtST4b homozygous mutants showed reduced sensitivity to different concentrations of cytokinins (Fig. 15A and 15C). The reduced sensitivity to cytokinin inhibition of root growth is also reported in plants with impaired-cytokinin signaling such as the Arabidopsis cytokinin receptor mutants (Higuchi et al. 2004; Nishimura et al. 2004), the Arabidopsis histidine phosphotransfer (AHPs) mutants (Hutchison et al. 2006), the Arabidopsis type-B ARR mutants (Mason et al. 2005) and the cytokinin-deficient tobacco plants that overexpress different cytokinin oxidases (CKXs) (Werner et al. 2001). However, the role of AtST4b in shoot and reproductive development is not consistent with the known cytokinin function. The AtST4b-KO plants produce more leaves, slightly bigger rosette, more siliques and bigger seeds than wild type plants. In
contrast, cytokinin $ahk2/3/4$ triple receptor mutants and cytokinin-deficient plants produce stunted shoots (with decreased leaf size and number), less seeds per siliques or infertile but bigger seeds (Werner et al. 2001; Werner et al. 2003; Higuchi et al. 2004; Nishimura et al. 2004; Riefler et al. 2006).

Several hypotheses can be proposed to explain the overall improved root and shoot growth of $AtST4b$-KO plants. In the first hypothesis, we can consider that cadabicine is a positive growth regulator and is inactivated by sulfonation. A similar inactivation function has been previously proposed for $AtST2a$ in controlling the biological activity of 12-hydroxyjasmonic acid in $A. thaliana$ (Gidda et al. 2003). Accumulation of polyamine conjugates has been shown to promote cell division, flower formation, organogenesis and tuber induction in plants (Facchini et al. 2002). In this scenario, it is the accumulation of cadabicine in $AtST4b$ mutant plants that induces plant growth. The results of mass spectrometry analyses support this hypothesis, and show that endogenous levels of free cadabicine is much higher in roots of $AtST4b$-KO plants compared to wild type plants (data not shown). Alternatively, cadabicine sulfate might be a general growth inhibitor and its absence would explain the improved growth observed for the $AtST4b$-KO plants.

In order to find out which of the two possibilities is right, a careful examination of the growth parameters of the $sct$ mutant which can not synthesize cadabicine and cadabicine sulfate will be required. $AtST4b$ can also be part of a pathway that copes with excess polyamines or polyamine conjugates in plants. Interestingly, the increase in the intracellular polyamines induces the spermidine/spermine N1-acetyltransferase (SSAT) gene which in turn inactivates the excess spermine and spermidine in humans (Moinard et al. 2005). To our knowledge, there is no SSAT in the genome of Arabidopsis.
However, a similar inactivation mechanism can be proposed in Arabidopsis in which spermidine dicoumaroyl transferase (SCT) and AtST4b control excessive spermidine accumulation in cells. Other conjugating enzymes might partly complement the missing SCT or AtST4b to prevent the toxic accumulation of spermidine but allowing an increase in spermidine accumulation sufficient to explain the improved growth phenotype.

In order to better understand the function of cadabicine and cadabicine sulfate in plants, several experiments will have to be conducted in the future.

- Comparison of the levels of spermidine and of its derivatives in AtST4b-KO, sct-KO and wild type plants. The results of these studies would allow to characterize other molecules that could affect growth in Arabidopsis and to have a better understanding of the dynamic of the accumulation of these important molecules.

- Study the tissue distribution of spermidine and its derivatives in Arabidopsis. SCT and AtST4b are only expressed in the root system. However, we could detect accumulation of cadabicine in the aerial parts suggesting a transport mechanism. The cytokinin-dependent sulfonation reaction taking place in the roots might block the transport of active cadabicine in the aerial parts. This would explain the increased growth of the shoot and root system of the AtST4b mutant plants.

- Careful examination of the growth behavior of the sct mutant. This mutant cannot synthesize cadabicine and cadabicine sulfate. A reduced growth phenotype would support a positive role for cadabicine. In contrast, an improved growth phenotype would support an inhibitory role of the final product, cadabicine sulfate.

- Construction and analysis of growth parameters of a transgenic line overexpressing AtST4b. The availability of this overexpressor line would allow a
cytokinin independent reduction of the pool of cadabicine in the aerial part of the plant and to study its effect on shoot growth and seed set.

4.2) Biochemical and biological characterization of AtST4a and AtST4c

Based on their amino acid sequence similarity, AtST4a and AtST4c are more closely related to each other (80% amino acid sequence identity) than to AtST4b. Analysis of the Genevestigator microarray database and our RT-PCR experimental results show that both genes are expressed mainly in roots and are down-regulated by cytokininins (Fig. 8 and Fig. 9C) (Zimmermann et al. 2004; Marsolais et al. 2007). Investigation of their promoter region also confirmed the presence of cytokinin regulatory motifs (Table 1). However, in contrast with AtST4b, AtST4a and AtST4c are repressed by cytokinins suggesting that they might have opposite functionalities.

Unfortunately, we were not able to identify the reaction products of AtST4a and AtST4c in vivo. The sulfated metabolome of AtST4a- and AtST4c-KO root extracts is almost identical to the wild type one, suggesting a redundant function for the two enzymes (Fig. 20, Fig. 21, and Fig. 22). However, the loss of function mutation of the two genes gave different phenotypes suggesting that the two enzymes might have different substrates with very similar chemical properties. Alternatively, the two genes might sulfonate the same substrate in different tissues explaining the different phenotypes.

Recently, AtST4a has been characterized in vitro and was shown to exhibit a relatively broad specificity toward brassinosteroids (Marsolais et al. 2007). However, the natural occurrence of sulfated brassinosteroids (BRs) has not been reported yet. BRs are widely distributed in plants, and play several regulatory roles during growth and development. Analysis of BR-deficient and -insensitive mutants confirmed their essential role for cell
elongation, male fertility, senescence and vascular differentiation (Clouse et al. 1996; Altmann 1998).

The crosstalk between cytokinins and BRs regulatory pathways is reported in several studies. For example, BRs interact with cytokinins and other phytohormones like auxin to regulate ethylene biosynthesis in plants (Arteca and Arteca 2008; Hansen et al. 2009). The physiological roles of cytokinins and BRs are similar in some aspects. BRs, like cytokinins, stimulate cell division through up-regulation of cyclin D3, a member of a family of proteins that allow progression through the cell cycle, and BRs can even substitute cytokinins in cell cultures of Arabidopsis. However, it is postulated that the pathway by which BRs regulate cyclin D3 is different from the one induced by cytokinins (Hu et al. 2000). The shoot growth promoting activity of BRs is also similar to the one induced by cytokinins (for a review see (Clouse and Sasse 1998). On the other hand, the effect of BRs on root growth is intriguing and contradictory results have been explained by the use of different experimental conditions. For example, even though low concentrations (<pM) of exogenous BRs stimulate root growth in wild type plants, higher concentrations were found to be inhibitory (Clouse et al. 1996; Mussig et al. 2003).

Based on the fact that CKs and BRs have inhibitory and stimulatory effects on growth, cytokinin-mediated repression of a brassinosteroid sulfotransferase can be part of a pathway in which both CKs and BRs are working together in order to regulate root and shoot growth in A.thaliana. This hypothesis will be confirmed only when we will be able to demonstrate that the AtST4a and/or AtST4c enzymes are sulfonating brassinosteroids in vivo.
Phenotypic analyses of the loss of function mutants revealed that, unlike *AtST4b*, *AtST4c* seems to play a positive function in root and shoot growth in Arabidopsis. *AtST4c*-KO plants produced shorter primary roots, reduced number of lateral roots, slightly smaller rosettes, reduced number of leaves, less seeds per siliques and finally smaller seeds. Based on the repression of *AtST4c* expression by cytokinins and the phenotype of the loss of function mutant, one can conclude that *AtST4c*, as opposed to *AtST4b*, positively regulates plant growth and that this positive effect is repressed by the cytokinin signaling pathway. In general, loss of *AtST4a* did not cause a strong phenotype in root and shoot growth. This mild effect is consistent with the weak expression of this gene seen in RT-PCR reactions (Fig. 8 and Fig. 9C).

*AtST4c*-KO plants flower earlier than wild type Arabidopsis. This phenotype was highly reproducible and observed in soil and *in vitro* grown plants. The late-flowering phenotype has also been reported in some cytokinin-deficient (Werner et al. 2003) and cytokinin-insensitive Arabidopsis mutants (Nishimura et al. 2004; Riefler et al. 2006) suggesting a positive role of cytokinins in flower induction. Furthermore, the relationship between *AtST4c* expression and flowering is strengthened by the fact that *AtST4c* (or *AtST4a*) is repressed by the *LEAFY (LFY)* gene (−9.84 and 13.14 fold induction in lfy-12 and 35S::amiR-lfy-1, respectively) (Zimmermann et al. 2004). *LEAFY* is a floral-meristem identity gene that regulates transition to flowering and the subsequent patterning of young floral meristems (Schultz and Haughn 1991; Huala and Sussex 1992). The link between the floral meristem-identity genes and flowering-time genes has been difficult to assess, but it is presumed that the former regulate onset of flowering in response to environmental stimuli and action of the genes that induce flowering time.
Taken together, these results suggest that the general growth reduction of the AtST4c-KO mutant growth might be the result of a premature switch from vegetative to reproductive growth, which is accompanied by a general growth arrest. In order to clarify the role of AtST4c in cytokinin-mediated growth control and its role in flowering, the following experiments should be performed in the future.

- Identify the endogenous substrate of AtST4c. Two approaches should be used: Overexpression of AtST4c to allow the identification of the sulfated product and analysis of the sulfated metabolome of the AtST4a/4c double mutant.

- Localize the site of expression of AtST4a and AtST4c in whole plants by the construction of promoter-GFP fusions.

- Study AtST4c expression in other meristem identity mutants of Arabidopsis.

- Study the regulation of meristem identity genes in the AtST4c-KO mutant.
References:


Yokoyama, A., T. Yamashino, Y. Amano, Y. Tajima, A. Imamura, H. Sakakibara and T. Mizuno (2007). "Type-B ARR transcription factors, ARR10 and ARR12, are..."

Table 1. Statistical data of root length phenotype analysis.

Root length analysis, DAG\(^1\)=2, Kruskal-Wallis Asymp. Sig. \(^2\)= 0.0

<table>
<thead>
<tr>
<th>Genotype (^3)</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number (^4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>315</td>
<td>0.455</td>
<td>Ctl =42 A=17</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>569.5</td>
<td>0.075</td>
<td>Ctl=42 B=35</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>268</td>
<td>2.44E(^{-5})</td>
<td>Ctl=42 C=29</td>
</tr>
</tbody>
</table>

Root length analysis, DAG=3, Kruskal-Wallis Asymp. Sig. = 0.0

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>268</td>
<td>0.001</td>
<td>Ctl =43 A=25</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>535</td>
<td>0.028</td>
<td>Ctl=43 B=35</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>99</td>
<td>3.41E(^{-10})</td>
<td>Ctl=43 C=31</td>
</tr>
</tbody>
</table>

Root length analysis, DAG=4, Kruskal-Wallis Asymp. Sig. = 0.0

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>403</td>
<td>0.087</td>
<td>Ctl =43 A=25</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>393</td>
<td>2.88E(^{-4})</td>
<td>Ctl=43 B=35</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>122</td>
<td>3.91E(^{-9})</td>
<td>Ctl=43 C=30</td>
</tr>
</tbody>
</table>

\(^1\) DAG= Days after Germination
\(^2\) Asymp. Sig. stands for asymptotic significance
\(^3\) Ctl, A, B and C stand for control, \(AtST4a\)-, \(AtST4b\)- and \(AtST4c\)-KO lines, respectively
\(^4\) Number shows the numbers of replicates in each mutant line
### Root length analysis, DAG=6, Kruskal-Wallis Asymp. Sig. = 0.0

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>426</td>
<td>0.076</td>
<td>Ctl =44 A=26</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>295</td>
<td>8.69E-6</td>
<td>Ctl=44 B=33</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>259</td>
<td>1.01E-5</td>
<td>Ctl=44 C=30</td>
</tr>
</tbody>
</table>

### Root length analysis, DAG=7, Kruskal-Wallis Asymp. Sig. = 0.0

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>379</td>
<td>0.06</td>
<td>Ctl =42 A=25</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>385</td>
<td>0.001</td>
<td>Ctl=42 B=33</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>281</td>
<td>6.74E-5</td>
<td>Ctl=42 C=30</td>
</tr>
</tbody>
</table>

### Root length analysis, DAG=8, Kruskal-Wallis Asymp. Sig. = 0.0

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>176</td>
<td>0.002</td>
<td>Ctl =40 A=18</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>343</td>
<td>0.001</td>
<td>Ctl=40 B=31</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>336</td>
<td>0.002</td>
<td>Ctl=40 C=30</td>
</tr>
</tbody>
</table>

### Root length analysis, DAG=9, Kruskal-Wallis Asymp. Sig. = 0.0

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>199</td>
<td>0.016</td>
<td>Ctl =37 A=18</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>349</td>
<td>0.015</td>
<td>Ctl=37 B=29</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>375</td>
<td>0.059</td>
<td>Ctl=37 C=28</td>
</tr>
</tbody>
</table>
### Table 2. Statistical data of hormone treatment analysis of root tissue.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Levene's Test for Equality of Variances sig.</th>
<th>t-test for Equality of Means sig.</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>0.068</td>
<td>0.014</td>
<td>Ctl=19 B=19</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>0.631</td>
<td>0.002</td>
<td>Ctl=20 B=18</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>73.00</td>
<td>0.011</td>
<td>Ctl=18 B=16</td>
</tr>
</tbody>
</table>

### Table 3. Statistical data of number of lateral root phenotype analysis.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Levene's Test for Equality of Variances sig.</th>
<th>t-test for Equality of Means sig.</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; B</td>
<td>0.5</td>
<td>0.149</td>
<td>Ctl=10 B=10</td>
</tr>
</tbody>
</table>
### Numbers of lateral roots, DAG=14, two-independent-samples test

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Levene's Test for Equality of Variances sig.</th>
<th>t-test for Equality of Means sig.</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; B</td>
<td>0.133</td>
<td>0.062</td>
<td>Ctl=9</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B=12</td>
</tr>
</tbody>
</table>

### Numbers of lateral roots, DAG=10, Kruskal-Wallis Asymp. Sig. = 0.001

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>120</td>
<td>0.836</td>
<td>Ctl =21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A=12</td>
</tr>
<tr>
<td>Ctrl &amp; C</td>
<td>48</td>
<td>4.49E-4</td>
<td>Ctl=21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C=15</td>
</tr>
</tbody>
</table>

### Numbers of lateral roots, DAG=13, Kruskal-Wallis Asymp. Sig. = 0.147

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctrl &amp; A</td>
<td>60.500</td>
<td>0.320</td>
<td>Ctl =20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A=8</td>
</tr>
<tr>
<td>Ctrl &amp; C</td>
<td>95.500</td>
<td>0.069</td>
<td>Ctl=20</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C=15</td>
</tr>
</tbody>
</table>
ANNEX 2

Table 1. Statistical data of rosette diameter analysis.

Rosette diameter, DAG=14, Kruskal-Wallis Asymp. Sig. = 0.38  
Rosette diameter, DAG=21, Kruskal-Wallis Asymp. Sig. = 0.015

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>334</td>
<td>0.035</td>
<td>Ctl =27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A=36</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>288</td>
<td>0.028</td>
<td>Ctl=27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B=32</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>451</td>
<td>0.76</td>
<td>Ctl=27</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C=35</td>
</tr>
</tbody>
</table>

Rosette diameter, DAG=26, Kruskal-Wallis Asymp. Sig. = 0.008

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>297</td>
<td>0.48</td>
<td>Ctl =21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A=32</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>296</td>
<td>0.867</td>
<td>Ctl=21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B=29</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>165</td>
<td>0.014</td>
<td>Ctl=21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C=29</td>
</tr>
</tbody>
</table>

Table 2. Statistical data of numbers of leaves phenotype analysis.

Number of leaves, DAG=14, Kruskal-Wallis Asymp. Sig. = 0.001

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>391</td>
<td>0.7</td>
<td>Ctl =21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>A=39</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>298</td>
<td>0.013</td>
<td>Ctl=21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>B=39</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>326</td>
<td>0.331</td>
<td>Ctl=21</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>C=35</td>
</tr>
</tbody>
</table>
Number of leaves, DAG=21, Kruskal-Wallis Asymp. Sig. = 0.00

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>298</td>
<td>0.010</td>
<td>Ctl =29 A=32</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>253</td>
<td>4.21E-4</td>
<td>Ctl=29 B=34</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>35</td>
<td>9.03E-12</td>
<td>Ctl=29 C=35</td>
</tr>
</tbody>
</table>
ANNEX 3

Table 1. Statistical data of number of seeds per siliques analysis.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>387</td>
<td>0.355</td>
<td>Ctl=30 A=30</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>321</td>
<td>0.056</td>
<td>Ctl=30 B=30</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>109</td>
<td>4.49E-7</td>
<td>Ctl=30 C=30</td>
</tr>
</tbody>
</table>

Table 2. Statistical data of number of siliques analysis.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>84.5</td>
<td>4.47E-5</td>
<td>Ctl=23 A=24</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>141</td>
<td>0.011</td>
<td>Ctl=23 B=22</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>0.00</td>
<td>1.47E-9</td>
<td>Ctl=23 C=27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; A</td>
<td>49</td>
<td>6.69E-6</td>
<td>Ctl=24 A=20</td>
</tr>
<tr>
<td>Ctl &amp; B</td>
<td>87</td>
<td>1.72E-4</td>
<td>Ctl=24 B=21</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>20</td>
<td>5.377E-8</td>
<td>Ctl=24 C=23</td>
</tr>
</tbody>
</table>
Table 3. Statistical data of seed size phenotype analysis.

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; B</td>
<td>393</td>
<td>8.57E-5</td>
<td>Ctl=40 B=40</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>490</td>
<td>0.003</td>
<td>Ctl=40 C=40</td>
</tr>
</tbody>
</table>

Seed length, Kruskal-Wallis Asymp. Sig. = 0.00

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; B</td>
<td>504</td>
<td>0.004</td>
<td>Ctl=40 B=40</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>341</td>
<td>8.57E-5</td>
<td>Ctl=40 C=40</td>
</tr>
</tbody>
</table>

Seed width, Kruskal-Wallis Asymp. Sig. = 0.00

<table>
<thead>
<tr>
<th>Genotype</th>
<th>Mann-Whitney U</th>
<th>Asymp. Sig. (2-tailed)</th>
<th>Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ctl &amp; B</td>
<td>416</td>
<td>2.19E-4</td>
<td>Ctl=40 B=40</td>
</tr>
<tr>
<td>Ctl &amp; C</td>
<td>313</td>
<td>2.77E-6</td>
<td>Ctl=40 C=40</td>
</tr>
</tbody>
</table>